

β -Hydroxylation of α -amino- β -hydroxybutanoyl-glycyluridine catalyzed by a nonheme hydroxylase ensures the maturation of caprazamycin

Saeid Malek Zadeh^{1,2,3}, Mei-Hua Chen^{1,2,4}, Zhe-Chong Wang¹, Elahe K. Astani⁵, I-Wen Lo¹, Kuan-Hung Lin¹, Ning-Shian Hsu¹, Kamal Adhikari^{1,6}, Syue-Yi Lyu¹, Hsin-Ying Tsai¹, Yuma Terasawa⁷, Miyuki Yabe⁷, Kazuki Yamamoto⁷, Satoshi Ichikawa^{7,8} & Tsung-Lin Li^{1,2,5,9}✉

Caprazamycin is a nucleoside antibiotic that inhibits phospho-*N*-acetylmuramyl-pentapeptide translocase (MraY). The biosynthesis of nucleoside antibiotics has been studied but is still far from completion. The present study characterized enzymes Cpz10, Cpz15, Cpz27, Mur17, Mur23 out of caprazamycin/muraymycin biosynthetic gene cluster, particularly the nonheme α KG-dependent enzyme Cpz10. Cpz15 is a β -hydroxylase converting uridine mono-phosphate to uridine 5' aldehyde, then incorporating with threonine by Mur17 (Cpz14) to form 5'-C-glycyluridine. Cpz10 hydroxylates synthetic 11 to 12 in vitro. Major product 13 derived from mutant Δ cpz10 is phosphorylated by Cpz27. β -Hydroxylation of 11 by Cpz10 permits the maturation of caprazamycin, but decarboxylation of 11 by Mur23 oriented to muraymycin formation. Cpz10 recruits two iron atoms to activate dioxygen with regio-/stereo-specificity and commit electron/charge transfer, respectively. The chemo-physical interrogations should greatly advance our understanding of caprazamycin biosynthesis, which is conducive to pathway/protein engineering for developing more effective nucleoside antibiotics.

¹Genomics Research Center, Academia Sinica, Taipei, Taiwan. ²Chemical Biology and Molecular Biophysics Program, Taiwan International Graduate Program, Academia Sinica, Taipei, Taiwan. ³Institute of Bioinformatics and Structural Biology, National Tsing Hua University, Hsinchu, Taiwan. ⁴Department of Chemistry, National Taiwan University, Taipei, Taiwan. ⁵Department of Chemistry, Faculty of Science, Tarbiat Modares University, Tehran, Iran. ⁶Molecular and Biological Agricultural Sciences Program, Taiwan International Graduate Program, Academia Sinica and National Chung Hsing University, Taipei, Taiwan. ⁷Graduate School of Pharmaceutical Sciences, Hokkaido University, Sapporo, Japan. ⁸Global Station for Biosurfaces and Drug Discovery, Global Institution for Collaborative Research and Education (GI-CoRE), Hokkaido University, Sapporo, Japan. ⁹Biotechnology Center, National Chung Hsing University, Taichung City, Taiwan. ✉email: tlili@gate.sinica.edu.tw

Nucleoside antibiotics are a group of natural products which specifically inhibit phospho-*N*-acetylmuramyl-pentapeptide translocase (MraY), an essential enzyme involved in the bacterial peptidoglycan biosynthesis¹. Up to date, no any antibiotics that inhibit MraY have been approved for clinical use, underscoring that MraY and nucleoside antibiotics respectively are drug target and drug candidate in pair as an ideal strategy to conquer antibiotic drug resistance². The pathway elucidation for potent nucleoside antibiotics has been a research hot spot³. Two approaches are generally employed in the field, namely, *in vivo* and *in vitro* approaches. Both are not exclusive but complementary to each other. Noticeably, collecting and purifying hydrophilic intermediates from given mutants are tedious and challenging⁴. On top of that, no known substrates or substrates unavailability are another difficulty in *in-vitro* study. These limitations considerably hamper the progress of biosynthetic pathway illustration of natural products. To unravel these shackles, an integrated approach by combining *in vivo*, *in vitro*, synthetic chemistry and computational chemistry is emerging that has been proven powerful⁵.

Nucleoside antibiotics, such as FR-900493 (1), caprazamycin (2), liposidomycin (3), sphaerimicin (4), and muraminomicin (5) (Fig. 1), all feature a uridine moiety in addition to an unusual 5'-amino-5'-deoxyribose (ADR) (Fig. 1), In view of structures, nucleoside antibiotics can be subdivided into two groups: the lipid I transition-like subgroup *e.g.* compounds 2–5 and the Park's nucleotide subgroup *e.g.* compounds 6 and 7. The uridine moiety is coupled with a short-chain amino acid to form an ADR-GlyU (5'-*C*-glycyluridine) core the drug warhead, which can be attached with a 6'-*N*-alkylamine side chain. This alkylamine is hydroxylated at its β -position (β -OH) for the biosynthesis toward the lipid I transition-like subgroup^{6,7}. On the other hand, the decarboxylation of the 6'-*N*-alkylamine side chain is pivotal toward the biosynthesis of the Park's nucleotide subgroup (see below). FR-900493 (1) acts as a rudimentary structure, where the ADR-GlyU disaccharide core is associated with an aminopropyl group and a methyl group at the glycine portion. By contrast, caprazamycin (2) includes a characteristic seven-membered diazepanone ring with two methyl groups each at one of two ring-nitrogen atoms (caprazol 8); this heterocyclic ring is decorated

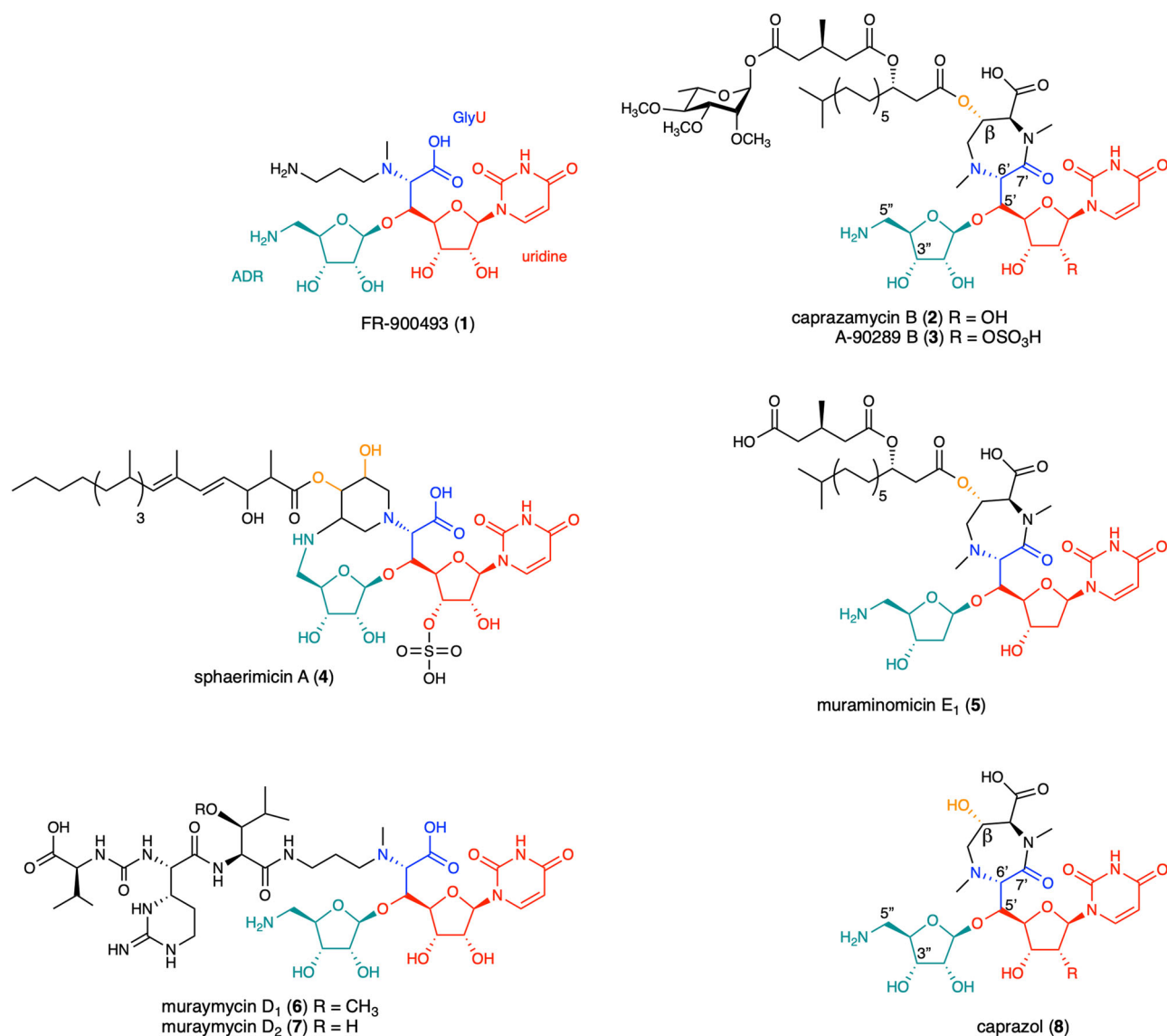


Fig. 1 Chemical structures of representative MraY inhibitors. Nucleoside antibiotics are composed of some characteristic moieties: uridine colored red, 5'-amino-5'-deoxyribose (ADR) colored green, the glycine moiety colored blue and the β -hydroxyl group colored orange. The numbering system referred to in the text is designated.

with a long lipid side chain at the β -OH, to which 3-methyl glutaric acid and 2,3,4-*O*-methyl-*L*-rhamnose are appended^{8–11}.

The addition of the fatty acyl chain to caprazol **8** has been determined in vivo using a gene-deleted mutant (Δ cpz23, coding for a lipase-like enzyme)¹², underlining the importance of the β -OH functionality. However, the enzyme that commits the β -hydroxylation remains elusive. In view of the biosynthetic gene clusters (BGCs) of lipo-nucleoside antibiotics caprazamycin (**2**) and liposidomycin (A-90289) (**3**), each individual BGC contains two α KG-dependent monooxygenases Cpz10,15 and LipG,L^{13,14}, the potential candidates. Nonheme iron-dependent enzymes play crucial roles in the diversification of secondary metabolites¹⁵. This type of enzymes exists across all domains of lives and has been studied extensively, for example, dopamine hydroxylase, deacetoxycephalosporin C synthase (DAOCS) and taurine dioxygenase (TauD)^{16–18}. These enzymes utilize a high-spin iron (Fe^{II}) to activate dioxygen into a reactive Fe^{IV} -oxo species at the expense of one molecule of α KG to oxygenate/oxidize/rearrange a wide variety of organic compounds^{17,19}. The active site is located within a core β -barrel domain in this class of enzymes¹⁶. Here we report that Cpz15 is the starter enzyme; Cpz10 is the β -hydroxylase recruiting an additional auxiliary iron to bring its activity into full play; Cpz27 is a phosphotransferase, alongside biochemical functionalities of Cpz13, 14 and Mur23. Our in vivo and in vitro results suggest that the phosphorylation at the 3'-OH takes place after glycosylation, methylation and alkylation at odds with the previously proposed scenario^{4,20}.

Results

Cpz10 and 15 are nonheme oxygenases. Two genes, *cpz10* and *cpz15*, that code for α KG dependent enzymes, exhibit in the BGC of caprazamycin. In silico analysis, *cpz15* is a homolog to the clavaminic acid synthetase (CAS)-like protein, which also appears in the BGCs of liposodomyin, jawsamycin, capuramycin (Supplementary Fig. 1) and muramycin (**6**, **7**). The one in muramycin is *mur16* (Supplementary Fig. 2), while the functions of Cpz15 and Mur16 were annotated differently. Cpz15 is designated as an oxygenase (Supplementary Fig. 3) responsible for the β -hydroxylation at a later step in contrast to Mur16 that engages in an oxidative dephosphorylation reaction in the very first step¹⁴. Rather, Cpz10 was assigned a role as that of Mur16. Recently, Cpz15 is re-proposed as the starter enzyme responsible for the oxidative dephosphorylation reaction²¹. To resolve this controversy, we set off biochemical examination for purified Cpz10 or Cpz15 versus possible substrates, including uridine, uridine monophosphate (UMP), uridine diphosphate (UDP) and uridine triphosphate (UTP). Of them, only did a new peak emerge on the LC trace concomitant with dwindling of UMP in the reactions with Cpz15 (Fig. 2A, and 3A). This new peak was subjected to MS analysis, which showed a mass unit ($M + H$)⁺ of $m/z = 243.1$ indicating formation of uridine 5'-aldehyde (U5'A **9**) (Supplementary Fig. 4)²², thereby redefining Cpz15 a nonheme-dependent oxidative de-phosphorylase rather than a β -hydroxylase.

As shown in Fig. 2A, given that UMP is oxidized by Cpz15 to U5'A (**9**)²², compound **9** would then be catalyzed by Cpz14 with threonine to form 5'-*C*-glycyluridine (5'GlyU **10**). To confirm this prediction, we expressed Cpz14 in *E. coli* or *Streptomyces*, while, unfortunately, it is insoluble in either system. To overcome this problem, we borrowed Mur17 from the muramycin BGC, an enzyme equivalent to Cpz14 and soluble in *E. coli*. The LC/MS analysis shows formation of compound **10** (expected mass unit at $m/z = 318.16$ ($M + H$)⁺) for the reactions with compound **9** and threonine plus Mur17 (Figs. 2A and 3B; and Supplementary Fig. 5) in agreement with previous results^{6,23}. Interestingly, not only threonine but also glycine can be accepted by Mur17 while assayed against all proteinogenic *L*-amino acids and their *D*-isomers.

Cpz10 is the β -hydroxylase. Since Cpz15 is not the β -hydroxylase, the likelihood now befalls Cpz10. In silico analysis, Cpz10 also shows protein sequence similarity/identity analogous to nonheme α KG dependent enzymes closest to aspartyl/asparaginyl β -hydroxylase (Supplementary Fig. 6). Concerning the reaction timing/order, the β -hydroxylation could occur on free amino acids, compound **11** (see below), or the seven-membered heterocyclic ring (e.g. compound **8** without β -OH). We first assayed Cpz10 against all *L/D*-amino acids, dopamine and pipercolic (Supplementary Fig. 7). Interestingly, *L*-methionine and *L*-selenomethionine were found subject to oxygenation, while the oxygenation takes place at the sulfur (MSO)/selenium atom instead of the expected β -carbon (β C; HR-MS, LC-MS/MS in conjunction with 2D NMR confirms the product is MSO with the ($M + H$)⁺ ion at $m/z = 166.05$, Figs. 2B and 3C and Supplementary Fig. 8). Next, *S*-adenosylmethionine (SAM) was examined, which was however ruled out as no product found (data not shown). It has been proposed that the four-carbon sidechain is transferred from SAM to compound **10** by Cpz13 to compound **11** (Figs. 2A and 3B, D and Supplementary Fig. 4a). Because compound **9** is highly unstable, the production of compounds **10** and **11** is therefore limited. These two compounds were then chemically synthesized (see Supplementary methods) following previously published methodologies in total synthesis of caprazamycin¹⁴. With the synthetic compound **11** in hand, Cpz10 were examined, of which a new peak was brought on LC traces (Fig. 2A, and Fig. 3D). HR-MS and LC-MS/MS analysis showed that this new peak has an ($M + H$)⁺ ion at $m/z = 435.1357$ in line with the predicted molecular weight of the hydroxylated compound **11** ($m/z = 435.1358$ for $\text{C}_{15}\text{H}_{22}\text{N}_4\text{O}_{11} + \text{H}$) (see Supplementary Note 1). X-ray crystallography, at long last, confirmed that the peak is α -amino- β -hydroxylbutanoyl-glycyluridine (compound **12**) (see below). Because of limited quantity of compound **11**, we took advantage of *L*-methionine (as Cpz10 can oxygenate it to MSO) to estimate the basic kinetics for Cpz10; the Michaelis constant and turnover rate were determined under the pseudo-first order condition: $K_m = 892.1 \pm 9 \mu\text{M}$ and $k_{\text{cat}} = 58.6 \pm 0.8 \text{ min}^{-1}$, which are comparable to those of typical nonheme monooxygenases (Supplementary Fig. 9).

Cpz10 X-ray crystal structures reveal two irons. To better understand the catalytic mechanism of Cpz10, we set out to determine crystal structures of Cpz10 in apo and complex forms (Supplementary Fig. 10). The purified Cpz10 was first crystallized and subjected to X-ray diffraction. The phase problem was solved using molecular replacement (MR), whereby the initial model obtained was used as the template for building complex structures. We determined five structures, including one apo structure (PDB code: 7V4N), two mutants (PDB codes: 7V4O and 7V4P) and two complexes (with α KG (PDB code: 7V4M), or with succinic acid (SA), compound **12** (PDB code: 7V4F) (see below, and Supplementary Fig. 11B for stereo views and density maps; the data-collection and refinement statistics are summarized in Table 1). Based on gel filtration chromatograph, Cpz10 is a homodimer in an aqueous solution (Supplementary Fig. 12). Consistently, two polypeptide chains are packed in an asymmetric unit (Supplementary Fig. 11) with immense protein-protein interactions at the dimer interface. A single polypeptide chain is folded in a twisted β -sheet jelly roll (cupin) domain, a common fold for a typical α KG-dependent nonheme enzyme²⁴. Unexpectedly, two iron metals were identified intrinsically existing in all structures determined except the apo form. The ferric iron, labeled Fe1, is held by an archetypal metal-chelating triad (His115, His160, Asp117), α KG that coordinates to Fe1 in a bidentate manner, and a water molecule (Wt91) that takes the

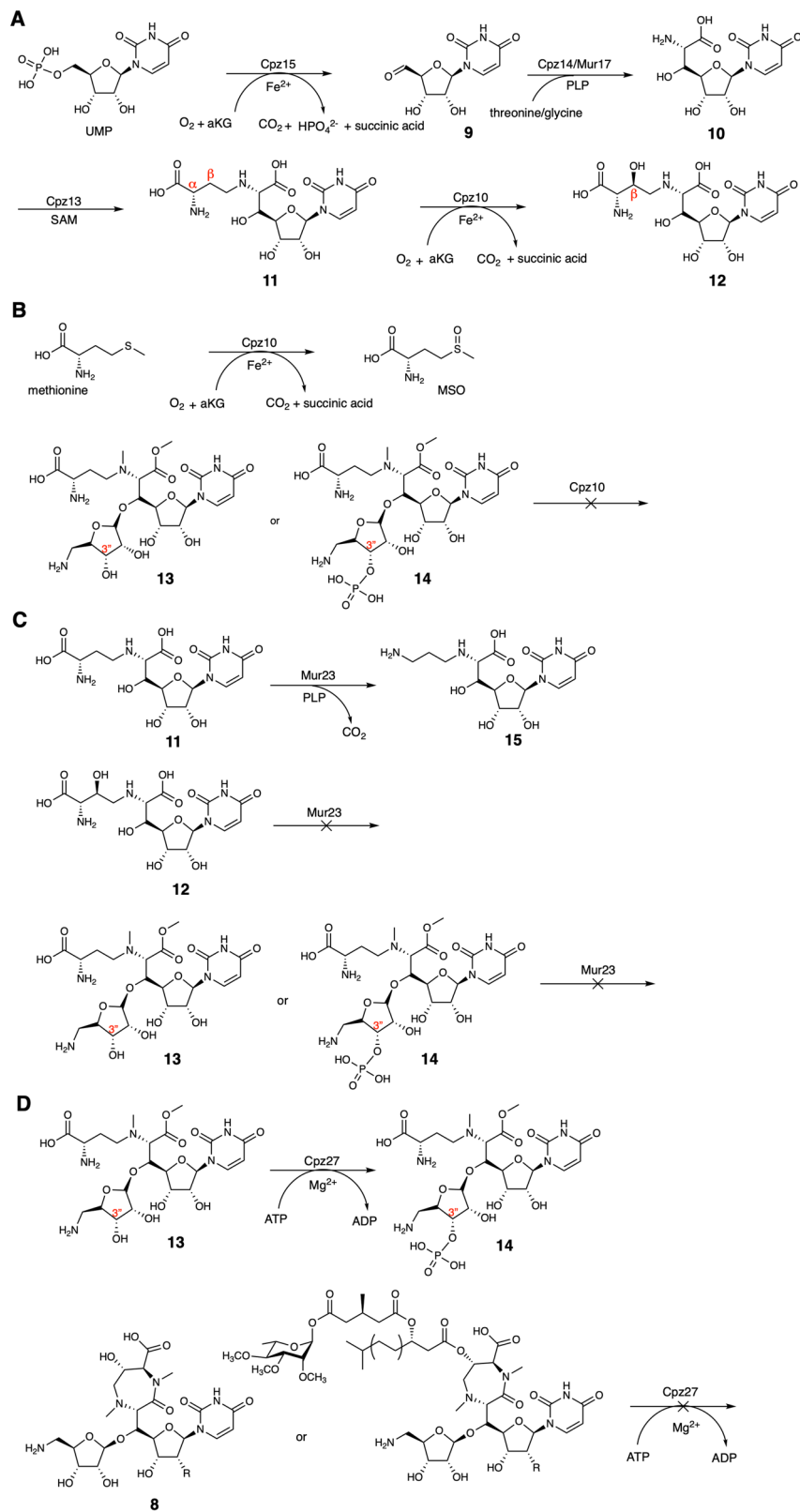


Fig. 2 Reaction scheme for selected enzymes referred to in this study. **A** Biosynthetic steps committed by Cpz15, Cpz14/Mur17, Cpz13 and Cpz10 in sequence. **B** Methionine or selenium can be oxygenated by Cpz10; compounds **13** or **14** cannot be accepted by Cpz10. **C** Compound **11** can be decarboxylated by Mur23 to compound **15** as opposed to compound **12** that cannot be consumed. **D** Compound **13** can be phosphorylated by Cpz27 to compound **14**, while compound **8** cannot be utilized.

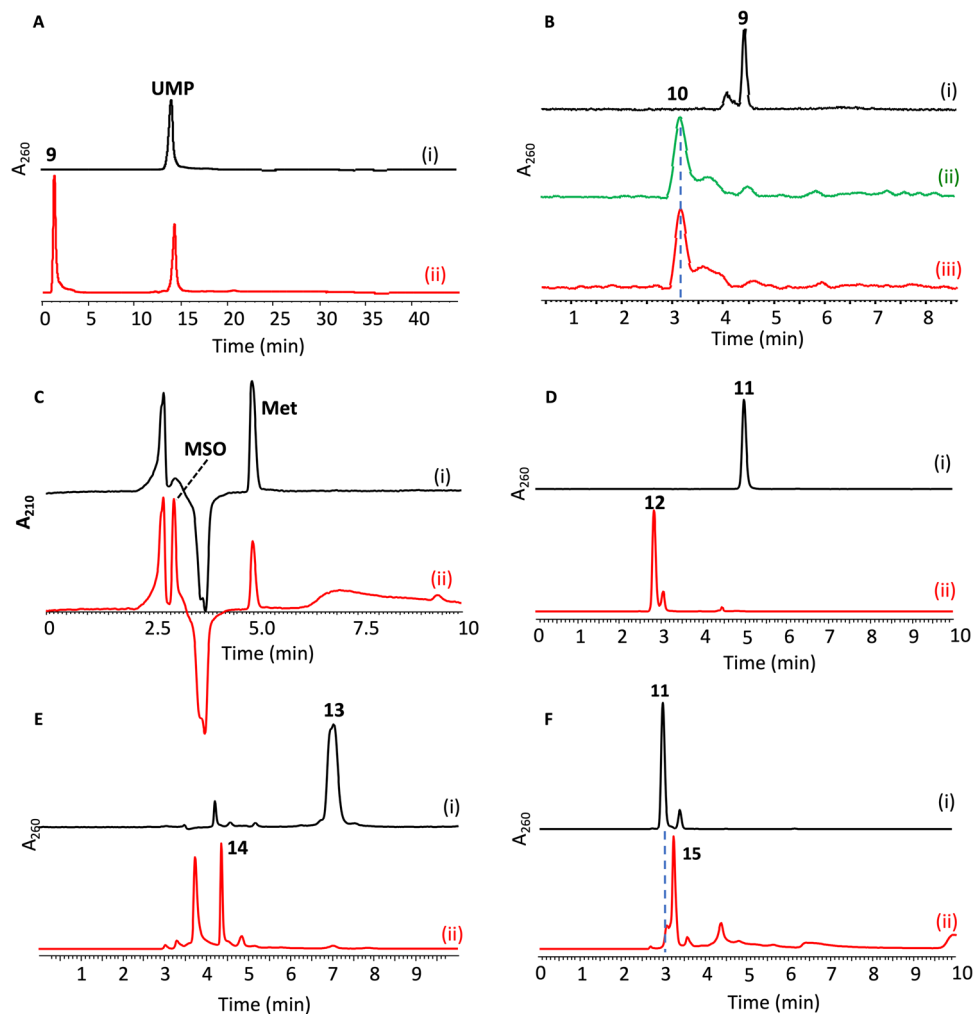


Fig. 3 LC traces for enzymatic reactions catalyzed by enzymes involved in the biosynthesis of nucleoside antibiotics. **A** Oxidative dephosphorylation reactions catalyzed by Cpz15, in which UMP (i, without Cpz15) is converted to compound **9** (ii, with Cpz15). **B** Coupling reactions catalyzed by Mur17, in which compound **9** (i, without enzyme) plus either threonine (ii) or glycine (iii) can be converted to compound **10**; both substrate and product were derivatized with dansyl chloride before subjected to LC analysis. **C** Oxygenation reactions catalyzed by Cpz10, in which methionine (i, without Cpz10) can be converted to MSO (ii, with Cpz10). **D** Hydroxylation reactions catalyzed by Cpz10, in which synthetic compound **11** (i, without Cpz10) can be converted to compound **12** (ii, with Cpz10). **E** Phosphorylation reactions catalyzed by Cpz27, in which collected compound **13** (i, without Cpz27) can be converted to compound **14** in the presence of ATP (ii, with Cpz27). **F** Decarboxylation reactions catalyzed by Mur23, in which compound **11** (i, without Mur23) can be converted to compound **15** (ii, with Mur23); both substrate and product were derivatized with dansyl chloride before subjected to LC analysis.

sixth coordinate altogether to form a tetragonal bipyramidal geometry (Fig. 4)^{17,25,26}. This 6-coordinate complex is akin to most reported complexes of α KG-dependent enzymes, where the carboxyl group of α KG coordinates *trans* to the distal histidine. As to the second iron, labeled Fe2, it is associated with two water molecules (Wt56 and Wt89) in close proximity to residues Arg170, Arg133, His144, Thr128, and Ile145 but away from Fe1 by 12 Å (Fig. 4 and Supplementary Fig. 13).

Based on sequence alignment, the metal-chelating residues His115, Asp117, and His160 are highly conserved as an iron-binding motif HRDX_nH in the nonheme oxygenase superfamily; Arg170, another conserved residue, recognizes and interacts with α KG through electrostatic forces. Interestingly, a cluster of aromatic or soft atom-containing residues, His172, His144, Thr128, His123, Tyr112, and Met104, are identified to form a cryptic electron/charge-transfer tunnel in a manner similar to that observed in Rieske-type oxygenases (Supplementary Fig. 14a, b)^{27–29}: residues His172 and His123 (2.8 Å) take a unique geographic position in a two-hand holding shape in close proximity to Fe2 akin to the one reported in *L*-proline *cis*-4-

hydroxylase (PDB CODE: 4P7X) (Supplementary Fig. 15)¹⁶; residues Thr128 and His144 are on par with residues His172 and His123 in a position to interact with Fe2. Fe1 that is held firmly by chelating residues in a short and strong coordinate covalent bond manner (2.0–2.2 Å); in contrast, Fe2 is held in a relatively loose way by nearby residues (3–4 Å) alongside a satellite water molecule (Wt56, 2.9 Å). These two irons are spatially and temporarily associated for electron/charge transfer by an electron/charge relay system made of a constellation of well-cooperative network Wt89, His172, His123, and α KG toward Fe1 (see below, and Fig. 5A, Supplementary Movie 1, Supplementary Movie 2), where the distances between each set of consecutive relay stations are <2.8 Å. To confirm the importance of residues His172 and His123 in this electron/charge transfer system, we performed site-directed mutagenesis (H172A and H123A). Both mutants, somehow, are insoluble, implying that they may assume a critical structural role in maintaining the integrity of the tunnel in addition to the proposed electron/charge relay station.

To determine the regioselectivity and stereoselectivity reactions catalyzed by Cpz10, we were prompted to pursue the quaternary

Table 1 Data collection and refinement statistics (molecular replacement).

	7V4N	7V4M	7V4F	7V4O	7V4P
<i>Data collection</i>					
Space group	P 4 ₃ 2 ₁ 2	C 2 2 2 ₁	P 4 ₃ 2 ₁ 2	C 1 2 1	P 4 ₃ 2 ₁ 2
Cell dimensions					
<i>a</i> , <i>b</i> , <i>c</i> (Å)	82.2, 82.2, 104.2	75.18, 77.56, 68.11	82.04, 82.04, 103.89	150.83, 39.03, 76.59	82.04, 82.04, 103.89
α , β , γ (°)	90, 90, 90	90, 90, 90	90, 90, 90	90, 119.33, 76.59	90, 90, 90
Resolution (Å)	29.83–2.20 (2.28–2.20)	19.86–1.90 (1.97–1.90)	26.48–1.97 (2.04–1.97)	25.03–1.65 (1.71–1.65)	26.46–1.94 (2.02–1.95)
<i>R</i> _{sym} or <i>R</i> _{merge}	0.078 (0.79)	0.039 (0.501)	0.040 (0.570)	0.051 (0.141)	0.066 (0.509)
<i>I</i> / σ <i>I</i>	24.9 (3.4)	23.6 (2.4)	33.5 (3.8)	32.3 (13.1)	36.9 (4.2)
Completeness (%)	100.0 (100.0)	100.0 (100.0)	96.2 (98.1)	99.0 (100.0)	100.0 (100.0)
Redundancy	14.4 (14.7)	5.3 (5.4)	9.8 (9.9)	7.0 (7.9)	15.7 (13.2)
<i>Refinement</i>					
Resolution (Å)	2.2	1.9	1.98	1.65	1.95
No. reflections	17812	15444	25285	46715	26621
<i>R</i> _{work} / <i>R</i> _{free}	0.1864/0.2393	0.1865/0.2331	0.1911/0.2510	0.1803/0.2138	0.1877/0.2413
No. atoms					
Protein	1403/1406	1374	1400/1390	1447/1447	1404/1399
Ligand/ion	2	18	45	17/17	2
Water	44	135	171	421	112
<i>B</i> -factors					
Protein	30.14/36.74	21.73	24.36/32.23	16.77/13.93	23.24/23.41
Ligand/ion	49.99	28.3	26.93	25.125	37.14
Water	25.97	32.41	30.75	27.21	20.29
R.m.s. deviations					
Bond lengths (Å)	0.0085	0.007	0.008	0.006	0.011
Bond angles (°)	1.3695	0.919	1.13	1.03	1.27

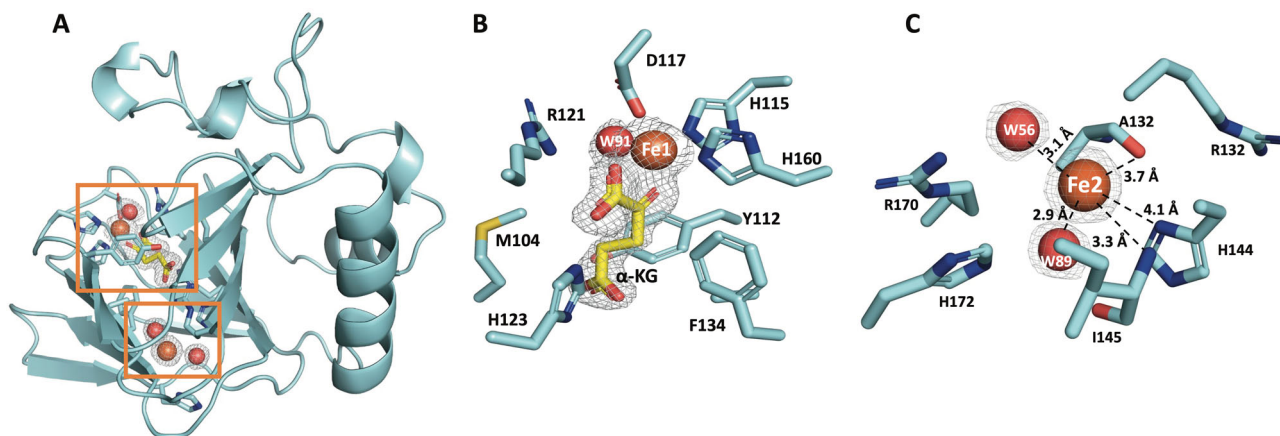


Fig. 4 Crystal structure of Cpz10 in complex with irons and α -KG. **A** The overall structure of Cpz10, of which the polypeptide chain (colored cyan) folds into a major distorted jelly roll β -sheet core and an α -helical-loop domain. Two irons and α -KG are positioned in the center of the β -sheet core. **B** The closed-up view of active site, where the active site iron (Fe1, colored orange) is held by a typical facial triad H115, H160, and D117 (cyan sticks) in association with the α -KG bidentate (yellow stick) and a water molecule (W91, colored red) to form a tetragonal bipyramidal coordination geometry. **C** The closed-up view of the second iron (Fe2, colored orange) which is located in the end of the β -sheet core. Fe2 is associated intimately with two water molecules (W89 and W56 in 2.9 and 3.1 Å, respectively) as well as with the backbone of R133 and in short distance with residues F134, H144, R170, and H172.

structure of Cpz10 (in complex with Fe- α KG and compound **11**). The complex that finally obtained shows that α KG soaked in the crystals has been converted to succinic acid likely undergone an oxidative decarboxylation reaction. Likewise, the soaked compound **11** has been converted to compound **12**, where a new hydroxyl group (β -OH) is installed at β C of the 2-amino-butanoic acid moiety in *S*-configuration (compound **12**) despite that the ADR-GlyU moiety cannot be flawlessly built because of lack of well-defined electron density as a result of exposed to bulk solvent (Fig. 6 and Supplementary Figs. 11 and 16). Based on the quaternary complex, the overall catalytic mechanism of Cpz10 is proposed as follows: α KG is ushered into the active site in

association with Fe1 (Fe^{III}) to form a high-spin iron-centered tetragonal bipyramidal complex with a water molecule atop the tetragonal plane. Fe1 is then reduced to Fe^{II} by Fe2 through the internal electron/charge relay system identified herein. When compound **11** is placed at its binding site, O_2 supersedes H_2O to form a Fe^{II} -superoxide anion (or a Fe^{III} -superoxide). The superoxide anion subsequently attacks the carbonyl carbon electrophile of α KG, resulting in formation of succinic acid (oxidative decarboxylation) and the reactive Fe^{IV} -oxo species (by heterolysis). Fe^{IV} -oxo then abstracts the nearby *pro-S* H-atom from the β C of compound **11** in situ to form Fe^{III} -OH and the corresponding compound **11** radical. Compound **12** is instantly

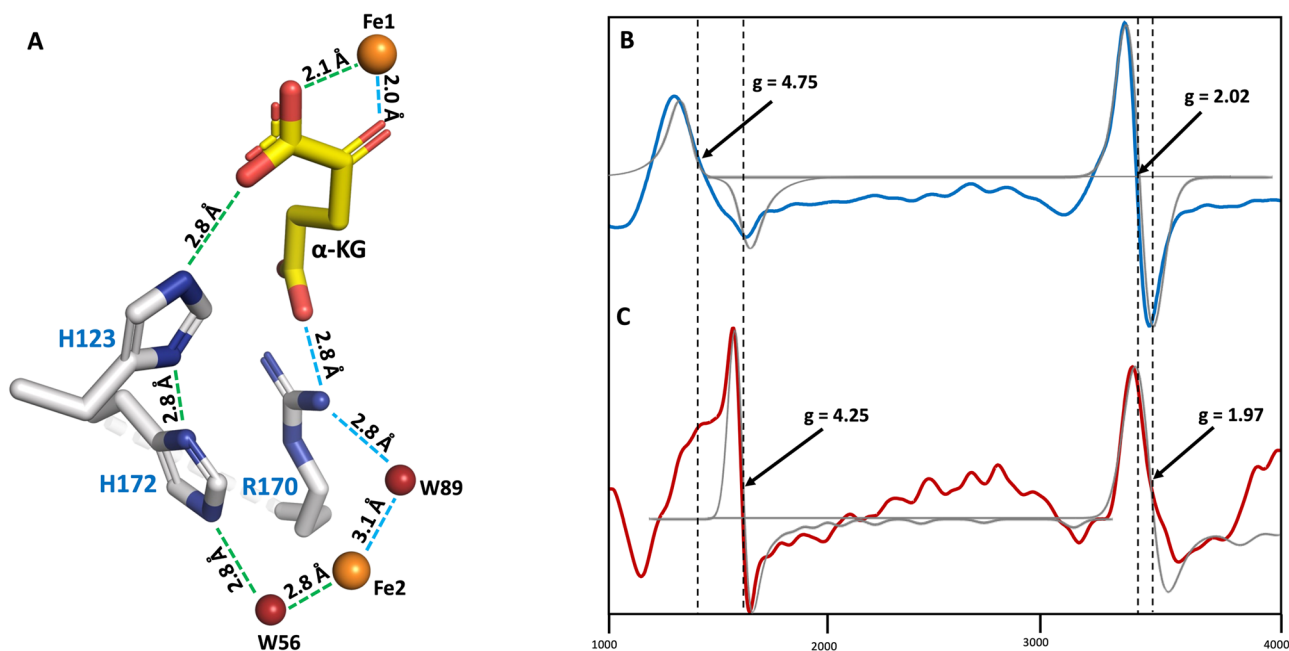


Fig. 5 Schematic representation of the CT Path in Cpz10 and EPR spectra. **A** The hidden CT conduit is schematically presented, where CT routes are proposed from Fe2 en route pass points of Wt89, His172, His123, and α KG to Fe1 or alternatively through Arg170, α KG to Fe1. **B** EPR spectrum (black line) for an enzymatic reaction of Cpz10 in a 50 mM potassium phosphate buffer solution at pH 7.5 containing 0.2 mM α KG and 0.2 mM ascorbic acid under an aerobic condition; simulated EPR spectrum (blue line). **C** EPR spectrum (black line) for an enzymatic reaction of Cpz10 in a 50 mM potassium phosphate buffer solution containing 0.2 mM α KG, 0.2 mM ascorbic acid, and 0.2 mM methionine at pH 7.5 under an aerobic condition; simulated EPR spectrum (red line).

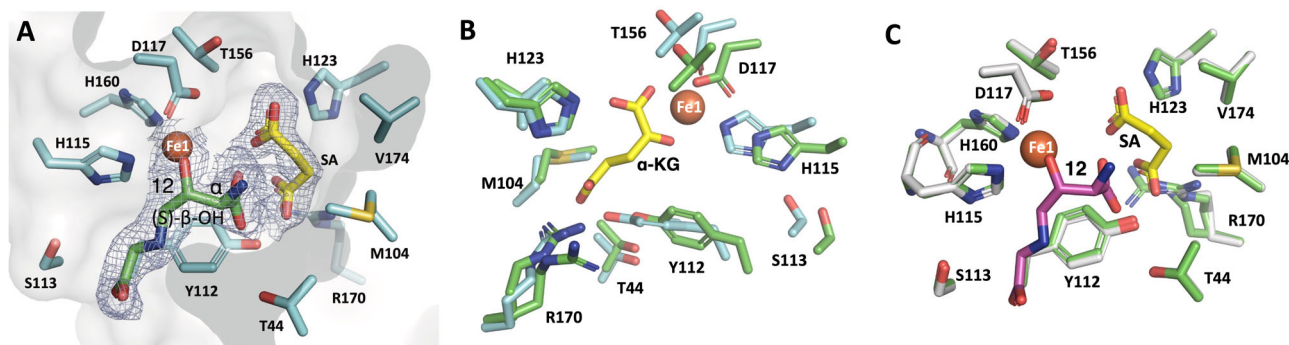


Fig. 6 Crystal structure of Cpz10 in complex with Fe1 and α -KG and 2-amino-3-hydroxybutanoic acid. **A** Closed-up view of the Cpz10 active site, where the active site Fe1 (colored orange) is coordinated with H106, H154, D108, succinic acid (SA) and 2-amino-3-hydroxybutanoic acid moiety of compound **12** (colored green). α KG has undergone decarboxylation to form SA (colored yellow); compound **11** that replaces water (W91; Fig. 4B) at the sixth coordinate has been converted to compound **12**, where the newly introduced OH group poses an *S*-configuration at β C in agreement with expected product. **B** Closed-up active-site view of Cpz10 apo (colored green) superimposed over that in complex with α -KG (colored yellow) with a r.m.s.d of 0.54, indicating Cpz10 undergoing some degree of conformational change. **C** Closed-up active-site view of Cpz10 apo (colored green) superimposed over that in complex with compound **12** (colored magenta) and SA (colored yellow) with a r.m.s.d of 0.1, indicating Cpz10 undergoing less extent of conformational change after formation of compound **12** and SA. $2F_o - F_c$ electron-density maps are contoured at 1σ (Supplementary Fig. 11 for stereo views).

formed following the rebound recombination reaction between the compound **11** radical and the iron-bound hydroxyl radical (Fig. 6)³⁰. As to the sulfoxide formation in *L*-methionine, it can be attributed to this mimetic substrate adapting an unnatural conformation as well as the preference for the soft sulfur base and soft Fe^{III}-OH acid pair (Fig. 6 and Supplementary Fig. 11)^{17,30–32}.

Nineteen mutants (primers/functions list, Supplementary Table 1A, B) of key residues referred above were made and assayed, most of which lose their enzymatic activity, indicating the importance of their roles proposed. Structurally, Tyr112 exhibits two different conformations in both apo/mutant and ternary complexes. The distances of the phenolic oxygen to the nonheme iron (Fe1) differs in 7.13 Å and 6.32 Å, where the

variation is analogous to its counterparts in isopenicillin N synthase or TauD, and is correlated with appearance or disappearance of α KG, succinic acid or n-oxalylglycine in the active site^{33,34}. Tyr112, however, seems to adapt a fixed conformation to Fe1 in the apo and mutant (R107A) structures (Supplementary Fig. 17). Two mutant structures R170A (at a resolution of 1.65 Å) and D117G (at a resolution of 1.95 Å) were determined: In the former, both irons are absent, α KG is no longer present, and Tyr112 takes a fixed conformation close to His160 (2.9 Å) (Supplementary Fig. 17). In the latter, the overall structure is similar to that of apo but retaining Fe2 in contrast to apo or R170A that contains no Fe2 (Supplementary Fig. 18), suggesting there is a dynamic interplay between Arg170 and Fe2.

It is worth noting that Fe1 is held tightly by the facial triad with short dative bonds (~ 2.2 Å), while Fe2 with two satellite water (Wt56 and Wt89) is less confined at a fixed position (Supplementary Fig. 16). This pair of irons may be spatially connected via a hidden electron/charge transfer conduit en route pass points of Wt56, His172, His123, and α KG to Fe1 or alternatively via Arg170, α KG to Fe1 (Fig. 5A) as a relay system reminiscent to that in Rieske-type oxygenases (the distances between the pass points are < 2.8 Å in contrast to > 3 Å of those in the apo form, where Fe2 is missing)^{27–29}.

EPR for Cpz10. To probe the oxidation/spin states of the two irons in Cpz10, the solid-state protein (freezing the sample in liquid nitrogen) at defined conditions was examined using electron paramagnetic resonance (EPR) spectroscopy at 77 K. The EPR spectra display two signals, which is correlated to the two irons in Cpz10. These two signals reflect the states in the presence of α KG (Fig. 5B) or in the same condition with additional ascorbic acid and *L*-methionine (Fig. 5C). In the former, the EPR spectrum shows a strong $S = 1/2$ low-spin Fe^{III} signal at *g*-value 2.02 and a relatively weak $S = 5/2$ high-spin Fe^{III} signal at *g*-value 4.75 (Fig. 5B). The presence of an axial low-spin signal indicates spin relaxation of ferromagnetic Fe^{III} (35–37) as a result of magnetic exchange, accounting for a less constrained metal ion^{24,37,38}. In contrast, the high-spin signal comparable to a typical nonheme Fe^{III} is relatively broader and less intensive with a *g*-value shifted from 4.25 to 4.75 likely due to nonheme iron electronic reconfiguration as well as coordination dynamics³⁹. This signal is relatively symmetric suggesting OH[−] ions in association to the fifth and/or sixth coordinates of the nonheme ferric iron⁴⁰. In the latter, the reduced ferrous iron (Fe^{II}) is usually insensitive to EPR ($S = 0$ or $S = 2$), while the appearance of a relatively weak signal indicates a portion of Fe^{II} have oxidized and moved to a different spatial position^{35,39,41,42}. Given that the Fe^{III} nonheme signal ($S = 3/2$) moved back to the typical *g*-value at 4.25⁴³ together with the low spin signal dwindled and shifted to 1.97, these phenomena suggest reactions are taking place and proteins are undergoing conformational changes in consistence with the reaction in the presence of substrates and reducing agents. Along with the relative distance change between the two irons (12–8 Å) and the overall protein conformational changes of Cpz10 with/without substrates (superimposition of structures with r.m.s.d. 0.54), a metal-coupled electron transfer (MCET) mechanism was proposed here, whereby electrons are transferred from the less-constrained ferrous iron to the nonheme ferric iron with assistance of an aromatic residues-relay system down the path (Supplementary movie, Fig. 5A). This reasoning is coherent with the EPR analysis where the peak shape is twisted and *g*-value is changed to *g* = 1.97 (Fig. 5C)^{27,44–47}. This spectrum is similar to the EPR spectra reported for proteins with an anti-ferromagnetically coupled Fe₂ (II/III) cluster that displays a broad axial signal with a *g*-value < 2 ($S_{\text{total}} = 1/2$)^{48,49}. Moreover, minor peaks between the two major peaks are visible as shown in Fig. 5C but undefinable, which should be attributed to formation of intermediate iron adducts such as ferric oxide, ferric hydroxy and paramagnetic ferric iron-substrate or product complexes^{37,50}.

Investigation of the nature of Fe^{III} and Fe^{II} EPR signals by NBO analysis. To better understand the charge/spin dynamics of active-site nonheme Fe^{III}/Fe^{II} during the reaction, we drew on the natural bond orbital (NBO) analysis. Fe^{III} in both intermediate-spin (IS) and high-spin (HS) states has the same natural electron configuration (NEC) as [core] 4s^{0.20} 3d^{6.33} 4p^{0.40} 4d^{0.02} 5p^{0.01}, in which the first three orbitals are considered the valance orbitals containing a total of 6.93 electrons as opposed to the last two

orbitals that are Rydberg orbitals with 0.03 electrons. The 6.33 valance electrons occupy the 3d orbitals as follows: $d_{xy}^{1.437}$ $d_{xz}^{1.230}$ $d_{yz}^{1.360}$ $d_{x^2-y^2}^{1.243}$ $d_{z^2}^{1.058}$. Furthermore, the natural charge of the Fe^{III} in both spin states is equal to +1.069 e. Since the NEC of free Fe^{III} cation is designated as [core] 4s 3d⁵ 4p with the natural charge +3.0 e, the increase of valance orbital occupancies and the significant change in the natural charge of the Fe^{III} cation bound to Cpz10 are ascribed to charge transfer (CT) interactions from six ligands, in which the 2s and 2p orbitals of oxygens and nitrogens of α KG, Wt91, His115, Asp117, and His160 coordinate with the 4s, 3d, and 4p orbitals (Supplementary Fig. 19).

During CT interactions, hybridization of valance orbitals on each cited atom is taken place in order to form the lone pair (LP) hybrid orbitals on either oxygen or nitrogen ligands as well as the lone pair* (LP*) hybrid orbitals on the Fe^{III}, including LP*(6), LP*(7), LP*(8), and LP*(9). In all of these CT interactions, electrons are donated from the LP hybrid orbitals of either the oxygen or the nitrogen ligands into nearly empty LP* hybrid orbitals of the Fe^{III}. NBO analysis indicates that not only the LP \rightarrow LP* interaction types identified in HS state are totally common to those characterized in IS state, but also the local orbital types of partner atoms in each of these interactions as well as their hybridization types and electronic occupancies are the same in both spin states (Supplementary Table 2). The energy of LP \rightarrow LP* interaction, called the second-order stabilization energy, $E^{(2)}$, is evaluated by the second-order perturbation theory according to the Eq. (1) below:

$$E^{(2)} = \Delta E_{CT} = \Delta E(n_i \rightarrow n_j^*) = -2 \frac{\langle n_i | F | n_j^* \rangle^2}{(\epsilon(n_j^*) - \epsilon(n_i))} \quad (1)$$

$\langle n_i | F | n_j^* \rangle$ is the Fock matrix element, while the denominator is the energy difference between the donor (LP) and the acceptor (LP*) orbitals.

It is evident from the results shown in Supplementary Table 2 that the second-order stabilization energy, $E^{(2)}$ (refs. 51–53), of a given CT interaction in IS state is only slightly different from that in HS state. On the other hand, the amount of the transferred charge, q_{CT} (refs. 52), between the interacting local orbitals in each common CT interaction of these two spin states are nearly the same (or the same). The stabilization energy of the LP \rightarrow LP* interaction is an appropriate criterion to evaluate the magnitude of the interaction strength between the LP* of Fe^{III} and the LP of its partner atom, approximately equal to $E^{(2)}$ values of these common CT interactions signifying their almost identical strengths in both spin states. Therefore, on the basis of NBO results, we conclude that the nature of the weak EPR signal detected at the *g* value of 4.75 *g* (Fig. 5B) can be derived from both the intermediate-spin state ($S = 3/2$) and the high-spin state ($S = 5/2$) of the Fe^{III} bound to the Cpz10 active site.

NBO analysis shows that LP (1) orbitals on N ϵ 2 atoms in His115 and His160 overlap with LP*(9) and LP*(8) of Fe^{III}, respectively. The result of these orbital overlaps is the formation of two bonds of N ϵ 2–Fe type in pairs of His115–Fe and His160–Fe (Supplementary Fig. 19). As there is only a slight difference between $E^{(2)}$ values of these CT interactions, the interaction strengths of the Fe^{III} with these two His are the same (Supplementary Table 2). As displayed in Supplementary Fig. 19, the Fe^{III} forms four bonds of O–Fe kind of with the oxygens of Asp117, α KG, and Wt91. In all cases, electrons of both LP (1) and LP (2) orbitals on oxygens are transferred to the LP* orbitals on the Fe^{III}. As can be seen from Supplementary Table 2, the interactions between the LP (2) and LP* orbitals are much stronger than those between the LP (1) and LP* orbitals because

$E^{(2)}$ values of the LP (2) \rightarrow LP* interactions are considerably greater than those of the LP (1) \rightarrow LP* interactions. Among them, the LP (2) O4 \rightarrow LP* (7) Fe^{III} interaction in α KG–Fe pair is the strongest CT interaction because it is having the highest values of $E^{(2)}$ (115.90 kJ mol⁻¹) and qCT (0.1116 e) compared to other interactions. In addition to O4, LP* (6) and LP* (8) of the Fe^{III} also accept electrons from LP (1) and LP (2) orbitals of its O2 atom. Of these, the LP (2) O2 \rightarrow LP* (6) Fe^{III} with an $E^{(2)}$ of 73.81 kJ mol⁻¹ and a qCT of 0.0624 e is the strongest CT interaction compared to other interactions of O2–Fe. Moreover, the LP (2) O δ 2 \rightarrow LP* (6) Fe^{III} in Asp117–Fe pair is its second strong interaction out of the surrounding atoms (Supplementary Table 2). Indeed, these three aforementioned CT interactions play essential roles in electron transfer process within Cpz10. In the case of interaction Fe^{III} with Wt91, the strongest CT belongs to the LP (2) O \rightarrow LP* (9) Fe^{III} interaction with an $E^{(2)}$ of 33.14 kJ mol⁻¹ and a qCT of 0.0273 e.

In order to emerge the EPR signal from the Fe^{II} cation, the number of unpaired electron(s) in its 3d orbitals in different spin states must be odd. As known, the NEC of a free Fe^{II} cation is as [core] 4s 3d⁶ 4p. Therefore, the free ferrous iron lacks unpaired electron in its 3d orbitals in the low-spin state ($S=0$), whereas these orbitals possess two and four unpaired electrons in the intermediate-spin state ($S=1$) and the high-spin state ($S=2$), respectively. It is hence expected that the free Fe^{II} is unable to generate the EPR signal. Nevertheless, the appearance of the EPR signal at g value of 2.02 signifies the presence of Fe^{II} as a ferrous iron. NBO analysis of the Fe^{II}–binding site within Cpz10 demonstrates that the Fe^{II} cation has different NECs in three states of LS ([core] 4s^{0.26} 3d^{6.99} 4p^{0.13} 5s^{0.01} 4d^{0.01} 5p^{0.02}), IS ([core] 4s^{0.30} 3d^{6.99} 4p^{0.13} 5s^{0.01} 5p^{0.02}), and HS ([core] 4s^{1.30} 3d^{6.05} 4p^{0.13} 4d^{0.01} 5p^{0.03}). As a result, due to the participation of Fe^{II} in CT interactions with surrounding atoms (Supplementary Fig. 20), it has an extra unpaired electron in its 3d orbitals in both LS and IS states, while the accepted electrons are mainly transferred into its 4s orbital in HS state. As a consequence, this extra unpaired electron in the 3d orbitals of Fe^{II}, which is observed in both LS and IS states, is the reason of the appearance of the EPR signal at 2.02 g . Moreover, due to the transfer of electrons from hybrid orbitals of the neighboring atoms into LP and LP* hybrid orbitals, NBO charges on the Fe^{II} cation in LS and IS states are 0.586 e and 0.542 e, respectively.

In both spin states, NBO analysis detects three CT interactions of the LP O \rightarrow LP* Fe type in pairs of Ala132–Fe, Wt56–Fe, and Wt89–Fe as well as the LP C δ 1 \rightarrow LP* Fe interaction in Ile145–Fe pair (Supplementary Table 3). Among them, the LP (2) O Wt89 \rightarrow LP* (4) Fe^{II} interaction with an $E^{(2)}$ of 13.22 kJ mol⁻¹ and a qCT of 0.0302 e is the strongest CT interaction inside the Fe^{II}–binding site of Cpz10. The LP (1) O Ala132 \rightarrow LP* (6) Fe^{II} is the second strong CT interaction of Fe^{II}. Thus, these two interactions play important roles in the electron transfer process within the Fe^{II}–binding site. It is evident (Supplementary Data 3) that the Fe^{III} not only involves in much more CT interactions than the Fe^{II}, but also it interacts more strongly with its neighboring atoms within the Cpz10 active site.

In vivo study using the *cpz10*-deletion mutant. To validate the function of Cpz10 in vivo, an in-frame *cpz10*-deletion mutant was prepared. The 34.7 kb *cpz* biosynthetic gene cluster (BGC) was cloned into pMKBAC02 to give pMKBAC02-CPZ-Int (Supplementary Fig. 21). In vitro CRISPR/Cas9 was performed to edit pMKBAC02-CPZ-Int to overcome the high GC content of *Streptomyces albus* J1074::ErmE*–crpsc⁵⁴. Both pMKBAC02-CPZ-Int and pMKBAC02-Cpz-Int Δ *cpz10* were individually integrated into the genome of *Streptomyces albus* J1074 for

heterologous expression of Cpz or Cpz Δ *cpz10* (Supplementary Fig. 22)⁵⁵. Liquid chromatography mass spectroscopy (LC–MS) analysis identified three Cpz aglycones respectively with molecular weights at m/z 930, 944, and 958 in consistence with previous reports (Supplementary Fig. 23)¹⁴. The caprazamycin production was totally abolished in the Δ *cpz10* mutant in contrast to that in Cpz-containing species where caprazamycin is produced. Δ *cpz10* was cultured in a scale of 30 liters, from which only was one compound (2 mg) purified (Supplementary Scheme 1) (Supplementary Fig. 24 and Supplementary Note 1)^{4,12,14,20,56}. This compound was subjected to MS and NMR analysis, confirming that it is an ADR–GlyU disaccharide species (compound **13**, where both C6' amine and C7' carboxyl groups are methylated) instead of the expected compound **11**. The chemical structure resolved is unexpected, while it is very informative pertaining to the cyclization of the seven-membered diazepanone ring. Our reasoning is that both the C6' amine and the C7' amide group of the diazepanone ring is methylated by the action of the same N-methyltransferase (Cpz11 or 26) while the methylation at the C7' amide group takes place after the ring closure (envisioning that both secondary α -amines have a similar spatial geometry and the ring is rotatable at C5' allowing molecular recognition by the same enzyme). The C7' carboxyl group, on the other hand, is methylated by the action of the O-methyltransferase (the remaining Cpz11 or 26), of which the resulting methoxyl group turns out to be a good leaving group for the subsequent ring cyclization to take place. This process differs from the canonical amino acid adenylation by reacting with ATP and the β -hydroxyl group appears to be critical for the ring to cyclize. The isolation of compound **13** further suggests that the proposed ribosyltransferase Cpz17 can transfer ADR onto compounds **11** and **12**, while the confirmation is beyond the scope of this study. Next, Cpz10 was examined against compound **13**, which, however, cannot be consumed by Cpz10, suggesting that it is not a substrate of Cpz10. Next, Cpz10 was examined against compound **14** (produced by Cpz27, see below), which was proposed the substrate in the biosynthesis of muraymycin and polyoxin. Again, we did not find any new product^{4,20}, concluding that both compounds **13** and **14** are not the substrates of Cpz10 in contradiction to the biosynthetic pathways proposed previously^{4,20} (Supplementary Scheme 2).

Cpz27 is responsible for phosphorylation of compound 13. In silico analysis, the *cpz* BGC holds two genes *cpz12* and *cpz27* both coding for kinase-like proteins (Supplementary Fig. 25). *cpz27* contains a walker ATP motif (GXXGXGKS/T) in opposite to *cpz12* that does not; *cpz27* also shows 24% protein sequence identity to tunicamycin-resistance protein (TmrD)⁵⁷. To elucidate their biochemical roles, we cloned and expressed both genes in *Streptomyces lividans* TK64 given as *E. coli* was not an appropriate host for these two proteins (Supplementary Fig. 26). Both gene products were purified from *Streptomyces* cultures and examined against compound **13** or tunicamycin. None of them accepts tunicamycin, while can only compound **13** be phosphorylated by Cpz27 to form compound **14** at the expense of one molecule of ATP (Fig. 2D, Fig. 3E and Supplementary Fig. 27). However, we could not isolate compound **14** from the culture of Δ *cpz10* likely due to low yield and/or high instability. As mention above, compounds **13** and **14** are not substrates of Cpz10 at odds with previous reports^{4,20}, where muraymycin, capuramycin and caprazamycin are phosphorylated at the corresponding position to abolish their antimicrobial activity (as a mechanism of self-resistance)^{5,12,58}. In the present study, this likelihood is low (see below) but still cannot be ruled out. Alternatively, the 3''-OH phosphorylation as a reaction switch was proposed in control of

the reaction order^{4,20,59}. To know if the phosphorylation is a reaction switch (it takes place before ring cyclization or at the late-stage products), we purified caprazamycins and prepared compound **8** from the pMKBAC02-CPZ-Int containing *Streptomyces*. Biochemical examination revealed that neither is the substrate of Cpz27, confirming that the phosphorylation takes place before the heterocyclic ring cyclization. As a result, Cpz27 is a phosphotransferase controlling the reaction order but is irrelevant to 7-membered ring cyclization. On the other hand, Mur23 a pyridoxal phosphate (PLP) dependent decarboxylate (Supplementary Fig. 28) was reported catalyzing decarboxylation of 3''-phospho-6'-*N*-2-aminobutanoyl-ADR-GlyU²⁰, where the 3''-phosphate is considered essential as Mur23 cannot react without it (Supplementary Scheme 3). This notion does not agree with our result as Mur23 is capable of decarboxylating compound **11** (without the 3''-phosphate group) to compound **15** (molecular ion (M + H)⁺ at *m/z* = 375.05 by 44) (Figs. 2C and 3F). Additionally, Mur23 was first assayed in the presence of compound **13**, then it was assayed in conjunction with Cpz27 against compound **14**. Whichever the condition there is no new peak emerged, suggesting compounds **13** and **14** are not the substrate of Mur23 (Fig. 2C). Taken together, the decarboxylation is independent of phosphorylation, which likely takes place after β -hydroxylation, methylation but prior to ring cyclization (Supplementary Fig. 29). Given the existence of muraymycin D4 (ref. ⁶⁰), one should agree that the decarboxylation of compound **11** is independent of the ADR moiety. Mur19 (the ADR transferase), however, cannot transform muraymycin D4 (Supplementary Fig. 1) to muraymycin D2 as well as compound **15** to ADR-**15** in the presence of UDP-ADR, suggesting that ribosylation takes place prior to decarboxylation. Plus, one previous report showed that Mur23 is incapable of decarboxylating compound **16** (without the 3''-phosphoryl group; Supplementary Fig. 1), indicating both ADR and 3''-phosphate groups are critical²⁰. These results somehow conflict each other, not least at odds with our result. This inconsistency is indeed not yet reconciled but awaits further clarification in the future.

Discussion

The present study elucidated the biochemical roles of selected enzymes involved in the biosynthesis of caprazamycin, particularly two nonheme α KG-dependent enzymes Cpz10 and Cpz15 (Supplementary Fig. 30). Cpz15 that was previously assigned as a β -hydroxylase turns out to be the starter enzyme converting uridine mono phosphate (UMP) to uridine 5' aldehyde (U5'A 9). In contrast, Cpz10 is the actual β -hydroxylase in a position to hydroxylate synthetic compound **11** to compound **12** in vitro, standing at a pivotal point to lead the biosynthesis toward caprazamycin. The importance of Cpz10 was parallelly examined in vivo using the in-frame gene-deletion mutant Δ *cpz10*, by which the major product was determined to be compound **13** agreeing with the in vitro result that compound **13** lacks the β -OH group. Compound **13** carries two unexpected methyl groups at 6'-amino and 7'-carboxyl in addition to one ADR at 3'-OH. None of compounds **13** and **14** was accepted by Cpz10, suggesting that β -hydroxylation takes place earlier than methylation, glycosylation and phosphorylation. Compound **11** is likely the physiological substrate for Cpz10; beyond that, Cpz10 oxygenates *L*-methionine to MSO as a sideline activity. MSO, a nonproteinogenic amino acid exhibits in *N*-acetylmureidomycin E, may be useful as a building unit concerning combinatorial biosynthesis as an approach to generate unnatural natural products (Supplementary Fig. 1)^{61,62}. In light of β -OH that bridges a long aliphatic side chain and permits the linear chain to form the seven-membered

heterocyclic ring, the Cpz10-mediated β -hydroxylation appears to be the determining step in the biosynthesis toward the successful production of caprazamycin. Concerning compound **14**, Cpz27 is the phosphotransferase to phosphorylate compound **13** at its 3''-OH with a 3''-phosphate group (compound **14**). Based on protein blasting, there are 29 BGCs for the class of nucleoside antibiotics reported thus far, while only do caprazamycin, A-90289 and liposidomycins contain two putative kinases⁴. Cpz27 out of the two putative kinases is the one directing the formation of compound **14**, while it conflicts with 3''-phospho-caprazol that was previously reported to be the product. In contrast, Mur23 is determined to decarboxylate compound **11** that also contradicts to the report of Cui, Z. et al., in which they specified that Mur23 can only accept the substrate with 3''-phosphate^{20,60}. Given caprazamycin (**2**) and muraymycin (**6**), one can appreciate that compound **11** is a watershed, of which β -hydroxylation catalyzed by Cpz10 permits the heterocyclic ring formation and subsequent acylation to a fully-active caprazamycin or decarboxylation catalyzed by Mur23 otherwise leads the synthesis to muraymycin (Fig. 7). To this end, we conclude that the β -hydroxylation of compound **11** takes place prior to 3'-glycosylation, 6'- and 7'-methylation, 3''-phosphorylation and the heterocyclic ring cyclization, and the formation of caprazol **8**.

X-ray crystal structures revealed that Cpz10 contains two irons (Fe1 and Fe2) in 12 Å away from each other. Fe1 sits at the active center held by the His115, His160 and Asp117 triad together with α KG and a water molecule (Wt91; at the sixth coordinate) to form a tetragonal bipyramidal coordination, a typical geometry in the nonheme α KG-dependent protein family. Compound **11** is placed above the sixth coordinate with its pro-S H atom of β -carbon pointing toward the sixth coordinate, thus accounting for the regioselectivity and stereoselectivity in the Cpz10-mediated reactions. Nonheme α KG-dependent enzymes that contain two irons are not uncommon but often ignored^{42,63–65}. The second iron (Fe2) does not appear in random but exists in a deliberately organized conduit ideally for electron/charge transfer. As a matter of fact, Fe2 is not confined in a fixed position but dynamically interacts with a constellation of residues in the tunnel; this observation is line with the NBO calculation. On top of that, EPR spectra reflect signal changes for both Fe2 (2.02 \rightarrow 1.97 g) and Fe1 (4.75 \rightarrow 4.25 g) in the absence or presence of substrates in the Cpz10-mediated reaction, echoing dynamic changes in electron/charge flow, oxidation status, coordination geometry, etc.⁴².

From the viewpoint of crystal field theory (CFT)^{66,67}, the five-degenerate d-orbitals in Fe^{II} and Fe^{III} are reorganized by the static electric field arising from the distribution of charges upon association with the coordinate residues, α KG, and dioxygen/water molecules. In the case of Fe^{III} with octahedral molecular geometry, the crystal field splits its five d-orbitals into two sets with different energy levels. The lower set, which is a collection of the three lower-energy orbitals of the d_{xy} , d_{xz} , and d_{yz} , is referred to as t_{2g} . The upper set, which consists of the two higher-energy orbitals of the $d_{x^2-y^2}$ and d_z^2 , is referred to as e_g (refs. ^{68,69}). From the quantum-mechanical point of view⁷⁰, five d-electrons in 3d orbitals of the Fe^{III} are distributed in three different spin states, including low-spin (LS) state ($S = 1/2$), intermediate-spin (IS) state ($S = 3/2$), and high-spin (HS) state ($S = 5/2$), in these t_{2g} and e_g orbitals. The spin states of 1/2, 3/2, and 5/2 are assigned to electron-spin multiplicities, $2S + 1$, of doublet, quartet, and sextet, respectively. The Fe^{II} has tetrahedral molecular geometry in which the splitting of 3d orbitals is the opposite of that in octahedral geometry, namely, three higher-energy orbitals of the d_{xy} , d_{xz} , and d_{yz} are located above two lower-energy orbitals of the $d_{x^2-y^2}$ and d_z^2 (refs. ^{68,69}). The occupancies of the six d-electrons of Fe^{II} in the e_g and t_{2g} orbitals are as three spin states of 0 (LS), 1

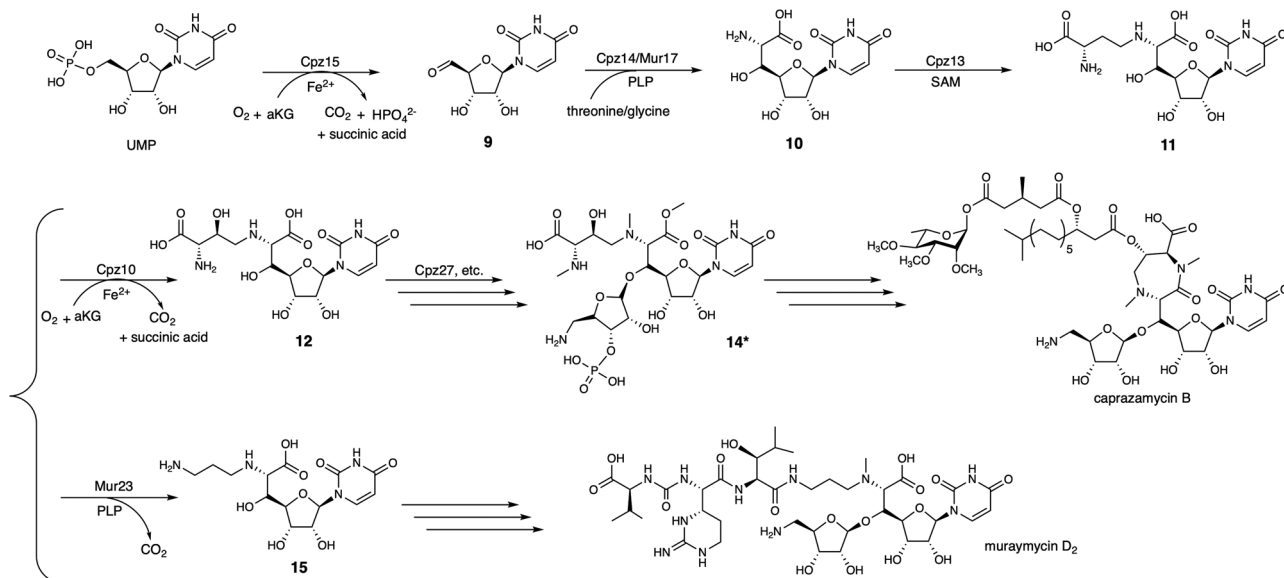


Fig. 7 Biosynthesis schemes of caprazamycin and muraymycin. **A** Compound **11** stands at a watershed, of which β -hydroxylation catalyzed by Cpz10 permits seven-membered diazepanone ring formation and subsequent acylation to a fully-active caprazamycin or decarboxylation catalyzed by Mur23 directs the synthesis otherwise toward muraymycin. Compound **14*** is compound **14** with a OH group at its β C position.

(IS), 2 (HS) attributing to multiplicities of singlet, triplet, and quintet, respectively⁷⁰. Indeed, on the basis of NBO results for Fe1 and Fe2, we conclude that another possible role of two iron cations is playing an essential role in an electron transfer process within the Cpz10. In this study, both complex structures and EPR examination have demonstrated that the second iron plays significant role to replace the succinic acid to new α KG. This new phenomenon not only controlling the reaction inhibitors but also accelerating the reaction. An important feature of iron that makes it appealing for biological redox processes is the easy one-electron interconversion of Fe^{II} and Fe^{III} (ref. ⁷¹). This discovery in the caprazamycin biosynthetic pathway has important implications for our comprehensive understanding of reaction mechanisms of nonheme metal iron-dependent enzymes, which is in support of our previous quantum chemistry analysis where key components of caprazamycin contribute significantly to the high binding affinity with MraY⁵.

In summary, the biosynthesis of nucleoside antibiotics has been studied to some extent while it is still far from completion^{8,72,73}. The present study concludes that the β -hydroxylation catalyzed by Cpz10, a nonheme dioxygenase, plays a critical role governing the subsequent heterocyclic ring formation as well as the long aliphatic sidechain acylation toward full-blown bioactivity of caprazamycin. By contrast, the decarboxylation of the 6'-N-alkylamine side chain catalyzed by Mur23, a PLP dependent decarboxylase, diverges the biosynthesis toward the Park's nucleotide subgroup. In conjunction with biochemical determination of other related enzymes, the *in vitro*, *in vivo* and biochemical/biophysical profiling has not only met our scientific interest with respect to the biosynthesis of caprazamycin but also paved a foundation for pathway engineering toward more effective nucleoside antibiotics in a way to help conquer the most formidable adversary multidrug-resistant pathogens that man is facing now.

Experimental section

General chemicals, reagents and analytical methods. Those chemicals are commercially unavailable were chemically (see Supplementary methods) and enzymatically synthesized, but most of the reagents were purchased from commercial sources

without further purification. NMR spectra were collected by the Bruker AV600-MHz spectrometer, a two-channel system equipped with 5 mm DCI Dual cryoprobe. NMR spectra were processed and analyzed with SpinWorks v.3.0. High-performance liquid chromatography (HPLC) analysis was conducted on an Agilent 1200 series HPLC system consisting of G1311A pump, G1316A column oven, G1367B autosampler and G1315D diode array equipped with an ODS Prodigy column (Phenomenex) or an Xbridge Amide column (5 μ m, 4.6 \times 250 mm, Waters). High resolution and high mass accuracy experiments were done on an LTQ Orbitrap XL ETD mass spectrometer (Thermo Fisher Scientific, San Jose, CA) equipped with a standard ESI ion source. 5 μ L of samples were flow injected at a rate of 50 μ L min⁻¹ in 80% ACN/H₂O with 0.1% formic acid to the Waters Acquity UPLC (Waters, Milford, MA). Full-scan MS condition is the mass range at *m/z* 200–2000, the resolution 60,000 at *m/z* 400. Electrospray voltage was maintained at 4 kV and capillary temperature was set at 275 $^{\circ}$ C. For LC-ESI-MS/MS, samples were detected by LC-ESI-MS on an Orbitrap Fusion mass spectrometer (Thermo Fisher Scientific, San Jose, CA) equipped with EASY-nLC 1200 system (Thermo, San Jose, CA, US) and EASY-spray source (Thermo, San Jose, CA, US). The digestion solution was injected (5 μ L) at 1 μ L min⁻¹ flow rate to an easy column (C₁₈, 0.075 mm \times 150 mm, ID 3 μ m; Thermo Scientific). Chromatographic separation was using 0.1% formic acid in water as mobile phase A and 0.1% formic acid in 80% acetonitrile as mobile phase B operated at 300 nL min⁻¹ flow rate. Briefly, the gradient employed was 2% buffer B at 2 min to 40% buffer B at 40 min. Full-scan MS condition: mass range *m/z* 375–1800 (AGC target 5E5) with lock mass, resolution 60,000 at *m/z* 200, and maximum injection time of 50 ms. Target *m/z* were isolated for CID with NCE35 and maximum injection time of 100 ms. Electrospray voltage was maintained at 1.8 kV and capillary temperature was set at 275 $^{\circ}$ C. High resolution mass spectrometry analyses were performed by Mass Core Facility of GRC, AS. Size exclusion column chromatography analyses were conducted on an AKTA purifier (GE Healthcare Life Sciences) with a Superdex 200 16/600 GL column (GE Healthcare Life Sciences, UPC-900 and P-920). pGUSRoIRPA3 was provided as a gift by the laboratory of Prof. Andriy Luzhetskyy (Helmholtz Institute for Pharmaceutical

Research Saarland, Saarland University). *E. coli* ET12567[pUZ8002] was provided as a gift by the laboratory of Prof. Chih-Hung Huang (Laboratory of Bacterial Genetic, National Taipei University of Technology, Taipei, Taiwan).

Bacteria strains and plasmids used in this study. *E. coli* strains were grown in Luria-Bertani (LB) medium supplemented with either apramycin (50 $\mu\text{g mL}^{-1}$), kanamycin (35 $\mu\text{g mL}^{-1}$) or chloramphenicol (35 $\mu\text{g mL}^{-1}$). The caprazamycin producing strain and the host strain for heterologous expression, *Streptomyces* sp. MK730-62F2 and *Streptomyces albus* J1074::ErME*-*crpsc*, were grown at 28 °C with MS medium agar (2% soy flour, 2% mannitol, 2% agar). apramycin (50 $\mu\text{g mL}^{-1}$) and nalidixic acid (25 $\mu\text{g mL}^{-1}$) were supplemented for conjugation. BAC clones with large inserts were isolated using the ZR BAC DNA Miniprep Kit (ZYMO research). Transformation of BAC plasmid into *E. coli* was performed using 10 G BAC-Optimized Electro-competent Cells (Lucigen).

Construction of pMKBAC02. The origin of transfer region (*oriT*) and apramycin resistance gene, *aacIII(IV)*, were amplified from plasmid pGUSRolRPA3⁷⁴ using the primer set 325-oriT-Am-F (5'-ATTACCACCTTAATTAAGGATCCGAATTCGAAGATCCTTTGATCTTTTC') and 326-oriT-Am-R (5'-ATTATTATTCCTAGGAGCTTGCATGCCTGCAGGTCGA-3'). The backbone of the pMKBAC02 vector was amplified from plasmid pBeloBAC11 (New England Biolabs, NEB) using primer set p327-BAC-F (5'-ATTATTATTCCTAGGGTTAAACAGGGCTTCCCGGTATCAAC-3') and p328-BAC-R (5'-ATTACCACCTTAA TAAAAGCTTGGTTACTCCGTTCTACAGGTTAC-3'). The 2.1 kb PCR product amplified from the first primer set and the 5.4 kb PCR product amplified from the second primer set were digested within AvrII and PacI individually and then ligated together to produce plasmid pMKBAC02.

Isolation of entire caprazamycin aglycone biosynthetic gene cluster into pMKBAC02. To isolate the CPZ BGC from the chromosome of *Streptomyces* sp. MK730-62F2, a homologous DNA fragment including a portion of CPZ amplified within 537-pMKBAC02-cpz-F (5'-ATTATTATTAAGCTTGTACACACAGCACTACCAGCACCCTTCG) and 538-pMKBAC02-cpz-R (5'-ATTATTATTGAATTCTCGTCCACGGTCGTCAGGAACAGC TGTA-3') was cloned into pMKBAC02 within HindIII and EcoRI restriction enzyme sites respectively to generate pMKBAC02-CpzH. Conjugation was later performed to integrate pMKBAC02-CpzH into the chromosomal DNA of *Streptomyces* sp. MK730-62F2 via homologous recombination as. Next, the qualified mutants were carefully chosen on apramycin containing MS agar⁷⁵, verified using PCR within primer sets, p574-check-genome-BAC-cpz-F (5'-TGCACATGAACCAAAGGATCTAGGTGAAGATCCTTTTTGATAA-3') and p575-check-genome-BAC-cpz-R (5'-GAGTCGTCGTCAGTTGCTCTTCGGACA-3') (Supplementary Fig. 22). This strain was cultured in TSB media for 1 day at 30 °C, after which its genomic DNA was prepared and then digested by restriction enzyme HindIII. The digested DNA fragments were purified and concentrated by ethanol precipitation before self-ligation using T4 ligase (Thermo; Fisher Scientific). The ligation mixture was used for electroporation of 10 G BAC-Optimized Electro-competent Cells. Recombinants were selected on apramycin-supplying LB medium gar, after which plasmids were verified by PCR using primers, P238-check-pMKBAC02-F (5'-TAATGTCCTTTGTTACAGGCCAGAAAGCATAA-3') and P239-check-cpz4 (5'-ATGAACGCCTCCTTCAAGGATCCGTT-3'), the BAC plasmid carrying CPZ aglycone BGC was identified as pMKBAC02-CPZ (Supplementary Fig. 22). The DNA fragment

containing attP-int Φ C31 moiety was amplified from plasmid pGUSRolRPA3 within primer sets, 329-AvrII-attP-int-phiC31-F (5'-ATTATTATTCCTAGGCTAGCGATTCCAGACGTCCCGAAG-3') and 330-AvrII-attP-int-phiC31-R (5'-ATTATTATTCCTAGGAAATCCCCAATGTCAAGCACTTCCGG-3'). The amplicon was digested and cloned into pMKBAC02-CPZ by AvrII restriction enzyme site for generating pMKBAC02-CPZ-Int.

Generation of cpz10-deficient mutants via in vitro CRISPR-Cas9 digestion. In vitro CRISPR-Cas9 digestion was performed as previous described⁷⁶. Briefly, the gene sequence to be edited was screened for the presence of a 20-bp guide and PAM sequence (5'-N20NGG-3') for sgRNA design. Transcription templates of the sgRNA were amplified by overlap extension PCR with the primers, 893-sgRNA-Dcpz10-1 (5'-GATCACTAATACGACTACTATAGTCTTTCCAACAACCTCCACGAGTTTTAGAGCTAGAAAATAGCAA-3'), 894-sgRNA-Dcpz10-2 (5'-GATCACTAATACGACTACTATAGCGTTCAAGCCCCATTTCGCACGTTTTAGAGCTAGAAAATAGCAA-3'), 229-sgRNA-F (5'-GTTTTAGAGCTAGAAAATAGCAAAGTTAAAATAAGGCTAGTC-3') and 230-sgRNA-R (5'-AAAAGCACCAGCTCGGTGCCACTTTTCAAGTTGATAACGGACTAGCCTTATTTTAACT-3').

In vitro transcription of sgRNA was performed using the TranscriptAid. T7 high-yield transcription kit (Thermo; Fisher Scientific) pursuant to the manufacturer's instructions. The sgRNA product was pretreated by heating at 95 °C for 10 min and slowly cooling down to room temperature.

In vitro Cas9-mediated editing was performed following the protocol of Cas9 nuclease, *S. pyogenes* (NEB). First, a 30 μL reaction mixture containing 30 nM Cas9 protein, 30 nM sgRNA and 3 nM DNA was prepared and incubated in buffer 3.1 (New England Biolabs) at 37 °C for 2 h.

The reaction was terminated by addition of 0.2 mg mL⁻¹ RNase and continued incubation at 37 °C for 15 min. Then the reaction mixture was treated with 1 mg mL⁻¹ proteinase K (Thermo; Fisher Scientific) and incubated at 37 °C for 30 min. Finally, the Cas9-digested DNA was recovered by ethanol precipitation and self-ligated in the ligation mixture and then the mixture was incubated at 16 °C for overnight.

The ligation mixture was used for electroporation of 10 G BAC-Optimized Electrocompetent Cells. Gene-deficient mutants were selected on apramycin-adding medium gar, after that the plasmids were verified by PCR using primers, 895-check-Dcpz10-F (5'-CTCATGACGCGGATCCGAGGATTCAT-3') and 896-check-Dcpz10-R (5'-AATCCTGGCAGGTCTTTCCCGCTT-3') (Supplementary Fig. 22).

Cloning for gene expression. All genes in caprazamycin and muramycin gene clusters were cloned from the genomic DNA of producing strains by applying to standard procedures (primers listed in Supplementary Table 1A, B). The PCR products for *cpz15* and *cpz10* were digested with EcoRI-HinDIII and ligated to the appropriate sites of pET28a to yield pET28a-*cpz15* and pET28a-*cpz10*, respectively. PCR results were confirmed by DNA sequencing. The PCR products for *cpz27* and *cpz12* were digested with NdeI-HinDIII, sequenced and inserted into a pET28a (pET28a-*cpz27* and pET28a-*cpz12*) vector by using ligation-independent cloning, following the commercial protocol. The PCR products for *mur16*, *mur17* and *mur23* were digested with NdeI-XhoI and ligated to the appropriate sites of pET28a to yield pET28a-*mur16*, pET28a-*mur17* and pET28a-*mur23*, respectively. Moreover, *cpz27* was cloned to PGM1202 plasmid to yield PGM1202-*cpz27* by using AseI-HinDIII enzymes.

Site-directed mutagenesis. R170A, D117G, D117A, M104A, M104L, D72A, D72V, R102A, R102V, R121A, and R121V point mutations of *cpz10* and a R259A point mutation of *cpz15* were generated by PCR amplification with high-fidelity, hot start, DNA polymerase using pET28-*cpz10* or pET28-*cpz15* as a template, respectively. The templates pET28a-*cpz10* and pET28a-*cpz15* were obtained by using ligation-independent cloning following the commercial kit protocol. The PCR products were digested with 10 units DpnI for overnight at room temperature and transformed into *E. coli* DH5 α -competent, then BL21-competent cells. The entire gene sequences, including of the correct point mutation sites, were confirmed by sequencing.

Recombinant protein production. Plasmids pGM1202-*cpz27* was transformed into *S. lividans* TK64 by applying PEG (poly ethylene glycol)-mediated protoplast transformation and plated on R5 medium containing 50 $\mu\text{g mL}^{-1}$ of apramycin. After the incubation at 28 °C for 6 days, successful transformants were affirmed by colony PCR. Successfully transformed strains were applied to inoculate 50 ml YEME medium containing 50 $\mu\text{g mL}^{-1}$ of apramycin, incubated for 3 days at 28 °C and then 4 mL was transferred to 400 mL fresh YEME medium including 50 $\mu\text{g mL}^{-1}$ of apramycin. After 36 h incubation at 28 °C, protein (Cpz27) expression was induced by the amount of thioestrepton (5 $\mu\text{g mL}^{-1}$) and before harvesting for more 36 h the culture was incubated. The cells from 400 ml of the culture were harvested by centrifugation. The pellet was completely re-suspended in ice-cold buffer D (100 mM Tris-HCl, 300 mM KCl, 50 mM arginine, 1% glycerol pH 8.0), and extra 4 mg mL^{-1} of lysozyme was later added to the suspension. After incubation at 30 °C for 30 min, the cell suspension was lysed and micro-fluidized at 150 MPa in a Microfluidizer processor model M-110 (Quadro Engineering Corp.) or sonicated with Misonix sonicator for 15 min. After centrifugation, the protein was purified by applying the affinity chromatography with IMAC sepharose 6 fast flow resin, and proteins were eluted with high concentrations of imidazole in buffer B (100 mM HEPES, 300 mM KCl, 50 mM arginine, 1% Glycerol, 400 mM imidazole, pH 8.0). Purified proteins were concentrated and buffer exchanged into buffer C (17 mM KH_2PO_4 , 33 mM K_2HPO_4 , pH 8.0) by using Amicon Ultra 10 K MWCO centrifugal filters and applying fresh for storage at -80 °C. Protein purity was assessed using sodium dodecylsulfate (SDS)-15 and 10% poly acrylamide gel electrophoresis; His₆-tagged proteins were used without further modifications⁷⁷.

All the plasmids; pET28a-*cpz10*, pET28a-*cpz10* (all mutants), pET28a-*cpz15*, pET28a-*cpz15* (R259A), pET28a-*cpz12*, pET28a-*mur16*, pET28a-*mur17* and pET28a-*mur23* were introduced into BL21(DE3), and the transformed strains were grown in LB supplemented with 50 $\mu\text{g mL}^{-1}$ of kanamycin. After inoculation of 30 ml lysogeny broth with 3 $\mu\text{g mL}^{-1}$ of kanamycin, the cultures were grown at 37 °C until the cell density reached an $A_{600} \approx 0.7$, when expression was induced with 0.3 or 0.5 mM isopropyl β -D-1-thiogalactopyranoside. Cells were harvested and processed as described for Cpz27 but with different buffers.

Reactions with Cpz15 and Mur16. Cpz15 and a mutant and Mur16 were lysed with buffer D: 100 mM Tris-HCl, 300 mM KCl, pH 8.0 and protein was eluted by with buffer E: 100 mM HEPES, 300 mM KCl, Imidazole 50 mM, pH 8.0. Reaction mixtures consisted of 100 mM HEPES (pH 8.0), 100 mM KCl, 1 mM UMP (uridine monophosphate), 1.5 mM αKG , 200 μM ascorbate, 100 μM FeSO_4 , and 100 ng Cpz15 and a mutant for 3 h at 28 °C, and the reactions were subsequently quenched by adding a volume of chloroform followed by centrifugation (15000 rpm, 30 min) to separate two phases then remove the precipitated proteins. The reactions were monitored by LC-MS with a prodigy

column (5 μm , 100 \AA , 250 mm \times 4.6 mm) and HPLC with a SOURCE™ 15Q 4.6/100 PE column for ion exchange chromatography. A series of linear gradients was developed from water with 0.1% Trifluoroacetic acid (mobile phase A) to acetonitrile with 0.1% Trifluoroacetic acid (mobile phase B) in the following manner (beginning time and ending time with linear change to percentage B): 0–6 min, 0.5% B; 10–13 min, 90% B; 15–20 min, 0.5% B. The flow rate was kept constant at 1 mL min^{-1} and elution was monitored at 260 nm.

Activity of Mur17. Reaction mixtures consisted of 50 mM potassium phosphate (pH 8.3), 2 mM U5'A, 2 mM Thr/Ser/Gly, 500 μM PLP and 200 nM Mur17 for 5 h at 30 °C, and the reaction was terminated by adding a volume of chloroform followed by centrifugation. The reaction components were analyzed by LC-MS with a prodigy column (5 μm , 100 \AA , 250 mm \times 4.6 mm) column as described in reactions with Cpz15 and Mur16.

Reaction with Cpz10. Cpz10 and all mutants were lysed with buffer D: 100 mM Tris-HCl, 300 mM KCl, pH 8.0 and protein was eluted with buffer F: 100 mM HEPES, 300 mM KCl, imidazole 400 mM, pH 7.5. Purified proteins were concentrated and buffer exchanged into buffer G (17 mM KH_2PO_4 , 33 mM K_2HPO_4 , pH 7.5) by using Amicon Ultra 10000 MWCO centrifugal filters. Reaction mixtures consisted of buffer G, 2 mM compound **9** or methionine, 1.5 mM αKG , 200 μM ascorbate, 100 μM FeSO_4 , 100 ng Cpz10 and the mutated proteins for 3 h at 28 °C, and the reactions were subsequently quenched by adding a volume of chloroform followed by centrifugation (21,000 g, 30 min) to remove the precipitated proteins. The reactions were monitored by LC-MS with a prodigy column (5 μm , 100 \AA , 250 mm \times 4.6 mm) and for HPLC with SOURCE™ 15Q 4.6/100 PE column for Ion exchange chromatography. A series of linear gradients was developed from water with 0.1% Trifluoroacetic acid (mobile phase A) to acetonitrile with 0.1% Trifluoroacetic acid (mobile phase B) in the following manner (beginning time and ending time with linear change to percentage B): 0–6 min, 0.5% B; 10–13 min, 90% B; 15–20 min, 0.5% B. The flow rate was kept constant at 1 mL min^{-1} and elution was monitored at 260 nm for compounds **11**, **12**, **13**, and for methionine and MSO was 210 nm.

Reactions with Cpz27 and Cpz12. Reaction mixtures comprised of 50 mM potassium phosphate (pH 8.0 for Cpz27 and pH 7.0 for Cpz12), 500 μM **13**, 1 mM ATP, 1 mM $\text{MgCl}_2/\text{MgSO}_4$, and 200 nM Cpz27 and Cpz12 for 3 h at 28 °C, and the reaction was later quenched by adding two volumes of acetonitrile followed by centrifugation (15000 rpm, 30 min) to remove the precipitated protein. The reaction was monitored by LC-MS or HPLC with a prodigy column (5 μm , 100 \AA , 250 mm \times 4.6 mm). A series of linear gradients was developed from water with 0.1% trifluoroacetic acid (mobile phase A) to acetonitrile with 0.1% Trifluoroacetic acid (mobile phase B) in the following: 0–10 min, 2.0% B; 10–20 min, 35% B; 24–34 min, 98% B; 35–40 min, 2% B. The flow rate was kept constant at 1 mL min^{-1} and elution was monitored at 260 nm.

Reaction with Mur23. Reaction mixtures consisted of 50 mM HEPES (pH 8.3), 1 mM **11**, 1 mM PLP and 200 nM Mur23 for 1 h at 30 °C, and the reaction was terminated by adding a volume chloroform followed by centrifugation. The reaction was monitored by LC-MS or HPLC with a prodigy column (5 μm , 100 \AA , 250 mm \times 4.6 mm). A series of linear gradients was developed from water with 0.1% trifluoroacetic acid (mobile phase A) to acetonitrile with 0.1% trifluoroacetic acid (mobile phase B) in the

following manner (beginning time and ending time with linear change to percentage B): 0–10 min, 2.0% B; 10–20 min, 35% B; 24–34 min, 98% B; 35–40 min, 2% B. The flow rate was kept constant at 1 mL min⁻¹ and elution was monitored at 260 nm.

Purification of cpz10-knockout intermediate. The compound **13** as a knockout intermediate was purified using culture condition pH (6–7), then centrifuge the culture broth for 30 min in 6000 rpm. After separating broth, its pH decrease to 4 for remain proteins aggregation. Centrifuge again same condition. Collected broth first mixed with ethyl acetate (EA) then butanol to separate hydrophobic compounds from the collected broth. Remain mix was added with Diaion HP20 resin (Sigma-Aldrich) for overnight in 4 °C, and washed the by water. The target compound was eluted out by the acetone, and dried with the rotary vapor. After rehydrating by the DD water, the intermediate was loaded to Sephadex LH20 (GE Healthcare) column for further purification. Eluate was collected, and loaded on CM Sephadex C-25 to separating sugar impurities. Collected remainder was injected to HPLC with a prodigy column (5 µm, 100 Å, 250 mm × 4.6 mm). A series of linear gradients was same as previously mentioned. The final compound was checked by HR-MS and NMR.

UV/Vis spectral analysis. Absorbance spectra of the analytes were recorded on a Beckman coulter DU 800-spectrophotometer using 30 µM indicated protein, with a range from 0.1 to 1 mM of the respective substrates (**9** and *L*-methionine) at room temperature in buffer G supplemented with 100 µM FeSO₄ and 200 µM ascorbate. The spectra were recorded every 1 min and 5 min.

Crystallization and data collection. Cpz10 and its mutants were crystallized by using hanging drop and sitting drop vapor diffusion methods at 20 °C, 16 °C and 4 °C. A total of 10 mg and 5 mg of Cpz10 in 10 mM HEPES (pH 7.0, 100 mM KCl) were mixed with the same volume of reservoir solution for the crystallization. Octahedron crystals (Supplementary Fig. 10) can be obtained in 5 days. The crystalized condition for the apo form protein is 100 mM calcium acetate, 100 mM MES (pH 6.5), 9% (v/v) PEG 4 K; the other condition for complex screening is 10 mg and 5 mg Cpz10 in 10 mM K₂HPO₄ (pH 7.0), 0.1 mM FeSO₄ and 6 mM αKG, contains 0.2 M potassium/sodium tartrate, 0.1 M Bis Tris propane (pH 6.5), 20% (w/v) PEG 3350. Moreover, for mutants, first Cpz10-D117G crystals were obtained in following crystallization condition: 0.2 M calcium acetate hydrate, 20% (w/v) PEG 3350, (pH 7.5) and for Cpz10-R170A: 0.1 M lithium sulfate monohydrate, 0.1 M ADA (pH 6.5), 12% PEG 4 K and 2% v/v 2-propanol. For the soaking condition, an overnight before collecting data, the crystals were exposed to glutaraldehyde. X-ray diffraction data sets were collected on an ADSC Quantum-315 or MX300HE CCD detectors at beamlines 13B1, 13C1, 15A1, 05 A of the National Synchrotron Radiation Research Center (Taiwan), or the beamline 44XU of the Spring-8 (Japan).

Structure determination and refinement. The data sets were processed and scaled with HKL-2000⁷⁸. The crystal structures of Cpz10 was determined by molecular replacement (MR) and a PHASER software in CCP4⁷⁹. The crystal structure of Hyps (Protein Data Bank (PDB) code 4P7W) was used as the search model. All structures were further refined using REFMAC⁸⁰. Subsequent iterative cycles of model building and refinement were performed by COOT⁸¹ and PHENIX⁸². All non-hydrogen atoms were refined with anisotropic displacement parameters (ADP). Both rigid body and real space refinement were applied. Both bulk-solvent correction and anisotropic scaling were also

implemented. Refinement was carried out using the conjugate-gradient least-squares method in SHELXL, in which non-geometric restraints were applied for Cpz10 conformers. The weights for stereochemical and ADP restraints were optimized by scaling in the PHENIX refinement. Structure models and electron density maps were presented using PyMOL⁸³.

Computational methods. To characterize accurately the physical nature of the intermolecular interactions in the bonds of residue–Fe, αKG–Fe, and Wt–Fe, the active site and the Fe-binding site of Cpz10 were undergone natural bond orbital (NBO) analysis^{51,52}. To attain this objective, we firstly designed two structural models, one of which includes the Fe (III) cation, Wt91, αKG, and the cited residues, named model I, and the other consists of the Fe (II) cation, Wt56 and Wt89 in addition with the specified residues, named model II. Because of the lack of coordinates of hydrogen atoms in the X-ray structure of Cpz10, each structural model was partly optimized to correctly determine the positions of hydrogen atoms by using the unrestricted hybrid meta-GGA density functional (UM06-2X)⁸⁴ and the standard basis sets of 6–31 G** (for C H N O atoms) and Lanl2dz (for the Fe cations)⁸⁵. The coordinates of non-hydrogen atoms were frozen during the geometry optimization. For the LS, IS, and HS states of Fe (III), the structural model I was separately optimized in three multiplicities of doublet, quartet, and sextet. Likewise, for the LS, IS, and HS states of Fe (II), the geometry optimization of the structural model II was discretely carried out in multiplicities of singlet, triplet, and quintet. Afterwards, NBO analysis was performed on each optimized structural model at the same computational level. All density functional theory (DFT) calculations were done by means of the Gaussian 09 software⁸⁶.

Setup CPZ–αKG Complex for MD simulation. The crystal structure of CPZ–αKG was used as initial coordinateness for molecular dynamics (MD) simulation utilizing NAMD program package⁸⁷. To attain this objective, firstly, the protein topology file of CHARMM36 force field⁸⁸ was applied to create topology entries of αKG and Fe ions as well as to add the missing hydrogen atoms to the non-hydrogen atoms of CPZ protein by employing the VMD software⁸⁹. Secondly, the CPZ–αKG complex was located in a periodic rectangular box comprising a 12 Å layer of TIP3P water molecules. Afterwards, counter sodium and chloride ions were added to the surface of protein to neutralize the total charge of the protein–water system. Thirdly, this neutral solvated CPZ–αKG system was initially simulated for 2 ns in the NPT ensemble at 310 K and 1 atm and then the MD simulation was continued for 10 ns in the NVT ensemble. Simulations in both ensembles were performed with time step 1 fs and by using the SHAKE algorithm. Particle Mesh Ewald (PME) approach was applied to calculate the electrostatic interactions, whereas a 12 Å cut off was imposed on van der Waals interactions modeled by the Lennard–Jones potential.

Data availability

Relevant data are available from the corresponding author on reasonable request. The atomic coordinates and structure factors of Cpz10 apo form (PDB ID code: 7V4N), Cpz10/αKG (PDB ID code: 7V4M), Cpz10/SA/compound **12** (PDB ID code: 7V4F), Cpz10 R170A mutant (PDB ID code: 7V4O), Cpz10 D117G mutant (PDB ID code: 7V4P). Two pdb files for CPZ–ligand configurations at initial and final states were provided in Supplementary Data 1 and Data 2, two DFT atomic coordinates of the optimized computational models were provided in Supplementary Data 3 and two Supplementary Movies were provided in Supplementary Movie 1 and Supplementary Movie 2.

Received: 7 October 2021; Accepted: 13 July 2022;
Published online: 28 July 2022

References

1. Winn, M., Goss, R. J., Kimura, K.-I. & Bugg, T. D. Antimicrobial nucleoside antibiotics targeting cell wall assembly: Recent advances in structure–function studies and nucleoside biosynthesis. *Nat. Prod. Rep.* **27**, 279–304 (2010).
2. Igarashi, M. et al. Caprazamycin B, a novel anti-tuberculosis antibiotic, from *Streptomyces* sp. *J. Antibiot.* **56**, 580–583 (2003).
3. McErlean, M., Liu, X., Cui, Z., Gust, B. & Van Lanen, S. Identification and characterization of enzymes involved in the biosynthesis of pyrimidine nucleoside antibiotics. *Nat. Prod. Rep.* (2021).
4. Draelos, M. M., Thanapipatsiri, A., Sucipto, H. & Yokoyama, K. Cryptic phosphorylation in nucleoside natural product biosynthesis. *Nat. Chem. Biol.*, 1–9 (2020).
5. Malek Zadeh, S. et al. Theoretical study of intermolecular interactions between critical residues of membrane protein MraYAA and promising antibiotic muraymycin D2. *ACS omega* **5**, 22739–22749 (2020).
6. Barnard-Britson, S. et al. Amalgamation of nucleosides and amino acids in antibiotic biosynthesis: discovery of an L-threonine: uridine-5'-aldehyde transaldolase. *J. Am. Chem. Soc.* **134**, 18514–18517 (2012).
7. Ushimaru, R. & Liu, H.-w. Biosynthetic origin of the atypical stereochemistry in the thioheptose core of albomycin nucleoside antibiotics. *J. Am. Chem. Soc.* **141**, 2211–2214 (2019).
8. Kaysser, L. et al. Formation and attachment of the deoxysugar moiety and assembly of the gene cluster for caprazamycin biosynthesis. *Appl. Environ. Microbiol.* **76**, 4008–4018 (2010).
9. Fujita, Y. et al. Muraminomicins, new lipo-nucleoside antibiotics from *Streptosporangium* sp. SANK 60501-structure elucidations of muraminomicins and supply of the core component for derivatization. *J. Antibiot.* **72**, 943–955 (2019).
10. Ishizaki, Y. et al. Inhibition of the first step in synthesis of the mycobacterial cell wall core, catalyzed by the GlcNAc-1-phosphate transferase WecA, by the novel caprazamycin derivative CPZEN-45. *J. Biol. Chem.* **288**, 30309–30319 (2013).
11. Dong, Y. Y. et al. Structures of DPAGT1 explain glycosylation disease mechanisms and advance TB antibiotic design. *Cell* **175**, 1045–1058. e1016 (2018).
12. Shiraishi, T., Hiro, N., Igarashi, M., Nishiyama, M. & Kuzuyama, T. Biosynthesis of the antituberculous agent caprazamycin: Identification of caprazol-3'-phosphate, an unprecedented caprazamycin-related metabolite. *J. Gen. Appl. Microbiol.* **62**, 164–166 (2016).
13. Yang, Z., Unrine, J., Nonaka, K. & Van Lanen, S. G. Fe (II)-dependent, uridine-5'-monophosphate α -ketoglutarate dioxygenases in the synthesis of 5'-modified nucleosides. *Methods Enzymol.* **516**, 153–168 (2012).
14. Kaysser, L. et al. Identification and manipulation of the caprazamycin gene cluster lead to new simplified liponucleoside antibiotics and give insights into the biosynthetic pathway. *J. Biol. Chem.* **284**, 14987–14996 (2009).
15. Solomon, E. I., Goudarzi, S. & Sutherlin, K. D. O₂ activation by non-heme iron enzymes. *Biochemistry* **55**, 6363–6374 (2016).
16. Koketsu, K. et al. Refined regio- and stereoselective hydroxylation of L-pipecolic acid by protein engineering of L-proline cis-4-hydroxylase based on the X-ray crystal structure. *ACS Synth. Biol.* **4**, 383–392 (2015).
17. Goudarzi, S. et al. Evaluation of a concerted vs. sequential oxygen activation mechanism in α -ketoglutarate-dependent nonheme ferrous enzymes. *Proc. Natl Acad. Sci. USA* **117**, 5152–5159 (2020).
18. O'Brien, J. R., Schuller, D. J., Yang, V. S., Dillard, B. D. & Lanzilotta, W. N. Substrate-induced conformational changes in *Escherichia coli* taurine/ α -ketoglutarate dioxygenase and insight into the oligomeric structure. *Biochemistry* **42**, 5547–5554 (2003).
19. Solomon, E. I., Decker, A. & Lehnert, N. Non-heme iron enzymes: contrasts to heme catalysis. *Proc. Natl Acad. Sci. USA* **100**, 3589–3594 (2003).
20. Cui, Z. et al. Pyridoxal-5'-phosphate-dependent alkyl transfer in nucleoside antibiotic biosynthesis. *Nat. Chem. Biol.* **16**, 904–911 (2020).
21. Wiker, F., Hauck, N., Grond, S. & Gust, B. Caprazamycins: biosynthesis and structure activity relationship studies. *Int. J. Med. Microbiol.* **309**, 319–324 (2019).
22. Yang, Z. et al. Characterization of LipL as a non-heme, Fe (II)-dependent α -ketoglutarate: UMP dioxygenase that generates uridine-5'-aldehyde during A-90289 biosynthesis. *J. Biol. Chem.* **286**, 7885–7892 (2011).
23. Scott, T. A., Heine, D., Qin, Z. & Wilkinson, B. An L-threonine transaldolase is required for L-threo- β -hydroxy- α -amino acid assembly during obafuorin biosynthesis. *Nat. Commun.* **8**, 1–11 (2017).
24. Yan, W. et al. Endoperoxide formation by an α -ketoglutarate-dependent mononuclear non-haem iron enzyme. *Nature* **527**, 539–543 (2015).
25. Light, K. M., Hangasky, J. A., Knapp, M. J. & Solomon, E. I. Spectroscopic studies of the mononuclear non-heme FeII enzyme FIH: Second-sphere contributions to reactivity. *J. Am. Chem. Soc.* **135**, 9665–9674 (2013).
26. Elkins, J. M. et al. X-ray crystal structure of *Escherichia coli* taurine/ α -ketoglutarate dioxygenase complexed to ferrous iron and substrates. *Biochemistry* **41**, 5185–5192 (2002).
27. Shanmugam, M., Quareshy, M., Cameron, A. D., Bugg, T. D. & Chen, Y. Light-activated electron transfer and catalytic mechanism of carnitine oxidation by rieske-type oxygenase from human microbiota. *Angew. Chem. Int.* **60**, 4529–4534 (2021).
28. Parales, R. E., Parales, J. V. & Gibson, D. T. Aspartate 205 in the catalytic domain of naphthalene dioxygenase is essential for activity. *J. Bacteriol.* **181**, 1831–1837 (1999).
29. Rogers, M. S. & Lipscomb, J. D. Salicylate 5-hydroxylase: Intermediates in aromatic hydroxylation by a Rieske monooxygenase. *Biochemistry* **58**, 5305–5319 (2019).
30. Martinez, S. & Hausinger, R. P. Catalytic mechanisms of Fe (II)- and 2-oxoglutarate-dependent oxygenases. *J. Biol. Chem.* **290**, 20702–20711 (2015).
31. Zhou, J. et al. Spectroscopic studies of substrate interactions with clavaminatase synthase 2, a multifunctional α -KG-dependent non-heme iron enzyme: correlation with mechanisms and reactivities. *J. Am. Chem. Soc.* **123**, 7388–7398 (2001).
32. Rabe, P., Kamps, J. J., Schofield, C. J. & Lohans, C. T. Roles of 2-oxoglutarate oxygenases and isopenicillin N synthase in β -lactam biosynthesis. *Nat. Prod. Rep.* **35**, 735–756 (2018).
33. Ryle, M. J. et al. O₂- and α -ketoglutarate-dependent tyrosyl radical formation in TauD, an α -keto acid-dependent non-heme iron dioxygenase. *Biochemistry* **42**, 1854–1862 (2003).
34. Bradley, J. M., Svistunenko, D. A., Moore, G. R. & Le Brun, N. E. Tyr25, Tyr58 and Trp133 of *Escherichia coli* bacterioferritin transfer electrons between iron in the central cavity and the ferroxidase centre. *Metallomics* **9**, 1421–1428 (2017).
35. de Sousa, D. P. et al. Redox- and EPR-active graphene diiron complex nanocomposite. *Langmuir* **35**, 12339–12349 (2019).
36. Tampieri, F., Silvestrini, S., Riccò, R., Maggini, M. & Barbon, A. A comparative electron paramagnetic resonance study of expanded graphites and graphene. *J. Mater. Chem. C* **2**, 8105–8112 (2014).
37. Bhattacharyya, A., Schmidt, M. P., Stavitski, E. & Martínez, C. E. Iron speciation in peats: chemical and spectroscopic evidence for the co-occurrence of ferric and ferrous iron in organic complexes and mineral precipitates. *Org. Geochem.* **115**, 124–137 (2018).
38. Weckhuysen, B. M., Heidler, R. & Schoonheydt, R. A. In *Characterization I* 295–335 (Springer, 2004).
39. Liberman, I. et al. Active-site modulation in an Fe-porphyrin-based metal-organic framework through ligand axial coordination: accelerating electrocatalysis and charge-transport kinetics. *J. Am. Chem. Soc.* **142**, 1933–1940 (2020).
40. Iwamoto, L. S., Serra, O. A., Manso, C., Prado, M. & Yamamoto, Y. One pot obtention of a tetrabutylammonium hydroxide solution for ironporphyrin-OH-interaction studies in organic solvents. *Química Nova* **22**, 277–279 (1999).
41. Hirst, J. et al. Replacement of the axial histidine ligand with imidazole in cytochrome c peroxidase. 2. Effects on heme coordination and function. *Biochemistry* **40**, 1274–1283 (2001).
42. Wolfe, M. D. & Lipscomb, J. D. Hydrogen peroxide-coupled cis-diol formation catalyzed by naphthalene 1, 2-dioxygenase. *J. Biol. Chem.* **278**, 829–835 (2003).
43. Bou-Abdallah, F. & Chasteen, N. D. Spin concentration measurements of high-spin ($g' = 4.3$) rhombic iron (III) ions in biological samples: theory and application. *J. Biol. Inorg. Chem.* **13**, 15–24 (2008).
44. Rajakovich, L. J. et al. A new microbial pathway for organophosphonate degradation catalyzed by two previously misannotated non-heme-iron oxygenases. *Biochemistry* **58**, 1627–1647 (2019).
45. Wörsdörfer, B. et al. Organophosphonate-degrading PhnZ reveals an emerging family of HD domain mixed-valent diiron oxygenases. *Proc. Natl Acad. Sci. USA* **110**, 18874–18879 (2013).
46. Fufezan, C., Zhang, C., Krieger-Liszka, A. & Rutherford, A. W. Secondary quinone in photosystem II of *Thermosynechococcus elongatus*: Semiquinone –Iron EPR signals and temperature dependence of electron transfer. *Biochemistry* **44**, 12780–12789 (2005).
47. Dong, M. et al. Organometallic and radical intermediates reveal mechanism of diphthamide biosynthesis. *Science* **359**, 1247–1250 (2018).
48. Xing, G. et al. A coupled dinuclear iron cluster that is perturbed by substrate binding in myo-inositol oxygenase. *Biochemistry* **45**, 5393–5401 (2006).
49. Doi, K., McCracken, J., Peisach, J. & Aisen, P. The binding of molybdate to uteroferrin. Hyperfine interactions of the binuclear center with ⁹⁵Mo, ¹H, and ²H. *J. Biol. Chem.* **263**, 5757–5763 (1988).
50. Oloo, W. N. et al. Identification of a low-spin acylperoxoiron (III) intermediate in bio-inspired non-heme iron-catalysed oxidations. *Nat. Commun.* **5**, 1–9 (2014).
51. Reed, A. E., Curtiss, L. A. & Weinhold, F. Intermolecular interactions from a natural bond orbital, donor-acceptor viewpoint. *Chem. Rev.* **88**, 899–926 (1988).
52. Weinhold, F. *Discovering chemistry with natural bond orbitals*. (John Wiley & Sons, 2012).

53. Alabugin, I. V., Manoharan, M., Peabody, S. & Weinhold, F. Electronic basis of improper hydrogen bonding: a subtle balance of hyperconjugation and rehybridization. *J. Am. Chem. Soc.* **125**, 5973–5987 (2003).
54. Alberti, F. & Corre, C. Editing streptomycete genomes in the CRISPR/Cas9 age. *Nat. Prod. Rep.* **36**, 1237–1248 (2019).
55. Kallifidas, D., Jiang, G., Ding, Y. & Luesch, H. Rational engineering of *Streptomyces albus* J1074 for the overexpression of secondary metabolite gene clusters. *Microb. Cell. Fact.* **17**, 1–14 (2018).
56. Ochi, K., Ezaki, M., Iwami, M., Komori, T. & Kohsaka, M. (Google Patents, 1990).
57. Kapp, U. et al. Structure of *Deinococcus radiodurans* tunicamycin-resistance protein (TmrD), a phosphotransferase. *Acta Crystallogr. Sect. F: Struct. Biol. Crystallization Commun.* **64**, 479–486 (2008).
58. Yang, Z. et al. Functional and kinetic analysis of the phosphotransferase CapP conferring selective self-resistance to capuramycin antibiotics. *J. Biol. Chem.* **285**, 12899–12905 (2010).
59. Cui, Z. et al. Self-resistance during muraymycin biosynthesis: a complementary nucleotidyltransferase and phosphotransferase with identical modification sites and distinct temporal order. *Antimicrob. Agents Chemother.* **62** (2018).
60. Cui, Z. et al. Enzymatic synthesis of the ribosylated glycol-uridine disaccharide core of peptidyl nucleoside antibiotics. *J. Org. Chem.* **83**, 7239–7249 (2018).
61. Jiang, L. et al. Identification of novel mureidomycin analogues via rational activation of a cryptic gene cluster in *Streptomyces roseosporus* NRRL 15998. *Sci. Rep.* **5**, 1–13 (2015).
62. Wright, G. D. Opportunities for natural products in 21 st century antibiotic discovery. *Nat. Prod. Rep.* **34**, 694–701 (2017).
63. Bjørnstad, L. G., Zoppellaro, G., Tomter, A. B., Falnes, P. Ø. & Andersson, K. K. Spectroscopic and magnetic studies of wild-type and mutant forms of the Fe (II)- and 2-oxoglutarate-dependent decarboxylase ALKBH4. *Biochemistry* **434**, 391–398 (2011).
64. Valegård, K. et al. The structural basis of cephalosporin formation in a mononuclear ferrous enzyme. *Nat. Struct. Mol. Biol.* **11**, 95–101 (2004).
65. Chang, C. Y. et al. Biosynthesis of streptolidine involved two unexpected intermediates produced by a dihydroxylation and a cyclase through unusual mechanisms. *Angew. Chem.* **126**, 1974–1979 (2014).
66. Bethe, H. Termäufspaltung in kristallen. *Ann. der Phys.* **395**, 133–208 (1929).
67. Van Vleck, J. Theory of the variations in paramagnetic anisotropy among different salts of the iron group. *Phys. Rev.* **41**, 208 (1932).
68. Schlapp, R. & Penney, W. G. Influence of crystalline fields on the susceptibilities of salts of paramagnetic ions. II. The iron group, especially Ni, Cr and Co. *Phys. Rev.* **42**, 666 (1932).
69. Penney, W. G. & Schlapp, R. The influence of crystalline fields on the susceptibilities of salts of paramagnetic ions. I. The rare earths, especially Pr and Nd. *Phys. Rev.* **41**, 194 (1932).
70. Levine, I. N., Busch, D. H. & Shull, H. *Quantum chemistry*. Vol. 6 (Pearson Prentice Hall Upper Saddle River, NJ, 2009).
71. Baker, H. M., Anderson, B. F. & Baker, E. N. Dealing with iron: common structural principles in proteins that transport iron and heme. *Proc. Natl Acad. Sci. USA* **100**, 3579–3583 (2003).
72. Tang, X. et al. A two-step sulfation in antibiotic biosynthesis requires a type III polyketide synthase. *Nat. Chem. Biol.* **9**, 610–615 (2013).
73. von Liponukleosid-Antibiotika, B. Biosynthesis of liponucleoside antibiotics in *Streptomyces*: Molecular and biochemical investigations of the caprazamycin and the liposidomycin gene cluster.
74. Horbal, L., Fedorenko, V. & Luzhetskyy, A. Novel and tightly regulated resorcinol and cumate-inducible expression systems for *Streptomyces* and other actinobacteria. *Appl. Environ. Microbiol.* **98**, 8641–8655 (2014).
75. Kieser, T., Bibb, M. J., Buttner, M. J., Chater, K. F. & Hopwood, D. A. *Practical streptomyces genetics*. Vol. 291 (John Innes Foundation Norwich, 2000).
76. Jiang, W. & Zhu, T. F. Targeted isolation and cloning of 100-kb microbial genomic sequences by Cas9-assisted targeting of chromosome segments. *Nat. Protoc.* **11**, 960–975 (2016).
77. Adhikari, K. et al. Chemoenzymatic synthesis and biological evaluation for bioactive molecules derived from bacterial benzoyl coenzyme A ligase and plant type III polyketide synthase. *Biomolecules* **10**, 738 (2020).
78. Otwinowski, Z. & Minor, W. [20] Processing of X-ray diffraction data collected in oscillation mode. *Meth. Enzymol.* **276**, 307–326 (1997).
79. McCoy, A. J. et al. Phaser crystallographic software. *J. Appl. Crystallogr.* **40**, 658–674 (2007).
80. Vagin, A. A. et al. REFMAC5 dictionary: organization of prior chemical knowledge and guidelines for its use. *Acta Crystallogr. Sect. D: Biol. Crystallogr.* **60**, 2184–2195 (2004).
81. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. Sect. D: Biol. Crystallogr.* **66**, 486–501 (2010).
82. Afonine, P. V. et al. Towards automated crystallographic structure refinement with phenix. *Acta Crystallogr. Sect. D: Biol. Crystallogr.* **68**, 352–367 (2012).
83. DeLano, W. L. The PyMOL molecular graphics system. <http://www.pymol.org> (2002).
84. Zhao, Y. & Truhlar, D. G. The M06 suite of density functionals for main group thermochemistry, thermochemical kinetics, noncovalent interactions, excited states, and transition elements: two new functionals and systematic testing of four M06-class functionals and 12 other functionals. *Theor. Chem. Acc.* **120**, 215–241 (2008).
85. Wadt, W. & P.J., H. Ab initio effective core potentials formolecular calculations. Potentials for K to Au including theoutermost core orbitals. *J. Chem. Phys.* **82**, 299–305 (1985).
86. Frisch, M. et al. Gaussian Inc. *Wallingford Ct* **2009** (2009).
87. Kalé, L. et al. NAMD2: greater scalability for parallel molecular dynamics. *J. Comput. Phys.* **151**, 283–312 (1999).
88. Best, R. B. et al. Optimization of the additive CHARMM all-atom protein force field targeting improved sampling of the backbone ϕ , ψ and side-chain χ 1 and χ 2 dihedral angles. *J. Chem. Theory Comput.* **8**, 3257–3273 (2012).
89. Humphrey, W., Dalke, A. & Schulten, K. VMD: visual molecular dynamics. *J. Mol. Graph.* **14**, 33–38 (1996).

Acknowledgements

This work was supported by funds from the Ministry of Science and Technology (MOST), Taiwan (MOST-108-2113-M-001-021-MY3 and MOST 110-0210-01-22-02) and Academia Sinica (AS-KPQ-109-BioMed, 109-0210-01-18-02 and AS-IA-109-L06). Portions of this research were carried out at the National Synchrotron Radiation Research Center (NSRRC), a national user facility supported by MOST of Taiwan, ROC. We thank NSRRC (Taiwan) and Spring-8 (Japan) for beam time allocations at beam lines 13B, 05 A, 15 A, and 44XU.

Author contributions

T.L.L. conceived the project. S.M.Z., M.H.C., K.H.L. N.S.H., K.A., and S.Y.L. collected and analyzed the data. S.M.Z., N.S.H. refined and solved the structures. I.W.L. determined NMR. For small molecules. S.M.Z. and Z.C.W. performed mutagenesis and biochemical assays. Z.C.W. made the *cpz*-BAC system and gene-deletion mutants. S.M.Z., M.H.C., K.H.L. and N.S.H. performed enzyme kinetics, EPR and enzymatic assays; E.K.A. performed chemical computation, molecular dynamics and analysis. S.M.Z. and H.Y.T. purified proteins and small molecules; Y.T., M.Y., K.Y., S.I. synthesized compounds; S.M.Z. and T.L.L. wrote the paper with the feedback from all authors.

Competing interests

The authors declare no competing interests.

Additional information


Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s42004-022-00703-6>.

Correspondence and requests for materials should be addressed to Tsung-Lin Li.

Peer review information *Communications Chemistry* thanks Binju Wang and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

Reprints and permission information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

 **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022