

Evolution-inspired engineering of anthracycline methyltransferases

Pedro Dinis ¹, Heli Tirkkonen ¹, Benjamin Nji Wandji , Vilja Siitonen, Jarmo Niemi , Thadée Grocholski and Mikko Metsä-Ketelä ^{*}

Department of Life Technologies, University of Turku, BioCity, Tykistökatu 6, FIN-20014 Turku, Finland

^{*}To whom correspondence should be addressed: E-mail: mianme@utu.fi

¹P.D. and H.T. contributed equally to this work.

Edited By: Christopher Dupont

Abstract

Streptomyces soil bacteria produce hundreds of anthracycline anticancer agents with a relatively conserved set of genes. This diversity depends on the rapid evolution of biosynthetic enzymes to acquire novel functionalities. Previous work has identified S-adenosyl-l-methionine-dependent methyltransferase-like proteins that catalyze 4-O-methylation, 10-decarboxylation, or 10-hydroxylation, with additional differences in substrate specificities. Here we focused on four protein regions to generate chimeric enzymes using sequences from four distinct subfamilies to elucidate their influence in catalysis. Combined with structural studies we managed to depict factors that influence gain-of-hydroxylation, loss-of-methylation, and substrate selection. The engineering expanded the catalytic repertoire to include novel 9,10-elimination activity, and 4-O-methylation and 10-decarboxylation of unnatural substrates. The work provides an instructive account on how the rise of diversity of microbial natural products may occur through subtle changes in biosynthetic enzymes.

Keywords: Enzyme evolution, protein engineering, *Streptomyces*, polyketide, natural product

Significance Statement

Soil bacteria produce tens of thousands of complex bioactive compounds that are used as antibiotics and anticancer agents. Key to this rich chemodiversity lies in the ability of the bacteria to rapidly evolve metabolic pathways to generate unique molecules. Here we show how small changes in enzymes responsible for the formation of natural products are sufficient to alter their functions and even generate unprecedented activities. We propose that this is a fundamental ability for enzymes involved in bacterial secondary metabolism that has allowed the formation of the tremendous chemical diversity observed in nature.

Introduction

Microbial natural products are found ubiquitously from the biosphere, where they are present in enormous numbers and display tremendous diversity in their chemical structures. In particular, *Streptomyces* soil bacteria have tens of thousands of secondary metabolic pathways for production of natural products (1). The potent bioactivities have made microbial natural products a cornerstone of modern medicine ever since the discoveries of penicillin and streptomycin (2). Approximately two-thirds of antibiotics and one-third of anticancer agents are natural products or their semi-synthetic derivatives (3). The full scope of microbial chemodiversity is still unknown, as the producing organisms have been noted to harbor an excess of biosynthetic gene clusters in relation to the number of metabolites observed under monocultures (4). Many secondary metabolic pathways have been thought to be silent and require specific environmental cues for activation (5). Microbial natural products typically function as chemical warfare agents against competing organisms (6), but a fraction are involved in more complex chemical communication interactions (7).

Both intra- and inter-species chemical communication, which may be either agonistic or antagonistic in nature, has been reported (6, 8). Even trans-kingdom interactions have been demonstrated, such as production of the terpenoid geosmin by *Streptomyces* to attract arthropods to assist in spore dispersal (9).

From the perspective of evolutionary biochemistry, the most remarkable aspect of microbial natural products is the divergent enzyme evolution responsible for the emergence of chemodiversity. Several models for the molecular evolution of secondary metabolism have been presented (10–12), which highlight important differences to the evolution of the conserved central or primary metabolic pathways. A key concept dictates that the appearance of novel bioactive molecules provides advantages for the producing organisms under shifting environmental conditions and selective pressures (8, 13). Proteins involved in secondary metabolism may be considered generalists (14, 15), which manifests as promiscuity (16) and slow reaction rates (17). The situation is in contradiction to central metabolism, where proteins catalyze their reactions both efficiently and specifically to meet the

Competing Interest: The authors declare no competing interest.

Received: August 24, 2022. **Revised:** October 18, 2022. **Accepted:** November 28, 2022

© The Author(s) 2023. Published by Oxford University Press on behalf of National Academy of Sciences. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

demands for high metabolic fluxes (10, 18). Furthermore, purifying selection has shaped central metabolism by improving the catalytic properties of ancestral generalist enzymes at the cost of reducing genetic variation (11, 19). However, resilience towards purifying selection has allowed enzymes involved in secondary metabolism to remain in a generalist state and acquire atypical catalytic abilities above and beyond what has been classically thought possible.

An example of microbial chemical diversity are the anthracyclines, which are glycosylated tetracyclic aromatic polyketides (20) (Fig. 1A). Anthracyclines are mainly produced by Actinobacteria and to date 408 metabolites have been described (21). The complexity of anthracyclines is generated in the so-called tailoring steps of the biosynthesis, where the carbon framework is modified in many ways, and in glycosylation steps that differ in the type and number of carbohydrate units attached to the aglycones (22). The importance of anthracyclines is in their potent cytotoxic activity against tumor cells, and metabolites such as aclacinomycin and doxorubicin are widely used in cancer chemotherapy (23). Although the biological activity is complex, poisoning of topoisomerases and intercalation to DNA are important contributors (24, 25). However, other factors such as the formation of reactive oxygen species (26), the ability to evict histones from chromatin (27), and proteolytic activation of transcription factors (28) have been noted.

Tailoring steps catalyzed by SAM methyltransferase-like proteins (29) have yielded particular insight into evolution of anthracyclines and generation of structural diversity. DnrK is a canonical S-adenosyl-L-methionine (SAM)-dependent methyltransferase that catalyzes 4-O-methylation (Fig. 1B) in daunorubicin biosynthesis (30). Recent studies have uncovered that DnrK is bifunctional and catalyzes additional 10-decarboxylation as a moonlighting activity (Fig. 1B) (31). However, the activity depends on the presence of a free 10-carboxyl group in the substrate generated by aclacinomycin 15-methylsterases such as EamC and DnrP (Fig. 1C) (29). DnrK is quite specific with respect to the length of the carbohydrate chain at C-7, accepting only monoglycosides (Fig. 1A).

Conversely, RdmB (53.4% sequence identity to DnrK) from the ϵ -rhomomycin pathway is an anthracycline 10-hydroxylase requiring SAM, molecular oxygen and a thiol reducing agent for activity (Fig. 1B) (32). RdmB harbors broad substrate specificity and is able to turn over aglycones, mono- and triglycosylated anthracyclines (Fig. 1A) (29). The 10-hydroxylation activity has evolved twice independently, as indicated by the discovery of a second clade of 10-hydroxylating enzymes, such as CalMB (50.8% sequence identity to DnrK), that have a similar tolerance towards glycosylated substrates as RdmB (29). Finally, the enzyme family was expanded by identification of EamK (52.4% sequence identity to DnrK) from the komodoquinone B pathway, which catalyzes solely 10-decarboxylation (Fig. 1B) (29). The specificity of EamK and related 10-decarboxylase TamK lies between DnrK and RdmB, since anthracycline aglycones and monoglycosylated compounds are accepted as substrates (Fig. 1A).

Here we have taken advantage of the natural diversity of anthracycline methyltransferases and employed extensive chimeragenesis (33) to probe factors that govern catalysis and substrate recognition. We have mimicked evolutionary processes and generated 24 chimeric proteins using two proteins scaffolds and sequences from four proteins. The chimeras expanded the catalytic repertoire to include 9,10-elimination activity, and broadened the substrate tolerance for 4-O-methylation and 10-decarboxylation (Fig. 1D). Functional and structural

characterization revealed a complex network of interactions that determine the molecular basis for the altered enzymatic activities. The engineering efforts demonstrated how minor changes in generalist proteins may induce drastic effects in their functions as biocatalysts.

Results

Selection of regions for chimeragenesis

Three regions have previously been noted as critical for the functional diversification of the 10-hydroxylase RdmB from the canonical 4-O-methyltransferase DnrK (31). The first region R1 (residues 286–299 in DnrK) consists of the N-terminal half of α 16 and the preceding loop region. Helix α 16 is situated parallel to the tetracyclic anthracycline substrate with R1 residing near the 10-carboxyl group (Fig. 2A, S1). The switch in activity from 4-O-methylation to 10-hydroxylation could be attributed to insertion of a single S298 residue in DnrK into the α 16 helix, which triggered rotation of the neighboring F297 towards the ligand (31). This shift resulted in a solvent-free active site architecture reminiscent of RdmB, where a channel leading to the surface of the protein is closed by F300. A mechanistic model for 10-hydroxylation has proposed that exclusion of water molecules from the active site allows an anthracycline 10-carbanion intermediate to react with molecular oxygen for 10-hydroxylation to occur (31). In turn, region R2 (residues 334–345) comprises the loop region between β 8 and β 9 that folds over the active site and is in contact with the ligand (Fig. 2A, S1). Finally, the loop region between α 11 and α 12 (residues 160–169), which interacts with the carbohydrate unit of the substrate, was defined as region R3 (Fig. 2A, S1). Regions R2 and R3 from RdmB were found to expand the substrate tolerance of DnrK towards triglycosylated anthracyclines (31). The regions R1–R3 described here are highly variable in the four clades of anthracycline methyltransferases-like proteins (Fig. 2A, S1) and hence remained unchanged in this study.

The expansion of the family by the 10-decarboxylases directed us to differences in the loop region between β 3 and α 13, which is located close to the binding site of the co-substrate SAM. The 4-O-methylation of DnrK has strict requirements for correct geometrical positioning of the substrate and SAM for an S_N2 reaction to occur (34) and misalignment of the ligands has been proposed to be the reason why RdmB does not have methylation activity (32). Since this may be triggered either by shifting the position of the substrate or the co-substrate, we selected the SAM binding site as region R4 (Fig. 2A, S1). Noteworthy, the 10-decarboxylases EamK and TamK harbor two point mutations G189A/K190P in R4 in comparison to DnrK, which is interesting as the presence of a proline restricts backbone conformations (35).

Due to the complex nature of the constructs, we devised a simple nomenclature to describe the chimeras. In this system, the scaffold enzyme is firstly mentioned, followed by a sequence of four letters, corresponding to the origin of the regions R1–R4 (D for DnrK, R for RdmB, C for CalMB, and T for TamK). To exemplify, TamK DRRD is a chimera with a TamK scaffold, regions R1 and R4 from DnrK and regions R2 and R3 from RdmB.

Enzymatic activities of the chimeric proteins

The great number of combinatorial possibilities with four regions (R1–R4) and four possible donor sequences (DnrK, RdmB, CalMB, and TamK) precluded a systematic survey. The chimeras selected for the study were designed based on structural information to probe research questions with 12 chimeras created using DnrK

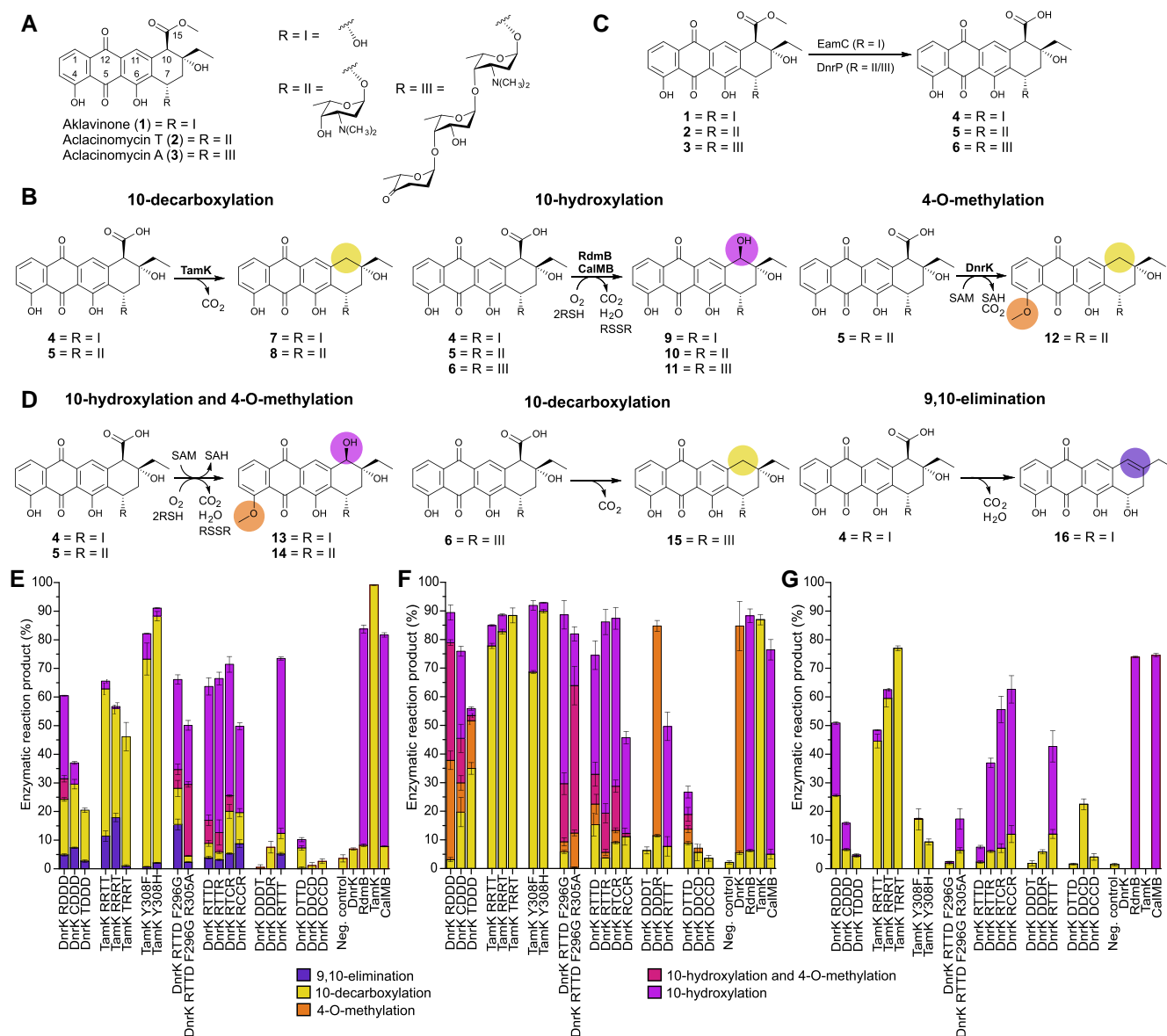


Fig. 1. Enzymatic activities of SAM-dependent methyltransferase-like proteins with anthracycline substrates. (A) Structures of aklavinone (**1**), aclacinomycin T (**2**) and aclacinomycin A (**3**). (B) Native reactions performed by native enzymes. (C) An initial 15-methylesterase activity is required for the modification reactions at C-10 position under study. EamC was used with **1**, while DnrP was used with both **2** and **3**. (D) Nonnatural reactions performed by the chimeric enzymes. Enzymatic reaction products with (E) **4**, (F) **5**, and (G) **6** as substrates. The reaction product yields are shown as mean value \pm standard deviation and were calculated from HPLC chromatogram traces by normalized peaks areas. The overall percentages may not add up to 100% in all samples meaning that some unreacted substrate remained in the reaction mixture.

as the main scaffold (Table S1). In addition, the viability of 10-decarboxylases to serve as scaffolds was tested with eight unique chimeras using TamK (Table S1). No hydroxylases were selected due to previous challenges in engineering the enzymatic activity of RdmB (31).

The relative enzymatic activities were tested with three compounds: the aglycone aklavinone (**1**); the monoglycosidic aclacinomycin T (**2**); and the triglycosidic aclacinomycin A (**3**) (Fig. 1A). In order to allow 10-decarboxylation and 10-hydroxylation to occur, all reactions were preceded by a 15-methylesterase reaction either by EamC (29) or DnrP (36). Subsequently intermediates **4**, **5**, and **6** were used as substrates for the chimeras together with the reducing agent dithiothreitol (DTT) and SAM.

Analysis of reaction products by HPLC revealed that 16 chimeras harbored diverse enzymatic activities (Fig. 1E–G, S2). The

majority of reaction products could be identified by comparison to standards obtained from previous studies, which included 10-decarboxylation (**7**, **8**, **15**), 10-hydroxylation (**9**, **10**, **11**), and 4-O-methylation (**12**) products, and compounds with both 4-O-methyl and 10-hydroxyl groups (**13**, **14**) (Figs. 1D, S3–S9). To confirm the identity of **15**, which was produced by several TamK-based chimeras when **6** was used as a substrate, the glycosidic units were acid hydrolyzed and the resulting aglycone **7** was compared to an authentic standard **7** (Fig. S10). MS measurements supported that **15** was a triglycosylated 10-decarboxylation reaction product ($[M+H]^+$, ESI+ obs. 754.3481, calc. 754.3433, Fig. S11). Similarly, four chimeras appeared to expand the substrate specificity for 4-O-methylation and 10-hydroxylation towards **4** based on MS analysis ($[2M+Na]^+$, ESI+ obs. 791.2254, calc. 791.2310, Fig. S12). In order to identify the novel product **13**, we removed

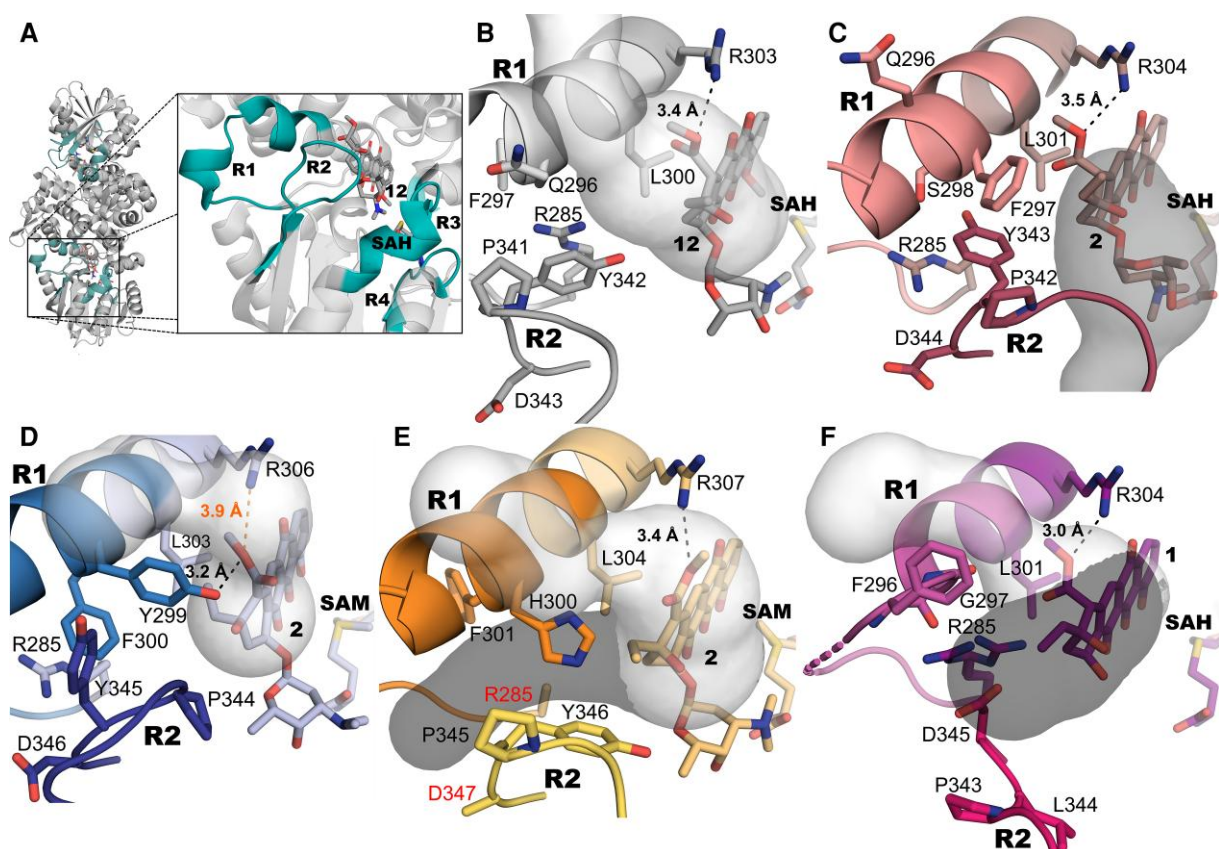


Fig. 2. Structural analysis of regions R1 and R2 from different clades. (A) Structure of DnrK, with exchanged regions highlighted (teal). The conformation of R1 and R2 changes the openness of the active site, with implications towards activity. The presence of a channel arising from R1 (light grey) is implicated in the absence of hydroxylation reaction, while the interface between the regions reveals an open/closed conformation of the active site (dark grey). (B) In DnrK (PDBID: 1TW2), the absence of an aromatic residue in the position of Q296 opens the channel (light grey) preventing 10-hydroxylation. The presence of the aromatic Y342 in R2 restricts the positioning of the conserved R285 away from the active site. (C) In DnrK DDDD S298 (PDBID: 4WXH), the addition of the serine leads to a repositioning of F297, with closure of the channel preventing access of bulk solvent to the active site. Despite a rearrangement of R2, the position of Y343 is maintained. (D) DnrK TDDD shows an aromatic residue (Y299) in a similar position to that of F297 in DnrK S98. However, the interaction of Y299 with the substrate **2** allows the solvent channel to remain open (light grey). (E) DnrK CDDD crystallized in an open conformation, with a substrate access channel (dark grey) visible between regions R1 and R2. Subtle movement of H300 and Y346 is required to close the solvent channel (light grey) in the closed conformation. Y346 in R2 reveals a different conformation, which likely allowed a higher degree of freedom to R285, not able to be built in this structure. (F) The dramatic rearrangements in DnrK RTTD F297G lead to repositioning of a neighboring residue F296 towards the solvent channel (light grey). This structure is in an open conformation with a visible substrate access channel (dark grey). The absence of an aromatic residue in R2 (Y342L), leads to a rearrangement of this region, with D345 being placed in a structurally equivalent position to that of the aromatic residues. This modification provides more space and a higher degree of freedom for R285, seen in two distinct conformations, allowing an interaction with the substrate.

the ϵ -rhodosamine aminosugar from standard **14** by acid hydrolysis and compared the products by LC-MS-EICS (Fig. S13).

We also noted that 10 chimeras produced a novel product **16**, which appeared to be an elimination product based on HR-MS measurement ($[M-H]^-$, ESI- obs. 335.0923, calc. 335.0925, Fig. S14), when **4** was used as a substrate. The highest yields were detected with TamK RRRT, which was used in large-scale reactions to obtain sufficient material for structure elucidation. NMR experiments (1H , ^{13}C , HSQCDE, COSY, and HMBC) revealed that an additional 9,10-elimination reaction had occurred in **16** (Table S2 and Figs. S15–S20).

The critical role of region R1 in reaction-type determination

Previous studies have demonstrated the importance of R1 from RdmB in gain-of-hydroxylation activity of DnrK RDDD with **5** and **6** (31). At the time, the activity could not be assayed for **4**, since the 15-methylesterase DnrP does not accept **1** as a substrate. However, the recent discovery of the 15-methylesterase EamC (29)

allowed us to show that DnrK RDDD is able to accept **4** with 29% hydroxylation reactivity (Fig. 1E).

Here we wished to focus on the second clade of 10-hydroxylating enzymes, since R1 of CalMB differs greatly from both DnrK and RdmB. Region R1 in CalMB is extended by an insertion of three residues and has a Phe to His substitution at the critical site in comparison to RdmB (Fig. S1). Despite these differences, DnrK CDDD behaved similar to its R1 provider with the highest yield of hydroxylation product (46%) shown with **5** (Fig. 1F).

Next, we introduced R1 from the 10-decarboxylase TamK, creating DnrK TDDD. The region is reminiscent to CalMB in terms of length, but contains a Tyr residue instead of His at the key junction. Activity assays confirmed the importance of R1 in determination of reaction chemistry, since the 10-decarboxylation activity of DnrK TDDD increased with all substrates, particularly with **5** (35%), in comparison to DnrK RDDD and DnrK CDDD (Fig. 1F).

The use of TamK as a protein scaffold proved to be moderately successful in highlighting the role of R1 in catalysis. TamK RRRT

and TamK RRRT acquired minor 10-hydroxylation activity as expected based on the R1 donor sequence, while TamK TRRT showed no 10-hydroxylation profile (Fig. 1E–1G). Similarly to experiments with the RdmB scaffold (31), engineering gain-of-methylation functionality was challenging, as chimeras TamK DRRD, TamK DRTD, TamK DDRD, TamK DDTD, and TamK DDDD did not display any measurable enzymatic activity.

Structural basis for 10-decarboxylation activity

All proteins from the four phylogenetic clades share 10-decarboxylation as a common feature. In addition to the true 10-decarboxylase TamK (29) and the moonlighting 10-decarboxylase activity of DnrK (31), formation of a 10-carbanion intermediate has been proposed to be a key step in the 10-hydroxylation activity of RdmB (32). A fully conserved arginine R303 (DnrK numbering) residue (Fig. 2B–2F, S1), which is within hydrogen bonding distance from the 10-carboxyl group, is considered to initiate the reaction (31, 32).

Structural determination of DnrK TDDD was carried out to explore the significant gain of decarboxylation observed by this chimera (Fig. 1F). The structure revealed that insertion of three additional residues in R1 extends the α 16 helix by one additional half-turn in comparison to DnrK and DnrK DDDD S298 (31) (Fig. 2B–2D). However, the longer helix protrudes away from the active site towards bulk solvent and is unlikely to have a significant influence in catalysis. The positioning of Q296 in DnrK allows access for solvent ions to reach the active site through an open channel (Fig. 2B), which is blocked by F297 in DnrK DDDD S298 (Fig. 2C). DnrK TDDD harbors Y299 in the structurally equivalent position (Fig. 2D), but the solvent channel remains open due to an additional interaction of Y299 with the ligand; a hydrogen bond between the tyrosine and the carbonyl oxygen of the substrate rotates the side chain of Y299 in a manner that prevents channel closure (Fig. 2D).

The residue Y299 is fully conserved in 10-decarboxylating enzymes and mutation of that residue, in TamK, into aromatic phenylalanine or histidine revealed the formation of hydroxylated products **9** and **10** (Fig. 1E–1F). It seems likely that the aforementioned interaction of Y299 with the substrate is required to prevent closing of the channel, and the disruption of the hydrogen bond may lead to a rotation of the functional group into a position that favors hydroxylation. The gain of hydroxylation by the single Y308H mutation (TamK numbering) seems particularly relevant, with CalMB showing a natural His in that position (Fig. S1), hinting at the evolutionary path taken by the TamK/CalMB clade.

Structural basis for 10-hydroxylation activity

To investigate if the previously proposed mechanism for 10-hydroxylation (31) holds true for the second clade of 10-hydroxylating enzymes, we solved the structure of DnrK CDDD. The structure demonstrates that two aromatic residues, H300 and Y346, are positioned adjacent to the channel. The analysis is complicated by the fact that DnrK CDDD crystallized in an open conformation where R2 hosting Y346 has moved away from R1 and, as a consequence, H300 does not fully close the solvent channel (Fig. 2E). However, it seems likely that movement of R2 to close the active site will lead to rearrangement of the H300:Y346 interaction. A similar effect can be observed in the reported structures of DnrK, where significant differences in R2 emerged between open and closed conformations (37).

To overcome the experimental uncertainty, we replaced the critical phenylalanine residue with glycine in DnrK RTTD F297G.

According to our mechanistic model, removal of the bulky hydrophobic side chain would allow bulk solvent to access the active site and prevent 10-hydroxylation (31). Unexpectedly the mutant displayed moderate 10-hydroxylation activity towards **4** and **5** with 38 and 79% conversion, respectively, with limited reactivity towards **6** (Fig. 1E–1G). The contradiction was cleared by structural data from DnrK RTTD F297G, which revealed significant rearrangements in the active site architecture. The introduction of G297 had induced a disruption to the α 16 helix secondary structure and, consequently, lead to rotation of a different phenylalanine residue (the preceding F296) towards the active site. While cavity analysis shows an open conformation also for this structure, F296 is ideally positioned to block the solvent channel in the closed conformation (Fig. 2F). The data therefore provide independent support for the importance of solvent exclusion in the 10-hydroxylation reaction. Interestingly, mutation of the phenylalanine to alanine in DnrK RTTD F297A led to comparable enzymatic activity to DnrK RTTD F297G (Fig. S21), but the active site architecture of this chimera could not be probed due to challenges in crystallization.

Structural basis for 9,10-elimination activity

The experiments with DnrK RTTD F297G fortuitously led to gain-of-elimination activity (Fig. 1D–1E) with production of **16** (15% conversion). The active site is significantly modified by the single mutation and an explanation for 9,10-elimination may be found within the interface of R1 and R2 (Fig. 2B–2F). Key to the switch in activity resides in the fully conserved residue R285 located directly adjacent to R1 (Fig. 2F), which has gained freedom to move towards the active site in DnrK RTTD F297G. The mobility is possible due to the Y342L mutation in R2, repositioning the neighboring D345, which removes bulky aromatic amino acids present in DnrK (Y342) and RdmB (F346) (Fig. S1) that prevent arginine–ligand interactions in the wild-type enzymes.

The structural data suggest that the initial step in the 9,10-elimination reaction is 10-decarboxylation by R305 (R303 in DnrK WT). The formation of a carbanion intermediate is likely to be essential for the activity (31), since DnrK RTTD F297G could not catalyze the reaction when the 10-decarboxylation product **7** was used as a substrate (Fig. S22). Interaction of R285 with the C9 hydroxyl group of the carbanion intermediate may lead to protonation and generation of a better leaving group resulting in the elimination product.

This proposed mechanism was tested by the mutation of the pivotal arginine into alanine, to create DnrK RTTD F297G R303A. The resulting enzyme suffered a seven times decrease in elimination activity (Fig. 1E), without impairing its hydroxylation ability with any of the substrates (Fig. 1E–1G), and in effect dramatically improving the hydroxylation of **6**, which shows a 14 times increase (Fig. 1G). The removal of the long chain amino acid had an additional effect of improving the methylation reaction rate with **4** and **5** (Fig. 1E–1F), likely due to a more accommodating active site.

Factors affecting substrate specificity

TamK has been shown to catalyze 10-decarboxylation of **4** and **5**, but not **6** (29). In order to influence reactivity towards triglycosidic substrates, we engineered R2 and R3 in the TamK scaffold. The use of amino acid sequences from RdmB had a notable synergistic effect in accepting **6**, with TamK TRRT showing the highest rate in the formation of the novel 10-decarboxylation product **14** (77% conversion, Fig. 1G). Minor gain-of-hydroxylation activity (3%

conversion, Fig. 1G) could be seen in the construct TamK RRRT, with slightly reduced 10-decarboxylation activity (60% conversion, Fig. 1G).

Next the chimeric pair DnrK RTTD and DnrK RTTR was constructed to examine the influence of the sugar-binding regions R2 and R3 from TamK, an enzyme with higher affinity towards **4** than RdmB (29). Both DnrK RTTD and DnrK RTTR became exceptional generalist enzymes, with the ability to use any of the substrates to catalyze reactions, yielding all potential products albeit at different yields (Fig. 1E–1G). Of particular interest is the significant gain-of-activity with the aglycone **4** in comparison to the DnrK (Fig. 1E). The presence of R1 from RdmB in both chimeras shifted the reactivity from 4-O-methylation towards 10-hydroxylation, with less than 15% of the native DnrK activity measured, compared to almost 60% 10-hydroxylation conversion.

Structures of DnrK RTTD and DnrK RTTR in complex with **1** revealed the molecular basis for the expanded substrate specificity. Regions R2 and R3 are observed in a closer interaction with each other when compared with all structures in complex with **2**, closing the active site to stabilize the smaller aglycone substrate. The mutations introduced by chimeragenesis have altered the hydrogen bonding network between the R2 and R3 interface (Fig. 3). The α 11 helix hosting R3 is extended by a single turn, which positions the helix in the space vacated by the absent carbohydrate unit. This structural change allows stabilization of the substrate via hydrogen bonding interactions between R163 and the 6-hydroxyl group of **1** (Figs. 3B and C).

In addition to R2 and R3, long-range secondary-shell effects via R4 were found to influence substrate specificity. As an example, DnrK RTTD and RTTR differ only in one amino acid, K190 and N190, respectively, that is located in R4 (Fig. S1). The activity profiles were generally similar for substrates **4** and **5**, but significant differences were observed with the triglycosylated **6** (Fig. 1E–1G). DnrK RTTD has minor total activity (8%) with **6**, whereas DnrK RTTR is able to turn over 37% of the substrate.

The interface between R3 and R4 in DnrK in complex with **2** demonstrates that K190 from R4 is within hydrogen bonding distance (3.1 Å) to S158 located in R3 (Fig. 3A). The same inter-region interaction (3.5 Å) may be observed in DnrK RTTD in complex with **1** (Fig. 3B), but the hydrogen bond is not present in DnrK RTTR in complex with **1** (Fig. 3C) due to the shorter side chain of N190 (5.6 Å). Unfolding of R3 has been observed in the complex structure of RdmB (32, 38) and appears to be essential for utilization of **6** as a substrate. The additional stabilization of helix α 11 provided by the interaction of K190 and S158 in DnrK and DnrK RTTD may prevent the use of triglycosylated substrates. The proposal is further supported by DnrK RTTT, which is an efficient 10-hydroxylase for **6** (31%) (Fig. 1G). These chimera harbors P190 in the equivalent position to K190 in DnrK RTTD (Table S1) and would not exhibit strong interactions between R3 and R4.

The influence of region R4 and loss-of-methylation activity

The SAM binding site is highly conserved in the enzyme family with the exception of the 10-decarboxylases that contain two mutations G189A/K190P (Fig. S1). In order to investigate what effects the differences in R4 may have, we constructed chimera DnrK DDDT, which displayed close to no reactivity towards any substrate and led to complete loss-of-methylation activity (Fig. 1E–1G). As a control, we generated DnrK DDDR, which contains a single conservative substitution K190N and the chimera behaved similarly to the wild-type enzyme. DnrK DDDR harbored

no reactivity towards **4** and **6**, while retaining close to 73% of 4-O-methylation and 11% 10-decarboxylation activity towards **5** (Fig. 1E–1G).

These experiments lent support to the idea that the introduction of the proline may compromise correct positioning of the co-factor and quench 4-O-methylation activity. The structure of the binary complex of wild-type TamK and SAM (Fig. S23) confirmed the interactions between the co-substrate and R4 in 10-decarboxylases. The overall fold of TamK remains very similar to that of the other members of the family with modest root-mean-square deviations to DnrK (3.7 Å) and to RdmB (3.8 Å). However, the distinct proline-alanine pair in R4 alters the conformation of the α main chain in the region and leads to a movement of 2.1 Å in comparison to DnrK. This change coincides with a 1.8 Å change in SAM positioning (Fig. S23), which is the most likely reason behind the lack of methylation activity in TamK. Similar, but less drastic, movement can be observed also in RdmB (32) (Fig. S23). The strict geometrical requirements for 4-O-methylation were also apparent in chimeras DnrK DTTD, DnrK DDCD, and DnrK DCCD, where the overall activity with **5** was severely impaired, without significant gains in 4-O-methylation of **4** or **6** (Fig. 1E–1G).

Generalist nature of the chimeric proteins

Several of the multi-chimeric proteins became truly generalist enzymes that may resemble their ancestral counterparts. DnrK RTTR can accept all substrates to carry out 10-decarboxylation, 4-O-methylation, 10-hydroxylation and 9,10-elimination (Fig. 1E–1G). The generalist nature is promoted by the mobility of **1**, which was observed in two different positions in the active site of the dimeric enzyme (Fig. 3C). In one of the monomers, the substrate is located close to the conserved R303 (DnrK numbering) critical for the 10-hydroxylation activity (32), whereas in the other monomer the substrate has moved 1.6 Å towards R163 located in the α 11 helix of R3.

Similarly, DnrK RTCR is adept at carrying out diverse reactions with improved acceptance of **6** (Fig. 1E–1G). The structure of DnrK RTCR with **2** shows that the sugar-binding regions R2 and R3 no longer interact with one another (Fig. S24), since the presence of the L-rhodamine unit of **2** has driven R3 away from the aglycone core of the substrate. The displaced amino acids in R3 remain without a defined secondary structure, which will likely provide a higher degree of freedom for the substrate to re-arrange its position.

Discussion

Evolutionary implications for the emergence of new-to-nature compounds

Diversification of enzyme function is fundamental to the evolution of secondary metabolism (12). The work presented here depicts the evolutionary events that have allowed SAM-dependent methyltransferases to diversify to catalyze atypical chemical transformations, and how similar events can be recreated in laboratory conditions to yield novel products. The experiments demonstrate how few point mutations in four hypervariable regions are sufficient to change protein functions (Fig. 4) and directly alter products of entire metabolic pathways.

We propose that the initial moonlighting 10-decarboxylation activity has been the seed for functional differentiation (Fig. 1B, S2), which has allowed the development of secondary catalytic activities (39, 40) without disturbances to 4-O-methylation. The

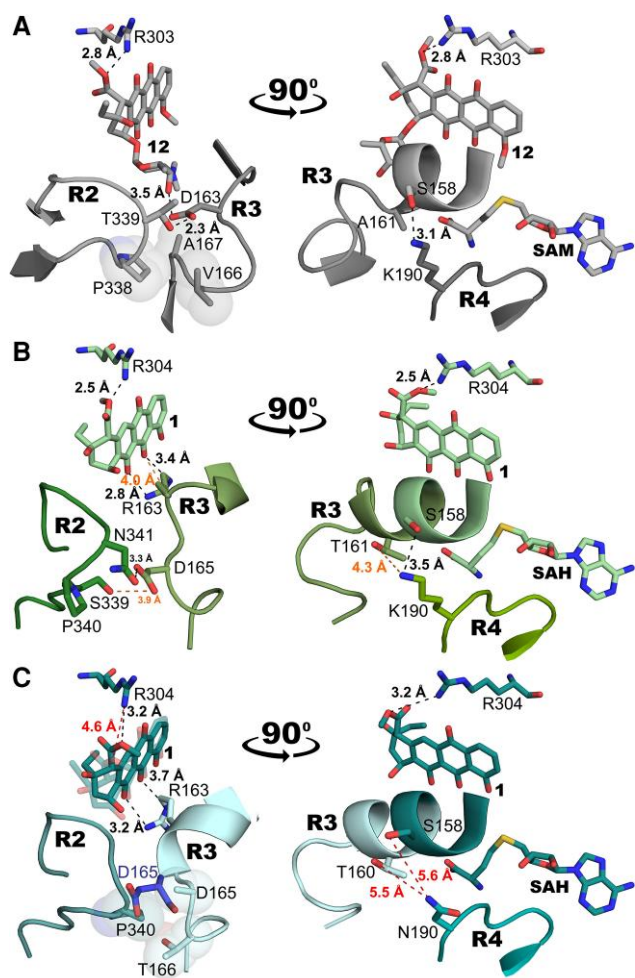


Fig. 3. Structural basis for substrate specificity and inter-region interactions between regions R2, R3 and R4. Binding of **1** leads to rearrangements in regions R2 and R3, while the interface between regions R3 and R4 may influence substrate specificity with **6**. (A) In DnrK (PDBID: 1TW2), D163 (R3) forms a hydrogen bond to the C3' hydroxyl group of **12**. The interface between regions R2 and R3 shows hydrogen bonding between T339 and D163. In addition, hydrophobic interactions are present between P338 (R2) and V166 and A167 (R3). The interaction between regions R3 and R4 is due to a hydrogen bond between K190 and S158. (B) In the absence of the sugar subunit, the extended α 11 helix in DnrK RTTD positions R163 (R3) to interact with **1**. Furthermore, a more complex hydrogen bonding network with interactions between S339 and N341 (R2) and D165 (R3) are observed. Similarly to DnrK, K190 (R4) is observed to interact weakly with residues from R3 in DnrK RTTD. (C) The structure of DnrK RTTR shows how N190 (R4) does not interact with residues from R3, which has likely implications for the reactivity with **6**. The structure of DnrK RTTR also demonstrates mobility of nonglycosylated **1** in the active site, where the ligand is found interacting either with the conserved R304 from helix α 16 (dark teal) or R163 (R3) (light teal).

moonlighting activity has emerged through evolution of anthracycline substrates with free 10-carboxyl groups (29, 41) in combination with conservation of R303 important for substrate binding (32, 37).

RdmB-type 10-hydroxylases (Fig. 1B, 4) have appeared in a single evolutionary step via insertion of S298 to DnrK-type ancestral proteins (31) (Fig. S1), leading to the repositioning of the aromatic F297 (DnrK numbering). A shared feature of DnrK and RdmB is the formation of an initial carbanion intermediate by R303. In DnrK, the intermediate reacts with H_3O^+ leading to 10-decarboxylation, whereas rotation of F300 and exclusion of solvent ions in RdmB

allows the carbanion to react with O_2 (Fig. S2). The hydroperoxy intermediate is stable enough to have been detected by mass spectrometry (32) and a reaction with cellular reductants such as glutathione yields the 10-hydroxylation product. The loosening of geometric requirements for the hydroxylation reactions, no longer constrained by the S_N2 -type mechanism of 4-O-methylation, allowed further diversification in the sugar-binding regions of the active site, expanding the role of the enzyme to accept triglycosidic substrates.

The changes that lead to acceptance of **6** are complex, but exploration of the structural networks in this project led to DnrK-based chimeras such as DnrK RCCR that demonstrate highly efficient 10-hydroxylation activity. The exploration of inter-regional effects led to the discovery of novel 10-decarboxylases able to use **6** as a substrate, DnrK DCCD and TamK TRRT (Fig. 1D and G).

The evolutionary events have taken a different route on the second clade of proteins (Fig. 4). The initial step has been subfunctionalization of TamK-type proteins to catalyze solely 10-decarboxylation (Fig. 1B). Two-point mutations G189A/K190P in R4 (Fig. 4, S1) are sufficient for misalignment of SAM and loss-of-methylation activity. In addition, the 10-decarboxylation activity has been improved via changes in R1 and positioning of Y308 in a manner that enables hydrogen bonding interactions with the 10-carboxyl group (Fig. 4, S2). Changes of Y308 onto other aromatic residues lead the protein onto a convergent evolution path of 10-hydroxylation (Fig. S1), as seen in the activity of TamK Y308F and Y308H (Fig. 1E–1G). Despite the different evolutionary paths of CalMB and RdmB, the chemistry appears to be mechanistically similar with exclusion of bulk solvent being a critical feature (Fig. 4). The chemistry is likely to hold true also for DnrK RTTD F297G (Fig. S2), which could be considered as a third alternative for gain-of-hydroxylation activity that Nature has not yet discovered.

Critical for the laboratory evolution of the novel 9,10-elimination activity (Fig. 1D) appears to be the removal of bulky aromatic amino acids in R2 (Fig. 4). Mutations such as Y342L in DnrK RTTD F297G (Fig. S1) allow repositioning of a fully conserved R285 (DnrK numbering) towards the active site and interaction with the substrate (Fig. 4). The 9,10-elimination is likely to proceed via the same carbanion intermediate that is also important for 10-decarboxylation and 10-hydroxylation activities (Fig. S2). The artificial evolution of 9,10-elimination functionality is encouraging as it demonstrates that the catalytic repertoire of natural product biosynthesis enzymes can be further engineered in the laboratory.

Materials and methods

Reagents

All reagents were bought from SigmaAldrich unless stated otherwise. TALON SuperFlow resin and PD-10 desalting columns were bought from GE Healthcare. SDS-PAGE Bolt 4–12% Bis-Tris Plus Gel and PageRuler Unstained Protein Ladder molecular weight standard were bought from Thermo Fisher Scientific. Phusion polymerase, restriction enzymes, T4 DNA ligases, and GeneJET Plasmid Miniprep Kit were bought from Thermo Fisher Scientific.

Production and purification of compounds

Compounds **1**, **2**, and **3** were produced and purified as described previously (29, 31). Briefly, aclacinomycin T (**2**) was obtained from a fermentation of strain *Streptomyces galilaeus* ATCC 31615

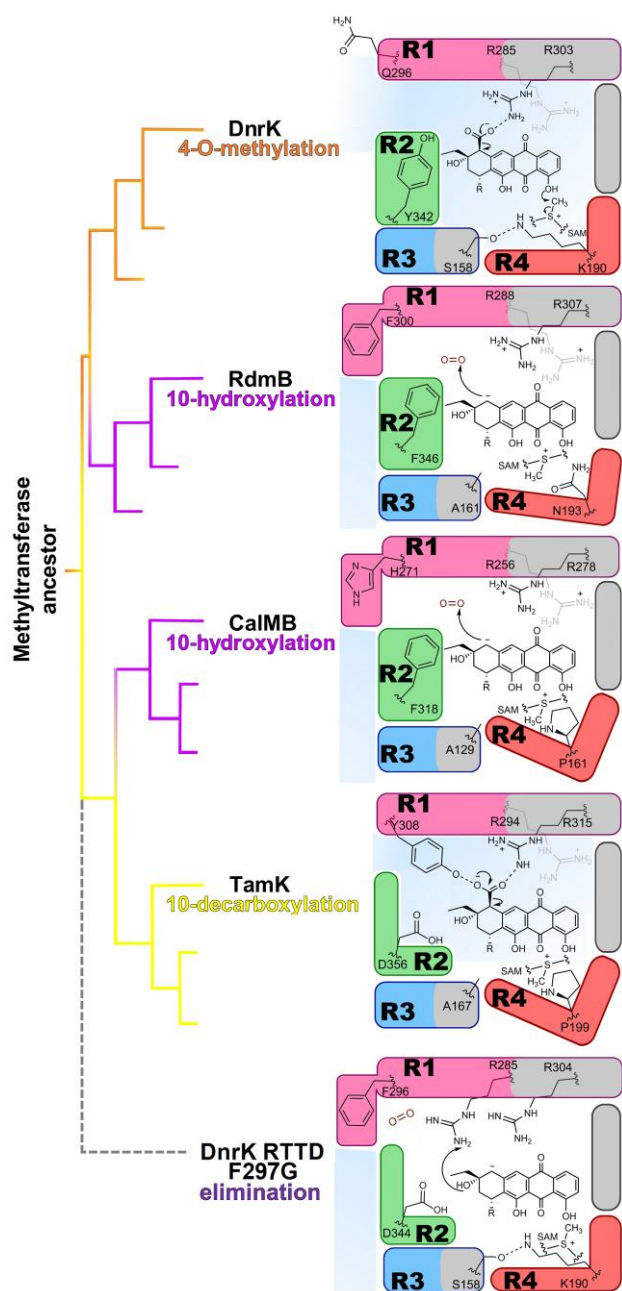


Fig. 4. Schematic representation of different activities. Active site evolution and mechanistic consequences. The active site of 4-O-methylase (orange) DnrK shows the shared decarboxylation mechanism. In addition, its open R1 cannot prevent bulk solvent access, SAM is positioned to catalyze methyl transfer and R3 and R4 interactions prevent active site flexibility for acceptance of **6**. Convergent evolution leads to two distinct 10-hydroxylases (pink). Both RdmB and CalMB share an aromatic residue in R1 that closes the active site to bulk solvent and noninteracting regions R3 and R4 allowing triglycosylated anthracyclines. In the case of CalMB, R4 resembles that of TamK. 10-decarboxylase (yellow) TamK has Y308 interacting with the leaving decarboxyl group. R4 leads to a different binding of SAM, with no methylation. The nonnatural (grey dots) “eliminase” possesses a decarboxylase-like R2, without an aromatic residue, allowing R285 to interact with the substrate.

mutant H038 (42). Aclacinomycin A (**3**) was obtained in a two-step fermentation process by first cultivating *Streptomyces galilaeus* ATCC 31615 mutant HO42 for production of aclacinomycin B, followed by biotransformation to **3** using *Streptomyces galilaeus* ATCC

31615 mutant HO26 (42). Aklavinone (**1**) was obtained through acid hydrolysis of **3** by using equal volumes of toluene and 2.5 M HCl at 60°C for 1 h 20 min. The compounds were purified by preparative HPLC using a LC-20AP/CBM-20A system with a diode array detector (Shimadzu) and EVO C18, 5 μ m, 100 Å, 250 \times 21.2 mm Kinetex column (Phenomenex). Purity of the compounds was estimated by HPLC and was at least 96% for **1**, 88% for **2** and 76% for **3**.

Protein expression and purification

Protein expression and purification were done as described previously (29, 31). Briefly, protein expression was carried out in *Escherichia coli* TOP10 cells. Protein purification was carried out through a polyHis-tag affinity chromatography with TALON SuperFlow resin and PD-10 desalting columns. Enzymes were concentrated using Amicon Ultra 0.5 mL centrifugation filters (10,000 nominal molecular weight limit). Purity of all the enzymes was estimated to be 90% or more by SDS-PAGE.

General DNA techniques

Modified pBAD/His B vector (Invitrogen) containing native *dnrK* and *tamK* (31, 37) were used as templates for chimeragenesis. The DnrK chimeras TDDD, CDDD, DTTD, RTTD, RTTR, and TamK chimeras DRRD, RRTT, DRTD, DDRD, TRRT, DDTD, RRRT, DDDR, and DDDD were ordered as synthetic DNA from GeneArt, Thermo Fisher or Sigma Aldrich (Strings DNA Fragments). The DnrK chimeras DDDT, DDCD, and DCCD were generated using the four-primer overhang extension method as described previously (31). In addition, both synthetic DNA and four-primer overhang extension method were utilized to construct the following DnrK chimeras: RTTT, RTCR, RCCR, and RTTD F297G. The obtained chimeras were cloned into pBAD/His B vectors. All DNA sequences were confirmed by sequencing before protein expression.

Enzyme activity measurements

Enzymatic activity measurements were carried out as previously described (29). Briefly, a two-step reaction was initiated by incubation of **1** (120 μ M) with EamC (9 μ M) or alternatively **2** (120 μ M) or **3** (120 μ M) with DnrP (130 μ M), to carry out an initial 15-methyl removal. The reaction products of this initial step were isolated as described previously (31). Activity measurements of the chimeric proteins were performed with the 15-demethylated compounds **4**, **5**, and **6** under the following conditions: 100 mM Tris-HCl (pH 7.5), 10 mM DTT, and 400 μ M SAM. The concentration of the native and chimeric enzymes was set to 6 μ M. The reactions were carried out at +37°C for 4 h.

All reactions were monitored by HPLC (SCL-10Avp/SpdM10Avp system with a diode array detector [Shimadzu]) using a C18 column (2.6 μ m, 100 Å, 4.6 \times 100 mm Kinetex column [Phenomenex]). The gradient elution used consisted of an initial aqueous buffer (5% (vol/vol) acetonitrile with 45% ammonium acetate—acetic acid [pH 3.6]) that was gradually changed to an elution buffer (100% acetonitrile). All compounds reported were confirmed by low-resolution MS (Agilent 6120 Quadrupole LCMS system; linked to an Agilent Technologies 1260 infinity HPLC system) with identical columns, gradient, and buffer systems as described previously (31).

HR-MS and NMR analysis of novel compounds

Analysis by HR-MS of **13**, **15**, and **16** was carried out by dissolving the compound in methanol, followed by direct injection to MicroTOF-Q high resolution MS (Bruker Daltonics). A sample of the same compound **16** was desiccated overnight, dissolved in deuterated chloroform and subjected to NMR analyses. The

spectra were recorded with a 500 MHz instrument (Bruker) with a liquid nitrogen cooled Prodigy BBO (CryoProbe), a 600 MHz instrument (Bruker) with a liquid nitrogen cooled Prodigy TCI (inverted CryoProbe) at 298 K. The signals were internally referenced to the solvent signals or tetramethylsilane. The experiments included 1D spectral analysis (^1H , ^{13}C) and 2D measurements (COSY, HMBC, and HSQCDE). Topspin (Bruker Biospin) was used for spectral analysis.

Crystallization, data collection and structure determination

Purified proteins were mixed with both the relevant anthracycline (100 mM stock in DMSO) and SAM (50 mM stock in buffer) to a 1:3:3 molar stoichiometry of protein:anthracycline:SAM, to a final protein concentration of 25 mg/mL. Crystals were set-up via the hanging drop method in a 1:1 ratio with the crystallization condition.

Crystals of DnrK RTTD, DnrK RTTR, and DnrK RTTR F297G were obtained in 0.1 M sodium acetate, pH 4.6; 8% PEG 8000, by co-crystallization with **1**. Attempts to co-crystallize with **2** yielded data with an aglycone present in the active site, likely caused by the acidic cleavage of the O-glycosidic bond. TamK crystals were grown in 0.04 M potassium phosphate monobasic (pH 7.5), 16% PEG 8000, 20% glycerol, in the presence of SAM and **1** as well as **2**, however no electron density for substrate was observed.

DnrK RTCR crystallized in 0.2 M magnesium chloride hexahydrate; 0.1 M HEPES 7.0; 20% w/v PEG 6000, in the presence of **2**. DnrK RTTT crystallized under the same condition, with varying concentrations of SAM, however no density for the cofactor was observed in the active site. The samples were incubated for 6 days at 21°C for crystal growth. DnrK TDDD crystallized in 0.2 M magnesium chloride hexahydrate; 0.1 M Tris-HCl (pH = 8.0); 20% w/v PEG 6000 in the presence of **2**. The samples were incubated for 14 days at 4°C for crystal growth. DnrK CDDD crystallized in 0.2 M Potassium thiocyanate, 0.1 M Bis-Tris propane (pH = 7.5), 20% w/v PEG 3350 in the presence of **2**. The samples were incubated for 21 days at 4°C for crystal growth. After crystal formation, the samples were collected and flash frozen in liquid nitrogen.

Diffraction data for DnrK RTTD, DnrK RTTR, DnrK RTTR F297G, DnrK RTCR, and DnrK RTTT were collected at the Max IV facilities (Lund, Sweden) at the BioMax beamline. Diffraction data for DnrK TDDD and DnrK CDDD were collected at the European Synchrotron Radiation Facilities (ESRF) (Grenoble, France), at the ID30A-1 beamline. Diffraction data for DnrK CDDD and DnrK TDDD were collected at the ESRF MASSIF-1 beamline. Diffraction data for TamK were collected at the ESRF (Grenoble, France), at the ID23-1 beamline. Details of the data collection and refinement statistics are listed in Table S3. Data were indexed, integrated, and scaled in XDS (43) and AIMLESS (44). Due to anisotropic diffraction in the DnrK RTTR dataset, ellipsoidal truncation followed by anisotropic scaling was carried out in the UCLA-DOE LAB Diffraction Anisotropy Server (45). Initial phases were obtained by molecular replacement with the DnrK WT structure (1tw2) as a starting model, using the Phenix Molecular Replacement program (Phaser) (46). Fo-Fc electron density maps demonstrated that the protein structures were ternary complexes (Fig. S25). Figures depicting protein structures were prepared using PyMOL (The PyMOL Molecular Graphics System, version 1.3; Schrödinger, LLC).

Acknowledgments

The European Synchrotron Radiation Facilities (Grenoble, France) and Max IV (Lund, Sweden) are acknowledged for access to

synchrotron radiation. This manuscript was posted on a preprint <https://doi.org/10.1101/2021.08.05.455194>.

Supplementary material

Supplementary material is available at PNAS Nexus online.

Funding

This study was supported by Academy of Finland Grant 340013 (to M.M.-K.), the Novo Nordisk Foundation Grant NNF19OC0057511 (to M.M.-K.), Turku University Foundation (to B.N.W.) and the Finnish Academy of Science and Letters (to B.N.W.).

Data availability

Crystallographic data is available from the Protein Data Bank (PDB) under the accession numbers: DnrK CDDD, 7OWB; DnrK TDDD, 7PHD; DnrK RTTD, 7PGA; DnrK RTTD F297G, 7PGJ; DnrK RTTR, 7PHF; DnrK RTCR, 7OY1; TamK, 7PG7. Enzyme activity assay raw data is available upon request from the authors.

References

- 1 Harir M, Bendif H, Bellahcene M, Fortas Z, Pogni R. 2018. Streptomycetes secondary metabolites. In: Enany S, editor. *Basic biology and applications of Actinobacteria*. London: IntechOpen. p. 99–122.
- 2 Newman DJ, Cragg GM, Snader KM. 2000. The influence of natural products upon drug discovery. *Nat Prod Rep*. 17:215–234.
- 3 Newman DJ, Cragg GM. 2016. Natural products as sources of new drugs from 1981 to 2014. *J Nat Prod*. 79:629–661.
- 4 Baral B, Akhgari A, Metsä-Ketelä M. 2018. Activation of microbial secondary metabolic pathways: avenues and challenges. *Synth Syst Biotechnol*. 3:163–178.
- 5 van der Meij A, Worsley SF, Hutchings MI, van Wezel GP. 2017. Chemical ecology of antibiotic production by actinomycetes. *FEMS Microbiol Rev*. 41:392–416.
- 6 Khalil ZG, Cruz-Morales P, Licona-Cassani C, Marcellin E, Capon RJ. 2019. Inter-Kingdom beach warfare: microbial chemical communication activates natural chemical defences. *ISME J*. 13:147–158.
- 7 Schmidt R, Ulanova D, Wick LY, Bode HB, Garbeva P. 2019. Microbe-driven chemical ecology: past, present and future. *ISME J*. 13:2656–2663.
- 8 Challis GL, Hopwood DA. 2003. Synergy and contingency as driving forces for the evolution of multiple secondary metabolite production by *Streptomyces* species. *PNAS*. 100:14555–14561.
- 9 Becher PG, et al. 2020. Developmentally regulated volatiles geosmin and 2-methylisoborneol attract a soil arthropod to *Streptomyces* bacteria promoting spore dispersal. *Nat Microbiol*. 5:821–829.
- 10 Noda-Garcia L, Tawfik DS. 2020. Enzyme evolution in natural products biosynthesis: target- or diversity-oriented? *Curr Opin Chem Biol*. 59:147–154.
- 11 Chevrette MG, et al. 2020. Evolutionary dynamics of natural product biosynthesis in bacteria. *Nat Prod Rep*. 37:566–599.
- 12 Fewer DP, Metsä-Ketelä M. 2020. A pharmaceutical model for the molecular evolution of microbial natural products. *FEBS J*. 287:1429–1449.
- 13 Metsä-Ketelä M. 2017. Evolution inspired engineering of antibiotic biosynthesis enzymes. *Org Biomol Chem*. 15:4036–4041.

- 14 Firm RD, Jones CG. 2003. Natural products—a simple model to explain chemical diversity. *Nat Prod Rep.* 20:382–391.
- 15 Bar-Even A, Salah Tawfik D. 2013. Engineering specialized metabolic pathways—is there a room for enzyme improvements? *Curr Opin Biotechnol.* 24:310–319.
- 16 Copley SD. 2015. An evolutionary biochemist’s Perspective on promiscuity. *Trends Biochem Sci.* 40:72–78.
- 17 Bar-Even A, et al. 2011. The moderately efficient enzyme: evolutionary and physicochemical trends shaping enzyme parameters. *Biochemistry.* 50:4402–4410.
- 18 Peracchi A. 2018. The limits of enzyme specificity and the evolution of metabolism. *Trends Biochem Sci.* 43:984–996.
- 19 Fischer I, et al. 2014. Impact of recurrent gene duplication on adaptation of plant genomes. *BMC Plant Biol.* 14:151.
- 20 Niemi J, Metsä-Ketelä M, Schneider G, Mäntsälä P. 2008. Biosynthetic anthracycline variants. *Top Curr Chem.* 282:75–99.
- 21 Elshahawi SI, Shaaban KA, Kharel MK, Thorson JS. 2015. A comprehensive review of glycosylated bacterial natural products. *Chem Soc Rev.* 44:7591–7697.
- 22 Metsä-Ketelä M, Niemi J, Mäntsälä P, Schneider G. 2008. Anthracycline biosynthesis: genes, enzymes and mechanisms. *Top Curr Chem.* 282:101–140.
- 23 Tacar O, Sriamornsak P, Dass CR. 2013. Doxorubicin: an update on anticancer molecular action, toxicity and novel drug delivery systems. *J Pharm Pharmacol.* 65:157–170.
- 24 Pommier Y, Leo E, Zhang H, Marchand C. 2010. DNA Topoisomerases and their poisoning by anticancer and antibacterial drugs. *Chem Biol.* 17:421–433.
- 25 Rabinovich GA, Toscano MA. 2009. Turning “sweet” on immunity: galectin-glycan interactions in immune tolerance and inflammation. *Nat Rev Immunol.* 9:338–352.
- 26 Vejpongsa P, Yeh ETH. 2014. Topoisomerase 2 β : a promising molecular target for primary prevention of anthracycline-induced cardiotoxicity. *Clin Pharmacol Ther.* 95:45–52.
- 27 Pang B, et al. 2013. Drug-induced histone eviction from open chromatin contributes to the chemotherapeutic effects of doxorubicin. *Nat Commun.* 4:1908.
- 28 Denard B, Lee C, Ye J. 2012. Doxorubicin blocks proliferation of cancer cells through proteolytic activation of CREB3L1. *Elife.* 1: e00090.
- 29 Grocholski T, et al. 2019. Evolutionary trajectories for the functional diversification of anthracycline methyltransferases. *ACS Chem Biol.* 14:850–856.
- 30 Madduri K, Torti F, Colombo AL, Hutchinson CR. 1993. Cloning and sequencing of a gene encoding carminomycin 4-O- methyltransferase from *Streptomyces peucetius* and its expression in *Escherichia coli*. *J Bacteriol.* 175:3900–3904.
- 31 Grocholski T, Dinis P, Niiranen L, Niemi J, Metsä-Ketelä M. 2015. Divergent evolution of an atypical S-adenosyl-L-methionine-dependent monooxygenase involved in anthracycline biosynthesis. *PNAS.* 112:9866–9871.
- 32 Jansson A, et al. 2005. Aclacinomycin 10-hydroxylase is a novel substrate-assisted hydroxylase requiring S-adenosyl-L-methionine as cofactor. *J Biol Chem.* 280:3636–3644.
- 33 Dinis P, Wandt BN, Grocholski T, Metsä-Ketelä M. 2019. Chimeragenesis for biocatalysis. In: Ram Sarup Singh CL, Singhanian RR, Pandey A, editors. *Biomass, biofuels, biochemicals: advances in enzyme technology.* Amsterdam: Elsevier. p. 389–418.
- 34 Kraus J. 1978. The biochemistry of adenosylmethionine. *Yale J Biol Med.* 51:669–670.
- 35 Gray VE, Hause RJ, Fowler DM. 2017. Analysis of large-scale mutagenesis data to assess the impact of single amino acid substitutions. *Genetics.* 207:53–61.
- 36 Dickens ML, Priestley ND, Strohl WR. 1997. In vivo and in vitro bioconversion of ϵ -rhodomycinone glycoside to doxorubicin: functions of DauP, DauK, and DoxA. *J Bacteriol.* 179:2641–2650.
- 37 Jansson A, Koskiniemi H, Mäntsälä P, Niemi J, 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.
- 38 Jansson A, Niemi J, Lindqvist Y, Mäntsälä P, Schneider G. 2003. Crystal structure of aclacinomycin-10-hydroxylase, a S-adenosyl-L- methionine-dependent methyltransferase homolog involved in anthracycline biosynthesis in *Streptomyces purpurascens*. *J Mol Biol.* 334:269–280.
- 39 Aharoni A, et al. 2005. The “evolvability” of promiscuous protein functions. *Nat Genet.* 37:73–76.
- 40 Copley SD. 2014. An evolutionary perspective on protein moonlighting. *Biochem Soc Trans.* 42:1684–1691.
- 41 Jansson A, Niemi J, Mäntsälä P, Schneider G. 2003. Crystal structure of aclacinomycin methylesterase with bound product analogues. Implications for anthracycline recognition and mechanism. *J Biol Chem.* 278:39006–39013.
- 42 Ylihonko K, Hakala J, Niemi J, Lundell J, Mantsala P. 1994. Isolation and characterization of aclacinomycin A-non-producing *Streptomyces galilaeus* (ATCC 31615) mutants. *Microbiology.* 140: 1359–1365.
- 43 Kabsch W, et al. 2010. XDS. *Acta Crystallogr Sect D Biol Crystallogr.* 66:125–132.
- 44 Evans PR, Murshudov GN. 2013. How good are my data and what is the resolution? *Acta Crystallogr Sect D Biol Crystallogr.* 69: 1204–1214.
- 45 Sawaya MR. 2014. Methods to refine macromolecular structures in cases of severe diffraction anisotropy. *Methods Mol Biol.* 1091: 205–214.
- 46 Adams PD, et al. 2010. PHENIX: a comprehensive python-based system for macromolecular structure solution. *Acta Crystallogr Sect D Biol Crystallogr.* 66:213–221.