JOC The Journal of Organic Chemistry

Featured Article

pubs.acs.org/joc

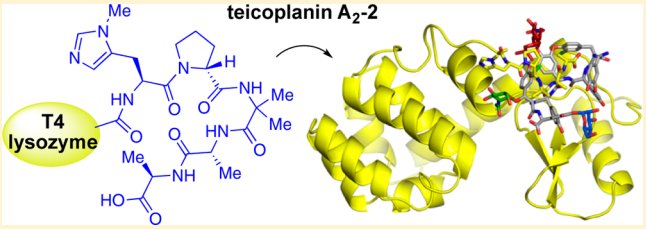
X-ray Crystal Structure of Teicoplanin A_2-2 Bound to a Catalytic Peptide Sequence via the Carrier Protein Strategy

Sunkyu Han,† Binh V. Le,† Holly S. Hajare, Richard H. G. Baxter, and Scott J. Miller*

Department of Chemistry, Yale University, New Haven, Connecticut 06511, United States

Supporting Information

ABSTRACT: We report the X-ray crystal structure of a site-selective peptide catalyst moiety and teicoplanin A_2-2 complex. The expressed protein ligation technique was used to couple T4 lysozyme (T4L) and a synthetic peptide catalyst responsible for the selective phosphorylation of the *N*-acetylglucosamine sugar in a teicoplanin A_2-2 derivative. The T4L-Pmh-DPro-Aib-DAla-DAla construct was crystallized in the presence of teicoplanin A_2-2. The resulting 2.3 Å resolution protein−peptide−teicoplanin complex crystal structure revealed that the nucleophilic nitrogen of *N*-methylimidazole in the Pmh residue is in closer proximity (7.6 Å) to the *N*-acetylglucosamine than the two other sugar rings present in teicoplanin (9.3 and 20.3 Å, respectively). This molecular arrangement is consistent with the observed selectivity afforded by the peptide-based catalyst when it is applied to a site-selective phosphorylation reaction involving a teicoplanin A_2-2 derivative.



INTRODUCTION

The development of site-selective catalysts that are able to deliver different products, derived from a common scaffold, is emerging as an important strategy to the diversification of biologically active scaffolds.[1] Significant chemical challenges arise when the scaffold to be modified is a complex natural product, especially in situations where the structure of the natural product contains multiple copies of the same functional group.[2] Polyhydroxylated compounds are but one example of molecules of this type; attempted derivatization of the hydroxyl group array (for example, through acylations,[3−7] phosphorylations,[8,9] sulfonylations,[10] thiocarbonylations,[11,12] phosphoramiditation,[13] glycosylations,[14] silylations,[15,16] inter alia) can lead to mixtures of different monofunctionalized products as well as mixtures of bis-, tris-, or even higher functionalized products. The chromatographic separation of complex mixtures, as well as the precise structural assignment of each component, can be a daunting task, independent of any attempt to describe a biological function to each new analogue that might be prepared in this way.

Catalysis provides an opportunity to simplify the challenge, provided that catalysts can be found that are selective for unique molecular products. Many types of catalysts have been applied to this problem, including enzymes[17] and, increasingly, small molecules.[2] Numerous catalyst discovery efforts of this type have involved a healthy dose of screening catalyst libraries.[18,19] Aside from often providing successful outcomes, nominal combinatorial techniques have been reasonable to apply to the problem, given the complexity of the reactions in terms of scaffold complexity, as well as the numerics of possible products that can be formed. On the other hand, so-called "rational design" of catalysts for the objective of site-selective modification of a complex scaffold may also be possible in situations where ample information may be available about how a particular natural product scaffold might associate with a sector of a catalyst candidate. We have been exploring the possibility of rational design of site-selective catalysts for the modification of the glycopeptide antibiotics vancomycin[12,20] and teicoplanin.[21] Our choice of these scaffolds was based on two considerations: (1) antibiotic resistance and the need for new antibiotics[22,23] and (2) the opportunity to exploit the well-known mechanism of action for inspiration in terms of catalyst design.[24]

In this context, we recently reported the peptide-based site-selective phosphorylation of teicoplanin derivative **2** (Figure 1).[25] From these studies, we identified three catalysts (**3**−**5**) that selectively phosphorylated three distinct sugar units present in teicoplanin derivative **2** (Table 1). Specifically, the design of catalysts **3** and **5** was inspired by the well-precedented interactions between the DAla-DAla moiety and glycopeptide natural products.[22,23] The X-ray crystal structure of the Lys-DAla-DAla peptide moiety bound to teicoplanin A_2-2 (**1**) reported by Loll and co-workers[26] served as a scaffold for the rational design of novel catalysts (Figure 2).

RESULTS AND DISCUSSION

While many experiments supported the strong correlation between the hypothesis and the observed results,[25] one element that was missing was any type of direct observation of the interactions that we postulated to account for the selectivity. A technique that we pursued was X-ray crystallographic analysis

Received: July 18, 2014
Published: August 22, 2014

   dx.doi.org/10.1021/jo501625f | J. Org. Chem. 2014, 79, 8550−8556

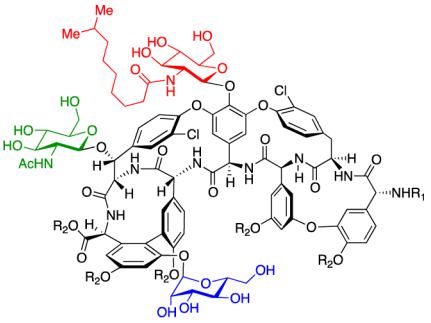
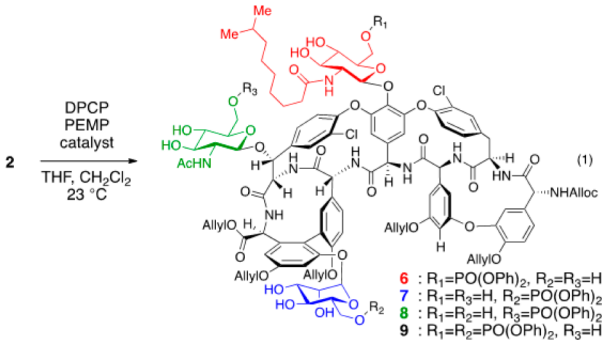


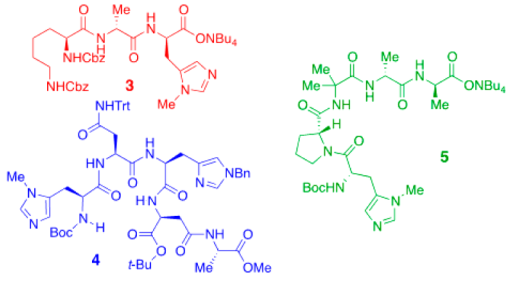
Figure 1. Structure of teicoplanin A_2-2 (**1**) and protected teicoplanin A_2-2 derivative **2**.

Table 1. Our Previous Studies on the Site-Selective Phosphorylation of Teicoplanin Derivative 2*



Entry	Catalyst	HPLC trace[a]
1	none	
2	NMI (100 mol%)	
3[b]	**3** (10 mol%)	
4[c],[d]	**4** (20 mol%)	
5[e]	**5** (30 mol%)	

*Reaction conditions: DPCP (6 equiv), PEMP (8 equiv), catalyst, THF, CH_2Cl_2, 23 °C unless noted otherwise. [a]HPLC traces were recorded at 280 nm. [b]**6** was isolated in 42% yield. [c]DPCP (3 equiv), PEMP (4 equiv), catalyst, THF, CH_2Cl_2, 23 °C. [d]**7** was isolated in 23% yield. [e]**8** was isolated in 41% yield. DPCP: diphenylchlorophosphate. PEMP: 1,2,2,6,6-pentamethylpiperidine.



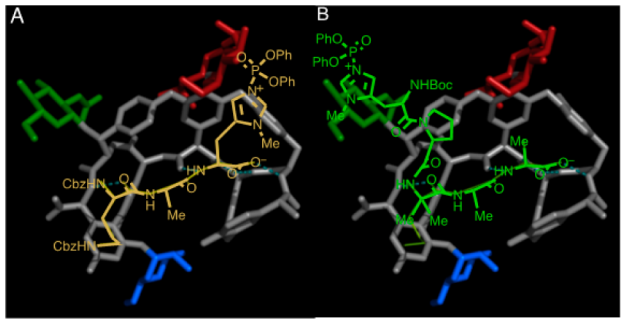
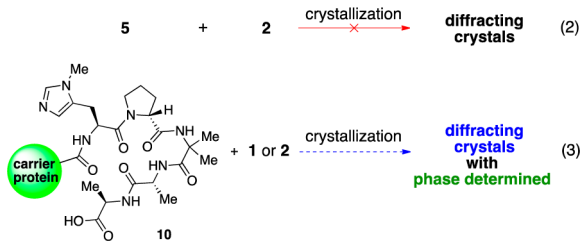


Figure 2. Predicted model of catalysts (**3** and **5**) and teicoplanin interaction based on the crystal structure of the Lys-D-Ala-D-Ala and teicoplanin A_2-2 complex (catalysts **3** and **5** were superposed to the crystal structure adopted from PDB 3VFJ). (a) Predicted diphenylphosphorylated Z-Lys(Z)-D-Ala-D-Pmh (**3**, yellow-orange) and teicoplanin A_2-2 (**1**) complex consistent with the observed *N*-dodecylglucosamine ("top/red" sugar) selectivity. (b) Predicted diphenylphosphorylated Boc-Pmh-D-Pro-Aib-D-Ala-D-Ala (**5**, green) and teicoplanin A_2-2 (**1**) complex consistent with the observed *N*-acetylglucosamine ("left/green" sugar) selectivity.

of catalyst/substrate complexes. While the observation of a complex of this nature is fraught with caveats,[27] we hoped to demonstrate at least plausibility of the models we proposed by X-ray crystallographic analysis. We describe below our studies that culminated in the direct observation of a crystal structure that supported many features of the initial hypothesis presented above for the catalyst targeted to selectively derivatize the most recalcitrant sugar on the scaffold, the *N*-acetylglucosamine now referred to as the "left/green" sugar.

We initially attempted the cocrystallization of catalyst **5** and teicoplanin derivative **2** (Scheme 1, eq 2). However, despite

Scheme 1. Crystallization Strategies for the Structure Determination of the Pmh-D-Pro-Aib-D-Ala-D-Ala Peptide Sequence and the Teicoplanin Complex



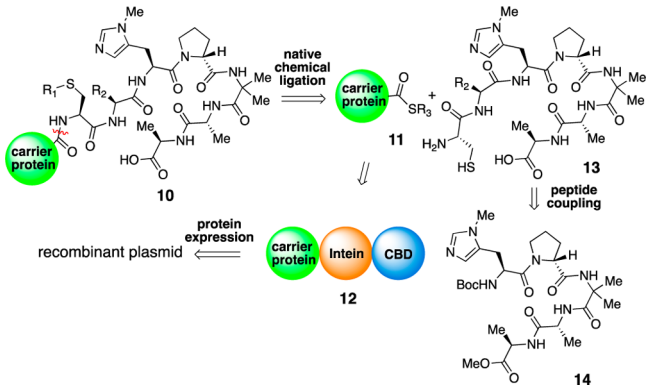
numerous attempts with various crystallization conditions, the formation of a single crystal suitable for the X-ray crystallographic analysis remained elusive. Hence, we envisioned incorporating a crystallization aid to the catalyst-teicoplanin complex. Loll and co-workers reported the carrier-protein strategy for the X-ray crystal structure determinations of the complexes of Lys-D-Ala-D-Ala peptide residues and various glycopeptide antibiotics.[26,28] In their studies, the Cys-Lys-D-Ala-D-Ala peptide moiety was conjugated to a carrier protein such as maltose-binding protein (MBP), ubiquitin, or T4L using expressed protein ligation.[29,30] The resulting recombinant protein construct was allowed to complex with glycopeptides such as dalbavancin, ristocetin, vancomycin, and teicoplanin, which resulted in crystals suitable for X-ray structure analysis. Inspired by these studies, we planned to covalently bind the peptide residues of catalyst **5** (Pmh-D-Pro-Aib-D-Ala-D-Ala) to the

dx.doi.org/10.1021/jo501625f | J. Org. Chem. 2014, 79, 8550−8556

carrier protein to obtain ligation product **10** and allow it to interact with teicoplanin A_2-2 (**1**) or its derivative **2** to obtain single crystals suitable for the X-ray structure determination (Scheme 1, eq 3).

We envisioned coupling the carrier protein and the synthetic peptide, which contains the functional residues of catalyst **5** via expressed protein ligation (Scheme 2).[29,30] The key peptide

Scheme 2. Retrosynthetic Analysis for the Formation of the Carrier Protein and Peptide Catalyst Conjugate 10

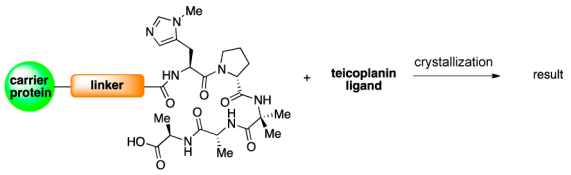


bond in the carrier protein−synthetic peptide conjugate **10** would be formed by native chemical ligation of the thioester of carrier protein **11** and the cysteine residue at the *N*-terminus of the synthetic peptide **13**.[31] Thioester **11** would be obtained by treating the carrier protein−intein−CBD (chitin binding domain) construct with a thiol reagent upon purification on chitin beads.[32] Protein construct **12** was to be expressed from a recombinant plasmid, which utilized the commercially available pTXB1 vector. Peptide **13** would be derived from the methyl ester derivative of "left/green" selective catalyst via two rounds of peptide coupling.

Based on our synthetic plan (Scheme 2), various carrier protein−peptide conjugates with different linkers were synthesized (Table 2). The initial crystallization attempt with ubiquitin as a carrier protein did not yield crystals suitable for X-ray diffraction under the conditions examined (Table 2, entries 1 and 2). Our first crystal was obtained from the MBP-based construct in the absence of ligand (Table 2, entry 3). While this crystal provided us with high resolution (1.3 Å) structural information on the MBP and the linker region, the catalytic peptide region (Pmh-DPro-Aib-DAla-DAla) was disordered (Table 2, entry 3). We reasoned that the longer linker was contributing to the disorder of the peptide and therefore deleted two serine residues from the linker. However, the crystal formation of the modified MBP−peptide−teicoplanin A_2-2 (**1**) complex remained elusive (Table 2, entry 4). In an attempt to lower the flexibility of the linker, we point mutated the glycine residue to an alanine residue and attempted crystallization with protected teicoplanin A_2-2 (**2**, Table 2, entry 5) and teicoplanin A_2-2 (**1**, Table 2, entry 6). The insolubility of protected teicoplanin derivative **2** in aqueous solution hampered the complexation with the protein−peptide conjugate. In spite of the high dilution and the prolonged mixing time of **2** with the MBP−peptide construct, the crystal obtained from this sample showed analogous unit cell dimensions to that without the ligand (Table 2, entry 3), and the peptide region was disordered, indicative of the absence of the ligand in the crystal. On the other hand, MBP-AAAAAC(carboxy methylated cysteine)G-Pmh-DPro-Aib-DAla-DAla and teicoplanin A_2-2 (**1**) complex formed a homogeneous solution with high binding affinity (K_d 90 nM).[33] However, diffraction-quality crystals were not obtained despite extensive optimization efforts.

Our first generation crystal structure including an ordered peptide region was obtained from the T4L-peptide ligated product and teicoplanin A_2-2 (**1**) complex (Table 2, entry 7, PDB 4PJZ). The data set obtained from this crystal was refined to an effective resolution of 1.9 Å and unambiguously showed the electron density map of T4L, teicoplanin A_2-2, and the DPro-Aib-DAla-DAla peptide moiety (Figure 3). However, the peptide residues from Asn 163 to Pmh 170 were disordered. We reasoned that the disorder resulted from a mismatch between the spatial requirement to place eight amino acids (from Asn 163 to Pmh 170) and the distance between Lys 162 and DPro 171 residues, which are 15.6 Å apart (Figure 3). Hence, we generated the second-generation T4L construct deleting the GSS motif from the C-terminus of the protein.

Table 2. Optimization of the Carrier Protein and the Linker



entry	carrier protein	linker[a]	crystallization[b]	ligand[c]	diffraction	note
1	ubiquitin	GSSCG	X	**1**		
2	ubiquitin	GSSCG	X	**2**		**2** is not soluble in water
3	maltose-binding protein	AAAAGSSCG	O	none	1.3 Å	structure is disordered in the peptide region
4	maltose-binding protein	AAAAGCG	X	**1**		
5	maltose-binding protein	AAAAACG	O	**2**	2.2 Å	**2** is not soluble in water; Structure disordered in the peptide region
6	maltose-binding protein	AAAAACG	Δ	**1**		small spherulitic crystals
7[d]	T4 lysozyme wt*	GSSCG	O	**1**	1.9 Å	Asn 163−Pmh 170 region is disordered
8[d]	T4 lysozyme wt*	CA	O	**1**	2.3 Å	ordered

[a]Cysteine was converted to *S*-carboxymethyl cysteine to prevent oxidative byproduct formation. [b]Screening was conducted using MBClass, MBClass II, PEGs, PEGs II suites, JCSG Core suite I, JCSG Core suite II, JCSG Core suite III, JCSG Core suite IV from Qiagen, Index from Hampton Research, and Wizard Classic 1 and 2 from Emerald Bio. O: diffracting crystals, X: no crystals, Δ: nondiffracting crystals. [c]Protein construct and ligand were mixted in 1:1.5 ratio. [d]Cysteine free T4 lysozyme was used as the carrier protein.

dx.doi.org/10.1021/jo501625f | J. Org. Chem. 2014, 79, 8550−8556

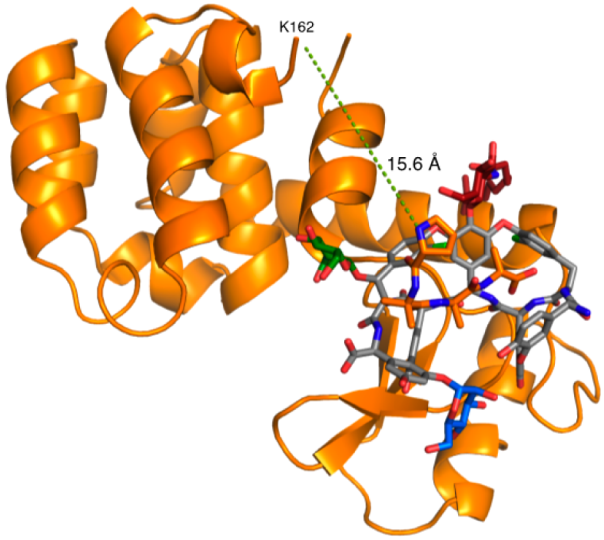


Figure 3. Ribbon diagram of the first-generation T4L−peptide conjugate and teicoplanin A_2-2 (**1**) complex (PDB 4PJZ). The teicoplanin A_2-2 ligand is displayed as sticks, carbon atoms colored gray. *N*-Dodecylglucosamine (top) carbon atoms colored in red, *N*-acetylglucosamine (left) carbon atoms colored in green, mannose (bottom) carbon atoms colored in blue.

Simultaneously, we mutated the spacer glycine residue between the cysteine and Pmh residues to alanine (G169A) to provide additional structural rigidity (Table 2, entry 8).

The crystal obtained from the second-generation T4L−peptide construct provided an X-ray structure to 2.3 Å resolution with all residues ordered (Figure 4, PDB 4PK0).

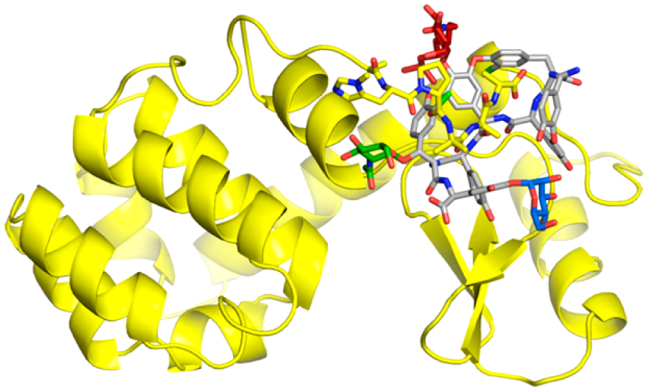


Figure 4. Ribbon diagram of the 2nd generation T4L-Pmh-DPro-Aib-DAla-DAla and teicoplanin A_2-2 complex (PDB 4PK0). The C-terminus appended by expressed protein ligation is displayed as sticks, carbon atoms colored yellow. The teicoplanin A_2-2 ligand is displayed as sticks, carbon atoms colored gray. *N*-dodecylglucosamine (top) carbon atoms colored in red, *N*-acetylglucosamine (left) carbon atoms colored in green, mannose (bottom) carbon atoms colored in blue.

Interestingly, the space group for both the first- and the second-generation crystal structures is *I*222, which has not been previously observed for T4 lysozyme and is distinct from the more common $P3_221$ symmetry in which the T4L-Lys-DAla-DAla-vancomycin complex (PDB 3RUN) was determined.[28] The asymmetric unit contains a single copy of protein−peptide conjugate and antibiotic. Teicoplanin rests against the loop following the first helix of T4L (residues 12−16). The teicoplanin *N*-acetylglucosamine lies 6 Å from the active site residue Glu 11 of the molecule to which it is bound, while the first residue of teicoplanin makes contact with Glu 22 and Arg 137 of a symmetry-related molecule.

The Aib-DAla-DAla residues are complexed to teicoplanin A_2-2 through five hydrogen bonds (Figure 5a), which is consistent

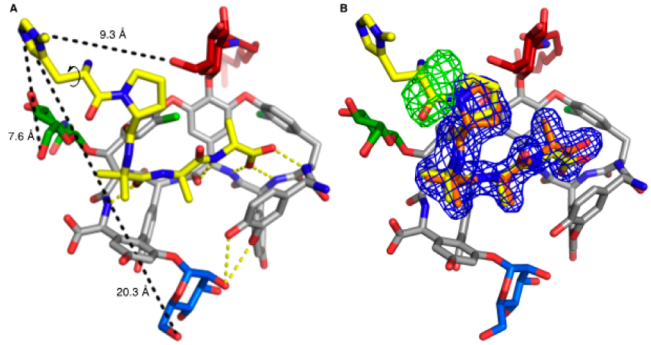


Figure 5. (a) Crystal structure of the second-generation Pmh-DPro-Aib-DAla-DAla and teicoplanin A_2-2 (**1**) complex (T4L is not shown for simplicity). (b) Superposition of the first-generation (orange) and the second-generation (yellow) of crystal structures of C-terminus peptide region and teicoplanin A_2-2 (**1**) complex. The teicoplanin A_2-2 coordinate is adopted from the first-generation crystal structure. Representative electron density is displayed for the refined model of the first-generation peptide. $2F_0$-F_c map (blue) contoured at 1.1σ and F_0-F_c map (green) contoured at 3.0σ for the unmodeled Pmh residue.

with our catalyst design. The *N*-terminus of the DPro residue bound next to the Aib residue is pointing toward the left side of the molecule and places the side chain of Pmh in close proximity to the "left/green" *N*-acetylglucosamine. On the basis of our crystal structure, the distance from the nucleophilic nitrogen to the primary alcohol of *N*-acetylglucosamine, *N*-dodecylglucosamine, and mannose are 7.6, 9.3, and 20.3 Å, respectively (Figure 5a). Notably, the close proximity of the nucleophilic moiety of Pmh to the "left/green" sugar is consistent with our observed selectivity in the phosphorylation of teicoplanin derivative **2** with catalyst **5**.[25]

Interestingly, our crystal structure reveals that the Pmh-DPro-Aib-DAla-DAla peptide moiety complexed to teicoplanin A_2-2 (**1**) does not adopt a β-hairpin structure. Literature precedents suggested that Boc-Pmh-DPro-Aib-DAla-DAla-ONBu$_4$ (**5**) could adopt a short β-turn, often associated with a DPro in the i+1 position of a tetrapeptide sequence.[34−38] However, in the presence of teicoplanin ligand, and while affixed to the carrier protein, DPro and Aib adopt ϕ and ψ values of an extended β-sheet (170, −130) and a left-handed α-helix (57, 39). The carboxymethylcysteine-alanine linker between the final residue of T4L residue Leu 164 and the catalytic Pmh adopt ϕ and ψ values consistent with a β-strand.[39] The superposition of the first and the second-generation crystal structures of the T4L−peptide conjugate and teicoplanin complex illustrates the conformational similarity between these two crystal structures (Figure 5b). Furthermore, the location of F_0-F_c density of the first-generation crystal structure is consistent with the refined position of Pmh in the second-generation crystal structure. These observations suggest the structural arrangement of catalyst peptide residues−teicoplanin complex is independent of the length and composition of the linker (Table 2, entries 7 and 8), at least in the context of the observed space group and packing within the unit cell.

dx.doi.org/10.1021/jo501625f | J. Org. Chem. 2014, 79, 8550−8556

A number of critical differences between the actual reaction conditions and the nature of the observed teicoplanin-catalytic sequence complex must be noted. The active catalytic species in eq 1 (Table 1) is presumed to be the highly reactive phosphorylated peptide catalyst, reminiscent of the phosphohistidine moiety implicated as an intermediate in the catalytic cycle carried out by histidine-dependent kinases.[40−42] Because of its high electrophilicity, we could not obtain a crystal structure of phosphorylated peptide and teicoplanin complex using a carrier protein strategy, which involves crystallization in aqueous solution. Notably, our crystal structure shows that the nucleophilic nitrogen in the *N*-methyl imidazole moiety of the peptide is pointing away from the "top/red" sugar and parallel to the "left/green" sugar. Hence, we surmise that the electrophilic phosphoryl group bound to the active catalytic species would be even further away (>9.3 Å) from the primary alcohol in the "top/red" sugar. Furthermore, the C6 of the *N*-acetylglucosamine residue is 4.2 Å from the Cβ of the Pmh amino acid. The Pmh side chain adopts a common outward-facing rotamer, but a 120° rotation around the Cα−Cβ bond would position the nucleophilic nitrogen of Pmh in closer proximity of the hydroxyl group it selectively derivatizes. This conformational flexibility of the Pmh side chain (Figure 5a, presented with rounded arrow) in conjunction with the free rotation involving the methylene of the "left/green" and "top/red" sugars, suggest potentially closer proximity of the imidazole moiety to the "left/green" sugar but further distance to the "top/red" sugar.

It is also of note that we obtained the crystal structure of the catalytic peptide sequence bound to teicoplanin A_2-2, without any adornment with protecting groups, which were necessary to achieve the site-selective phosphoryl transfer reactions. Attempted cocrystallization with the "allylyated" teicoplanin **2** has as yet been unsuccessful. Yet, it is striking that both the actual asymmetric reactions and the crystal structure of the nominal catalyst substrate complex give a self-consistent result.

CONCLUSION

In summary, a carrier protein strategy was effectively applied for the X-ray crystal structure determination of an intriguing site-selective catalyst that is effective for the modification of a teicoplanin A_2-2 derivative. Crystallographic analysis revealed a complex between the catalytically competent peptide sequence and teicoplanin itself that proved to be quite consistent with the original hypothesis that led to its study. A catalytic peptide−T4L conjugate based on the "left/green-selective catalyst" was prepared using expressed protein ligation. The resulting X-ray crystal structure reveals that the *N*-methylimidazole moiety, which serves to transfer the phosphoryl group in the O−P bond-forming reaction, is in closer proximity to the *N*-acetylglucosamine than the other sugar units, consistent with the observed selectivity in the phosphorylation event of our prior studies. While consideration of transition state structures based on analysis of intrinsically ground-state crystallographic complexes must always be exercised with caution,[43] structural insight obtained from this catalyst scaffold is expected to serve as a platform for the design of novel catalysts capable of selectively derivatizing glycopeptide antibiotics. Furthermore, we speculate that the use of a carrier protein strategy could be a broadly useful tool to obtain structural information in the field of asymmetric catalysis in situations where information on catalyst−substrate complexes are difficult to obtain otherwise.

EXPERIMENTAL SECTION

Synthesis of Boc-(Cys)$_2$-Ala-Pmh-D-Pro-Aib-D-Ala-D-Ala-OMe (S6). Structures of S1−S6 are shown in Supplementary Figure 1, Supporting Information).

To a flask charged with a known compound **14**[25] (253 mg, 0.416 mmol, 1 equiv) was added a 4 M solution of hydrochloric acid in 1,4-dioxane (4.6 mL) and methanol (2.2 mL) at 23 °C. After 3 h, the reaction mixture was concentrated under reduced pressure (caution: HCl gas release) and exposed to high vacuum until residual solvent and HCl were removed to yield **S1** as a white solid. **S1** was moved to the next step without further purification.

To a flask charged with protonated amine **S1** (226 mg, 0.416 mmol, 1 equiv), alanine derivative **S2** (78.7 mg, 0.416 mmol, 1.00 equiv), 1-ethyl-3-(3-(dimethylamino)propyl)carbodiimide hydrogen chloride (EDC·HCl, 87.8 g, 0.458 mmol, 1.10 equiv), and hydroxybenzotriazole hydrate (HOBt·H_2O, 70.14 mg, 0.458 mmol, 1.10 equiv) was added dichloromethane (4.8 mL) at 23 °C to yield a heterogeneous reaction suspension. After 3 min, *N*,*N*-diisopropylethylamine (165 μL, 0.915 mmol, 2.2 equiv) was added to the heterogeneous reaction mixture, and within 1 min, the reaction mixture became homogeneous. After 12 h, the reaction mixture was diluted with dichloromethane (4 mL). Subsequently, saturated aqueous sodium bicarbonate solution (4 mL) was added, and the layers were separated. The organic layer was dried over anhydrous sodium sulfate, filtered, and concentrated under reduced pressure to give **S3** as a white solid mixture. The sample of the crude residue of **S3** was purified by MPLC (SNAP-C18 120-g column with a 12-g samplet: solvent A = water, solvent B = acetonitrile, flow = 40 mL/min, λ = 210, 254 nm, loaded crude **S3** in methanol; equilibrated at 5% B for 3 column volumes (CV), then ramped to 95% B over 10 column volumes (CV)) to give methyl ester **S3** (208 mg, 74%) as a white solid.

To a flask charged with **S3** (208 mg, 0.306 mmol, 1 equiv) was added 4 M solution of hydrochloric acid in 1,4-dioxane (4 mL) and methanol (2 mL) at 23 °C. After 1.5 h, the reaction mixture was concentrated under reduced pressure (caution: HCl gas release) and exposed to high vacuum until residual solvent and HCl were removed to yield **S4** as a white foam. **S4** was moved to the next step without further purification.

To a flask charged with ClNH$_3$-Ala-Pmh-D-Pro-Aib-D-Ala-D-Ala-OMe (**S4**, 199 mg, 0.306 mmol, 2.00 equiv), **S5** (67.4 mg, 0.153 mmol, 1.00 equiv), 1-ethyl-3-(3-(dimethylamino)propyl)carbodiimide hydrogen chloride (EDC·HCl, 64.6 mg, 0.337 mmol, 2.20 equiv), and hydroxybenzotriazole hydrate (HOBt·H_2O, 51.6 mg, 0.337 mmol, 2.20 equiv) was added dichloromethane (4 mL) at 23 °C to yield a heterogeneous reaction suspension. After 3 min, *N*,*N*-diisopropylethylamine (118 μL, 0.673 mmol, 4.4 equiv) was added to the heterogeneous reaction mixture, and within 1 min, the reaction mixture turned homogeneous. After 24 h, the reaction mixture was diluted with dichloromethane (6 mL), saturated aqueous sodium bicarbonate solution (6 mL) was added, and the layers were separated. The organic layer was dried over anhydrous sodium sulfate, filtered, and concentrated under reduced pressure to give **S6** as a yellow solid mixture. The sample of the crude residue of **S6** was purified by MPLC (SNAP-C18 120-g column with a 12-g samplet: solvent A = water, solvent B = acetonitrile, flow = 40 mL/min, λ = 210, 254 nm, loaded crude **S6** in methanol; equilibrated at 5% B for 3 column volumes (CV), then ramped to 95% B over 10 column volumes (CV)) to give methyl ester **S6** (150.5 mg, 63%) as a white solid. ^{1}H NMR (500 MHz, $CDCl_3$, 21 °C): δ 8.18 (app-d, J = 38.9 Hz, 4H), 7.52 (s, 2H), 7.46 (d, J = 7.4 Hz, 2H), 7.38 (d, J = 7.6 Hz, 2H), 7.18 (s, 2H), 6.86 (s, 2H), 5.95 (s, 2H), 4.54−4.26 (m, 10H), 4.14 (app-s, 2H), 3.81−3.62 (m, 4H), 3.62 (s, 6H), 3.57 (s, 6H), 3.21 (br-s, 2H), 3.04 (br-s, 4H), 2.96 (br-s, 2H), 2.09 (br-s, 2H), 1.95 (br-s, 4H), 1.78 (br-s, 2H), 1.49 (s, 6H), 1.44 (s, 6H), 1.38 (app-s, 30H), 1.31 (d, J = 7.4 Hz, 6H). ^{13}C NMR (151 MHz, $CDCl_3$, 21 °C): δ 174.7, 173.4, 173.3, 173.1, 172.0, 171.0, 170.8, 155.8, 138.2, 127.3, 127.0, 80.4, 62.4, 57.3, 54.1, 52.4, 52.2, 49.6, 49.4, 48.4, 48.0, 42.1, 31.8, 29.1, 26.5, 25.2, 24.8, 18.0, 17.6, 17.2. FTIR (neat) cm^{-1}: 3326 (w), 1651 (s), 1528 (m), 1450

dx.doi.org/10.1021/jo501625f | J. Org. Chem. 2014, 79, 8550−8556

(w), 1162 (m). HRMS (ESI) (m/z): calcd for $C_{68}H_{108}N_{18}O_{20}S_2$, $[M + H]^+$ 1561.7507, found 1561.7487.

Synthesis of $ClNH_3$-$(Cys)_2$-Ala-Pmh-DPro-Aib-DAla-DAla-OH (S8). Structures of **S7** and **S8** are shown in Supplementary Figure 2, Supporting Information).

To a vial charged with **S6** (79.7 mg, 51.1 μmol, 1 equiv) was added a 4 M solution of hydrochloric acid in 1,4-dioxane (1.0 mL) and methanol (1.0 mL) at 23 °C. After 1 h, the reaction mixture was concentrated under reduced pressure (caution: HCl gas release) and exposed to high vacuum until residual solvent and HCl were removed to yield **S7** as a white solid. **S7** was moved to the next step without further purification.

Sodium hydroxide solution (1.0 M) in water (300 μL, 0.3 mmol, 6 equiv) was added via syringe to a solution of **S7** (76.9 mg, 51.1 μmol, 1 equiv) in methanol (1 mL) at 23 °C. After 2 h, 10% aqueous solution of hydrochloric acid (1 mL) was added to the reaction mixture, and the resulting mixture was concentrated under reduced pressure. The sample of the crude residue of **S8** was purified by MPLC (SNAP-C18 60-g column with a 12-g samplet: solvent A = water, solvent B = acetonitrile, flow = 30 mL/min, λ = 210, 254 nm, loaded crude **S8** in 10% aqueous solution of hydrochloric acid; equilibrated at 1% B for 3 column volumes (CV), then run at 1% B for 11.9 column volumes (CV), then ramped to 95% B over 12 column volume (CV)) to product **S8** (51.4 mg, 68%) as a pale yellow solid. ^{1}H NMR (600 MHz, D_2O, 21 °C):[44] δ 8.64 (s, 2H), 7.32 (s, 2H), 5.00 (dd, J = 8.8, 5.8 Hz, 2H), 4.41 (dd, J = 9.2, 4.7 Hz, 2H), 4.36–4.29 (m, 5H), 4.29–4.24 (m, 3H), 3.85 (s, 6H), 3.60–3.56 (m, 4H), ^{1}H 3.47 (dd, J = 15.2, 4.6 Hz, 2H), 3.21 (dd, J = 16.1, 5.8 Hz, 2H), 3.09 (td, J = 16.0, 8.9 Hz, 4H), 2.30–2.22 (m, 2H), 2.01 (p, J = 5.7 Hz, 2H), 1.96–1.85 (m, 4H), 1.47–1.44 (s, 6H), 1.44–1.39 (m, 12H), 1.39–1.35 (m, 12H). ^{13}C NMR (151 MHz, D_2O, 21 °C): δ 176.5, 176.0, 174.6, 173.8, 173.2, 169.0, 167.5, 135.1, 130.0, 118.1, 60.9, 56.5, 51.3, 49.8, 49.6, 48.5, 48.0, 37.2, 33.2, 29.0, 24.7, 24.6, 24.4, 23.8, 16.7, 16.6, 16.1. MS (ESI) (m/z): calcd for $C_{56}H_{89}N_{18}O_{16}S_2$, $[M + H]^+$ 1333.6145, found 1333.7275.

Construction of Recombinant Vectors. The gene encoding cysteine-free T4 lysozyme (T4L) was PCR-amplified from plasmid pET11b-T4L* with 5′ NdeI/3′ SapI restriction sites (Table S1, Supporting Information). The PCR product was digested, gel purified, and ligated into NdeI/SapI-digested pTXB1. The SapI restriction site introduced a C-terminal sequence GSS. The GSS sequence was subsequently deleted by site-directed mutagenesis using asymmetric primers. Recombinant plasmids were confirmed by restriction digest and DNA sequencing.

Synthesis of T4L–Peptide Conjugate. Both T4L-GSS (first-generation T4 lysozyme) and T4L (second-generation T4 lysozyme) were expressed in Rosetta (DE3) cells in terrific broth (TB) for 24 h at 20 °C postinduction. All steps of purification were carried out at 4 °C unless otherwise indicated. Cells were harvested by centrifugation and stored at −80 °C. The cell pellet was thawed in buffer A (500 mM NaCl, 25 mM Tris pH 7.5) and lysed by sonication, and the lysate was clarified by centrifugation at 100000g for 30 min. The supernatant was loaded onto a chitin column equilibrated with buffer A and washed until OD_{280} < 0.01. Intein-cleavage was initiated by adding buffer B (buffer A with 500 mM MESNA) to the column overnight at room temperature. T4L protein was eluted, concentrated, and exchanged against buffer C (100 mM HEPES pH 8.0, 500 mM NaCl, 500 mM MESNA). Native protein–peptide ligation was carried out with a 4-fold molar excess of synthesized peptide **S8**.[45] The ligation products were monitored by mass spectrometry to confirm the complete conversion (second-generation T4L: calculated mass, 19250.6; observed mass, 19251). The reaction was then loaded into a desalting column twice using buffer D (100 mM sodium borate pH 8.0, 150 mM NaCl, 5 mM EDTA). The thiol group of the protein–peptide fusion was protected by adding a 5-fold molar excess of iodoacetic acid for 2 h at room temperature. The protein–peptide chimera was further purified by size-exclusion chromatography using a HiLoad 16/60 Superdex 75 column with buffer E (20 mM HEPES pH 7.5, 25 mM NaCl).

Crystallization. The purified protein–peptide fusion was mixed with teicoplanin at 1:1.5 ratio. The sample was concentrated to 20 mg/mL (T4L) and screened in 96-well format with commercial sparse matrix screens using an automated liquid handling system. Initial crystallization conditions of both complexes were found in the PEG I screen with 20% PEG 3350 and 0.2 M dibasic potassium and sodium phosphate. The crystals were optimized in 24-well hanging drop plates. Diffraction-quality crystals grew as thin plates (Supplementary Figure 4, Supporting Information) at 19% PEG 3350 and 0.2 M dibasic potassium phosphate solution. Crystals were cryoprotected by transfer to 20% glycerol in mother liquor and flash-frozen in liquid nitrogen.

■ ASSOCIATED CONTENT

Supporting Information

General information, copies of ^{1}H and ^{13}C NMR spectra of synthetic peptides, crystallographic data, and supplementary discussion. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: scott.miller@yale.edu

Author Contributions

†These authors contributed equally to this work.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We are grateful to the National Institutes of General Medical Sciences of the National Institute of Health Foundation (GM-068649) for support of this project. This work included research conducted at the Advanced Photon Source on the Northeastern Collaborative Access Team beamlines, which are supported by a grant from the National Institute of General Medical Sciences (P41 GM103403) from the National Institutes of Health. This research used resources of the Advanced Photon Source, a U.S. Department of Energy (DOE) Office of Science User Facility operated for the DOE Office of Science by Argonne National Laboratory under Contract No. DE-AC02-06CH11357.

■ REFERENCES

(1) For a recent review on catalytic selective synthesis, see: Mahatthananchai, J.; Dumas, A. M.; Bode, J. W. *Angew. Chem., Int. Ed.* **2012**, *51*, 10954–10990.

(2) For a recent review on chemo- and site-selective derivatizations of natural products, see: Robles, O.; Romo, D. *Nat. Prod. Rep.* **2014**, *31*, 318–334.

(3) Lewis, C. A.; Miller, S. J. *Angew. Chem., Int. Ed.* **2006**, *45*, 5616–5619.

(4) Lewis, C. A.; Chiu, A.; Kubryk, M.; Balsells, J.; Pollard, D.; Esser, C. K.; Murry, J.; Reamer, R. A.; Hansen, K. B.; Miller, S. J. *J. Am. Chem. Soc.* **2006**, *128*, 16454–16455.

(5) Lewis, C. A.; Longcore, K. E.; Miller, S. J.; Wender, P. A. *J. Nat. Prod.* **2009**, *72*, 1864–1869.

(6) Wilcock, B. C.; Uno, B. E.; Bromann, G. L.; Clark, M. J.; Anderson, T. M.; Burke, M. D. *Nat. Chem.* **2012**, *4*, 996–1003.

(7) Ueda, Y.; Mishiro, K.; Yoshida, K.; Furuta, T.; Kawabata, T. *J. Org. Chem.* **2012**, *77*, 7850–7857.

(8) Sculimbrene, B. R.; Morgan, A. J.; Miller, S. J. *Chem. Commun.* **2003**, 1781–1785.

(9) Sculimbrene, B. R.; Xu, Y.; Miller, S. J. *J. Am. Chem. Soc.* **2004**, *126*, 13182–13183.

(10) Fiori, K. W.; Puchlopek, A. L. A.; Miller, S. J. *Nat. Chem.* **2009**, *1*, 630–634.

(11) Sánchez-Roselló, M.; Puchlopek, A. L. A.; Morgan, A. J.; Miller, S. J. *J. Org. Chem.* **2008**, *73*, 1774−1782.
(12) Fowler, B. S.; Laemmerhold, K. M.; Miller, S. J. *J. Am. Chem. Soc.* **2012**, *134*, 9755−9761.
(13) Jordan, P. A.; Miller, S. J. *Angew. Chem., Int. Ed.* **2012**, *51*, 2907−2911.
(14) Beale, T.; Taylor, M. S. *Org. Lett.* **2013**, *15*, 1358−1361.
(15) Zhao, Y.; Rodrigo, J.; Hoveyda, A. H.; Snapper, M. L. *Nature* **2006**, *443*, 67−70.
(16) Sun, X.; Lee, H.; Lee, S.; Tan, K. L. *Nat. Chem.* **2013**, *5*, 790−795.
(17) Gonzalez-Sabin, J.; Moran-Ramallal, R.; Rebolledo, F. *Chem. Soc. Rev.* **2011**, *40*, 5321−5335.
(18) Robbins, D. W.; Hartwig, J. F. *Science* **2011**, *333*, 1423−1427.
(19) Lichtor, P. A.; Miller, S. J. *ACS Comb. Sci.* **2011**, *13*, 321−326.
(20) Pathak, T. P.; Miller, S. J. *J. Am. Chem. Soc.* **2012**, *134*, 6120−6123.
(21) Pathak, T. P.; Miller, S. J. *J. Am. Chem. Soc.* **2013**, *135*, 8415−8422.
(22) Kahne, D.; Leimkuhler, C.; Lu, W.; Walsh, C. T. *Chem. Rev.* **2005**, *105*, 425−448.
(23) Walsh, C. T.; Wencewicz, T. A. *J. Antibiot.* **2014**, *67*, 7−22.
(24) Barna, J. C.; Williams, D. H. *Annu. Rev. Microbiol.* **1984**, *38*, 339−357.
(25) Han, S.; Miller, S. J. *J. Am. Chem. Soc.* **2013**, *135*, 12414−12421.
(26) Economou, N. J.; Zentner, I. J.; Lazo, E.; Jakoncic, J.; Stojanoff, V.; Weeks, S. D.; Grasty, K. C.; Cocklin, S.; Loll, P. J. *Acta Crystallogr.* **2013**, *D69*, 520−533.
(27) Fraudenfelder, H.; Sligar, S. G.; Wolynes, P. G. *Science* **1991**, *254*, 1598−1603.
(28) Economou, N. J.; Nahoum, V.; Weeks, S. D.; Grasty, K. C.; Zentner, I. J.; Townsend, T. M.; Bhuiya, M. W.; Cocklin, S.; Loll, P. J. *J. Am. Chem. Soc.* **2012**, *134*, 4637−4645.
(29) Muir, T. W.; Sondhi, D.; Cole, P. A. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 6705−6710.
(30) For a review on expressed protein ligation, see: Muir, T. W. *Annu. Rev. Biochem.* **2003**, *72*, 249−289.
(31) Dawson, P. E.; Muir, T. W.; Clark-Lewis, I.; Kent, S. B. *Science* **1994**, *266*, 776−779.
(32) Evans, T. C., Jr.; Benner, J.; Xu, M.-Q. *Protein Sci.* **1998**, *7*, 2256−2264.
(33) See the Supporting Information for details.
(34) Awasthi, S. K.; Raghothama, S.; Balaram, P. *Biochem. Biophys. Res. Commun.* **1995**, *216*, 375−381.
(35) Haque, T. S.; Little, J. C.; Gellman, S. H. *J. Am. Chem. Soc.* **1996**, *118*, 6975−6985.
(36) Stanger, H. E.; Gellman, S. H. *J. Am. Chem. Soc.* **1998**, *120*, 4236−4237.
(37) Fisk, J. D.; Powell, D. R.; Gellman, S. H. *J. Am. Chem. Soc.* **2000**, *122*, 5443−5447.
(38) For the crystal structure of structurally relevant peptides, see: Blank, J. T.; Miller, S. J. *Pept. Sci.* **2006**, *84*, 38−47.
(39) Teicoplanin A_2-2 (**1**) adopted an open conformation in our crystal structure. Comparison of our crystal structure to Loll's MBP−peptide and teicoplanin complex structure (ref 26) showed a high degree of structural similarity in the teicoplanin aglycon region but revealed slight differences in the more flexible sugar moieties.
(40) Bilwes, A. M.; Alex, L. A.; Crane, B. R.; Simon, M. I. *Cell* **1999**, *96*, 131−141.
(41) Mizuguchi, H.; Cook, P. F.; Tai, C. H.; Hasemann, C. A.; Uyeda, K. *J. Biol. Chem.* **1999**, *274*, 2166−2175.
(42) Fraser, M. E.; James, M. N.; Bridger, W. A.; Wolodko, W. T. *J. Mol. Biol.* **1999**, *285*, 1633−1653.
(43) Halpern, J. *Science* **1982**, *217*, 401−407.
(44) Solvent residual peak was referenced to 4.79 ppm.
(45) Peptide **S8** was added as an aqueous solution (20.6 mg/mL). Reduced thiol substrate **13** was generated by a thiol−disulfide interchange reaction. Alternatively, dithiothreitol (10 equiv) was added to the peptide **S8** solution before addition to the protein thioester solution.