View Online
Dynamic Article Links

Molecular BioSystems

Cite this: *Mol. BioSyst.*, 2011, **7**, 1224–1231

www.rsc.org/molecularbiosystems

PAPER

Regioselective deacetylation based on teicoplanin-complexed Orf2* crystal structures†

Hsiu-Chien Chan,[abc] Yu-Ting Huang,[a] Syue-Yi Lyu,[a] Chuen-Jiuan Huang,[a] Yi-Shan Li,[a] Yu-Chen Liu,[a] Chia-Cheng Chou,[d] Ming-Daw Tsai*[ab] and Tsung-Lin Li*[ae]

Received 9th December 2010, Accepted 6th January 2011
DOI: 10.1039/c0mb00320d

Lipoglycopeptide antibiotics are more effective than vancomycin against MRSA as they carry an extra aliphatic acyl side chain on glucosamine (Glm) at residue 4 (r4). The biosynthesis of the r4 *N*-acyl Glc moiety at teicoplanin (Tei) or A40926 has been elucidated, in which the primary amine nucleophile of Glm is freed from the r4 GlcNac pseudo-Tei precursor by Orf2* for the subsequent acylation reaction to occur. In this report, two Orf2* structures in complex with β-D-octyl glucoside or Tei were solved. Of the complexed structures, the substrate binding site and a previously unknown hydrophobic cavity were revealed, wherein r4 GlcNac acts as the key signature for molecular recognition and the cavity allows substrates carrying longer acyl side chains in addition to the acetyl group. On the basis of the complexed structures, a triple-mutation mutant S98A/V121A/F193Y is able to regioselectively deacetylate r6 GlcNac pseudo-Tei instead of that at r4. Thereby, novel analogs can be made at the r6 sugar moiety.

Introduction

The emergence of multidrug-resistant Gram-positive pathogens is a serious public health concern.[1] Glycopeptide antibiotics, such as vancomycin and teicoplanin (Tei) (**1**, **2**, Fig. 1), were the drugs of the last resort to treat methicillin-resistant *Staphylococcus aureus* (MRSA) infections. Vancomycin-resistant *Enterococcus* (VRE) and *S. aureus* (VRSA), which are immune to glycopeptide antibiotics, have emerged.[2] New antibiotics with better efficacy against these intractable pathogens therefore are urgently demanded.[3,4]

Glycopeptide antibiotics were known to inhibit the formation of the bacterial cell wall by binding to the *N*-acyl-D-Ala-D-Ala terminal of the lipid II precursor, thereby terminating subsequent transpeptidation and transglycosylation reactions in the late stages of peptidoglycan synthesis. Lipoglycopeptides, Tei and A40926 (**3**, Fig. 1) were reported to be more effective than vancomycin against MRSA, likely as they carry an extra aliphatic acyl side chain on glucosamine (Glm) at residue 4 (r4). Such a modification may enable lipoglycopeptides anchoring themselves on the lipid-layer of the bacterial membrane to be close to the cell wall precursors.[5,6] Modifications of glycopeptide antibiotics at this position were explored and shown hopeful. For examples, telavancin and oritavancin (**4** and **5**, Fig. 1), two semisynthetic derivatives different from vancomycin as they carry an extra decylaminoethyl[7] or *N*-alkyl-*p*-chlorophenylbenzyl[8] moiety on vancosamine at r4, respectively, are more effective against VRE. Additionally, dalbavancin (**6**, Fig. 1), a semisynthetic derivative of A40926, carries an extra short diamine chain at residue 7 of the parent compound and is active against MRSA.[9] Such modifications, *i.e.* addition of a long aliphatic side chain and/or a new modification position, may have shed light for developing new generations of glycopeptide antibiotics.

The biosyntheses of *N*-acyl Glc and GlcA at r4 of Tei and A40926 have been elucidated. For Tei, three gene products Orf10*, Orf2* and Orf11* in the biosynthetic gene cluster of Tei were characterized as glycosyltransferase, deacetylase and acyltransferase, respectively, to generate *N*-acyl Glc pseudo-Tei from the heptapeptide core (**7**) in sequence (Fig. 2A; Fig. S1, ESI†). For A40926, three gene products Dbv9, Dbv21 and Dbv8 in the biosynthetic gene cluster of A40926 correspond to Orf10*, Orf2* and Orf11*, respectively, along with Dbv29, a hexose oxidase, to form *N*-acyl GlcA peudo-A40926 (**3**, Fig. 1). Orf1, a glycosyltransferase encoded in the Tei biosynthetic gene cluster but absent in that of A40926, adds additional

[a] *Genomics Research Center, Academia Sinica, Taipei, Taiwan. E-mail: tlli@gate.sinica.edu.tw; Fax: +886 2 2789 8811; Tel: +886 2 2787 1235*
[b] *Institute of Biological Chemistry, Academia Sinica, Taipei, Taiwan. E-mail: mdtsai@gate.sinica.edu.tw; Fax: +886 2 2788-9759; Tel: +886 2 2785-5696 ext 1013*
[c] *Institute of Bioinformatics and Structural Biology, National Tsing Hua University, Hsinchu, Taiwan*
[d] *National Synchrotron Radiation Research Center, Hsinchu, Taiwan*
[e] *Biotechnology Center, National Chung Hsing University, Taichung, Taiwan*

† Electronic supplementary information (ESI) available: Supplementary text, experimental procedures, figures, and tables. See DOI: 10.1039/c0mb00320d

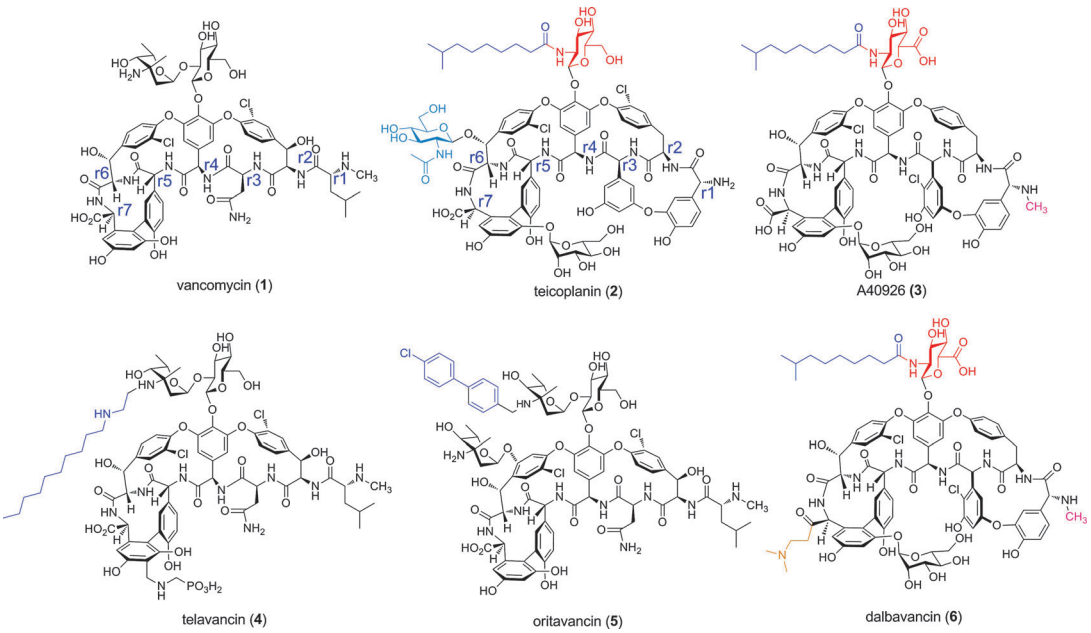

Fig. 1 Structures of glycopeptide antibiotics and derivatives thereof. The structures of vancomycin (**1**), teicoplanin (**2**), A40926 (**3**), telavancin (**4**), oritavancin (**5**), and dalbavancin (**6**). The r6 GlcNac in teicoplanin is colored in light blue. The r4 Glm/GlmA in teicoplanin, A40926 or dalbavancin is colored in red. The aliphatic acyl, aminoacyl, and chlorobiphenyl moieties in individual glycopeptides are colored in blue. The methyl and diamine moieties in A40926 or dalbavancin are colored in magenta and orange, respectively.

GlcNac from UDP-GlcNac to r6 of Tei.[10–14] Of these enzymes, Orf2*/Dbv21 is the key in this modification as it deprotects r4 GlcNac to Glm to enable a subsequent long aliphatic chain addition; it was suggested also to be involved in lipoglycopeptide degradation.[15]

To improve the antimicrobial activity, glycopeptide antibiotics still possess many unexplored positions for further modifications. However, this remains a considerable challenge to medicinal chemists. One way to stride over this hurdle is to take the advantages of the solved biosynthetic pathways of Tei and A40926, both similar in backbone but dissimilar in decorations, by manipulating/engineering the genes and/or gene products in the pathways to create hybrids or new analogs. Here, we attempted to reposition *N*-acyl Glc from r4 to r6 as one of the pursuits. Given that Orf1 can add GlcNac at r6 of the aglycone, a specialized enzyme then is needed to free the primary amine from GlcNac for making new analogs. Orf2*, a stable and efficient enzyme, was chosen to go through the protein engineering with a hope to regioselectively deacetylate GlcNac at r6 rather than at r4. The structure-based approach was used. Thus, solving the Tei-bound Orf2* structure was deemed as the key to the success of the pursuit (Fig. 2C).

Results and discussion

Overall structure

While we were initiating the work, the recently released structures of native and decanoic acid (DAK)-bound Orf2*[16] were not available at that time in addition to lack of the Tei-bound structure. So, the crystal structure‡ of the β-D-octyl glucoside (BOG)-bound Orf2* was first solved to 1.7 Å resolution using the method of multi-wavelength anomalous scattering (MAD) for selenomethionine-derived crystals and refined to an R/R_{free} factor of 0.18/0.23. The first eight amino acids and the chain segments, 112–119 and 238–241, of the protein were excluded in the final structure model as they are unrecognizable in the electron density map. Second, the crystal structure of Tei-bound Orf2* mutant (H164N) was also determined to be 1.7 Å resolution by molecular replacement (MR) using the solved BOG-bound Orf2* structure as the searching template. The first eight amino acids and the chain segment 233–243 of the protein were not contained in the final model, likewise due to lack of well defined electron density. This structure was refined to an R/R_{free} factor of 0.20/0.24. Detailed data statistics is listed in Table 1.

An asymmetric unit of the BOG-bound Orf2* crystal ($P2_1$ space group) contains a single polypeptide chain. In contrast, an asymmetric unit of the Tei-bound Orf2* mutant crystal ($P1$ space group) includes four polypeptide chains (Fig. S2a, ESI†), where each polypeptide chain holds one molecule of Tei. The PDBePISA[17] analysis suggested that Orf2* structures are monomeric in solution as they lack significant surface contacts or interactions.

Both Orf2* structures are made of a single Rossmann-like domain, which includes nine α helices (αA–αI) and nine β strands (β1–β9) interconnected with loops in various lengths. In detail, this domain is composed of a central parallel seven-stranded β sheet (in the order of β3, β2, β1, β6, β7, β9, and β8 except β9 that is in antiparallel) with four helices on one side (αA, αC, αH and αI) and two on the other side (αE and αF) plus the αB helix that runs over the central sheet at its C termini. Such a setting shapes the substrate binding site, in which the central sheet sits at the bottom with α helices and loops extending from the C-termini of the central sheet.

‡ Crystal coordinates can be accessed with 2X9L for BOG-bound Se-derived Orf2* and 2XAD for a Tei-bound Orf2* mutant.

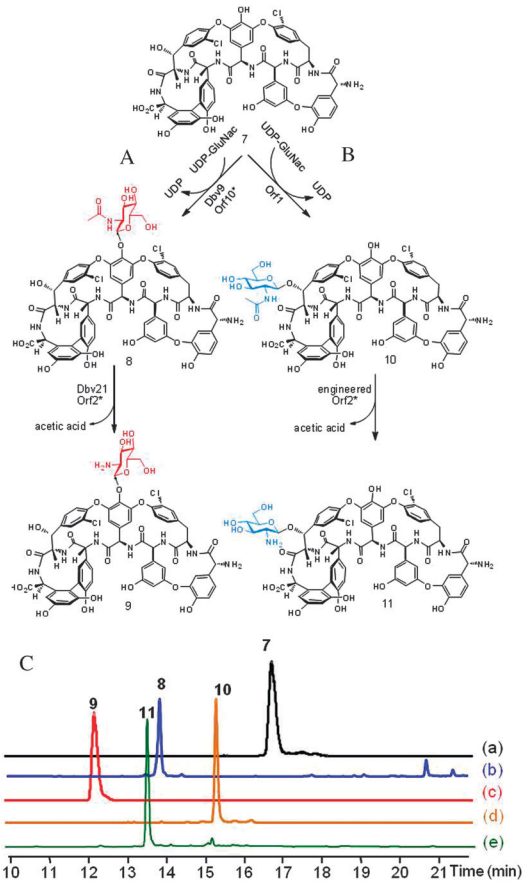

Fig. 2 Synthetic routes and LC traces of enzymatic reactions. (A) First two steps in the biosynthesis of r4 *N*-acyl Glc/*N*-acyl GlcA for Tei/A40926. (B) The proposed synthetic route for r6 Glm pseudo-Tei with the engineered Orf2*. (C) LC traces for: (a) aglycone (**7**); (b) r4 GlcNac pseudo-Tei (**8**), the product of Orf10*; (c) r4 Glm pseudo-Tei (**9**), the product of Orf2*; (d) r6 GlcNac pseudo-Tei (**10**), the product of Orf1; (e) r6 Glm pseudo-Tei (**11**), the product of the engineered Orf2*.

Both structures of BOG-bound Orf2* (Fig. S3a, ESI†) and Tei-bound H164N Orf2* (Fig. 3a) do not cause substantial conformational changes as the rmsd is as low as 0.428 Å for 233 backbone carbons.

Substrate binding site

In the BOG-bound Orf2* electron density map, an extra electron density that does not fit the polypeptide chain was deconvoluted into BOG lipoglycoside (Fig. S3b, ESI†). BOG is adjacent to a zinc ion which is in coordination with residues H16, D19, H164 as well as a water molecule to form a triangular pyramid, wherein the active site in question was identified (see below).

In the Tei-bound H164N Orf2* map, a swath of cohesive electron densities are apparent and well defined, which allowed building of Tei and confined the substrate-binding site (Fig. 3b). Tei was identified to be highly exposed to the solvent, and the substrate binding site is a long but shallow cleft, formed with a two-layer closure. r6 GlcNac and r7 mannose are free in solution with no apparent contacts with protein residues. In contrast, r4 GlcNac is buried deep inside

Table 1 Data collection, phasing and refinement statistics

	SeOrf2*-BOG			H164N-Teico		
Data collection						
Space group	$P2_1$			$P1$		
Cell dimension						
a, *b*, *c*/Å	42.49	64.01	50.30	58.78	70.71	76.31
α, β, γ/°	90.00	99.06	90.00	113.89	108.29	98.20
	Peak	*Inflection*	*Remote*			
Wavelength	0.9809	0.9811	0.9630	0.96		
Resolution/Å	1.75	1.76	1.73	1.7		
R_{sym} or R_{merge}	0.06	0.054	0.052	0.042		
$I/\sigma I$	24.0	17.18	23.6	36.14		
Completeness(%)	97.5	98.0	97.5	95.8		
Redundancy	5.4	2.7	2.7	4.0		
Refinement						
Resolution/Å		27–1.7		20–1.7		
No. reflections		26 082		101 452		
R_{work}/R_{free}		0.18/0.23		0.20/0.24		
No. atoms						
Protein		2270		8040		
Ligand/ion		20/1		528/—		
Water		257		814		
B-factors						
Protein		17.8		23.05		
Ligand/ion		21.6/27.8		32.26/—		
Water		31.1		29.0		
Rms deviations						
Bond lengths/Å		0.004		0.025		
Bond angles/°		1.0		2.3		

the protein. The setting of Tei in the enzyme explains why pseudo-Tei with r6 GlcNac or r7 mannose or both can serve as substrates of Orf2* in addition to r4 GlcNac pseudo-Tei. The complexed structure also revealed that the r4 *N*-acyl Glc moiety plays an important role in molecular recognition: residues D97, S98, H161 and D163 together interact with the

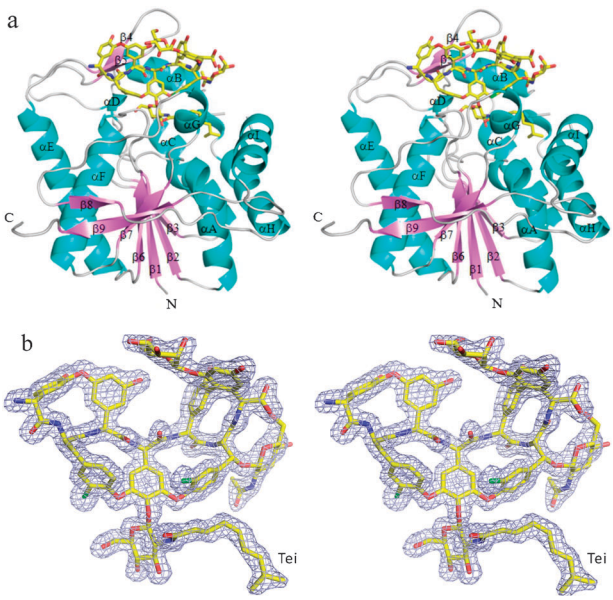

Fig. 3 Overall structure of Tei-bound H164N-Orf2*. (a) Stereo view of the overall Tei-bound H164N-Orf2* structure. Nine α helices (αA–αI; colored in cyan) and nine β strands (β1–β9; colored in magenta) are interconnected by loops in various lengths. Tei is colored in yellow. (b) Stereo view of the $2F_o - F_c$ electron density map of Tei, contoured at 1.0σ (gray mesh).

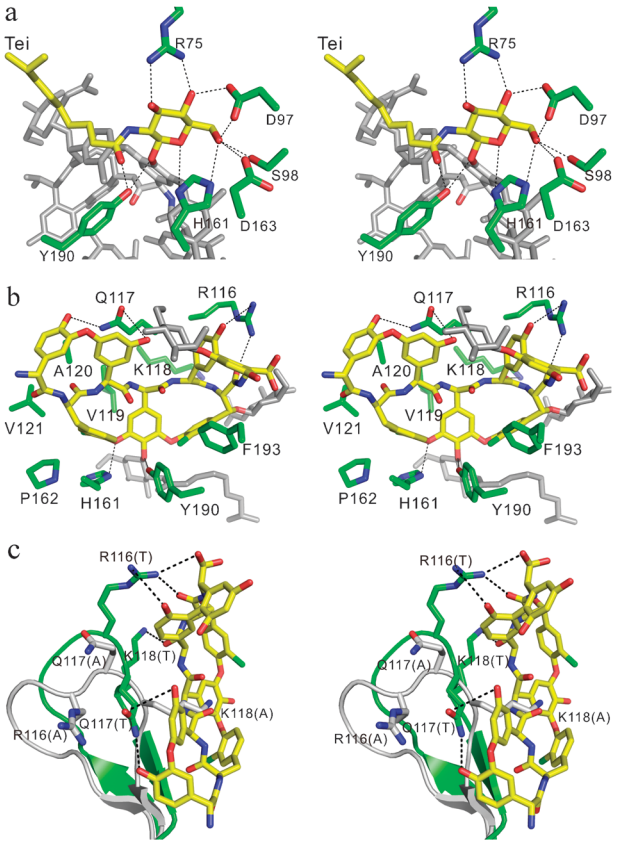

Fig. 4 Binding interactions of Tei with Orf2*. (a) The residues that interact with r4 *N*-acyl Glc in the Tei binding site are indicated and colored in green, in which *N*-acyl Glc and aglycone are colored in yellow and gray, respectively. (b) The residues that interact with aglycone (colored in yellow) in the Tei binding site are indicated and colored in green. (c) The loop regions, residues 110–120, in the presence or absence of Tei are shown by superimposing apo and Tei-bound Orf2* structures, which are colored in gray and green, respectively. The residues in apo and Tei-bound structures are affixed with letters A and T, respectively.

C6 hydroxyl group of *N*-acyl Glc through four hydrogen bonds; residues R75 and D97 interact with the C4 hydroxyl groups through two hydrogen bonds; R75 links with the C3 hydroxyl group through one hydrogen bond; Y190 otherwise is in a close contact with both the oxygen atoms of the *N*-acyl side chain (carbonyl oxygen) and the ether group (O1 oxygen) of Glm through hydrogen bonds (Fig. 4a). As a result, these residues not only recognize the r4 *N*-acyl Glc signature but also anchor the aminosugar in place for the reaction to proceed. As to the aglycone core (Fig. 4b), hydrophobic interactions and hydrogen bonds between the heptapeptide backbone and the polypeptide are the major sources of forces. The hydrophobic interactions are substantiated by residues L119, A120, V121, P162, and F193; the hydrogen bonds are provided by residues R116, Q117, K118, H161, and Y190.

Capping loop

Free and Tei-bound structures were superimposed well with 0.536 of rmsd for 234 backbone carbons, suggesting that Orf2* bound with Tei does not provoke substantial conformational transitions, except the loop region, residues 110–120 (Fig. S2b, ESI†), which was suggested to be a cap in controlling the substrate entry (Fig. 4c).[16] This loop is disordered in BOG-bound Orf2* structures but well defined in the Tei-bound structure. In fact, residues R116 and Q117 are in a close contact with the phenol moieties of r5,6,7 and r1,3 of Tei through 3 and 2 hydrogen bonds, respectively; residue K118 also interacts with the carbonyl oxygen of r6 and r7. These residues, however, flip away in the substrate-free structure. CD analysis agreed with these results (Fig. S4, ESI†): in general, the spectra are well aligned for either the native or the Orf2* mutant (H164N) with/without Tei, except the region of 190–205 nm (a random coil region) where the mean residues ellipticity slightly dropped in the samples added with Tei. By contrast, the mean residues ellipticity does not drop significantly for the double-mutation mutant (R116A and Q117A) added with Tei (Fig. S4, ESI†). In addition, biochemical assays showed no apparent difference in catalytic efficiency between the native and mutants (Table 2). Taken together, the capping loop is able to interact with the substrate but contributes little to the enzyme efficiency.

Catalytic mechanism

By superimposing BOG- and Tei-bound structures, both Glc of BOG and r4 Glm are well matched (Fig. 5a): the C2 oxygen atom of Glc holds the same position as the nitrogen atom of r4 Glm, while the aliphatic chain of BOG takes the spaces that the phenol part of the aglycone core resides (see below). The active site Zn^{2+} ion is located at the bottom of the Tei binding site and is in coordination with three residues H16 (Zn^{2+}–Nδ1, distance = 2.16 Å), D19 (Zn^{2+}–Oδ2, distance = 2.08 Å), and H164 (Zn^{2+}–Nδ1, distance = 2.14 Å) as well as a water molecule (distance from Zn^{2+} of 2.41 Å) to form a triangular pyramid (Fig. 5b). The space of the active site Zn^{2+} ion in the Tei-bound structure is otherwise occupied by a water molecule (Fig. 5c). The Zn^{2+} ion may serve as a Lewis acid together with D18 to activate a water molecule to attack the acyl carbonyl carbon as well as to stabilize the resulting oxyanion transition. Residue Y190 was also identified to put the carbonyl oxygen of the acylamide in position, because the enzyme has no activity for the substrate lacking the oxo group. This structure-based working model was verified by mutagenesis and biochemical assays (Table 2). In short, single mutants H16A, D19A and H164A showed no enzyme activity, which may be as a result of disrupting the active-site coordination chemistry. For the r4 *N*-acyl Glc recognition, single mutants R75Q, D97N, H161A, D163N and Y190F also showed no enzyme activity; interestingly, the S98A mutant however enhanced the activity by 40%.

Lipid cavity

Since both Tei and GlcNac pseudo-Tei were known to be the substrates of Orf2*, a lipid cavity for Tei in the enzyme was speculated.[16] In fact, this lipid cavity was discovered in the Tei-bound structure (Fig. 5d), which is curved and slender. The long aliphatic side chain of Tei fits nicely in the cavity, as it is enclosed by hydrophobic residues V21, L22, L188, P189, Y190, V226, Y229, and L243 as well as stabilized by

Table 2 Activities and roles of key residues in Orf2*

Mutants	(Tei, **2**; %)	(Compound **10**; %)	Role of residue
WT	100	100	Deacylating r4 GlcNac
R75Q	0	nd	Sugar binding/activating D18 for activity
D97N	0	nd	Sugar binding
S98A	142	72	Sugar binding/substrate specificity
R116A	110	77	Capping loop; substrate affinity
Q117A	89	75	Capping loop; substrate affinity
R116A/Q117A	131	55	Capping loop; substrate affinity
A120S	80	97	Substrate specificity
V121A	26	110	Substrate specificity
H161A	0	0	Sugar binding
D163N	0	nd	Sugar binding; interacts with S98
H164N	0	nd	In coordination with zinc
Y190F	0	nd	Oxyanion stabilization
F193Y	392	114	Substrate specificity
V121A/F193Y	126	120	Substrate specificity
S98A/V121A/F193Y	145	130	Substrate specificity
S98A/F193Y	230	120	Substrate specificity

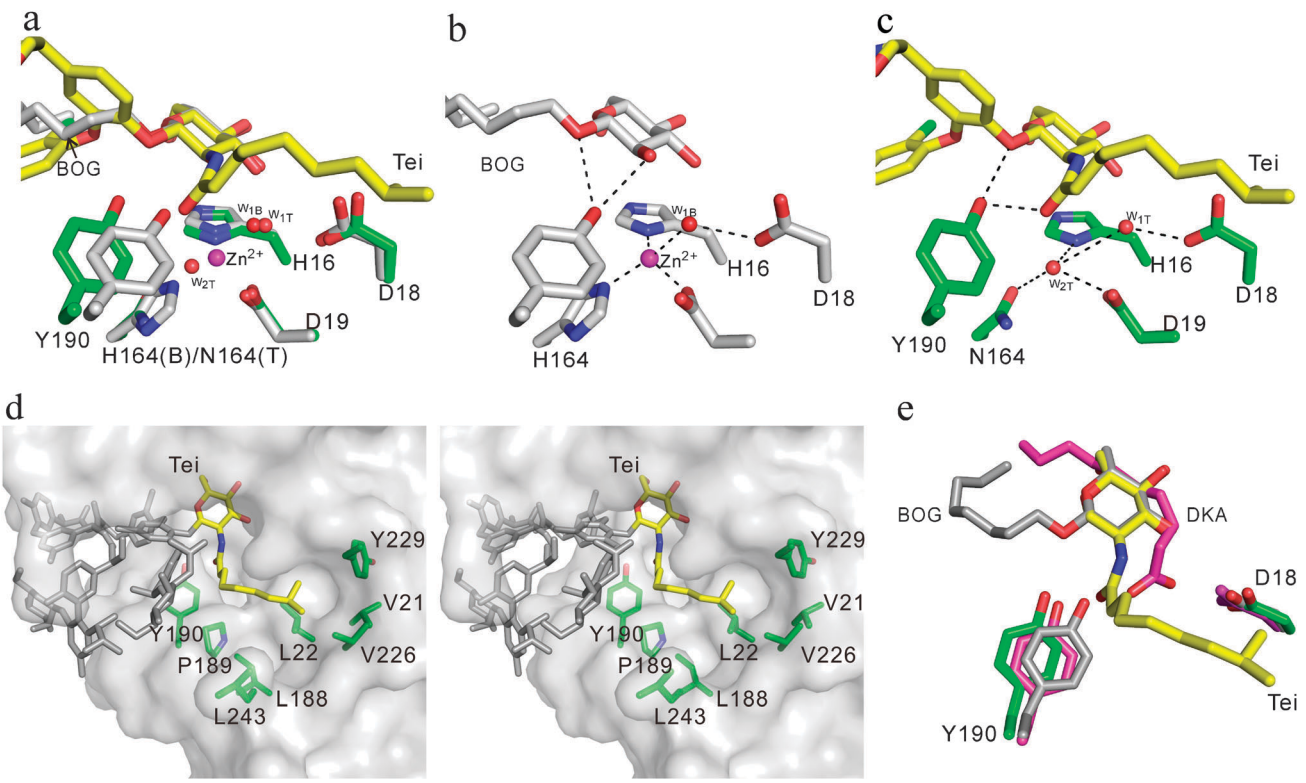

Fig. 5 Active site and lipid cavity of Orf2*. (a) The active site of Orf2*, which is presented by superimposing BOG- and Tei-bound structures. r4 *N*-acyl-Glc of Tei and BOG are colored in yellow and gray, respectively. Catalytically relevant residues are shown in stick: H164(B) and N164(T) represent corresponding residues in BOG- and Tei-bound Orf2* structures, respectively. Zinc ion is shown in magenta, and water molecules in BOG- and Tei-bound are shown in red and affixed with letters B and T, respectively. (b) Close-up of the active site in the BOG-bound Orf2* structure. (c) Close-up of the active site in the Tei-bound H164N-Orf2* structure. (d) Stereo view of the lipid cavity. (e) Superimposed structures for ligands BOG, DKA, and Tei.

van der Waal forces. In contrast, the space (surrounded by residues M59, W63, I99, L119 and V121) that DKA lies is not the lipid cavity suggested by Zou *et al*. (Fig. 5e); it however is the space resided by the sugar and the central tri-phenyl backbone of Tei. As a result, Orf2* can deacylate either a long or a short side chain of *N*-acyl Glc pseudoaglycone. The lipid cavity identified here was further considered to be facilitating the substrate binding. Binding assays were carried out to measure the binding constants by using an isothermal titration calorimeter (ITC) for Orf2* H164N *versus* some prepared compounds. Four pseudoaglycones were enzymatically synthesized, each of which contains a lipid side chain in a given carbon length (C4, 6, 8 and 10).[12–14] The dissociation constants (K_d) for *N*-butyryl- (C4), -hexyl- (C6), -octyl- (C8) and -decanoyl- (C10) Glc Tei-pseudoaglycones were determined to be 14.9, 7.8, 5.4 and 1.3 μM, respectively. In other words,

Table 3 ITC analysis and enzyme kinetics

	Butyryl- (C4)	Hexanoyl- (C6)	Octanoyl- (C8)	Decanoyl- (C10)
K_a	6.7×10^4	1.28×10^5	1.84×10^5	8.32×10^5
K_d/μM	14.9	7.8	5.4	1.3
ΔH/cal mol^{-1}	−8289	−4973	−6755	−4448
ΔS/cal mol^{-1} deg^{-1}	−5.72	6.69	1.43	12.2

the affinity of C10, C8 or C6 to enzyme is 10-, 3-, or 2-fold, respectively, higher than that of C4.[14] Furthermore, kinetic studies showed that the enzyme specificity (k_{cat}/K_m) of C10 is about 10-fold higher than that of C2 (Table 3; Fig. S5, ESI†). Thereby, N-C10 Glc pseudo-Tei was concluded to be a better substrate than the N-C2 counterpart.

Given that Tei can serve as a substrate of Orf2*, an efficient exporting system in the producing strain was reasoned, because a futile cycle of de- and re-acylation may result when the synthesis of Tei is completed in the cell. Alternatively, Orf2* may be involved in Tei degradation/detoxification as a part of self-protection or as a salvage process when Tei is accumulated in excess in the cell. In contrast, a distinct system had evolved in the A40926 biosynthesis, where *N*-acyl GlcA pseudo-A40926 (the product of Dbv29 hexose oxidase) is no longer the substrate of Dbv21 (the Orf2* homolog) so that it avoids mature products from the irregular deacylation.

Structure-based protein engineering

With these insights into the substrate–enzyme interaction, we identified that *N*-acyl Glc at r4 is the key signature in molecular recognition. In order to allow the deacetylation reaction to occur at r6, Tei was rotated by 90° in the Tei-bound Orf2* structure (Fig. S6, ESI†). This rotated ligand was manually optimized from colliding into surrounding residues and was subjected to the CHARMM energy minimization routine. The r4 *N*-acyl Glc moiety determined needs to be removed, because: (1) GlcNac at r6 is less competitive than that at r4 for the same active site when both sugars are present, and (2) a considerable steric hindrance would result after 90° rotation.

r6 GlcNac pseudo-Tei (compound **10**) was first prepared by adding Orf1 to a reaction solution containing aglycone and UDP-GlcNac. When compound **10** (Fig. 2B) was subjected to the Orf2* deacylation reaction, remarkably a new but weak peak emerged on the LC trace (data not shown). Since the deacetylation at r6 of compound **10** by Orf2* is successful, mutagenesis and biochemical assays then were performed (Table 2; column 2).

In this working model, residues R116, Q117, A120, V121, and F193 that may be involved in aglycone binding were mutated into R116A, Q117A, A120S, V121A and F193Y, respectively. Interestingly, these mutants carry various extents of activity (Table 2), in which V121A (110%) and F193Y (114%) showed improved activities. V121A located at the capping loop region may favor the re-orientated substrate or facilitate substrate reorientation to help r6 GlcNac towards the active site. F193Y may interact with the r6 chlorine through the hydroxyl group in tyrosine for better anchoring the substrate. Next, we tested the combination of these

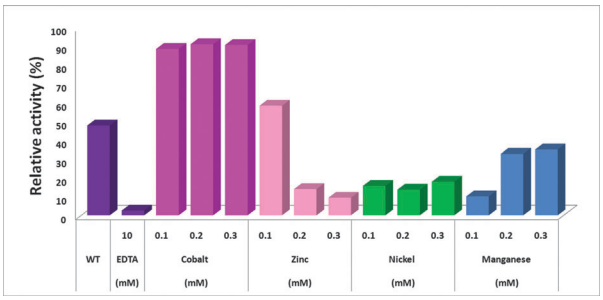

Fig. 6 Effects of selected divalent metal ions on Orf2* enzyme activity. The catalytic efficiency was determined by deacetylating r4 GlcNac pseudoTei. Orf2* was pre-treated with EDTA (10 mM) to rid off the zinc ion (purple). Only can cobalt (magenta), zinc (pink), nickel (green) and manganese (blue) restore the enzyme activity (see text).

mutations. Double mutation V121A/F193Y or S98A/F193Y showed an improved activity by 20% in comparison to WT; the triple mutation (S98A/V121A/F193Y) surprisingly showed an additional 30% improvement, likely by a synergistic effect (Fig. S7–S9, ESI†).

Orf2* would lose its activity once Zn^{2+} is removed from the active site. So, Orf2* is a Zn^{2+}-dependent metalloenzyme. To further enhance the enzyme efficiency, Zn^{2+} replacement with a different metal ion was attempted. To rid Zn^{2+} off the enzyme requires rigorous EDTA (10 mM) and rounds of dialysis treatments, because Zn^{2+} is tightly bound with the enzyme. An array of divalent metal ions (Ca^{2+}, Mg^{2+}, Fe^{2+}, Cu^{2+}, Co^{2+}, Zn^{2+}, Ni^{2+}, Mn^{2+}) then were examined. The protein that carried either Ca^{2+}, Mg^{2+}, Fe^{2+}, or Cu^{2+} did not show any detectable enzyme activity. However, the protein containing either Co^{2+}, Ni^{2+} or Mn^{2+} can restore the activity in the activity order of Co^{2+} > Zn^{2+} > Mn^{2+} > Ni^{2+} (Fig. 6). Interestingly, Orf2*–Zn^{2+} was found sensitive to the concentration of zinc; the enzyme activity is inversely related to the concentration of zinc. In contrast, the enzyme activity is proportional to the concentrations of cobalt, nickel and manganese. Remarkably, the triple mutant (S98A/V121A/F193Y) plus Co^{2+} provides a 2-fold increase in efficiency. Now, the deacetylated compound **11** (Fig. 2B; Fig. S10, ESI†) can be produced in quantity.

Conclusions

The emergence of multi-drug resistant bacteria has alerted us that there is not much time left to prepare more effective warheads to combat drug-resistant pathogens. In this study, we solved the ligand-bound Orf2* crystal structures, located the substrate binding site and identified the previously unknown hydrophobic cavity in the enzyme. With this information, we were able to expand the new activity for

the enzyme—the activation of glucosamine at residue 6 of teicoplanin pseudoaglycone, by which new Tei analogs can be made. Importantly, this would not be possible without the high-resolution ligand-bound crystal structures.

Experimental

Protein expression and purification

Wild-type and SeMet-substituted proteins were produced in an *Escherichia coli* expression system according to the previous description[18] and induced with 1 mM IPTG for 16 h at 16 °C. The lysate was purified by centrifugation and the resulting supernatant was loaded onto a Ni^{2+}-affinity column equilibrated in a buffer containing 20 mM Tris (pH 8.0), 500 mM KCl, 10% glycerol and 10 mM imidazole. Protein was eluted using a gradient of imidazole (10–400 mM). The eluted protein was further purified by gel-filtration chromatography on a Superdex200 column (Amersham Biosciences) in 20 mM Tris (pH 8.0), and 100 mM NaCl. Purity of the protein was verified by SDS-PAGE.

Mutagenesis

Site-directed mutagenesis was carried out by using QuickChange® and the wild-type Orf2*. All mutations were confirmed by DNA sequencing. Mutant enzymes were expressed and purified with the same protocol as wide-type protein. Primers used for corresponding mutants are shown in Table S1 (ESI†).

Enzymatic activity assay

The deacetylation was determined by HPLC. The assay mix containing enzyme (20 μg) and the corresponding substrate (0.5 mM) in buffer (50 mM HEPES, pH 7.2, 100 mM NaCl, 1 mM DTT) (total volume 100 μl) was incubated for 3 h at 30 °C. Each reaction mixture was then centrifuged at 16 000*g* for 5 min (Heraeus Biofuge Pico) and filtered on an ultra-centrifugal filter unit (5 kDa cut-off membrane, Millipore). The filtrate was directly subjected to HPLC-ESI/Q-Tof (Waters HPLC 2695 interfaced with a ESI source coupled to a Micromass micro Q-Tof mass spectrometer) or HPLC-ESI-LTQ (Agilent 1200 Series interfaced with a ESI source coupled to a Thermo-Finnigan LTQ-XL ion trap spectrometer), using a gradient of 0 to 60% acetonitrile in 0.1% TFA in water over 30 min. On-line LC-MS spectra were recorded by MassLynx™ or Xcalibur™.

ITC analysis

Isothermal titration calorimetry (ITC) experiments[19] were performed at 25 °C using an iTC_{200} microcalorimeter from MicroCal, Inc. (Northampton, MA). The final buffer solution for ITC experiments was 10 mM HEPES (pH 7.5). The Orf2* or Orf2* mutant (H164N) was prepared as above and rebuffered. Titrations consisted of 16 injections of 2 μl and were separated by 150 s. The cell stirring speed was 1000 rpm. Each titration contained 0.1 mM enzyme in the sample cell (250 μl). Teicoplanin analog solutions (1 mM) were loaded in the injection syringe (40 μl). The reference cell was filled with water throughout all experiments. The heat of dilution was determined by titrating the ligand into the buffer without enzyme and subtracted prior to curve fitting. Data analysis was done with the Origin 7.0 software provided by MicroCal.

Crystallization and data collection

Different shapes of crystals were formed for selenomethionine (SeMet)-substituted and BOG-bound Orf2* and teicoplanin-complexed H164N-Orf2* mutants by the hanging drop method at 277 K. Crystallization conditions for each were provided as follows: for the SeMet-substituted Orf2*, the crystallization solution included SeMet-substituted Orf2* (25 mg ml^{-1}), HEPES (10 mM, pH 7.5), polyethylene glycol (PEG, 10 000 precipitant (22–30% (w/v)) and *N*-octyl-β-D-glucoside (0.05%). For the teicoplanin-complexed Orf2* mutant (H164N), the crystallization solution included the mutated Orf2* (5 mg ml^{-1}), HEPES (10 mM, pH 8.0), teicoplanin (0.5 mM), PEG4000 (25% w/v), potassium iodide (0.2 M) and MES (0.1 M). Prior to data collection, crystals were transferred to cryoprotectant solutions that were made of individual mother liquids plus 20% glycerol. Multi-wavelength anomalous dispersion (MAD) data sets for SeMetsubstituted Orf2* and H164N-Orf2* were all collected to 1.7 Å resolution on an ADSC Q315 detector using the synchrotron radiation X-ray source at beamline 13B1 of the National Synchrotron Radiation Research Center (NSRRC) in Taiwan. Diffraction intensities were integrated and scaled with programs DENZO and SCALEPACK in the HKL2000 package.[20]

Structure determination and refinement

Multi-wavelength anomalous dispersion (MAD) was used to obtain phase information, for which the methionine residues in Orf2* were substituted with SeMet. Three selenium sites in Orf2* were identified with the program SOLVE.[21] These heavy atom sites were used to calculate the initial phase using the program RESOLVE,[22] yielding an interpretable map for a single protein in an asymmetric unit at space group $P2_1$. Iterative cycles of manual rebuilding with XtalView[23] along with refinements with CNS[24,25] and PHENIX[26] resulted in a model at 1.7 Å resolution with $R = 0.18$ and $R_{free} = 0.23$. Water molecules were added with a water-pick routine in the CNS program. The H164N mutant structure in complex with teicoplanin was solved by molecular replacement using the SeMet substituted Orf2* as the search model. The model was further refined with the CCP4i package.[27] The resulting model was refined at 1.7 Å resolution with $R = 0.20$ and $R_{free} = 0.24$. No residues appear in the disallowed region in Ramachandran plots for all protein structures. Structure presentations were generated using PyMol.[28] Detailed refinement statistics are given in Table 1.

Acknowledgements

We thank Academia Sinica and National Science Council of Taiwan (96-2628-B-001-026-MY3 and 98-2311-B-001-014-MY3 to T.L.L.) for the financial support. Portions of this research were carried out at the National Synchrotron Radiation Research Center, Taiwan. The Synchrotron

Radiation Protein Crystallography Facility is supported by the National Research Program.

Notes and references

1 H. Breithaupt, *Nat. Biotechnol.*, 1999, **17**, 1165–1169.
2 G. M. Anstead and A. D. Owens, *Curr. Opin. Infect. Dis.*, 2004, **17**, 549–555.
3 C. A. Arias and B. E. Murray, *N. Engl. J. Med.*, 2009, **360**, 439–443.
4 D. J. Payne, M. N. Gwynn, D. J. Holmes and D. L. Pompliano, *Nat. Rev. Drug Discovery*, 2007, **6**, 29–40.
5 M. A. Cooper and D. H. Williams, *Chem. Biol.*, 1999, **6**, 891–899.
6 S. D. Dong, M. Oberthür, H. C. Losey, J. W. Anderson, U. S. Eggert, M. W. Peczuh, C. T. Walsh and D. Kahne, *J. Am. Chem. Soc.*, 2002, **124**, 9064–9065.
7 D. L. Higgins, R. Chang, D. V. Debabov, J. Leung, T. Wu, K. M. Krause, E. Sandvik, J. M. Hubbard, K. Kaniga, D. E. Schmidt-Jr, Q. Gao, R. T. Cass, D. E. Karr, B. M. Benton and P. P. Humphrey, *Antimicrob. Agents Chemother.*, 2005, **49**, 1127–1134.
8 T. I. Nicas, D. L. Mullen, J. E. Flokowitsch, D. A. Preston, N. J. Snyder, M. J. Zweifel, S. C. Wilkie, M. J. Rodriguez, R. C. Thompson and R. D. Cooper, *Antimicrob. Agents Chemother.*, 1996, **40**, 2194–2199.
9 P. K. Linden, *Expert Rev. Anticancer Ther.*, 2008, **6**, 917–928.
10 S. Donadio and M. Sosio, *Curr. Top. Med. Chem.*, 2008, **8**, 654–666.
11 S. Donadio, M. Sosio, E. Stegmann, T. Weber and W. Wohlleben, *Mol. Genet. Genomics*, 2005, **274**, 40–50.
12 J. Y. Ho, Y. T. Huang, C. J. Wu, Y. S. Li, M. D. Tsai and T. L. Li, *J. Am. Chem. Soc.*, 2006, **128**, 13694–13695.
13 T. L. Li, F. Huang, S. F. Haydock, T. Mironenko, P. F. Leadlay and J. B. Spencer, *Chem. Biol.*, 2004, **11**, 107–119.
14 Y. S. Li, J. Y. Ho, C. C. Huang, S. Y. Lyu, C. Y. Lee, Y. T. Huang, C. J. Wu, H. C. Chan, C. J. Huang, N. S. Hsu, M. D. Tsai and T. L. Li, *J. Am. Chem. Soc.*, 2007, **129**, 13384–13385.
15 A. W. Truman, L. Robinson and J. B. Spencer, *ChemBioChem*, 2006, **7**, 1670–1675.
16 Y. Zou, J. S. Brunzelle and S. K. Nair, *Chem. Biol.*, 2008, **15**, 533–545.
17 E. Krissinel and K. Henrick, *J. Mol. Biol.*, 2007, **372**, 774–797.
18 W. A. Hendrickson, J. R. Horton and D. M. LeMaster, *EMBO J.*, 1990, **9**, 1665–1672.
19 Y. T. Huang, S. Y. Lyu, P. H. Chuang, N. S. Hsu, Y. S. Li, H. C. Chan, C. J. Huang, Y. C. Liu, C. J. Wu, W. B. Yang and T. L. Li, *ChemBioChem*, 2009, **10**, 2480–2487.
20 Z. Otwinowski and W. Minor, in *Methods in enzymology, macromolecular crystallography, part A*, ed. C. W. C. Jr and R. M. Sweet, Academic Press, 1997, pp. 307–326.
21 T. C. Terwilliger and J. Berendzen, *Acta Crystallogr., Sect. D: Biol. Crystallogr.*, 1999, **55**, 849–861.
22 T. C. Terwilliger, *Acta Crystallogr., Sect. D: Biol. Crystallogr.*, 2003, **59**, 38–44.
23 D. E. McRee, *J. Struct. Biol.*, 1999, **125**, 156–165.
24 A. T. Brünger, P. D. Adams, G. M. Clore, W. L. DeLano, P. Gros, R. W. Grosse-Kunstleve, J.-S. Jiang, J. Kuszewski, M. Nilges, N. S. Pannu, R. J. Read, L. M. Rice, T. Simonson and G. L. Warren, *Acta Crystallogr., Sect. D: Biol. Crystallogr.*, 1998, **54**, 905–921.
25 G. N. Murshudov, A. A. Vagin and E. J. Dodson, *Acta Crystallogr., Sect. D: Biol. Crystallogr.*, 1997, **53**, 240–255.
26 P. D. Adams, P. V. Afonine, G. Bunkoczi, V. B. Chen, I. W. Davis, N. Echols, J. J. Headd, L.-W. Hung, G. J. Kapral, R. W. Grosse-Kunstleve, A. J. McCoy, N. W. Moriarty, R. Oeffner, R. J. Read, D. C. Richardson, J. S. Richardson, T. C. Terwilliger and P. H. Zwart, *Acta Crystallogr., Sect. D: Biol. Crystallogr.*, 2010, **66**, 213–221.
27 COLLABORATIVE COMPUTATIONAL PROJECT NUMBER 4, *Acta Crystallogr., Sect. D: Biol. Crystallogr.*, **50**, 1994, 760–763.
28 W. L. DeLano The PyMOL molecular graphics system ⟨http://www.pymol.org⟩ (DeLano Scientific).