

Structural and Computational Insights into the Noncanonical Aromatization in Fungal Polyketide Biosynthesis

Hang Wang,[#] Chao Peng,[#] Xiao-Xuan Chen,[#] Hao-Yang Wang, Run Yang, Hao Xiang, Qiu-Fen Hu, Ling Liu, Lung Wa Chung,* Yudai Matsuda,* and Wei-Guang Wang*



Cite This: ACS Catal. 2024, 14, 10796–10805



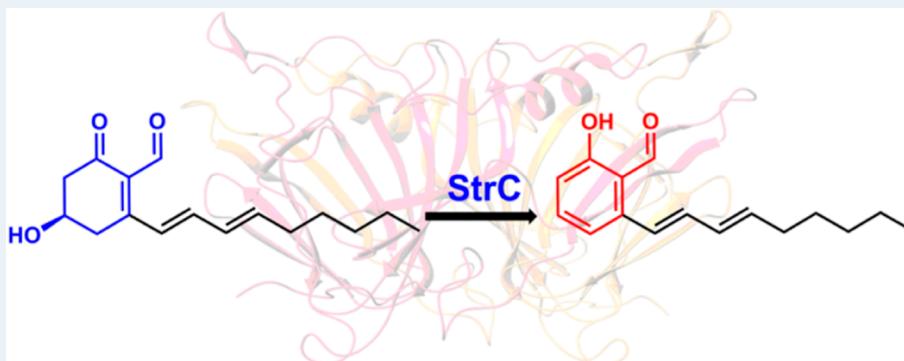
Read Online

ACCESS |

Metrics & More

Article Recommendations

Supporting Information



ABSTRACT: Aromatic polyketides, with their diverse structures and biological activities, have attracted significant scientific interest. Aromatization reactions are typically performed by aromatases and cyclases associated with bacterial type II polyketide synthases (PKSs) or product template and Claisen cyclase domains found in fungal type I nonreducing PKSs. Intriguingly, our recent study discovered a noncanonical aromatization mechanism involving a cupin domain-containing enzyme, StrC, in fungal alkyl salicylaldehyde derivative biosynthesis. Nevertheless, the catalytic function of StrC largely remained enigmatic. In this study, an in-depth *in vitro* characterization of StrC was performed to gain a better understanding of the reaction mechanism for aromatization. In addition, the X-ray crystal structure of StrC in the complex with a substrate analogue was obtained to elucidate the unique active site. Subsequent mutational experiments identified important amino acid residues for aromatase activity. Furthermore, a quantum mechanics study provided a plausible reaction mechanism for aromatization and highlighted the vital role of StrC in the biosynthetic process. Thus, our study unveils an unknown paradigm for aromatization in polyketide biosynthesis.

KEYWORDS: natural products, biosynthesis, polyketides, aromatases, cupin domain-containing enzymes

INTRODUCTION

Polyketides represent a diverse group of natural products and have profoundly impacted human health through the development of therapeutics,¹ including anti-infectives like erythromycin² and amphotericin B;³ immunosuppressants such as tacrolimus⁴ and rapamycin;⁵ and lipid-lowering agents with lovastatin⁶ being a primary example. The central enzymes for polyketide biosynthesis are polyketide synthases (PKSs), which are further classified into types I,⁷ II,⁸ and III⁹ and contribute to the generation of a diverse array of molecular architectures.

Aromatic polyketides, ubiquitously distributed in nature, are one of the major classes of polyketides and are synthesized by all three types of PKSs. Aromatization mechanisms to yield aromatic polyketides vary among the PKS systems. In bacteria, aromatic polyketides are generally formed by type II PKSs, in which aromatases/cyclases play a key role in aromatization, as exemplified by the biosynthesis of tetracenomycin¹⁰ (Figure 1A). On the other hand, the majority of fungal aromatic

polyketides are synthesized by the involvement of type I nonreducing PKSs (NR-PKSs), which adopt an intrinsic product template (PT) domain,¹¹ occasionally along with a thioesterase (TE)-like Claisen cyclase (CLC) domain,¹² to control the cyclization mode of a polyketomethylene chain, as seen in the formation of norsolorinic acid anthrone (Figure 1B). Furthermore, partially reducing PKSs (PR-PKSs), as represented by the 6-methylsalicylic acid synthase (6-MSAS), are also known to afford aromatic polyketides,¹³ although the aromatization mechanisms employed by PR-PKSs remain elusive.

Received: February 16, 2024

Revised: June 3, 2024

Accepted: June 20, 2024

Published: July 3, 2024



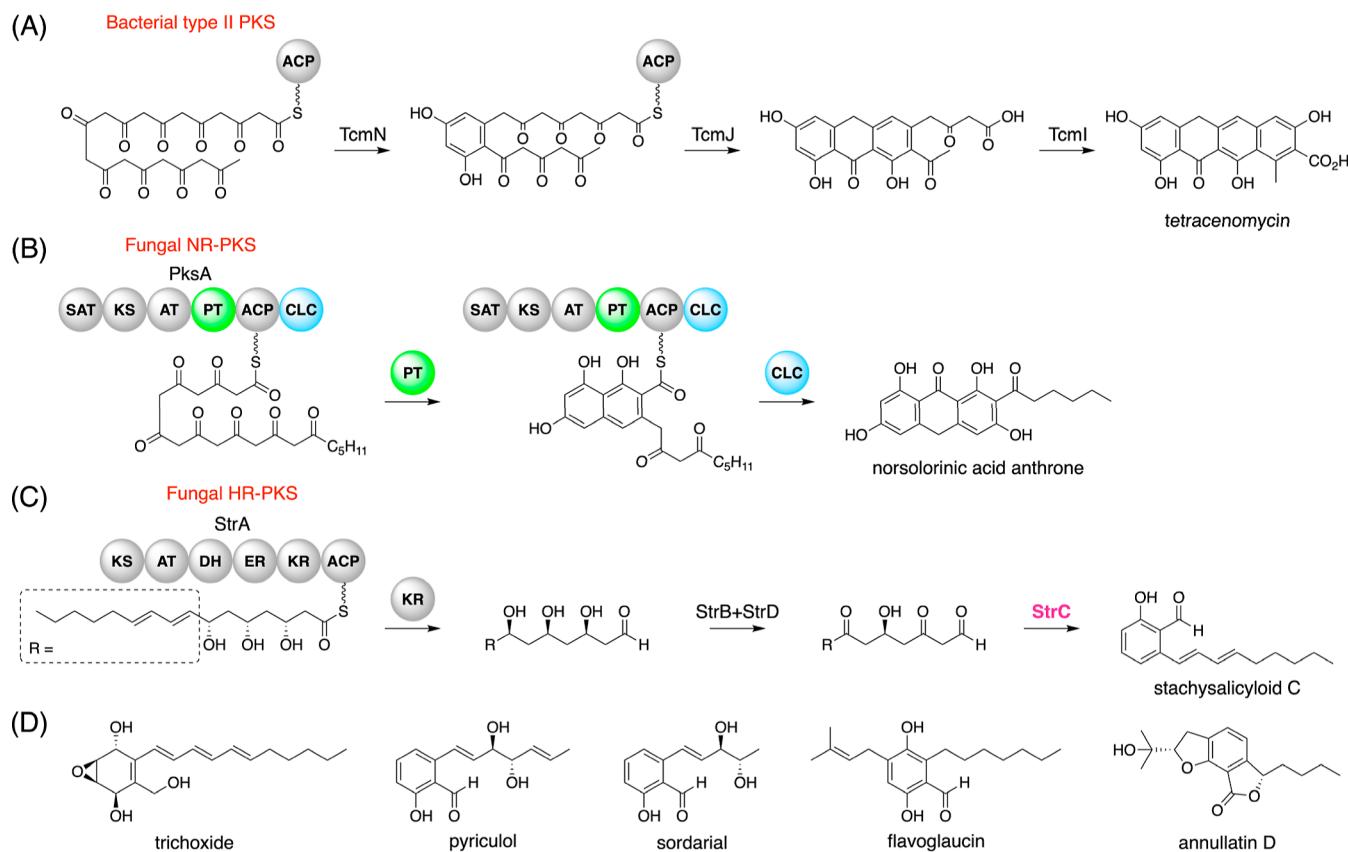


Figure 1. (A–C) Selected aromatization reactions in microbial polyketide biosynthesis. (A) Biosynthesis of tetracenomycin, which involves three cyclases/aromatases: TcmN, TcmJ, and TcmI. (B) Formation of norsolorinic acid anthrone by the NR-PKS PksA. (C) Biosynthesis of stachysalicyloid C. (D) Representative fungal alkyl salicylaldehyde derivatives.

In the past few years, several highly reducing PKSs (HR-PKSs) have been identified as a core enzyme for the biosynthesis of fungal alkyl salicylaldehyde derivatives, such as pyriculol,¹⁴ sordarial,¹⁵ trichoxide,¹⁶ flavoglaucin,¹⁷ annulatin D,¹⁸ and stachysalicyloids¹⁹ (Figure 1C,D). Currently, the mechanism of the salicylaldehyde scaffold of these molecules is generally accepted as follows. First, HR-PKS yields a linear polyketide aldehyde with a triol unit, after which two short-chain dehydrogenases/reductases (SDRs) perform two rounds of alcohol dehydrogenation. The resultant diketo aldehyde then undergoes an aldol condensation and subsequent dehydration to afford the salicylaldehyde core. Intriguingly, our recent work indicated that the ketoreductase (KR) domain of StrA, the HR-PKS for stachysalicyloid biosynthesis, is involved in the reductive release of the polyketide chain.¹⁹ Furthermore, we revealed that the cupin domain-containing enzyme StrC serves as an aromatase, which is critical for the efficient production of alkyl salicylaldehyde.¹⁹ Nevertheless, the details of the aromatization process, including the actual catalytic role of StrC, have yet to be clarified. Therefore, deciphering the enigma of aromatization by StrC should facilitate our understanding and potential applications of polyketide biosynthesis.

In this study, we carried out an in-depth investigation of the catalytic function of StrC by a multidisciplinary approach including in vitro enzymatic assays, X-ray crystallographic analysis, and computational studies. Our study identified the nonenzymatic and enzymatic stages of the aromatization process and elucidated how StrC engages in the formation of the salicylaldehyde core.

■ RESULTS

Investigation of the StrC-Catalyzed Step in the

Aromatization Process. We previously found that the linear aldehyde **1** is readily transformed into the alkyl salicylaldehyde stachysalicyloid C (**5**) in the presence of both the SDR StrD and the cupin domain-containing protein StrC.¹⁹ Meanwhile, the production of **5** was significantly hampered in the absence of StrC. Instead, several probable precursors of **5**, which were believed to be compounds **2–4** (Figure 2A), have been detected by liquid chromatography–mass spectrometry (LC–MS) analysis (Figure S1).¹⁸ These observations indicated that StrC was involved in the aromatization process to yield **5**. However, it remained unclear in which reaction step(s) of the aromatization process StrC plays a catalytic role. Accordingly, we sought to identify the reaction catalyzed by StrC as well as nonenzymatic transformations, if any. To this end, we initially reacted **1** with StrD for 15 min and extracted the reaction products, which were then incubated with StrC for another 15 min after the removal of StrD. As a result, the production of the aromatized product **5** was clearly observed without detectable levels of any reaction intermediates. On the contrary, a reaction intermediate, which is predicted to be cyclohexenone **4**, was predominantly observed in the absence of StrC (Figure 2B, traces i and ii). This observation suggested that the substrate of StrC is **4** rather than **2** or **3**. To further support this hypothesis, **1** was reacted with StrD for extended periods, and the reaction products were analyzed at various time intervals (Figure S2). After the 30 min reaction, LC–MS analysis detected only the major product, expected to be **4**, with apparently no traces of **2** or **3** (Figure S2,

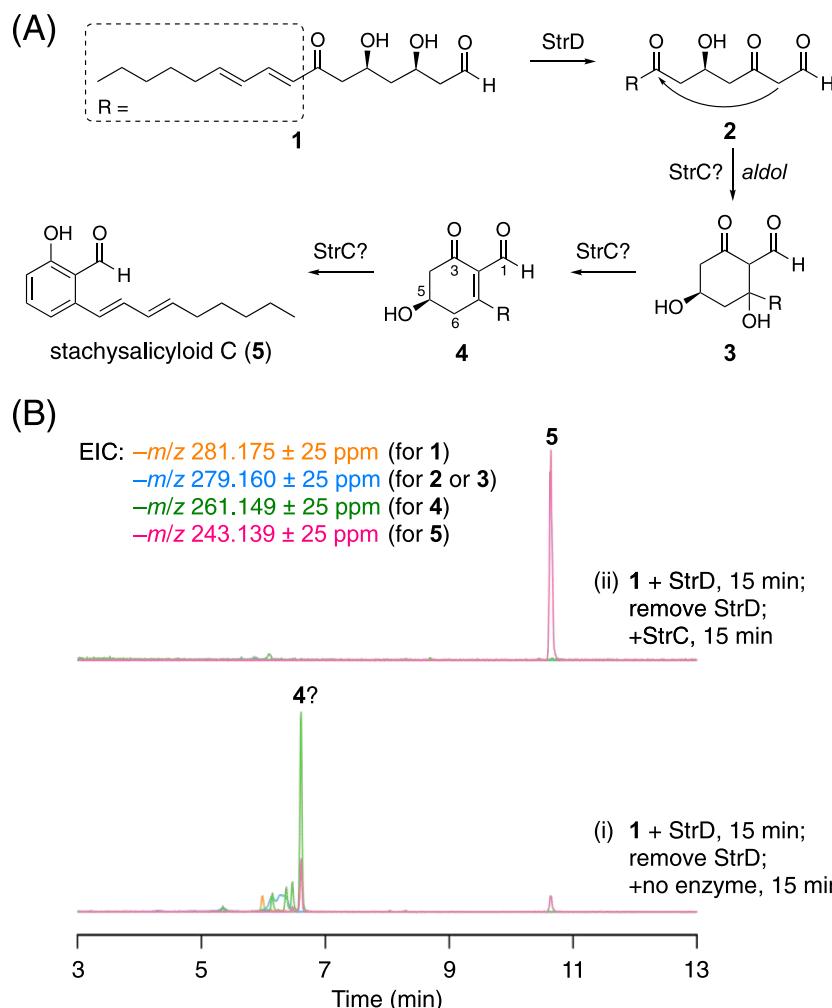


Figure 2. (A) Proposed biosynthetic pathway of the alkyl salicylaldehyde **5**. (B) LC–MS profiles of the products from the enzymatic reactions of StrD and StrC.

trace i). Furthermore, no conversion of **4** to **2** or **3** was observed over time (Figure S2, traces i–iv). Nevertheless, **5** was successfully generated when StrC was added (Figure S2, trace v). Collectively, these results suggest that StrC accepts **4** as its substrate to perform the dehydration reaction and that consumption of **4** drives the conversion of **2** or **3** to **4** by shifting the equilibrium toward the product.

Structural Characterization of StrC and MD Simulation. Next, we turned our attention to the structural basis of the StrC-catalyzed aromatization reaction. Pleasingly, we successfully obtained the X-ray crystal structures of StrC in its apo form and in complex with a substrate analogue **6**, which was obtained in our previous study¹⁹ and does not readily undergo dehydration even in the presence of StrC. The crystallographic analysis revealed that StrC forms a dodecamer, essentially consisting of two hexamers, as supported by size-exclusion chromatography (SEC) analysis (Figure S3). Each hexamer further consists of three dimer pairs. The structural integrity of StrC is maintained in various domains, especially its core cupin structure, which is characterized by a jelly roll pattern of 11 β -strands. These strands assemble into two antiparallel β sheets, with the N-terminal segment of StrC, formed by β 1– β 2, interacting with another cupin domain, resulting in the formation of a dimer. The β 3– β 10 strands give rise to the

characteristic cupin fold with a presumable substrate-binding pocket (Figure 3A,B).

As some cupin domain-containing enzymes are known to be metal-dependent,²⁰ our previous study identified potential metal-binding residues as His117, His157, and Asp123.¹⁹ However, the crystal structures of StrC revealed the absence of a metal ion in the vicinity of these residues. Instead, a water molecule was found alongside these residues. In the StrC-**6** complex, Asp123, Gln152, and His157 form hydrogen bonds with the C-5 hydroxy group of **6** directly or via the aforementioned water molecule (Figure 3D). In addition, Arg104 interacts with the C-3 carbonyl group. Meanwhile, His157 and Trp159 are located adjacent to the C-4 position (Figure 3D), from which deprotonation occurs during the dehydration process. It should be noted that these residues that presumably interact with the substrate are conserved among StrC and its known homologues, except that, in AnuC,¹⁸ the residue corresponding to Gln152 was found as asparagine (Figure S4).

Next, we performed molecular dynamics (MD) simulations using the crystal structure of StrC, in which **6** was replaced with the genuine substrate **4**, to investigate the important role of the amino acid residues identified above (Figures S5–S7). Our MD simulations suggested that Arg104 formed hydrogen bonding with the C-1 and C-3 carbonyl oxygens with a relative frequency

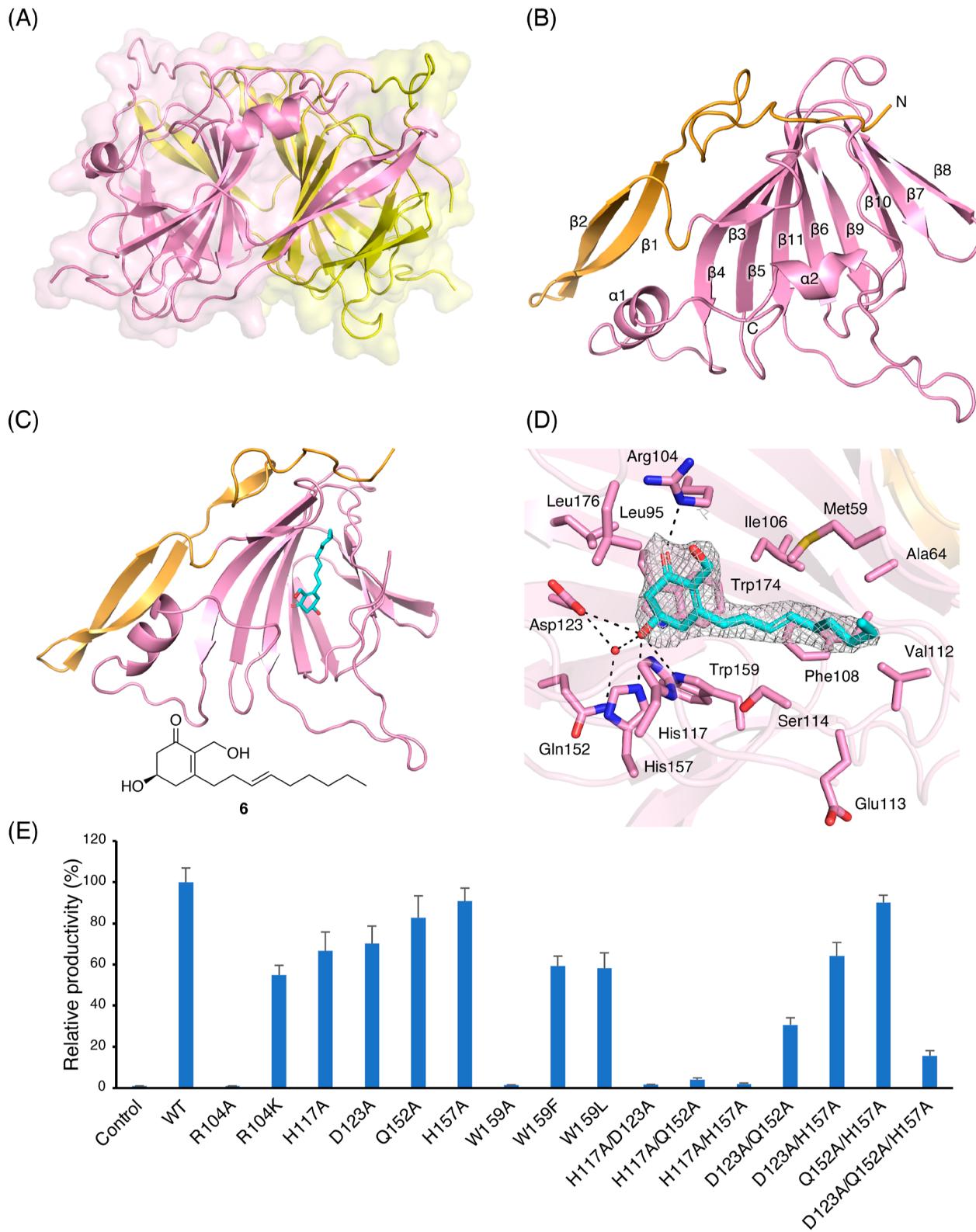


Figure 3. (A) Overall crystal structure of StrC. One of the dimers in the asymmetric unit is shown. (B) Secondary structure representations of StrC. The dimerization and cupin core domains are shown in orange and pink, respectively. (C) Crystal structure of StrC complexed with **6**. One of the monomers in the asymmetric unit is shown. (D) Key interactions of **6** with the active-site residues of StrC. The electron density map for **6** (Fo-Fc polder omit map shown in gray mesh) is contoured at 4σ . (E) Relative productivity of **5** from each StrC variant (also see Table S4; Control: reaction without StrC or its variants; WT: reaction with the wild-type StrC).

of approximately 106 and 124%, respectively. Concurrently, interactions involving Asp123 and Gln152 with the C-5 hydroxy group were observed at a frequency of around 95 and 64%,

respectively. Thus, these residues might play critical roles in the aromatization process.

Mutational Experiments on StrC. To explore a more mechanistic understanding of the StrC-catalyzed reaction, we conducted mutational experiments targeting the possible catalytic residues of StrC. Single-site mutation experiments were first performed by replacing the six active-site residues, namely, Arg104, His117, Asp123, His157, Gln152, and Trp159, with alanine. The activities of the resultant StrC variants were evaluated based on the production level of **5** in the reaction with StrD and **1** (Figures 3E, S8, and S9). The R104A and W159A variants almost completely abolished the formation of compound **5**, whereas the other four mutations did not significantly impact StrC activity. Moreover, when Arg104 was mutated to lysine, the resultant variant was still active with somewhat reduced activity, supporting the critical role of Arg104 in forming hydrogen bonds with the substrate carbonyls. In addition, Trp159 was also substituted with either phenylalanine or leucine. Both the W159F and W159L variants exhibited a similar productivity to **5**, implying that the bulkiness, rather than the aromaticity, of the amino acid residue at this position is important for the enzyme activity.

Individual mutations at His117, Asp123, His157, and Gln152 did not crucially affect the StrC activity. However, given the proximity of these residues to the eliminating hydroxy group and the deprotonation site of the substrate, these residues were nevertheless expected to play some roles in the aromatization process. Thus, we introduced multiple-site mutations to these four residues and evaluated their activities (Figures 3E, S8, and S9). Among the six double variants, those with the H117A mutation (i.e., H117A/D123A, H117A/Q152A, and H117A/H157A) showed a dramatically reduced productivity of **5**. In contrast, the other three, namely, the D123A/Q152A, D123A/H157A, and Q152A/H157A variants, retained the activity, although somewhat reduced productivity was observed from those with the D123A mutation. The productivity of **5** in the D123A/Q152A variant was less than half that in the D123A/H157A variant. These results suggest that Gln152 could work together with Asp123 or that Gln152 could compensate for the function of Asp123 upon its inactivation. Interestingly, the D123A/Q152A/H157A triple variant remained active and produced ~15% **5** compared with the wild-type enzyme. Collectively, it appears that His117, Asp123, and Gln152 collaboratively work to achieve the aromatization, with His117 being the major contributor to the process.

Computational Study of the Aromatization Steps. To gain more insights into the reaction mechanism of the aromatization process, a systematic computational study [CPCM DLPNO-CCSD(T)/ma-def2-TZVP(-f)//SMD M06-2X/6-31G* method as the main QM methods and SMD M06-2X/6-311+G**//SMD M06-2X/6-31G* method] was conducted by using a slightly modified substrate (**m2**) in aqueous solution and using the genuine substrate **4** inside the active site of StrC (Figures 4, 5, and S19–S22; see the Supporting Information for the detailed computational methods and results).²¹ Prior to aromatization, an aldol condensation reaction should occur as an initial step. Unfortunately, unstable intermediates **2–4** pose hurdles in our experimental investigation on the feasibility of the aldol reaction, which, however, can be resolved computationally. Our computational results show that **m2** first undergoes tautomerization to generate neutral and reactive enol isomers [**m2-enol** and **m2-enol'**; $\Delta E + ZPE_{soln} = \sim 1.3\text{--}4.0$ kcal/mol (CPCM DLPNO-CCSD(T)//SMD M06-2X method); Figures 4A and S10]. Then, an aldol reaction from **m2-enol'** takes place to form a cyclic intermediate

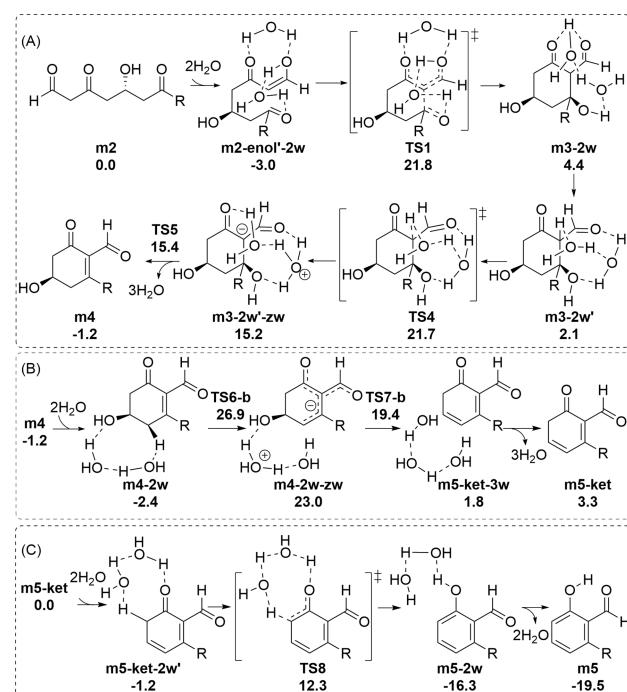


Figure 4. Possible mechanisms of (A) uncatalyzed aldol condensation and the (B,C) two subsequent aromatization processes in aqueous solution (without StrC protein) with the computed relative electronic energy with zero-point energy correction (in kcal/mol) by the CPCM DLPNO-CCSD(T)/ma-def2-TZVP(-f)//SMD M06-2X/6-31G* method [where R stands for (CH)₃]. Another high-energy water elimination pathway from **m4** is given in Figure S18C.

m3–2w in the presence of two explicit water molecules, in which the new C–C and O–H bonds are formed via **TS1**. The reaction barrier for this process in an aqueous solution is approximately 24.8 kcal/mol above **m2-enol'-2w** [CPCM DLPNO-CCSD(T)//SMD M06-2X method], which allows this step to occur in a solution without the need for an enzyme. Similarly, the related anionic enolate or neutral enamine²² form of **m2** requires a lower reaction barrier (Figures S11–S12). A cyclic intermediate **m3–2w'** can proceed with the subsequent stepwise dehydration to afford **m4** in the presence of two explicit water molecules, which act as a proton relay. The reaction barriers via the two eight-membered-ring transition states **TS4** and **TS5** are about 24.7 and 18.4 kcal/mol, respectively.

However, two subsequent water-assisted water elimination pathways from **m4** (or **m4–2w**) have to overcome much higher barriers in an aqueous solution (at least 29.3 kcal/mol above **m4–2w**; Figures 4B and S18C). Moreover, another tautomerization process from **m4** to give its enol isomer (**m4-enol**) was computed to be highly thermodynamically unstable (~11.5 kcal/mol; Figure S13). Therefore, these computational results support the uncatalyzed aldol condensation of **2** in an aqueous solution, but override the uncatalyzed aromatization of **4** in an aqueous solution, thus pinpointing the critical role of StrC in the aromatization process of **4**.

The mechanism of the aromatization process from **4** in StrC was further examined by using classical docking, followed by the active-site model at the CPCM DLPNO-CCSD(T)//SMD M06-2X level. As shown in Figure 5, **4** can form two different types of hydrogen bonds (F1 and F2) with its nearby Arg104, in which Arg104 can form hydrogen bonds with two or one carbonyl group of **4** in **IntA-w-F1** or **IntA-w-F2**, respectively.

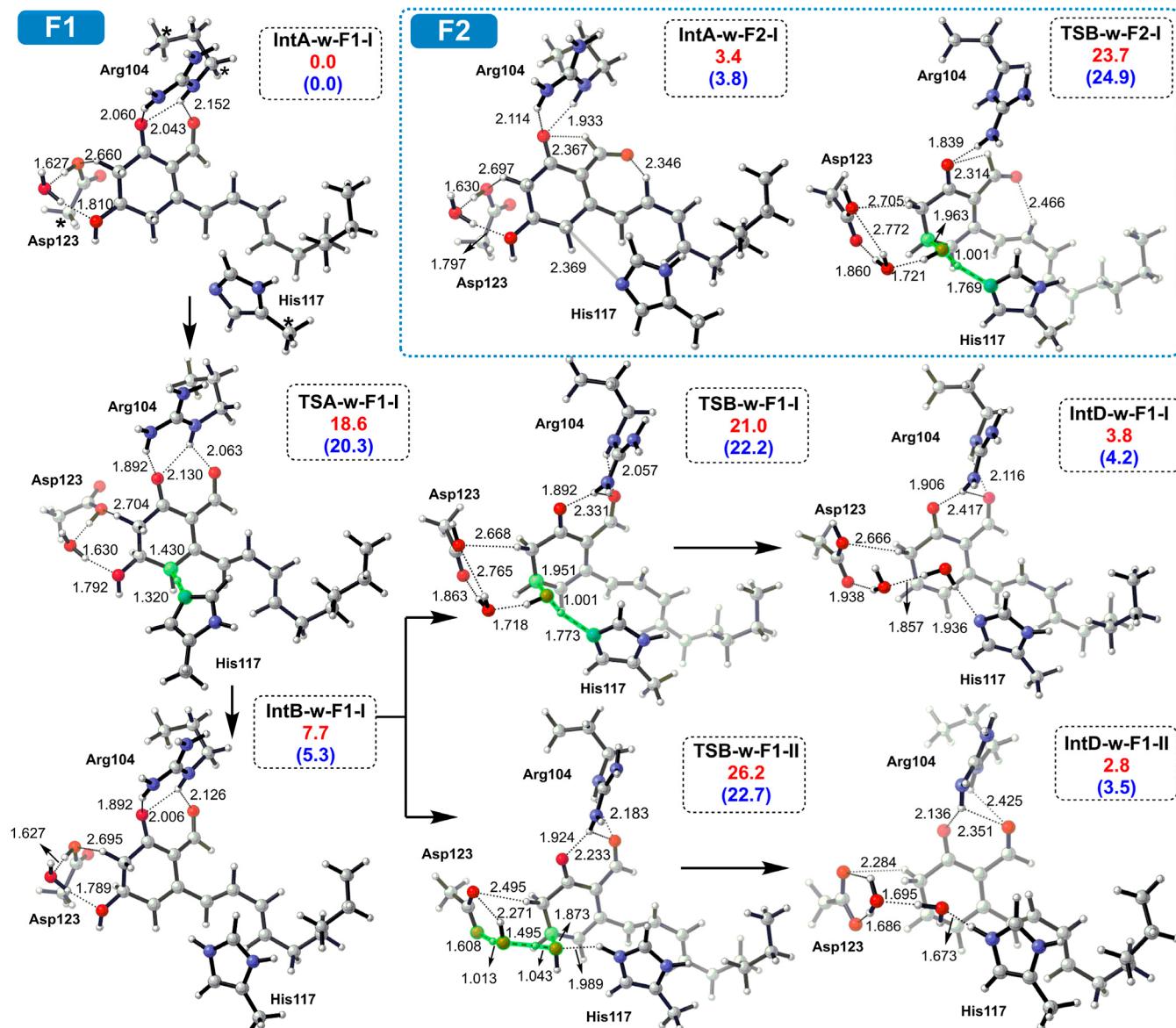


Figure 5. Computed relative electronic energy with zero-point energy correction (in kcal/mol) for the key catalyzed aromatization step from the key intermediate 4 in StrC by the CPCM DLPNO-CCSD(T)/ma-def2-TZVP(-f)//SMD M06-2X/6-31G* (in red) and SMD M06-2X/6-311+G**//SMD M06-2X/6-31G* (in blue) methods. The fixed atoms in all structures are marked with an asterisk (*) during the optimization process, as shown in **IntA-w-F1-I** as an example.

The two carbonyl groups of 4 cannot adopt a *cis* relationship in the latter case due to its lone-pair/lone-pair repulsion. In addition, our QM results reveal that, for these two hydrogen-bonding forms of 4, **IntA-w-F1** is more stable than **IntA-w-F2** by 3.4 kcal/mol, and the subsequent water elimination processes for the **F1** form generally have lower barriers (**Figures 5, S21, and S22**). Moreover, His117 can readily deprotonate from the C-6 position of 4 to give the anionic intermediate **IntB-w-F1** and protonated His117 (**Figure 5**). The reaction barrier for this deprotonation in StrC via **TSA-w-F1** is not high (approximately 18.6 kcal/mol).

Furthermore, our QM results suggest two possible succeeding proton-transfer-catalyzed dehydration pathways from **IntB-w-F1**. For instance, the cationic His117 can return the proton to catalyze the C–O bond cleavage (at the C-5 position) in concert with water elimination with a reaction barrier of 21.0 kcal/mol via **TSB-w-F1-I** above **IntA-w-F1**. Alternatively, neutral

Asp123, whose pK_a value was predicted to be 7.3 by the PROPKA program,²³ could also donate a proton to catalyze the same process via **TSB-w-F1-II** with a slightly higher reaction barrier of 26.2 kcal/mol at the CPCM DLPNO-CCSD(T)//SMD M06-2X level. Meanwhile, **TSB-w-F1-I** and **TSB-w-F1-II** have very similar energy barriers (22.2–22.7 kcal/mol) by the SMD M06-2X/6-311+G**//SMD M06-2X/6-31G* method. These computational results imply that His117 and Asp123 can serve as proton sources to trigger the key aromatization process.

In comparison, our DFT results show that the initial proton transfer from Arg104 to 5 to form a cationic intermediate is unlikely. Finally, intermediate **m5-ket** (or **m5-ket-2w'**) can readily undergo water-assisted tautomerization to form a much more thermodynamically stable phenol-form product **m5** in an aqueous solution with a low barrier of 13.5 kcal/mol above **m5-ket-2w'** (**Figures 4C and S18D**). Overall, our computational study supports the vital role of StrC in the aromatization process.

from the key intermediate **4** and the feasibility of uncatalyzed pathways for the other steps in the water solution.

■ DISCUSSION

Cupin superfamily proteins exhibit extensive prevalence and diverse functions across species from bacteria, fungi, and plants to humans,²⁴ encompassing roles as oxidoreductases,²⁵ dehydratases,²⁶ decarboxylases,²⁷ isomerases,²⁸ lyases,²⁹ and nitrogen–nitrogen bond-forming enzymes.³⁰ In this study, we have characterized cupin domain-containing enzyme StrC, which is the central enzyme for aromatization in stachysalicyloid biosynthesis, through structural biology and computational approaches. Thus, this work should add a new repertoire of cupin superfamily proteins. Our study has also revealed that the linear polyketide **2** undergoes a spontaneous aldol condensation to yield **4**, which is then accepted by StrC. In the active site of StrC, His117 serves as a catalytic base to deprotonate from the C-6 position of **4**, and the deprotonation is facilitated by the two hydrogen bonds between Arg104 and **4**, which increase the acidity of the C-6 position. Additionally, His117 or Asp123 protonates the C-5 hydroxy group to promote water elimination. The three catalytic residues of StrC identified herein are all conserved among known StrC homologues involved in fungal alkyl salicylaldehyde biosynthesis (Figure S4). Thus, these StrC homologues should adopt a similar mechanism to achieve aromatization.

The discovery and characterization of StrC mark a significant deviation from the conventional pathways involving fungal type I NR-PKSs and bacterial type II PKSs (Figure 1A,B). NR-PKSs employ a PT domain or a CLC domain integrated within themselves for cyclization and aromatization. In contrast, fungal alkyl salicylaldehyde biosynthesis adopts two SDRs and a cupin domain-containing enzyme, which are independent of the PKS, to aromatize the off-loaded linear PKS product. Accordingly, StrC and its homologues exhibit minimal sequence homology with PT domains despite their functional similarities. Furthermore, analysis of 92 cupin superfamily proteins from the UniProt database illuminates that StrC-like proteins are phylogenetically somewhat related to TcmJ-like cyclases associated with bacterial type II PKSs^{8c} (Table S5 and Figure S26). A primary distinction between these bacterial aromatases and StrC is their substrate recognition mechanisms; the former generally interacts with an ACP-bound substrate, whereas StrC acts on an offloaded and modified substrate. Notably, many of the cupin superfamily proteins analyzed herein possess 3-His-1-Glu (H-H-E-H) or 4-His metal-binding motifs. However, the second histidine residue is not conserved in StrC and its close homologues, and the aspartate is found in the replacement of the glutamate residue (Figures S26 and S27). Consistently, the *in vitro* enzymatic reactions of StrC and StrD treated with an excess of EDTA³¹ did not alter the activity of StrC nor the yield of the product **5**, indicating the metal-independent nature of StrC (Figure S28).

Moreover, the comparison of the crystal structures of StrC with TcmJ and RemF, metal-binding type II PKS aromatases for tetracenomycin³² and resistomycin³³ biosynthesis, respectively, reveals that StrC has a similar overall structure to these two enzymes. However, as mentioned above, the metal-binding motif is not completely conserved in StrC, eventually making StrC a metal-independent enzyme. Further analysis reveals that StrC structurally resembles a range of enzymes, including the cysteine dioxygenase type 1 (CDO1),³⁴ the oxalate oxidase 1 (OXO1),³⁵ the hydroxypropylphosphonic acid epoxidase

(HppE),³⁶ the acireductone dioxygenase (ARD),³⁷ the hydroxynitrile lyase GtHNL,³⁸ the dimethylsulfoniopropionate lyase DddK,³⁹ the phosphoglucose isomerase PfPGI,⁴⁰ and the dehydratase Pac13²⁶ (Figure S29). Notably, these enzymes are all metal-dependent except for Pac13, whose function is the most analogous to that of StrC. Collectively, StrC represents a new class of polyketide aromatases, which are evolutionarily distinct from known PKS domains, thus expanding the repertoire of core enzymatic reactions for fungal polyketide biosynthesis.

■ CONCLUSIONS

This study positions the cupin domain-containing enzyme StrC as a new class of polyketide aromatase for stachysalicyloid biosynthesis. Along with our recent characterization of HR-PKS StrA,¹⁹ a universal yet unusual biosynthetic mechanism for fungal alkyl salicylaldehydes can now be proposed, with (i) KR domain-catalyzed polyketide chain release and (ii) cupin enzyme-mediated aromatization being the key features. Our study thus expands our understanding of fungal polyketide biosynthesis and opens new avenues for further exploring overlooked biosynthetic reactions behind polyketide scaffold synthesis.

■ ASSOCIATED CONTENT

SI Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/acscatal.4c01043>.

Experimental details, analytical data, gene information, primers used and plasmids constructed data, X-ray data, LC–MS quantitative analysis results, genetic information used for evolutionary analysis, computational results, mass spectra, LC–MS profiles, protein molecular weight analysis, sequence alignment, MD simulation results, SDS-PAGE analysis, HPLC profiles, and detailed computational data (PDF)

8X0U data (CIF)

8X0V data (CIF)

■ AUTHOR INFORMATION

Corresponding Authors

Lung Wa Chung – Department of Chemistry, Shenzhen Grubbs Institute and Guangdong Provincial Key Laboratory of Catalysis, Southern University of Science and Technology, Shenzhen 518055, China;  orcid.org/0000-0001-9460-7812; Email: oscarchung@sustech.edu.cn

Yudai Matsuda – Department of Chemistry, City University of Hong Kong, Kowloon, Hong Kong SAR, China;  orcid.org/0000-0001-5650-4732; Email: ymatsuda@cityu.edu.hk

Wei-Guang Wang – Key Laboratory of Natural Products Synthetic Biology of Ethnic Medicinal Endophytes, State Ethnic Affairs Commission; Key Laboratory of Chemistry in Ethnic Medicinal Resources, Ministry of Education, Yunnan Minzu University, Kunming 650031 Yunnan, China;  orcid.org/0000-0003-3395-767X; Email: wwg@live.cn

Authors

Hang Wang – Key Laboratory of Natural Products Synthetic Biology of Ethnic Medicinal Endophytes, State Ethnic Affairs Commission; Key Laboratory of Chemistry in Ethnic Medicinal Resources, Ministry of Education, Yunnan Minzu University, Kunming 650031 Yunnan, China; School of Life

Science and Technology, China Pharmaceutical University, Nanjing 210009, China

Chao Peng – Key Laboratory of Natural Products Synthetic Biology of Ethnic Medicinal Endophytes, State Ethnic Affairs Commission; Key Laboratory of Chemistry in Ethnic Medicinal Resources, Ministry of Education, Yunnan Minzu University, Kunming 650031 Yunnan, China

Xiao-Xuan Chen – Department of Chemistry, Shenzhen Grubbs Institute and Guangdong Provincial Key Laboratory of Catalysis, Southern University of Science and Technology, Shenzhen 518055, China

Hao-Yang Wang – Department of Chemistry, Shenzhen Grubbs Institute and Guangdong Provincial Key Laboratory of Catalysis, Southern University of Science and Technology, Shenzhen 518055, China

Run Yang – Key Laboratory of Natural Products Synthetic Biology of Ethnic Medicinal Endophytes, State Ethnic Affairs Commission; Key Laboratory of Chemistry in Ethnic Medicinal Resources, Ministry of Education, Yunnan Minzu University, Kunming 650031 Yunnan, China

Hao Xiang – Key Laboratory of Natural Products Synthetic Biology of Ethnic Medicinal Endophytes, State Ethnic Affairs Commission; Key Laboratory of Chemistry in Ethnic Medicinal Resources, Ministry of Education, Yunnan Minzu University, Kunming 650031 Yunnan, China;  orcid.org/0000-0002-0299-1069

Qiu-Fen Hu – Key Laboratory of Natural Products Synthetic Biology of Ethnic Medicinal Endophytes, State Ethnic Affairs Commission; Key Laboratory of Chemistry in Ethnic Medicinal Resources, Ministry of Education, Yunnan Minzu University, Kunming 650031 Yunnan, China

Ling Liu – State Key Laboratory of Mycology, Institute of Microbiology, Chinese Academy of Sciences, Beijing 100101, China;  orcid.org/0000-0002-4882-0396

Complete contact information is available at:

<https://pubs.acs.org/10.1021/acscatal.4c01043>

Author Contributions

#H.W., C.P., and X.-X.C. contributed equally.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

We are grateful to Prof. Katsuya Gomi (Tohoku University) and Prof. Jun-ichi Maruyama (The University of Tokyo) for kindly providing the expression vectors and the fungal strain. We thank Dr. Xudong Qu (Shanghai Jiao Tong University) for his valuable suggestions on the experimental design. We are grateful to the staff at beamlines BL17U1 and BL19U1 of the Shanghai Synchrotron Radiation Facility for their assistance during data collection. This work was supported by grants from the National Natural Science Foundation of China (project nos. 22377105, 21907083, 21933003, 22193020, 22193023, 32022002, and 21977113), grants from the National Natural Science Foundation of Yunnan Province and Jiangsu Province (nos. 202101AS070022, 202401AV070003, and BK20210434), the Yunnan Ten-thousand Talents Program to W.-G.W., the General Research Fund grant (Project no. 11301321) from the Research Grants Council (RGC) of Hong Kong, the Shenzhen Nobel Prize Scientists Laboratory Project (C17783101), and the Guangdong Provincial Key Laboratory of Catalytic Chemistry (2020B121201002). We thank the

Center for Computational Science and Engineering at the Southern University of Science and Technology and CHEM HPC at SUSTech for partly supporting this work.

REFERENCES

- (a) Larsen, E. M.; Wilson, M. R.; Taylor, R. E. Conformation–activity relationships of polyketide natural products. *Nat. Prod. Rep.* **2015**, *32* (8), 1183–1206. (b) Pham, J. V.; Yilmaz, M. A.; Feliz, A.; Majid, M. T.; Maffetone, N.; Walker, J. R.; Kim, E.; Cho, H. J.; Reynolds, J. M.; Song, M. C.; Park, S. R.; Yoon, Y. J. A Review of the Microbial Production of Bioactive Natural Products and Biologics. *Front. Microbiol.* **2019**, *10*, 1404.
- (2) Cortes, J.; Haydock, S. F.; Roberts, G. A.; Bevitt, D. J.; Leadlay, P. F. An unusually large multifunctional polypeptide in the erythromycin-producing polyketide synthase of *Saccharopolyspora erythraea*. *Nature* **1990**, *348* (6297), 176–178.
- (3) Caffrey, P.; Lynch, S.; Flood, E.; Finn, S.; Oliynyk, M. Amphotericin biosynthesis in *Streptomyces nodosus*: deductions from analysis of polyketide synthase and late genes. *Chem. Biol.* **2001**, *8* (7), 713–723.
- (4) Motamedi, H.; Shafiee, A. The biosynthetic gene cluster for the macrolactone ring of the immunosuppressant FK506. *Eur. J. Biochem.* **1998**, *256* (3), 528–534.
- (5) Schwecke, T.; Aparicio, J. F.; Molnár, I.; König, A.; Khaw, L. E.; Haydock, S. F.; Oliynyk, M.; Caffrey, P.; Cortés, J.; Lester, J. B. The biosynthetic gene cluster for the polyketide immunosuppressant rapamycin. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92* (17), 7839–7843.
- (6) Kennedy, J.; Auclair, K.; Kendrew, S. G.; Park, C.; Vedera, J. C.; Richard Hutchinson, C. Modulation of polyketide synthase activity by accessory proteins during lovastatin biosynthesis. *Science* **1999**, *284* (5418), 1368–1372.
- (7) (a) Fischbach, M. A.; Walsh, C. T. Assembly-line enzymology for polyketide and nonribosomal peptide antibiotics: Logic, machinery, and mechanisms. *Chem. Rev.* **2006**, *106* (8), 3468–3496. (b) Herbst, D. A.; Townsend, C. A.; Maier, T. The architectures of iterative type I PKS and FAS. *Nat. Prod. Rep.* **2018**, *35* (10), 1046–1069. (c) Nivina, A.; Yuet, K. P.; Hsu, J.; Khosla, C. Evolution and diversity of assembly-line polyketide synthases. *Chem. Rev.* **2019**, *119* (24), 12524–12547. (d) Wang, J.; Deng, Z.; Liang, J.; Wang, Z. Structural enzymology of iterative type I polyketide synthases: Various routes to catalytic programming. *Nat. Prod. Rep.* **2023**, *40* (9), 1498–1520.
- (8) (a) Hertweck, C.; Luzhetsky, A.; Rebets, Y.; Bechthold, A. Type II polyketide synthases: gaining a deeper insight into enzymatic teamwork. *Nat. Prod. Rep.* **2007**, *24* (1), 162–190. (b) Wang, J.; Zhang, R.; Chen, X.; Sun, X.; Yan, Y.; Shen, X.; Yuan, Q. Biosynthesis of aromatic polyketides in microorganisms using type II polyketide synthases. *Microb. Cell Fact.* **2020**, *19* (1), 110. (c) Xie, S.; Zhang, L. Type II polyketide synthases: A bioinformatics-driven approach. *ChemBioChem* **2023**, *24*, No. e202200775.
- (9) (a) Abe, I.; Morita, H. Structure and function of the chalcone synthase superfamily of plant type III polyketide synthases. *Nat. Prod. Rep.* **2010**, *27* (6), 809–838. (b) Morita, H.; Wong, C. P.; Abe, I. How structural subtleties lead to molecular diversity for the type III polyketide synthases. *J. Biol. Chem.* **2019**, *294* (41), 15121–15136.
- (10) (a) Bao, W.; Wendt-Pienkowski, E.; Hutchinson, C. R. Reconstitution of the iterative type II polyketide synthase for tetracenomycin F2 biosynthesis. *Biochemistry* **1998**, *37* (22), 8132–8138. (b) Thompson, T. B.; Katayama, K.; Watanabe, K.; Hutchinson, C. R.; Rayment, I. Structural and functional analysis of tetracenomycin F2 cyclase from *Streptomyces glaucescens*: A TYPE II POLYKETIDE CYCLASE. *J. Biol. Chem.* **2004**, *279* (36), 37956–37963. (c) Ames, B. D.; Korman, T. P.; Zhang, W.; Smith, P.; Vu, T.; Tang, Y.; Tsai, S.-C. Crystal structure and functional analysis of tetracenomycin ARO/CYC: Implications for cyclization specificity of aromatic polyketides. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105* (14), 5349–5354.
- (11) (a) Crawford, J. M.; Korman, T. P.; Labonte, J. W.; Vagstad, A. L.; Hill, E. A.; Kamari-Bidkorpeh, O.; Tsai, S.-C.; Townsend, C. A. Structural basis for biosynthetic programming of fungal aromatic

- polyketide cyclization. *Nature* **2009**, *461* (7267), 1139–1143.
- (b) Barajas, J. F.; Shakya, G.; Moreno, G.; Rivera, H.; Jackson, D. R.; Topper, C. L.; Vagstad, A. L.; La Clair, J. J.; Townsend, C. A.; Burkart, M. D.; Tsai, S.-C. Polyketide mimetics yield structural and mechanistic insights into product template domain function in nonreducing polyketide synthases. *Proc. Natl. Acad. Sci. U.S.A.* **2017**, *114* (21), E4142–E4148.
- (12) Korman, T. P.; Crawford, J. M.; Labonte, J. W.; Newman, A. G.; Wong, J.; Townsend, C. A.; Tsai, S.-C. Structure and function of an iterative polyketide synthase thioesterase domain catalyzing Claisen cyclization in aflatoxin biosynthesis. *Proc. Natl. Acad. Sci. U.S.A.* **2010**, *107* (14), 6246–6251.
- (13) (a) Moriguchi, T.; Kezuka, Y.; Nonaka, T.; Ebizuka, Y.; Fujii, I. Hidden function of catalytic domain in 6-methylsalicylic acid synthase for product release. *J. Biol. Chem.* **2010**, *285* (20), 15637–15643. (b) Sun, H.; Ho, C. L.; Ding, F.; Soehano, I.; Liu, X.-W.; Liang, Z.-X. Synthesis of (R)-mellein by a partially reducing iterative polyketide synthase. *J. Am. Chem. Soc.* **2012**, *134* (29), 11924–11927.
- (14) Jacob, S.; Grötsch, T.; Foster, A. J.; Schüffler, A.; Rieger, P. H.; Sandjo, L. P.; Liermann, J. C.; Opitz, T.; Thines, E. Unravelling the biosynthesis of pyriculol in the rice blast fungus Magnaporthe oryzae. *Microbiology* **2017**, *163* (4), 541–553.
- (15) Zhao, Z.; Ying, Y.; Hung, Y.-S.; Tang, Y. Genome mining reveals *neurospora crassa* can produce the salicylaldehyde sordarial. *J. Nat. Prod.* **2019**, *82* (4), 1029–1033.
- (16) Liu, L.; Tang, M.-C.; Tang, Y. Fungal highly reducing polyketide synthases biosynthesize salicyldehydes that are precursors to epoxycyclohexenol natural products. *J. Am. Chem. Soc.* **2019**, *141* (50), 19538–19541.
- (17) Nies, J.; Ran, H.; Wohlgemuth, V.; Yin, W.-B.; Li, S.-M. Biosynthesis of the prenylated salicylaldehyde flavoglaucin requires temporary reduction to salicyl alcohol for decoration before reoxidation to the final product. *Org. Lett.* **2020**, *22* (6), 2256–2260.
- (18) Xiang, P.; Kemmerich, B.; Yang, L.; Li, S.-M. Biosynthesis of annullatin D in *Penicillium roqueforti* implies oxidative lactonization between two hydroxyl groups catalyzed by a BBE-like enzyme. *Org. Lett.* **2022**, *24* (32), 6072–6077.
- (19) Yang, R.; Feng, J.; Xiang, H.; Cheng, B.; Shao, L.-D.; Li, Y.-P.; Wang, H.; Hu, Q.-F.; Xiao, W.-L.; Matsuda, Y.; Wang, W.-G. Ketoreductase domain-catalyzed polyketide chain release in fungal alkyl salicylaldehyde biosynthesis. *J. Am. Chem. Soc.* **2023**, *145* (20), 11293–11300.
- (20) Dunwell, J. M.; Purvis, A.; Khuri, S. Cupins: The most functionally diverse protein superfamily? *Phytochemistry* **2004**, *65* (1), 7–17.
- (21) (a) Sheng, X.; Himo, F. The quantum chemical cluster approach in biocatalysis. *Acc. Chem. Res.* **2023**, *56* (8), 938–947. (b) Tantillo, D. J. Interrogating chemical mechanisms in natural products biosynthesis using quantum chemical calculations. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2020**, *10* (3), No. e1453. (c) Zhang, B.; Wang, K. B.; Wang, W.; Wang, X.; Liu, F.; Zhu, J.; Shi, J.; Li, L. Y.; Han, H.; Xu, K.; Qiao, H. Y.; Zhang, X.; Jiao, R. H.; Houk, K. N.; Liang, Y.; Tan, R. X.; Ge, H. M. Enzyme-catalysed [6 + 4] cycloadditions in the biosynthesis of natural products. *Nature* **2019**, *568* (7750), 122–126. (d) Ye, Y.; Du, L.; Zhang, X.; Newmister, S. A.; McCauley, M.; Alegre-Requena, J. V.; Zhang, W.; Mu, S.; Minami, A.; Fraley, A. E.; Adrover-Castellano, M. L.; Carney, N. A.; Shende, V. V.; Qi, F.; Oikawa, H.; Kato, H.; Tsukamoto, S.; Paton, R. S.; Williams, R. M.; Sherman, D. H.; Li, S. Fungal-derived brevianamide assembly by a stereoselective semipinacolase. *Nat. Catal.* **2020**, *3* (6), 497–506. (e) Cheng, L.; Li, D.; Mai, B. K.; Bo, Z.; Cheng, L.; Liu, P.; Yang, Y. Stereoselective amino acid synthesis by synergistic photoredox-pyridoxal radical biocatalysis. *Science* **2023**, *381* (6656), 444–451. (f) Lan, J.; Li, X.; Yang, Y.; Zhang, X.; Chung, L. W. New insights and predictions into complex homogeneous reactions enabled by computational chemistry in synergy with experiments: Isotopes and mechanisms. *Acc. Chem. Res.* **2022**, *55* (8), 1109–1123.
- (22) (a) Hajos, Z. G.; Parrish, D. R. Asymmetric synthesis of bicyclic intermediates of natural product chemistry. *J. Org. Chem.* **1974**, *39* (12), 1615–1621. (b) Eder, U.; Sauer, G.; Wiechert, R. New type of asymmetric cyclization to optically active steroid CD partial structures. *Angew. Chem., Int. Ed.* **1971**, *10* (7), 496–497. (c) Clemente, F. R.; Houk, K. N. Theoretical studies of stereoselectivities of intramolecular aldol cyclizations catalyzed by amino acids. *J. Am. Chem. Soc.* **2005**, *127* (32), 11294–11302.
- (23) (a) Søndergaard, C. R.; Olsson, M. H. M.; Rostkowski, M.; Jensen, J. H. Improved treatment of ligands and coupling effects in empirical calculation and rationalization of pK_a values. *J. Chem. Theory Comput.* **2011**, *7* (7), 2284–2295. (b) Olsson, M. H. M.; Søndergaard, C. R.; Rostkowski, M.; Jensen, J. H. Propka3: Consistent treatment of internal and surface residues in empirical pK_a predictions. *J. Chem. Theory Comput.* **2011**, *7* (2), 525–537.
- (24) Dunwell, J. M.; Culham, A.; Carter, C. E.; Sosa-Aguirre, C. R.; Goodenough, P. W. Evolution of functional diversity in the cupin superfamily. *Trends Biochem. Sci.* **2001**, *26* (12), 740–746.
- (25) (a) Hu, Z.; Awakawa, T.; Ma, Z.; Abe, I. Aminoacyl sulfonamide assembly in SB-203208 biosynthesis. *Nat. Commun.* **2019**, *10* (1), 184. (b) Simmons, C. R.; Liu, Q.; Huang, Q.; Hao, Q.; Begley, T. P.; Karplus, P. A.; Stipanuk, M. H. Crystal structure of mammalian cysteine dioxygenase: A NOVEL MONONUCLEAR IRON CENTER FOR CYSTEINE THIOL OXIDATION. *J. Biol. Chem.* **2006**, *281* (27), 18723–18733.
- (26) Michailidou, F.; Chung, C.-w.; Brown, M. J. B.; Bent, A. F.; Naismith, J. H.; Leavens, W. J.; Lynn, S. M.; Sharma, S. V.; Goss, R. J. M. Pac13 is a small, monomeric dehydratase that mediates the formation of the 3'-deoxy nucleoside of pacidamycins. *Angew. Chem., Int. Ed.* **2017**, *56* (41), 12492–12497.
- (27) Anand, R.; Dorrestein, P. C.; Kinsland, C.; Begley, T. P.; Ealick, S. E. Structure of oxalate decarboxylase from *Bacillus subtilis* at 1.75 Å resolution. *Biochemistry* **2002**, *41* (24), 7659–7669.
- (28) (a) Qiu, X.; Zhu, W.; Wang, W.; Jin, H.; Zhu, P.; Zhuang, R.; Yan, X. Structural and functional insights into the role of a cupin superfamily isomerase in the biosynthesis of Choi moiety of aeruginosin. *J. Struct. Biol.* **2019**, *205* (3), 44–52. (b) Vorobjeva, N. N.; Kurilova, S. A.; Petukhova, A. F.; Nazarov, T. I.; Kolomijtseva, G. Y.; Baykov, A. A.; Rodina, E. V. A novel, cupin-type phosphoglucose isomerase in *Escherichia coli*. *Biochim. Biophys. Acta Gen. Subj.* **2020**, *1864* (7), 129601.
- (29) Wang, S.-Y.; Zhang, N.; Teng, Z.-J.; Wang, X.-D.; Todd, J. D.; Zhang, Y.-Z.; Cao, H.-Y.; Li, C.-Y. A new dimethylsulfoniopropionate lyase of the cupin superfamily in marine bacteria. *Environ. Microbiol.* **2023**, *25* (7), 1238–1249.
- (30) (a) Ng, T. L.; Rohac, R.; Mitchell, A. J.; Boal, A. K.; Balskus, E. P. An N-nitrosating metalloenzyme constructs the pharmacophore of streptozotocin. *Nature* **2019**, *566* (7742), 94–99. (b) Zhao, G.; Peng, W.; Song, K.; Shi, J.; Lu, X.; Wang, B.; Du, Y.-L. Molecular basis of enzymatic nitrogen–nitrogen formation by a family of zinc-binding cupin enzymes. *Nat. Commun.* **2021**, *12* (1), 7205.
- (31) (a) Li, C.-Y.; Zhang, D.; Chen, X.-L.; Wang, P.; Shi, W.-L.; Li, P.-Y.; Zhang, X.-Y.; Qin, Q.-L.; Todd, J. D.; Zhang, Y.-Z. Mechanistic insights into dimethylsulfoniopropionate lyase DddY, a new member of the cupin superfamily. *J. Mol. Biol.* **2017**, *429* (24), 3850–3862. (b) Vasconcellos, V.; Tardioli, P.; Giordano, R.; Farinas, C. Addition of metal ions to a (hemi) cellulolytic enzymatic cocktail produced in-house improves its activity, thermostability, and efficiency in the saccharification of pretreated sugarcane bagasse. *New Biotech.* **2016**, *33* (3), 331–337.
- (32) Axelrod, H. L.; Kozbial, P.; McMullan, D.; Krishna, S. S.; Miller, M. D.; Abdubek, P.; Acosta, C.; Astakhova, T.; Carlton, D.; Caruthers, J.; Chiu, H.-J.; Clayton, T.; Deller, M. C.; Duan, L.; Elias, Y.; Feuerhelm, J.; Grzechnik, S. K.; Grant, J. C.; Han, G. W.; Jaroszewski, L.; Jin, K. K.; Klock, H. E.; Knuth, M. W.; Kumar, A.; Marciano, D.; Morse, A. T.; Murphy, K. D.; Nigoghossian, E.; Okach, L.; Oommachen, S.; Paulsen, J.; Reyes, R.; Rife, C. L.; Tien, H. J.; Trout, C. V.; van den Bedem, H.; Weeks, D.; White, A.; Xu, Q.; Zubietta, C.; Hodgson, K. O.; Wooley, J.; Elsliger, M.-A.; Deacon, A. M.; Godzik, A.; Lesley, S. A.; Wilson, I. A. Conformational changes associated with the binding of zinc acetate at the putative active site of XcTcmJ, a cupin from *Xanthomonas*

campestris pv. *Acta Crystallogr., Sect. F: Struct. Biol. Cryst. Commun.* **2010**, *66* (10), 1347–1353.

(33) Silvennoinen, L.; Sandalova, T.; Schneider, G. The polyketide cyclase RemF from *Streptomyces resistomycificus* contains an unusual octahedral zinc binding site. *FEBS Lett.* **2009**, *583* (17), 2917–2921.

(34) Levin, E. J.; Kondrashov, D. A.; Wesenberg, G. E.; Phillips, G. N. Ensemble refinement of protein crystal structures: Validation and application. *Structure* **2007**, *15* (9), 1040–1052.

(35) Opaleye, O.; Rose, R.-S.; Whittaker, M. M.; Woo, E.-J.; Whittaker, J. W.; Pickersgill, R. W. Structural and spectroscopic studies shed light on the mechanism of oxalate oxidase. *J. Biol. Chem.* **2006**, *281* (10), 6428–6433.

(36) Yun, D.; Dey, M.; Higgins, L. J.; Yan, F.; Liu, H.-w.; Drennan, C. L. Structural basis of regiospecificity of a mononuclear iron enzyme in antibiotic fosfomycin biosynthesis. *J. Am. Chem. Soc.* **2011**, *133* (29), 11262–11269.

(37) Ju, T.; Goldsmith, R. B.; Chai, S. C.; Maroney, M. J.; Pochapsky, S. S.; Pochapsky, T. C. One protein, two enzymes revisited: A structural entropy switch interconverts the two isoforms of acireductone dioxygenase. *J. Mol. Biol.* **2006**, *363* (4), 823–834.

(38) Hajnal, I.; Łyskowski, A.; Hanefeld, U.; Gruber, K.; Schwab, H.; Steiner, K. Biochemical and structural characterization of a novel bacterial manganese-dependent hydroxynitrile lyase. *FEBS J.* **2013**, *280* (22), 5815–5828.

(39) Schnicker, N. J.; De Silva, S. M.; Todd, J. D.; Dey, M. Structural and biochemical insights into dimethylsulfoniopropionate cleavage by cofactor-bound DddK from the prolific marine bacterium *Pelagibacter*. *Biochemistry* **2017**, *56* (23), 2873–2885.

(40) Berrisford, J. M.; Hounslow, A. M.; Akerboom, J.; Hagen, W. R.; Brouns, S. J. J.; van der Oost, J.; Murray, I. A.; Michael Blackburn, G.; Walther, J. P.; Rice, D. W.; Baker, P. J. Evidence supporting a cis-enediol-based mechanism for *Pyrococcus furiosus* phosphoglucose isomerase. *J. Mol. Biol.* **2006**, *358* (5), 1353–1366.