

國立臺灣大學生命科學院生化科學研究所



博士論文

Graduate Institute of Biochemical Sciences

College of Life Science

National Taiwan University

Doctoral Dissertation

醣胜肽抗生素生合成酵素蛋白晶體結構及反應機制為基礎的  
新化學結構生物活性分子設計

Structure- and mechanism-based enzyme design for biologically  
active new chemical entities on glycopeptide antibiotics

劉祐禎

Yu-Chen Liu

指導教授：李宗璘 博士

蔡明道 博士

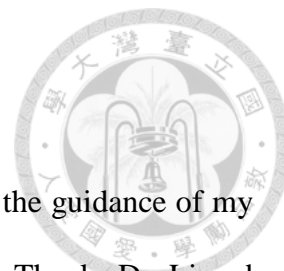
Advisor: Tsung-Lin Li, Ph.D.

Ming-Daw Tsai, Ph.D.

中華民國 102 年 7 月

July, 2013

## Acknowledgement



First, I would like to appreciate my advisor Dr. Tsung-Lin Li for the guidance of my Ph.D. research. He provides unlimited support to my Ph.D. training. Thanks Dr. Li and Dr. Ming-Daw Tsai for providing supervision on my research work. I also thank committee members Dr. Po-Huang Liang, Dr. Che Ma, and Dr. Su-Chang Lin for their efforts on my thesis evaluation as well as many valuable suggestions and comments. With all these professors' help I can complete this dissertation.

Second, I would like to acknowledge my lab members Ms. Yi-Shan Li, Mr. Syue-Yi Lyu, Mr. Li-Jen Hsu, Dr. Yu-Hou Chen, Ms. Yu-Ting Huang, Dr. Hsiu-Chien Chan, Ms. Chuen-Jiuan Huang, Dr. Gan-Hong Chen, Dr. Chia-Cheng Chou, Ms. Xiao-Wei Zou, Mr. Ning-Shian Hsu, Dr. Chin-Yuan Chang, Mr. Hsien-Wei Yeh and Mr. Kuan-Hung Lin for their help on my experiments. Especially thank Yi-Shan for her biochemical experiments, Syue-Yi for his MIC assay, Li-Jen for his animal study, Yu-Hou for his AUC analysis in Dbv29 project. For MppJ project, I especially appreciate Xiao-Wei for her mutagenesis and activity assay, Ning-Shian for his chiral analysis and binding affinity assay. I cannot finish my experimental works by myself without their assistance.

Finally, I would like to thank my parents and my wife Ms. Pei-Ju Wang for their support and accompany. They are always on my side sharing my happiness and sadness. With their encouragement I can graduate from school within four years. Now I get my Ph.D. degree and I will move on to the next stage as a postdoctoral research fellow. I hope they will be proud of me and that I can share my success with them in the near future.

**Section 1. Interception of teicoplanin oxidation intermediates yields new antimicrobial scaffolds.**

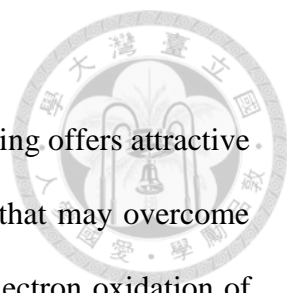


## 摘要

細菌抗藥性問題日益嚴重，使得發展新型抗生素更加困難，藉由改變抗生素合成路徑而改變抗生素結構，提供我們一個可能克服細菌抗藥性問題的機會。Dbv29 為參與抗生素 A40926 生合成之六碳糖氧化酶，利用 X 光蛋白質結晶學與生物化學的方法，我們解析出 Dbv29 的蛋白質結構以及催化作用機制，並且發現一對酪胺酸活性基團在輔酶蛋白共價結合、酵素活性及維持蛋白質結構上扮演非常重要的角色。更特別的是，受質在 Dbv29 與 teicoplanin 複合體蛋白質結構中以反應中間產物呈現，使我們得以進一步利用 Dbv29 合成各種不同化學結構的衍生物。在抑菌測試中，部分的衍生物對於具有抗藥性的腸球菌(VRE)比抗生素 vancomycin 與 teicoplanin 表現出更好的抑菌效果。因此利用這嶄新的方法我們可以發展出更多不同型態的抗生素衍生物來解決細菌抗藥性問題。

關鍵字: 醣肽類抗生素、teicoplanin 衍生物、氧化還原酶、還原胺化、抗萬古黴素腸球菌(VRE)。

## Abstract



In the search for new efficacious antibiotics, biosynthetic engineering offers attractive opportunities to introduce minor alterations to antibiotic structures that may overcome resistance. Dbv29, a flavin-containing oxidase, catalyzes the four-electron oxidation of a vancomycin-like glycopeptide to yield A40926. Structural and biochemical examination of Dbv29 now provides insights into residues that govern flavinylation and activity, protein conformation and reaction mechanism. In particular, the serendipitous discovery of a reaction intermediate in the crystal structure led us to identify an unexpected opportunity to intercept the normal enzyme mechanism at two different points to create new teicoplanin analogs. Using this method, we synthesized families of antibiotic analogs with amidated and aminated lipid chains, some of which showed marked potency and efficacy against multidrug resistant pathogens. This method offers a new strategy for the development of chemical diversity to combat antibacterial resistance.

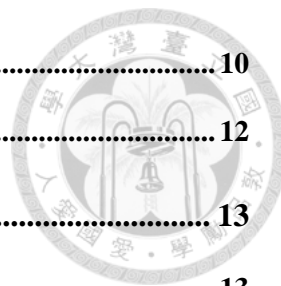
Keywords: glycopeptide antibiotics, teicoplanin analogs, oxidoreductase, reductive amination, vancomycin-resistant *Enterococcus* (VRE).

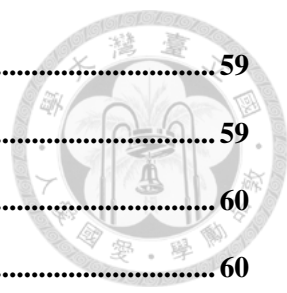


## Table of Contents

Verification letter from the oral examination committee.....	i
Acknowledgement.....	ii
Section 1. Interception of teicoplanin oxidation intermediates yields new antimicrobial scaffolds. ....	iii
Chinese abstract.....	iv
Abstract .....	v
Table of Contents.....	vi
List of Figures.....	ix
List of Tables.....	x
<b>1. Introduction .....</b>	<b>1</b>
1.1 Emerging of resistant pathogens .....	1
1.2 Glycopeptide antibiotics .....	1
1.3 Modifications of glycopeptide antibiotics by biosynthetic enzymes .....	3
1.4 Dbv29 .....	3
<b>2. Materials and Methods .....</b>	<b>6</b>
2.1 Gene cloning and protein purification .....	6
2.2 Crystallization and data collection .....	6
2.3 Structure determination and refinement.....	7
2.4 Enzymatic activity assay.....	8
2.5 Mutagenesis .....	8
2.6 Circular dichroism spectroscopy.....	8
2.7 Analytical ultracentrifuge analysis.....	9
2.8 Synthetic conditions for new analogs .....	9

2.9 Compound characterization .....	10
2.10 <i>In vivo</i> study.....	12
<b>3. Results.....</b>	<b>13</b>
3.1 Dbv29 is a FAD-containing dimer .....	13
3.2 Contributions to cofactor- and substrate-binding sites .....	13
3.3 A diol intermediate leads to catalytic redirection .....	15
3.4 Analogs provide alternate antimicrobial protection.....	17
3.5 Dbv29 has a dynamic quaternary structure .....	18
<b>4. Discussion .....</b>	<b>20</b>
<b>References.....</b>	<b>48</b>
<b>Section 2. Structure and mechanism of non-heme iron-SAM dependent methyltransferase and its engineering to hydratase.....</b>	<b>xi</b>
<b>Chinese abstract.....</b>	<b>xii</b>
<b>Abstract .....</b>	<b>xiii</b>
<b>List of Figures.....</b>	<b>xiv</b>
<b>List of Tables.....</b>	<b>xv</b>
<b>1. Introduction .....</b>	<b>52</b>
1.1 Mannopeptimycin .....	52
1.2 Methyltransferases and methylation.....	54
1.3 MppJ .....	55
<b>2. Materials and Methods .....</b>	<b>57</b>
2.1 Gene cloning .....	57
2.2 Protein expression and purification .....	57
2.3 Crystallization and data collection .....	58
2.4 Structure determination and refinement.....	58





2.5 Enzymatic reaction .....	59
2.6 Site-directed mutagenesis.....	59
2.7 Analytical ultracentrifugation analysis.....	60
2.8 Isothermal titration calorimetry analysis .....	60
2.9 X-ray absorption spectroscopy .....	60
2.10 Electron paramagnetic resonance spectroscopy .....	61
2.11 Chiral HPLC analysis.....	61
<b>3. Results.....</b>	<b>62</b>
3.1 Protein crystallization and structure determination .....	62
3.2 Overview of the structure.....	63
3.3 Conformational change and SAM/SAH-binding .....	64
3.4 Ppy-binding site and coordination chemistry.....	65
3.5 Reaction mechanism .....	66
3.6 No non-heme oxygenase activity.....	68
3.7 Stereochemistry.....	70
3.7.1 Methylation.....	70
3.7.2 Post-translational modification.....	71
3.7.3 Methoxylation.....	72
<b>4. Discussion .....</b>	<b>73</b>
<b>References.....</b>	<b>93</b>
<b>Appendix .....</b>	<b>xvi</b>



## List of Figures



<b>Figure 1. Structures of relevant glycopeptide antibiotics. ....</b>	<b>21</b>
<b>Figure 2. The structure of Dbv29 and its enzymatic mechanism.....</b>	<b>22</b>
<b>Figure 3. Intermediate trapping strategies to probe the catalytic mechanism. ....</b>	<b>23</b>
<b>Figure 4. Variations in acyl chain length regulate the resultant glycopeptide analogue.....</b>	<b>24</b>
<b>Figure 5. Mice infection test for analog 25 in blood bacterial clearance.....</b>	<b>25</b>
<b>Figure 6. Structural analysis of dimeric and monomeric Dbv29. ....</b>	<b>26</b>
<b>Figure 7. Analytical ultracentrifugation (AUC) analyses for Dbv29 and mutants.</b>	<b>28</b>
<b>Figure 8. Circular dichroism (CD) analyses for Dbv29 and mutants.....</b>	<b>30</b>
<b>Figure 9. CD analysis of protein stability.....</b>	<b>31</b>
<b>Figure 10. MS and MSMS spectra of the benzylamine-Tei aminated analog (25).</b>	<b>32</b>
<b>Figure 11. NMR information for de-r4-Tei. NMR spectra include <math>^1\text{H}</math>, <math>^{13}\text{C}</math>, HSQC, HMBC, and COSY. ....</b>	<b>35</b>
<b>Figure 12. NMR information for analog 25. NMR spectra include <math>^1\text{H}</math>, <math>^{13}\text{C}</math>, HSQC, HMBC, and COSY. ....</b>	<b>38</b>
<b>Figure 13. HRMS data for representative compounds. ....</b>	<b>39</b>



## List of Tables

<b>Table 1. MICs of vancomycin, teicoplanin and analogs against tested strains. ....</b>	<b>40</b>
<b>Table 2. Data collection and refinement statistics for Dbv29 structures. ....</b>	<b>41</b>
<b>Table 3. Relative enzymatic activities of Dbv29 mutants. ....</b>	<b>42</b>
<b>Table 4. Product yields of oxidized, aminated or amidated teicoplanin analogs synthesized by Dbv29. ....</b>	<b>44</b>
<b>Table 5. NMR assignments for compound 25 as determined from spectra shown in Figures 11 and 12. ....</b>	<b>47</b>



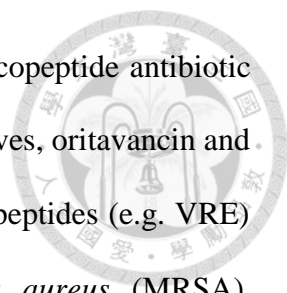
## 1. Introduction

### 1.1 Emerging of resistant pathogens

The rise of bacterial resistance against long-standing drugs of last resort, such as vancomycin, has created an urgent demand for the development of new antibiotics with enhanced or broadened antimicrobial efficacy<sup>1-5</sup>. Recent strategy for finding new classes of antibacterial compounds based on targets identified from bacterial genomics has not yet proved successful<sup>6</sup>. This has led us to consider rationally manipulating potential natural products. Natural products have historically been invaluable as a source of antibacterial drugs, such as penicillins, macrolides and glycopeptides, reflecting their evolutionary origin as ‘weapons’ that bacteria use against each other. Glycopeptide, vancomycin, had long been known as the ‘last resort’ in treating serious infections caused by bacteria resistant to several other antibiotics, such as the ‘super bug’ methicillin-resistant *Staphylococcus aureus* (MRSA), but even vancomycin resistance has now emerged, further underlining the need for new antibiotics. Thus, compounds with activity against clinically important Gram-positive pathogens, including antibiotic-resistant strains such as MRSA, penicillin resistant *Streptococcus pneumoniae*, and vancomycin resistant *Enterococci* (VRE) are highly demanded<sup>7,8</sup>.

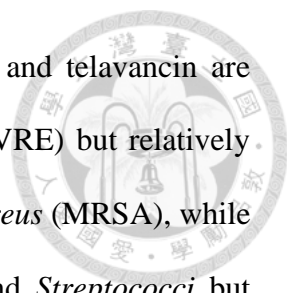
### 1.2 Glycopeptide antibiotics

Vancomycin (Figure 1) that has outpaced multidrug resistance of methicillin-resistant *Staphylococcus aureus* (MRSA) for decades is facing mighty backlashes by emerging vancomycin-intermediate *S. aureus* (VISA) and vancomycin-resistant *S. aureus* (VRSA) in addition to having lost battle to vancomycin-resistant *Enterococci* (VRE). Dalbavancin (Figure 1), a new member of the vancomycin family antibiotics known for active against *Staphylococci* (e.g. MRSA) and *Streptococci* while weak against



*Enterococci* (e.g. VRE), was developed based on the drug lead glycopeptide antibiotic A40926 (Figure 1). Two other renowned vancomycin-based derivatives, oritavancin and telavancin, are active against strains resistant to conventional glycopeptides (e.g. VRE) but relatively weak against methicillin-resistant *Staphylococcus aureus* (MRSA). Seeking new compounds with enhanced and/or broader antimicrobial spectrum for refractory bacterial infections are never over demanded. A40926 differs from teicoplanin (Tei, **1**, Figure 1) mainly in lacking the *N*-Ac glucosaminyl on residue 6 (r6) and the replacement of the *N*-acyl glucosaminyl C-6 OH substituent on residue 4 (r4) with a COOH group (Figure 1). These structural features make it an ideal system to exploit the manipulation of genes and enzymes involved in the biosynthetic pathways of both antibiotics to create hybrid structures and new analogs. The close structural similarity of vancomycin and related glycopeptide antibiotics including A40926 and teicoplanin have prompted a variety of studies to determine whether late-stage tailoring enzymes that act on one of the family members will cross-react with other family members to create new hybrid analogs; these strategies have yielded several compounds with altered and improved antimicrobial profiles, demonstrating the utility of this approach. Though many hybrid structures have been developed using these methods or have been identified from natural sources, modifications at the sugars of the residue 4 and 6 are rare.

Three new glycopeptides are currently close to common clinical usage, namely, oritavancin (Targanta Therapeutics Inc) and telavancin (Theravance Inc/Astellas Pharm Inc) which are analogs of vancomycin, and dalbavancin (Pfizer Inc) (Figure 1) which in fact is a A40926 analog. All three compounds have some commonality in structure but with difference in pharmacokinetic and pharmacodynamic properties and in vitro spectrum. In general, the three compounds all have been well tolerated in clinical trials



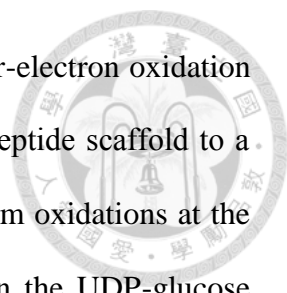
to date but with pros and cons individually. In short, oritavancin and telavancin are active against strains resistant to conventional glycopeptides (e.g. VRE) but relatively weak against strains such as methicillin-resistant *Staphylococcus aureus* (MRSA), while dalbavancin features active against *Staphylococci* (e.g. MRSA) and *Streptococci* but weak against *Enterococci* (e.g. VRE) or *Staphylococci* harboring the vanA gene cluster.

### 1.3 Modifications of glycopeptide antibiotics by biosynthetic enzymes

Previous reports have shown that even modest structural modifications of vancomycin can overcome resistance. We have elucidated the biological functions for enzymes Dbv8, 9, 21, and 29. They together were characterized to be responsible for the synthesis of the N-acyl glucosamine moiety of A40926 in the filamentous *Actinomycete Nonomuraea* sp. ATCC39727. Given the “diversity-oriented synthesis” and “synthesis on privileged compound” as new concepts in the field, we are keen to build up a useful platform using native and engineered Dbv8, 9, 21, and 29 as biocatalysts. New analogs by modifications of teicoplanin will be dissimilar to any reported compound (they are almost vancomycin and A40926 derivatives). New analogs demonstrated superior in anti-VRE activity will be expected to be at least not less than oritavancin and/or telavancin to the given strains tested and they are very likely to be more effective to some oritavancin- and telavancin-insensitive strains. As a consequence, diversification of the “privileged compound” teicoplanin may help conquer bacterial resistance and ease the strong demand for new antibiotics.

### 1.4 Dbv29

Though many hybrid structures have been developed using these methods or have been identified from natural sources, modifications at the C6 position are rare. We have been interested in Dbv29, a hexose oxidase that catalyzes the last step in A40926 biosynthesis in the filamentous *Actinomycete Nonomuraea* sp. ATCC39727<sup>9, 10</sup>. Our

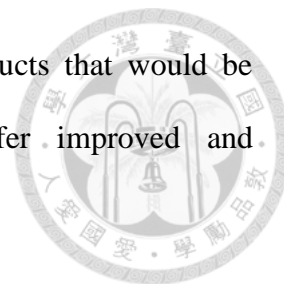


previous efforts demonstrated that this flavoenzyme performs a four-electron oxidation to convert the glucosaminyl C6 alcohol found on the ‘top’ of the peptide scaffold to a carboxylic acid. Dbv29 is unusual among enzymes known to perform oxidations at the C6 site, both in the coenzyme used (for example, NAD is used in the UDP-glucose dehydrogenase family) and in lacking a typical cysteine residue in the active site<sup>11-13</sup>.

Our previous studies showed that Dbv29 activity is dependent on the presence of an acyl group on the C2 amine of the same carbohydrate residue<sup>9, 10, 14, 15</sup>, though the exact nature of this acyl group proved somewhat flexible, as Dbv29 could oxidize both teicoplanin—containing an extended acyl chain—and analogs of A40926 with an acetyl or acyl group attached (and lacking a distant N-acetyl glucosamine substituent). Bioinformatics analysis pointed to two tyrosine residues, Tyr165 and Tyr473, as potentially important for the reaction mechanism<sup>16</sup>, but single phenylalanine mutants did not significantly affect activity in an overnight incubation experiment. Mutagenesis of His91 and Cys151 did provide strong support that the flavin cofactor, characterized at that time as flavin mononucleotide (FMN), was covalently bound to the enzyme through these residues; unexpectedly, the double mutant of these two enzymes retained some level of activity, unlike other flavoenzymes that covalently link to their cofactor<sup>9, 16-19</sup>. Thus, the function of this enzyme remained enigmatic.

In this study, we sought to both understand the mechanism of this unusual enzyme and—given our prior demonstration of acyl chain promiscuity—explore whether it could be employed to generate non-natural glycopeptide analogs. We report crystallographic, biophysical and biochemical characterization of Dbv29 that demonstrates the critical role of Tyr165 and Tyr473 in the enzyme reaction. Additionally, through the serendipitous discovery of a reaction intermediate in the crystal structure, we design a new synthetic strategy to rationally intercept the reaction

mechanism. This approach allows access to new classes of products that would be extremely difficult to obtain by other means and that offer improved and complementary profiles for antimicrobial drug discovery efforts.





## 2. Materials and Methods

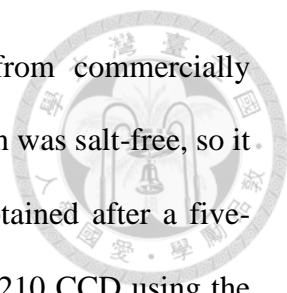
### 2.1 Gene cloning and protein purification

The dbv29 gene was amplified from genomic DNA by PCR. The products were each ligated according to the manufacturers' instruction (Novagene) into the pET-28a(+) expression vector which provides an N-terminal His<sub>6</sub>-tagged protein. *E. coli* cells were grown at 37°C in 1 L of LB medium until an A<sub>600</sub> of about 0.7 was reached. Protein expression was induced with 0.2 mM IPTG at 16°C, and, after 12 h, the cells were harvested by centrifugation, resuspended in 10 mM imidazole-HCl buffer, pH 7.8 (binding buffer), and ruptured using a microfluidizer. The cell-free extract prepared by centrifugation was applied to an Ni<sup>2+</sup>-NTA resin column, which was then washed successively with binding buffer and wash buffer before eluting the target protein with 200 mM imidazole-HCl. Gel filtration was performed using an Äkta FPLC system equipped with an S-200 Superdex column (Amersham Bioscience) under isocratic conditions (20 mM Tris, pH 7.6, 100 mM KCl). The buffer was exchanged for 50 mM HEPES buffer, pH 7.2 using Millipore centrifugal filters. Protein concentrations were estimated using the Bradford assay. The purified protein was confirmed by SDS-PAGE, Western blotting, and electrospray mass spectrometry (ESI-MS).

### 2.2 Crystallization and data collection

Dbv29 was crystallized by using hanging drop vapor diffusion method at 20 °C. 10 mg ml<sup>-1</sup> Dbv29 in 20 mM TRIS (pH 8.0, 100 mM NaCl) was mixed with the same volume of reservoir solution. Hexagonal crystals can be obtained in two crystallization conditions. One contains 50 mM calcium chloride, 100 mM BIS-TRIS (pH 6.5), 30% (v/v) PEGMME 550; the other contains 200 mM imidazole malate (pH 5.5), 33% (v/v) PEG 600. Crystals were obtained after one-week incubation. A different crystal shape was achieved in 200 mM di-ammonium hydrogen citrate (pH 5.0), 17% (w/v) PEG



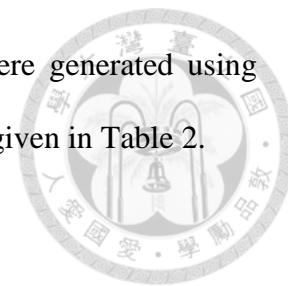


3350, C10 teicoplanin (5 mM; C10 teicoplanin was purified from commercially available teicoplanin mixtures by HPLC; the purified C10 teicoplanin was salt-free, so it required 10% (v/v) DMSO to dissolve), whereby a crystal was obtained after a five-month incubation. Data sets were collected on an ADSC Quantum-210 CCD using the synchrotron radiation X-ray source tuned at a wavelength of 0.97315 Å and an operating temperature of 100 K at beamline 13C1 of the National Synchrotron Radiation Research Center in Taiwan. Data were indexed and scaled with the HKL2000 package<sup>20</sup>. The hexagonal crystal belongs to P<sub>6</sub><sub>1</sub>22 space group with unit cell dimensions  $a = b = 66.1 \text{ \AA}$  and  $c = 790.5 \text{ \AA}$ . Because of the large cell dimensions, the maximum resolution that can be reached is 3.2 Å. The crystal obtained from the second condition belongs to P2<sub>1</sub> space group with unit cell dimensions  $a = 61.8 \text{ \AA}$ ,  $b = 150.8 \text{ \AA}$ ,  $c = 124.9 \text{ \AA}$  and  $\beta = 98.4^\circ$  and was diffracted to 1.93 Å resolution. The contents of both asymmetric units were estimated from the Matthews coefficient<sup>21</sup>. The data suggest that a value of  $2.08 \text{ \AA}^3 \text{ Da}^{-1}$  with 40.8% solvent corresponds to two molecules per asymmetric unit in the P<sub>6</sub><sub>1</sub>22 crystal, and a value of  $2.44 \text{ \AA}^3 \text{ Da}^{-1}$  with 49.7% solvent content indicates four molecules per asymmetric unit in the P2<sub>1</sub> crystal.

### 2.3 Structure determination and refinement

The molecular replacement method was used to obtain phase information, and MolRep<sup>22</sup> was used to find the phase solution. Initial phases were obtained at 4 Å resolution using the coordinates of AknOx (PDB ID: 2IPI) as the search model. Phase extension yielded electron density maps into which a polypeptide model was built with the program XtalView<sup>23</sup>. The model was further refined with CNS and REFMAC<sup>24, 25</sup>. Water molecules were added with a water-pick routine in the CNS program. The final model has an R factor of 16.0% for all reflections between 30 Å and 1.93 Å resolution and an  $R_{\text{free}}$  of 20.5% under 5% randomly distributed reflections. In the Ramachandran

plot, 99.5% of all residues are in the allowed region. Figures were generated using PyMOL (<http://www.pymol.org>). Detailed refinement statistics are given in Table 2.



#### 2.4 Enzymatic activity assay

Dbv29 activity was determined by LC-MS. The assay mix containing enzyme (10  $\mu\text{g}$ ) and the corresponding substrate (1 mM) in buffer (50 mM HEPES pH 7.2, 100 mM NaCl, 1 mM DTT) (total volume 150  $\mu\text{l}$ ) was incubated for 2 h at 37  $^{\circ}\text{C}$ . Each reaction mixture was then centrifuged at 16,000 g for 5 min (Heraeus Biofuge Pico) and filtered on an ultracentrifugal filter unit (5 kDa cut-off membrane, Millipore). The filtrate was directly subjected to HPLC-ESI/Q-ToF (Waters HPLC 2695 interfaced with an ESI source coupled to a Micromass micro Q-ToF mass spectrometer) or HPLC-ESI-LTQ (Agilent 1200 Series interfaced with an ESI source coupled to a Thermo-Finnigan LTQ XL ion trap spectrometer), using a gradient of 0–60% acetonitrile in 0.1% TFA in water over 30 min. Online LC-MS spectra were recorded by MassLynx (Waters) or Xcalibur (Thermo Fisher Scientific, Inc.).

#### 2.5 Mutagenesis

Site-directed mutagenesis was carried out by using QuickChange (Stratagene), and the wild-type Dbv29 was used as the template for single mutation. For double mutation, the single or double mutants, respectively, served as templates. All mutations were confirmed by DNA sequencing. The pET/His plasmid was used for protein expression. Mutant enzymes were purified with the same protocol as recombinant wild-type Dbv29.

#### 2.6 Circular dichroism spectroscopy

For structural analysis, far UV circular dichroism spectra of wild-type Dbv29 and mutants were recorded between 240 nm and 190 nm on a Jasco circular dichroism spectropolarimeter at 25  $^{\circ}\text{C}$ . Proteins were present at 1 mg ml $^{-1}$  in a standard buffer

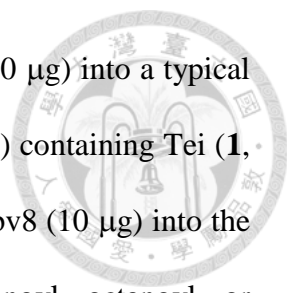
solution. Each spectrum represents an average of three scans, wherein the buffer background is subtracted. For stability analysis, proteins of wild-type Dbv29 and Y165F/Y473F mutant were dissolved in a standard buffer solution at  $1 \text{ mg ml}^{-1}$  for thermal melting experiments. Circular dichroism measurement was carried out with a Jasco J-815 spectropolarimeter in a temperature range of 20–100 °C at 222 nm wavelength. Melting temperatures,  $T_m$ , were determined by using two-state approximation.

## 2.7 Analytical ultracentrifuge analysis

The sedimentation velocity experiments were performed with a Beckman-Coulter XL-I analytical ultracentrifuge. Samples and buffers were loaded into 12-mm standard double-sector Epon charcoal-filled centerpieces and mounted in an An-60 Ti rotor. We introduced 400  $\mu\text{l}$  of a  $1 \text{ mg ml}^{-1}$  sample into the cell. Sedimentation velocity experiments were performed at rotor speed of 40,000 r.p.m. at 20 °C. The signals of samples were monitored at 280 nm and collected every 3 min for 6 h. The raw data of experiments were calculated using SedFit software<sup>26</sup>. The density and viscosity of buffer were calculated using Sednterp software (<http://www.jphilo.mailway.com/default.htm>). All samples were visually checked for clarity after ultracentrifugation, and no indication of precipitation was observed.

## 2.8 Synthetic conditions for new analogs

Teicoplanin (a mixture of five analogs) was purchased from AAPIN Chemicals Ltd (UK). The mixture was subjected to HPLC purification to obtain C10-Teicoplanin with purity >95%. 5-Azide pentylamine (99%) was obtained from Dr. Chung-Yi Wu (Academia Sinica) as a gift. All other chemicals were obtained from Sigma/Aldrich Chemical Co. (St. Louis, MO, USA) without further purification unless otherwise stated. Teicoplanin analogs (**2-6**) were enzymatically synthesized as described previously<sup>9, 10, 14</sup>.



<sup>15</sup>. In brief, compound **2** was prepared by adding Dbv21 or Orf2\* (10  $\mu$ g) into a typical buffer solution (50 mM HEPES, pH 7.2, 100 mM NaCl, 1 mM DTT) containing Tei (**1**, 1 mM) for 5 h at 37°C; compounds **3-6** were prepared by adding Dbv8 (10  $\mu$ g) into the same buffer solution but containing compound **2** and butanoyl-, hexanoyl-, octanoyl-, or decanoyl-CoA for compounds **3-6**, respectively, at 25°C overnight. For the final reaction, Teicoplanin analogs (**9**, **10**, **11**, **25** and **32**, which showed relatively higher yields from initial tests) were enzymatically synthesized using the optimized conditions described below. Enzyme stability (and, indirectly, the extent of exposure of the aldehyde for functionalization) was tested in an array of organic solvents (MeCN, MeOH, EtOH, DHF, DMF, DCM, DMSO etc.); DMSO turned out to be the most appropriate solvent as Dbv29 is highly stable in up to 90% DMSO. 50% DMSO, however, is included in the current protocol as it gives relatively higher yields. Other conditions, such as the amounts of alkylamine and reductant<sup>27</sup>, were also optimized one at a time against the fixed concentration of Tei (0.5 mM) for a better yield (below). In general, 10-fold of alkylamine and the reducing agent versus Tei was determined to be most appropriate for the current protocol, which is summarized as follows: Tei (0.5 mM), alkylamine (5 mM; carbon number >6), and Na(CN)BH<sub>3</sub> (5 mM) in 50% DMSO buffer solution at 37°C overnight incubation.

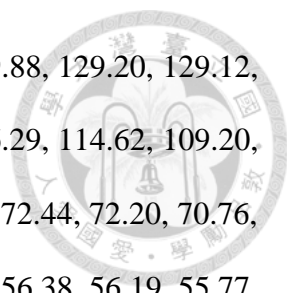
## 2.9 Compound characterization

All compounds were characterized by MS and HPLC. For select compounds, NMR analyses were performed on a Bruker Avance 600 spectrometer equipped with CryoProbe<sup>TM</sup>, with tetramethylsilane (TMS) as an internal standard. Compounds were dissolved in deuterated dimethyl sulfoxide (DMSO-d<sub>6</sub>) if not otherwise stated, and spectra were recorded at room temperature. NMR peaks of reference compound (**de-r4-**

**Tei**) and compound **25** are listed below (also see Table 5); original spectra can be found in Figures 11 and 12.

Reference compound (**de-r4-Tei**). HRMS ES(-): 780.1677 [M-2H]<sup>-2</sup>, calc. for C<sub>72</sub>H<sub>68</sub>O<sub>28</sub>N<sub>8</sub>Cl<sub>2</sub> 780.1760 [M-2H]<sup>-2</sup> (Figure 13). <sup>1</sup>H-NMR (600 MHz, DMSO-d<sub>6</sub>) (Figure 11): 9.68 (m, 2H), 9.27 (m, 1H), 8.65 (m, 1H), 8.53 (s, 1H), 8.42 (s, 1H), 7.92 (m, 2H), 7.80 (m, 1H), 7.26 (m, 3H), 7.18 (m, 2H), 7.10 (m, 4H), 6.97 (m, 1H), 6.75 (m, 3H), 6.46 (m, 2H), 6.23 (m, 1H), 5.27 (m, 2H), 5.13 (m, 2H), 4.42 (m, 2H), 4.34 (m, 1H), 4.28 (d, J 11.5, 1H), 4.12 (m, 1H), 3.06 (m, 2H), 2.95 (m, 1H), 1.88 (m, 2H), 1.26 (m, 1H). <sup>13</sup>C NMR (600 MHz, DMSO-d<sub>6</sub>): 172.79, 172.72, 169.87, 169.63, 169.32, 168.59, 168.44, 167.36, 167.27, 167.19, 158.76, 158.67, 158.56, 158.49, 158.29, 158.08, 157.87, 157.46, 157.38, 155.52, 155.09, 150.27, 150.23, 149.06, 148.56, 148.49, 147.63, 143.51, 135.58, 134.67, 134.41, 128.26, 128.19, 127.13, 126.20, 125.66, 125.25, 123.43, 120.71, 120.19, 120.08, 118.29, 116.30, 105.66, 99.26, 99.13, 97.52, 96.65, 76.91, 76.81, 73.80, 73.66, 73.00, 72.55, 70.59, 70.00, 69.92, 69.81, 65.99, 65.92, 63.11, 61.13, 60.75, 59.25, 56.12, 56.07, 55.86, 55.79, 55.17, 54.59, 23.21, 23.12. An expanded version of HSQC, HMBC and <sup>1</sup>H COSY spectra of the reference compound is shown in Figure 11.

Compound **25**. HRMS ES(+): 984.3033 [M+2H]<sup>2+</sup>, calc. for C<sub>95</sub>H<sub>104</sub>O<sub>32</sub>N<sub>10</sub>Cl<sub>2</sub> 984.3098 [M+2H]<sup>2+</sup> (Figure 13). <sup>1</sup>H-NMR (600 MHz, DMSO-d<sub>6</sub>) (Figure 12): 9.67 (m, 2H), 8.61 (m, 1H), 8.40 (m, 1H), 7.93 (m, 2H), 7.77 (m, 2H), 7.18 (m, 14H), 6.71 (m, 3H), 6.48 (m, 2H), 6.23 (m, 1H), 5.24 (m, 7H), 4.96 (m, 1H), 4.62 (m, 1H), 4.39 (m, 5H), 4.08 (m, 1H), 3.70 (m, 2H), 3.00 (m, 4H), 2.14 (m, 1H), 1.99 (m, 1H), 1.86 (m, 3H), 1.48 (m, 2H), 1.37 (m, 1H), 1.15 (m, 12H), 0.84 (d, J 6.5, 6H). <sup>13</sup>C NMR (600 MHz, DMSO-d<sub>6</sub>): 172.87, 172.79, 172.41, 172.36, 169.94, 169.81, 169.39, 169.24, 168.88, 168.76, 168.35, 167.28, 158.54, 158.41, 158.21, 158.00, 157.80, 157.67, 157.48, 157.41, 155.59, 148.27,




135.73, 135.62, 134.68, 131.88, 131.82, 130.21, 130.09, 129.96, 129.88, 129.20, 129.12, 128.86, 128.80, 128.38, 125.02, 120.79, 120.74, 120.20, 118.28, 116.29, 114.62, 109.20, 98.91, 96.62, 76.95, 76.83, 73.80, 73.73, 72.85, 72.80, 72.71, 72.56, 72.44, 72.20, 70.76, 70.68, 70.09, 69.81, 66.02, 63.13, 61.30, 61.20, 61.08, 60.76, 56.58, 56.38, 56.19, 55.77, 54.57, 51.74, 51.00, 50.16, 35.98, 31.37, 29.23, 29.10, 29.02, 28.97, 28.92, 28.87, 28.80, 27.47, 27.40, 26.76, 26.71, 26.67, 25.11, 24.97, 23.25, 23.17, 22.65, 22.18, 14.09. An expanded version of HSQC, HMBC and <sup>1</sup>H COSY spectra of the compound **25** is shown in Figure 12.

#### 2.10 *In vivo* study

ICR female mice were purchased from the National Laboratory Animal Breeding and Research Center, Taipei, Taiwan. Mice with average body weight of 27 to 30 g were subjected to infection via intravenous (i.v.) injection with  $1.3 \times 10^5$  cfu/mouse of *E. faecalis* (ATCC 51559) at day 0. For treatment study, mice were randomized into four groups at the start of the experiment and administered 10 mg/kg of either vancomycin, teicoplanin, benzylamine-teicoplanin (**25**) or saline (control) by i.v. twice a day for three days (from day 1 to day 3, a total of 6 doses). Mice were subjected to anesthesia; whole blood was then sampled from orbital sinus on day 1, day 2, and day 3. The whole blood underwent serial dilutions with PBS, which were plated on Brain Heart infusion agar (BHI agar; Difco, Detroit, MI, USA) for enumeration of cfu (colony formation unit). The graphs and statistical analyses were performed using SigmaPlot® and SigmaStat® . Differences were considered significant if the P value was < 0.05.

### 3. Results

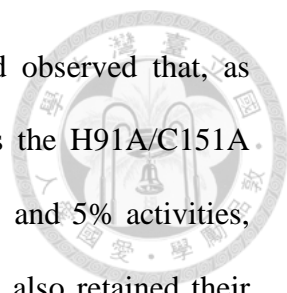
#### 3.1 Dbv29 is a FAD-containing dimer



To begin our investigations, we solved crystal structures of Dbv29 free and in complex with substrate. The free structure superimposes very well with the complexed form (r.m.s. deviation of 0.291 Å for 498 C $\alpha$ ), suggesting that binding of ligand does not induce substantial conformational transitions. As shown for the bound enzyme (Figure 2a), the structure consists of two domains: the F domain, which binds the flavin cofactor, and the S domain, which recognizes substrates, confirming that Dbv29 belongs to the flavin-dependent *p*-cresol methylhydroxylase (PCMH) superfamily<sup>19</sup>. The crystal structure also offers clear evidence that the flavin cofactor is flavin adenine dinucleotide (FAD) and not FMN. We suspect that the previous suggestion of an FMN-containing enzyme was likely because of hydrolysis of FAD in solution. The major difference between the two structures is that apo Dbv29 appears to form a dimer on the basis of a large subunit-subunit interface (1,346 Å<sup>2</sup> per monomer). However, in the complex, an asymmetric unit contains four molecules but does not demonstrate substantial protein-protein interaction, suggesting that the bound form exists as a monomer. This is in contrast to three other members of the PCMH family, glucooligosaccharide oxidase (GOOX)<sup>17</sup>, aclacinomycin oxidoreductase (AknOx)<sup>16</sup> and berberine bridge enzyme (BBE)<sup>28</sup>, which were solved recently in monomeric forms.

#### 3.2 Contributions to cofactor- and substrate-binding sites

As discussed above, our prior studies had highlighted unusual roles for four residues in controlling flavin binding and reactivity. The crystal structure now reveals a possible explanation for these functions (Figure 2). Though the isoalloxazine ring of flavin is covalently linked to the side chains of His91 and Cys151, it is closely connected to Tyr165 and Tyr473 via the substrate and a bridging water (Figure 2c). We revisited our



enzyme activity studies to obtain improved semi-kinetic data and observed that, as previously reported, single H91A and C151A mutants, as well as the H91A/C151A double mutant sequence, retained low levels of activity (11%, 23% and 5% activities, respectively, relative to wild type; Table 3). The mutated enzymes also retained their characteristic yellow color, in stark contrast to the corresponding mutants of GOOX, AknOx and BBE, which lost their color and activities. In contrast to our earlier work, however, the improved kinetic analysis demonstrated that Y165F and Y473F mutations did substantially affect enzyme activity, reducing function to 70% or 23% of the wild-type function, respectively; the Y165F/Y473F double mutant further led to loss of the yellow color and total loss of catalytic activity. Notably, the Y473E mutation similarly abolished activity; we suspect the lower pKa of this side chain or the alternate conformation adopted by the nonplanar ligand might explain this observation. We additionally examined the impact of mutations at the nearby positions Arg360, Ser364 and Thr366, as well as six other tyrosine residues to search for other contributing interactions. However, aside from the Y135F mutant protein, which could not be expressed, none of these mutations disrupted flavin binding (Table 3).

The substrate-binding site in the S domain is somewhat hidden by a long loop (residues 351–362) that crosses over the binding pocket. The electron density for the aglycone shows some discontinuity, whereas that for the N-acyl amino sugar moiety is apparent and clear. The sugar ring anchors against the *si* face of the isoalloxazine ring of FAD and exposes the C5 carbon close to the N5 nitrogen of FAD, while the C4 hydroxyl group forms a hydrogen bond with Asn427. The N-acyl group extends from the sugar ring into a hydrophobic lipid cavity (Figure 2b). We tested the importance of packing interactions in this lipid cavity by comparing the enzymatic turnover of two substrates with short (C4, **3**) and long (C10, **6**) alkane chains; the longer chain substrate



was preferred by approximately ten-fold, indicating contributions from but not a strict requirement for hydrophobic packing.

Beyond Tyr165 and Tyr473, which contact the C6 hydroxyl group, most residues that line the ligand-binding site are located in different loop regions. Mutational analysis of some of these residues showed that Ser364, Thr366, Tyr370 and Tyr470—though they did not impact flavin binding—did significantly alter enzyme function, with resultant activities of 4–21% of wild type. Additionally, Trp399 and Ile401 proved to be critical for function, as the mutated enzymes displayed either no activity or no more than 10% of wild-type function (Table 3).

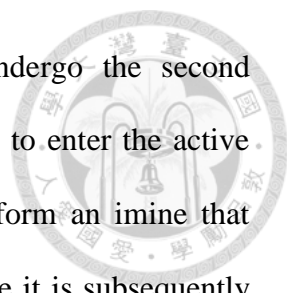
### 3.3 A diol intermediate leads to catalytic redirection

Although the structure of the complex was solved by cocrystallization with teicoplanin, the electron density is best fitted with neither teicoplanin (the substrate) nor oxo-teicoplanin (the product). Instead, the appearance of extra out-of-plane electron density on the C6 terminus of the amino sugar was best explained by the presence of two OH groups in  $sp^3$  configuration or a water-coordinated diol (Figure 2c). We had previously suggested a basic series of reactions needed to convert the C6 alcohol to a carboxylate in which a diol intermediate, or aldehyde equivalent, is used as the transformed substrate for the second half of the oxidation. The structural information obtained here now allows us to build on our prior analysis to propose a specific model for Dbv29 function (Figure 2d): in the first oxidation step, the C6-OH is deprotonated by the hydroxyl group of Tyr473, which in turn is activated by the hydroxyl group of Tyr165. These residues together drive the pro-R hydride transfer from the C6 carbon to the N5 of the isoalloxazine ring and yield the aldehyde intermediate corresponding to the diol found in the crystal structure. In the second oxidation, a second hydride transfer to FAD is likely facilitated by a proton relay network involving the tyrosine pair, the

substrate diol and the diol-coordinated water molecule. In each two-electron oxidation reaction, the reduction of FADH<sub>2</sub> back to FAD converts one molecule of molecular oxygen to hydrogen peroxide.

Beyond the mere presence of this unexpected chemical group on the protein substrate, we also discovered that this moiety appeared to be directly exposed to solvent. To test this hypothesis, we conducted two sets of reactions using isotopically labeled reactants. When the teicoplanin and Dbv29 were mixed in <sup>18</sup>O-labeled water, a dominant mass of M+4 instead of M+2 was found, indicating both oxygens of the diol intermediate were labeled with <sup>18</sup>O (Figure 3a). The simplest explanation for this result is that the aldehyde is accessible to water and that the aldehyde and diol are in equilibrium. In addition, when we reduced teicoplanin with NaBD<sub>4</sub> in the presence and absence of Dbv29, MS analyses of the product recovered from the reaction with protein showed a mass increased by 1 mass unit (Figure 3b). Again, the simplest explanation for this result is that the NaBD<sub>4</sub> was able to access the oxidized aldehyde in the reactive site and reduce it to a deuterated alcohol species.

The accessibility of the aldehyde further led us to explore the possibility of trapping and modifying this intermediate as a way to produce new antibiotic analogs. First, we tested the reductive amination of an aldehyde with an amine<sup>27</sup>. Enzymatic reactions were carried out in the presence of teicoplanin, Dbv29, sodium cyanoborohydride (Na(CN)BH<sub>3</sub>) and either n-decylamine or n-hexylamine. Remarkably, the LC-MS traces (Figure 3c) indicated that although reaction with n-decylamine resulted in the expected formation of n-decylaminated teicoplanin (**10**), the reaction with n-hexylamine led to n-hexylamidated teicoplanin (**11**), both of which were confirmed by MS (Figure 3d,e). These results can be explained by differential insertion of the amine into the active site. Thus, the hexylamine would react with the aldehyde intermediate, and the resulting

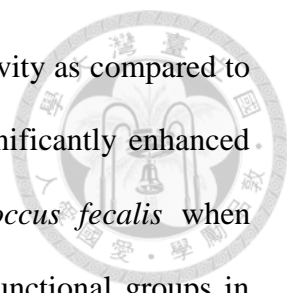


gem-hydroxyl-alkylamine could remain in the active site to undergo the second oxidation reaction. In contrast, decylamine appears to be too bulky to enter the active site and so instead reacts with the aldehyde functional group to form an imine that remains exposed to solution or is even released into solution, where it is subsequently reduced. Thus, the reactive intermediates can be trapped at varying stages during the catalytic course.

To take advantage of this reactivity, we constructed a chemoenzymatic strategy to create additional analogs with variations in both the length and the structure of the lipid chains at both C2- and C6-positions of teicoplanin and oxo-teicoplanin (Figure 4). Our protocol used Dbv21, an enzyme catalyzing deacetylation of the GlcNac moiety<sup>10, 14</sup>, and Dbv8<sup>15</sup>, an enzyme catalyzing the acylation of the glucosamine moiety, in combination with Dbv29 and various coenzyme A derivatives. When mixed in the presence of organic solvents (for example, 50% DMSO) and selective reducing agents (for example, 5 mM sodium cyanoborohydride), we were able to obtain a series of new analogs **12–31**. The enzymes were quite tolerant to new functional groups, as demonstrated by the inclusion of amantidine and related compounds as well as alkyne and azide handles that can be further manipulated through ‘click’ chemistry for applications such as identification of a second mode of action for broader spectrum analogs<sup>29, 30</sup>.

### 3.4 Analogs provide alternate antimicrobial protection

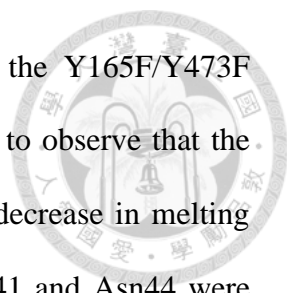
Having gained access to new chemical space with our modified teicoplanin analogs, we wanted to examine whether we had similarly reached new biological space; accordingly, we determined the minimum inhibition concentrations (MICs) of several of the compounds against collections of vancomycin-resistant *Enterococcus* (VRE) strains (Table 1). We observed that inhibition is proportional to the length of acyl side chains in



compounds **2–7**, with the carboxylate (**9**) showing no change in activity as compared to teicoplanin. The bi-lipid analogs **10** and **11**, however, showed significantly enhanced bactericidal activities against all five tested strains of *Enterococcus fecalis* when compared to vancomycin and teicoplanin. Notably, the different functional groups in compounds **10** and **11** yielded opposite outcomes for the different strains: the shorter-chain analog **11** is more effective against antibiotic-sensitive strains (for example, ATCC 33186), whereas the longer-chain analog **10** is more effective against drug-resistant strains (for example, ATCC 51559 and 700221). Strikingly, analog **25**, tailored with benzylamine (Figure 10), is equally effective against sensitive and resistant strains with dose potency one to two orders of magnitudes lower than vancomycin and teicoplanin and thus may complement compounds such as the A40926 derivative dalbavancin, which is active against *Staphylococci* (for example, MRSA) and *Streptococci* but weak against VRE<sup>31-34</sup>. As a first proof of principle for in vivo efficacy, we tested analog **25** against mice infected with VRE (ATCC 51559) via intravenous injection and found that this molecule was more effective than either vancomycin or teicoplanin at reducing blood bacterial counts (Figure 5).

### 3.5 Dbv29 has a dynamic quaternary structure

We remained intrigued by the different quaternary forms of the protein observed in the apo and bound forms (Figure 6). To investigate this matter further, we biophysically characterized several constructs, including the wild-type protein, the double mutant that lacks covalent linkages to flavin (H91A/C151A), the nonfunctional double mutant (Y165F/Y473F) and an additional construct with alanine mutations at each of two residues located at the dimeric interface (R41A/N44A). The latter two constructs did not form dimers as monitored by gel filtration analysis or analytical ultracentrifuge experiments and showed some minor deviations from wild type by circular dichroism

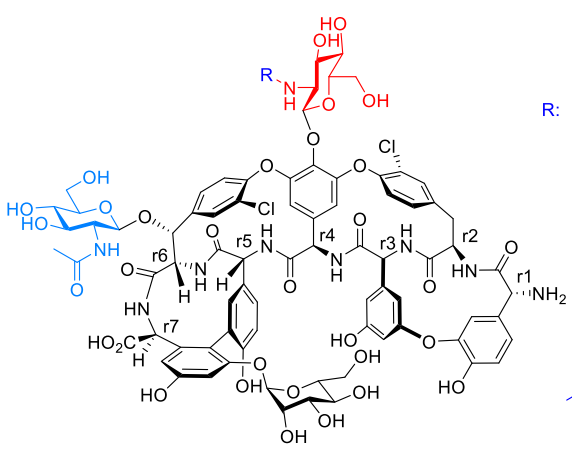
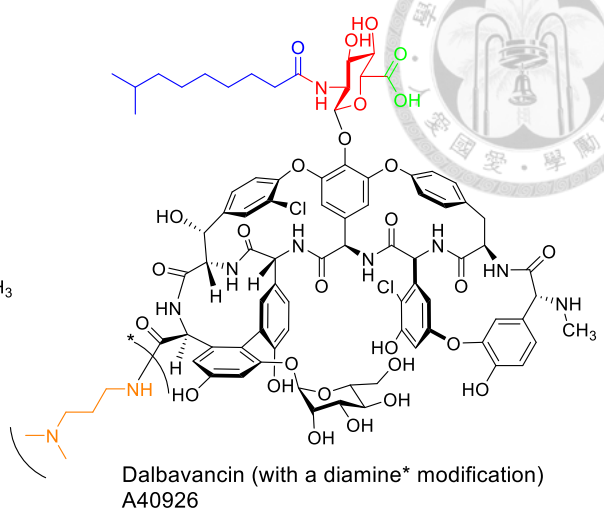
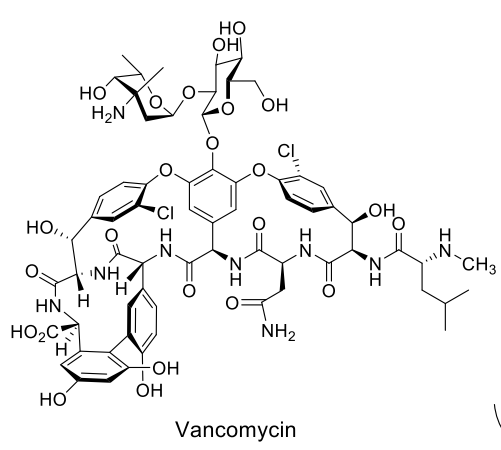
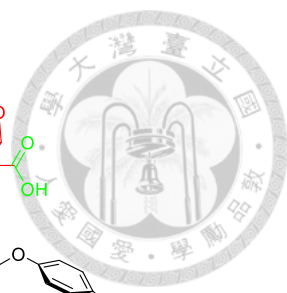


(Figures 7, 8). Our earlier functional test had demonstrated that the Y165F/Y473F double mutant lacked catalytic activity, but we were also surprised to observe that the enzyme structure was significantly destabilized, as indicated by a decrease in melting temperature of almost 35 °C (Figure 9). Additionally, though Arg41 and Asn44 were not anticipated to participate in the enzyme mechanism directly, the catalytic efficiency of the double mutant is substantially reduced (59% of wild-type activity; Table 3). Although the mechanistic importance of this is not fully understood, we speculate that the protein quaternary structure dynamics may be correlated to subtle structure changes when a ligand is loaded on or off the enzyme.

#### 4. Discussion

The development of new chemical entities for antibacterial applications remains a pressing and challenging goal. Our chemoenzymatic strategy toward this end was made possible by elucidating the reaction mechanism of Dbv29, coupled with the identification of several unusual features of this catalytic process. In particular, we took advantage of a promiscuous lipid cavity as well as a solvent-exposed reaction intermediate to create a small library of new compounds, several of which show promising results against a panel of bacterial strains. Because we focused our attention on an unusual site of modification—C6—we anticipate that the inevitable resistance profiles of these molecules should differ from other glycopeptides for which modifications occur at a dissimilar position, thus extending the effective lifetime of both classes of compounds and suggesting combination treatment as an attractive therapeutic strategy. Toward this end, our combined synthetic protocol provides a cost-effective and environment-friendly method that presages scalable production of these and other next generation drugs.





- R:
- Teicoplanin A2-2 (1)
  - Deacyl-teicoplanin (2)
  - C<sub>4</sub>-teicoplanin (3)
  - C<sub>6</sub>-teicoplanin (4)
  - C<sub>8</sub>-teicoplanin (5)
  - C<sub>10</sub>-teicoplanin (6)

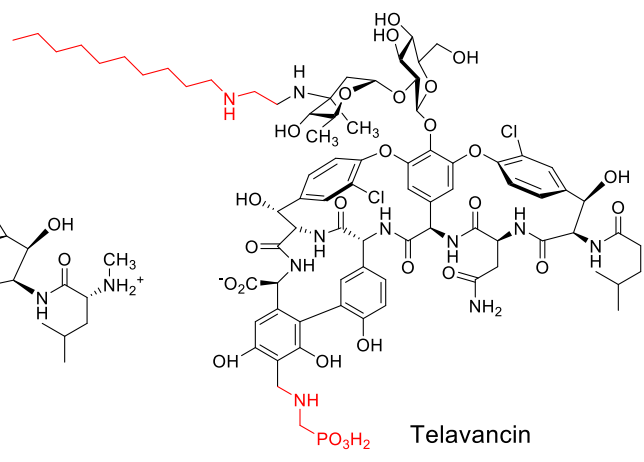
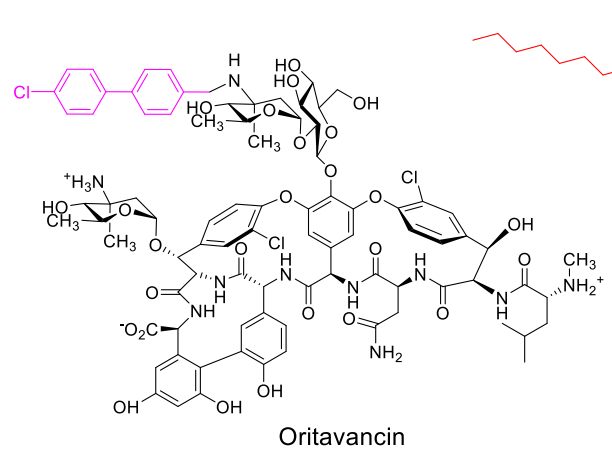
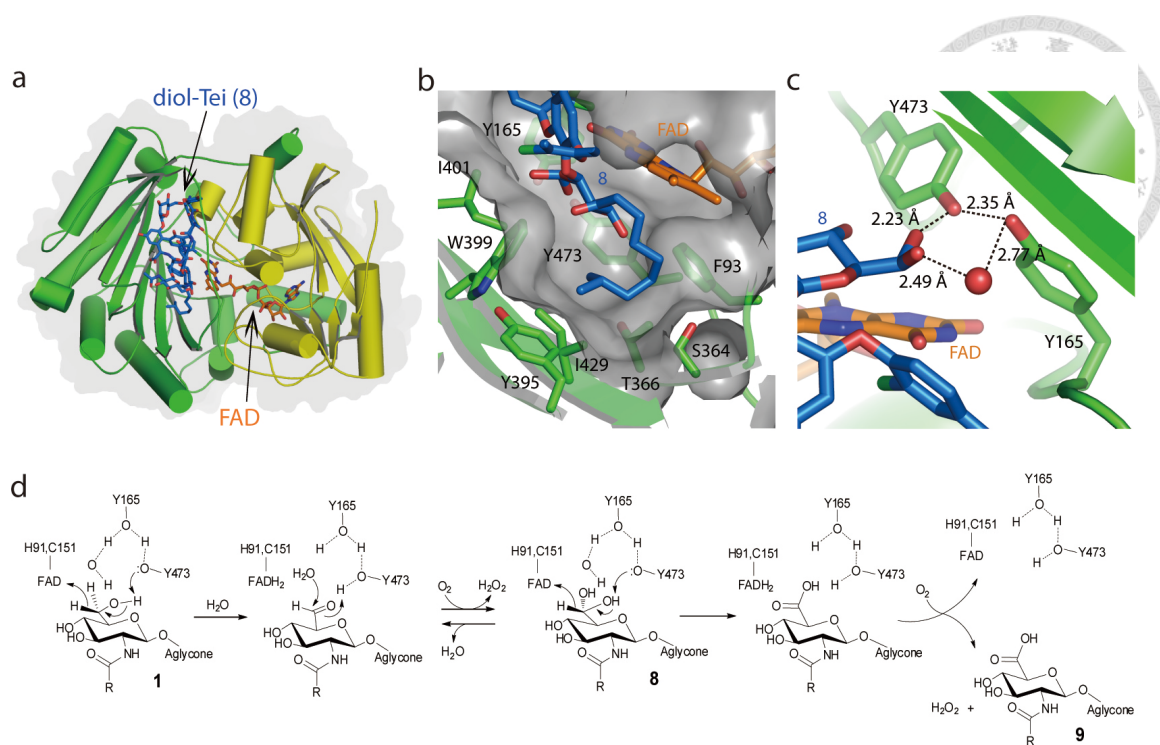


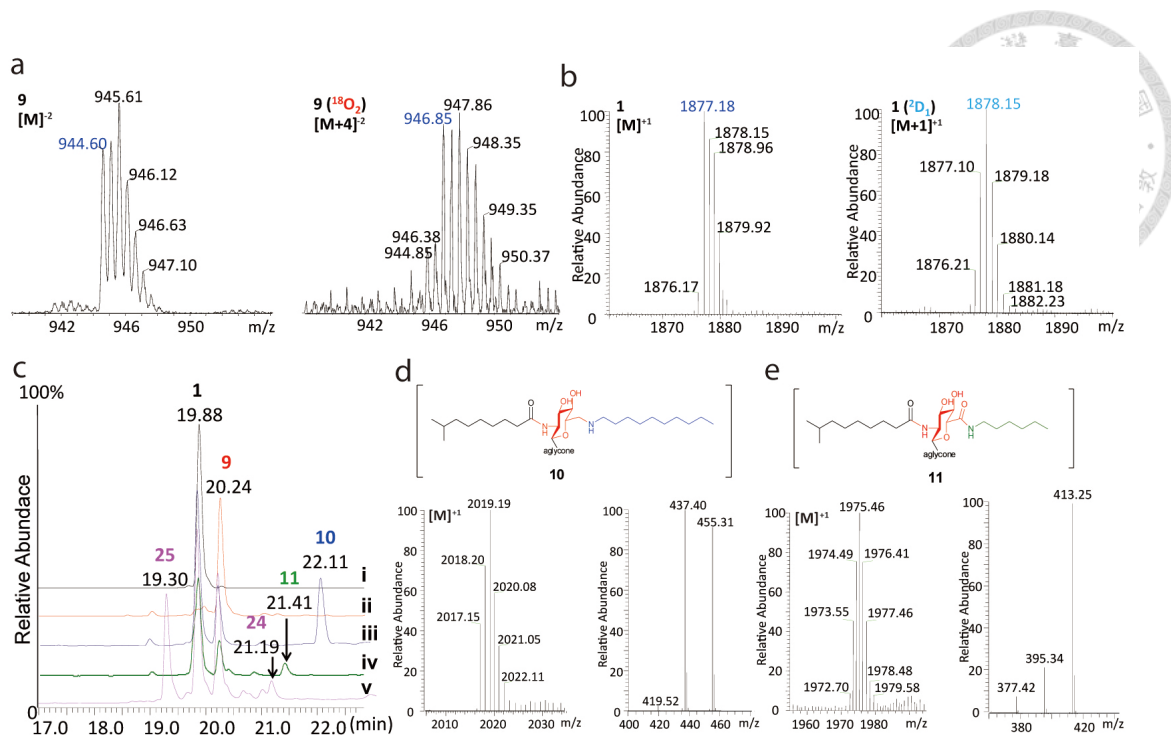
Figure 1. Structures of relevant glycopeptide antibiotics.



**Figure 2. The structure of Dbv29 and its enzymatic mechanism.**

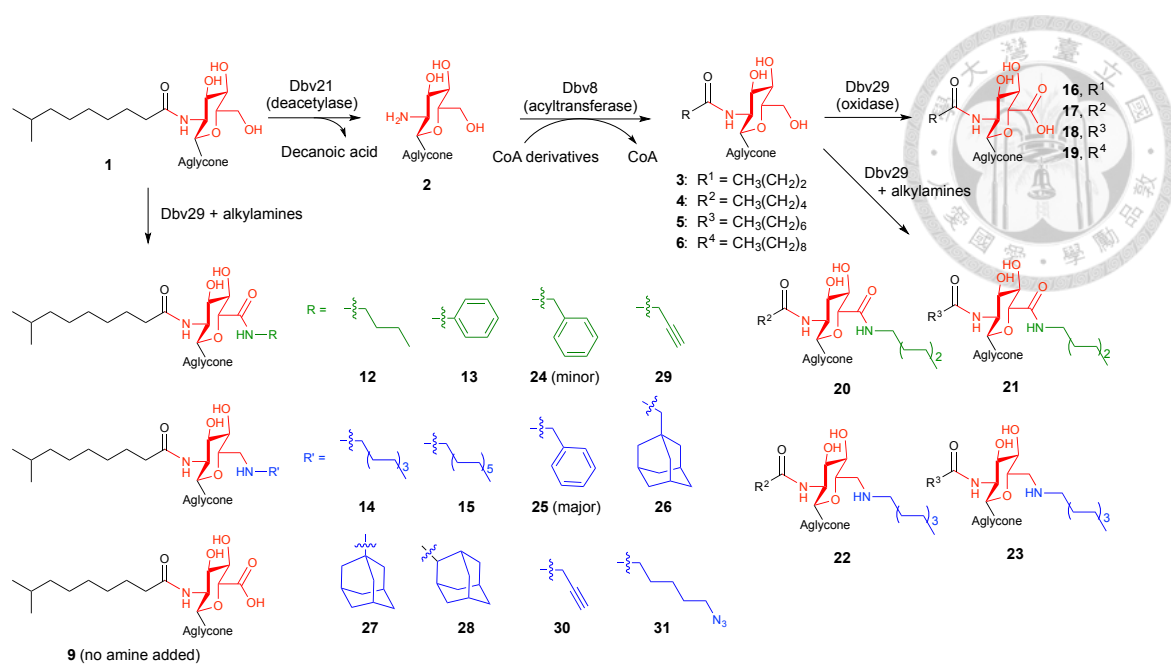
(a) The ligand-bound structure of Dbv29, in which the ‘F’ domain is colored in yellow and the ‘S’ domain is colored in green. (b) View of the lipid tunnel in Dbv29 and substrate binding residues. (c) View of the Tei-binding site, wherein the residues Tyr-473 and -165 interact with the water-coordinated diol intermediate. (d) The proposed catalytic course.





**Figure 3. Intermediate trapping strategies to probe the catalytic mechanism.**

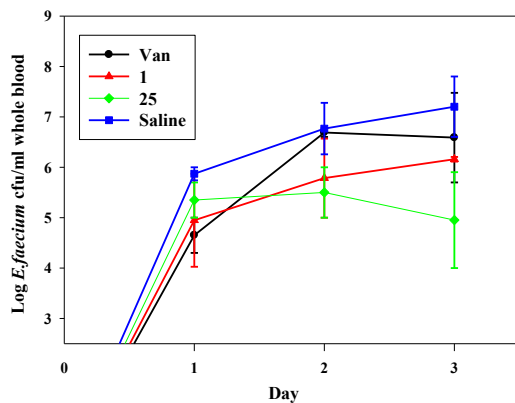
(a) Mass spectra of Dbv29 enzymatic assays carried out in a  $H_2^{16}O$  or  $H_2^{18}O$  buffer solution. (b) Mass spectra of the substrate with  $NaBD_4$  in the absence (left) and presence (right) of Dbv29. (c) HPLC traces demonstrating the formation of n-decylaminated Tei (**10**) and n-hexylamidated Tei (**11**); the reactions were conducted in the presence of  $Na(CN)BH_3$ , Dbv29, and decylamine (trace iii) or hexylamine (trace iv); controls contained either no Dbv29 (trace i) or no  $Na(CN)BH_3$  (trace ii). (d) MS and MS/MS spectra of n-decylaminated Tei (**10**). An  $m/z$  of 455 represents the mixed alkyl-acyl glucosaminyl moiety. (e) MS and MS/MS spectra of bis-acyl oxo-Tei (**11**). An  $m/z$  of 413 represents the bis-acyl aminoglucuronyl moiety.



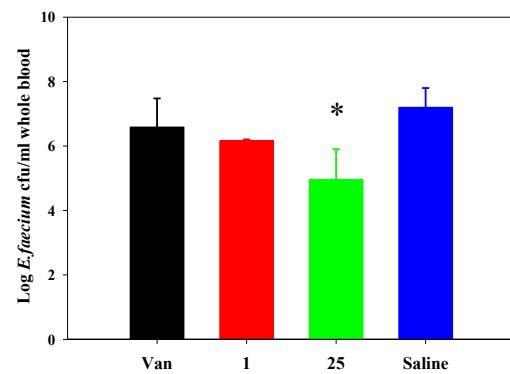
**Figure 4. Variations in acyl chain length regulate the resultant glycopeptide analogue.**

Reaction with large alkylamines yield amine products, smaller amines yield amide products, and the absence of an amine yields the carboxylic acid. Black, green and blue colors denote variations in carbon length at C2, amidations at C6 and aminations at C6, respectively. See Table 4 for full descriptions of acyl groups used and product yields.

a

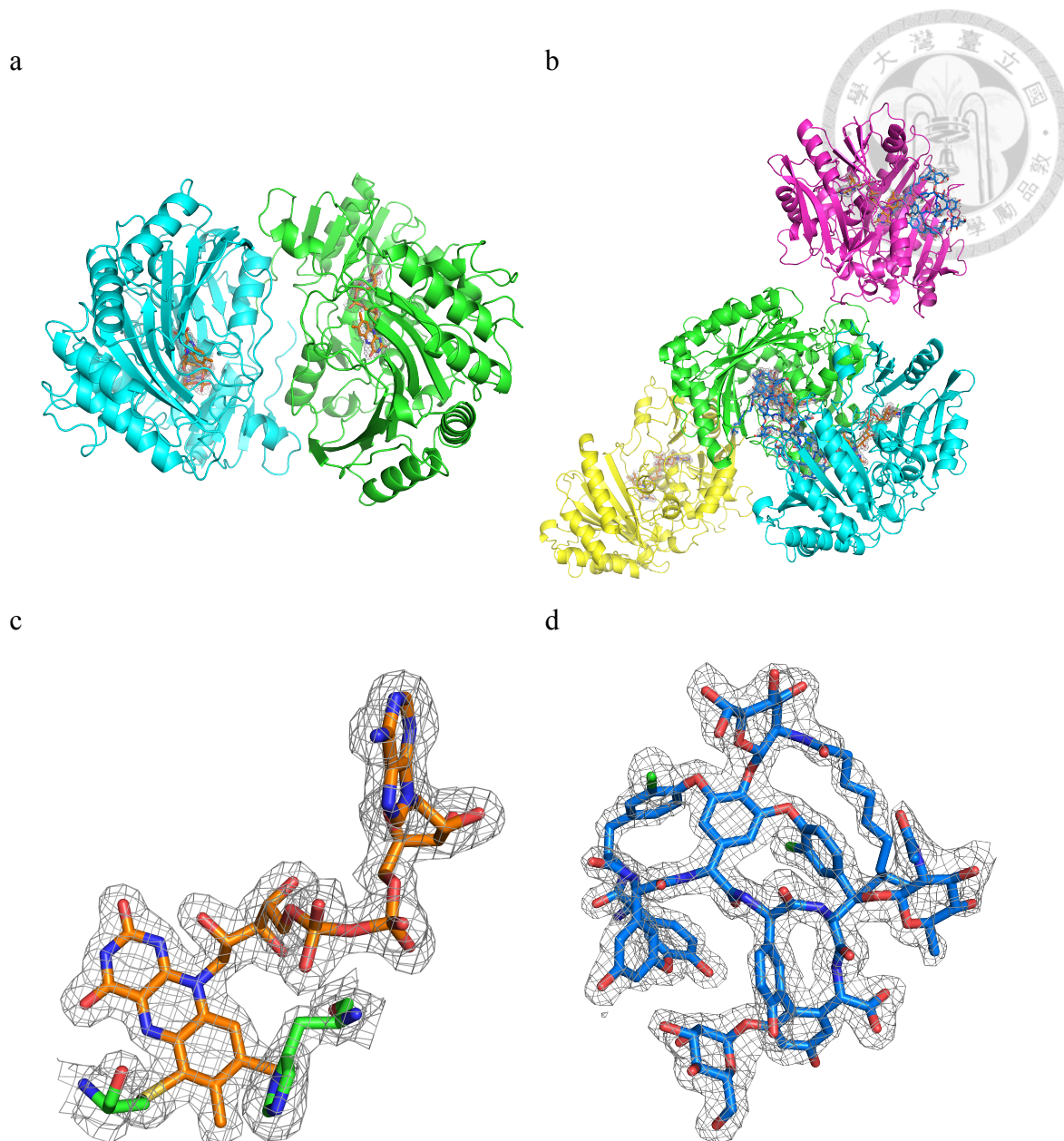


b



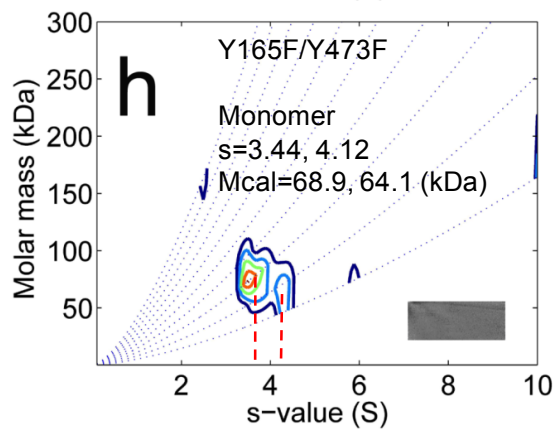
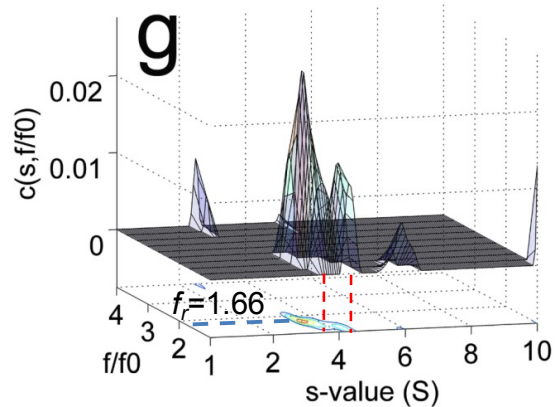
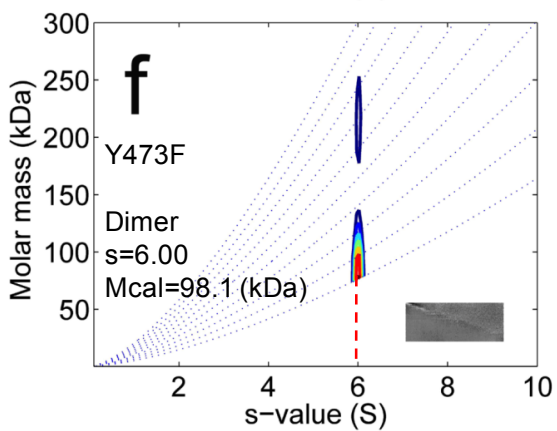
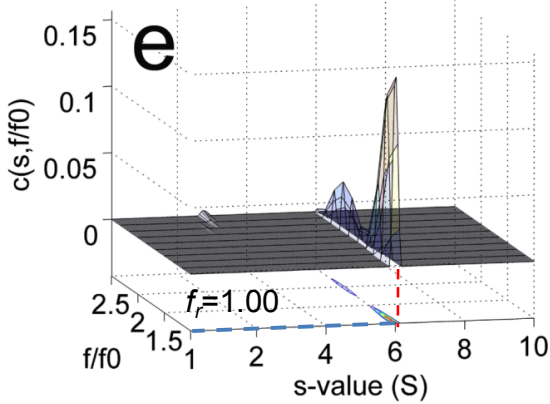
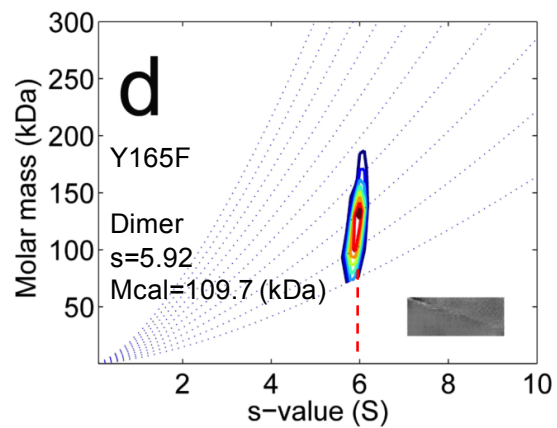
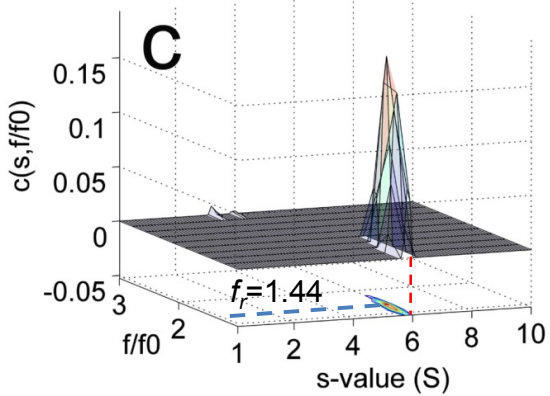
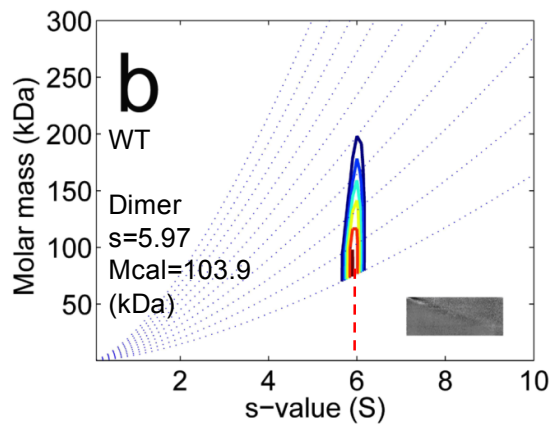
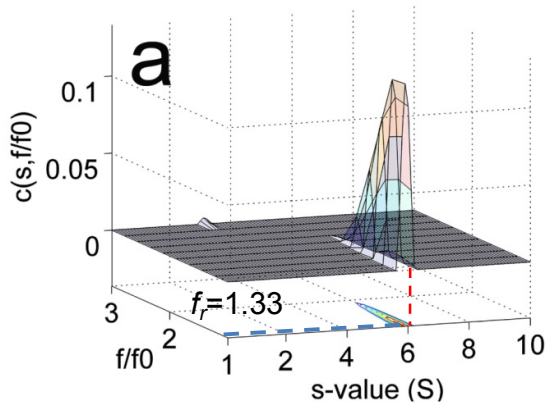
**Figure 5. Mice infection test for analog 25 in blood bacterial clearance.**

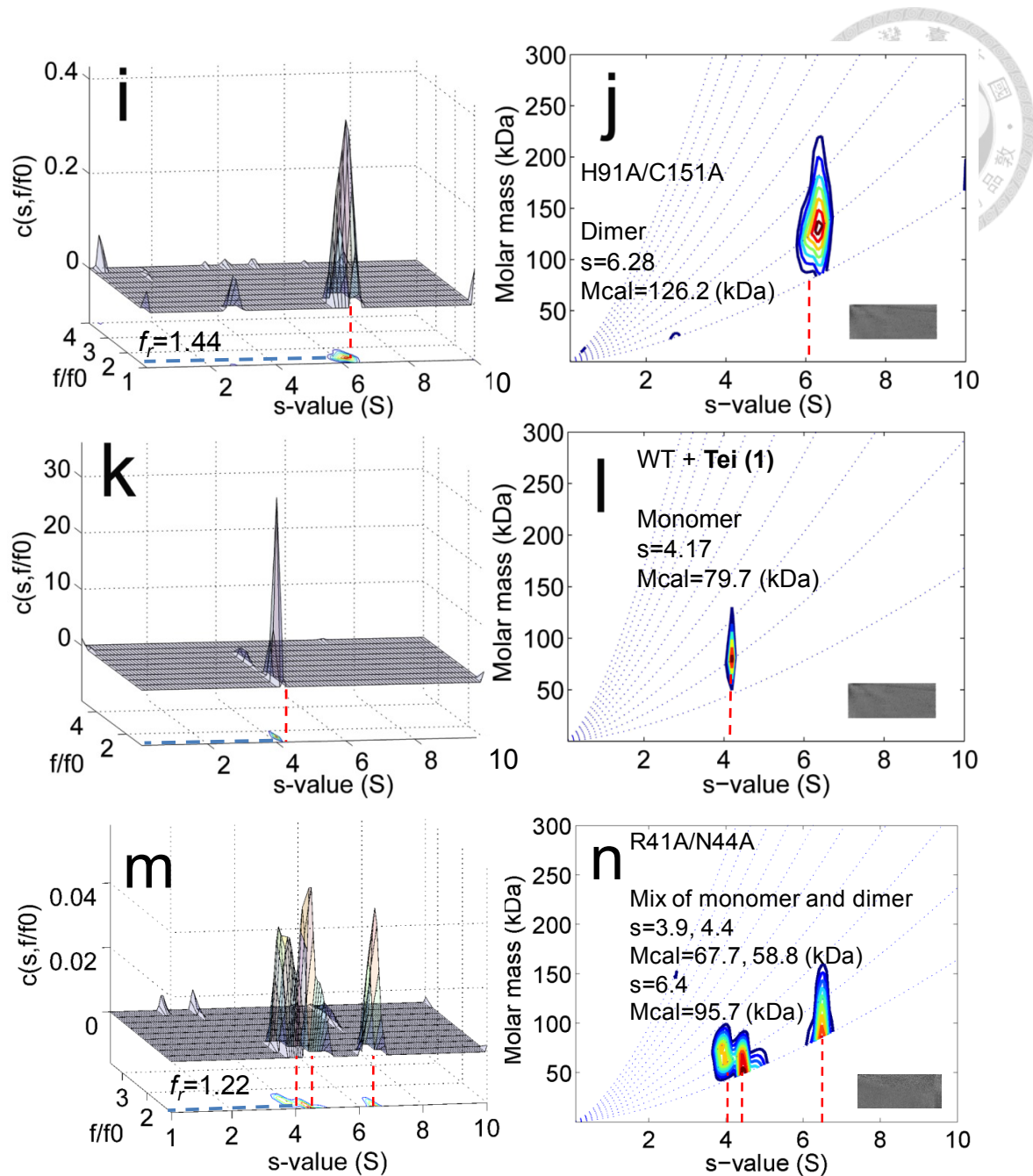
(a) The plot shows the whole blood bacterial counts of mice infected by *E. faecalis* (ATCC 51559) and treated with vancomycin (Van, in black) or teicoplanin (**1**, in red) or analog **25** (in green) or saline (control, in blue) for three days. The bacterial counts for the control are steadily increased over time, while the counts for groups treated with drugs are all suppressed but in various extents. Teicoplanin and analog **25** are more prominent than vancomycin in terms of overall drug efficacy. (b) The bar chart shows the bacterial counts at day 3 (end of treatment). In general, these three drugs all show positive effects against the infection, while only analog **25** displays significant. The asterisk (\*) indicates significant difference ( $P < 0.05$ ); data were expressed as mean  $\pm$  SD for a group of three mice)



**Figure 6. Structural analysis of dimeric and monomeric Dbv29.**

(a) The dimer of Dbv29 in a  $P6_122$  space group (two polypeptide chains in an asymmetric unit); FAD cofactors are represented as stick models. (b) Monomer of Dbv29 in complex with a water-coordinated-diol intermediate in a  $P2_1$  space group (an asymmetric unit consists of four polypeptide chains in which there are no typical protein-protein interactions with one another); the FAD cofactors and the Tei intermediate are represented as stick models in orange and blue, respectively. (c) The  $2F_o-F_c$  electron density map of FAD contoured at  $1.0 \sigma$ . (d) The  $2F_o-F_c$  electron density map of the diol intermediate contoured at  $1.0 \sigma$ .

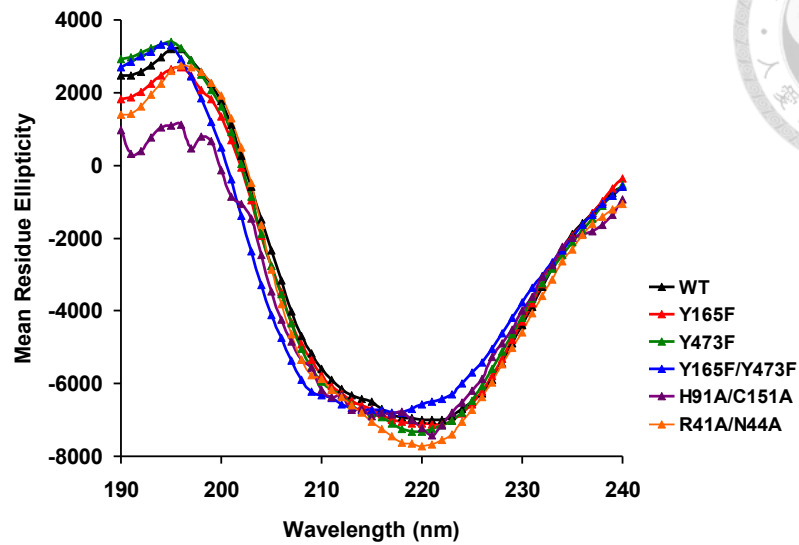




**Figure 7. Analytical ultracentrifugation (AUC) analyses for Dbv29 and mutants.**

(a-n) The raw experimental data were analyzed by Sedfit (<http://www.analyticalultracentrifugation.com/default.htm>) and the plots were generated by MATLAB (MathWork, Inc.). The spectra in left panels (a, c, e, g, i, k, m) are the calculated  $c(s, f_r)$  distributions that are shown as two-dimensional distribution with grid lines representing the  $s$  and  $f_r$  grid in the thermograph. Below this  $c(s, f_r)$  surface a contour plot of the distribution is projected into the  $s$ - $f_r$  plane, where the magnitude of  $c(s, f_r)$  is indicated by contour lines at constant  $c(s, f_r)$  in equidistant intervals of  $c$ . Contour plots in the right panels (b, d, f, h, j, l, n) are transformed from the calculated

$c(s, f_r)$  distribution and are shown as  $c(s, M)$  distribution. The dotted lines (colored in blue) indicate lines of  $f_r$  (frictional ratio). The signal of the  $c(s, M)$  distribution is indicated by the color temperature. The insert grayscale bars in the right panels indicate the residuals bitmap of each fit. **(a, b)** Dbv29 WT (dimer,  $s=5.97$  (red dashed line),  $f_r=1.33$  (blue dashed line)); **(c, d)** Dbv29 Y165F (dimer,  $s=5.92$ ,  $f_r=1.44$ ); **(e, f)** Dbv29 Y473F (dimer,  $s=6$ ,  $f_r=1.00$ ); **(g, h)** Dbv29 Y165F/Y473F (monomer,  $s=3.44$ ,  $f_r=1.66$ ); **(i, j)** Dbv29 H91A/C151A (dimer,  $s=6.28$ ,  $f_r=1.44$ ); **(k, l)** Dbv29 WT in the presence of 10 mM Tei (**1**) (monomer,  $s=4.17$ ,  $f_r=1.44$ ); **(m, n)** Dbv29 R41A/N44A (monomer,  $s=3.9$ ,  $f_r=1.22$ ). The wild-type enzyme and the two single mutants (Y165F and Y473F) **(a-f)** remain associated as dimers as opposed to the double mutant that dissociates into monomers **(g, h)**. The mutant also has a larger frictional ratio ( $f_r$ ), suggesting its molecular shape has been altered considerably.



	Helix (%)	Strand (%)	Turns (%)	Unordered (%)
WT	8	38	11	43
Y165F	9	36	11	43
Y473F	8	37	12	43
Y165F/Y473F	16	31	13	40
H91A/C151A	8	35	11	44
R41A/N44A	16	26	20	39

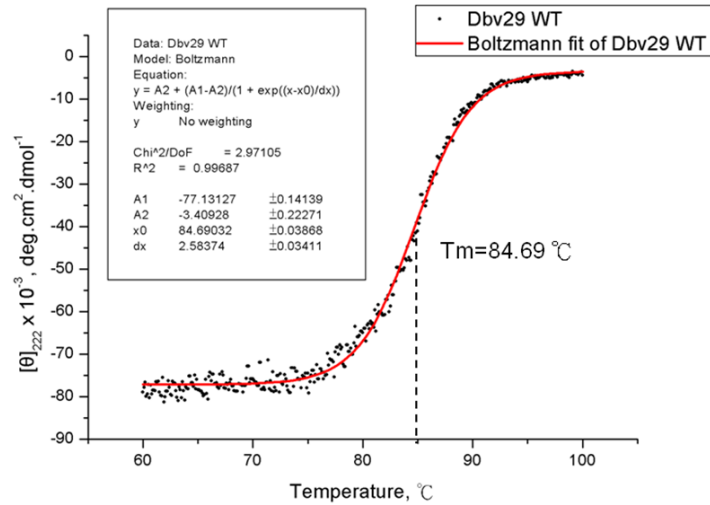
**Figure 8. Circular dichroism (CD) analyses for Dbv29 and mutants.**

CD spectra of WT and mutants (Y165F, Y473F, Y165F/Y473F, H91A/C151A and R41A/N44A), in which Y165F/Y473F and R41A/N44A double-mutations are somewhat deviated from others as shown in secondary structure distributions.

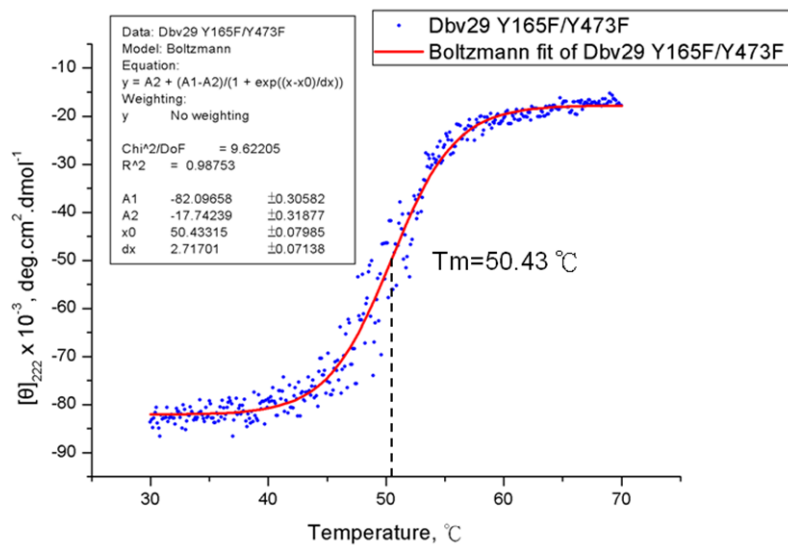




a



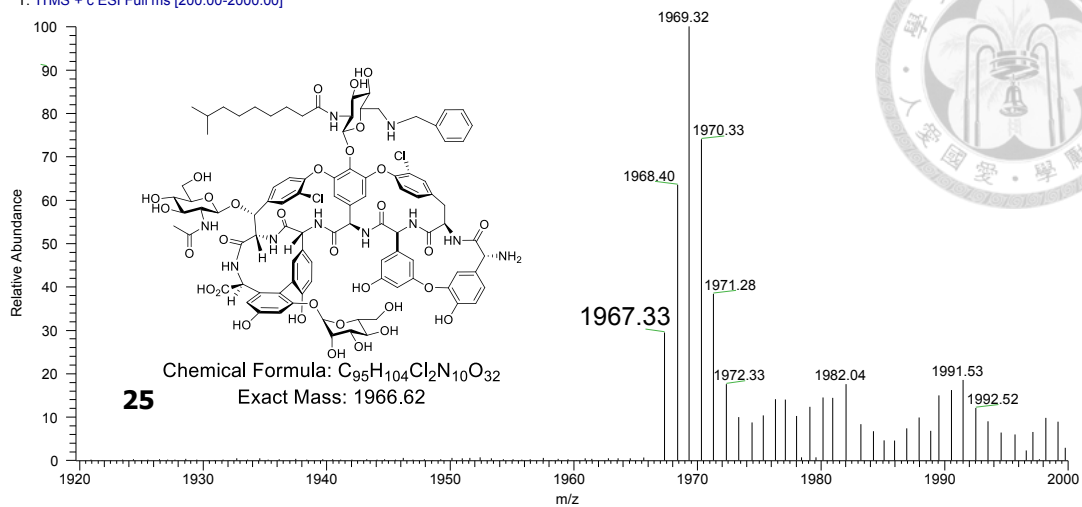
b



**Figure 9. CD analysis of protein stability.**

(a) The denaturation temperature ( $T_m$ ) for WT was determined by CD to be 84.7°C, meaning Dbv29 is unusually stable. (b) The denaturation temperature ( $T_m$ ) for the Y165F/Y473F mutant was determined to be 50.4°C; the protein was significantly destabilized by lacking the two oxygen atoms on the tyrosine pair. Therefore, residues Y165 and Y473 have considerable influences on protein stability. ( $\Delta T_m = 34.3^\circ\text{C}$ )

090723-tei+dbv29+dms0+nacnbh3+benzylamine #6163-6200 RT: 19.43-19.54 AV: 38 NL: 2.36E5  
T: ITMS + c ESI Full ms [200.00-2000.00]



090723-tei+dbv29+dms0+nacnbh3+benzylamine #6155-6237 RT: 19.40-19.65 AV: 83 NL: 5.28E4  
T: ITMS + c ESI Full ms [200.00-2000.00]

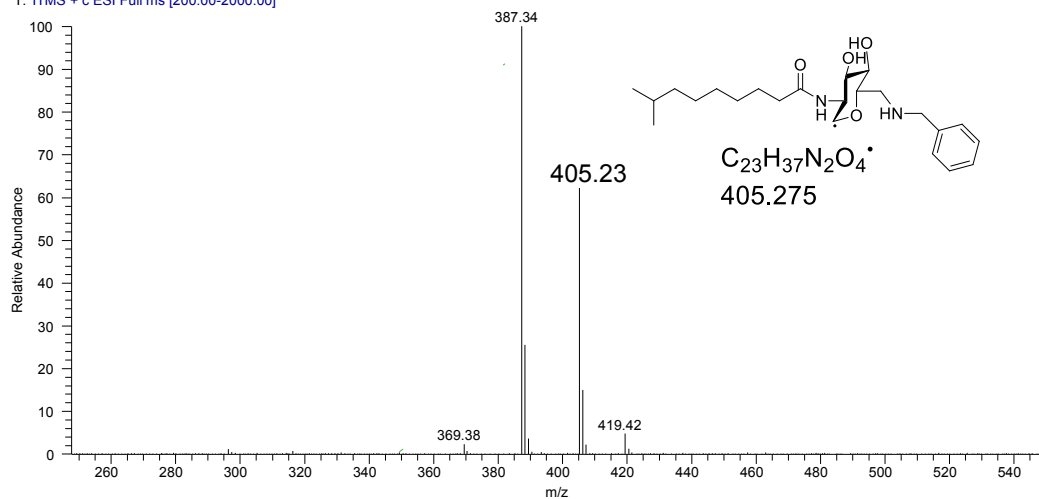
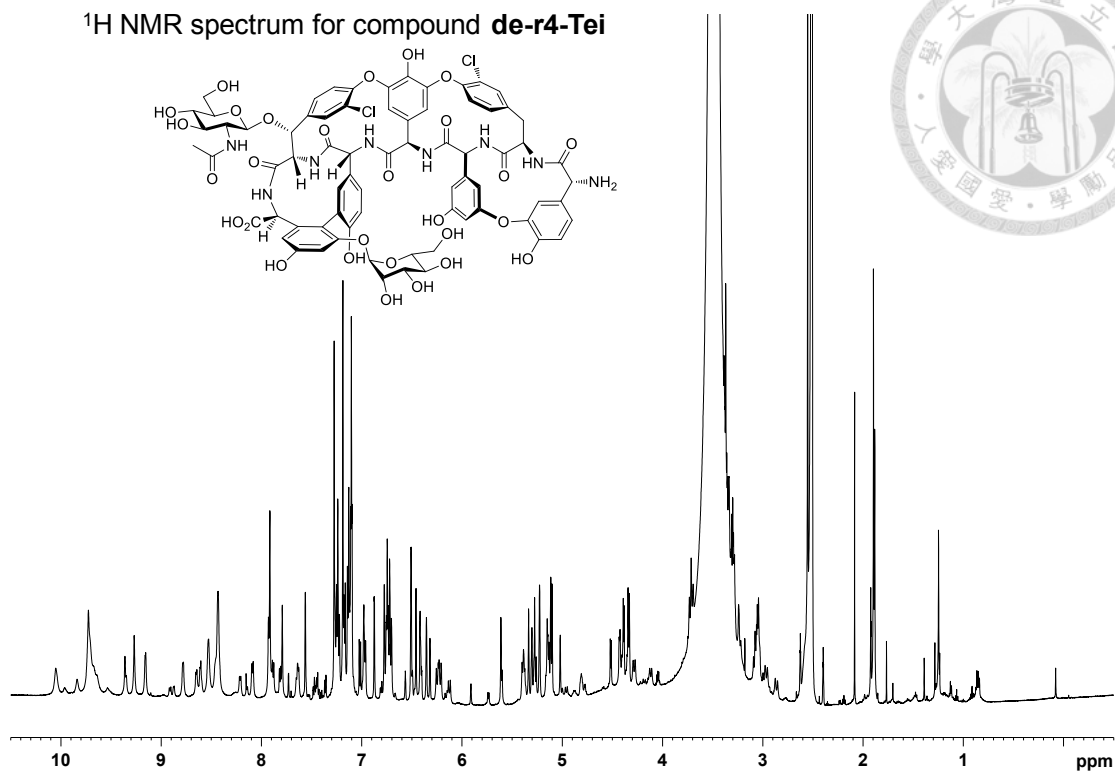


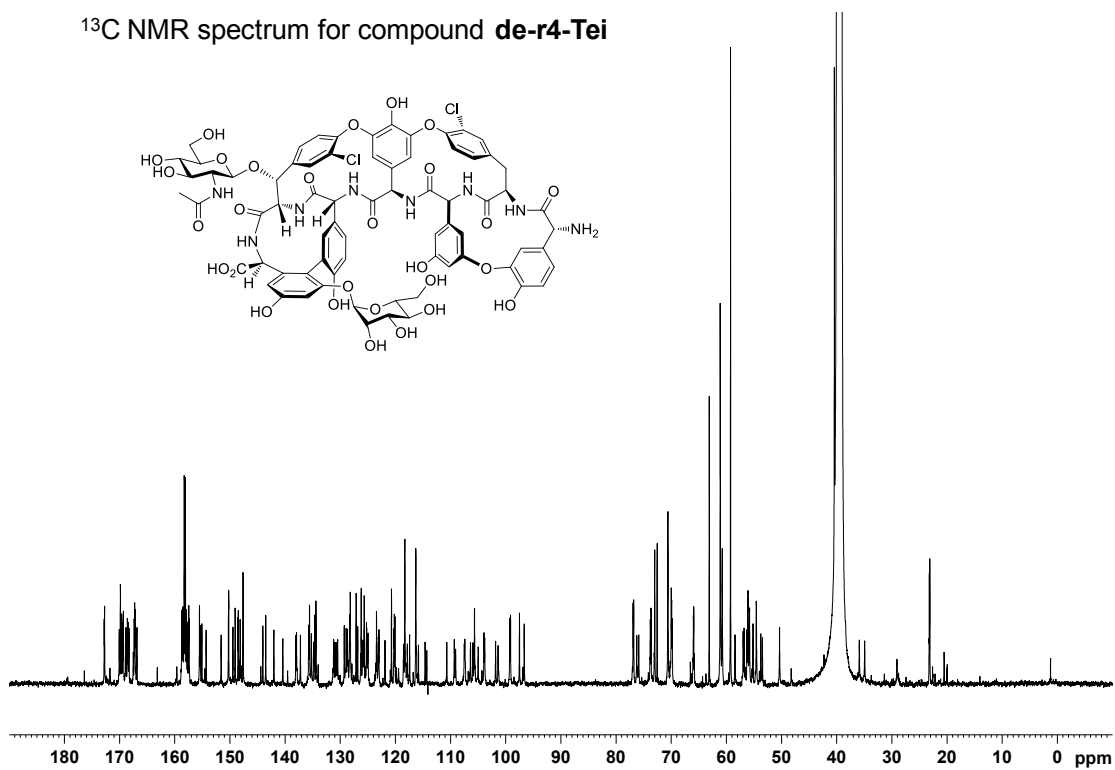
Figure 10. MS and MSMS spectra of the benzylamine-Tei aminated analog (25).

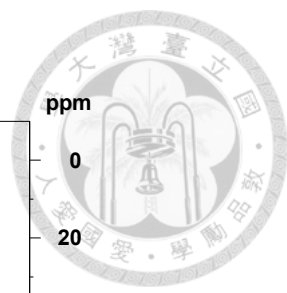


$^1\text{H}$  NMR spectrum for compound **de-r4-Tei**

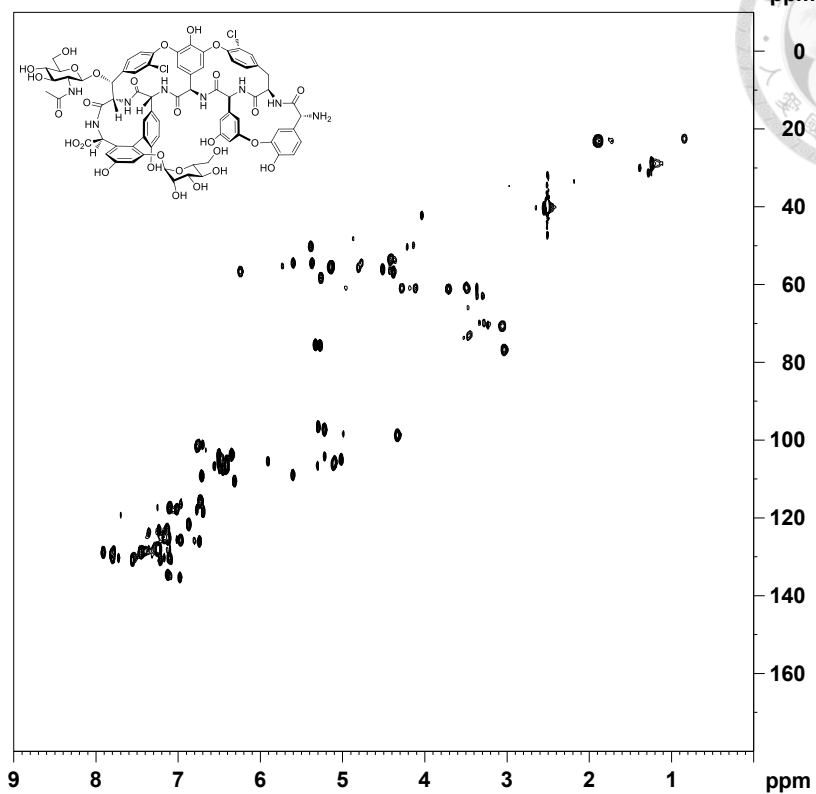


$^{13}\text{C}$  NMR spectrum for compound **de-r4-Tei**

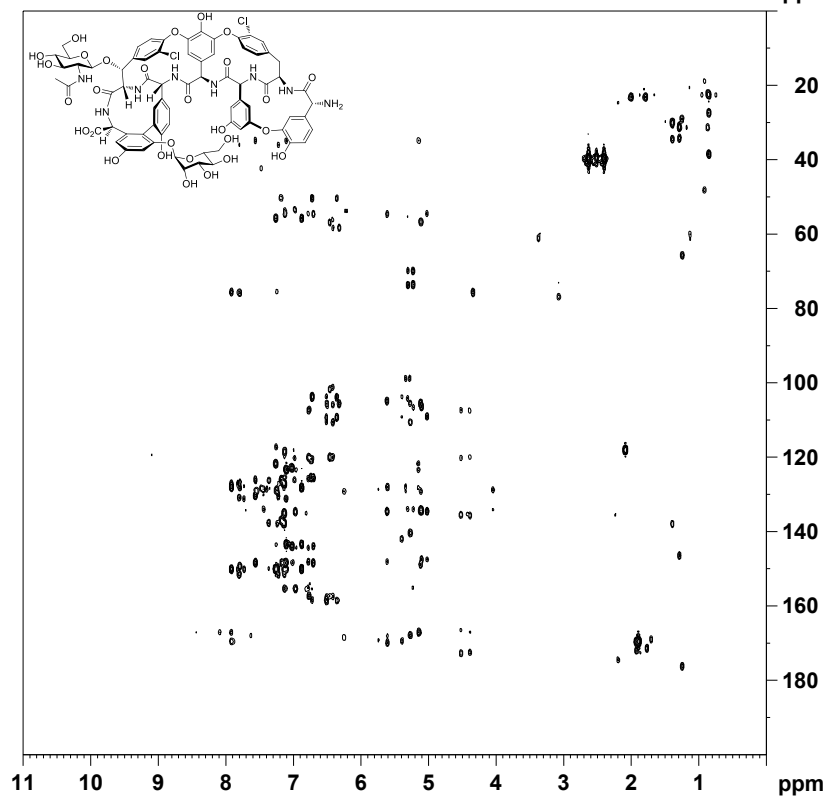




$^1\text{H}$ - $^{13}\text{C}$  HSQC spectrum for compound **de-r4-Tei**



$^1\text{H}$ - $^{13}\text{C}$  HMBC spectrum for compound **de-r4-Tei**



$^1\text{H}$  COSY spectrum for compound **de-r4-Tei**

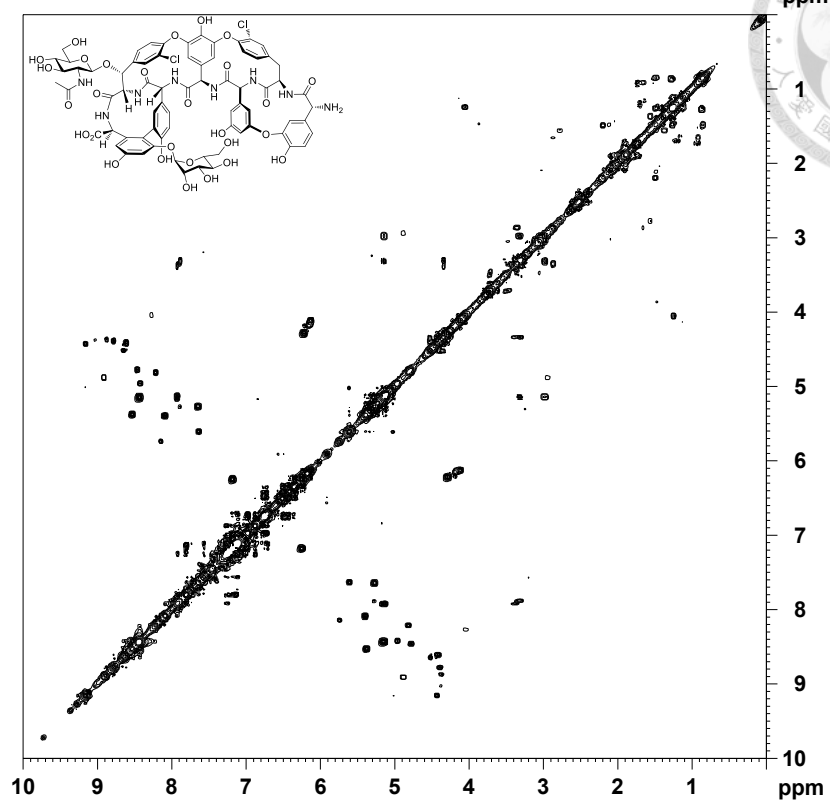
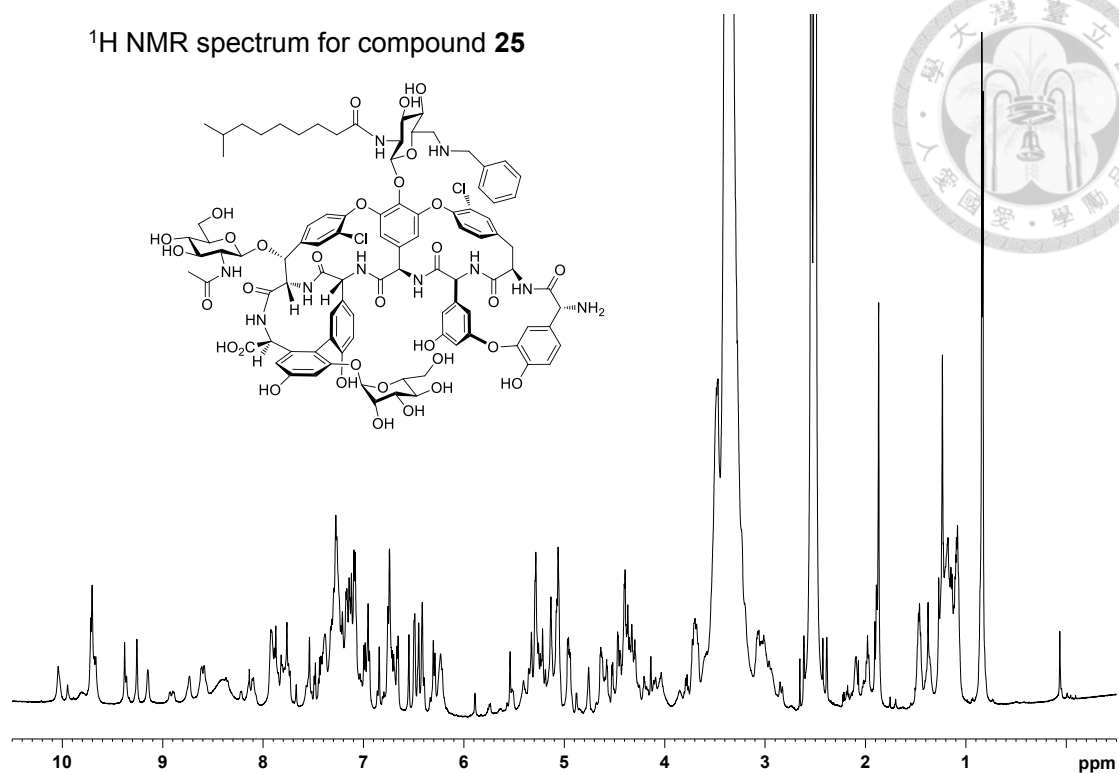
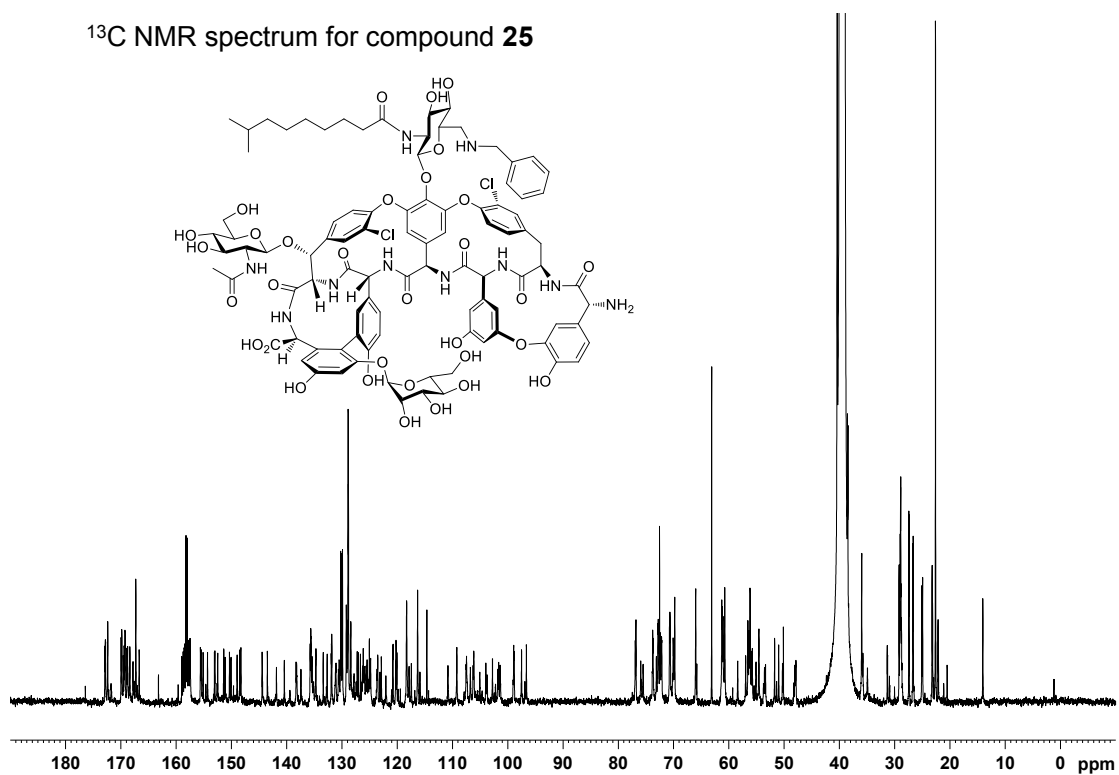


Figure 11. NMR information for **de-r4-Tei**. NMR spectra include  $^1\text{H}$ ,  $^{13}\text{C}$ , HSQC, HMBC, and COSY.

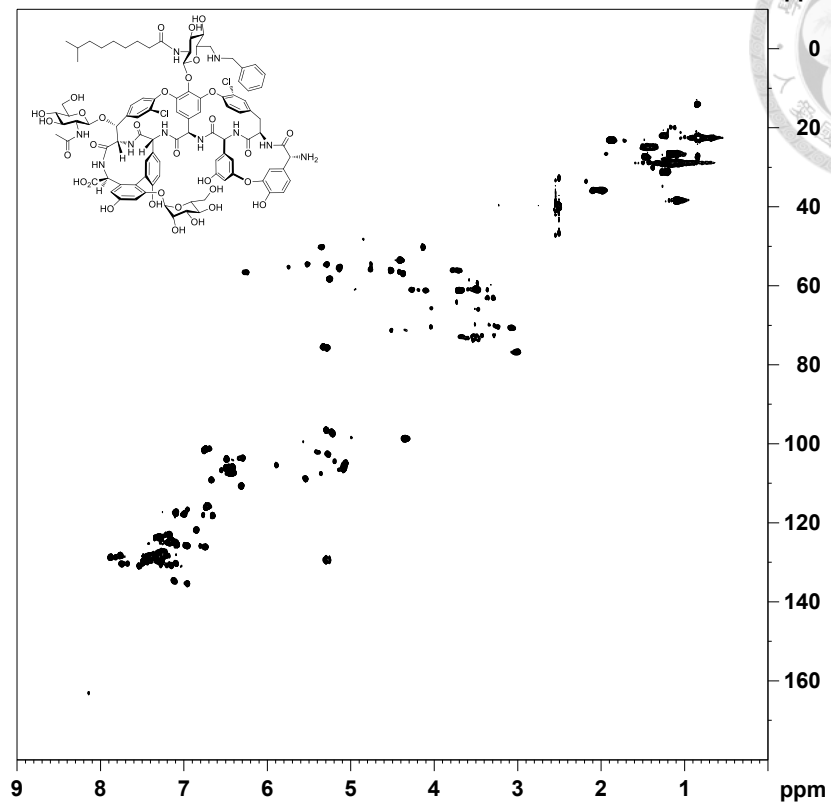
$^1\text{H}$  NMR spectrum for compound **25**



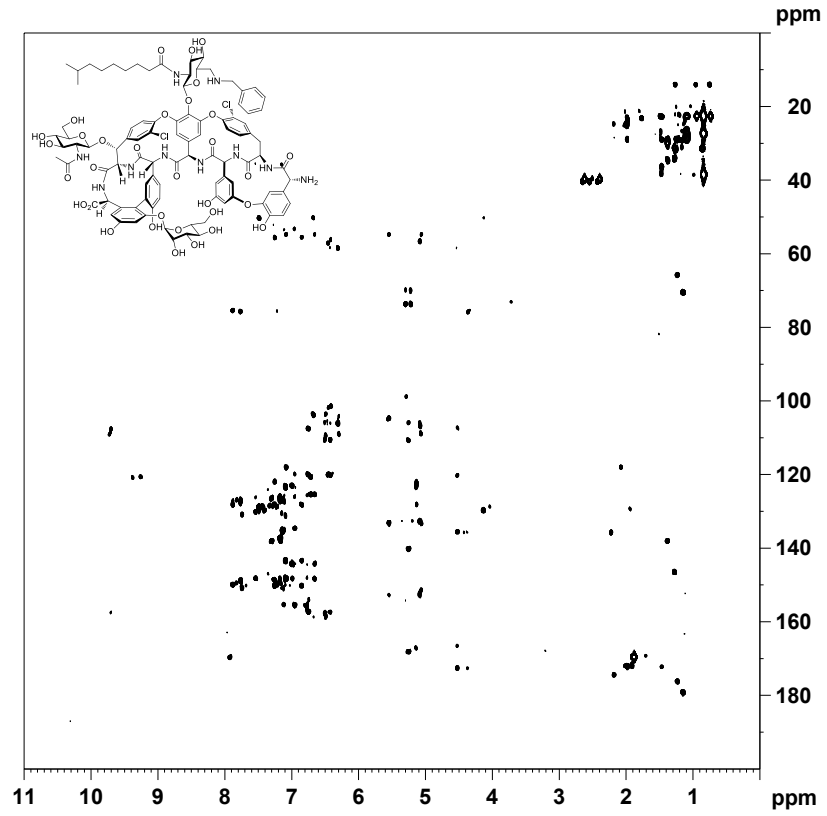
$^{13}\text{C}$  NMR spectrum for compound **25**



$^1\text{H}$ - $^{13}\text{C}$  HSQC spectrum for compound **25**



$^1\text{H}$ - $^{13}\text{C}$  HMBC spectrum for compound **25**



$^1\text{H}$  COSY spectrum for compound 25

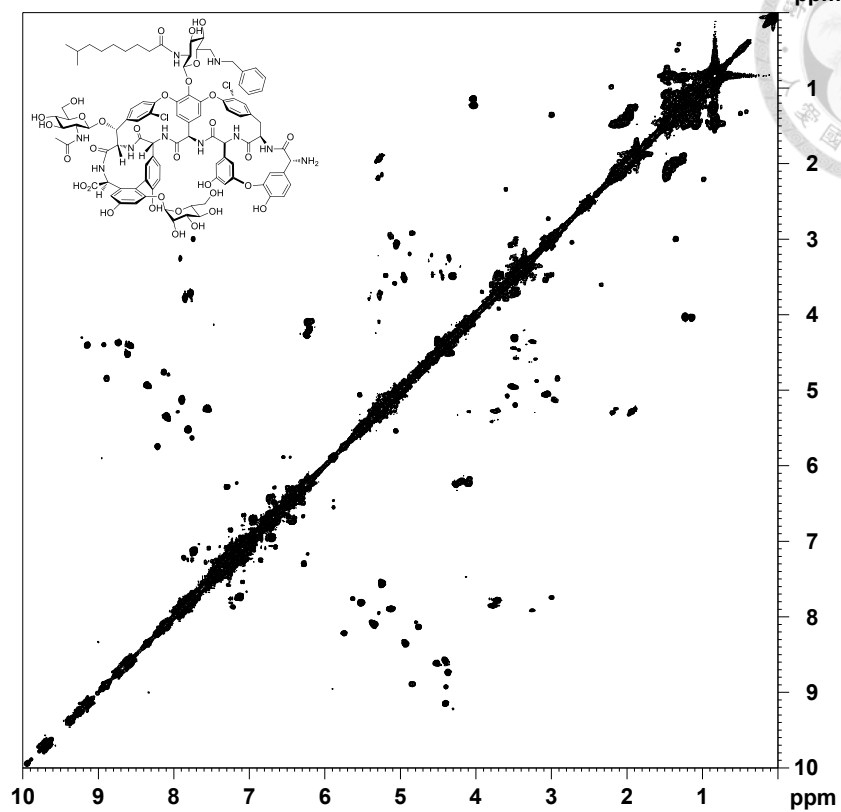
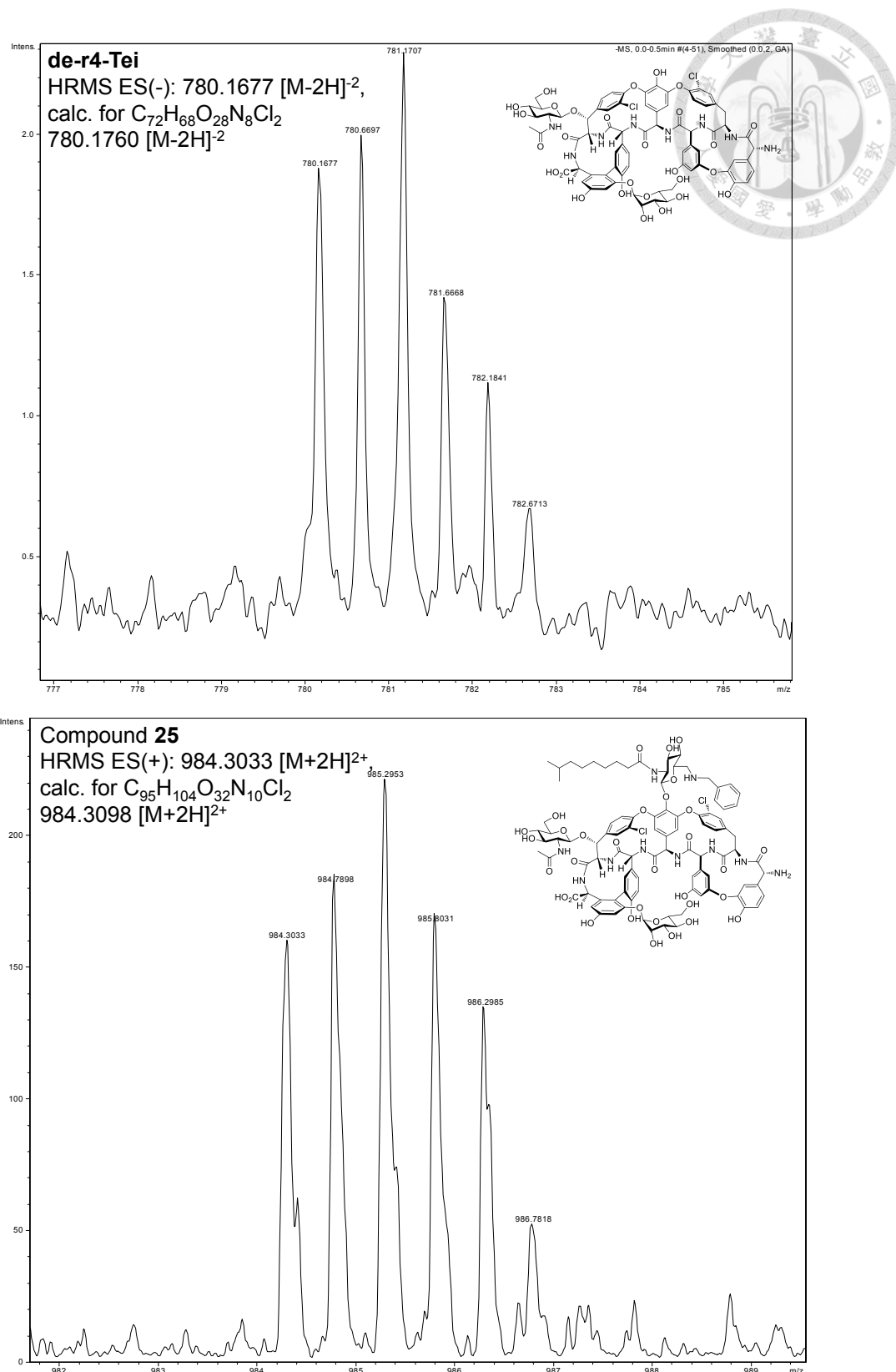


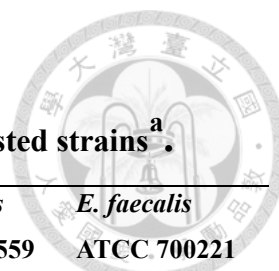
Figure 12. NMR information for analog 25. NMR spectra include  $^1\text{H}$ ,  $^{13}\text{C}$ , HSQC, HMBC, and COSY.





**Figure 13. HRMS data for representative compounds.**

High resolution mass spectra were obtained for compounds **de-r4-Tei** and **25**, in which experimental and theoretical values are designated.

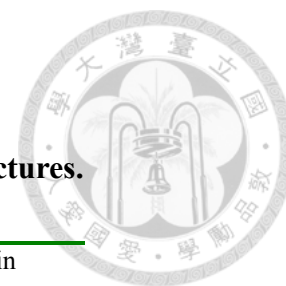


**Table 1. MICs of vancomycin, teicoplanin and analogs against tested strains<sup>a</sup>.**

strains <sup>b</sup> cmpds	<i>E. faecalis</i> ATCC 29302	<i>E. faecalis</i> ATCC 33186	<i>E. faecalis</i> ATCC 51299	<i>E. faecalis</i> ATCC 51559	<i>E. faecalis</i> ATCC 700221
Vancomycin	1	4	>64	>32	>64
Tei (1)	0.125	0.25	0.5	>32	>32
<b>2</b>	8	>32	>32	>32	>32
<b>3</b>	8	>32	>32	>32	>32
<b>4</b>	8	>32	>32	>32	>32
<b>5</b>	2	>32	>32	>32	>32
<b>6</b>	0.125	0.25	0.25	>32	>32
<b>7</b>	>64	>64	>64	>64	>64
Oxo-Tei (9)	0.125	0.25	0.25	>32	>32
<b>10</b>	0.5	0.5	0.5	1	4
<b>11</b>	0.0625	0.0625	0.125	8	32
<b>25</b>	0.125	0.25	0.5	0.25	2

a: The concentration unit is  $\mu\text{g/mL}$ ; MICs were determined as concentrations for which no growth was observed against the strain tested. Experiments were performed in double duplicate.

b: ATCC 29302: a standard strain; ATCC 33186: an antibiotics sensitive strain; ATCC 51299: a low-level vancomycin-resistant (VRE, VanB type); ATCC 51559: a multidrug-resistant strain (VRE, ampicillin, ciprofloxacin, gentamicin, rifampin, teicoplanin, and vancomycin); ATCC 700221: a strain resistant to vancomycin (VRE).



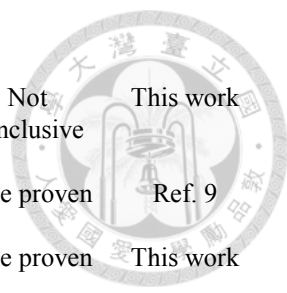
**Table 2. Data collection and refinement statistics for Dbv29 structures.**

	Dbv29	Dbv29/Teicoplanin
<b>Data collection</b>		
Space group	P6 <sub>1</sub> 22	P2 <sub>1</sub>
Cell dimensions		
<i>a</i> , <i>b</i> , <i>c</i> (Å)	66.09, 66.09, 790.46	61.79, 150.78, 124.85
$\alpha$ , $\beta$ , $\gamma$ (°)	90.00, 90.00, 120.00	90.00, 98.40, 90.00
Resolution (Å)	30.00-3.21 (3.32-3.21) *	30.00-1.93 (1.96-1.93)
<i>R</i> <sub>merge</sub>	8.3 (16.0)	9.1 (57.3)
<i>I</i> / $\sigma$ <i>I</i>	24.8 (17.0)	15.0 (2.9)
Completeness (%)	95.7 (92.0)	97.8 (97.1)
Redundancy	9.7 (9.8)	4.6 (4.6)
Wavelength (Å)	0.97315	1
Temperature (K)	100	100
<b>Refinement</b>		
Resolution (Å)	3.21	1.93
No. reflections	16838	155629
<i>R</i> <sub>work</sub> / <i>R</i> <sub>free</sub>	0.237/0.291	0.160/0.205
No. atoms		
Protein	7710	16104
Ligand/ion	106	615
Water	69	1400
<i>B</i> -factors		
Protein	29.2	30.9
Ligand/ion	24.3	57.7
Water	9.0	37.7
R.m.s. deviations		
Bond lengths (Å)	0.005	0.019
Bond angles (°)	0.837	2.030
PDB code	2WDW	4K3T

\*Values in parentheses are for highest-resolution shell.

**Table 3. Relative enzymatic activities of Dbv29 mutants.**

Mutants	Expression	Flavinylolation/ FAD binding	Relative activity <sup>a</sup>	Rationale	Outcome	Source of result
C26A	+	+	1.11	Putative point of active site general base	Role disproven	Ref. 9
C92A	+	+	0.95	Putative point of active site general base	Role disproven	Ref. 9
C151A	+	+	0.23	Putative point of active site general base/ Putative point of covalent attachment	Role disproven/ Role proven	Ref. 9
C161A	+	+	0.84	Putative point of active site general base	Role disproven	Ref. 9
H91A	+	+	0.11	Putative point of covalent attachment	Role proven	Ref. 9
I401A	+	+	0.09	Potential ligand-binding residue	Role proven	This work
I401W	+	+	0.02	Potential ligand-binding residue	Role proven	This work
N44A	+	+	0.87	Potential interface interacting residue	Role proven	This work
R41A	+	+	0.72	Potential interface interacting residue	Role proven	This work
R360E	+	+	0.76	Putative point affecting FAD attachment	Role disproven	This work
R360L	+	+	0.23	Putative point affecting FAD attachment	Role disproven	This work
S364K	+	+	0.04	Potential ligand-binding residue/ Putative point affecting FAD attachment	Role proven/ Role disproven	This work
S364R	+	+	0.08	Potential ligand-binding residue/ Putative point affecting FAD attachment	Role proven/ Role disproven	This work
T366A	+	+	0.11	Potential ligand-binding residue/ Putative point affecting FAD attachment	Role proven/ Role disproven	This work
T366E	+	+	0.05	Potential ligand-binding residue/ Putative point affecting FAD attachment	Role proven/ Role disproven	This work
T366L	+	+	0.13	Potential ligand-binding residue/ Putative point affecting FAD attachment	Role proven/ Role disproven	This work
W399A	+	+	—	Potential ligand-binding residue	Role proven	This work
W399F	+	+	0.16	Potential ligand-	Role proven	This work

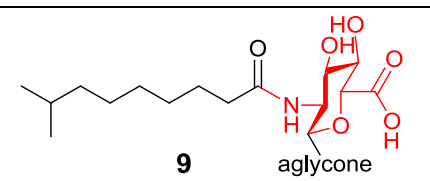
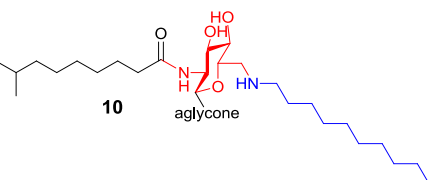
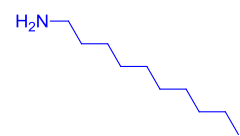
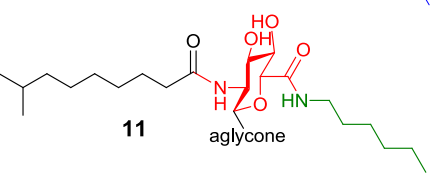
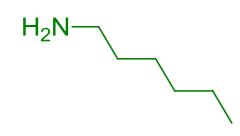
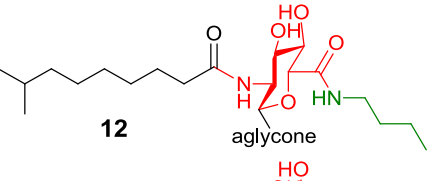
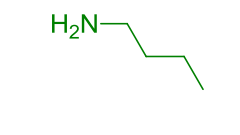
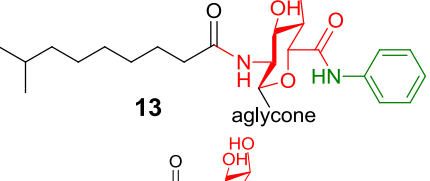
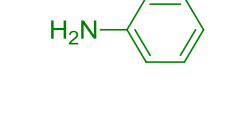
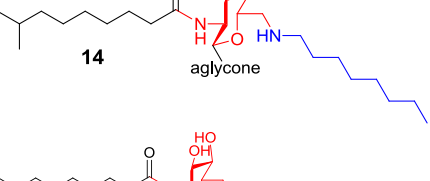

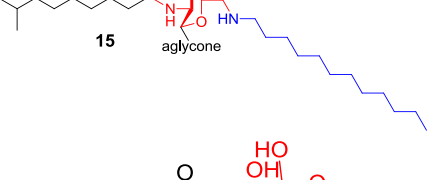
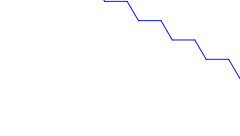
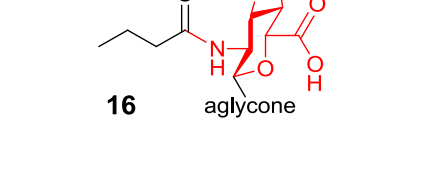


Y135F	—	—	—	binding residue Putative point affecting FAD attachment	Not conclusive	This work
Y165F	+	+	0.70	Putative point of active site general base	Role proven	Ref. 9
Y165W	+	+	0.03	Putative point of active site general base	Role proven	This work
Y370F	+	+	0.10	Putative point affecting FAD attachment	Role disproven	This work
Y403F	+	+	0.65	Putative point affecting FAD attachment	Role disproven	This work
Y428F	+	+	1.00	Putative point affecting FAD attachment	Role disproven	This work
Y453F	+	+	0.82	Putative point affecting FAD attachment	Role disproven	This work
Y470F	+	+	0.21	Putative point affecting FAD attachment	Role disproven	This work
Y473F	+	+	0.23	Putative point of active site general base	Role proven	Ref. 9
Y473E	+	+	—	Putative point of active site general base	Role proven	This work
H91A/C151A	+	+	0.05	Putative point of covalent attachment	Role proven	Ref. 9
R41A/N44A	+	+	0.59	Potential interface interacting residue	Role proven	This work
R360E/T366E	+	+	0.03	Putative point affecting FAD attachment	Role disproven	This work
R360L/T366A	+	+	0.03	Putative point affecting FAD attachment	Role disproven	This work
R360L/T366E	+	+	0.01	Putative point affecting FAD attachment	Role disproven	This work
Y165F/Y473F	+	—	—	Putative point of active site general base	Role proven	Ref. 9
Y428F/Y453F	+	+	1.08	Putative point affecting FAD attachment	Role disproven	This work
Y370F/Y403F	+	+	0.05	Putative point affecting FAD attachment	Role disproven	This work

a. The products of mutants were analyzed by HPLC; peak areas were integrated and averaged for 3 triplicate points in 1.5 hours; the reaction rates were calculated using linear regression equation; relative activities were determined by dividing individual reaction rate with that of WT; the relative activity of WT is 1.0.

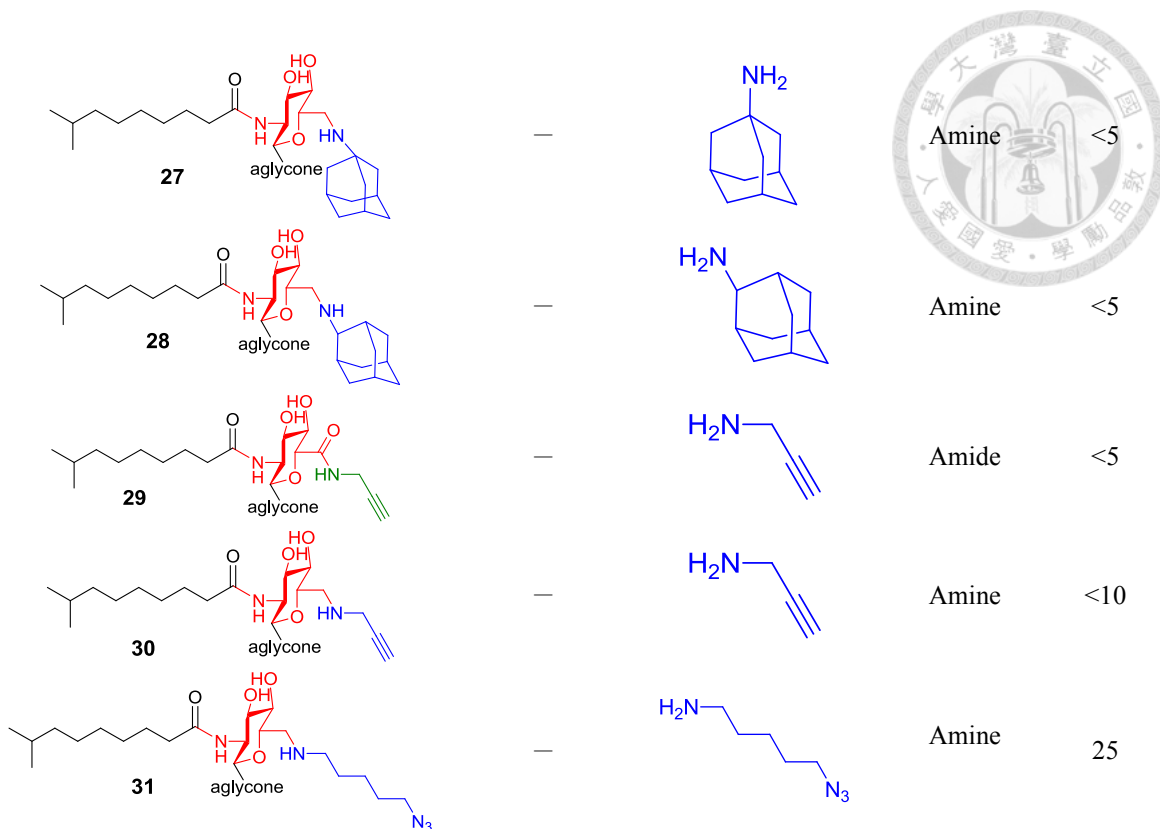


**Table 4. Product yields of oxidized, aminated or amidated teicoplanin analogs synthesized by Dbv29.**

Product	Coenzyme A derivative (Carbon No.) <sup>1</sup>	Amine added	Product type: amide or amine	Yield (%) <sup>2</sup>
 <p><b>9</b> aglycone</p>	—	—	—	100
 <p><b>10</b> aglycone</p>	—		Amine	25
 <p><b>11</b> aglycone</p>	—		Amide	10
 <p><b>12</b> aglycone</p>	—		Amide	<10
 <p><b>13</b> aglycone</p>	—		Amide	10
 <p><b>14</b> aglycone</p>	—		Amine	25
 <p><b>15</b> aglycone</p>	—		Amine	17
 <p><b>16</b> aglycone</p>	C4	—	—	100



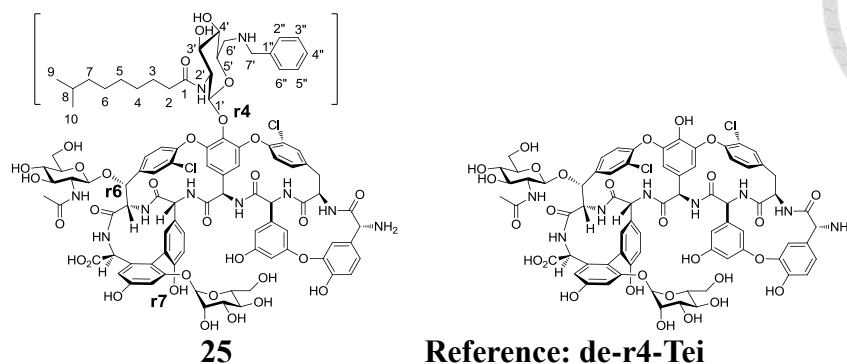
17		C6	—	—	100
18		C8	—	—	100
19		C10	—	—	100
20		C6		Amide	10
21		C8		Amide	10
22		C6		Amine	30
23		C8		Amine	30
24		—		Amide	<10
25		—		Amine	46
26		—		Amine	<5



1. Dbv8 was used to add various lengths of acyl side chains from corresponding coenzyme A derivatives to the C2 amine group. If no CoA derivative is listed, Teicoplanin was used directly.
2. The product yields were determined by HPLC; yield = (peak area of aminated or amidated product)/(peak areas of aminated and amidated and oxidized products) × 100%.



**Table 5. NMR assignments for compound 25 as determined from spectra shown in Figures 11 and 12.**

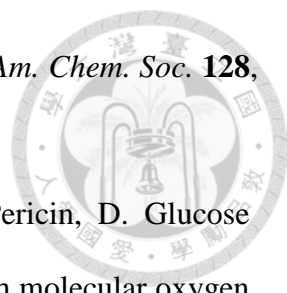


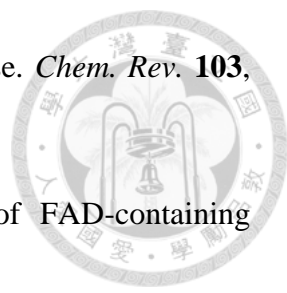
Position	$^{13}\text{C-NMR}$ ( $\delta$ ppm)	$^1\text{H-NMR}$ ( $\delta$ ppm)
<b>C-1</b>	172.4	—
<b>C-2</b>	28.9	1.18, 1.18 m (multiplicity)
<b>C-3</b>	28.9	1.08, 1.08 m
<b>C-4</b>	26.8	1.0 , 1.08 m
<b>C-5</b>	26.7	1.17, 1.17 m
<b>C-6</b>	25.1	1.46, 1.46 m
<b>C-7</b>	25.0	1.38, 1.38 m
<b>C-8</b>	27.4	1.46 m
<b>C-9/10</b>	22.7	0.84, 0.84, 0.84 d (J 6.5)
<b>C-1'</b>	102.8	5.27 m
<b>C-2'</b>	72.4	3.28 m
<b>C-3'</b>	72.8	3.67 m
<b>C-4'</b>	72.2	3.28 m
<b>C-5'</b>	75.5	5.42 m
<b>C-6'</b>	51.7	3.85, 3.85 m
<b>C-1''</b>	133.3	—
<b>C-2''/C-6''</b>	129.2	7.27 m
<b>C-3''/C-5''</b>	130.1	7.48 m
<b>C-4''</b>	128.9	7.32 m
<b>C-7''</b>	48.1	4.83, 4.83 m

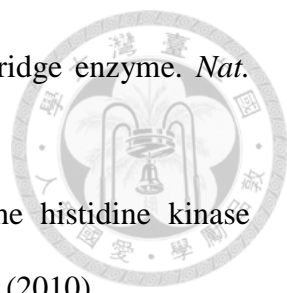


## References

1. Fischbach, M.A. & Walsh, C.T. Antibiotics for Emerging Pathogens. *Science* **325**, 1089-1093 (2009).
2. Silver, L.L. Multi-targeting by monotherapeutic antibacterials. *Nat Rev Drug Discov* **6**, 41-55 (2007).
3. Payne, D.J., Gwynn, M.N., Holmes, D.J. & Pompliano, D.L. Drugs for bad bugs: confronting the challenges of antibacterial discovery. *Nat Rev Drug Discov* **6**, 29-40 (2007).
4. Louis B, R. Antimicrobial resistance in gram-positive bacteria. *Am. J. Infect. Control* **34**, S11-S19 (2006).
5. Murray, B.E. Vancomycin-Resistant Enterococcal Infections. *New Engl. J. Med.* **342**, 710-721 (2000).
6. Chan, P.F., Holmes, D.J. & Payne, D.J. Finding the gems using genomic discovery: antibacterial drug discovery strategies – the successes and the challenges. *Drug Discovery Today: Therapeutic Strategies* **1**, 519-527 (2004).
7. Khare, M. & Keady, D. Antimicrobial therapy of methicillin resistant *Staphylococcus aureus* infection. *Expert Opin Pharmacother* **4**, 165-177 (2003).
8. Marsden, A.F.A. *et al.* Engineering Broader Specificity into an Antibiotic-Producing Polyketide Synthase. *Science* **279**, 199-202 (1998).
9. Li, Y.-S. *et al.* A Unique Flavin Mononucleotide-Linked Primary Alcohol Oxidase for Glycopeptide A40926 Maturation. *J. Am. Chem. Soc.* **129**, 13384-13385 (2007).
10. Ho, J.-Y. *et al.* Glycopeptide Biosynthesis: Dbv21/Orf2\* from dbv/tcp Gene Clusters Are N-Ac-Glm Teicoplanin Pseudoaglycone Deacetylases and Orf15

- 
- from cep Gene Cluster Is a Glc-1-P Thymidyltransferase. *J. Am. Chem. Soc.* **128**, 13694-13695 (2006).
11. Leskovac, V., Trivic, S., Wohlfahrt, G., Kandrac, J. & Pericin, D. Glucose oxidase from *Aspergillus niger*: the mechanism of action with molecular oxygen, quinones, and one-electron acceptors. *Int. J. Biochem. Cell Biol.* **37**, 731-750 (2005).
  12. Snook, C.F., Tipton, P.A. & Beamer, L.J. Crystal Structure of GDP-Mannose Dehydrogenase: A Key Enzyme of Alginate Biosynthesis in *P. aeruginosa*<sup>†,‡</sup>. *Biochemistry (Mosc)*. **42**, 4658-4668 (2003).
  13. Campbell, R.E., Mosimann, S.C., van de Rijn, I., Tanner, M.E. & Strynadka, N.C.J. The First Structure of UDP-Glucose Dehydrogenase Reveals the Catalytic Residues Necessary for the Two-fold Oxidation<sup>†,‡</sup>. *Biochemistry (Mosc)*. **39**, 7012-7023 (2000).
  14. Chan, H.-C. *et al.* Regioselective deacetylation based on teicoplanin-complexed Orf2\* crystal structures. *Molecular BioSystems* **7**, 1224-1231 (2011).
  15. Li, T.-L. *et al.* Biosynthetic Gene Cluster of the Glycopeptide Antibiotic Teicoplanin: Characterization of Two Glycosyltransferases and the Key Acyltransferase. *Chem. Biol.* **11**, 107-119 (2004).
  16. Alexeev, I., Sultana, A., Mäntsälä, P., Niemi, J. & Schneider, G. Aclacinomycin oxidoreductase (AknOx) from the biosynthetic pathway of the antibiotic aclacinomycin is an unusual flavoenzyme with a dual active site. *Proc. Natl. Acad. Sci. USA* **104**, 6170-6175 (2007).
  17. Huang, C.-H. *et al.* Crystal Structure of Glucooligosaccharide Oxidase from *Acremonium strictum*: A NOVEL FLAVINYLLATION OF 6-S-CYSTEINYLL, 8 $\alpha$ -N1-HISTIDYLL FAD. *J. Biol. Chem.* **280**, 38831-38838 (2005).

- 
18. Whittaker, J.W. Free Radical Catalysis by Galactose Oxidase. *Chem. Rev.* **103**, 2347-2364 (2003).
19. Dym, O. & Eisenberg, D. Sequence-structure analysis of FAD-containing proteins. *Protein Sci.* **10**, 1712-1728 (2001).
20. Otwinowski, Z. & Minor, W. [20] Processing of X-ray diffraction data collected in oscillation mode, in *Methods in Enzymology*, Vol. Volume **276**. (ed. Charles W. Carter, Jr.) 307-326 (Academic Press, 1997).
21. Kantardjieff, K.A. & Rupp, B. Matthews coefficient probabilities: Improved estimates for unit cell contents of proteins, DNA, and protein–nucleic acid complex crystals. *Protein Sci.* **12**, 1865-1871 (2003).
22. Vagin, A. & Teplyakov, A. MOLREP: an Automated Program for Molecular Replacement. *J. Appl. Crystallogr.* **30**, 1022-1025 (1997).
23. McRee, D.E. XtalView/Xfit—A Versatile Program for Manipulating Atomic Coordinates and Electron Density. *J. Struct. Biol.* **125**, 156-165 (1999).
24. Brünger, A.T. *et al.* Crystallography & NMR System: A New Software Suite for Macromolecular Structure Determination. *Acta Crystallographica Section D* **54**, 905-921 (1998).
25. Murshudov, G.N., Vagin, A.A. & Dodson, E.J. Refinement of Macromolecular Structures by the Maximum-Likelihood Method. *Acta Crystallographica Section D* **53**, 240-255 (1997).
26. Schuck, P. Size-Distribution Analysis of Macromolecules by Sedimentation Velocity Ultracentrifugation and Lamm Equation Modeling. *Biophys. J.* **78**, 1606-1619 (2000).
27. Borch, R.F., Bernstein, M.D. & Durst, H.D. Cyanohydrinborate anion as a selective reducing agent. *J. Am. Chem. Soc.* **93**, 2897-2904 (1971).

- 
28. Winkler, A. *et al.* A concerted mechanism for berberine bridge enzyme. *Nat. Chem. Biol.* **4**, 739-741 (2008).
29. Koteva, K. *et al.* A vancomycin photoprobe identifies the histidine kinase VanSsc as a vancomycin receptor. *Nat Chem Biol* **6**, 327-329 (2010).
30. Higgins, D.L. *et al.* Telavancin, a Multifunctional Lipoglycopeptide, Disrupts both Cell Wall Synthesis and Cell Membrane Integrity in Methicillin-Resistant *Staphylococcus aureus*. *Antimicrob. Agents Chemother.* **49**, 1127-1134 (2005).
31. Roecker, A.M. & Pope, S.D. Dalbavancin: a lipoglycopeptide antibacterial for Gram-positive infections. *Expert Opinion on Pharmacotherapy* **9**, 1745-1754 (2008).
32. Linden, P.K. Vancomycin resistance: are there better glycopeptides coming? *Expert review of anti-infective therapy* **6**, 917-928 (2008).
33. Bailey, J. & Summers, K.M. Dalbavancin: a new lipoglycopeptide antibiotic. *Am. J. Health. Syst. Pharm.* **65**, 599-610 (2008).
34. Kim, A., Kuti, J.L. & Nicolau, D.P. Review of dalbavancin, a novel semisynthetic lipoglycopeptide. *Expert opinion on investigational drugs* **16**, 717-733 (2007).

**Section 2. Structure and mechanism of non-heme iron-SAM dependent methyltransferase and its engineering to hydratase.**

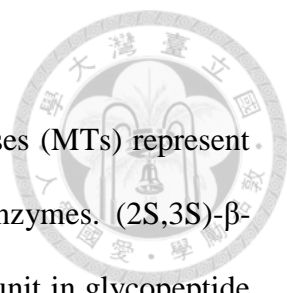


## 摘要

S-腺苷甲硫氨酸(SAM)必需之甲基轉移酶(methyltransferases)為最具多樣性及重要生物活性功能的酵素之一。(2S,3S)- $\beta$ -甲基苯丙氨酸( $\beta$ MePhe)為非蛋白氨基酸，是組成醣肽類抗生素 mannopeptimycin 六環肽類中的部分化學分子。我們先前的研究發現，mannopeptimycin 生合成 *mppJ* 基因利用 S-腺苷甲硫氨酸(SAM)將苯丙酮酸(Ppy)上苯甲基的碳甲基化而產生 $\beta$ -甲基苯丙酮酸( $\beta$ -MePpy)，雖然其苯甲基的碳具有部分酸性，但是並不表示這碳原子一定會進行親和性取代反應。特別是純化之 MppJ 蛋白質呈現藍綠色，意指 MppJ 有金屬離子參與催化反應，但不同於目前已知鎂( $Mg^{2+}$ )或鈣( $Ca^{2+}$ )離子必需之甲基轉移酶以及自由基反應機制之甲基轉移酶。我們利用 X 光蛋白質結晶學與生物化學的方法，解析出 MppJ 的蛋白質結構以及催化作用機制，藉由分析蛋白質與受質(substrates)及產物(products)複合體蛋白結構，我們發現 MppJ 為目前唯一鐵離子( $Fe^{3+}$ )及 S-腺苷甲硫氨酸(SAM)必需之甲基轉移酶，並且知道 MppJ 如何在碳原子上發生甲基化反應的催化機制。多個複合體蛋白結構使我們了解 MppJ 利用鐵離子與 $\alpha$ 酮基酸( $\alpha$ -ketoacid)結合配位與活化，並且發展出兩個水分子裝置來控制酵素專一性進行碳上甲基化反應。如此發現使我們更進一步將甲基轉移酶改變成結構與功能不相關的新酵素。我們藉由改變金屬離子的配位化學，使得原本酵素變成具有水合酶及甲基轉移酶的雙重功能，進而產生具有立體特異性的新化學分子。

關鍵字：mannopeptimycin，非蛋白氨基酸，甲基轉移酶，S-腺苷甲硫氨酸(SAM)，苯丙酮酸(Ppy)。

## Abstract



The S-adenosyl-L-methionine (SAM)-dependent methyltransferases (MTs) represent one of the most diverse and biologically important classes of enzymes. (2S,3S)- $\beta$ -methylphenylalanine, a nonproteinogenic amino acid, is a building unit in glycopeptide antibiotic mannopeptimycin. The *mppJ* gene product in the mannopeptimycin biosynthetic gene cluster was determined to be the committed MT to methylate the benzylic carbon of phenylpyruvate (Ppy) into  $\beta$ -MePpy. The benzylic carbon of Ppy has some extent of acidity, but its acting as an operational nucleophile was not concluded. The purified MppJ displays a turquoise color implicating involvement of a metal ion. The solved crystal structures revealed MppJ the first ferric ion/SAM-dependent MT. Additional four structures in binary and ternary complexes illustrated the molecular mechanism for the metal ion dependent MT reaction. MppJ has evolved a non-heme iron center to bind, orientate and activate the  $\alpha$ -ketoacid substrate and meanwhile developed a sandwiched bi-water device to avoid formation of the unwanted reactive oxo-Fe(IV) species during the C-MT reaction. This unprecedented discovery further prompted us to convert the MT into a structurally/functionally unrelated new enzyme. Through stepwise-protein engineering and manipulation of coordination chemistry MTs were engineered to perform both non-heme iron dependent hydration and methyltransferation reactions for stereo-specific new compounds. The process was validated by five crystal structures.

Keywords: mannopeptimycin, nonproteinogenic amino acid, methyltransferase, S-adenosyl-L-methionine (SAM), phenylpyruvic acid (Ppy).



## List of Figures



<b>Figure 1. Chemical structures of mannopeptimycins. ....</b>	<b>74</b>
<b>Figure 2. Multiple sequence alignments for MppJ and homologues. ....</b>	<b>75</b>
<b>Figure 3. Crystal structures of MppJ complexed with substrates or products. ....</b>	<b>77</b>
<b>Figure 4. Gel filtration and analytical ultracentrifugation analyses of MppJ. ....</b>	<b>78</b>
<b>Figure 5. Schematic topology of MppJ. ....</b>	<b>79</b>
<b>Figure 6. Conformational change and SAM/SAH-binding. ....</b>	<b>80</b>
<b>Figure 7. Ppy binding site and metal (Fe<sup>3+</sup>) coordination in MppJ structure. ....</b>	<b>81</b>
<b>Figure 8. Metal ion determination by X-ray absorption spectroscopy (XAS) and electron paramagnetic resonance (EPR). ....</b>	<b>82</b>
<b>Figure 9. Active site structure of MppJ and proposed catalytic mechanism. ....</b>	<b>83</b>
<b>Figure 10. Isothermal titration calorimetry (ITC) analyses of MppJ. ....</b>	<b>84</b>
<b>Figure 11. HPLC traces of enzymatic reactions for MppJ WT, D244L, and D244E mutants. ....</b>	<b>85</b>
<b>Figure 12. Active site structures of MppJ D244E and D244L mutants. ....</b>	<b>86</b>
<b>Figure 13. Chiral HPLC analyses of MppJ. ....</b>	<b>87</b>
<b>Figure 14. Ppy/4HPpy are covalently linked to Cys319. ....</b>	<b>88</b>
<b>Figure 15. Active site structures of MppJ mutants and proposed mechanism for new compound. ....</b>	<b>89</b>



## List of Tables

**Table 1. Data collection, phasing and refinement statistics for MppJ structures. .... 90**

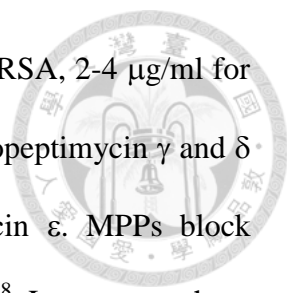
**Table 2. Relative enzymatic activities (MT), enzyme colors and proposed residue functions. .... 92**



## 1. Introduction

### 1.1 Mannopectimycin

Mannopectimycins (MPPs) are a group of metabolites, including mannopectimycin  $\alpha$ ,  $\beta$ ,  $\gamma$ ,  $\delta$ , and  $\epsilon$ , which are produced by *Streptomyces hygroscopicus*, LL-AC98 and discovered in 1958 by scientists at Wyeth<sup>1, 2</sup>. MPPs are active against Gram-positive bacteria but not Gram-negative bacteria. Recently, MPPs were reexamined in a program searching for new classes of antibiotics to combat multidrug resistant pathogens<sup>3</sup>. Remarkably, MPPs exhibited in vitro activity against clinically important pathogens like methicillin-resistant *Staphylococcus aureus* (MRSA) and vancomycin-resistant *Enterococci* (VRE)<sup>4</sup>. This prompted further studies to investigate the structures of the MPPs, their pharmacological activities and mechanisms, biosynthesis and chemical modifications. The mannopectimycin complex is composed of five major members, designated mannopectimycin  $\alpha$ - $\epsilon$  (Figure 1). The series is based on a cyclic hexapeptide core containing D, L stereoisomers of an unprecedented amino acid,  $\alpha$ -amino- $\beta$ -[4<sup>7</sup>-(2<sup>7</sup>-iminoimidazolidinyl)]- $\beta$ -hydroxypropionic acid (Aiha, also termed  $\beta$ -hydroxyenduracididine,  $\beta$ -OH-End). Mannopectimycin  $\alpha$ ,  $\gamma$ ,  $\delta$ , and  $\epsilon$  are O-glycosylated with mannose disaccharide moieties on the tyrosine phenolic OH and N-glycosylated with mannose monosaccharide moiety on the D-Aiha residue. In mannopectimycin  $\gamma$ - $\epsilon$  the terminal mannose bears an isovalerate ester. The unprecedented structures of these molecules illustrate the challenges facing structure elucidation and lead optimization. MPP  $\alpha$ ,  $\gamma$  and  $\delta$  all show moderate to good antibacterial activities. The MPP  $\epsilon$  is most active of the family against MRSA and VRE<sup>2, 5-8</sup>. It is indicated that the dimannosyl moiety is required for MPPs activity and the isovaleryl group on the mannose affects the activity. The minimum inhibitory



concentrations (MICs) are 2-4  $\mu\text{g/ml}$  for *Staphylococci* including MRSA, 2-4  $\mu\text{g/ml}$  for *Streptococci* and 4-32  $\mu\text{g/ml}$  for *Enterococci* including VRE. Mannopectimycin  $\gamma$  and  $\delta$  demonstrate two to four folds less activity than mannopectimycin  $\epsilon$ . MPPs block bacteria cell wall synthesis by inhibiting peptidoglycan formation<sup>2, 5-8</sup>. Later research on the cellular target of MPPs revealed that MPPs inhibit transglycosylase by binding to its substrate lipid II, which is similar to the mechanism of ramoplanin and vancomycin<sup>2, 5-8</sup>. MPPs don't bind to the D-Ala-D-Ala terminus of lipid II which is the target of vancomycin and explains why mannopectimycin has antibacterial activity against vancomycin-resistant bacteria<sup>2, 5-9</sup>.

Antibacterial and mechanistic studies with Gram-positive and Gram-negative bacteria demonstrated that the primary cellular target was bacterial cell wall biosynthesis, with a specific targeted step different from those targeted by known agents. It has been shown that mannopectimycins act by inhibiting cell wall biosynthesis via a unique mode of action, by binding the membrane bound cell wall precursor lipid II in a fashion different from other lipid II binders such as ramoplanin, mersacidin, and vancomycin. In this binding mode, the mannopectimycins inhibit the essential transglycosylation steps necessary for incorporation of lipid II into nascent peptidoglycan but in a manner that serves to concentrate the antibiotic on the cell surface. These finding further validate lipid II as a viable target for continued antibiotic discovery in that no fewer than four different classes of antibiotics are now known to target lipid II, each with a different detailed mechanism<sup>3, 4, 10</sup>.

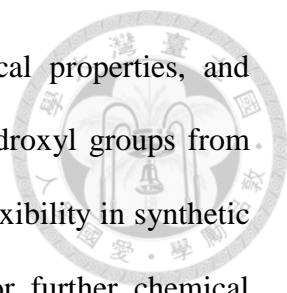
In parallel to chemical semisynthesis, directed biosynthetic and biosynthetic engineering techniques were used to elaborate novel analogues. These approaches were based on the successful cloning, sequencing, and manipulation of biosynthetic gene clusters. The MPP biosynthesis gene cluster was cloned and sequenced by Wyeth

Research from *Streptomyces hygrosopicus* NRRL 30439 and the sequence information was deposited in NCBI GenBank (accession number AY735112). Encoded by 27 genes within the mannopeptimycin biosynthetic gene cluster are enzymes responsible for the generation of the hexapeptide core (nonribosomal peptide synthetases [NRPS]) and tailoring reactions (mannosylation, isovalerylation, hydroxylation, and methylation)<sup>11</sup>.

## 1.2 Methyltransferases and methylation

The S-adenosyl-L-methionine (SAM)-dependent methyltransferases (MTs) represent one of the most diverse and biologically important classes of enzymes<sup>12</sup>. These enzymes utilize the ubiquitous SAM as the methyl donor to methylate a wide variety of such biomaterials as nucleic acids, proteins, lipids, and small molecules<sup>13, 14</sup>. Analysis of SAM-dependent MTs structures indicates that the majority are fundamentally similar, having core folds reminiscent of the canonical Rossmann nucleotide binding fold, with a central, mainly parallel,  $\beta$ -sheet flanked on both sides by  $\alpha$ -helices<sup>13, 14</sup>. Nevertheless, in many cases, these cores have been embellished with additional structural elements tailored to accept a plethora of substrates. The structures of MTs solved so far can be grouped by topology into five classes (classes I–V)<sup>13, 14</sup>. However, even within a class, sequence similarity can be as low as 10%. In SAM-dependent enzymes, the only residues that are loosely conserved lie in the core fold and comprise a glycine-rich sequence (referred to as motif I), which interacts with the amino acid portion of SAM, and an acidic loop (denoted as motif II), which interacts with the ribose hydroxyls<sup>13, 14</sup>. This domain however is highly variable at both the structural and sequence levels<sup>13, 14</sup>, which is consistent with the fact that these enzymes methylate a variety of distinct substrates.

Glycopeptide, macrolide antibiotics and other natural products are frequently modified with nonproteinogenic amino acids and deoxy sugars, many of which are



methylated<sup>15, 16</sup>. Methylation can affect solubility, pharmacological properties, and stability of the natural products<sup>17</sup>. Methylation also can protect hydroxyl groups from undesired modification<sup>17</sup>, and control of methylation can provide flexibility in synthetic schemes when natural products are used as starting materials for further chemical reactions. As with all steps involved in natural product assembly, a thorough understanding of the mechanisms and specificity regarding methylation will facilitate efforts to engineer and modify existing natural product biosynthetic pathways with an eye toward optimizing pharmacological properties or using natural products as a starting point for further synthetic modifications<sup>18-20</sup>.

### 1.3 MppJ

(2*S*,3*S*)- $\beta$ -methylphenylalanine ( $\beta$ MePhe), a nonproteinogenic amino acid, is a component in mannopeptimycin (produced by *Streptomyces hygroscopicus* NRRL3085)<sup>2</sup> and occasionally appears in other secondary metabolites<sup>21</sup>. Mannopeptimycin happens to be a potent drug lead, whose biosynthetic gene cluster (*mpp*) encoding essential proteins for biosynthesis of mannopeptimycin has been identified<sup>11</sup>. In the gene cluster, the *mppJ* gene containing a motif in binding the methyl-donating cofactor SAM was biochemically elucidated as the MT for the benzylic C-methylation on phenylpyruvate (Ppy) instead of the previously thought phenylalanine or mannopeptimycin aglycone<sup>22</sup>. Though the benzylic carbon of Ppy may exhibit certain extent of acidity, turning the carbon a good nucleophile may not be as straightforward as one thought. The fact is when the methyl acceptors are carbon atoms other than oxygen and nitrogen atoms (O- and NMT, respectively), the reaction mechanisms become more complicated and depend on the electronic properties of the acceptor molecules in SAM-dependent alkylation reactions<sup>23</sup>. This notion makes MppJ unprecedented as the purified protein is turquoise, implicating involvement of metal

ions in the reaction mechanism but at odds with all known  $\text{Ca}^{2+}/\text{Mg}^{2+}$ -dependent OMTs<sup>24-26</sup> as well as the radical SAM superfamily<sup>27, 28</sup>.



## 2. Materials and Methods

### 2.1 Gene cloning

The *mppJ* gene was amplified from *Streptomyces hygroscopicus* NRRL3085 genomic DNA by polymerase chain reaction (PCR). The product was ligated according to the manufacturers' instruction (Novagene) into the pET-28a(+) expression vector which affords an N-terminal His<sub>6</sub>-tagged protein.

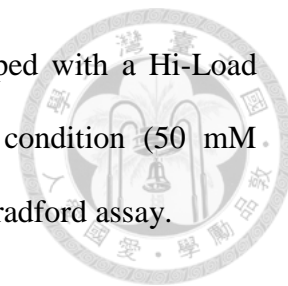
### 2.2 Protein expression and purification

Single colony was grown overnight in Luria broth (LB) medium added with antibiotic kanamycin at 37°C. The overnight culture was inoculated with 50 times volume of pre-warmed LB medium and grew at 37°C until OD<sub>600</sub> 0.6. One millimolar isopropyl-D-thiogalactoside (IPTG) was used to induce the protein expression. Cell pellets were collected and suspended in lysis buffer of 20 mM Tris-HCl (pH 8.0), 500 mM NaCl, 10 % glycerol, 10 mM imidazole, and 0.5 mM phenylmethanesulfonyl fluoride (PMSF). Sonicator or microfluidizer was used to lyse the cells. The His<sub>6</sub>-tagged target protein can be extracted from the lysates by batch Ni<sup>2+</sup>-NTA agarose (Qiagen). The mixtures of lysate-Ni<sup>2+</sup>-NTA agarose were loaded into the column equilibrated with lysis buffer. Wash buffer of 20 mM Tris-HCl (pH 8.0), 500 mM NaCl, 10 % glycerol and 40 mM imidazole was used to remove background proteins and the same buffer with 250 mM imidazole was used to elute the desired proteins. The purity was checked by denaturing sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) and visualized by Coomassie Brilliant Blue staining. The purified proteins were concentrated by ultrafiltration and buffer was exchanged to remove excess imidazole and ions by several times of the dilution/reconcentration procedure. The final protein with desired concentration that contained glycerol up to 50% was stored at -20°C. Gel





filtration analysis was performed using Äkta FPLC system equipped with a Hi-Load Superdex 200 column (Amersham Bioscience) in the isocratic condition (50 mM HEPES, pH 8.0). Protein concentrations were estimated using the Bradford assay.



### 2.3 Crystallization and data collection

The purified protein in a buffer solution (50 mM HEPES at pH 7.5, 100 mM CaCl<sub>2</sub>) was concentrated to 24 mg ml<sup>-1</sup> and added with 6 mM substrates (SAM and Ppy) for crystallization. The initial crystallization condition was determined with a sparse crystallization matrix at 20°C using the hanging drop-vapor diffusion technique. The optimized crystallization condition consisted of 16% PEG 3350, 200 mM NaI (1 µL of protein solution: 1 µL of the reservoir solution). Crystals can be obtained within 10 days. Crystals were transferred into the crystallization buffer with additional 20 % ethylene glycol and flash-frozen in liquid nitrogen before data collection. X-ray diffraction data sets were collected on an operating temperature of 100 K and an ADSC Quantum-315 or Quantum-210 CCD detectors at beamlines 13B1, 13C1 and 15A1 of the National Synchrotron Radiation Research Center (Taiwan), beamlines 12B2 and 44XU of SPring-8 (Japan), and in-house X-ray crystallography facilities. Data were indexed and scaled with the HKL2000 package<sup>29</sup>. The contents of asymmetric units were estimated from the Matthews coefficient<sup>30</sup>. The data suggest that a value of 2.29 Å<sup>3</sup> Da<sup>-1</sup> with 46.3 % solvent content corresponds to two molecules per asymmetric unit in the crystal.

### 2.4 Structure determination and refinement

The Multiple wavelength anomalous dispersion (MAD) method was used to obtain phase information, and CRANK<sup>31</sup> (in CCP4 package) was used to find the phase solution. The CRANK pipeline started with substructure detection and ended with model building, including the procedures of substructure detection by AFRO/CRUNCH2<sup>32</sup>, substructure refinement by BP3<sup>33</sup>, Hand determination and

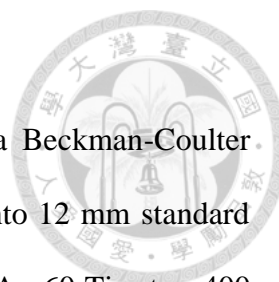
density modification by SOLOMON<sup>34</sup>, and model building by BUCCANEER<sup>35, 36</sup>. Phase extension yielded electron density maps into which a polypeptide model was built with the program COOT<sup>37, 38</sup>. The model was further refined with REFMAC<sup>39</sup>. Other native structures were solved by molecular replacement method using the Se-Met substituted MppJ as the search model. Figures were generated using PyMOL (<http://www.pymol.org>). Detailed refinement statistics are given in Table 1.

## 2.5 Enzymatic reaction

Typical assays for the enzymes (2.5 ng each) contain: PBS buffer (50 mM, pH 7.4), Ppy (3 mM), and SAM (1.5 mM) are mixed (150  $\mu$ l final assay volume) for 4.5 hr incubation at 25°C. After incubation, the reaction mixtures are centrifuged at 16000 g for 5 min and then spin in an ultracentrifugal filter unit fitted with a 5 kDa cut-off membrane (Millipore). The resulting filtrate is directly subjected to HPLC-ESI/MS analysis with a gradient solvent system of increasing acetonitrile in water (0.1% TFA) from 0 to 60% over 30 min. On-line LC-ESI/MS spectra are obtained from an Agilent HPLC 1200 series instrument coupled to a Thermo-Finnigan LTQ-XL ion trap mass spectrometer fitted with a positive mode ESI source.

## 2.6 Site-directed mutagenesis

Site-directed mutagenesis was carried out by using QuickChange (Stratagene), and the wild-type MppJ was used as the template for single mutation. For double mutation, the single or double mutants, respectively, served as templates. All mutations were confirmed by DNA sequencing. The pET/His plasmid was used for protein expression. Mutant enzymes were purified with the same protocol as recombinant wild-type MppJ.



## 2.7 Analytical ultracentrifugation analysis

The sedimentation velocity experiments were performed with a Beckman-Coulter XL-I analytical ultracentrifuge. Samples and buffers were loaded into 12 mm standard double-sector Epon charcoal-filled centerpieces and mounted in an An-60 Ti rotor. 400  $\mu\text{l}$  of 1  $\text{mg ml}^{-1}$  MppJ protein sample were loaded into the cell. Sedimentation velocity experiments were performed at rotor speed of 42,000 r.p.m. at 20°C. The signals of samples were monitored at 280 nm and collected every 3 min for 6 h. The raw data of experiments were calculated using SedFit software<sup>40</sup>. The density and viscosity of buffer were calculated using Sednterp software (<http://www.jphilo.mailway.com/default.htm>). All samples were visually checked for clarity after ultracentrifugation, and no indication of precipitation was observed.

## 2.8 Isothermal titration calorimetry analysis

ITC experiments were performed at 25°C using an iTC<sub>200</sub> microcalorimeter from MicroCal, Inc. (Northampton, MA). 50 mM HEPES buffer (pH 8.0) was used for ITC experiments. Titrations consisted of 15 injections of 2  $\mu\text{l}$  and were separated by 150 s. Each titration contained 0.1 mM MppJ in the sample cell (250  $\mu\text{l}$ ). 1 mM substrates (SAM, Ppy, and SAH) were loaded in the injection syringe (40  $\mu\text{l}$ ). The reference cell was filled with water throughout all experiments. The heat of dilution was determined by titrating the substrates into the buffer without enzyme and subtracted prior to curve fitting. Data analysis was performed by using the Origin 7.0 software provided by MicroCal.

## 2.9 X-ray absorption spectroscopy

X-ray absorption spectra were recorded at beamline 17C1 of National Synchrotron Radiation Research Center (Taiwan). 2 mM MppJ in buffer of 50 mM HEPES (pH 8.0)

was used for data collection. A cold-air device was used to lower down temperature to about 10°C for reducing photodamage. The absorption spectra were collected in fluorescence mode with solid-state detector or Lytle detector. The first inflection point at 7112 eV of Fe foil spectrum was used for energy calibration. Raw X-ray absorption data were analyzed following standard procedures, including background subtractions and normalization. Data were processed by using ATHENA software<sup>41</sup>.

#### 2.10 Electron paramagnetic resonance spectroscopy


EPR spectra were recorded at the X band (9.5 GHz) by using a Bruker X-band E500CW spectrometer (Bruker Biospin). 2 mM MppJ samples were loaded into EPR sample tube and immersed into a liquid nitrogen Cold Finger Dewar to maintain the temperature at 77K for data collection.

#### 2.11 Chiral HPLC analysis

A Phenomenex Chirex (D)-penicillamine phase 3126 (4.6 mm × 250 mm) ligand exchange column (Torrence, CA, USA) was used for resolving the four diastereomers of β-methylphenylalanine standards and the enzymatic products. The β-methylphenylalanine (βMePhe) standards (Sigma-Aldrich) and enzymatic products were prepared in 1:1 H<sub>2</sub>O/CH<sub>3</sub>CN, and their chiral HPLC analyses were conducted isocratically on the Chirex 3126 column using CuSO<sub>4</sub> (2 mM) in 85:15 H<sub>2</sub>O/CH<sub>3</sub>CN as the mobile solution.

### 3. Results

#### 3.1 Protein crystallization and structure determination



Recombinant MppJ, a 38-kDa protein consisting of 337 amino acids, was expressed in *E. coli* as an N-terminal His<sub>6</sub>-tagged protein and purified sequentially by Ni<sup>2+</sup> affinity chromatography/ion exchange chromatography (the latter is for metal ion determination) and gel filtration chromatography. MppJ was cocrystallized in the presence of both SAM and Ppy. Though the sequence similarity of MppJ is analogous to the MT protein family (Figure 2), native crystals cannot be solved by molecular replacement (MR). The selenium–methionine labeled protein–crystal diffracted to 2.0 Å resolution. X-ray diffraction data collected at the selenium (Se) peak wavelength were used to obtain initial phase information by multiple wavelength anomalous dispersion (MAD). The phase information from the MAD experiment with the Se peak data were used to build the initial model (CRANK)<sup>31</sup>, which consists of two molecules in the asymmetric unit (Figure 3a). The final model of a single MppJ consists of 332 residues, and it displays overall excellent stereochemistry judged by the Ramachandran plot, which shows that 99.5% of the residues are in the allowed region. Data collection and refinement statistics are summarized in Table 1. In the ternary complex, the electron density for the polypeptide chain and the bound ligands is well defined, while substrates SAM/Ppy turned out to be products SAH/MePpy (Figure 3b,c). From collated crystal diffraction data, the MppJ–SAM–Ppy ternary complex was found and solved by MR (Phaser)<sup>42</sup> to 2.4 Å resolution. Both ternary structures were refined to an R<sub>cryst</sub> and R<sub>free</sub> of 17.1% and 23.8% for MppJ–SAM–Ppy and 20.1% and 25.6% for MppJ–SAH–MePpy (Table 1).

### 3.2 Overview of the structure

The final model of MppJ is a densely packed homo-dimer in the asymmetric unit, which is in accord with the analysis of gel filtration and analytical ultracentrifugation (determined MW 70.2 kDa; calculated MW for the dimer, 74.4 kDa) (Figure 4). The majority of the SAM-dependent MTs structures are fundamentally similar, having core folds reminiscent of the Rossmann nucleotide binding fold<sup>13, 14</sup>. This core has been elaborated with additional structural elements evolved to accept assorted substrates. Topologically, the structures of MTs can be grouped into five classes (classes I–V)<sup>13, 14</sup>. These folds are highly variable at both structural and sequence levels, while they are consistent with the fact that this family of enzymes methylates a vast variety of substrates. MppJ features the class I fold, of which an individual subunit is made of two domains (Figure 5) with an extended N-terminal dimerization domain (DD, residues 1–160) and a C-terminal methyltransferase core domain (MT, residues 161–337). The DD domain has an antiparallel three-stranded  $\beta$ -sheet core ( $\beta$ 1– $\beta$ 3) cramped by forceps-like  $\alpha$ -helices ( $\alpha$ 1 and a fractured long helix  $\alpha$ 5– $\alpha$ 8). These two  $\alpha$ -helices extended away from the DD domain break into the DD domain of the other subunit, promoting dimerization. The contact area of dimer interface is 3135 Å<sup>2</sup>, mainly dominated by hydrophobic interactions along with 23 hydrogen bonds and 8 salt bridges. The DD also helps create a substrate-binding pocket at the occlusion between the DD and the MT domains in the same subunit. The MT domain is folded into an  $\alpha/\beta$  sandwich fold (or a Rossmann fold), comprising a seven-stranded  $\beta$ -sheet core ( $\beta$ 4– $\beta$ 10) sandwiched by three  $\alpha$ -helices on one face ( $\alpha$ 10– $\alpha$ 12) and another three on the other ( $\alpha$ 13,  $\alpha$ 14 and  $\alpha$ 16) (Figure 5). The central sheet is parallel with the exception of strand  $\beta$ 10 juxtaposed between  $\beta$ 8 and  $\beta$ 9. One short  $3_{10}$ -helical segment ( $\eta$ 3) between  $\beta$ 7 and  $\alpha$ 14 is what is typically an  $\alpha$ -helix in typical MT folds. The other two  $3_{10}$ -helical segments ( $\eta$ 1

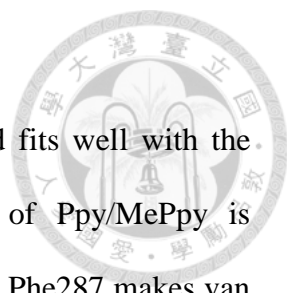
between  $\alpha 8$  and  $\alpha 9$ , and  $\eta 2$  between  $\beta 6$  and  $\alpha 13$ ) are additional elements not found in typical MT folds.



### 3.3 Conformational change and SAM/SAH-binding

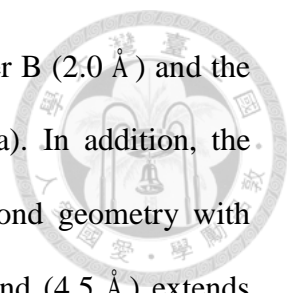
Superposition of binary (Ppy) and ternary (Ppy and SAM) structures (RMSD = 0.968 for 334 C $\alpha$  of MppJ) revealed that the MT domain undergoes a modest conformational change when SAM was uploaded. Upon binding of SAM,  $\beta 5$ ,  $\beta 6$  and  $\alpha 12$  together move toward  $\alpha 9$  resulting an overall 3.9 Å displacement (Figure 6a). Some conserved motifs lie in the core fold and comprise a glycine-rich sequence (motif I), which interacts with the methionine portion of SAM, and an acidic loop (motif II), which interacts with the ribose moiety<sup>13, 14</sup>. The former (DXGXGXG) is located between  $\beta 4$  and  $\alpha 11$ , while the latter is placed between  $\beta 5$  and  $\alpha 12$ . Asp164 may interact with the amino group of methionine (see Figure 2 for sequence alignment with representative homologs)<sup>13, 14</sup>. In both ternary complex structures, the cofactor SAM or its counterpart SAH is well defined by electron density (Figure 3), in which SAM/SAH lies across the top of the MT fold, with the ribose above the carboxyl-end of  $\beta 4$ , the methionine moiety extending to the area between  $\alpha 9$  and  $\alpha 11$ , and the adenine ring extending to  $\beta 6$  and  $\eta 2$ . The carboxylate group of SAM/SAH forms an electrostatic interaction with side chains of Arg172 and backbone nitrogen of Arg137. The terminal amino group participates in a hydrogen bonding network in direct interactions with backbone oxygens of Gly166 and Phe239. The adenine ring is cramped between the side chains of Ile190 and Met241, and Asp217 and Arg219 form hydrogen bonds to the N6 amino group (Figure 6b). The ribose moiety is anchored to MppJ through hydrogen bonds of the C2 and C3 hydroxyl groups to the backbone nitrogens of Gly168 and Ala191. Finally, the SAM/SAH sulfur makes a hydrophobic interaction with the side chain of Cys136.

### 3.4 Ppy-binding site and coordination chemistry



Adjacent to SAM/SAH, a chunk of electron density observed fits well with the substrate/product Ppy/MePpy (Figure 7a). The phenyl portion of Ppy/MePpy is surrounded by hydrophobic and van der Waals residues (Figure 7b): Phe287 makes van der Waals contacts with the phenol ring of Ppy/MePpy ( $\sim 3.1 \text{ \AA}$ ), and along with Ile139, Met240, Phe287, Phe291, Leu328 forms a predominantly hydrophobic pocket. Off side of the  $\alpha$ -ketoacid moiety of Ppy/MePpy, we observed a round bulk of additional electron density that best fits a mononuclear metal ion (Figure 7a). This agreed with the metal dependence on activity<sup>22</sup> as no product was formed in the presence of 5 mM EDTA. By means of UV/Vis, inductively coupled plasma mass spectrometry (ICP-MS), and X-ray absorption spectroscopy (XAS), the metal ion was determined to be a ferric ion (Figure 8a), the first case of its kind. In addition to binding Ppy/MePpy, the mononuclear ferric ion is in coordination with residues His243, Asp244, and His295, of which His243 and Asp244 are positioned on  $\eta 3$  immediately C-terminal of  $\beta 7$  and His295 is located at the C-terminal end of  $\alpha 15$  (Figure 7a). This 2-His-1-carboxylate facial triad localizes the active site metal ion through direct coordination, in common to all non-heme Fe(II) dependent oxygenases<sup>43, 44</sup>. The non-heme Fe(II)  $\alpha$ -ketoacid dependent oxygenases typically have a jellyroll fold motif (or a cupin fold)<sup>43,44</sup>, and can be categorized into three groups based on the spacing of these three ligands in the primary structure. HX(D/E)X<sub>50-210</sub>H, the most general spacing motif, is similar to HDX<sub>51</sub>H (HD juxtaposition) the motif in MppJ<sup>43, 45-48</sup>. Ppy/MePpy makes a distinctly asymmetric bidentate contact with the active site metal ion through one of carboxylate oxygens ( $2.1 \text{ \AA}$ ) and its adjacent  $\alpha$ -oxylate/keto oxygen ( $1.9 \text{ \AA} / 2.2 \text{ \AA}$ ), completing what is an almost ideal tetragonal bipyramidal arrangement (octahedral coordination) of ferric ion ligands: the  $\epsilon 2$  nitrogen of His243 ( $2.2 \text{ \AA}$ ), the  $\alpha$ -ketoacid bidentate of

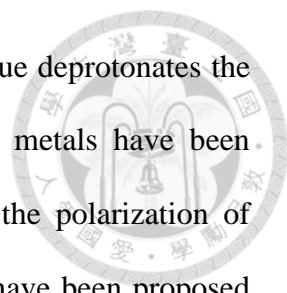




Ppy/MePpy and water A (2.0 Å) form the equatorial plane and water B (2.0 Å) and the  $\epsilon 2$  nitrogen of His295 (2.1 Å) occupy apical positions (Figure 9a). In addition, the benzylic carbon of Ppy (enol) forms almost an ideal hydrogen-bond geometry with Trp99 (3.5 Å) from the *si* face. The second potential hydrogen bond (4.5 Å) extends from Cys136 to the *re* face benzylic carbon. The third apparent hydrogen-bonding interaction is an electrostatic interaction between the carboxylic group of Ppy/MePpy and the guanidino group of Arg127 (2.7 Å) (Figure 9a). In analogy to non-heme Fe(II)  $\alpha$ -ketoacid dependent oxygenases, the coordination of Ppy with Fe(III) leads to the formation of thermodynamically stable ferric-Ppy complexes in rhombic octahedral symmetry with a high-spin ferric ion ( $S=5/2$ , isotropic  $g\sim 4.3$ ) (Figure 8b). Unlike non-heme Fe(II)  $\alpha$ -ketoacid dependent oxygenases (a five-coordinate species), Asp244 makes no direct contact with the metal ion. Two water molecules, sandwiched by the metal ion and the carboxylic group of Asp244 leading to a six-coordinate species, is likely irreplaceable (particularly with molecular oxygen as coordinative saturation), inasmuch as neither hydroxylated products (in the presence/absence of ascorbic acid) nor  $H_2O_2$  (peroxidase assay using an iron and xylenol orange reagent) nor radicals (2,2-diphenyl-1-picrylhydrazyl (DPPH) radical scavenging assay) as leaky byproducts were detected. The mechanistic role for these two water molecules in the MT reaction is likely to prevent occurrence of non-heme Fe(II)  $\alpha$ -ketoacid dependent oxygenase-like reactions.

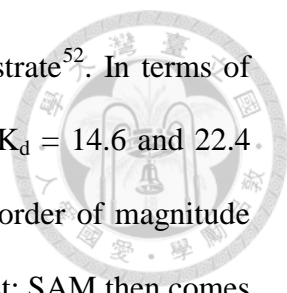
### 3.5 Reaction mechanism

In general, methyltransferase reactions involve direct transfer of the SAM methyl group to the substrate in an  $S_N2$ -like mechanism. This process requires that the OH/ $NH_2$  group of an acceptor be deprotonated by means of acid-base chemistry prior to attack of the SAM methyl group from the resulting nucleophile. In metal-independent OMTs



such as chalcone OMT, isoflavine OMT<sup>17</sup>, and RebM<sup>49</sup>, a His residue deprotonates the given hydroxyl in the substrates. In metal-dependent OMTs, the metals have been proposed to have roles in both the binding of the substrate and the polarization of hydroxyl to facilitate its deprotonation. Residues of different type have been proposed to serve as the general base in different MT families: The active-site structure of MycE supports a mechanism in which a conserved His acts as a general base, and the Mg<sup>2+</sup> ion helps position the methyl acceptor and stabilize a hydroxylate intermediate. The structures of rat catechol OMT (rOMT)<sup>50</sup> and plant flavonoid OMT (PFOMT)<sup>51</sup> share a common structural motif, whereby methyl transfer proceeds with one ionized hydroxyl group, facilitated by the Mg<sup>2+</sup> ion, and a conserved Lys as a catalytic base, to attack the electron-deficient methyl group of SAM<sup>52</sup>. NovP catalyzes the penultimate step in the biosynthesis of the aminocoumarin antibiotic novobiocin. Along with a proposed metal ion, a conserved Asp residue likely acts as the general base that initiates the reaction by deprotonating the 4-OH group of the noviose unit.

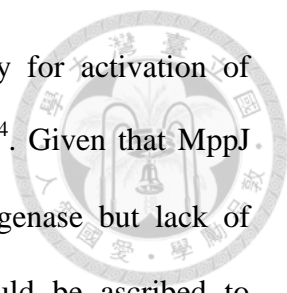
In the MppJ-SAM-Ppy complex, the benzylic carbon is anchored between the methyl group of SAM and residue Trp99, which is optimally positioned and assisted by both Ser104 and carboxylate of Ppy for proton abstraction. Together with the MppJ-SAH-MePpy ternary structure, we devised a reaction mechanism (Figure 9b): First, Ppy is anchored in place by binding its  $\alpha$ -keto acid to the ferric ion and Arg127. Trp99, the general base, removes a proton from the benzylic carbon of Ppy to convert conformation of Ppy from keto to enol (making the phenyl, enol and carboxyl moieties of Ppy coplanar), whereby the benzylic carbon is directed toward the methylium of SAM. With collapse of the  $\alpha$ -oxylate ion to the keto isomer, the benzylic-carbanion attacks the methylium of SAM to form MePpy (keto form) via S<sub>N</sub>2 nucleophilic substitution. Additionally, Cys136 is adjacent to the alkylated sulfur of SAM may



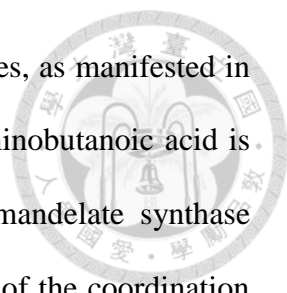
facilitate transfer of the methylum from SAM to the acceptor substrate<sup>52</sup>. In terms of binding affinity to MppJ, both Ppy and SAM are about the same ( $K_d = 14.6$  and  $22.4$   $\mu\text{M}$ , respectively), while SAH is higher than Ppy or SAM by one order of magnitude ( $K_d = 2.5$   $\mu\text{M}$ ; Figure 10). Taken together Ppy may bind to MppJ first; SAM then comes in and elicits a conformational change to better align the methyl transfer path. Once the reaction completed, MePpy comes off prior to SAH. Then MppJ retakes the initial state for next cycle of the MT reaction. To verify the proposed mechanism, we made an array of single-residue substitutions H243F, D244L, H295F, W99F, S104A, C136A, R127L, M240L, M241L (Table 2). Single mutants H243F, H295F, D244L and double mutation mutant H243F/H295F all lose the color and activity, suggesting proper ligand-metal coordination is an essential part of an effective reaction. The activity of the W99F and S104A mutants dropped dramatically (46% and 36%, respectively, Table 2), confirming Trp99 acts as the general base assisted by Ser104. Though fairly weak the W99F mutant is still active, suggesting the ferric ion has a role exerting the benzylic carbon a competent nucleophile in addition to orientating substrate and stabilizing the transition state. The relative activity of C136A mutant turns out to be 75% of wild-type, suggesting the Cys136 thiol group properly stabilizes the  $\text{CH}_3\text{-S}$  transition between SAM and Ppy. The assay for the mutant R127L indicated Arg127 helps recognize and orientate Ppy as the enzyme activity slightly dropped 6% when the electrostatic interaction was disrupted. Enzyme activities of M240L and M241L also dropped significantly (10%), suggesting their roles in SAM binding.

### 3.6 No non-heme oxygenase activity

In non-heme  $\alpha$ -ketoacid-dependent oxygenases, the 2-His-1-carboxylate facial triad motif provides a scaffold to bind the Fe(II) center while retaining flexibility to bind additional exogenous electron-rich ligands (e.g.  $\alpha$ -ketoacid) to the Fe(II)



coordination sphere (forming a five-coordinate species) necessary for activation of molecular oxygen (as the sixth coordinate) on the Fe(II) center<sup>53, 54</sup>. Given that MppJ possesses traits of a typical non-heme  $\alpha$ -ketoacid-dependent oxygenase but lack of intrinsic dioxygenase activities, we considered the inactivity could be ascribed to coordinative saturation. The fact is two water molecules are confined between the iron center and Asp244, where the distances for both iron-water (2.0 Å) and water-carboxylate (2.8 Å) are in the ranges of ‘short-strong hydrogen bond’ (SSHB) and hydrogen bond, respectively (Figure 9a)<sup>55, 56</sup>. It is likely the two water molecules are held as tight as irreplaceable allowing no O<sub>2</sub> to take the sixth coordinate. To examine this speculation mutants D244E and D244L were made to see their coordination influence. Interestingly, the original turquoise color becomes deepened for D244E, but the color of D244L turns to olive green, suggesting the sandwiched water is related to ligand coordination and charge transfer. The non-heme oxygenase assays (conversion of Ppy to mandelate/homogentisate) were performed, of which no products were found in the presence of either mutant. On the other hand, the relative yield of  $\beta$ -MePpy increased in the D244E assay, while the co-substrate SAM seemed to undergo fragmentation in the D244L assays (Figure 11b). D244E and D244L in complex with SAM and Ppy were then subjected to crystallization in a bid to explain the above observations. In the D244E structure, the distance of the water-carboxylate is apparently shortened to a ‘low-barrier hydrogen bond’ (LBHB) distance (Figure 12a) in correlation with the color enhancement, and the  $\alpha$ -keto acid remains bidentately chelated with the iron while having been methylated. In the D244L structure, the coordination chemistry is somehow changed to a five-coordination, wherein surprisingly dioxygen instead of water is side-on the iron center (Figure 12b). On the other hand, this is well in line with the color alteration. We reasoned fragmentation of



SAM is likely as result of hydrogen abstraction by a peroxide species, as manifested in the density map where the bond between the sulfur atom and 2-aminobutanoic acid is partially discontinued. In an effort to convert MppJ to hydroxymandelate synthase (HMS)<sup>46</sup>, five mutants were made by permutation and combination of the coordination triads to obtain the best coordination arrangement, most of which however lose the iron-binding ability and activities (Table 2). Additionally, residue Val300 in MppJ that geometrically may be more close to Glu in the 2-His-1-carboxylate facial triad of HMS was mutated to Glu (V300E). Still, there is no HMS activity for the double mutant D244A/V300E, while interestingly the methylation activity of the single mutant V300E is enhanced (Table 2). Provided that the iron center is commonly located at the hydrophobic center of the  $\beta$ -barrel in a typical non-heme dioxgenase, we reasoned in addition to the ‘coordinative saturation’ the no activity could also be ascribed to the iron center that is highly exposed to solvent, so that no sooner had oxo-Fe(IV) formed than it is annihilated by bulk solvent.

### 3.7 Stereochemistry

#### 3.7.1 Methylation

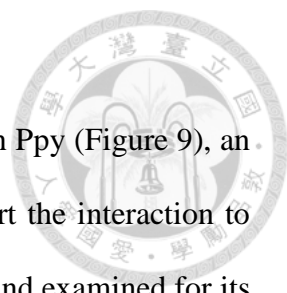
The third residue in mannopeptimycin is a diastereomer (2*S*,3*S*)- $\beta$ -methylphenylalanine ( $\beta$ MePhe)<sup>2</sup>. The methylated product in the solved structure turned out to be (3*R*)- $\beta$ MePpy (Figure 7). This puzzled us, as the MT was previously characterized to be ‘*S*’-configuration specific<sup>22</sup>. We therefore re-determined the chirality of the MT product  $\beta$ MePpy. The reactions added with MppJ and (2*S*)-aromatic amino acid-specific transaminase TyrB were performed, whereby  $\beta$ MePpy was converted to  $\beta$ MePhe and subsequently resolved by chiral HPLC. LC traces showed ratios of 5:1 and 1:1 for (2*S*,3*R*)- $\beta$ MePhe vs. (2*S*,3*S*)- $\beta$ MePhe at two reaction courses 4 and 8 hr, respectively (Figure 13). This analysis alongside the structural information enabled us

to redefine that the MT is '*R*'-configuration specific. We summarized that (3*R*)- $\beta$ MePpy was formed first, which underwent racemization in solution due to the relative acidity of the benzylic carbon. The racemic (3*R/S*)- $\beta$ MePpy were subsequently converted to (2*S,3R*)/(2*S,3S*)- $\beta$ MePhe by aromatic amino acid-specific transaminases. Given that the transamination is reversible, the higher level of (2*S,3R*)- $\beta$ MePhe would tilt to (2*S,3S*)- $\beta$ MePhe to establish a new equilibrium, and so does when the latter is drawn out from the system for the biosynthesis of mannopeptimycin (Figure 1).

### 3.7.2 Post-translational modification

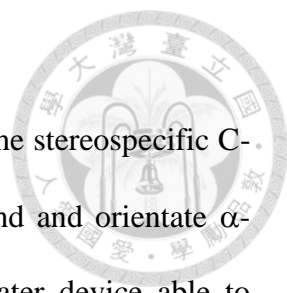
On the other hand, there is an extra electron density placed in the cleft formed by  $\alpha$ 16 and  $\beta$ 9 (Figure 14a,c). It is best fitted with Ppy, while it is covalently linked to Cys319. The linkage was determined to be a diastereomer (2*S,3R*)- $\beta$ -thioether- $\alpha$ -hydroxyl-phenylpropionic acid, which is likely formed through the thiolate of Cys319 addition to  $\beta$ -carbon of enol Ppy from the *Re* face. Interestingly, the Ppy adduct looks a projection of the active site (3*R*)- $\beta$ MePpy (by 22.3 Å translational displacement) symmetrical to the  $\beta$ -sheet core. Provided if this distant modification influences the MT activity, Cys319 was mutated to alanine and subjected to activity examination and protein crystallization. Strangely, C319A could not be crystallized suggesting Cys319 may play a structural role; the MT activity also slumped considerably (61% relative to WT) suggesting the residue may be related to an allosteric regulation. To further understand where the modification occurs in cells or in test tubes, MppJ in complex with 4-hydroxyphenylpyruvate, 4HPpy, was crystallized. The structure clearly showed 4HPpy appearing at both sites (Figure 14b,d), one chelating with the iron center and the other covalently linked to Cys319. Apparently, PTM took place in the test tube, suggesting the modification is substrate-dependent and the MT activity is positively regulated.

### 3.7.3 Methoxylation



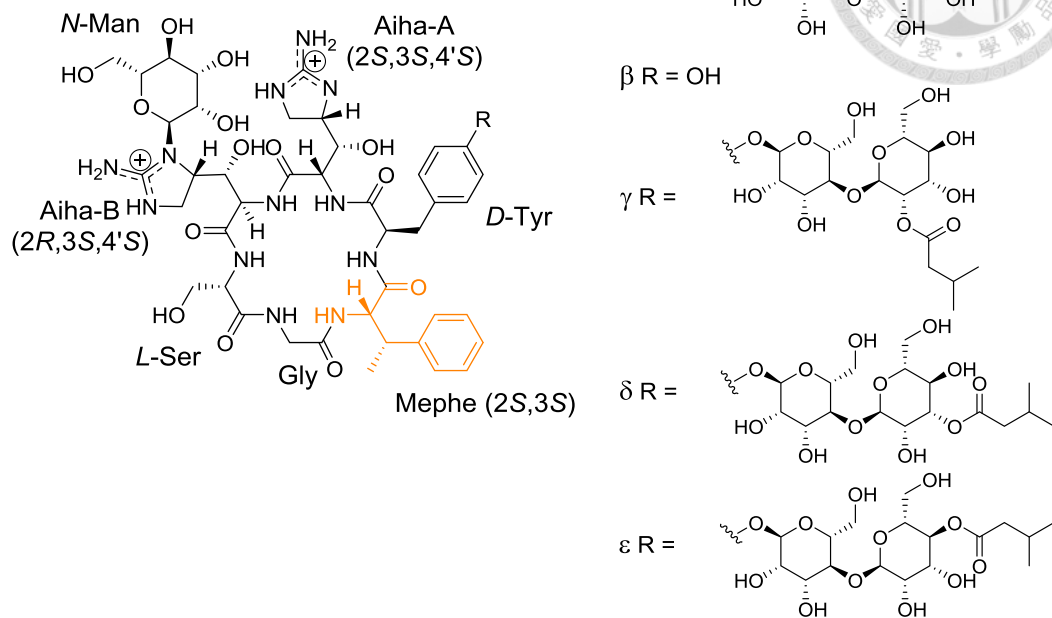
As mentioned above, Arg127 is electrostatically associated with Ppy (Figure 9), an interaction not seen in HMS. We considered the residue might exert the interaction to confine the enzyme reaction. Mutant R127L/D244E was first made and examined for its enzyme activities. The experiments showed the mutant retains the methylation activity (70%), indicating Arg127 has a modest effect on the Ppy binding. The active-site coordination in the structure of R127L/D244E is slightly dissimilar to that of D244E, where two water molecules are no longer sandwiched by D244E while  $\alpha$ -keto acid of Ppy remains chelation with the iron center (Figure 15a). To alter the bidentate binding mode, we re-introduced V300E into the R127L mutants. Two mutants R127L/V300E and R127L/D244A/V300E were made and subjected to enzyme examination. The results showed both mutants yield no  $\beta$ MePpy, while Ppy seemed consumed considerably. These two mutants were further crystallized to glean clues on this discrepancy (Figure 15b,c). To our surprise, a new chemical was identified in both structures, which was determined to be (2*R*,3*R*)- $\alpha$ -hydroxy- $\beta$ -methoxy-phenylpropionic acid (Figure 15d). In the structures, Ppy adopts a new conformation with merely the carboxyl group engaging in the coordination. We reasoned Trp99 likely serves as a general base abstracting the *pro-S* proton from the benzylic carbon of Ppy to form an enol conformer. A water molecule assisted by Cys136 then attacks the benzylic carbon from *Re*-face and Trp99 reprotonates the  $\alpha$ -carbon to form a hydrated adduct. Cys136 further deprotonates the  $\beta$ -hydroxyl group to attack the methyl group from SAM to form the unusual  $\alpha,\beta$ -vicinal diastereomer (Figure 15e).

#### 4. Discussion

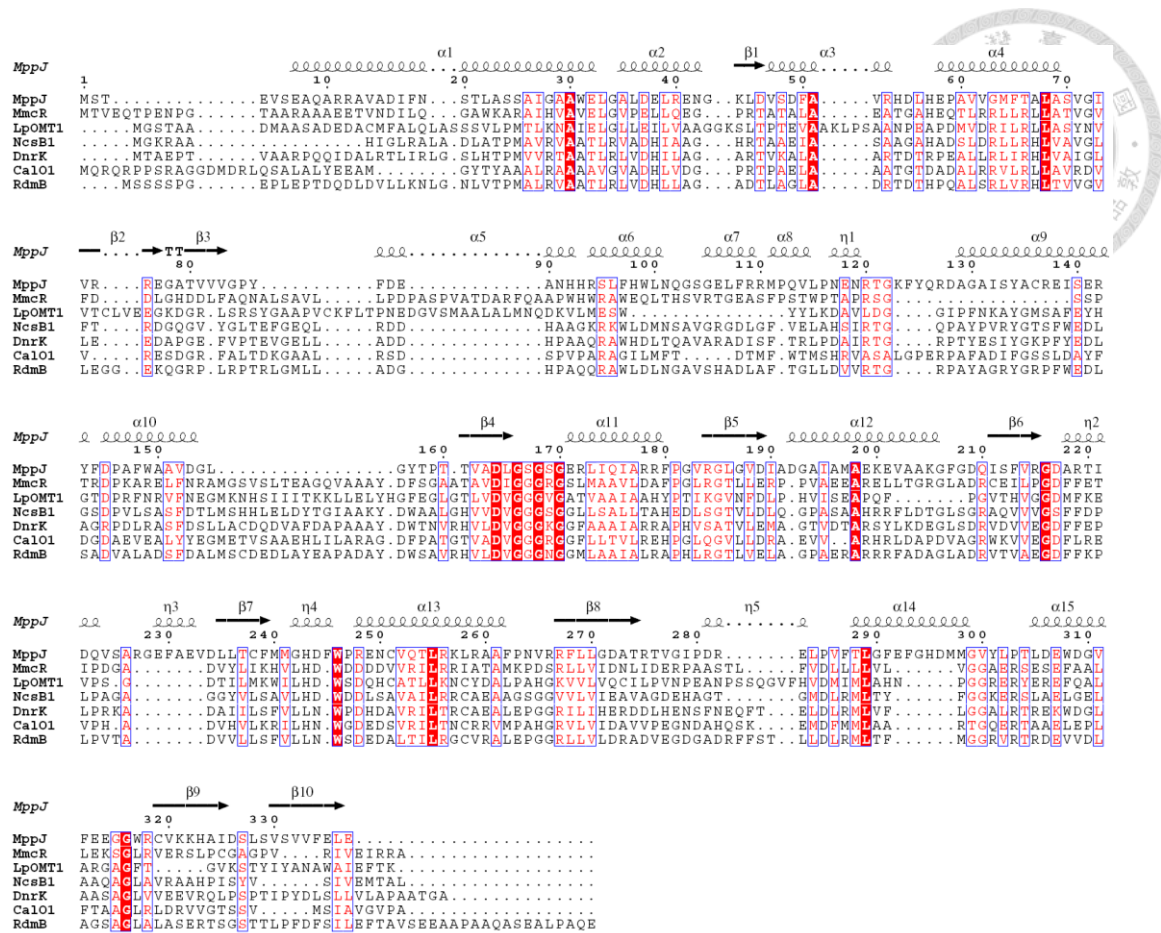


Having solved multiple complexed structures, we realized for the stereospecific C-MT reaction MppJ has evolved a non-heme iron center able to bind and orientate  $\alpha$ -ketoacid substrates and meanwhile developed a sandwiched bi-water device able to avoid formation of the reactive oxo-Fe(IV) species. In typical non-heme oxygenase, the iron center is commonly located at the hydrophobic center of the  $\beta$ -barrel. In opposition to non-heme oxygenase, the iron center of MppJ is solvent exposed. Maybe we can create a hydrophobic environment near the active site of MppJ by introducing a long loop between  $\eta 1$  and  $\alpha 9$ . The engineered MppJ with long loop “cap” structure could have non-heme oxygenase activity, converting Ppy to mandelate or homogentisate. In addition, MppJ mutants were engineered for the first time able to perform hydration and methylation reactions through altering coordination chemistry for stereospecific new compounds. Glycopeptides and their derivatives are still the most attracting candidates in the development of new types of anti-MRSA or anti-VRE compounds. Biosynthetic engineering techniques, for example by addition of modified genes into producing strains, may also generate more powerful or equivalent glycopeptides but in a more diversified and economical fashion.





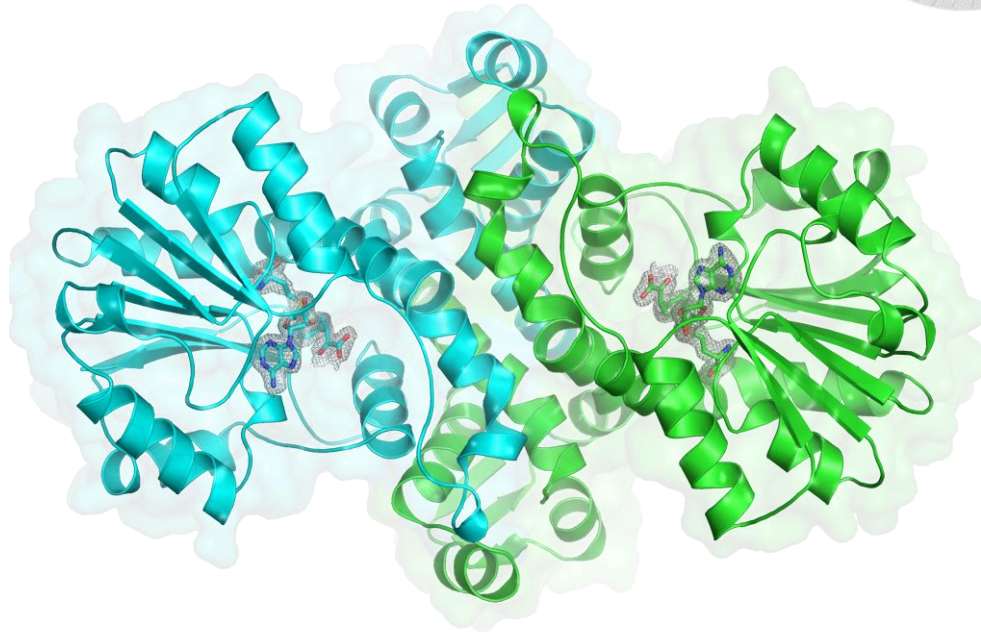
**Figure 1. Chemical structures of mannopeptimycins.**



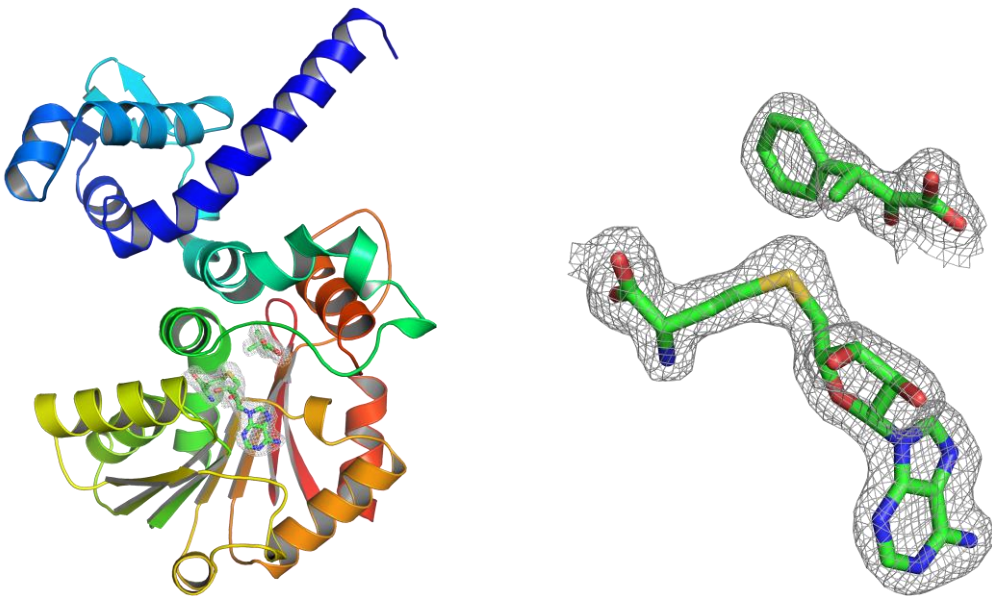
**Figure 2. Multiple sequence alignments for MppJ and homologues.**

Sequence alignments were performed by using ClustalW software (<http://www.ebi.ac.uk/Tools/msa/clustalw2/>). The plot was generated by ESPript (<http://esprict.ibcp.fr/ESPript/ESPript>). Secondary structure of MppJ was shown on the top of the sequence. Accession codes for MppJ (Q643C8), MmcR (Q9X5T6), LpOMT1 (Q9ZTU2), NcsB1 (Q84HC8), DnrK (Q06528), CalO1 (Q8KNE5), and RdmB (Q54527) are designated by the UniProtKB databank (<http://www.uniprot.org/>).

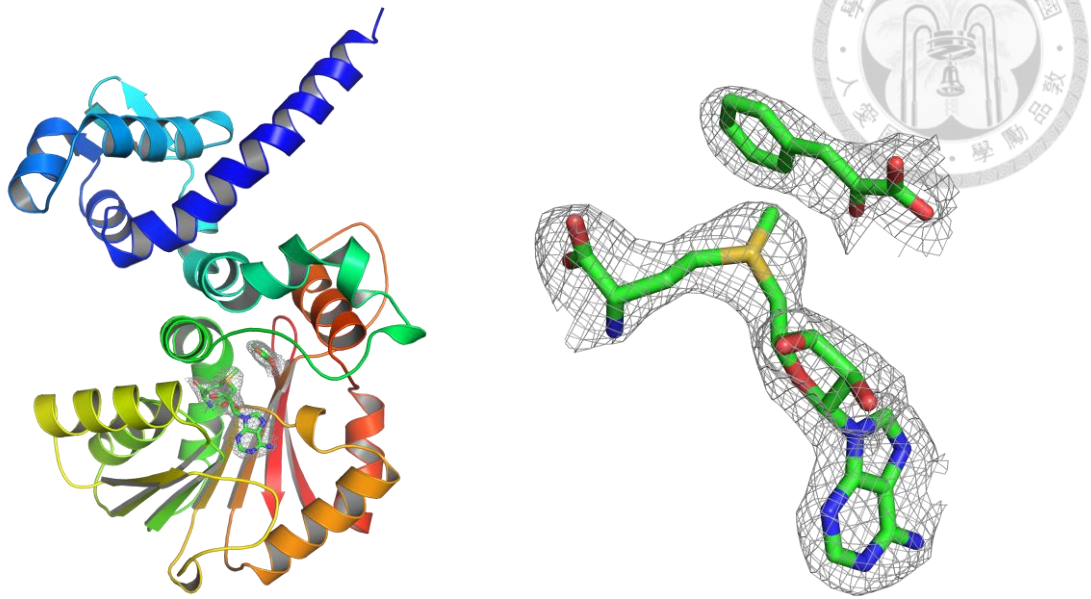
a



b

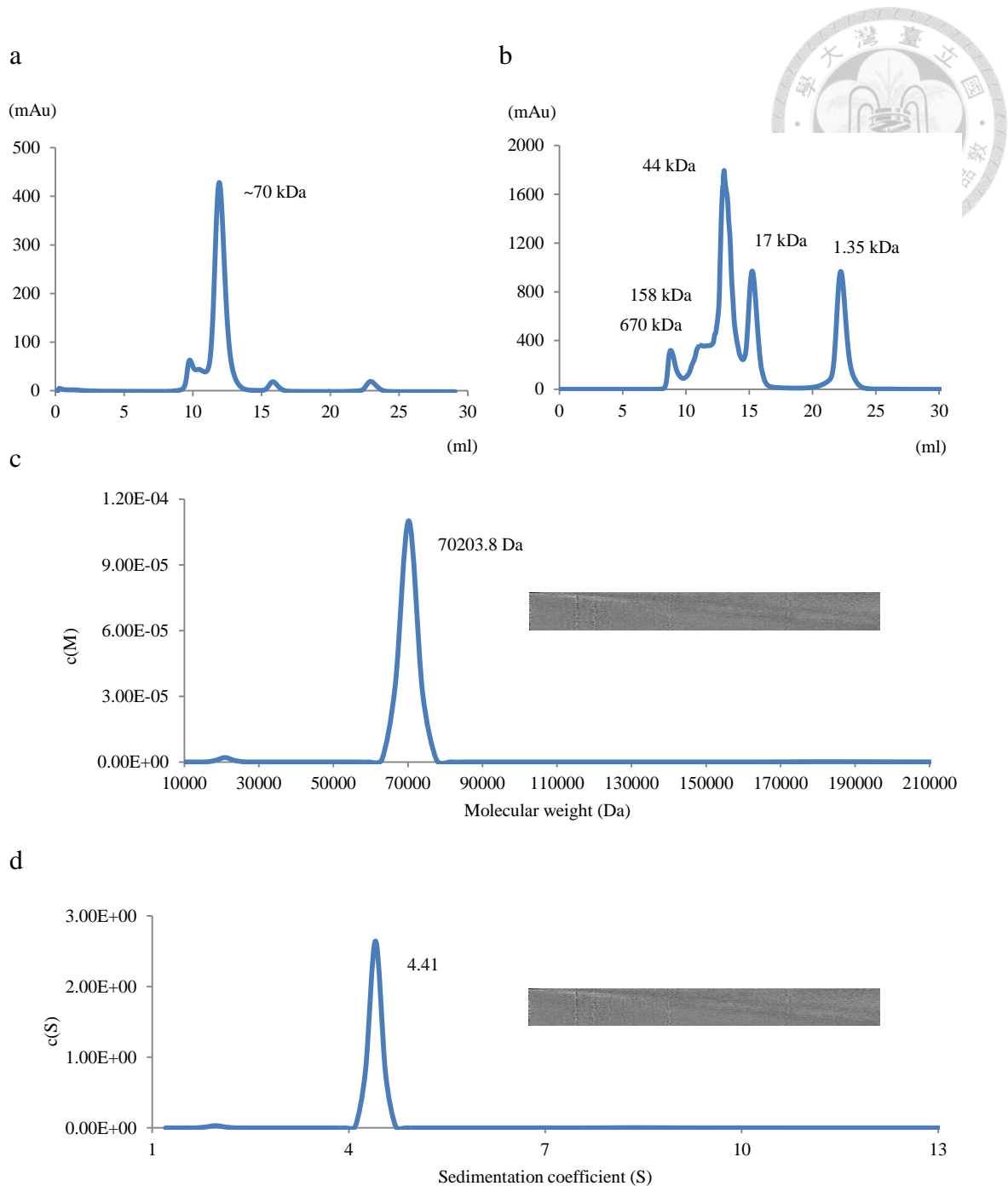


c



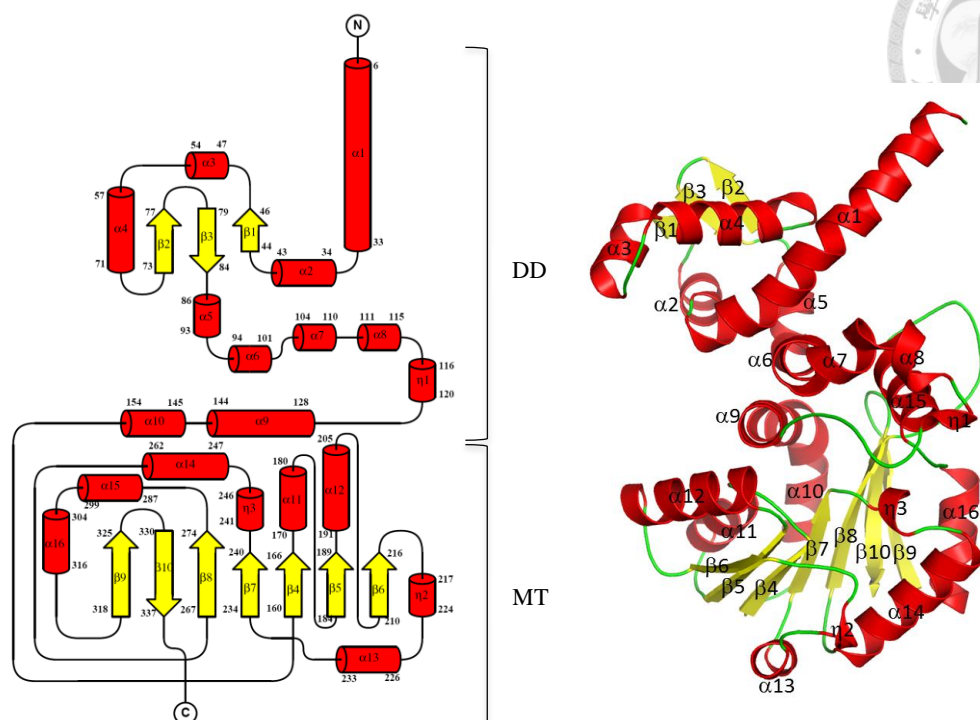
**Figure 3. Crystal structures of MppJ complexed with substrates or products.**

(a) Overall structure of MppJ complexed with products SAH/MePpy. The structure of MppJ is a homo-dimer in the asymmetric unit. The bound ligands are well defined as shown in the electron density map. (b) Structure of MppJ-SAH-MePpy complex.  $2F_o-F_c$  electron density map of products SAH/MePpy contoured at  $1 \sigma$ . (c) Structure of MppJ-SAM-Ppy complex.  $2F_o-F_c$  electron density map of substrates SAM/Ppy contoured at  $1 \sigma$ .



**Figure 4. Gel filtration and analytical ultracentrifugation analyses of MppJ.**

(a) Size exclusion chromatography analysis of MppJ by using Superdex 75 10/300 column. (b) Size exclusion chromatography analysis of molecular weight standards in the same condition. The AUC data were analyzed by SedFit (<http://www.analyticalultracentrifugation.com/default.htm>). The calculated  $c(M)$  distribution (c) and  $c(S)$  distribution (d) as shown in the panels indicated MppJ forms a dimer. The insert grayscale bars indicate the residuals bitmap of each fit.

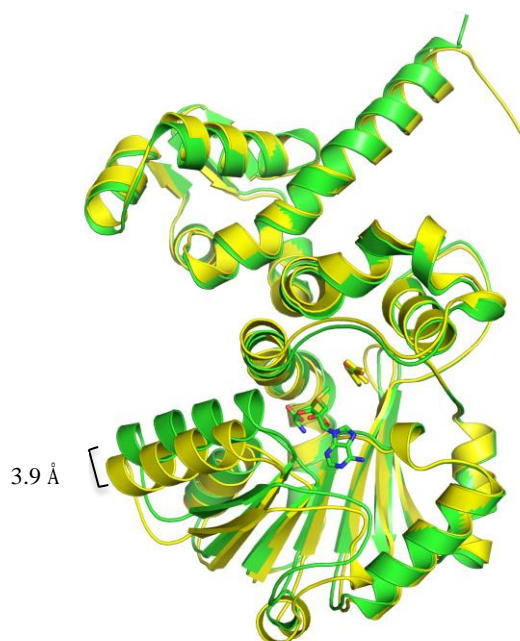


**Figure 5. Schematic topology of MppJ.**

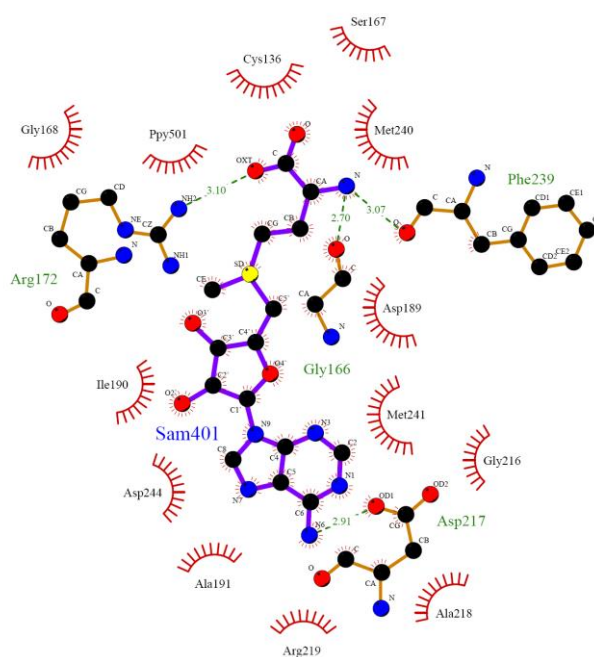
Topology diagram was generated by TopDraw program. The secondary structure of  $\alpha$ -helix and  $\beta$ -sheet are shown in red and yellow, respectively. A subunit of MppJ is composed of two domains, which are N-terminal DD domain (residues 1-160) for protein dimerization and C-terminal MT catalytic domain (residues 161-337) for methyltransferation.



a

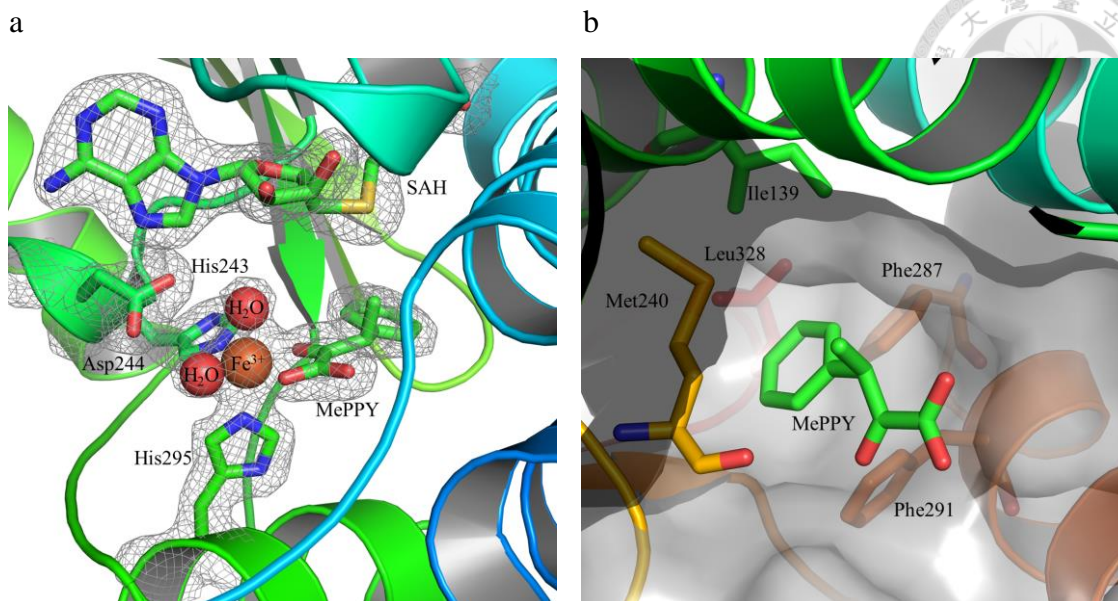


b



**Figure 6. Conformational change and SAM/SAH-binding.**

(a) Superposition of binary (Ppy) and ternary (Ppy and SAM) structures revealed a protein conformational change (RMSD=0.968 for 344 C $\alpha$  of MppJ). The structures of binary and ternary structures are shown in yellow and green, respectively. (b) Interactions between SAM and residues in MppJ structure. The figure was generated by LigPlot.

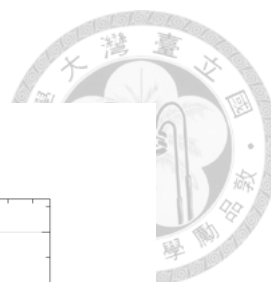


**Figure 7. Ppy binding site and metal ( $\text{Fe}^{3+}$ ) coordination in MppJ structure.**

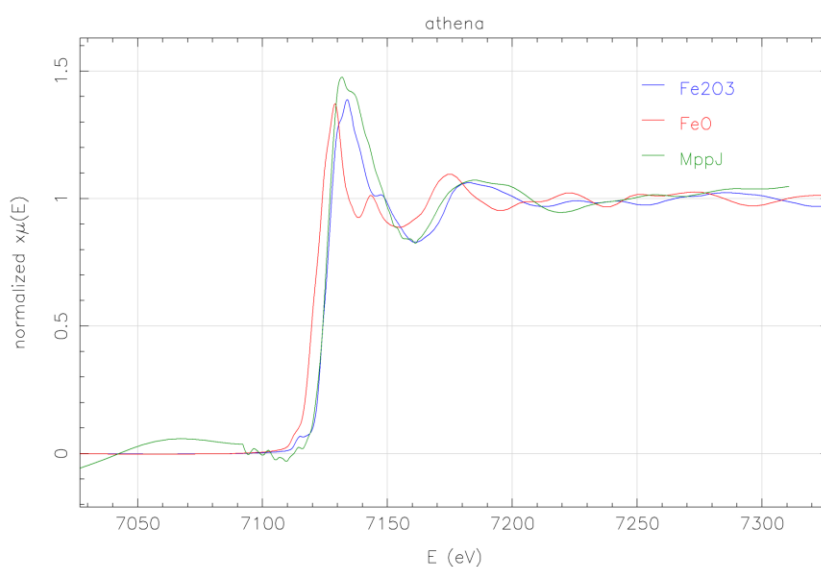
(a) Ppy makes a distinctly asymmetric bidentate contact with ferric ion ( $\text{Fe}^{3+}$ ) through one of carboxylate oxygens and its adjacent  $\alpha$ -oxylate/keto oxygen. The  $\epsilon 2$  nitrogen of His243 and His295, the  $\alpha$ -ketoacid bidentate of Ppy, two water molecules sandwiched by the metal ion and the carboxylic group of Asp244 leading to a six-coordinate species.

(b) The phenyl portion of Ppy is located in a hydrophobic cavity formed by residues Phe287, Ile139, Met240, Phe287, Phe291, and Leu328.

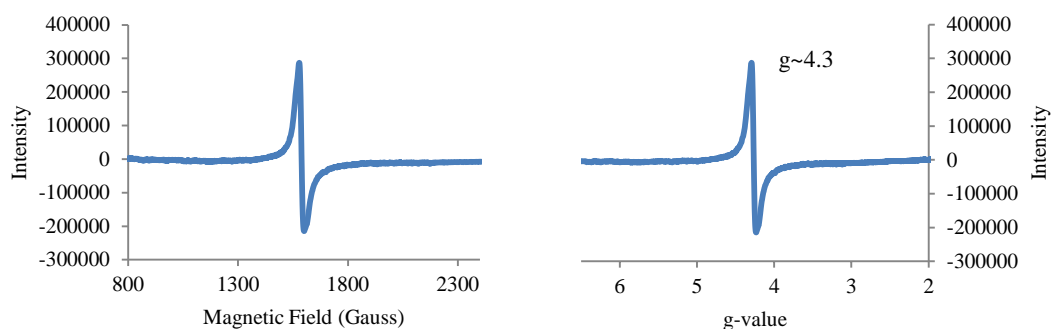




a

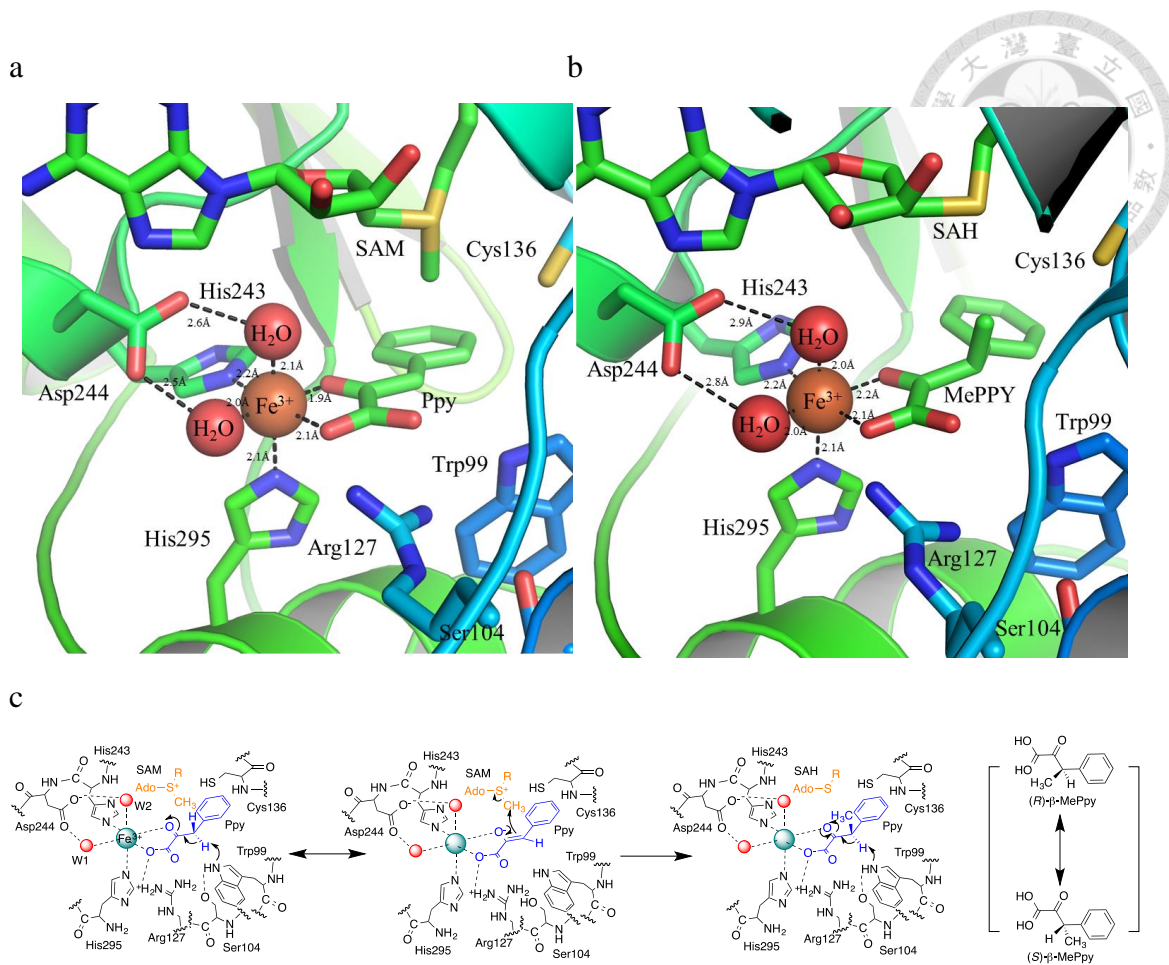


b



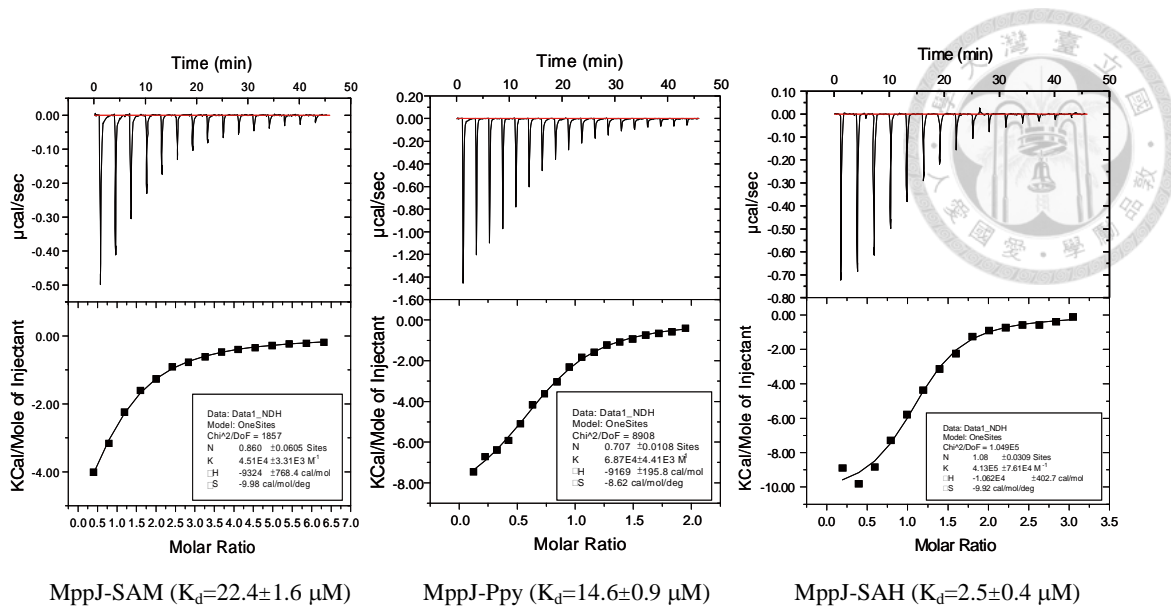
**Figure 8. Metal ion determination by X-ray absorption spectroscopy (XAS) and electron paramagnetic resonance (EPR).**

(a) The XAS spectrum of MppJ is shown in green. The spectra of standard Fe<sup>3+</sup> and Fe<sup>2+</sup> are colored in blue and red, respectively. (b) EPR analyses of MppJ. The metal is determined to be a high-spin ferric ion ( $S=5/2$ ,  $g\sim 4.3$ ).



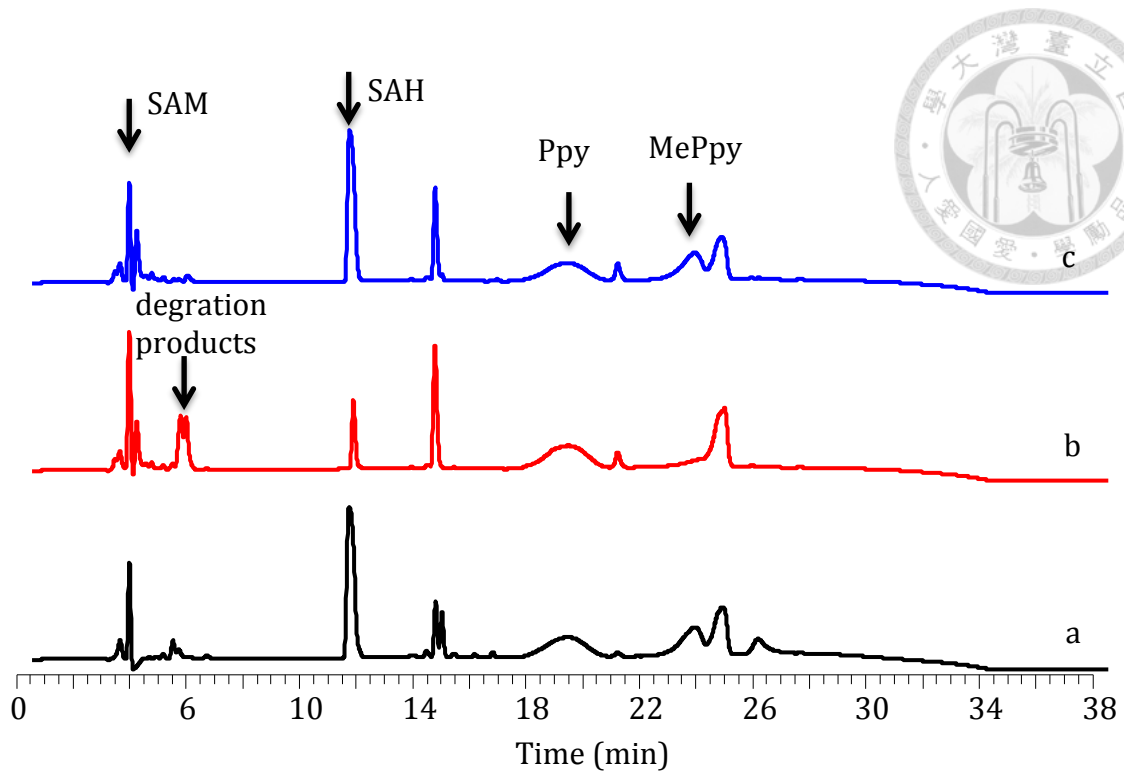
**Figure 9. Active site structure of MppJ and proposed catalytic mechanism.**

(a) Active site arrangement of MppJ/SAM/Ppy complex. (b) Active site arrangement of MppJ/SAH/MePpy complex. The benzylic carbon of Ppy/MePpy forms hydrogen bonds with Trp99 and Cys136 from the *si* face and *re* face, respectively. The carboxylic group of Ppy/MePpy interacts with the guanidino group of Arg127. (c) The proposed methyltransferation mechanism of MppJ.



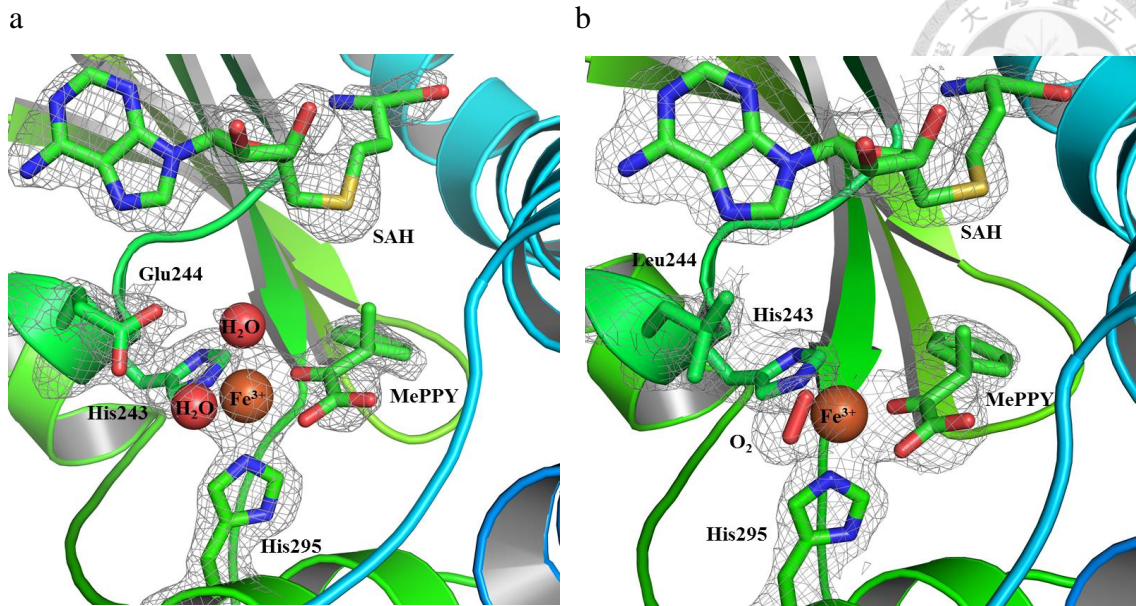
**Figure 10. Isothermal titration calorimetry (ITC) analyses of MppJ.**

ITC thermogram for MppJ binds with substrates (SAM and Ppy) or product (SAH). Each exothermic heat pulse corresponds to injection of 2 µl of ligands (1 mM) into the protein solution (0.1 mM); integrated heat areas constituted a differential binding curve that was fit to a standard single-site binding model (Origin 7.0, MicroCal iTC<sub>200</sub>).



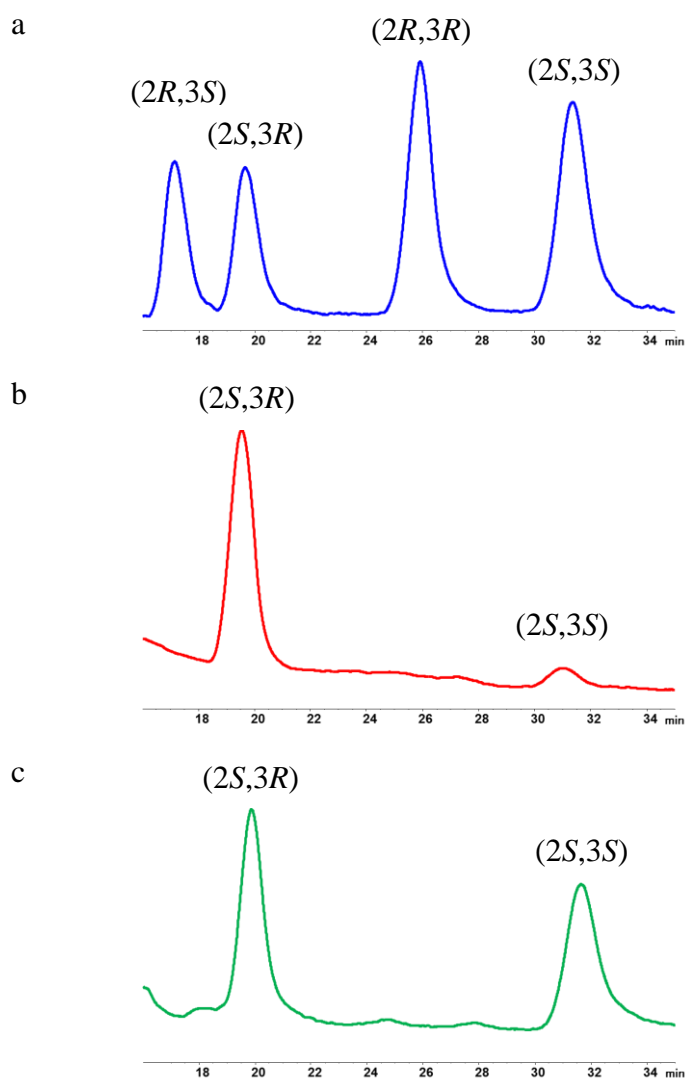
**Figure 11. HPLC traces of enzymatic reactions for MppJ WT, D244L, and D244E mutants.**

LC traces of enzymatic reactions for (a) WT MppJ, (b) D244L mutant, and (c) D244E mutant.



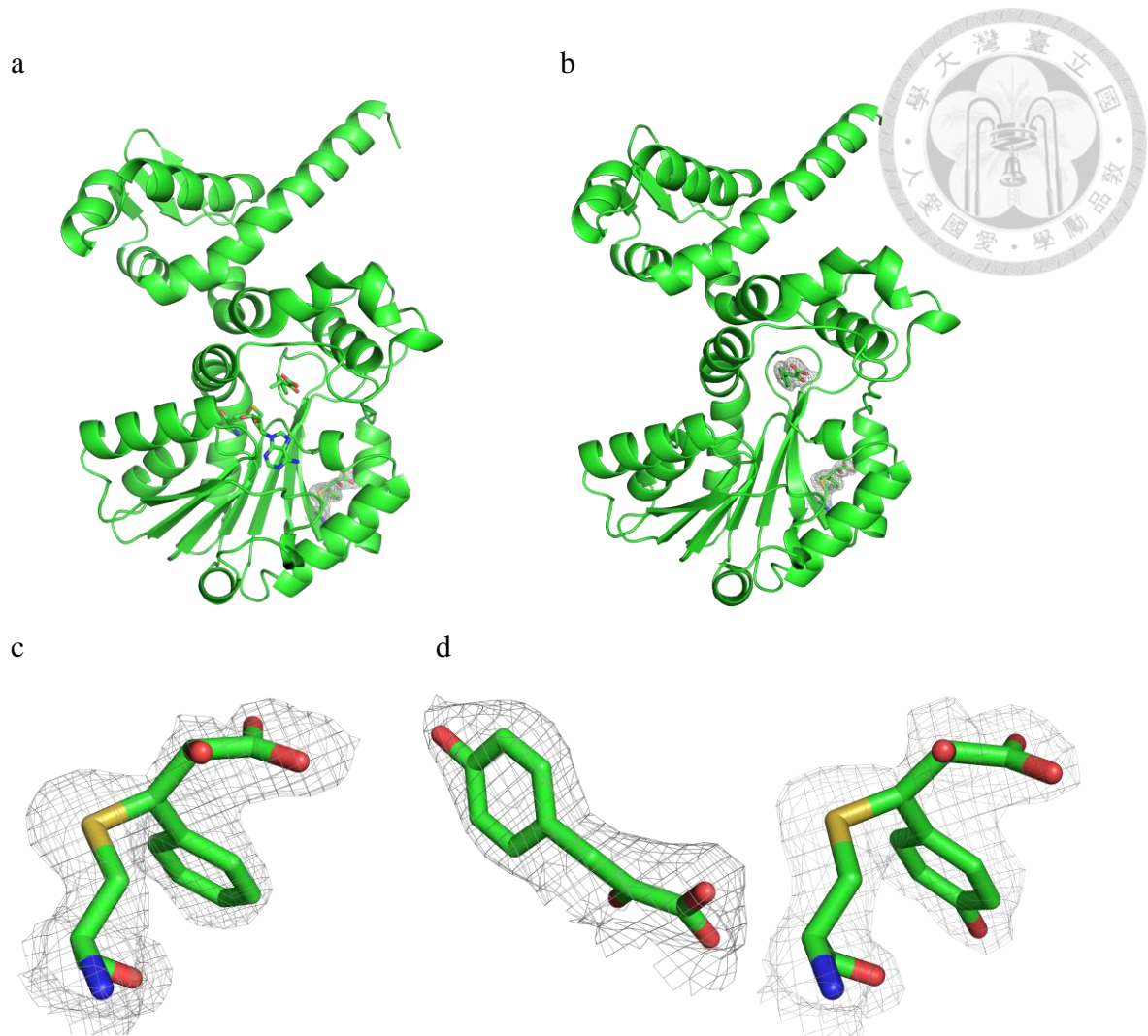
**Figure 12. Active site structures of MppJ D244E and D244L mutants.**

(a) Active site arrangement of MppJ D244E mutant. (b) Active site arrangement of MppJ D244L mutant, where dioxygen is side-on the iron center.



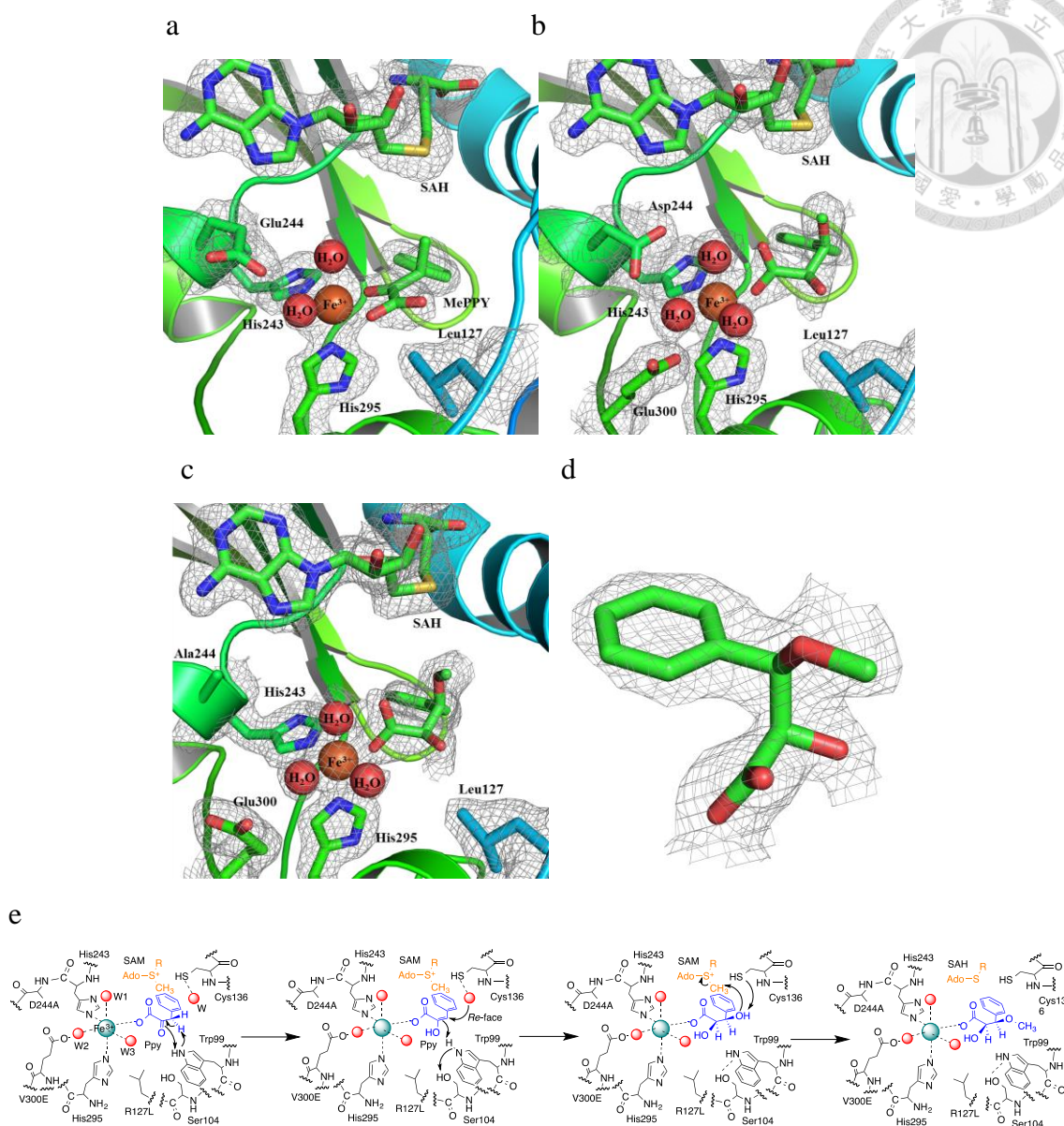
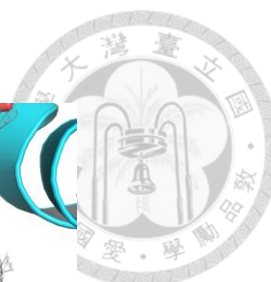
**Figure 13. Chiral HPLC analyses of MppJ.**

(a) Four enantiomers of  $\beta$ -methylphenylalanine ( $\beta$ MePhe) resolved by chiral HPLC. (b) Enzymatic products of MppJ and TyrB (aromatic amino acid-specific transaminase) at reaction courses 4 hr. (c) Enzymatic products of MppJ and TyrB at reaction courses 8 hr.



**Figure 14. Ppy/4HPpy are covalently linked to Cys319.**

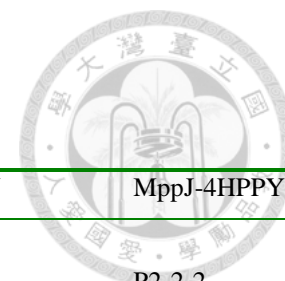
(a) Structure of MppJ-SAH-MePpy complex. (b) Structure of MppJ-4HPpy complex. (c)  $2F_o-F_c$  electron density map of covalently linked Ppy contoured at  $1\sigma$ . (d)  $2F_o-F_c$  electron density maps of 4HPpy in the active site and covalently linked 4HPpy contoured at  $1\sigma$ .



**Figure 15. Active site structures of MppJ mutants and proposed mechanism for new compound.**

(a) Active site structure of MppJ R127L/D244E mutant. (b) Active site structure of MppJ R127L/V300E mutant. (c) Active site structure of MppJ R127L/D244A/V300E mutant. (d)  $2F_o - F_c$  electron density map of (2R,3R)- $\alpha$ -hydroxy- $\beta$ -methoxy-phenylpropionic acid contoured at  $1 \sigma$ . (e) The proposed hydration-methylation mechanism.





**Table 1. Data collection, phasing and refinement statistics for MppJ structures.**

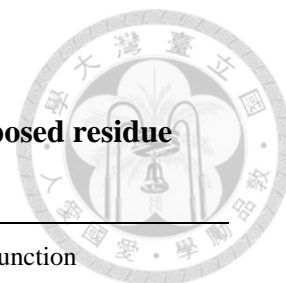
	SeMppJ-SAH-MePPY			MppJ-SAM-PPY	MppJ-PPY	MppJ-4HPPY
<b>Data collection</b>						
Space group	P2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>			P2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>	P2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>	P2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>
Cell dimensions						
<i>a</i> , <i>b</i> , <i>c</i> (Å)	64.3, 78.0, 138.2			57.5, 93.5, 142.2	57.4, 88.7, 137.0	59.5, 81.1, 137.4
$\alpha$ , $\beta$ , $\gamma$ (°)	90.0, 90.0, 90.0			90.0, 90.0, 90.0	90.0, 90.0, 90.0	90.0, 90.0, 90.0
	<i>Peak</i>	<i>Inflection</i>	<i>Remote</i>			
Wavelength	0.97874	0.97893	0.96353	1.54	0.90	0.97
Resolution (Å)	30.0-2.0(2.07-2.0)	30.0-2.0(2.07-2.0)	30.0-2.0(2.07-2.0)	19.0-2.45(2.54-2.45)	30.0-2.50(2.59-2.50)	30.0-2.40(2.49-2.40)
<i>R</i> <sub>sym</sub> or <i>R</i> <sub>merge</sub>	11.0(72.3)	10.5(72.2)	11.0(75.1)	8.5(61.4)	12.2(62.0)	13.7(58.9)
<i>I</i> / $\sigma$ <i>I</i>	23.3(3.25)	16.5(2.24)	16.3(2.29)	15.7(2.06)	13.3(2.67)	12.5(2.94)
Completeness (%)	99.9(100.0)	99.9(99.8)	99.9(99.8)	97.1(90.4)	99.1(98.6)	93.5(96.0)
Redundancy	10.3(10.2)	5.2(5.1)	5.1(5.1)	4.2(3.9)	4.9(5.0)	5.0(5.1)
<b>Refinement</b>						
Resolution (Å)	2.00			2.45	2.50	2.40
No. reflections	45227			27511	23457	23694
<i>R</i> <sub>work</sub> / <i>R</i> <sub>free</sub>	0.201/0.256			0.171/0.238	0.174/0.248	0.172/0.248
No. atoms						
Protein	5200			5186	5246	5170
Ligand/ion	102/7			102/16	60/6	52/11
Water	625			299	347	358
B-factors						
Protein	29.2			42.9	36.6	33.0
Ligand/ion	27.7/34.0			37.3/44.8	36.6/43.4	31.3/42.7
Water	39.1			39.3	33.7	33.8
R.m.s deviations						
Bond lengths (Å)	0.009			0.013	0.012	0.012
Bond angles (°)	1.400			1.649	1.569	1.542

\*Highest resolution shell is shown in parenthesis.



	MppJ D244E	MppJ D244L	MppJ R127L/D244E	MppJ R127L/V300E	MppJ R127L/D244A/V300E
<b>Data collection</b>					
Space group	P2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>	P2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>	P2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>	P2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>	P2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>
Cell dimensions					
<i>a</i> , <i>b</i> , <i>c</i> (Å)	57.0, 90.2, 136.5	57.2, 89.6, 136.6	60.8, 93.7, 137.4	60.9, 97.1, 135.2	61.0, 96.8, 135.4
$\alpha$ , $\beta$ , $\gamma$ (°)	90.0, 90.0, 90.0	90.0, 90.0, 90.0	90.0, 90.0, 90.0	90.0, 90.0, 90.0	90.0, 90.0, 90.0
Wavelength	0.97	0.90	1.00	1.00	1.00
Resolution (Å)	30.0-2.30(2.38-2.30)	30.0-2.50(2.59-2.50)	30.0-2.20(2.28-2.20)	30.0-2.10(2.18-2.10)	30.0-2.00(2.07-2.00)
<i>R</i> <sub>sym</sub> or <i>R</i> <sub>merge</sub>	9.9(59.5)	9.5(66.9)	6.9(54.6)	8.7(52.1)	5.7(47.2)
<i>I</i> / $\sigma$ <i>I</i>	16.6(3.0)	19.1(3.2)	25.0(3.3)	19.4(3.7)	26.6(3.5)
Completeness (%)	100.0(100.0)	99.9(100.0)	99.9(100.0)	99.9(100.0)	99.9(100.0)
Redundancy	4.9(4.9)	6.1(6.2)	6.1(6.2)	6.2(6.2)	5.8(5.3)
<b>Refinement</b>					
Resolution (Å)	2.30	2.50	2.20	2.10	2.00
No. reflections	30264	24097	34611	43121	48442
<i>R</i> <sub>work</sub> / <i>R</i> <sub>free</sub>	0.165/0.227	0.182/0.255	0.162/0.229	0.147/0.196	0.155/0.211
No. atoms					
Protein	5232	5246	5173	5182	5169
Ligand/ion	102/7	102/6	102/17	104/10	104/16
Water	427	331	529	759	788
B-factors					
Protein	34.9	52.4	31.2	26.3	29.4
Ligand/ion	43.6/58.8	56.1/70.7	32.8/42.8	36.6/37.9	34.0/40.3
Water	40.5	47.5	35.8	39.6	41.8
R.m.s deviations					
Bond lengths (Å)	0.010	0.009	0.016	0.019	0.020
Bond angles (°)	1.415	1.380	1.775	1.946	2.032

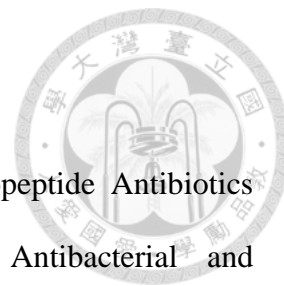
\*Highest resolution shell is shown in parenthesis.



**Table 2. Relative enzymatic activities (MT), enzyme colors and proposed residue functions.**

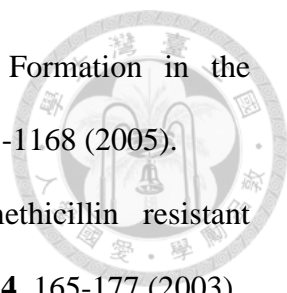
Mutants	Relative activity <sup>a</sup>	Color	Function
M240L	0.88	green	SAM/SAH binding
M241L	0.91	green	SAM/SAH binding
M240L/M241L	0.70	green	SAM/SAH binding
W99F	0.46	colorless	Active site base
S104A	0.36	colorless	Active site base
C136A	0.75	green	SAM/SAH stabilization
C136H	0.01	green	SAM/SAH stabilization
C319A	0.61	green	Allosteric regulation
R127L	0.94	yellow	Ppy binding
R127L/D244E	0.70	lime	Ppy binding/water interaction
R127L/D244L	0.00	lime	Ppy binding/water interaction
R127L/V300E	0.00	lime	Hydrotase and methyltransferase
R127L/D244A/V300E	0.00	lime	Hydrotase and methyltransferase
H243F	0.00	colorless	Metal coordination
H295F	0.01	colorless	Metal coordination
D244L	0.02	olive green	Metal coordination
D244E	1.20	green	Metal coordination
V300E	1.37	lime	Metal coordination
D244A/V300E	0.26	lime	Metal coordination
H243D/D244H	0.00	colorless	Metal coordination
H243E/D244H	0.01	colorless	Metal coordination
D244H/H295D	0.00	colorless	Metal coordination
D244H/H295E	0.02	green	Metal coordination
R127L/C136F/D244A/V300E	0.00	lime	Ppy binding/water activation/metal coordination/metal coordination
R127L/D244A/F291L/V300E	0.01	lime	Ppy binding/metal coordination/Ppy binding/metal coordination
R127L/C136F/D244A/F291L/V300E	0.00	lime	Ppy binding/water activation/metal coordination/Ppy binding/metal coordination

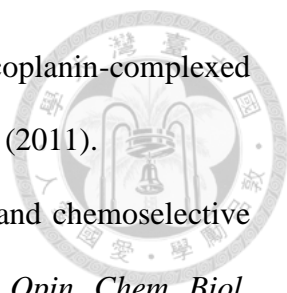
a. The reactions of mutants were analyzed by HPLC; peak areas in triplicate were integrated and averaged; the reaction rates were calculated using linear regression equation; relative activities were determined by dividing individual reaction rate with that of WT; the relative activity of WT is 1.0.

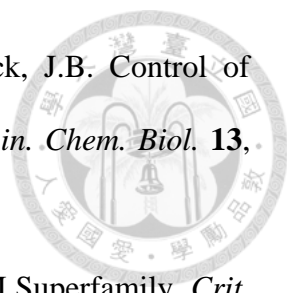


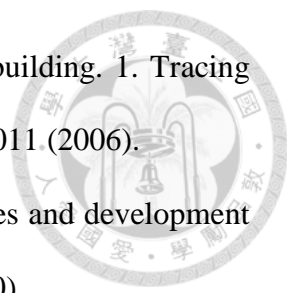
## References

1. Singh, M.P. *et al.* Mannopectimycins, New Cyclic Glycopeptide Antibiotics Produced by *Streptomyces hygroscopicus* LL-AC98: Antibacterial and Mechanistic Activities. *Antimicrob. Agents Chemother.* **47**, 62-69 (2003).
2. He, H. *et al.* Mannopectimycins, Novel Antibacterial Glycopeptides from *Streptomyces hygroscopicus*, LL-AC98. *J. Am. Chem. Soc.* **124**, 9729-9736 (2002).
3. Breukink, E. & de Kruijff, B. Lipid II as a target for antibiotics. *Nat Rev Drug Discov* **5**, 321-323 (2006).
4. Koehn, F.E. New Strategies and Methods in the Discovery of Natural Product Anti-Infective Agents: The Mannopectimycins. *J. Med. Chem.* **51**, 2613-2617 (2008).
5. Malabarba, A. & Goldstein, B.P. Origin, structure, and activity in vitro and in vivo of dalbavancin. *J. Antimicrob. Chemother.* **55** Suppl 2, ii15-20 (2005).
6. Van Bambeke, F. Glycopeptides in clinical development: pharmacological profile and clinical perspectives. *Curr Opin Pharmacol* **4**, 471-478 (2004).
7. Ruzin, A. *et al.* Mechanism of Action of the Mannopectimycins, a Novel Class of Glycopeptide Antibiotics Active against Vancomycin-Resistant Gram-Positive Bacteria. *Antimicrob. Agents Chemother.* **48**, 728-738 (2004).
8. Petersen, P.J., Wang, T.Z., Dushin, R.G. & Bradford, P.A. Comparative In Vitro Activities of AC98-6446, a Novel Semisynthetic Glycopeptide Derivative of the Natural Product Mannopectimycin  $\alpha$ , and Other Antimicrobial Agents against Gram-Positive Clinical Isolates. *Antimicrob. Agents Chemother.* **48**, 739-746 (2004).

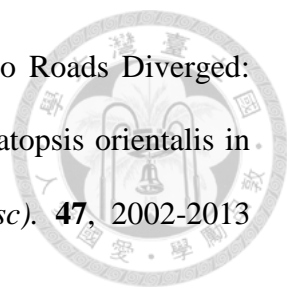
- 
9. Haltli, B. *et al.* Investigating  $\beta$ -Hydroxyenduracididine Formation in the Biosynthesis of the Mannopeptimycins. *Chem. Biol.* **12**, 1163-1168 (2005).
  10. Khare, M. & Keady, D. Antimicrobial therapy of methicillin resistant *Staphylococcus aureus* infection. *Expert Opin Pharmacother* **4**, 165-177 (2003).
  11. Magarvey, N.A., Haltli, B., He, M., Greenstein, M. & Hucul, J.A. Biosynthetic Pathway for Mannopeptimycins, Lipoglycopeptide Antibiotics Active against Drug-Resistant Gram-Positive Pathogens. *Antimicrob. Agents Chemother.* **50**, 2167-2177 (2006).
  12. Fontecave, M., Atta, M. & Mulliez, E. S-adenosylmethionine: nothing goes to waste. *Trends Biochem. Sci.* **29**, 243-249 (2004).
  13. Schubert, H.L., Blumenthal, R.M. & Cheng, X. Many paths to methyltransfer: a chronicle of convergence. *Trends Biochem. Sci.* **28**, 329-335 (2003).
  14. Martin, J.L. & McMillan, F.M. SAM (dependent) I AM: the S-adenosylmethionine-dependent methyltransferase fold. *Curr. Opin. Struct. Biol.* **12**, 783-793 (2002).
  15. Thibodeaux, C.J., Melançon, C.E. & Liu, H.-w. Natural-Product Sugar Biosynthesis and Enzymatic Glycodiversification. *Angew. Chem. Int. Ed.* **47**, 9814-9859 (2008).
  16. Thibodeaux, C.J., Melancon, C.E. & Liu, H.-w. Unusual sugar biosynthesis and natural product glycodiversification. *Nature* **446**, 1008-1016 (2007).
  17. Zubieta, C., He, X.-Z., Dixon, R.A. & Noel, J.P. Structures of two natural product methyltransferases reveal the basis for substrate specificity in plant O-methyltransferases. *Nat Struct Mol Biol* **8**, 271-279 (2001).
  18. Liu, Y.C. *et al.* Interception of teicoplanin oxidation intermediates yields new antimicrobial scaffolds. *Nat. Chem. Biol.* **7**, 304-309 (2011).


- 
19. Chan, H.-C. *et al.* Regioselective deacetylation based on teicoplanin-complexed Orf2\* crystal structures. *Molecular BioSystems* **7**, 1224-1231 (2011).
20. Li, T.-L., Liu, Y.-C. & Lyu, S.-Y. Combining biocatalysis and chemoselective chemistries for glycopeptide antibiotics modification. *Curr. Opin. Chem. Biol.* **16**, 170-178 (2012).
21. Houge-Frydrych, C.S., Gilpin, M.L., Skett, P.W. & Tyler, J.W. SB-203207 and SB-203208, two novel isoleucyl tRNA synthetase inhibitors from a *Streptomyces* sp. II. Structure determination. *J. Antibiot. (Tokyo)*. **53**, 364-372 (2000).
22. Huang, Y.-T. *et al.* In vitro Characterization of Enzymes Involved in the Synthesis of Nonproteinogenic Residue (2S,3S)- $\beta$ -Methylphenylalanine in Glycopeptide Antibiotic Mannopeptimycin. *ChemBioChem* **10**, 2480-2487 (2009).
23. Lin, H. S-Adenosylmethionine-dependent alkylation reactions: When are radical reactions used? *Bioorg. Chem.* **39**, 161-170 (2011).
24. Akey, D.L. *et al.* A New Structural Form in the SAM/Metal-Dependent O-Methyltransferase Family: MycE from the Mycinamicin Biosynthetic Pathway. *J. Mol. Biol.* **413**, 438-450 (2011).
25. Gómez García, I. *et al.* The Crystal Structure of the Novobiocin Biosynthetic Enzyme NovP: The First Representative Structure for the TylF O-Methyltransferase Superfamily. *J. Mol. Biol.* **395**, 390-407 (2010).
26. Kopycki, J.G. *et al.* Biochemical and Structural Analysis of Substrate Promiscuity in Plant Mg<sup>2+</sup>-Dependent O-Methyltransferases. *J. Mol. Biol.* **378**, 154-164 (2008).

- 
27. Duschene, K.S., Veneziano, S.E., Silver, S.C. & Broderick, J.B. Control of radical chemistry in the AdoMet radical enzymes. *Curr. Opin. Chem. Biol.* **13**, 74-83 (2009).
28. Frey, P.A., Hegeman, A.D. & Ruzicka, F.J. The Radical SAM Superfamily. *Crit. Rev. Biochem. Mol. Biol.* **43**, 63-88 (2008).
29. Otwinowski, Z. & Minor, W. [20] Processing of X-ray diffraction data collected in oscillation mode, in *Methods in Enzymology*, Vol. Volume **276**. (ed. Charles W. Carter, Jr.) 307-326 (Academic Press, 1997).
30. Kantardjieff, K.A. & Rupp, B. Matthews coefficient probabilities: Improved estimates for unit cell contents of proteins, DNA, and protein–nucleic acid complex crystals. *Protein Sci.* **12**, 1865-1871 (2003).
31. Ness, S.R., de Graaff, R.A.G., Abrahams, J.P. & Pannu, N.S. Crank: New methods for automated macromolecular crystal structure solution. *Structure* **12**, 1753-1761 (2004).
32. De Graaff, R.A.G., Hilge, M., Van Der Plas, J.L. & Abrahams, J.P. Matrix methods for solving protein substructures of chlorine and sulfur from anomalous data. *Acta Crystallographica Section D* **57**, 1857-1862 (2001).
33. Pannu, N.S. & Read, R.J. The application of multivariate statistical techniques improves single-wavelength anomalous diffraction phasing. *Acta Crystallographica Section D* **60**, 22-27 (2004).
34. Abrahams, J.P. & Leslie, A.G.W. Methods used in the structure determination of bovine mitochondrial F1 ATPase. *Acta Crystallographica Section D* **52**, 30-42 (1996).
35. Cowtan, K. Fitting molecular fragments into electron density. *Acta Crystallographica Section D* **64**, 83-89 (2008).

- 
36. Cowtan, K. The Buccaneer software for automated model building. 1. Tracing protein chains. *Acta Crystallographica Section D* **62**, 1002-1011 (2006).
37. Emsley, P., Lohkamp, B., Scott, W.G. & Cowtan, K. Features and development of Coot. *Acta Crystallographica Section D* **66**, 486-501 (2010).
38. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallographica Section D* **60**, 2126-2132 (2004).
39. Murshudov, G.N., Vagin, A.A. & Dodson, E.J. Refinement of Macromolecular Structures by the Maximum-Likelihood Method. *Acta Crystallographica Section D* **53**, 240-255 (1997).
40. Schuck, P. Size-Distribution Analysis of Macromolecules by Sedimentation Velocity Ultracentrifugation and Lamm Equation Modeling. *Biophys. J.* **78**, 1606-1619 (2000).
41. Ravel, B. & Newville, M. ATHENA, ARTEMIS, HEPHAESTUS: data analysis for X-ray absorption spectroscopy using IFEFFIT. *Journal of Synchrotron Radiation* **12**, 537-541 (2005).
42. McCoy, A.J. *et al.* Phaser crystallographic software. *J. Appl. Crystallogr.* **40**, 658-674 (2007).
43. Dunwell, J.M., Purvis, A. & Khuri, S. Cupins: the most functionally diverse protein superfamily? *Phytochemistry* **65**, 7-17 (2004).
44. Dunwell, J.M. Cupins: a new superfamily of functionally diverse proteins that include germins and plant storage proteins. *Biotechnol. Genet. Eng. Rev.* **15**, 1-32 (1998).
45. Conrad, J.A. & Moran, G.R. The interaction of hydroxymandelate synthase with the 4-hydroxyphenylpyruvate dioxygenase inhibitor: NTBC. *Inorg. Chim. Acta* **361**, 1197-1201 (2008).



- 
46. Brownlee, J., He, P., Moran, G.R. & Harrison, D.H.T. Two Roads Diverged: The Structure of Hydroxymandelate Synthase from *Amycolatopsis orientalis* in Complex with 4-Hydroxymandelate<sup>†,‡</sup>. *Biochemistry (Mosc)*. **47**, 2002-2013 (2008).
47. Purpero, V. & Moran, G. The diverse and pervasive chemistries of the  $\alpha$ -keto acid dependent enzymes. *J Biol Inorg Chem* **12**, 587-601 (2007).
48. Moran, G.R. 4-Hydroxyphenylpyruvate dioxygenase. *Arch. Biochem. Biophys.* **433**, 117-128 (2005).
49. Singh, S. *et al.* Structure and Mechanism of the Rebeccamycin Sugar 4'-O-Methyltransferase RebM. *J. Biol. Chem.* **283**, 22628-22636 (2008).
50. Vidgren, J., Svensson, L.A. & Liljas, A. Crystal structure of catechol O-methyltransferase. *Nature* **368**, 354-358 (1994).
51. Männistö, P.T. & Kaakkola, S. Catechol-O-methyltransferase (COMT): Biochemistry, Molecular Biology, Pharmacology, and Clinical Efficacy of the New Selective COMT Inhibitors. *Pharmacol. Rev.* **51**, 593-628 (1999).
52. Schluckebier, G., O'Gara, M., Saenger, W. & Cheng, X. Universal Catalytic Domain Structure of AdoMet-dependent Methyltransferases. *J. Mol. Biol.* **247**, 16-20 (1995).
53. Costas, M., Mehn, M.P., Jensen, M.P. & Que, L. Dioxygen Activation at Mononuclear Nonheme Iron Active Sites: Enzymes, Models, and Intermediates. *Chem. Rev.* **104**, 939-986 (2004).
54. Solomon, E.I. *et al.* Geometric and Electronic Structure/Function Correlations in Non-Heme Iron Enzymes. *Chem. Rev.* **100**, 235-350 (2000).

- 
55. Decker, A. & Solomon, E.I. Dioxygen activation by copper, heme and non-heme iron enzymes: comparison of electronic structures and reactivities. *Curr. Opin. Chem. Biol.* **9**, 152-163 (2005).
56. Gerlt, J.A., Kreevoy, M.M., Cleland, W.W. & Frey, P.A. Understanding enzymic catalysis: the importance of short, strong hydrogen bonds. *Chem. Biol.* **4**, 259-267 (1997).

## Appendix



### Curriculum Vitae

#### Personal Information

Name: Yu-Chen Liu

Date of Birth: 1979/09/20

Gender: male

Nationality: Taiwan

E-mail: ycl0920@gate.sinica.edu.tw

Contact Number: +8862-27898070

Contact Address: 3L08, Genomics Research Center, Academia Sinica, Taipei 115, Taiwan.

#### Professional Profile

Current Research: Structural and functional studies of enzymes involved in the biosynthesis of natural products.

Affiliated Institution:

1. Genomics Research Center (GRC), Academia Sinica, Taipei, Taiwan.
2. Chemical Biology and Molecular Biophysics (CBMB) Program, Taiwan International Graduate Program (TIGP), Academia Sinica, Taipei, Taiwan.
3. Institute of Biochemical Sciences (IBS), National Taiwan University, Taipei, Taiwan.

#### Education and Experience

B.S., Department of Chemistry, National Cheng-Kung University, 1998- 2002.

M.S., Department of Biochemistry, National Cheng-Kung University Medical College, 2002- 2004.

R.A., Institute of Molecular Biology (IMB), Academia Sinica, 2004-2005.

R.A., Genomics Research Center (GRC), Academia Sinica, 2007-2009.

Ph.D., Chemical Biology and Molecular Biophysics (CBMB) Program, Taiwan.

International Graduate Program (TIGP), Academia Sinica, 2009-2013.




### **Conferences and Workshops**

1. **Natural Products Conference 2012-Natural Products Synthesis and Biosynthesis**, Puerto Calero, Lanzarote, Spain. (February 10-13, 2012)
2. **48<sup>th</sup> International Conference on Medicinal Chemistry-RICT 2012-Interfacing Chemical Biology and Drug Discovery**, Poitiers, France. (July 4-6, 2012)
3. **1<sup>st</sup> OIST CCP4 school-From data processing to structure refinement and beyond**, Okinawa, Japan. (December 5-9, 2011)

### **Academic Honors and Awards**

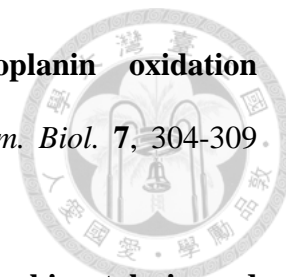
1. Outstanding performance for the 2010 CBMB, IBC, and IBS retreat. (The second prize of poster and oral competition award)
2. Outstanding performance for the 2011 16<sup>th</sup> Biophysics Conference. (The first prize of poster and oral competition award)
3. Outstanding performance for the 2011 Symposium on Frontiers of Biomedical Sciences. (The first prize of poster and oral presentation award)-Prof. Jung-Yaw Lin Science and Education Foundation award.
4. Outstanding student research award, College of Life Science, National Taiwan University (2011).
5. Outstanding performance for the 2011 NSRRC 17<sup>th</sup> Users' Meeting & Workshop. (The first prize of poster and oral presentation in biological science award-The Glory of Taiwan)
6. Outstanding student research award, 21<sup>st</sup> Wang Ming-Ning Memorial Foundation award (2011).

- 
7. Outstanding student research award, 1<sup>st</sup> GRC research staff award (2011).
  8. TIGP student conference travel grant award. (The first prize of the grant award for the period of January-April 2012)
  9. Outstanding student research award, 2<sup>nd</sup> GRC research staff award (2012).

### **Publications**

1. Jia-Hau Shiu, Chiu-Yueh Chen, Long-Sen Chang, Yi-Chun Chen, Yen-Chin Chen, Yu-Hui Lo, Yu-Chen Liu, and Woei-Jer Chuang\*, **Solution Structure of  $\gamma$ -Bungarotoxin: The Functional Significance of Amino Acid Residues Flanking the RGD Motif in Integrin Binding**, *Proteins* **57**, 839-849 (2004).
2. Chiu-Yueh Chen, Jia-Hau Shiu, Yao-Hsun Hsieh, Yu-Chen Liu, Yen-Chin Chen, Yi-Chun Chen, Wen-Yih Jeng, Ming-Jer Tang, Szecheng J. Lo, and Woei-Jer Chuang\*, **Effect of D to E mutation of the RGD motif in rhodostomin on its activity, structure, and dynamics: Importance of the interactions between the D residue and integrin**, *Proteins* **76**, 808-821 (2009).
3. Yu-Ting Huang, Syue-Yi Lyu, Pei-Hsuan Chuang, Ning-Shian Hsu, Yi-Shan Li, Hsiu-Chien Chan, Chuen-Jiuan Huang, Yu-Chen Liu, Chang-Jer Wu, Wen-Bin Yang, and Tsung-Lin Li\*, **In vitro Characterization of Enzymes Involved in the Synthesis of Nonproteinogenic Residue (2S,3S)- $\beta$ -Methylphenylalanine in Glycopeptide Antibiotic Mannopeptimycin**, *ChemBioChem* **10**, 2480-2487 (2009).
4. Hsiu-Chien Chan, Yu-Ting Huang, Syue-Yi Lyu, Chuen-Jiuan Huang, Yi-Shan Li, Yu-Chen Liu, Chia-Cheng Chou, Ming-Daw Tsai\*, and Tsung-Lin Li\*, **Regioselective deacetylation based on teicoplanin-complexed Orf2\* crystal structures**, *Mol. BioSyst.* **7**, 1224-1231 (2011).
5. Yu-Chen Liu, Yi-Shan Li, Syue-Yi Lyu, Li-Jen Hsu, Yu-Hou Chen, Yu-Ting Huang, Hsiu-Chien Chan, Chuen-Jiuan Huang, Gan-Hong Chen, Chia-Cheng Chou, Ming-

Daw Tsai, and Tsung-Lin Li\*, **Interception of teicoplanin oxidation intermediates yields new antimicrobial scaffolds**, *Nat. Chem. Biol.* **7**, 304-309. (2011).



6. Tsung-Lin Li\*, Yu-Chen Liu, and Syue-Yi Lyu, **Combining biocatalysis and chemoselective chemistries for glycopeptide antibiotics modification**, *Curr. Opin. Chem. Biol.* **16**, 170-178 (2012).
7. Hai-Chen Wu, Yi-Shan Li, Yu-Chen Liu, Syue-Yi Lyu, Chang-Jer Wu, and Tsung-Lin Li\*, **Chain elongation and cyclization in type III PKS DpgA**, *ChemBioChem* **13**, 862-871 (2012).
8. Xiao-Wei Zou<sup>†</sup>, Yu-Chen Liu<sup>†</sup>, Ning-Shian Hsu, Chuen-Jiuan Huang, Syue-Yi Lyu, Hsiu-Chien Chan, Chin-Yuan Chang, Hsien-Wei Yeh, Kuan-Hung Lin, Chang-Jer Wu, Ming-Daw Tsai, and Tsung-Lin Li\*, **Structure and mechanism of non-heme iron-SAM dependent methyltransferase and its engineering to hydratase**. (†These authors contribute equally) (submitted)
9. Syue-Yi Lyu, Yu-Chen Liu, Chuen-Jiuan Huang, Chin-Yuan Chang, Ning-Shian Hsu, Kuan-Hung Lin, Chang-Jer Wu, Ming-Daw Tsai, and Tsung-Lin Li\*, **Multi-phased acyltransferase crystal structures lead to discovery of new chemistry and chemicals**. (submitted)
10. Chin-Yuan Chang, Syue-Yi Lyu, Yu-Chen Liu, Ning-Shian Hsu, Chih-Chung Wu, Cheng-Fong Tang, Kuan-Hung Lin, Jin-Yuan Ho, Chang-Jer Wu, Ming-Daw Tsai, and Tsung-Lin Li\*, **Unexpected dihydroxyarginine and hydroxycapreomycin are the actual biosynthetic intermediacy toward streptolidine**. (submitted)

## Patents

1. **Structural and mechanistic basis for novel compound biosynthesis using the 4-electron hexose oxidase**, Tsung-Lin Li, Yu-Chen Liu, Yi-Shan Li, Syue-Yi Lyu,

USA/TAIWAN (application No.: 12/896,747; publication No.: US2012/0108498A1)

2. **Crystal-structures-based development for new glycopeptide antibiotics**, Tsung-

Lin Li, Yu-Chen Liu, Syue-Yi Lyu, USA/TAIWAN (Academia Sinica file No.:

28A-991215)

