1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

**Three recently diverging duplicated methyltransferases exhibit
substrate-dependent regioselectivity essential for xantholipin
biosynthesis**

Lingxin Kong[†], Qing Wang[†], Weinan Yang[†], Jufang Shen[†], Yan Li[†], Xiaoqing Zheng[†],
Lu Wang[‡], Yiwen Chu[‡], Zixin Deng[†], Yit-Heng Chooi^{§,*}, Delin You^{†,*}

[†]State Key Laboratory of Microbial Metabolism, Joint International Research
Laboratory of Metabolic and Developmental Sciences, School of Life Sciences &
Biotechnology, Shanghai Jiao Tong University, Shanghai 200030, China.

[‡]Sichuan Industrial Institute of Antibiotics, Chengdu University, Chengdu 610052,
China.

[§]School of Molecular Sciences, University of Western Australia, Perth, Western
Australia 6009, Australia.

*Correspondence: yitheng.chooi@uwa.edu.au and dlyou@sjtu.edu.cn.

Abstract

Polycyclic xanthones are characterized by highly oxygenated, angular hexacyclic frameworks and exhibit diverse biological activities. Although many of them have been isolated and chemically synthesized, the detailed biosynthetic machinery awaits discovery. Recently, xanthone construction in the xantholipin (**1**) pathway was shown to involve cryptic demethoxylation. This suggested a rationale for the existence of three O-methyltransferase (OMT) genes in the gene cluster, although there are only two O-methyl groups in the structure of **1**. Here, *in vivo* and *in vitro* analysis have been used to show that the three paralogous OMTs, XanM1-M3, introduce individual methyl groups at specific points in the biosynthetic pathway. Each OMT can to some extent take over the role of the other OMTs, although they exhibit highly substrate-dependent regiospecificity. In addition, phylogenetic analysis suggests their evolution from a common ancestor. Four putative ancestral proteins were constructed and one of them performed all the functions of XanM1-M3, while the others possessed more limited catalytic functions. The results suggest that a promiscuous common ancestor may have been able to catalyze all three reactions prior to gene duplication and functional divergence. The characterization of XanM1-M3 expands the enzyme inventory for polycyclic xanthone biosynthesis and suggests novel directed evolution approaches to diversifying natural product pathways.

Key words

polycyclic xanthone antibiotics, xantholipin, methyltransferase, methylation, biosynthesis

Introduction

Polycyclic xanthone antibiotics (Figure 1A and Figure S1) are an expanding family of aromatic polyketides characterized by a highly oxygenated, angular hexacyclic core¹. This class of natural products has intrigued numerous research groups not only for the unique chemical structures but also the diverse biological activities, such as antibacterial activities (mainly Gram-positive bacteria), antifungal activities, antitumor activities, and anticoccidial activities. The pharmacological properties of simaomicin²⁻³, kigamicin⁴, citreamicin ϵ ⁵ and kibdelones⁶ have fueled the interest in total synthesis of these molecules due to their significant antiproliferative activity. However, it remains a challenge to establish viable synthetic routes, mainly due to the partially saturated polycyclic xanthenes and the chirality on the polycyclic framework. Despite the wide interest in these molecules, the biosynthetic study of polycyclic xanthenes has been limited. Up to date, biosynthetic gene clusters for xantholipin⁷ (**1**), lysolipin⁸⁻⁹, FD-594¹⁰, arixanthomycin¹¹ (Figure 1A) have been reported. The detailed biosynthetic machinery is not fully understood, mainly due to the limited knowledge of enzymes involved in the formation of the unique molecular skeleton and in different late-stage modifications.

Xantholipin (**1**) showed strong antibiotic activity against several Gram-positive bacteria (MIC~0.02-5 $\mu\text{g/ml}$) and potent cytotoxic activity against several tumor cells (such as leukemia cell line HL60 (IC₅₀ < 0.3 mM) and the oral squamous carcinoma cell line KB (IC₅₀ < 2 nM)). Meanwhile, it exhibited an inhibitory effect on heat shock protein 47 (HSP47) and could inhibit collagen production (IC₅₀ of 27 nM)¹². The biosynthetic gene cluster of **1** has been identified and genetically verified by our group previously⁷. More recently, the characteristic xanthone nucleus in **1** has been shown to originate from the anthraquinone intermediate **3** via a novel demethoxylation-dependent enzymatic transformation¹³ (Figure 1B). This mechanism appears to be common to the biosynthesis of the polycyclic xanthone family as supported by the characterization of homologous proteins from other polycyclic

xanthone gene clusters¹³. However, several questions remain unanswered in the biosynthetic pathway specifically regarding the roles of the *O*-methyltransferase (OMT) genes. Although the final structure of **1** only contains two methyl groups, the gene cluster of **1** (*xan* cluster) encodes three homologous OMT genes. Analogously, the structure of lysolipin contains five methyl groups but the gene cluster harbours six OMT genes⁸. Among the most known common tailoring reactions involved in natural product biosynthesis, methylations are responsible for the diversification of a wide range of natural products¹⁴⁻¹⁵, improving compound bioactivity and bioavailability (e.g. polycarcin V¹⁶), modulating the programming of megasynthetases¹⁷ and contributing to protecting reactive functional groups from undesired spontaneous modifications, and thus suppress the production of off-pathway products (e.g. griseofulvin¹⁸). Methylation may also play critical roles in the programming of iterative PKS and NRPS. For example, methylation of a triketide intermediate is necessary for the downstream action of enoylreductase LovC during the biosynthesis of lovastatin¹⁹; and an α -methylation step is a prerequisite for the acyl transfer between the nonribosomal peptide synthetase (NRPS) and polyketide synthase (PKS) modules during the biosynthesis of pseurotin A¹⁷. MTs usually utilize *S*-adenosyl-methionine (SAM) as the methyl donor for the production of *O*-, *N*-, *S*-, and *C*-tethered methyl groups²⁰. Methylations can occur either at the building blocks, or on the polyketide/nonribosomal peptide aglycon, and on the sugar residues either prior to or after the glycosyl transfer steps²¹. Methylation is prevalent in most of the polycyclic xanthenes, so unveiling the machinery of methylations is important for the goal of chemical innovation and diversification of polycyclic xanthone antibiotics.

We initiated the study by analyzing genes encoding three homologous OMTs. By *in vivo* gene disruption and *in vitro* biochemical assays, we demonstrated the respective roles of the three OMTs and found that they exhibited functional overlaps with high substrate-dependent regio-specificities. A putative ancestral protein displaying the function of all three OMTs was identified. It appears that such an

ancestral OMT may have displayed catalytic promiscuity before divergence. These findings illustrate the role of gene duplication of tailoring genes in generating structural diversity.

Results and Discussion

The *xan* Cluster Encodes Three Paralogous OMTs.

Bioinformatics analysis of the *xan* cluster revealed that the three genes *xanM1-M3* encode three OMTs XanM1-M3. The three OMTs each comprise about 340 amino acids, suggesting them as members of metal-independent class II OMTs¹⁴. The bioinformatics analysis assigned them to the methyltransf_2 superfamily (Pfam 00891) with the highly conserved glycine-rich sequence E/DXGXGXG within the proposed SAM binding region²² (Figure 2A). The three OMTs exhibited sequence similarity with several class II OMTs, such as DnrK²³ (30% identity with XanM2), MmcR²⁴ (33% identity with XanM1) and NcsB1²⁵ (30% identity with XanM1). Meanwhile, the three OMTs exhibited higher sequence identity with each other (XanM1 & XanM2 41%, XanM1 & XanM3 44%, XanM2 & XanM3 40%) (Table S1). Typically, MTs share little sequence identity but incorporate a highly conserved structural fold²². The secondary structure of the three OMTs XanM1-M3 was analyzed by PSIPRED as shown in Figure 2A. The three OMTs possess the prototypical core SAM-MT fold, which incorporates alternating α helices and β strands topologically arranged as α - β -(α - β)₅- β ²². Meanwhile, additional modifications can be found within the SAM-MT fold regions of the three OMTs (Figure 2A). The phylogenetic analysis (Figure 2B) showed that XanM1-M3 are closely related to each other and homologous proteins of XanM1-3 can be found in other characterized polycyclic xanthone antibiotic biosynthetic gene clusters, such as lysolipin⁸ and arixanthomycin¹¹ (Figure1, Figure S1 and Table 1). The lysolipin cluster (*llp* cluster) has four such paralogous OMTs (LlpMII-IV). Especially LlpMII, LlpMIII and LlpMIV from the *llp* cluster showed high similarity with XanM3, XanM2 and XanM1, respectively, and occupied the same clade with homologous OMTs (Figure 2B). However, the arixanthomycin cluster (*arx*

cluster) encodes only one (Arx34) methyltransferase and is at the basal of these paralogous OMTs in lysolipin and **1** gene cluster. The informatics analysis suggested similar functions of homologous OMTs for the biosynthesis of polycyclic xanthenes (such as **1**, lysolipin and arixanthomycins).

Deciphering the *in vivo* Functions of the Three OMTs.

To determine the *in vivo* role of the three OMTs, these genes were deleted separately (Figure S2). Liquid chromatography-mass spectrometry (LC-MS) analysis and high-resolution accurate mass spectrometry Q-TOF MS (HR-MS) were used to analyze the fermentation products. Surprisingly, the $\Delta xanM1$ mutant abolished the ability to produce **1** and did not accumulate any new intermediates (Figure 3A). Since *xanM1* is within a group of 14 genes that form an operon, inactivation of downstream genes due to polar effects on transcription might have contributed to the phenotype of the $\Delta xanM1$ mutant. To rule this out, we complemented the mutant by expressing *xanM1* under the control of *permE** (Figure S3), which partially restored the production of **1** (Figure 3A). This suggests that deletion of *xanM1* was primarily responsible for the loss of **1** production in $\Delta xanM1$ mutant. We hypothesized that the lack of accumulated intermediate could be due to a methyl group that is required for the release of the mature xanthone polyketide product from the PKS, or the unmethylated hypothetical intermediate **2** may be unstable *in vivo*. Therefore, XanM1 is likely involved in an indispensable O-methylation occurring early in the pathway, possibly when the nascent polyketide product is still tethered on the ACP of the PKS complex (ACP-**2**→**3**) (Figure 1).

The $\Delta xanM2$ mutant accumulated a yellow compound **8** (**1** was yellow) (Figure 3A) as the major product with small quantities of **1**. The UV-Vis spectrum suggested that **8** contains a xanthone chromophore like **1** (Figure 3B), while the relative isotopic abundance (9:3:3:1) of **8** as detected by LC-MS indicated the existence of a chlorine atom (Figure 3C). HR-MS analysis showed that **8** has an *m/z* of 508.0800 with the predicted molecular formula $C_{26}H_{18}ClNO_8$ (Calc *m/z* 508.0794) (Figure S4C).

Extensive NMR spectra analysis of **8** (Table S2) showed that the structure resembled **1**, but the methylenedioxy-bridge was missing and the ring E remained at the lower oxidation state in **8** as shown in Figure 1. The accumulation of **8** suggests that XanM2 is primarily responsible for the methylation of the C13-OH in **8**. The small amount of **1** present in the $\Delta xanM2$ mutant (Figure 3A), however, suggests that another enzyme may partially take over its function. The increased production of **1** in the complementary strain $\Delta xanM2::xanM2$ confirmed the importance of this methylation (Figure 3A).

The $\Delta xanM3$ mutant accumulated compounds **4**, **3** and **3f** (Figure 3A). The accumulation of **4** in the $\Delta xanM3$ mutant and the reported methylation of **4** by XanM3 both supported the proposed function of XanM3 for the remethylation of the C17-OH in **1** biosynthesis¹³. Compound **3f** is an off-pathway C17-OH anthraquinone, which cannot be processed by XanO4 to the xanthone **4** and has previously been reported in the *in vitro* reaction system of **3** with XanO4-R45A, where the arginine residue important for FAD-binding has been mutated^{13, 26}. These findings here supported the previous findings on the role of XanM3 in the remethylation of C17-OH (Figure 1).

The *in vitro* Functional Overlaps of OMTs with High Substrate-dependent Regio-specificities.

Based on the above results, the chief role of the three OMTs could be deduced. To characterize the methylation activities *in vitro*, we expressed each OMT in *Escherichia coli* BL21(DE3)/pLysE and purified them to near homogeneity (Figure S5). First, we tested the ability of the three OMTs to methylate C13-OH in **8**. XanM2 could methylate **8** to give a new product **9** (Figure 4A). The identical UV-Vis spectrum of **9** with **8** (Figure 4A) and the 14 Da mass shift for **9** from **8** supported the loading of a methyl group (Figure 4A). The HR-MS analysis of **9** (Figure S4C) showed that it has an *m/z* of 522.1010 with the predicted molecular formula $C_{27}H_{20}ClNO_8$ (Calc *m/z* 522.0956), which was identical with the compound accumulated in the previously reported $\Delta xanO2$ strain⁷. When $^{13}CH_3$ -SAM was used to replace $^{12}CH_3$ -SAM in the *in*

vitro enzymatic reaction systems, MS analysis of the reaction mixtures revealed an increase in mass of 1 Da for **9**, confirming the incorporation of a single labeled carbon atom (Figure S4D). Considering the homology of the three OMTs, compound **8** was also tested as a substrate for the other two OMTs. Incubation of **8** with XanM1 and XanM3 gave partial methylation of **8**, albeit the conversion is lower compared to XanM2 (Figure 4A). This activity of XanM1 and XanM3 might explain the slight production of **1** in the $\Delta xanM2$ mutant (Figure 3A).

We next tested the ability of each of the three OMTs in methylating the C17-OH. XanM3 fully converted **3f** into **3** *in vitro* (Figure 4B). XanM2 could convert **3f** to **3** to a small extent, but XanM1 was not active (Figure 4B). When **3f** was used, as the substrate for a one-pot reaction with XanO4, XanM3 converted all the substrates to **5**, while XanM2 yielded a small amount of **5**, and XanM1 catalyzed no conversion (Figure S6A). In our previous work, we demonstrated that xanthone formation by XanO4 is dependent on the demethoxylation of C17-OMe, as it could not proceed with **3f** (which lack C17-OMe) but only with the 17-O-methoxy **3**¹³. This suggests that XanM3, and to some extent XanM2, could redirect any **3f**, which may be formed as a result of the partial conversion of **3** by XanO4, back into the pathway towards **1** by remethylating the C17-OH to form **3** again.

We then next used compound **4** as a substrate for the three OMTs. When **4** was used as a substrate for XanM3, almost all of **4** was transformed into **5** within 1 hr at 30 °C (Figure 4C). For the same reaction, XanM2 could yield **5** as well, but not XanM1 (Figure 4C). The conversion steps from **3** to **4** and then **4** to **5** could be observed in the one-pot reaction systems of XanM2/XanO4 and XanM3/XanO4, but not XanM1/XanO4 (Figure S6B). Meanwhile, a similar process could be monitored when titrating the one-pot reaction across increasing SAM concentration (Figure S7). Thus, we can conclude that the primary function of XanM3 is for methylation or remethylation of the C17-OH on **4** and **3f**, and that XanM2 has a partial crossover of

1
2
3 this function in the pathway. These functional crossovers between the OMTs are
4 summarized in Table 1. An intriguing observation from these *in vitro* assays is that
5 although both XanM2 and XanM3 can catalyze methylation at both C13-OH and
6 C17-OH on the angular xanthone core, in the one-pot assays with XanO4 and **3/3f**, no
7 C13-OH derivative of **5** is observed (Figure S7). This suggests that the regioselectivity
8 of XanM2 and XanM3 is highly dependent on the substrate. To further corroborate
9 this finding, we tested all three OMTs with **6** containing the C17-MeO and chlorine
10 substitution as substrate, and none of the OMT could methylate C13-OH. In fact,
11 except **8**, XanM1-3 cannot methylate the C13-OH on any of the C17-OMe-containing
12 substrates (**3**, **5** and **6**) (Figure S8). This suggests that the methylation on C13-OH by
13 XanM2/M3 is highly selective toward the pyridone moiety, which is present in **8**.

24
25
26 Based on the observations above, XanM2 is more specialized in methylation of
27 C13-OH on **8**, XanM3 is best in methylation of C17-OH on **4** and **3f**, while XanM1 is
28 likely involved in an indispensable O-methylation that occurred early in the pathway,
29 possibly when the nascent polyketide product is still tethered on the ACP of the PKS
30 complex (Figure 1 and Figure 5). At the same time, XanM2 and XanM3 could carry
31 out the function of each other to a small extent and XanM1 also retained some
32 function of XanM2. Taken together, all the three OMTs are multifunctional and they
33 each possess at least one major catalytic activity and one or two minor catalytic
34 activities that overlap with other OMTs, thus exhibiting redundancy to a certain extent.
35 This redundancy could serve as a mechanism to ensure that all intermediates in the
36 pathway are directed toward the production of the final product **1**. Perhaps the most
37 intriguing discovery from this study is the substrate-dependent regioselectivity of
38 these multifunctional OMTs. This required that the different substrates be positioned
39 in a different manner in the active sites of the OMTs. To the best of our knowledge,
40 this is unprecedented for OMTs in microbial secondary metabolite biosynthetic
41 pathways. Specific regiospecific methylation catalyzed by separate and consecutive
42 OMTs have been reported in griseofulvin¹⁸ and elloramycin biosynthesis²⁷,

respectively. However, substrate-dependent regioselectivity appears to be more commonly reported for multifunctional oxygenases. Many P450s have been demonstrated to possess a fascinating ability to catalyze multiple reactions in a highly ordered fashion at separate sites starting with a single substrate molecule, such as MycG²⁸, AurH²⁹, TamI³⁰ and GfsF³¹. It has been reported that the substrate-dependent regioselectivity exhibited by some of these P450s is due to the variation in binding conformation of the different substrates in the active site and changes in the protein binding pockets as well when bound to different substrates^{29, 32}. It is possible that the multifunctional OMTs in this study use similar strategies to alter the regioselectivity. Multiple methylations have been reported in natural products biosynthetic pathways, such as that in aminoglycoside gentamicin³³ and spinosyn³⁴. The methyltransferases GenN, GenD1, and GenK involved in the biosynthesis of the aminoglycoside gentamicin comprise a methylation network in which early intermediates are ectopically modified³³. The OMTs SpnH, SpnI, and SpnK are responsible for the methylations of the 4'-, 2'-, and 3'-OH in rhamnose, respectively. It has been found that only one route is productive for the permethylation of the rhamnose moiety³⁴. Meanwhile, minor activities of tailoring enzymes have been reported to contribute to the branched parallel biosynthetic pathways³⁵⁻³⁷, so similarly, it is possible that the minor activities of the three OMTs, XanM1-M3, might help to direct branched pathway products back to the biosynthetic pathway toward **1**.

The substrate-specific OMTs are diversified from a versatile ancestor.

Ancestral sequence reconstruction (ASR) has been used to infer the evolutionary precursors of extant functional proteins³⁸, reveal the evolution process of altered enzyme substrate specificity³⁹, explore the evolution route for loss of enzymatic activity⁴⁰, investigate the key transition from noncatalytic ancestor to active enzyme⁴¹ and many of other protein features⁴²⁻⁴⁴. The observations of OMTs mentioned above raised the interesting question the origin of a homologous set of enzymes that exhibit catalytic specialisation yet retain minor catalytic overlaps. The phylogenetic analysis

of OMTs suggested that XanM1-M3 are likely evolved from a common multifunctional ancestor shared with Arx34. Meanwhile, similar OMT gene amplification derived from the same common ancestor as XanM1-M3 occurred in lysolipin biosynthetic gene cluster as well. To investigate whether these three OMTs (XanM1-3) might represent duplicated genes undergoing functional divergence, ASR was conducted for XanM1-M3, LlpMII-MIII, LpMIV and LlpMVI using FastML⁴⁵. Four of the most probable ancestral proteins inferred from the analysis, named as aXLpMI-MIV, were expressed and purified (Figure S5). Using the compounds **3f**, **3**, **4** and **8** as substrate, these four ancestral proteins were tested for methylation activities. As can be seen from the *in vitro* assay results (Figure 6 and Figure S9), only aXLpMIII can catalyze all of the reactions (**3f**→**3**, **4**→**5**, and **8**→**9**), suggesting that this protein is one of the more promiscuous ancestral OMTs that possess all of the activities XanM1-M3 collectively possess. On the other hand, all the other three ancestral OMTs, aXLpMI, MII and MIV, can convert **4** to **5** at a level that is similar to aXLpMIII but not **3f** to **3** or **8** to **9** (Figure S9). Meanwhile, integrative plasmids expressing *aXLpMI-aXLpMIV* from the *permE** were introduced into the $\Delta xanM1$ mutant, respectively, for the construction of complementation strains (Figure S10). As can be seen from the data shown in Figure S10, all the resultant four complementation strains partially restored production of **1**, supporting the conclusion the ancestral OMTs possess the putative indispensable function of ACP-**2**→**3** (Figure 5). This is in line with the phylogenetic analysis as C17-OMe is a common feature of arixanthomycins, lysolipin and **1**. Specifically, we expect that the common ancestor of XanM1-M3 to be quite similar to Arx34, which -based on the structure of arixanthomycin- should be responsible for the 17-O-methylation.

The finding that aXLpMIII possesses all three catalytic activities suggests that the common ancestor of this group of homologous OMTs from *xan* and *llp* cluster may have evolved new catalytic activities (i.e. **3f**→**4** and **8**→**9**) before further functional divergence. If so, these findings might fit well with the

innovation-amplification-divergence (IAD) model, in which the ancestral gene first acquired a new secondary minor function that confers selective advantages (Innovation), the selection for the minor function resulted in duplications of the genes (Amplification), which allowed subsequent functional specialisation (Divergence) ⁴⁶. We proposed that, similar to the IAD model, the common ancestor of the *xan//lp* OMTs (aXLpM) first had undergone an innovation phase, which led to new minor secondary catalytic activities (i.e. **3f**→**4** and **8**→**9**, as in aXLpMIII) that conferred a selective advantage to the host. With the new minor activities under selection pressure, this led to gene amplification, and the multiple copies of the new multifunctional OMTs provided the opportunities for an individual copy to diverge into the more specialized OMTs as we see today in XanM1-M3 (Figure 7). Interestingly, the *//p* cluster encodes an additional OMT, LlpMIV, within the XanM1-M3 clade, while each of the other three Llp-OMTs is paired with XanM1-M3 (Figure 4A and Figure S1). Based on the structure of **1** and lysolipin, LlpMIV could be responsible for methylation of C19-OH, which is absent in **1**. This suggests that there might be an additional round of IAD that led to *//pMIV* with the acquisition of a gene encoding C19 hydroxylase in the lysolipin producer or the *//pMIV* homolog in the ancestral *xan* cluster was lost due to the loss of a C19 hydroxylase gene. This could be tested using the ancestral OMTs on C19-OH lysolipin intermediates in the future. However, the possibility that the common precursor might be simply nonspecific or promiscuous before gene duplication still can't be excluded.

Conclusion

In summary, this study clarified the respective roles of three homologous multifunctional OMTs for the biosynthesis of **1**. The three OMTs exhibited functional overlaps and highly substrate-dependent regioselectivity. ASR was successfully conducted and a universal ancestor possessing all the functions of XanM1-M3 was identified. The phylogenetic analysis and ASR support that XanM1-M3 might be diversified from a common ancestor via a mechanism similar to the proposed IAD

model. This study implies that IAD of tailoring enzymes contributes to the evolution of structural diversification in natural product biosynthetic pathways. Combining IAD with directed pathway evolution (for example, involving multiple copies of genes encoding ancestral promiscuous enzyme) could be a potential approach to generate new structural diversity and enhancing compound bioactivities.

Associated Content

Supporting information

The Supporting Information is available free of charge *via* the internet. Additional data on the methods for plasmids construction, generation of *Streptomyces* mutants, sequence alignment and phylogenetic analysis of the OMTs, *in vitro* biochemical assays, reconstruction of the ancestral proteins are included.

Accession number

Genbank accession numbers for XanM1, XanM2, XanM3 and XanO4 are ADE22294.1., ADE22299.1., ADE22301.1 and ADE22300.1, respectively.

Author information

Corresponding Authors

*E-mail: yitheng.chooi@uwa.edu.au.

*E-mail: dlyou@sjtu.edu.cn.

ORCID

Lingxin Kong: 0000-0002-0382-9282

Yit-Heng Chooi: 0000-0001-7719-7524

Delin You: 0000-0003-1644-9825

Author contributions

L.X.K., Y.H.C., and D.L.Y designed the research; L.X.K., Q.W, J.F.S., Y.L., X.Q.Z. performed the research; L.X.L., Y.H.C., and D.L.Y. analyzed the data; L.W. and Y.W.C. analyzed the NMR spectra; L.X.K., Y.H.C., and D.L.Y. wrote the manuscript.

Notes

The authors declare no competing financial interest.

Acknowledgements

This work was supported by grants from the National Natural Science Foundation of China (31630002, 31700029, 31770038, 31470183, 21661140002 and 31170085); the Ministry of Science and Technology; Shanghai Pujiang Program from the Shanghai Municipal Council of Science and Technology (12PJD021); China Postdoctoral Science Foundation (2017M620151); Australian Research Council (ARC) Future Fellowship (FT160100233).

References

- (1) Winter, D. K., Sloman, D. L. and Porco, J. A. (2013) Polycyclic xanthone natural products: structure, biological activity and chemical synthesis. *Nat. Prod. Rep.* 30, 382-391.
- (2) Arai, M., Sato, H., Kobayashi, H., Suganuma, M., Kawabe, T., Tomoda, H. and Ōmura, S. (2004) Selective inhibition of bleomycin-induced G2 cell cycle checkpoint by simaomicin α . *Biochem. Biophys. Res. Commun.* 317, 817-822.
- (3) Ishiyama, A., Otoguro, K., Namatame, M., Nishihara, A., Furusawa, T., Takahashi, Y., Masuma, R., Shiomi, K. and Omura, S. (2008) Simaomicin alpha: effects on the cell cycle of synchronized, cultured *Plasmodium falciparum*. *J. Antibiot.* 61, 254-257.
- (4) Turner, P. A., Griffin, E. M., Whatmore, J. L. and Shipman, M. (2011) Tetrahydroxanthones by sequential Pd-catalyzed C–O and C–C bond construction and use in the identification of the “antiausterity” pharmacophore of the kigamicins. *Org. Lett.* 13, 1056-1059.
- (5) Zhu, T., Cheng, X., Liu, Y., Deng, Z. and You, D. (2013) Deciphering and engineering of the final step halogenase for improved chlortetracycline biosynthesis in industrial *Streptomyces aureofaciens*. *Metab. Engin.* 19, 69-78.
- (6) Sloman, D. L., Mitasev, B., Scully, S. S., Beutler, J. A. and Porco, J. A. (2011) Synthesis and biological evaluation of ABCD ring fragments of the kibdelones. *Angew. Chem. Int. Ed.* 50, 2511-2515.

- (7) Zhang, W. K., Wang, L., Kong, L. X., Wang, T., Chu, Y., Deng, Z. X. and You, D. L. (2012) Unveiling the post-PKS redox tailoring steps in biosynthesis of the Type II polyketide antitumor antibiotic xantholipin. *Chem. Biol.* **19**, 422-432.
- (8) Lopez, P., Hornung, A., Welzel, K., Unsin, C., Wohlleben, W., Weber, T. and Pelzerb, S. (2010) Isolation of the lysolipin gene cluster of *Streptomyces tendae* Tü 4042. *Gene* **461**, 5-14.
- (9) Bockholt, H., Udvarnoki, G., Rohr, J., Mocek, U., Beale, J. M. and Floss, H. G. (1994) Biosynthetic studies on the xanthone antibiotics lysolipins X and I. *J. Org. Chem.* **59**, 2064-2069.
- (10) Kudo, F., Yonezawa, T., Komatsubara, A., Mizoue, K. and Eguchi, T. (2011) Cloning of the biosynthetic gene cluster for naphthoxanthene antibiotic FD-594 from *Streptomyces* sp. TA-0256. *J. Antibiot.* **64**, 123-132.
- (11) Kang, H.-S. and Brady, S. F. (2014) Arixanthomycins A-C: phylogeny-guided discovery of biologically active eDNA-derived pentangular polyphenols. *ACS Chem. Biol.* **9**, 1267-1272.
- (12) Terui, Y., Chu, Y., Li, J. Y., Ando, T., Yamamoto, H., Kawamura, Y., Tomishima, Y., Uchida, S., Okazaki, T. and Munetomo, E. (2003) Xantholipin, a novel Inhibitor of HSP47 gene expression produced by *Streptomyces* sp. *Tetrahedron Lett.* **34**, 5427-5430.
- (13) Kong, L. X., Zhang, W. K., Chooi, Y.-H., Wang, L., Cao, B., Deng, Z. X., Chu, Y.W. and You, D. L. (2016) A multifunctional monooxygenase XanO4 catalyzes xanthone formation in xantholipin biosynthesis via a cryptic demethoxylation. *Cell Chem. Biol.* **23**, 508-516.
- (14) Cho, J. H., Park, Y., Ahn, J. H., Lim, Y. and Rhee, S. (2008) Structural and functional insights into O-methyltransferase from *Bacillus cereus*. *J. Mol. Biol.* **382**, 987-997.
- (15) Cabry, M. P., Offen, W. A., Saleh, P., Li, Y., Winzer, T., Graham, I. A. and Davies, G. J. (2019) Structure of *Papaver somniferum* O-methyltransferase 1 reveals initiation

of noscapine biosynthesis with implications for plant natural product methylation. *ACS Catal.* **9**, 3840-3848.

(16) Chen, J.-M., Shepherd, M. D., Horn, J., Leggas, M. and Rohr, J. (2014) Enzymatic methylation and structure-activity-relationship studies on polycarcin V, a gilvocarcin-type antitumor agent. *ChemBioChem* **15**, 2729-2735.

(17) Zou, Y., Xu, W., Tsunematsu, Y., Tang, M., Watanabe, K. and Tang, Y. (2014) Methylation-dependent acyl transfer between polyketide synthase and nonribosomal peptide synthetase modules in fungal natural product biosynthesis. *Org. Lett.* **16**, 6390-6393.

(18) Cacho, R. A., Chooi, Y.-H., Zhou, H. and Tang, Y. (2013) Complexity generation in fungal polyketide biosynthesis: a spirocycle-forming P450 in the concise pathway to the antifungal drug griseofulvin. *ACS Chem. Biol.* **8**, 2322-2330.

(19) Ames, B. D., Nguyen, C., Bruegger, J., Smith, P., Xu, W., Ma, S., Wong, E., Wong, S., Xie, X., Li, J. W. H., Vederas, J. C., Tang, Y. and Tsai, S.-C. (2012) Crystal structure and biochemical studies of the trans-acting polyketide enoyl reductase LovC from lovastatin biosynthesis. *Proc. Natl. Acad. Sci. U. S. A.* **109**, 11144-11149.

(20) Liscombe, D. K., Louie, G. V. and Noel, J. P. (2012) Architectures, mechanisms and molecular evolution of natural product methyltransferases. *Nat. Prod. Rep.* **29**, 1238-1250.

(21) Olano, C., Mendez, C. and Salas, J. A. (2010) Post-PKS tailoring steps in natural product-producing actinomycetes from the perspective of combinatorial biosynthesis. *Nat. Prod. Rep.* **27**, 571-616.

(22) Martin, J. L. and Mcmillan, F. M. (2002) SAM (dependent) I AM: the S-adenosylmethionine-dependent methyltransferase fold. *Curr. Opin. Struc. Biol.* **12**, 783-793.

(23) Jansson, A., Koskiniemi, H., Mäntsälä, P., Niemi, J. and Schneider, G. (2004) Crystal structure of a ternary complex of DnrK, a methyltransferase in daunorubicin biosynthesis, with bound products. *J. Biol. Chem.* **279**, 41149-41156.

- (24) Singh, S., Chang, A., Goff, R. D., Bingman, C. A., Grüşchow, S., Sherman, D. H., Phillips, G. N., Jr. and Thorson, J. S. (2011) Structural characterization of the mitomycin 7-O-methyltransferase. *Proteins* 79, 2181-8.
- (25) Cooke, H. A., Guenther, E. L., Luo, Y., Shen, B. and Bruner, S. D. (2009) Molecular basis of substrate promiscuity for the SAM-dependent O-methyltransferase NcsB1, involved in the biosynthesis of the enediyne antitumor antibiotic neocarzinostatin. *Biochemistry* 48, 9590-9598.
- (26) Beam, M. P., Bosserman, M. A., Noinaj, N., Wehenkel, M. and Rohr, J. (2009) Crystal structure of Baeyer-Villiger monooxygenase MtmOIV, the key enzyme of the mithramycin biosynthetic pathway. *Biochemistry* 48, 4476-4487.
- (27) Patallo, E. P., Blanco, G., Fischer, C., Braña, A. F., Rohr, J., Méndez, C. and Salas, J. A. (2001) Deoxysugar methylation during biosynthesis of the antitumor polyketide elloramycin by *Streptomyces olivaceus*: characterization of three methyltransferase genes. *J. Biol. Chem.* 276, 18765-18774.
- (28) Anzai, Y., Tsukada, S.-I., Sakai, A., Masuda, R., Harada, C., Domeki, A., Li, S., Kinoshita, K., Sherman, D. H. and Kato, F. (2012) Function of cytochrome P450 enzymes MycCl and MycG in *Micromonospora griseorubida*, a producer of the macrolide antibiotic mycinamicin. *Antimicrob. Agents Chemother.* 56, 3648-3656.
- (29) Zocher, G., Richter, M. E. A., Mueller, U. and Hertweck, C. (2011) Structural fine-tuning of a multifunctional cytochrome P450 monooxygenase. *J. Am. Chem. Soc.* 133, 2292-2302.
- (30) Carlson, J. C., Li, S., Gunatilleke, S. S., Anzai, Y., Burr, D. A., Podust, L. M. and Sherman, D. H. (2011) Tirandamycin biosynthesis is mediated by co-dependent oxidative enzymes. *Nat. Chem.* 3, 628-633.
- (31) Miyanaga, A., Takayanagi, R., Furuya, T., Kawamata, A., Itagaki, T., Iwabuchi, Y., Kanoh, N., Kudo, F. and Eguchi, T. (2017) Substrate recognition by a dual-function P450 monooxygenase GfsF involved in FD-891 biosynthesis. *ChemBioChem* 18, 2179-2187.

- (32) Li, S., Tietz, D. R., Rutaganira, F. U., Kells, P. M., Anzai, Y., Kato, F., Pochapsky, T. C., Sherman, D. H. and Podust, L. M. (2012) Substrate recognition by the multifunctional cytochrome P450 MycG in mycinamicin hydroxylation and epoxidation reactions. *J. Biol.Chem.* 287, 37880-37890.
- (33) Li, S., Guo, J., Reva, A., Huang, F., Xiong, B., Liu, Y., Deng, Z., Leadlay, P. F. and Sun, Y. (2018) Methyltransferases of gentamicin biosynthesis. *Proc. Natl. Acad. Sci. U. S. A.* 115, 1340-1345.
- (34) Kim, H. J., White-Phillip, J. A., Ogasawara, Y., Shin, N., Isiorho, E. A. and Liu, H. W. (2010) Biosynthesis of spinosyn in *Saccharopolyspora spinosa*: synthesis of permethylated rhamnose and characterization of the functions of SpnH, SpnI, and SpnK. *J. Am. Chem. Soc.* 132, 2901-2903.
- (35) Ban, Y. H., Shinde, P. B., Hwang, J. Y., Song, M. C., Kim, D. H., Lim, S. K., Sohng, J. K. and Yoon, Y. J. (2013) Characterization of FK506 biosynthetic intermediates involved in post-PKS elaboration. *J. Nat. Prod.* 76, 1091-1098.
- (36) Chen, D., Zhang, L., Pang, B., Chen, J., Xu, Z., Abe, I. and Liu, W. (2013) FK506 maturation involves a cytochrome P450 protein-catalyzed four-electron C-9 oxidation in parallel with a C-31 O-methylation. *J. Bacteriol.* 195, 1931-1939.
- (37) Gao, S. S., Hothersall, J., Wu, J., Murphy, A. C., Song, Z., Stephens, E. R., Thomas, C. M., Crump, M. P., Cox, R. J., Simpson, T. J. and Willis, C. L. (2014) Biosynthesis of mupirocin by *Pseudomonas fluorescens* NCIMB 10586 involves parallel pathways. *J. Am. Chem. Soc.* 136, 5501-5507.
- (38) Gumulya, Y. and Gillam, E. M. J. (2017) Exploring the past and the future of protein evolution with ancestral sequence reconstruction: the 'retro' approach to protein engineering. *Biochemical J.* 474, 1-19.
- (39) Steindel, P. A., Chen, E. H., Wirth, J. D. and Theobald, D. L. (2016) Gradual neofunctionalization in the convergent evolution of trichomonad lactate and malate dehydrogenases. *Protein Sci.* 25, 1319-1331.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

(40) Anderson, D. P., Whitney, D. S., Hanson-Smith, V., Woznica, A., Campodonico-Burnett, W., Volkman, B. F., King, N., Thornton, J. W. and Prehoda, K. E. (2016) Evolution of an ancient protein function involved in organized multicellularity in animals. *eLife* 5, e10147.

(41) Kaltenbach, M., Burke, J. R., Dindo, M., Pabis, A., Munsberg, F. S., Rabin, A., Kamerlin, S. C. L., Noel, J. P. and Tawfik, D. S. (2018) Evolution of chalcone isomerase from a noncatalytic ancestor. *Nat. Chem. Biol.* 14, 548-555.

(42) Lynch, V. J., May, G. and Wagner, G. P. (2011) Regulatory evolution through divergence of a phosphoswitch in the transcription factor CEBPB. *Nature* 480, 383.

(43) Natarajan, C., Hoffmann, F. G., Weber, R. E., Fago, A., Witt, C. C. and Storz, J. F. (2016) Predictable convergence in hemoglobin function has unpredictable molecular underpinnings. *Science* 354, 336-339.

(44) Tzul, F. O., Vasilchuk, D. and Makhatadze, G. I. (2017) Evidence for the principle of minimal frustration in the evolution of protein folding landscapes. *Proc. Natl. Acad. Sci. U. S. A.* 114, E1627-E1632.

(45) Ashkenazy, H., Penn, O., Doron-Faigenboim, A., Cohen, O., Cannarozzi, G., Zomer, O. and Pupko, T. (2012) FastML: a web server for probabilistic reconstruction of ancestral sequences. *Nucleic Acids Res.* 40, W580-W584.

(46) Näsvall, J., Sun, L., Roth, J. R. and Andersson, D. I. (2012) Real-time evolution of new genes by innovation, amplification, and divergence. *Science* 338, 384-387.

Figures

Figure 1. Molecular structures of representative polycyclic xanthones (A) and biosynthetic pathway of 1 (B).

Figure 2. Informatics analysis of OMTs within biosynthetic gene cluster of **1**. (A) Multiple sequence alignment of OMTs XanM1-M3. The alignment was generated using ClustalX. Light and dark shading indicated $\geq 70\%$ similarity and identity at that position, respectively. Secondary structure elements indicated below the alignment showed the conserved SAM binding motif and are based on the crystal structure of (PDB: 2IP2), with α -helices represented as cylinders, β -strands as arrows. The conserved SAM binding motif was marked with green rectangles. The proposed core

SAM-MT fold was indicated with gray rectangle and the varied 'modification' was depicted with wavy line. (B) The phylogenesis analysis of XanM1-M3 with other homologous OMTs. The neighbor-joining phylogenetic tree was constructed by using MEGA V5.10 with 500 bootstrap replicates. The scale 0.2 is the genetic distance.

Figure 3. HPLC analysis and characterization of fermentation products in the corresponding OMT gene deletion and complemented mutants. (A) HPLC profile of the fermentation products in mutants. $\Delta xanM1$ (ii), $\Delta xanM1::xanM1$ (iii), $\Delta xanM2$ (iv), $\Delta xanM2::xanM2$ (v) and $\Delta xanM3$ (vi) were analysed with WT strain (i) as control. The standard compounds 1, 3f, 3 and 4 were also used as control (vii). (B) UV-Vis spectra of compound 8. (C) MS analysis of compound 8.

Figure 4. *In vitro* methylations catalyzed by the three OMTs. **(A)** HPLC profile of the methylation of the hydroxyl group at C13 in compound **8** by XanM1 (ii), XanM2 (iii) and XanM3 (iv), respectively. Standard compound **8** (i) was used as control. **(B)** HPLC profile of the methylation of the hydroxyl group at C17 in compound **3f** by XanM1 (i), XanM2 (iii) and XanM3 (ii), respectively. Standard compounds **3f** and **3** (iv) were used as controls. **(C)** HPLC profile of the methylation of the hydroxyl group at C17 in compound **4** by XanM1 (i), XanM2 (ii) and XanM3 (iii), respectively. Standard compounds **4** and **5** (i) were used as controls. The data shown here represent results derived from three different independent experiments.

Figure 5. The revised biosynthetic pathway of **1**. The methoxy groups at C17 and C13 in compounds **3**, **5** and **9** are highlighted with yellow ellipses, and the hydroxyl group at C17 in compound **4** is marked with a green ellipse. The carbonyl group of C15 in compound **3** is highlighted with a red ellipse, while the corresponding oxygen atom in compound **4** is marked in blue ellipse.

Figure 6. *In vitro* activity assay of ancestral protein aXLpMIII. HPLC profiles of the reactions catalyzed by aXLpMIII with compounds **3f**, **4** and **8**. Reactions catalyzed by XanM3 with compounds **3f**, **4** and XanM2 with compound **8** were used as controls. Notice that the different retention time of the peaks from other figures resulted from the change of column and parameters used in HPLC manipulations. The data shown here represent results derived from three different independent experiments.

Figure 7. The proposed evolutionary innovation-amplification and divergence (IAD) model for the OMTs.

Table 1. Compounds accumulated in mutants and reactions catalyzed by OMTs.

* indicates the first methylation of C17-OH on a proposed potential ACP-tethered precursor of compound **3**, like ACP-2.

Strains/OMTs	Compounds accumulated	Catalysing reactions
$\Delta xanM1/XanM1$	-	*ACP-2 $\rightarrow$ 3 ? [C17-O]; 8 $\rightarrow$ 9 [C13-O]
$\Delta xanM2/XanM2$	1, 8	8 $\rightarrow$ 9 [C13-O] (3f $\rightarrow$ 3 ; 4 $\rightarrow$ 5) [C17-O]
$\Delta xanM3/XanM3$	3, 3f, 4	3f $\rightarrow$ 3 ; 4 $\rightarrow$ 5 [C17-O]; (8 $\rightarrow$ 9) [C13-O]

Table 2. Information for OMTs involved in this study.

Genes	Biosynthetic gene cluster	Source organism	GenBank number
<i>xanM1, M2, M3</i>	xantholipin (<i>xan</i>) cluster	<i>S. flavogriseus</i> SIIA-A02191	GQ421798.1
<i>lIpMII, MIII-IV, MVI</i>	lysolipin (<i>lIp</i>) cluster	<i>S. tendae</i> Tu 4042	AM492533.1
<i>arx34</i>	arixanthomycin (<i>arx</i>) cluster	Uncultured bacterium	KF931341.1

TOC graphic