Divergent evolution of an atypical *S*-adenosyl-L-methionine–dependent monooxygenase involved in anthracycline biosynthesis

Thadée Grocholski, Pedro Dinis, Laila Niiranen, Jarmo Niemi, and Mikko Metsä-Ketelä[1]

Department of Biochemistry, University of Turku, FIN-20014 Turku, Finland

Edited by Frank M. Raushel, Texas A&M University, College Station, TX, and accepted by the Editorial Board July 7, 2015 (received for review January 27, 2015)

Bacterial secondary metabolic pathways are responsible for the biosynthesis of thousands of bioactive natural products. Many enzymes residing in these pathways have evolved to catalyze unusual chemical transformations, which is facilitated by an evolutionary pressure promoting chemical diversity. Such divergent enzyme evolution has been observed in *S*-adenosyl-L-methionine (SAM)-dependent methyltransferases involved in the biosynthesis of anthracycline anticancer antibiotics; whereas DnrK from the daunorubicin pathway is a canonical 4-*O*-methyltransferase, the closely related RdmB (52% sequence identity) from the rhodomycin pathways is an atypical 10-hydroxylase that requires SAM, a thiol reducing agent, and molecular oxygen for activity. Here, we have used extensive chimeragenesis to gain insight into the functional differentiation of RdmB and show that insertion of a single serine residue to DnrK is sufficient for introduction of the monooxygenation activity. The crystal structure of DnrK-Ser in complex with aclacinomycin T and *S*-adenosyl-L-homocysteine refined to 1.9-Å resolution revealed that the inserted serine S297 resides in an α-helical segment adjacent to the substrate, but in a manner where the side chain points away from the active site. Further experimental work indicated that the shift in activity is mediated by rotation of a preceding phenylalanine F296 toward the active site, which blocks a channel to the surface of the protein that is present in native DnrK. The channel is also closed in RdmB and may be important for monooxygenation in a solvent-free environment. Finally, we postulate that the hydroxylation ability of RdmB originates from a previously undetected 10-decarboxylation activity of DnrK.

enzyme evolution | protein engineering | *Streptomyces* | polyketide | natural products

Anthracyclines are a medically important class of natural products, which belong to the type II family of aromatic polyketides (1–3). These compounds, produced by *Streptomyces* bacteria, are among the most effective anticancer drugs available and are currently included in 500 clinical trials worldwide (4). Doxorubicin is widely used as a first-choice chemotherapeutic agent for many tumors, including various carcinomas and sarcomas (5), whereas aclacinomycin A has been used in the treatment of hematological malignancies (6). The biological activity of doxorubicin and aclacinomycin is mediated through topoisomerase II (7), while additional mechanisms of these drugs include the recently discovered histone eviction activity (4). Despite the success of anthracyclines in cancer chemotherapy, improved compounds are still urgently needed, because the clinical efficacy of these metabolites has been limited by severe side effects (8, 9).

The great diversity of anthracyclines is generated during tailoring steps in their biosynthesis, where glycosylations, hydroxylations, decarboxylations, and methylations modify the common 7,8,9,10-tetrahydro-5,12-naphthacenoquinone aglycone chromophore to obtain the biologically active molecules of the various pathways (1–3). A detailed understanding of these biosynthetic steps is essential for the rational design of improved bioactive metabolites by pathway and protein engineering. In particular, the discovery of thousands of unknown biosynthetic pathways by next-generation sequencing and the emergence of synthetic biology holds great promise for increasing the chemical space of natural products (10). However, identification of the function of gene products residing in these pathways by bioinformatic analysis alone is complicated by the atypical chemistry the corresponding enzymes may catalyze. Unexpected chemical transformations appear to be especially common in the biosynthesis of aromatic polyketides, where biochemical studies have frequently revealed unusual reaction mechanisms (2, 11), which includes instances where the reactivity of the substrates themselves are used in catalysis (12–14). In addition, homologous enzymes have been noted to catalyze diverse chemical transformations (15). Recently, the use of protein engineering for altering the reaction specificities of biosynthetic enzymes has emerged as another developing field for modification of natural products (16, 17).

This unusually rapid differentiation of enzyme function may be related to the evolution of secondary metabolism, where traits enabling a swift increase in the diversity of natural products are beneficial (18). To promote chemical diversity rather than formation of a single product, specialized metabolism enzymes appear to have evolved as generalists, which is manifested as promiscuity and slow reaction rates typical to these proteins (19). One intriguing example of such divergent evolution is the enzyme pair DnrK and RdmB from the daunorubicin and rhodomycin

Significance

Natural products produced by *Streptomyces* are widely used in the treatment of various medical conditions. Over the years, thousands of metabolites with complex chemical structures have been isolated from cultures of these soil bacteria. An evolutionary pressure that promotes chemical diversity appears to be critical for generation of this rich source of biologically active compounds. This is reflected in the biosynthetic enzymes, where functions of similar proteins may greatly differ. Here, we have clarified the molecular basis of how a classical methyltransferase has evolved into an unusual hydroxylase on the biosynthetic pathways of two anthracycline anticancer agents. Detailed understanding of enzymes involved in antibiotic biosynthesis will facilitate future protein engineering efforts for generation of improved bioactive natural products.

Author contributions: T.G., J.N., and M.M.-K. designed research; T.G., P.D., L.N., and M.M.-K. performed research; T.G., P.D., L.N., J.N., and M.M.-K. analyzed data; and M.M.-K. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission. F.M.R. is a guest editor invited by the Editorial Board.

Data deposition: The atomic coordinates and structure factors have been deposited in the Protein Data Bank, www.pdb.org (PDB ID code 4WXH).

[1]To whom correspondence should be addressed. Email: mikko.mk@gmail.com.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1501765112/-/DCSupplemental.

pathways, respectively, that are homologous to classical *S*-adenosyl-L-methionine (SAM)-dependent methyltransferases.

Methyltransferase DnrK is a true methyltransferase that catalyzes the 4-*O*-methylation of carminomycin (**1**; Fig. 1) in one of the final steps in the biosynthesis of the antitumor drug daunorubicin (**2**; Fig. 1) in *Streptomyces peucetius* (20, 21). The enzyme possesses rather relaxed substrate specificity in regard to modifications in the polyaromatic anthracycline ring system, but it is quite specific with respect to the length of the carbohydrate chain at C-7, accepting only monoglycosides (22). In addition, DnrK has even been shown to be able to methylate various flavonoids (23). The crystal structure of the ternary complex of DnrK revealed that the mechanism of methylation is consistent with canonical S_N2-type reactions of SAM-dependent methyltransferases, although proximity and orientation effects rather than acid/base catalysis appear to be more critical in the case of DnrK (22).

On the other hand, RdmB from the β-rhodomycin pathway in *Streptomyces purpurascens* is a very unusual enzyme that is evolutionarily closely related to DnrK with a sequence identity of 52% (24). Determination of the crystal structure of RdmB confirmed that the enzyme binds SAM (25) and that the ternary complex is indeed very similar to that of DnrK with a root-mean-square deviation of 1.14 Å for 335 equivalent Cα atoms (26). Despite these features, RdmB completely lacks methyltransferase activity (24, 27, 28); it is an anthracycline 10-hydroxylase, which requires SAM, molecular oxygen, and a thiol reducing agent for activity (26). Contrary to DnrK, RdmB is able to use both monoglycosylated and triglycosylated anthracyclines as substrates.

Furthermore, both rhodomycin and daunorubicin gene clusters encode homologous 15-methylesterases RdmC (28) and DnrP (21), respectively. On the rhodomycin pathway, RdmB works strictly in conjunction with RdmC, which first converts ε-rhodomycin (**3**; Fig. 1) into 15-demethoxy-ε-rhodomycin (**4**; Fig. 1), whereas RdmB is responsible for conversion of the RdmC reaction product into β-rhodomycin (**5**; Fig. 1) (28). This is in contrast to the daunorubicin pathway, where the reaction order of the two enzymes DnrP and DnrK is more relaxed, as DnrK is able to methylate at least five biosynthetic intermediates (21).

To gain insight into how a methyltransferase has evolved into a monooxygenase, we report here the generation of several chimeric DnrK/RdmB proteins that were designed based on the available crystal structures. These initial experiments guided us in further structure/function studies, where we were able to demonstrate that the insertion of a single amino acid is sufficient for introduction of the 10-hydroxylase activity into the DnrK scaffold, thus providing an especially illustrative example of divergent enzyme evolution. Protein crystallography and further biochemical experiments were then conducted to reveal the molecular basis for the altered enzymatic activity.

Results and Discussion

Identification of Protein Regions for Chimeragenesis. Structural studies of DnrK and RdmB (22, 25) have confirmed that both enzymes consist of an N-terminal dimerization domain, an α-helical middle domain, and a C-terminal catalytic domain that binds the cofactor SAM (Fig. 1). The substrate binding site is formed between the C-terminal domain, which has a Rossmann-like α/β-fold typical to nucleotide-binding proteins, and the middle domain. As an initial experiment, we fused the dimerization domain of DnrK onto the catalytic domain of RdmB, generating the enzyme variant RdmB-CT. Because the activity of RdmB-CT was not altered and the enzyme catalyzed exclusively 10-hydroxylation (Fig. 2), we were able to focus on amino acid residues residing in the catalytic subunit in subsequent experiments. In the absence of a high-throughput screening system, we decided to create chimeric enzymes by interchanging key subdomain regions ranging from 10 to 12 aa to probe the functional differentiation of the enzyme pair.

Comparison of the two enzymes revealed that most of the amino acid residues residing in the vicinity of the binding site for SAM were highly conserved and that differences could be mainly attributed to three key regions around the putative entrance to the active site cavity (Fig. 1). The first region (R1; Fig. S1) selected to the study resides near the C-10 position of the polyaromatic anthracycline ligand and consists of a loop region and the N-terminal half of α16 from the middle domain (Fig. 1). This loop region is disordered in the RdmB–SAM binary complex (25), whereas it is well defined in the ternary complex containing 11-deoxy-β-rhodomycin A (**6**; Fig. 1) (26). In addition, in comparison with DnrK (22), the loop region has moved closer to the ligand in RdmB, suggesting that R1 may be important for the 10-hydroxylation activity and/or the entry of the ligand into the active site. Second, moderate differences can be observed in the loop region between β8 and β9 (R2; Fig. S1), which folds over the ligands (Fig. 1), and as such might be responsible for assuring the correct orientation of the substrate. Finally, the region between α11 and α12 defines a restricted space for binding of the carbohydrate unit of

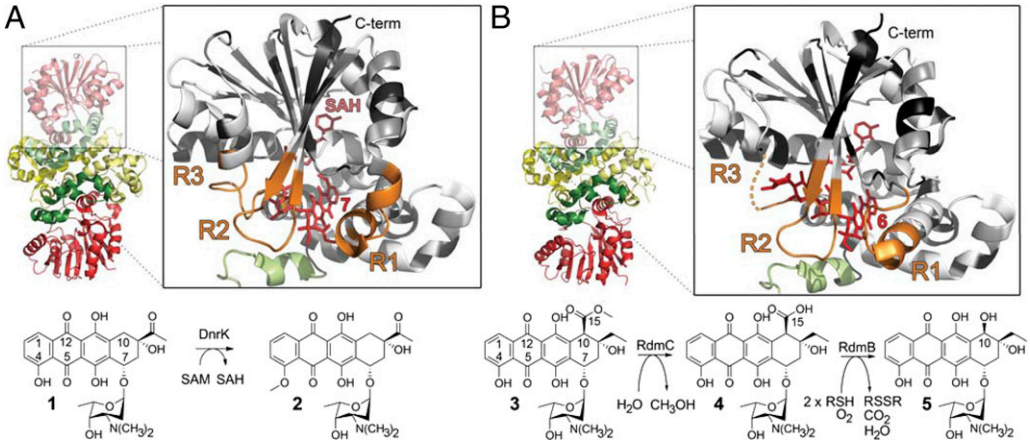

Fig. 1. Similar structures, distinct functions. The structures and reactions catalyzed by (*A*) DnrK and (*B*) RdmB. In the overall structures, the N-terminal dimerization, the middle, and the catalytic domains are shown in yellow, green, and red, respectively. In the *Inset*, the catalytic domains are color coded based on dissimilar (black), similar (gray), and identical (white) residues. SAH and the ligands are shown in red, whereas the three regions selected for chimeragenesis (R1, R2, and R3) are shown in orange. The R3 region of RdmB is disordered in the crystal structure and is shown as a dashed orange line in the *Inset*. In the reaction scheme, "R" represents a thiol reducing agent such as DTT or glutathione.

4-methoxy-ε-rhodomycin T (**7**; Fig. 1) in the DnrK structure (Fig. 1), which has been noted as the likely explanation for why the enzyme only accepts monoglycosylated substrates (22). The corresponding region in RdmB is rather different and the carbohydrate chain extends from the binding pocket into the bulk solution in the RdmB ternary complex (26), which prompted us to select this region (R3; Fig. S1) as the third segment for chimeragenesis.

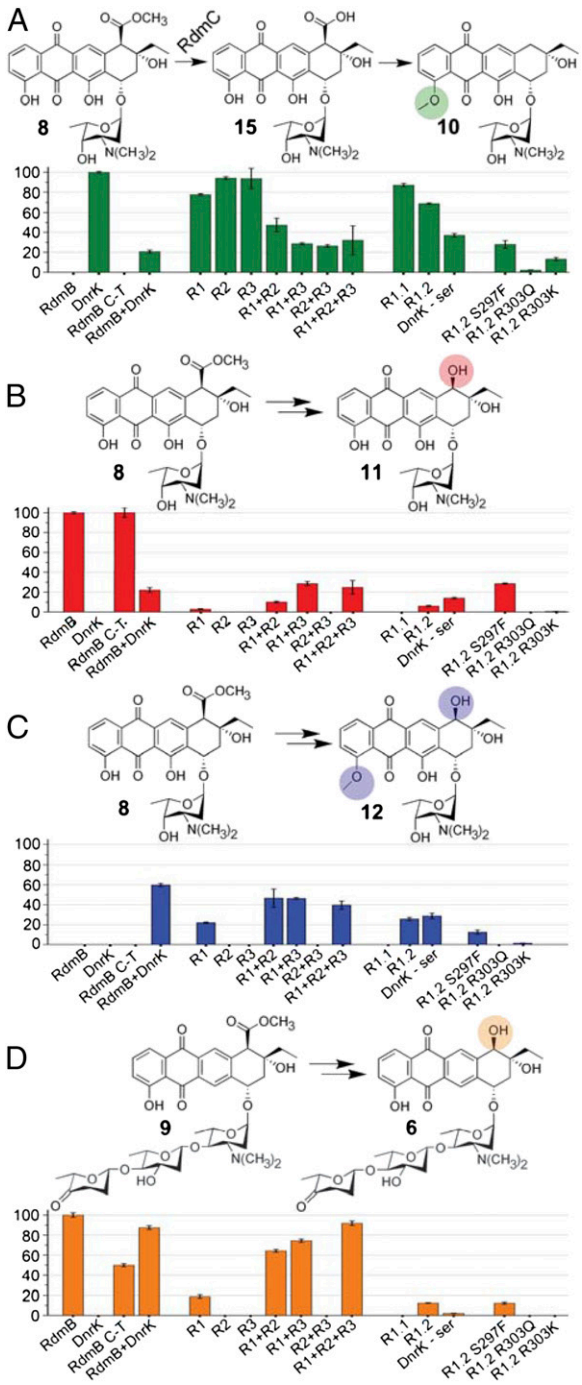

Fig. 2. Relative activities of native and engineered enzymes in a coupled reaction with RdmC. The reactions with **8** lead to the accumulation of three products (*A*–*C*), whereas in the case of **9** only one product was observed (*D*). The columns present the formation of (*A*) 4-*O*-methyl-15-decarboxyaclacinomycin T (**10**), (*B*) 11-deoxy-β-rhodomycin T (**11**), (*C*) 4-*O*-methyl-11-deoxy-β-rhodomycin T (**12**), and (*D*) 11-deoxy-β-rhodomycin A (**6**) by the enzymes. The overall percentages may not add up to 100% in all samples because some unreacted substrate **15** was left with chimeras harboring poor activity (e.g., DnrK R2 + R3). The excess substrate was converted to the corresponding 10-decarboxylated compounds by exposure to light (overall conversion of **8** to **13** and **9** to **14**; Fig. S2) to facilitate the HPLC analyses.

Enzymatic Activities of the Native Enzymes. The activity assays were performed in a coupled reaction together with the 15-methylesterase RdmC using either the monoglycosylated aclacinomycin T (**8**; Fig. 2) or the triglycosylated aclacinomycin A (**9**; Fig. 2) as substrates. The reaction mixtures also included SAM, which is essential as a cosubstrate for DnrK (22) and as a cofactor for RdmB (26), and a reducing agent (DTT) that is required solely by RdmB (26). The reaction with native DnrK resulted in a single main product (Fig. 2*A*), which was identified as the 4-*O*-methylated and 10-decarboxylated aclacinomycin T (**10**; Figs. S2 and S3, and Table S1), when **8** was used as a substrate, whereas no methylation activity could be observed with the triglycosylated **9**. In agreement with previous experiments (27), addition of native RdmB to the assays with **8** and **9** led to the accumulation of the expected 10-hydroxylated products 11-deoxy-β-rhodomycin T (**11**; Fig. 2*B*) and **6** (Fig. 2*D*), respectively.

Discovery of a Moonlighting 10-Decarboxylation Activity of DnrK. The RdmC and DnrP reaction products with a free 10-carboxyl group (compounds such as **4** in Fig. 1) have been described as unstable and undergo spontaneous 10-decarboxylation on the rhodomycin and daunorubicin pathways, respectively (21, 29). However, indirect evidence has suggested that the decarboxylation might also be enzymatically catalyzed by DnrP (21). During our assays with **8**, we noted that the RdmC reaction product, 15-demethoxy-aclacinomycin T (**15**; Fig. 2) is stable if the enzymatic reactions are conducted in the dark, but exposure to light is sufficient to promote the 10-decarboxylation reaction to proceed to near completion in 60 min resulting in 10-decarboxymethylaclacinomycin T (**13**; Fig. S3). To our surprise, we next observed that if **15** is isolated and incubated in the dark together with DnrK, both 4-*O*-methylation and 10-decarboxylation reactions occur, because **10** is observed as the sole product (Fig. S3). The result therefore demonstrated directly that it is DnrK instead of DnrP that harbors the moonlighting 10-decarboxylation activity, which in our view has important evolutionary implications for the emergence of 10-hydroxylation activity (see below).

Enzymatic Activities of the Chimeric Proteins. The single chimeric DnrK enzymes, where the amino acid residues in the R1, R2, or R3 regions were changed to correspond to RdmB, behaved much like native DnrK in the activity measurements and led to the accumulation of **10** (Fig. 2*A*) as the main product. However, the reaction mixture of DnrK R1 contained a small amount of the expected RdmB reaction product, **11** (Fig. 2*B*), and significant quantities of a compound that was identified as the double DnrK/RdmB reaction product, 4-*O*-methyl-11-deoxy-β-rhodomycin T (**12**; Fig. 2*C*). Neither DnrK nor RdmB alone produced **12**, but when both native enzymes were incubated together with RdmC, the same metabolite was observed (Fig. 2*C*). The double and triple chimeric enzymes reinforced our initial results; chimeric proteins that contained the R1 region were able to catalyze 10-hydroxylation, whereas the R2 + R3 chimera was only able to perform 4-*O*-methylation. These multiple exchanges enhanced the RdmB-like activity in comparison with the single R1 chimera, with the R3 region having a slightly greater effect than the R2 region (Fig. 2*B*).

A similar trend was observed when **9** was used as a substrate and the 10-hydroxylated reaction product **6** was observed only in chimeras containing the R1 region from RdmB. It is noteworthy that the increased rate of 10-hydroxylation correlated exceedingly well with the number of amino acid exchanges from the RdmB template (Fig. 2*D*), which highlights the importance of

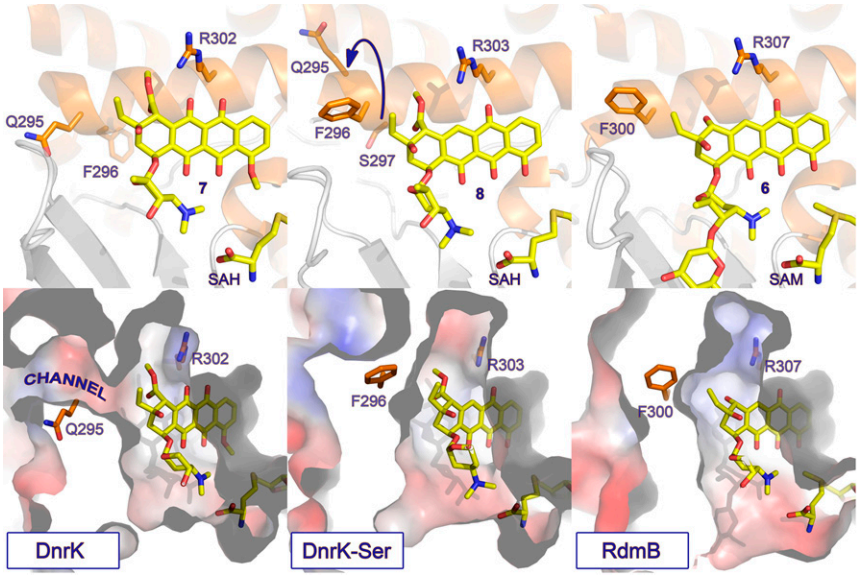

Fig. 3. Cartoon (*Top*) and surface (*Bottom*) representations of the active sites of DnrK in complex with 4-methoxy-ε-rhodomycin T (**7**) and SAH, DnrK-Ser with bound **8** and SAH, and the ternary complex of RdmB, 11-deoxy-β-rhodomycin (**6**), and SAM. Insertion of S297 in DnrK-Ser rotates F296 toward the ligand, which reshapes the DnrK active site toward the RdmB architecture. In particular, DnrK-Ser F296 blocks a channel to the surface of the protein that is open in native DnrK, but closed in RdmB. A conserved arginine residue that is important for both methylation and hydroxylation activities is also shown.

the R2 and R3 regions in correct alignment of the triglycosidic **9** substrate in the active site. Previous studies have shown that the R3 region of DnrK folds around the amino sugar moiety (22), and it was surprising to find that exchange of the R3 region was not essential for the gain of the 10-hydroxylation activity. Despite these changes, no 4-*O*-methylation activity was detected with this substrate in any of the chimeric enzymes, which would indicate that the aglycone is still not correctly positioned in terms of geometry in the active-site cavity. The lack of methylation activity in native RdmB has been suggested to occur due to wrong alignment of the substrate for a methyl transfer reaction using a S_N2 mechanism, where the methyl group has to be in line with the oxygen of the substrate as well as the sulfur atom of SAM for the reaction to proceed (26).

In contrast to the engineered DnrK enzymes, when the equivalent changes were made to the RdmB template, no methylation activity could be detected. The chimeric RdmB enzymes either behaved in a manner similar to the native enzyme or simply lost all enzymatic activity.

Mapping the Minimal Region Required for Gain of 10-Hydroxylation Activity in DnrK. In an attempt to further pinpoint the origin of the 10-hydroxylation activity in the DnrK R1 chimera, we divided the R1 region into two segments corresponding to the loop region and the following α16 helix, which resulted in two additional mutants denoted as DnrK R1.1 and DnrK R1.2, respectively. The 10-hydroxylation activity could be attributed solely to the α16 helix (Fig. 2), because the activity of DnrK R1.2 was similar to that of DnrK R1 (i.e., both methylation and hydroxylation of **8** and hydroxylation of **9**), whereas DnrK R1.1 behaved like native DnrK (i.e., methylation of **8** and no activity with **9**). Inspection of the amino acid sequences of the R1.2 region (Fig. S1) revealed that the RdmB sequence contains an additional serine insertion in this area in comparison with DnrK. Remarkably, introduction of this single S297 residue to the DnrK scaffold was sufficient to persuade the 10-hydroxylation activity to emerge as seen from the activity measurements of the DnrK-Ser variant (Fig. 2).

Structural Basis for the Altered Enzymatic Activity of the DnrK-Ser Mutant. The DnrK-Ser mutant crystallized in a different space group than native DnrK, and consequently the diffraction data could be refined to an improved resolution of 1.9 Å (Table S2). Electron density could be traced for most of the polypeptide chain in the two subunits found in the asymmetric unit with the exception of two amino acids residing in loop regions (D85 and H293) in chain B, where the density in general was less well defined than in chain A. Based on electron density, one molecule of **8** and *S*-adenosyl-L-homocysteine (SAH) was bound to each subunit (Fig. S4).

The overall structure of DnrK-Ser was exceedingly similar to native DnrK with an r.m.s.d. of 0.90 Å over 327 residues, and the main differences were mostly restricted to the R1 region. In short, introduction of the additional S297 residue in R1.2 appeared to trigger a transition of all preceding residues in helix α16 toward the surface of the protein, which altered the structure of the R1.1 loop region. Unexpectedly, the side chain of the inserted serine was found pointing away from the active site, but at the same time the preceding phenylalanine (F296) had rotated inward adjacent to the ligand near the site of monooxygenation (Fig. 3). Comparison of the structures of DnrK-Ser and RdmB reveals that both enzymes have a phenylalanine, F296 and F300, respectively, next to the anthracycline ligand, which is in contrast to native DnrK, which contains glutamine Q295 at the equivalent position (Fig. 3). The implication of these changes is that the glutamine in DnrK allows bulk solvent to access the active site cavity via an open channel, whereas the bulkier hydrophobic phenylalanine residues block this route to the surface of the protein in DnrK-Ser and RdmB (Fig. 3). The structural analysis therefore suggested that critical to the shift in activity is rotation of F296 and that the amino acid inserted at position 297 would not necessarily have to be serine. Consistent with this hypothesis, reintroduction of the original phenylalanine observed in this structural position in native DnrK led to hydroxylation of both substrates by DnrK R1.2 S297F (Fig. 2 *B* and *D*).

Mechanistic Divergence of the DnrK and RdmB Reactions. The chimeragenesis studies of DnrK and RdmB permit us to propose the following mechanistic proposal for the divergent reactions catalyzed by the two enzymes (Fig. 4). It is tempting to suggest that a common initial step is the decarboxylation of the substrate **15** to generate a carbanion intermediate. The more open active site of DnrK allows facile protonation of the carbanion by solvent molecules and results in the formation of a neutral 10-decarboxylated and 4-*O*-methylated anthracycline **10**. However, such compounds are no longer suitable as substrates for RdmB (26), which avoids the protonation step by closing the channel to the surface of the protein and excluding solvent ions from the active site. The carbanion may be stabilized through distribution

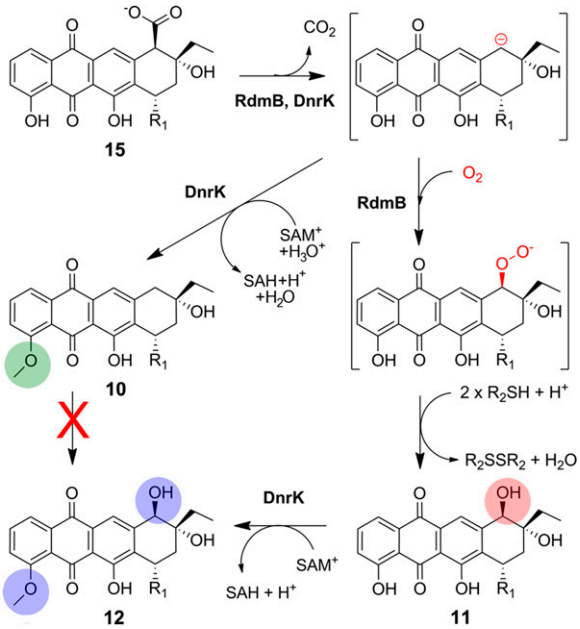

Fig. 4. Mechanistic proposal for the reactions catalyzed by DnrK and RdmB. Both DnrK and RdmB are postulated to catalyze the decarboxylation of **15** as a first step. In DnrK, the open active site would allow protonation of the resulting carbanion by solvent molecules and 4-*O*-methylation to generate **10**. In contrast, RdmB stabilizes the carbanion and catalyzes the formation of an O_2–anthracycline caged radical pair in the closed active site. Consequent formation of a peroxyl intermediate and subsequent reduction of the peroxide by intracellular thiols possibly outside the active site of the enzyme yields the 10-hydroxylated product **11**. The reaction sequence leading to the double-product **12** is only possible if the RdmB reaction happens first. R_1, L-rhodosamine.

of the negative charge throughout the polyphenolic ring system and the adjacent positive charge of SAM, which is the reason why the enzyme requires the cofactor for catalysis (26). The delocalization of electrons enables substrate-assisted activation of molecular oxygen to overcome the spin-barrier of dioxygen reactivity in manner similar to other cofactor-independent monooxygenases (12–14, 30, 31). The exclusion of solvent ions from the active site may also be essential for this step. Once the carbanion has reacted with dioxygen, the 10-peroxy anthracycline formed is reduced by a thiol reducing reagent possibly outside the active-site cavity to generate the 10-hydroxylated end product **11** (26). It is important to note that formation of the DnrK/RdmB double-reaction product **12** must proceed through initial 10-hydroxylation followed by 4-*O*-methylation (Fig. 4), because neutral anthracyclines are suitable substrates only for DnrK (21).

It has been proposed that arginine R307 near the R1 region in RdmB (Fig. 3), which is also conserved in DnrK (R302), has an important role for the initial decarboxylation of the substrate (26). The discovery of the 10-decarboxylation activity for DnrK prompted us to study the importance of this residue for the activity of DnrK R1.2. Expectedly, the mutants DnrK R1.2 R303K and R303Q displayed greatly reduced activities for both 4-*O*-methylation and 10-hydroxylation. The 10-hydroxylation activity was completely lost in the R303Q mutant, whereas only trace activities remained in the R303K mutant when **8** was used as a substrate (Fig. 2). The 4-*O*-methylation activity was also affected, although both mutants still harbored ~10% of their activities. In addition, changing the arginine R302 to isoleucine in DnrK has been described to greatly reduce the methylation activity of flavonoids (23), confirming the importance of this residue for activity.

Evolutionary Implications. The intimate connection between the 15-methylesterase reaction to both 4-*O*-methylation and 10-hydroxylation may be observed from analysis of eight gene clusters putatively involved in anthracycline biosynthesis (Fig. S5). In all cases, the gene order is conserved, indicating possible coevolution of the proteins. In our view, the ancestral configuration may have been a daunorubicin-type system where one enzyme catalyzes anthracycline 15-demethylation, whereas another protein promotes 4-*O*-methylation and 10-decarboxylation under light-deprived conditions in the soil. The 10-hydroxylation activity may then have evolved from the moonlighting 10-decarboxylation activity upon changes in the R1 region leading to generation of more diversity in anthracycline scaffolds. Further changes in other parts of the protein have then led to full conversion of the enzymatic activity.

Our hypothesis is supported by sequence analysis of the R1 area, which indicates that the region is an evolutionary hot spot in these enzymes (Fig. 5). Three sequences, including RdmB, are rhodomycin-type with a serine insertion, whereas DnrK appears to be the sole daunorubicin-type sequence. Interestingly, three of the proteins of unknown function contain an insertion of "TSDLY" where the tyrosine may be positioned in an equivalent manner to the phenylalanine in DnrK-Ser, and these proteins may represent an alternative evolutionary path to 10-hydroxylated anthracyclines. The final sequence is even more distantly related with an insertion of "TADLH" with a histidine at the critical position, making prediction of the activity of this protein challenging in silico. Investigations into this previously unidentified group of putative anthracycline methyltransferases are currently in progress in our laboratory.

Materials and Methods

General DNA Techniques. Plasmids containing native *dnrK* and *rdmB* (22, 25) were used as templates for all of the PCRs. DnrK R1.2 S297F, R303Q, and R303K mutants were ordered as synthetic genes from GeneArt (Strings DNA Fragments). The remaining DnrK mutants and RdmB-CT were generated using the four-primer overhang extension method (32). Details regarding the cloning experiments, including generation of the RdmC expression construct, are presented in *SI Materials and Methods*. Oligonucleotides used are listed in Table S3. The gene encoding RdmB-CT was cloned into the pBAD/His B vector (Invitrogen), whereas all other enzymes were expressed from the modified pBHBΔ vector (33). All DNA sequences were confirmed by sequencing before protein expression.

Protein Expression and Purification. *Escherichia coli* cell strain and expression conditions were performed as previously described (34), except LB media was used and cells were collected 5 h after induction. Purification, yields, and storage of all enzymes are described in *SI Materials and Methods*.

Enzyme Activity Measurements. All activity measurements were performed in 100 mM Tris·HCl (pH 7.5), 10 mM DTT, and 350 μM SAM with 140 μM **8** or **9**. The concentration of RdmC was set to 100 nM in the reactions, whereas all other enzymes were used at 5.8 μM. Further details are described in *SI Materials and Methods*.

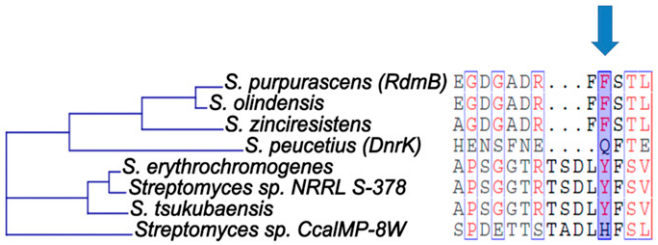

Fig. 5. A phylogenetic tree of eight putative SAM-dependent methyltransferases residing in anthracycline biosynthesis gene clusters and multiple sequence alignment of the R1 area. Amino acids marked with an arrow in the blue column are predicted to be situated at the critical position near C-10 of the anthracycline ligands. Further details regarding these unknown gene clusters and their accession numbers are shown in Fig. S5.

Analysis of Metabolites. The various reaction products were analyzed and identified as described in *SI Materials and Methods*, Fig. S2, and Tables S1 and S4. All reactions were monitored by HPLC using a SCL-10Avp/SpdM10Avp system with a diode array detector (Shimadzu). High-resolution MS (micrOTOF-Q; Bruker Daltonics; linked to an Agilent Technologies 1200 series HPLC system) was used for confirmation of the molecular formulae.

Crystallization, Data Collection, and Structure Determination of DnrK-Ser. Crystallization and structure determination of DnrK-Ser is described in *SI Materials and Methods*. Diffraction data were collected at the European Synchrotron Radiation Facility (ESRF) (Grenoble, France) at beamline ID23-1. Details of the data collection and refinement statistics are listed in Table S2. Figures depicting protein structures were prepared using PyMOL (The PyMOL Molecular Graphics System, version 1.3; Schrödinger, LLC).

ACKNOWLEDGMENTS. We thank Elina Taipalus and Jonna Pörsti for assistance with protein crystallization, and Vilja Siitonen, MSc, for the high-resolution mass spectrometry measurements. We acknowledge access to synchrotron radiation at ESRF (Grenoble, France). This study was supported by Academy of Finland Grant 136060 (to M.M.-K.).

1. Metsä-Ketelä M, Niemi J, Mäntsälä P, Schneider G (2008) Anthracycline biosynthesis: Genes, enzymes and mechanisms. *Anthracycline Chemistry and Biology I: Biological Occurence and Biosynthesis, Synthesis and Chemistry*, ed Krohn K (Springer, Berlin), pp 101–140.
2. Hertweck C, Luzhetskyy A, Rebets Y, Bechthold A (2007) Type II polyketide synthases: Gaining a deeper insight into enzymatic teamwork. *Nat Prod Rep* 24(1):162–190.
3. Hutchinson CR (1997) Biosynthetic studies of daunorubicin and tetracenomycin C. *Chem Rev* 97(7):2525–2536.
4. Pang B, et al. (2013) Drug-induced histone eviction from open chromatin contributes to the chemotherapeutic effects of doxorubicin. *Nat Commun* 4:1908.
5. Weiss RB (1992) The anthracyclines: Will we ever find a better doxorubicin? *Semin Oncol* 19(6):670–686.
6. Wei G, et al. (2011) A meta-analysis of CAG (cytarabine, aclarubicin, G-CSF) regimen for the treatment of 1029 patients with acute myeloid leukemia and myelodysplastic syndrome. *J Hematol Oncol* 4:46.
7. Nitiss JL (2009) Targeting DNA topoisomerase II in cancer chemotherapy. *Nat Rev Cancer* 9(5):338–350.
8. Singal PK, Iliskovic N (1998) Doxorubicin-induced cardiomyopathy. *N Engl J Med* 339(13):900–905.
9. Kaye S, Merry S (1985) Tumour cell resistance to anthracyclines—a review. *Cancer Chemother Pharmacol* 14(2):96–103.
10. Mitchell W (2011) Natural products from synthetic biology. *Curr Opin Chem Biol* 15(4): 505–515.
11. Metsä-Ketelä M, Oja T, Taguchi T, Okamoto S, Ichinose K (2013) Biosynthesis of pyranonaphthoquinone polyketides reveals diverse strategies for enzymatic carbon-carbon bond formation. *Curr Opin Chem Biol* 17(4):562–570.
12. Oja T, et al. (2012) Biosynthetic pathway toward carbohydrate-like moieties of alnumycins contains unusual steps for C-C bond formation and cleavage. *Proc Natl Acad Sci USA* 109(16):6024–6029.
13. Grocholski T, et al. (2010) Crystal structure of the cofactor-independent monooxygenase SnoaB from *Streptomyces nogalater*: Implications for the reaction mechanism. *Biochemistry* 49(5):934–944.
14. Siitonen V, Blauenburg B, Kallio P, Mäntsälä P, Metsä-Ketelä M (2012) Discovery of a two-component monooxygenase SnoaW/SnoaL2 involved in nogalamycin biosynthesis. *Chem Biol* 19(5):638–646.
15. Patrikainen P, et al. (2012) Tailoring enzymes involved in the biosynthesis of angucyclines contain latent context-dependent catalytic activities. *Chem Biol* 19(5): 647–655.
16. Williams GJ, Gantt RW, Thorson JS (2008) The impact of enzyme engineering upon natural product glycodiversification. *Curr Opin Chem Biol* 12(5):556–564.
17. Zabala AO, Cacho RA, Tang Y (2012) Protein engineering towards natural product synthesis and diversification. *J Ind Microbiol Biotechnol* 39(2):227–241.
18. Firn RD, Jones CG (2000) The evolution of secondary metabolism—a unifying model. *Mol Microbiol* 37(5):989–994.
19. Bar-Even A, Salah Tawfik D (2013) Engineering specialized metabolic pathways—is there a room for enzyme improvements? *Curr Opin Biotechnol* 24(2):310–319.
20. 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(12):3900–3904.
21. 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(8):2641–2650.
22. 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(39):41149–41156.
23. Kim NY, et al. (2007) *O*-Methylation of flavonoids using DnrK based on molecular docking. *J Microbiol Biotechnol* 17(12):1991–1995.
24. Niemi J, Mäntsälä P (1995) Nucleotide sequences and expression of genes from *Streptomyces purpurascens* that cause the production of new anthracyclines in *Streptomyces galilaeus*. *J Bacteriol* 177(10):2942–2945.
25. 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(2):269–280.
26. 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(5): 3636–3644.
27. Wang Y, et al. (2000) Modifications of aclacinomycin T by aclacinomycin methyl esterase (RdmC) and aclacinomycin-10-hydroxylase (RdmB) from *Streptomyces purpurascens*. *Biochim Biophys Acta* 1480(1-2):191–200.
28. Wang Y, Niemi J, Mäntsälä P (2002) Modification of aklavinone and aclacinomycins in vitro and in vivo by rhodomycin biosynthesis gene products. *FEMS Microbiol Lett* 208(1):117–122.
29. 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(40):39006–39013.
30. Thierbach S, et al. (2014) Substrate-assisted O_2 activation in a cofactor-independent dioxygenase. *Chem Biol* 21(2):217–225.
31. Fetzner S, Steiner RA (2010) Cofactor-independent oxidases and oxygenases. *Appl Microbiol Biotechnol* 86(3):791–804.
32. Ho SN, Hunt HD, Horton RM, Pullen JK, Pease LR (1989) Site-directed mutagenesis by overlap extension using the polymerase chain reaction. *Gene* 77(1):51–59.
33. Kallio P, Sultana A, Niemi J, Mäntsälä P, Schneider G (2006) Crystal structure of the polyketide cyclase AknH with bound substrate and product analogue: Implications for catalytic mechanism and product stereoselectivity. *J Mol Biol* 357(1):210–220.
34. Koskiniemi H, Grocholski T, Schneider G, Niemi J (2009) Expression, purification and crystallization of the cofactor-independent monooxygenase SnoaB from the nogalamycin biosynthetic pathway. *Acta Crystallogr Sect F Struct Biol Cryst Commun* 65(Pt 3): 256–259.
35. Ylihonko K, Hakala J, Niemi J, Lundell J, Mäntsälä P (1994) Isolation and characterization of aclacinomycin A-non-producing *Streptomyces galilaeus* (ATCC 31615) mutants. *Microbiology* 140(Pt 6):1359–1365.
36. Niemi J, et al. (1994) Hybrid anthracycline antibiotics: Production of new anthracyclines by cloned genes from *Streptomyces purpurascens* in *Streptomyces galilaeus*. *Microbiology* 140(Pt 6):1351–1358.
37. Kunnari T, Niemi J, Ylihonko K, Mäntsälä P, Hakala J (1997) Hybrid anthracyclines by a genetically engineered *Streptomyces galilaeus* mutant. *Bioorg Med Chem Lett* 7(6): 725–726.
38. Gasteiger E, et al. (2005) Protein identification and analysis tools on the ExPASy server. *The Proteomics Protocols Handbook*, ed Walker JM (Humana, Totowa, NJ), pp 571–607.
39. Kabsch W (2010) Integration, scaling, space-group assignment and post-refinement. *Acta Crystallogr D Biol Crystallogr* 66(Pt 2):133–144.
40. Collaborative Computational Project, Number 4 (1994) The CCP4 suite: Programs for protein crystallography. *Acta Crystallogr D Biol Crystallogr* 50(Pt 5):760–763.
41. Long F, Vagin AA, Young P, Murshudov GN (2008) BALBES: A molecular-replacement pipeline. *Acta Crystallogr D Biol Crystallogr* 64(Pt 1):125–132.
42. Langer G, Cohen SX, Lamzin VS, Perrakis A (2008) Automated macromolecular model building for X-ray crystallography using ARP/wARP version 7. *Nat Protoc* 3(7): 1171–1179.
43. Murshudov GN, Vagin AA, Dodson EJ (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr D Biol Crystallogr* 53(Pt 3): 240–255.
44. Schüttelkopf AW, van Aalten DMF (2004) PRODRG: A tool for high-throughput crystallography of protein-ligand complexes. *Acta Crystallogr D Biol Crystallogr* 60(Pt 8): 1355–1363.
45. Emsley P, Lohkamp B, Scott WG, Cowtan K (2010) Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* 66(Pt 4):486–501.
46. Afonine PV, et al. (2012) Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr D Biol Crystallogr* 68(Pt 4):352–367.
47. Chen VB, et al. (2010) MolProbity: All-atom structure validation for macromolecular crystallography. *Acta Crystallogr D Biol Crystallogr* 66(Pt 1):12–21.
48. Sultana A, et al. (2004) Structure of the polyketide cyclase SnoaL reveals a novel mechanism for enzymatic aldol condensation. *EMBO J* 23(9):1911–1921.
49. Hutchinson CR, Colombo AL (1999) Genetic engineering of doxorubicin production in *Streptomyces peucetius*: A review. *J Ind Microbiol Biotechnol* 23(1):647–652.
50. Sullivan MJ, Petty NK, Beatson SA (2011) Easyfig: A genome comparison visualizer. *Bioinformatics* 27(7):1009–1010.