Flavoenzyme CrmK-mediated substrate recycling in caerulomycin biosynthesis†

Yiguang Zhu,†‡ a Marie-Eve Picard,†b Qingbo Zhang,a Julie Barma,b Xavier Murphy Després,b Xiangui Mei,c Liping Zhang,a Jean-Baptiste Duvignaud,b Manon Couture,b Weiming Zhu,c Rong Shi‡b and Changsheng Zhang§a

Substrate salvage or recycling is common and important for primary metabolism in cells but is rare in secondary metabolism. Herein we report flavoenzyme CrmK-mediated shunt product recycling in the biosynthesis of caerulomycin A (CRM A1), a 2,2 ‐bipyridine ‐ containing natural product that is under development as a potent novel immunosuppressive agent. We demonstrated that the alcohol oxidase CrmK, belonging to the family of bicovalent FAD‐binding flavoproteins, catalyzed the conversion of an alcohol into a carboxylate via an aldehyde. The CrmK‐mediated reactions were not en route to biosynthesis but played an unexpectedly important role by recycling shunt products back to the main pathway of 1. Crystal structures and site‐directed mutagenesis studies uncovered key residues for FAD‐binding, substrate binding and catalytic activities, enabling the proposal for the CrmK catalytic mechanism. This study provides the first biochemical and structural evidence for flavoenzyme‐mediated substrate recycling in secondary metabolism.

Introduction

The substrate salvage or recycling pathway is important and common in primary metabolism, such as for nucleosides,1 uracil,2 coenzyme B1,3 tetrahydrobiopterin,4 methionine,5 thiamine,6 and sphingolipids.7,8 The existence of a salvage pathway guarantees the preservation of cell integrity,9 and allows the recycling of metabolites derived from normal cellular catabolism, in compensation to the general de novo biosynthetic pathway.10 However, there are rare examples of the salvage pathway in secondary metabolism.11–13 In this study, we provided evidence of flavoenzyme CrmK‐mediated substrate recycling in caerulomycin biosynthesis. Caerulomycins (CRMs) feature a 2,2 ‐bipyridine core structure and are produced by terrestrial Streptomyces caeruleus14 and marine‐derived Actinoalloteichus cyanogriseus WH1‐2216‐6.15,16 CRMs are of considerable medicinal interest due to their diverse bioactivities.17 Particularly, CRM A1 (Fig. 1) has shown remarkable promise in immunosuppression of both T‐ and B‐cells, being ten times more potent than the known immunosuppressant cyclosporin A,18 and is currently under development as a novel immunosuppressive agent for the treatment of autoimmunity by Nostrum Pharmaceuticals.19,20

Recently, biosynthetic studies on CRM A1 and its closely‐related analogue collismycin A (Fig. 1) have revealed a common origin of a hybrid polyketide synthase (PKS)/non‐ribosomal peptide synthetase (NRPS) system for bipyridine core formation.
with unclear mechanisms. The tailoring steps for collismycin biosynthesis have been demonstrated by characterizing a number of collismycin analogues from gene-disrupted mutants. In parallel, the CRM A1 biosynthetic pathway has been well established by biochemically elucidating four biosynthetic enzymes, including an unusual amidohydrolase CrmL, an aminotransferase CrmG, a two-component monooxygenase CrmH, and a methyltransferase CrmM (Fig. 1). Hence, the conversion of CRM O3 to CRM M4 (Fig. 1) remains the only step en route to CRM A1 without experimental evidence, and was speculated to be catalyzed by a pair of GirC/GirD-like dehydrogenase components CrmN/CrmO.

Interestingly, a flavoenzyme-encoding genecrmK was found to flank the 1 biosynthetic gene cluster, and has been demonstrated to be inessential for CRM A1 biosynthesis, given the production of 1 in the ΔcrmK mutant. In this study, our biochemical characterizations demonstrate that CrmK plays an unexpectedly important role in 1 biosynthesis by catalyzing the conversion of an alcohol shunt product 7 to an aldehyde 4 and then to a carboxylate 3, thus recycling or salvaging the shunt product 7 into the main pathway of 1 (Fig. 1). Moreover, the crystal structures of CrmK in apo form and in complex with the substrate CRM P7 and site-directed mutagenesis studies provide insights into the catalytic mechanisms and the molecular basis of CrmK in shunt product recycling.

Results and discussion

In vivo characterization of the crmK gene

CrmK belongs to a growing family of alcohol oxidase flavoproteins, which share a conserved bicovalent flavin adenine dinucleotide (FAD) binding domain and function to convert an alcohol to an aldehyde or a ketone. CrmK resembles a number of flavin-dependent oxidoreductases in natural product biosynthesis, such as TrdL or TamL, Dbv29, Spnj, AknOx, GilR, and Pac11. These flavoenzymes contain invariant histidine and cysteine residues that are responsible for bicovalent binding of the cofactor FAD (Fig. S1f). HPLC analysis showed that the ΔcrmK mutant accumulated CRM P7 and CRM F8 (Fig. 2 and 3A), two products observed in the ΔcrmG mutant, along with the major product 1. The yield of 1 in the ΔcrmK mutant was 52% of the yield of 1 in the wild type (WT) strain. In contrast, the total yield of CRM (1, 7 and 8) was comparable (91%) to the yield of 1 in the WT (Fig. 2). Given that the ΔcrmK mutant could produce two alcohol-containing shunt products CRM P7 and CRM F8, we reason that CrmK might act as an alcohol oxidase to convert the alcohols CRM P7 and CRM F8 to their aldehyde derivatives (Fig. 3A).

Biochemical characterization of CrmK

To probe the exact function of CrmK in vitro, the crmK gene was overexpressed in E. coli to yield a soluble N-(His)6-tagged protein (Fig. S2†). Although the purified CrmK protein was yellow in color, no flavin cofactors could be released by simple heat denaturation, consistent with the predicted covalent binding of flavin cofactors. Subsequently, a time course of in vitro CrmK assays revealed that 7 was quickly converted to 4, and 4 began to be slowly converted to 3 after most of 7 was consumed (Fig. 3B). The identities of enzymatic products 3 and 4 were confirmed by LC-HRESIMS (Fig. S3f) and comparison with the standards (Fig. 3B). In a similar manner to the turnover of 7 to 3, CrmK was shown to be capable of catalyzing the conversion of 8 to 10 via the intermediate 9 (Fig. 3C), which was confirmed by comparison with the standard 9 (Fig. 3C). The product 10 was isolated and purified from 100 mL of a CrmK enzyme reaction, and was confirmed by HRESIMS (C12H16N2O3, m/z [M + H]+ 231.0769) and NMR analyses (Fig. S4f). These experiments demonstrated that CrmK was a dedicated enzyme in CRM A1.
biosynthesis that catalyzed two sequential oxidizing steps to convert an alcohol to a carboxyl through an aldehyde intermediate, and thus to recycle an alcohol shunt product back to the main pathway of CRM A1 (Fig. 1).

To better understand the CrmK-catalyzed two-step oxidative reactions, kinetic parameters were determined by using the alcohol 7 and the aldehyde 4 as substrates. CrmK displayed much higher binding affinity toward 7 ($K_m$ 8.4 μM, Fig. 4A) than 4 ($K_m$ 2384.0 μM, Fig. 4B). Comparison of the $k_{cat}/K_m$ values revealed that CrmK was almost 230-fold more efficient in catalyzing the first oxidation step from an alcohol 7 to an aldehyde 4 ($k_{cat}/K_m$ 6.94 μM$^{-1}$ min$^{-1}$) than in catalyzing the second oxidation step from an aldehyde 4 to a carboxylate 3 ($k_{cat}/K_m$ 0.0297 μM$^{-1}$ min$^{-1}$).

**Overall crystal structure of CrmK**

To better understand the enzymatic reaction and substrate binding of CrmK, the crystal structures of CrmK alone and in complex with the substrate CRM P7 were resolved at 1.84 Å and 2.15 Å (Table 1), respectively. In both crystals, each asymmetric unit contains four subunits of CrmK that form two stable dimers with a dimerization interface of 1695 Å$^2$, comparable to that in the TamL dimer (1680 Å$^2$) and more than that in the Dbv29 dimer (~1300 Å$^2$). The closest structural homologs of CrmK include TamL, GilR, AknOx, and Dbv29. Each CrmK subunit contains two domains (Fig. 5A): the cofactor FAD binding domain (F) and the substrate binding domain (S).

**FAD-binding site**

The electron density maps clearly showed that CrmK is a bicovalently flavinylated enzyme as the FAD isoalloxazine ring forms two covalent bonds via its 8z-methyl group and C6 carbon atom with the N81 atom of His64 and the sulfur atom of Cys124 (Fig. 5B), respectively. Removing either of the covalent linkages by site-directed mutagenesis (e.g. H64A or C124A) leads to severely impaired mutants with only residual activities, and the double mutant (H64A/C124A) completely abolishes CrmK activities (Table 2). Notably, Y446 from the S domain forms a hydrogen bond with the O2 atom of the isoalloxazine moiety of FAD (Fig. 5C). Consequently, Y446F kept only ~20% activity of wild type CrmK for both steps (Table 2).

**Table 1 X-ray data collection and refinement statistics (values in parentheses are for the highest-resolution shell)**

<table>
<thead>
<tr>
<th>Structure</th>
<th>CrmK-FAD</th>
<th>CrmK-FAD-CRM P7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Space group</td>
<td>P1</td>
<td>P1</td>
</tr>
<tr>
<td>$a$, $b$, $c$ (Å)</td>
<td>63.6, 95.7, 98.4</td>
<td>63.1, 95.2, 98.0</td>
</tr>
<tr>
<td>$\alpha$, $\beta$, $\gamma$ (°)</td>
<td>95.2, 97.0, 104.4</td>
<td>94.7, 96.6, 105.2</td>
</tr>
<tr>
<td>Wavelength (Å)</td>
<td>0.9795</td>
<td>0.9795</td>
</tr>
<tr>
<td>Resolution (Å)</td>
<td>48.7–1.84 (1.94–1.84)</td>
<td>45.6–2.15 (2.27–2.15)</td>
</tr>
<tr>
<td>Observed $hkl$</td>
<td>729 153 (103 972)</td>
<td>278 363 (40 448)</td>
</tr>
<tr>
<td>Unique $hkl$</td>
<td>184 445 (26 576)</td>
<td>114 678 (16 652)</td>
</tr>
<tr>
<td>Redundancy</td>
<td>4.0 (3.9)</td>
<td>2.4 (2.4)</td>
</tr>
<tr>
<td>Completeness (%)</td>
<td>95.9 (94.7)</td>
<td>97.0 (96.1)</td>
</tr>
<tr>
<td>$R_{merge}$</td>
<td>0.118 (0.828)</td>
<td>0.162 (0.699)</td>
</tr>
<tr>
<td>$CC_{1/2}$</td>
<td>0.993 (0.498)</td>
<td>0.950 (0.400)</td>
</tr>
<tr>
<td>$I/\sigma(I)$</td>
<td>7.5 (2.1)</td>
<td>5.0 (2.8)</td>
</tr>
<tr>
<td>Wilson $B$ ($\AA^2$)</td>
<td>22.5</td>
<td>23.7</td>
</tr>
<tr>
<td>$R_{work}$ (# $hkl$)</td>
<td>0.161 (175 148)</td>
<td>0.168 (108 921)</td>
</tr>
<tr>
<td>$R_{free}$ (# $hkl$)</td>
<td>0.194 (9296)</td>
<td>0.213 (5754)</td>
</tr>
</tbody>
</table>

**B-factors (# atoms)**

| Protein | 29.8 (15 629) | 35.0 (15 577) |
| Solvent  | 37.5 (1386)  | 37.3 (600) |
| Ligand (FAD) | 19.9 (212) | 23.8 (212) |
| Ligand (7 or 4) | 47.1 (60) | 47.1 (60) |

**Ramachandran**

| Allowed (%) | 99.9 |
| Generous (%) | 0.1 |
| Disallowed (%) | 0 |

**rmsd**

| Bonds ($\AA$) | 0.014 | 0.015 |
| Angles (°)     | 1.67  | 1.71  |
| PDB code       | 5H1V  | 5H1W  |

![Fig. 4](image-url) Determination of CrmK kinetic parameters for substrate 7 with 5–200 μM 7 (A) and for substrate 4 with 500–8000 μM 4 (B).
G66, and the aliphatic moiety of R340. The importance of these residues was confirmed by site-directed mutagenesis studies. The CrmK mutants F403A, F403E, F123A and R340A led to impaired or largely reduced enzymatic activities toward both oxidative steps (Table 2). Remarkably, the putative catalytic residue Y449 forms the only H-bond (2.6–2.9 Å) between the CRM P7 and CrmK residues, contributing to anchoring the C7 hydroxyl group of the substrate (Fig. 5C). Interestingly, this hydroxyl group of CRM P7 is also within the H-bonding distance (2.6–3.0 Å) of the N5 atom on the FAD isoalloxazine ring. The only major conformational change in CrmK upon the substrate binding (compared with the substrate-free form, Fig. S5†) is the swing of the Y449 side chain to enable H-bond (2.6–3.0 Å) formation with the Y138 hydroxyl group. Notably, no water molecules could be identified in the binding pocket to mediate additional H-bond interactions between the substrates.
and CrmK. Other residues that line the active site pocket also include E376 and I378, which are ~4.3–5.5 Å away from the ligand (Fig. 5C).

### Proposed catalytic mechanism of CrmK

Most characterized alcohol oxidases catalyze a two-electron oxidation to convert an alcohol to a ketone (e.g. TrdL/TamL, SpnJ, GilR, and Pac11). However, a four-electron oxidation process was observed for CrmK catalysis, similar to the glycopeptide hexose oxidase Dbv29, to catalyze the conversion of an alcohol to a carboxylate. In Dbv29, this conversion was proposed to undergo an aldehyde intermediate (only trapable by chemical derivatization) and a diol species. In sharp contrast to the Dbv29-catalyzed reaction, an aldehyde 4 was readily available as a CrmK reaction product and was also trapped in one subunit of the CrmK–ligand complex structure, providing solid evidence for the first two-electron oxidation process to convert an alcohol to an aldehyde. Interestingly, the mass unit of a diol intermediate (m/z 219 [M + H]+) was readily detectable by simply dissolving 4 (m/z 201 [M + H]+) in water (Fig. S6f). When 4 was dissolved in 18O-labeled water, a series of mass units were observed to correspond to 18O-labeled aldehyde 4 (m/z 203 [M + H]+), single 18O-labeled diol species (m/z 221 [M + H]+), and double 18O-labeled diol species (m/z 223 [M + H]+). These observations indicated a spontaneous hydration/dehydration equilibrium in water for an interconversion between the aldehyde 4 and the diol species (Fig. S6f). When the CrmK reaction with 7 was performed in Tris–HCl buffer prepared with 18O-labeled water, the final product exhibited a mass unit (m/z 221 [M + H]+) for double 18O-labeled 3 (Fig. S6f). Cumulatively, these data unequivocally demonstrated that CrmK catalyzed two oxidative steps, to convert the alcohol 7 to the carboxylate 3 via the aldehyde 4 (or its diol species), mechanistically analogous to the Dbv29 catalysis.

The overall arrangement of active site residues in the CrmK structure in complex with CRM P 7 suggested that Y449, Y138 and E376 are important catalytic residues (Fig. 5C and D). Site-directed mutagenesis studies revealed that the mutants Y449F, Y138F and E376A impaired both oxidative steps (Table 2). Especially, Y449F completely lost the ability to catalyze the second step from an aldehyde to a carboxylate. The double mutant Y138F/Y449F maintains 18% activity for the first step. In sharp contrast, the double mutants of tyrosine residues (Y138/ Y449 equivalents) in AknOx and Dbv29 completely lost their catalytic activity, and the single mutant Y448F in GilR led to a 400-fold reduction in activity. This unique feature of CrmK catalysis may be attributed to the presence of a second catalytic base E376, alternative to Y449. E376 is unique in CrmK and is not found in its homologous enzymes (Fig. S1†). The E376A mutant keeps only 35% of the CrmK activity, similar to that of Y449F (29%). Structurally, E376 could form an H-bond with the C7 hydroxyl group of 7 that adopts a different orientation from the one observed in the crystal structure (Fig. 5D). Interestingly, AknOx has a dual active site (Y450/Y144 and Y378) catalyzing two different dehydrogenations (Fig. S7†). Superimposition of CrmK and AknOx structures reveals that the position of the carboxylate of E376 in the active site of CrmK matches well with the position of the hydroxyl group of Y378 in AknOx (1.2–1.3 Å away, Fig. S8†), thus also indicating the potential role of E376 as a catalytic base in CrmK.

Taken together, a catalytic mechanism for CrmK was proposed (Fig. 5E). The first-step oxidation is likely initiated via proton abstraction from the C7 hydroxyl group by Y449 and/or E376 with concomitant hydride transfer from C7 to the N5 atom of FAD. The proton transfer by Y449 can be aided by Y138, given the H-bond between their hydroxyl groups, similar to the tyrosine pairs in TamL, AknOx, Dbv29, and GilR. In Dbv29 catalysis, two well-defined water molecules were found in the vicinity of diol groups that may facilitate the formation of the diol intermediate. In CrmK, the equivalent position is occupied by the F123 side chain, thus excluding the presence of water there. The lack of well-defined water molecules in the CrmK active site may suggest that the first half reaction product (CRM M 4) is released to the solvent to become a diol intermediate prior to re-binding to the active site. As such, CRM M 4 became an available product during CrmK catalytic cycles.

### Substrate recycling in caerulomycin biosynthesis

In this study, the flavoenzyme CrmK was demonstrated to play an unexpectedly important role in reflexing shunt products back to the main pathway of CRM A 1 (Fig. 6). Our previous characterization of CrmG kinetic parameters revealed that the conversion of CRM M 4 to CRM N 5 (Kcat/Km 14.5 µM, kcat/Km 0.0345 µM⁻¹ min⁻¹) was a rather slow process, owing to an intrinsically complex ping-pong mechanism for amino transfer. We have also previously shown that CrmH could catalyze the conversion of CRM N 5 into both E- and Z-configured CRM H 6, while the unstable Z-configured CRM H 6 could be spontaneously transformed into the aldehyde CRM M 4 (Fig. 6). These data indicated the possibility of the accumulation of the
aldehyde CRM M 4 in cells. However, cells usually could not tolerate the accumulation of aldehydes because aldehydes are reactive electrophiles that form adducts with cellular protein thiols and amines and thus can cause toxicity.42 Thus, aldehydes are detoxified in cells by enzymatic or spontaneous oxidation to acids and/or reduction to alcohols.42 Consistent with this statement,42 aldehydes 7 and 8 were accumulated in the ΔcrmG mutant (Fig. 2), and the expected aldehydes 4 and 9 were not found but were confirmed in vitro to be spontaneously converted to 7 and 8, putatively via a Cannizzaro reaction.26 Similar results were also observed in the aminotransferase gene disrupted mutants for studying collisycin A and borrelidin biosynthesis.24,41 It would be a great energy waste for cells to produce alcohols 7 and 8 in 1 biosynthesis. As such, the flavoenzyme CrmK was incorporated into the 1 biosynthetic pathway, to convert 7 and 8 to their corresponding aldehydes, which could be further processed by downstream enzymes to make the final product 1 (Fig. 6). Indeed, CRM P 7 was almost completely converted to 1 when fed to the ΔcrmA (a hybrid PKS/NRPS encoding gene) mutant, in which the production of CRMs was blocked (Fig. S9†).21 CRM F 8 was also converted to 1 by the ΔcrmA mutant, but in a ratio of less than 5% (Fig. S9†), indicating that the tailoring enzymes (e.g. CrmG26 and CrmH27) could not efficiently process the derivatives of 8. It should be also noted that, on the basis of kinetic parameters of CrmK and CrmG, the aldehyde CRM M 4 would be captured by the aminotransferase CrmG (the \(K_m\) value of the aldehyde 4 for CrmG26 is 160-fold lower than that for the second step oxidation by CrmK). That means, under physiological conditions, the second step oxidation of 4 (\(K_m = 2384 \mu M\)) to 3 by CrmK (\(k_{cat}/K_m = 0.0297 \mu M^{-1} \text{ min}^{-1}\)) is unlikely to happen (Fig. 6). By this way, nature has developed a smart way in 1 biosynthesis to avoid wasting energy by kinetically controlling the flow of substrates, both in recycling or salvaging shunt products (from 7 to 4), and in preventing over-oxidation (from 4 to 3). Elegantly, CrmK has evolved to have two independent catalytic residues (Y449 and E376) to enforce the recycling of 7 to 4. Similarly, in collisycin A biosynthesis, the ΔclmD2 (crmK homologue) mutant accumulated alcohol-containing products, indicating a similar recycling role of ClmD2 in the collisycin pathway.24 The homology model of ClmD2 constructed based on the structure of CrmK reveals the existence of a less bulky Val residue (V343, equivalent to I328 in CrmK) to better accommodate the extra –SCH\(_3\) group on collisycin DH, the natural substrate of ClmD2 (Fig. S10†).24

Natural product biosynthetic gene clusters (BGCs) are well known to often carry additional genes for resistance,44–46 transport,47 regulation,48,49 or carrying out natural protection group chemistry.21,50,51 Now the functional characterization of CrmK in caerulomycin biosynthesis adds a recycling mechanism that may operate in secondary metabolism. The majority of antibiotic BGCs carry genes with unknown or unassigned functions, and this work indicates that a salvage/recycling role could be a hypothesis to test for some of such genes. The existence of a “saving pathway” was reported for the biosynthesis of deoxysugars in neomycin11 and landomycins,12 putatively via a short sugar-activation pathway that was previously found for L-fucose.22 The dTDP-glucose synthase-like protein LanZ2 was speculated to be involved in this short sugar-activation pathway but further genetic and biochemical evidence was required to fully address its function.33 With the exponential growth in characterization of BGCs for secondary metabolites, it is expected that more and more enzymes with a salvaging/recycling function could be found in the future and explored for the yield improvement of secondary metabolites.

Conclusions

The substrate salvage pathway is common and plays an important role in controlling primary metabolism. However, rare examples of salvage pathways have been reported in secondary metabolism. In this study, we provided biochemical and structural insights into the flavoenzyme CrmK in CRM A 1 biosynthesis, and set an unusual example of a substrate recycling/salvaging pathway for secondary metabolites.

Acknowledgements

This work is supported in part by NSFC (31470204, 31290233, 31125001), CAS (XDA11030403), the Administration of Ocean and Fisheries of Guangdong Province (GD2012-D01-002), and by a NSERC discovery grant (436202) to R. Shi. We are grateful to the analytical facility at SCSIO for recording spectroscopic data. The X-ray crystallographic studies were performed at the Canadian Light Source, which is supported by the Canada Foundation for Innovation, Natural Sciences and Engineering.
Research Council of Canada, the University of Saskatchewan, the Government of Saskatchewan, Western Economic Diversification Canada, the National Research Council Canada, and the Canadian Institutes of Health Research. We thank the anonymous reviewers for their insightful comments.

References