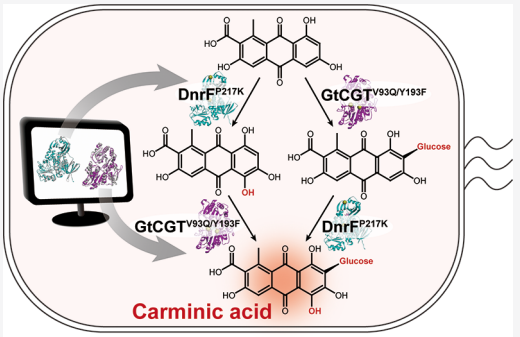
JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY
pubs.acs.org/JACS
Article

Production of Carminic Acid by Metabolically Engineered *Escherichia coli*

Dongsoo Yang, Woo Dae Jang, and Sang Yup Lee*

Cite This: *J. Am. Chem. Soc.* 2021, 143, 5364−5377
Read Online

ABSTRACT: Carminic acid is an aromatic polyketide found in scale insects (i.e., *Dactylopius coccus*) and is a widely used natural red colorant. It has long been produced by the cumbersome farming of insects followed by multistep purification processes. Thus, there has been much interest in producing carminic acid by the fermentation of engineered bacteria. Here we report the complete biosynthesis of carminic acid from glucose in engineered *Escherichia coli*. We first optimized the type II polyketide synthase machinery from *Photorhabdus luminescens*, enabling a high-level production of flavokermesic acid upon coexpression of the cyclases ZhuI and ZhuJ from *Streptomyces* sp. R1128. To discover the enzymes responsible for the remaining two reactions (hydroxylation and *C*-glucosylation), biochemical reaction analyses were performed by testing enzyme candidates reported to perform similar reactions. The two identified enzymes, aklavinone 12-hydroxylase (DnrF) from *Streptomyces peucetius* and *C*-glucosyltransferase (GtCGT) from *Gentiana triflora*, could successfully perform hydroxylation and *C*-glucosylation of flavokermesic acid, respectively. Then, homology modeling and docking simulations were performed to enhance the activities of these two enzymes, leading to the generation of beneficial mutants with 2−5-fold enhanced conversion efficiencies. In addition, the GtCGT mutant was found to be a generally applicable *C*-glucosyltransferase in *E. coli*, as was showcased by the successful production of aloesin found in *Aloe vera*. Simple metabolic engineering followed by fed-batch fermentation resulted in 0.63 ± 0.02 mg/L of carminic acid production from glucose. The strategies described here will be useful for the design and construction of biosynthetic pathways involving unknown enzymes and consequently the production of diverse industrially important natural products.

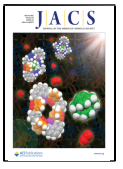


INTRODUCTION

Polyketides are an attractive class of natural products with diverse biological activities, which are used as food supplements, nutraceuticals, cosmetics, and drugs.[1−3] Polyketides are produced by polyketide synthases (PKSs) which are classified into three major types according to their modes of action and biosynthetic mechanisms. In contrast to macrolide polyketides produced by type I modular PKSs,[4] aromatic polyketides are produced by either type II or type III PKSs (Figure 1a,b).[5,6] A notable example of aromatic polyketides is carminic acid (**6**), which is a red colorant widely used for food, cosmetics, and pharmaceuticals. Naturally produced from scale insects including female cochineals (*Dactylopius coccus*), **6** confers the red hues of various food products including ketchup, strawberry milk, and candies or cosmetics including eyeshadow, nail polish, and lipstick.

Although the commercial production of **6** solely depends on extraction from the cochineals, an increase in the production scale is limited due to the slow growth of the cochineals and their limited cultivable region; cochineals live in symbiosis on prickly pear cacti, which can be only cultivated in hot and dry areas (mainly Peru or the Canary Islands). Also, the **6** extraction process from desiccated cochineals is inefficient, as it is estimated that 70000 female cochineals are required to obtain one pound of **6**.[7−10] Due to its finite supply despite the rapidly increasing demand of naturally colored food, the price of **6** per ton has risen by 73% from 2013 to 2018.[10] Regardless of these economic issues, many people (especially vegans or vegetarians) are uncomfortable with consuming insect-derived products. Moreover, although rare, it was reported that food products containing **6** provoked allergic reactions in some people,[11] possibly due to the insect protein contaminants.[11,12] This is thought to be due to the significant variations in the purity of **6** between different batches because of the nonstandardized extraction methods.[13] One the other hand, **6** itself has been generally recognized as safe. As a response toward such issues, a number of franchises have started to use alternative red colorants (e.g. lycopene and betanin) instead of

Received: November 27, 2020
Published: April 2, 2021



https://doi.org/10.1021/jacs.0c12406
J. Am. Chem. Soc. 2021, 143, 5364−5377

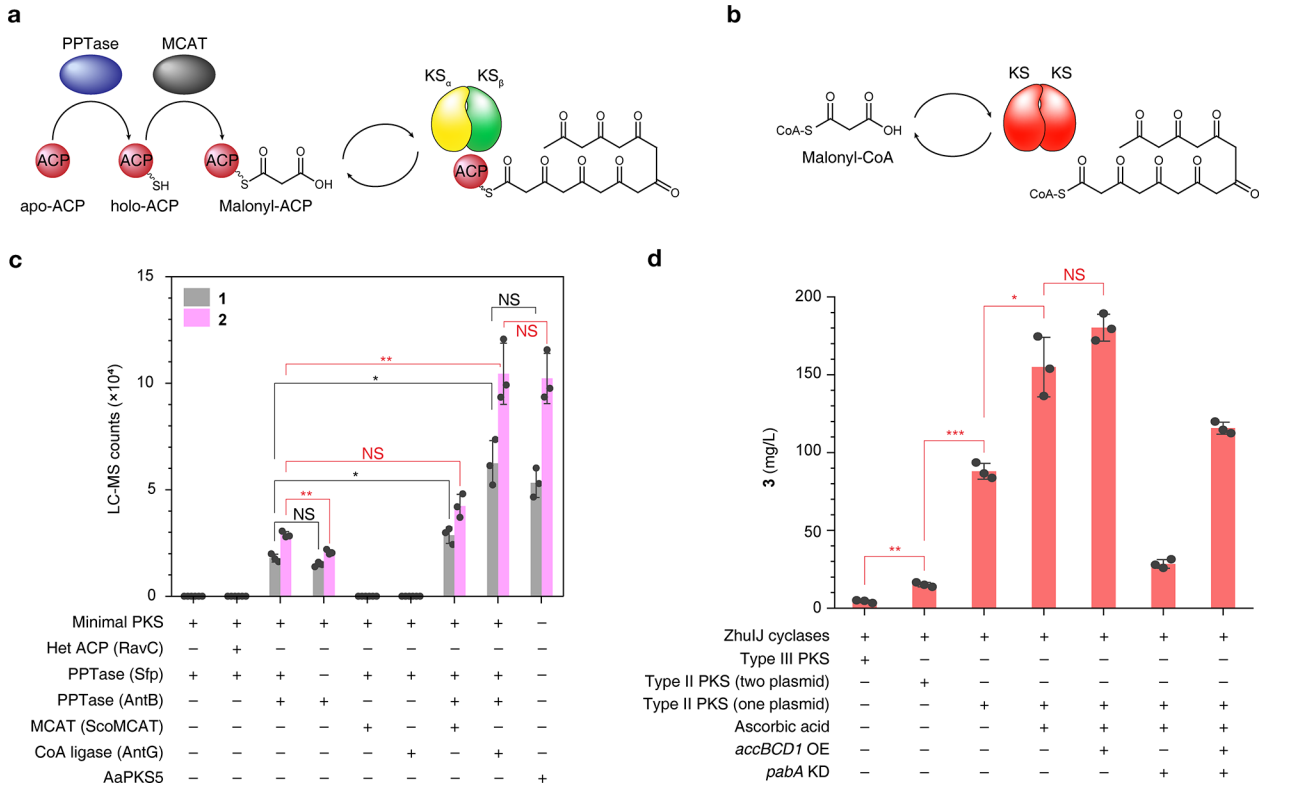


Figure 1. Employment of type II and type III PKS machineries for the production of octaketide shunt products (**1** and **2**) and flavokermesic acid (**3**). (a) Schematic description of the mechanism of the type II PKS system. Abbreviations: PPTase, phosphopantetheinyl transferase; MCAT, malonyl-CoA:ACP malonyltransferase; ACP, acyl-carrier protein; KS, ketosynthase. (b) Schematic description of the mechanism of the type III PKS system. (c) Production of **1** and **2** by different combinations of the type II PKS components. Abbreviation: Het ACP, heterologous ACP. n = 3. (d) Production of **3** by applying different metabolic engineering strategies. Type III PKS denotes AaPKS5 (expressed together with ZhuIJ). Type II PKS denotes AntDEFBG and ZhuIJ expressed from a single plasmid (pFK) or two plasmids (pDS00-antDEFBG and pCDF-zhuIJ). Abbreviations: *accBCD1* OE, *C. glutamicum accBCD1* (encoding acetyl-CoA carboxylase) overexpression; KD, knockdown. Error bars denote mean ± standard deviation (SD; n = 3). (c, d) $^{*}P < 0.0125$, $^{**}P < 0.0025$, $^{***}P < 0.00025$, determined by a two-tailed Student's t-test. P value thresholds were adjusted using a Bonferroni correction (corrected significance levels represented as α/m; α, original significance level; m, number of hypotheses). NS denotes not significant.

6.[10] However, **6** remains one of the most stable natural red colorants in comparison to these alternatives; thus, the demand for the development of sustainable and insect-free production of **6** continues to increase.

Systems metabolic engineering can be a good strategy to solve the aforementioned problems. Systems metabolic engineering has been successfully employed for the sustainable production of bulk and fine chemicals from renewable biomass with high yields.[14−17] Also, the microbial production of an increasing number of important natural products has been reported, including phenylpropanoids (e.g. resveratrol[18]), isoprenoids (e.g. artemisinin[19] and paclitaxel[20]), polyketides (e.g. erythromycin[4] and 6-MSA[5]), alkaloids (e.g. opioids[21] and scopolamine[22]), and cannabinoids.[23] Heterologous **6** production was previously achieved by the filamentous fungus *Aspergillus nidulans* after combining a plant type III PKS (AaOKS), two cyclases (ZhuI and ZhuJ) from *Streptomyces*, a *C*-glucosyltransferase from *D. coccus*, and unknown inherent monooxygenases in *A. nidulans*,[24,25] yet detailed information such as the titer, productivity, or yield was not provided. Furthermore, the monooxygenase required in the later steps of **6** biosynthesis has not been reported.

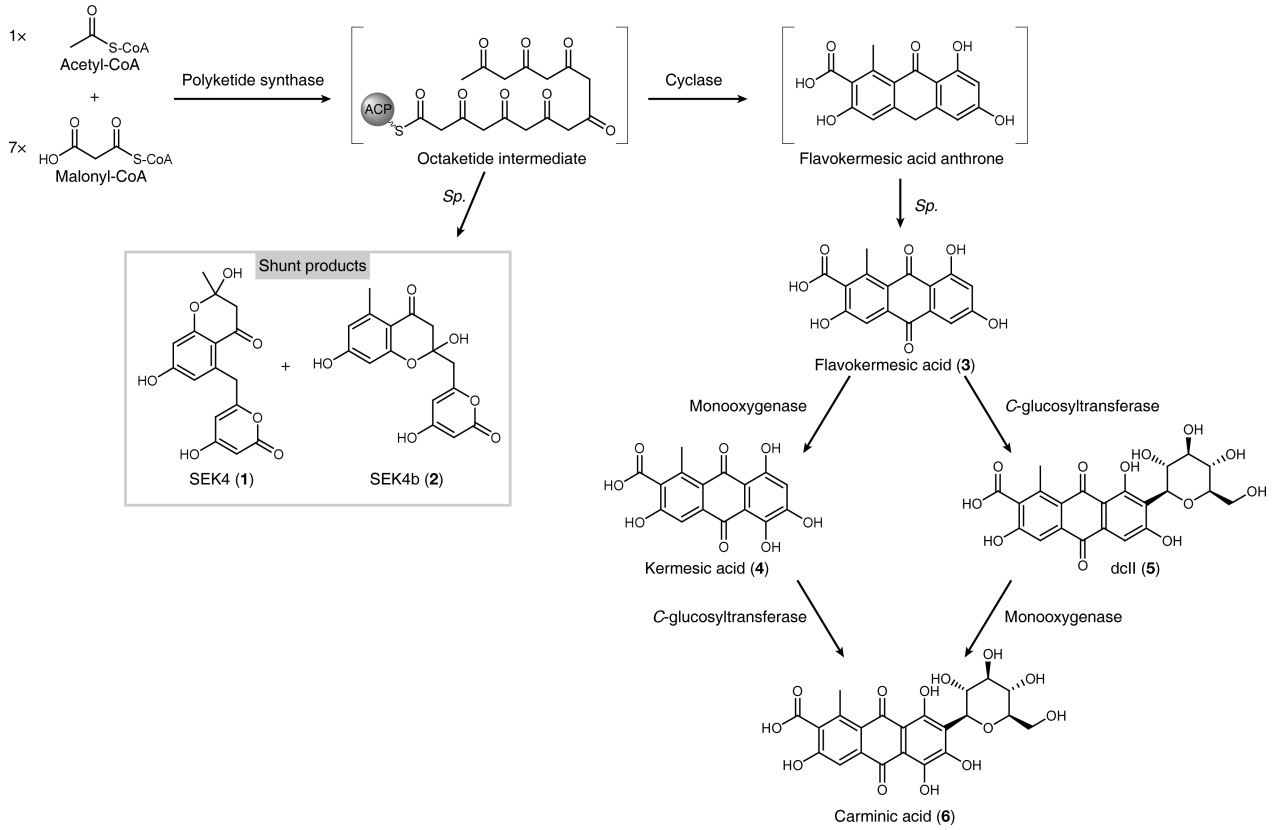
In this study, we report the development of an engineered *Escherichia coli* strain that is capable of the complete biosynthesis of **6** from glucose. To biosynthesize the aromatic polyketide backbone of **6**, the type II and III PKS machineries were compared. Also, biochemical reaction analyses were performed to identify the enzymes responsible for the later steps in the **6** biosynthetic pathway. To enhance the activities of the identified enzymes, *in silico* homology modeling and docking simulations were employed to predict beneficial enzyme mutants, leading to the development of improved enzymes. The final engineered *E. coli* strain developed by combining all these strategies was able to produce **6** from glucose.

RESULTS AND DISCUSSION

Employing Type II PKS and Cyclases for Flavokermesic Acid Production. Although the complete biosynthetic pathway of **6** has not been discovered, **6** is classified as a *C*-glucosylated anthraquinone (a polycyclic aromatic compound with three carbon rings wherein the central ring harbors keto groups) on the basis of its structural features (Scheme 1).[26] To be more specific, flavokermesic acid (**3**) is an octaketide (an aromatic polyketide synthesized from eight ketide units) that can be produced by either type II or III PKSs.[27] The type III PKS machinery requires a single ketosynthase (KS) for carbon chain biosynthesis (Figure 1b), whereas the type II PKS machinery requires a group of enzymes called minimal PKSs comprising ketosynthase (also referred to as KS_{α}), chain-length

https://doi.org/10.1021/jacs.0c12406
J. Am. Chem. Soc. 2021, 143, 5364−5377

Scheme 1. Proposed Biosynthetic Pathway for the Production of Carminic Acid (6)[a]



[a]Abbreviations: ACP, acyl-carrier protein; *Sp.*, spontaneous reaction.

factor (CLF; also referred to as KS_β), and acyl-carrier protein (ACP) (Figure 1a).[6,28] KS_α and KS_β form a heterodimer, which has been reported to be mostly expressed as inclusion bodies in heterologous microbial hosts other than *Streptomyces* species, frustrating numerous previous attempts of heterologously expressing the type II minimal PKS machinery.[28] A number of studies attempted to circumvent the aforementioned type II minimal PKS solubility issue in model microorganisms such as *E. coli*.[29,30] However, none of these studies could solve the underlying PKS solubility problem, leading to inefficient polyketide production with low titers. More recently, type II minimal PKSs that are phylogenetically close to *E. coli* were identified by employing the MiBiG (minimum information about a biosynthetic gene cluster) repository.[31] Another type II minimal PKS from *Streptomyces ambofaciens* was found to be soluble and functional in *E. coli* when the native transcriptional couplings of the genes encoding KS_α and KS_β were maintained, enabling successful production of dehydrorabelomycin.[32] A type III PKS was also employed for the production of aromatic polyketides with type II PKS origin.[33] In this study, we sought to compare both type II and type III PKS machineries for the biosynthesis of the **6** backbone.

First, the type II minimal PKS (AntDEF) from *Photorhabdus luminescens* capable of producing octaketides (AQ256 and other octaketide intermediates) was tested, as it was reported to be soluble and functional in *E. coli*.[31] The *E. coli* BAP1 strain,[5] which is a BL21(DE3) derivative integrated with *Bacillus subtilis* phosphopantetheinyl transferase (PPTase; encoded by *sfp*[34]), was used as the host strain. Two additional auxiliary enzymes were needed: malonyl-CoA:ACP malonyltransferase (MCAT), which converts malonyl-CoA into malonyl-ACP, and PPTase, which activates inactive *apo*-ACP into active *holo*-ACP (Figure 1a). Different components of the type II PKS system were tested for enhanced production of octaketide shunt products **1** and **2** (for details, see Supplementary Note S1). As a result, the expression of *antDEF* (*antD* encoding KS_α; *antE* encoding KS_β; *antF* encoding ACP), *antB* (encoding a PPTase), and *antG* (encoding a CoA ligase[35]) in *E. coli* BAP1 resulted in the highest **1** and **2** titers ($P = 2.1 \times 10^{-3}$ and 8.2×10^{-4}, respectively). Thus, the *antDEFBG* construct introduced into the BAP1 strain was chosen for further studies. The authenticity of **1** and **2** was confirmed by LC-MS/MS (Figure S1a and Table S5) and NMR (Figures S20 and S21 and Table S6) analyses and comparison of the spectra with those of the previously published data.[24,36,37]

In addition to the type II PKS machinery optimization, type III PKSs were also tested for the production of **1** and **2**. We previously tested the production of **1** and **2** from three representative type III PKSs capable of producing octaketides (AaOKS, AaPKS4, and AaPKS5 from *Aloe arborescens*), of which AaPKS5 resulted in the highest production of **1** and **2**.[5] Thus, the optimized type II PKS platform (AntDEFBG) and the type III PKS (AaPKS5) were compared for the production of **1** and **2** (Figure 1c). As the type II and type III PKS machineries showed similar productions of **1** and **2** ($P = 0.28$ and 0.85, respectively), both constructs were examined for the production of **3**. To produce **3**, pCDF-zhuIJ harboring *zhuI* and *zhuJ* from *Streptomyces* sp. R1128 (both encoding cyclases)[24,38] was introduced to BAP1 harboring the type II PKS system (encoded by pDS00-antDEFBG) or the type III

https://doi.org/10.1021/jacs.0c12406
J. Am. Chem. Soc. 2021, 143, 5364−5377

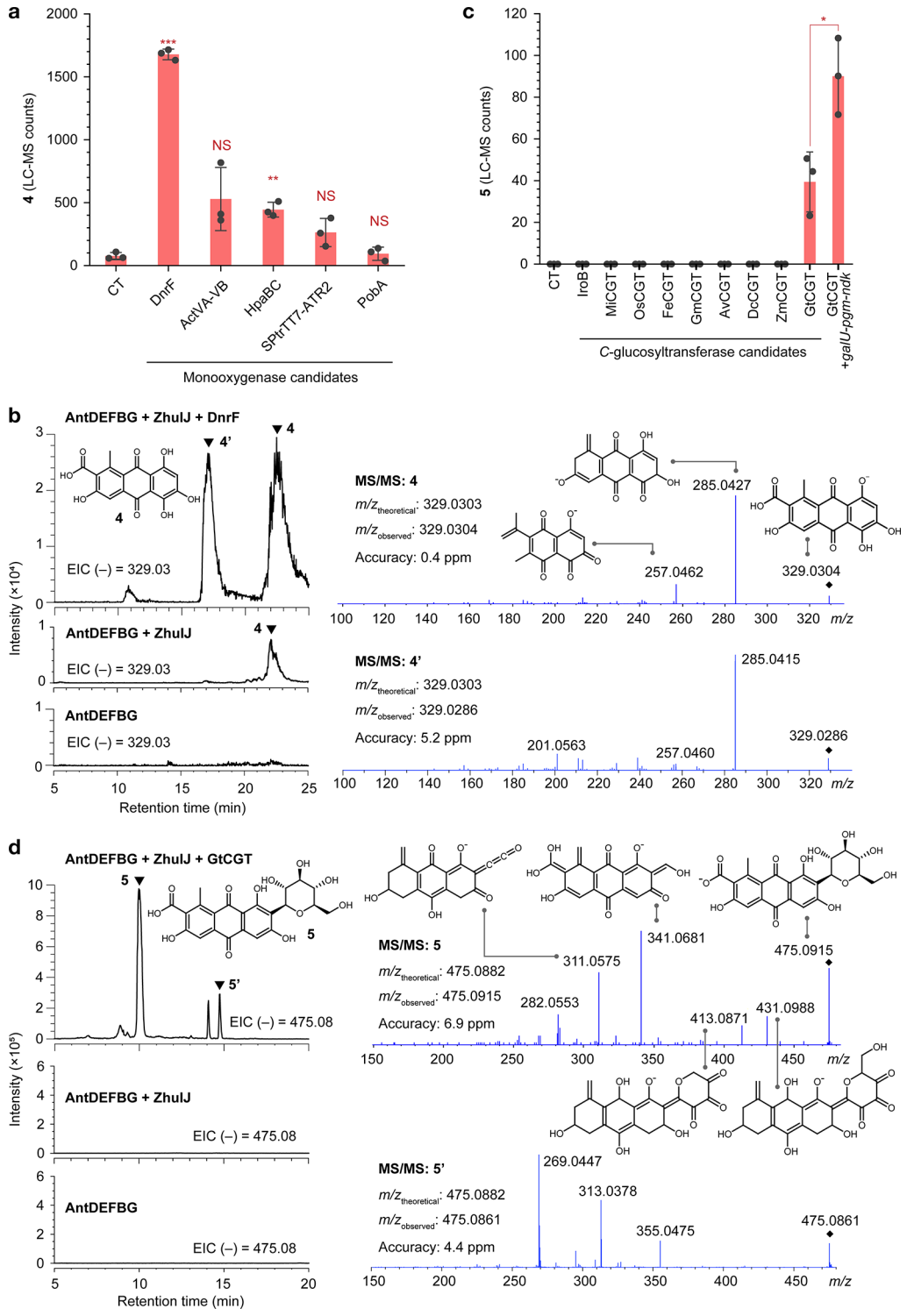


Figure 2. Production of carminic acid precursors (**4** and **5**). (a) Production of kermesic acid (**4**) by employing different monooxygenase candidates identified by a biochemical reaction analysis. *P < 0.01, ***P < 0.0002, determined by a two-tailed Student's t-test. P value thresholds were adjusted using a Bonferroni correction ($P < \alpha/m$). NS denotes not significant. (b) Extracted ion chromatograms (EICs) generated by LC-MS and MS/MS fragmentation patterns of **4** and its isomer **4′** produced by the FK strain additionally harboring *dnrF*. (c) Production of dcII (**5**) by employing different *C*-glucosyltransferase candidates identified by a biochemical reaction analysis. *P < 0.05, determined by a two-tailed Student's t-test. (d) EICs generated by LC-MS and MS/MS fragmentation patterns of **5** and its isomer **5′** produced by the FK strain additionally harboring *GtCGT*. (a, c) Error bars denote mean ± SD (n = 3). CT denotes the control FK strain. (b, d) Data are representative of three independent replicates. Peak annotation and fragment structures were modeled using the CFM-ID tool.[49] Orientations of the fragment ions were adjusted to that of the precursor ion so that fragmentation could be easily identified. Parent ions are marked with filled diamonds. The $m/z_{\text{theoretical}}$ and m/z_{observed} values noted are for the parent ions $[M - H]^-$.

https://doi.org/10.1021/jacs.0c12406
J. Am. Chem. Soc. 2021, 143, 5364−5377

PKS system (encoded by pET-AaPKS5; for more details on the spontaneous conversion of flavokermesic acid anthrone to **3**, see Supplementary Note S1). As a result, BAP1 harboring the type II PKS system resulted in a significantly higher production of **3** (15.1 mg/L) from glucose in comparison with that (4.3 mg/L) produced by BAP1 harboring the type III PKS system (Figure 1d, $P = 4.1 \times 10^{-4}$; for more details on the reasons for higher production of **3** by the type II PKS system, see Supplementary Note S1). The authenticity of **3** was confirmed by LC-MS/MS (Figure S1b and Table S5) and NMR (Figure S22; Table S6) analyses and comparison of the spectra with those of the previously published data.[24]

In order to reduce the metabolic burden of having two plasmids, one plasmid expression system was tested by cloning *zhuIJ* into pDS00-antDEFBG to make pFK. *E. coli* BAP1 harboring pFK produced 88.0 mg/L ($P = 1.8 \times 10^{-5}$; Figure 1d) of **3**, showing significant enhancement. Colocalization of the genes encoding minimal PKS and cyclases is thought to have further increased the chance for the formation of a transient complex between minimal PKS and cyclases by producing the enzymes in proximity. As the culture gradually turned from bright red to brownish (Figure S1d), it was thought that **3** might be polymerized into melanin-like compounds as has been reported for other similar aromatic polyketides.[39] To prevent excess oxidation causing the polymerization of **3**, ascorbic acid was supplemented to the medium, which increased the titer of **3** to 154.9 mg/L ($P = 4.2 \times 10^{-3}$; Figure 1d). To further enhance the production of **3**, increasing the malonyl-CoA pool was pursued by the overexpression of *Corynebacterium glutamicum accBCD1* (encoding acetyl-CoA carboxylase complex) or knockdown of *pabA* (encoding *p*-aminobenzoate synthetase) as we recently reported.[5] Overexpression of *accBCD1* further increased the **3** titer to 180.3 mg/L ($P = 0.105$; Figure 1d and Figure S1c).

Identification of a Monooxygenase for Kermesic Acid Production. Along with **3**, kermesic acid (**4**) was produced in a trace amount, possibly due to the inherent *E. coli* oxygenases (Figure 2a,b). However, this conversion efficiency is too low for the successful production of **6**, and the monooxygenase responsible for this conversion has not yet been reported. Thus, a biochemical reaction analysis was performed on the basis of previous literature and chemical databases; enzyme candidates that could possibly perform the corresponding hydroxylation reaction were identified. Five enzyme candidates were selected from different enzyme categories (polyketide monooxygenase, cytochrome P450, aromatic compound monooxygenase): aklavinone 12-hydroxylase DnrF from *Streptomyces peucetius*,[40] a flavin mononucleotide-dependent monooxygenase (ActVA-ORF5) and its reductase (ActVB) from *Streptomyces coelicolor*,[41] a plant P450 oxygenase (TT7) and its reductase (ATR2) from *Arabidopsis thaliana*,[42] 4-hydroxyphenylacetate 3-monooxygenase (HpaB) and its reductase (HpaC) from *E. coli* BL21,[43] and *p*-hydroxybenzoate hydroxylase (PobA)[44] from *Pseudomonas fluorescens* (Figure S3). DnrF and ActVA-ORF5 are polyketide monooxygenases capable of hydroxylating aromatic polyketides at the position *para* to the pre-existing hydroxyl group (Figure S3). Although the hydroxylation reactions performed by the last three enzymes (TT7-ATR2, HpaB-HpaC, and PobA) are not highly similar to the desired reaction for the production of **4**, these enzymes were chosen for their capabilities to hydroxylate a substituted phenol ring at the position *ortho* to the pre-existing hydroxyl group (Figure S3). Five plasmids (pDnrF, pAct-VAVB, pTT7, pHpaBC, pPobA) each harboring the corresponding gene(s) were separately introduced into the FK strain (BAP1 harboring pFK) and were cultured in flasks. The strain with DnrF produced the highest titer of **4** ($P = 6.6 \times 10^{-7}$; Figure 2a). Production of **4′**, a putative isomer of **4**, was also identified (Figure 2b and Figure S4). Conversion of **3** into **4** and **4′** by DnrF was confirmed by an *in vitro* enzyme assay (Figure S5). The authenticity of **4** was confirmed by LC-MS/MS analysis and a comparison of the MS/MS fragmentation pattern with that of the previously published data (Table S5 and Figure 2b).[45]

Identification of a *C*-Glucosyltransferase for dcII Production. Conversion of **3** into **6** can take one of two pathways, via **4** or via **5**, both requiring a monooxygenase and a *C*-glucosyltranserase (Scheme 1). As glucosylation generally enhances the solubility, stability, and bioavailability of an aglycone (particularly *C*-glucosylation due to the formation of a more stable carbon−carbon bond), the production of **6** is more favorable than the production of its aglycone **3**.[46,47] DcUGT2 from *D. coccus* was recently identified to be responsible for the conversion of **3** to **5** (or from **4** to **6**) and was shown to be active in *Saccharomyces cerevisiae*.[24,45] However, as this enzyme is heavily glycosylated in its active form and contains a C-terminal transmembrane helix structure and a eukaryotic N-terminal signal peptide which generally do not work in *E. coli*, functional expression of *DcUGT2* in *E. coli* was expected to be difficult; indeed, expression of *DcUGT2* in the FK strain failed to produce **5**. To test if cytoplasmic expression of *DcUGT2* leads to the production of **5**, the elimination of its N-terminal signal peptide (Ntr-DcUGT2), the C-terminal transmembrane helix (Ctr-DcUGT2), or both (Ntr-Ctr-DcUGT2) was performed. However, these constructs did not allow the production of **5**. Next, the *E. coli* OmpA signal peptide was tethered to the N-termini of Ntr-DcUGT2 and Ntr-Ctr-DcUGT2 to properly guide DcUGT2 to the inner membrane and restore its activity.[48] However, the production of **5** was not observed. Thus, it was not possible for us to functionally express *DcUGT2* in *E. coli*.

In comparison to the better studied *O*-glucosylation, the *C*-glucosylation of natural products has been more difficult in bacteria such as *E. coli*.[50] To identify *C*-glucosyltransferases (CGTs) that can convert **3** to **5** (or **4** to **6**), a biochemical reaction analysis was performed to identify enzyme candidates with a high probability of performing the corresponding reaction. UDP-glycosyltransferases that have been previously reported to perform *C*-glucosylation in *E. coli* were selected with high priority. Eight enzyme candidates were selected as follows: IroB (EnCGT) from *E. coli* Nissle,[51] UGT708A6 (ZmCGT) from *Zea mays* which acts as a dual *C*/*O*-glucosyltransferase,[52] UGT708C2 (FeCGT) from *Fagopyrum esculentum*,[53] MiCGT from *Mangifera indica*,[54] OsCGT from *Oryza sativa*,[55] UGT708D1 (GmCGT) from *Glycine max*,[56] GtUF6CGT1 (GtCGT) from *Gentiana triflora*,[57] and AvCGT from *Aloe vera* (Figure S6).[58] Despite the low similarity of the reaction performed by IroB to the desired reaction for **5** production, IroB was chosen because it is from *E. coli*. Thus, IroB was expected to be functional in *E. coli* BAP1, which would perform the desired reaction if it is promiscuous enough to accommodate **3** as its substrate. Among these eight candidates, only GtCGT could successfully convert **3** into **5** (Figure 2c,d).

With **5** as the major product, its putative isomer **5′** was also observed (Figure 2d). A comparison of the MS/MS

https://doi.org/10.1021/jacs.0c12406
J. Am. Chem. Soc. 2021, 143, 5364−5377

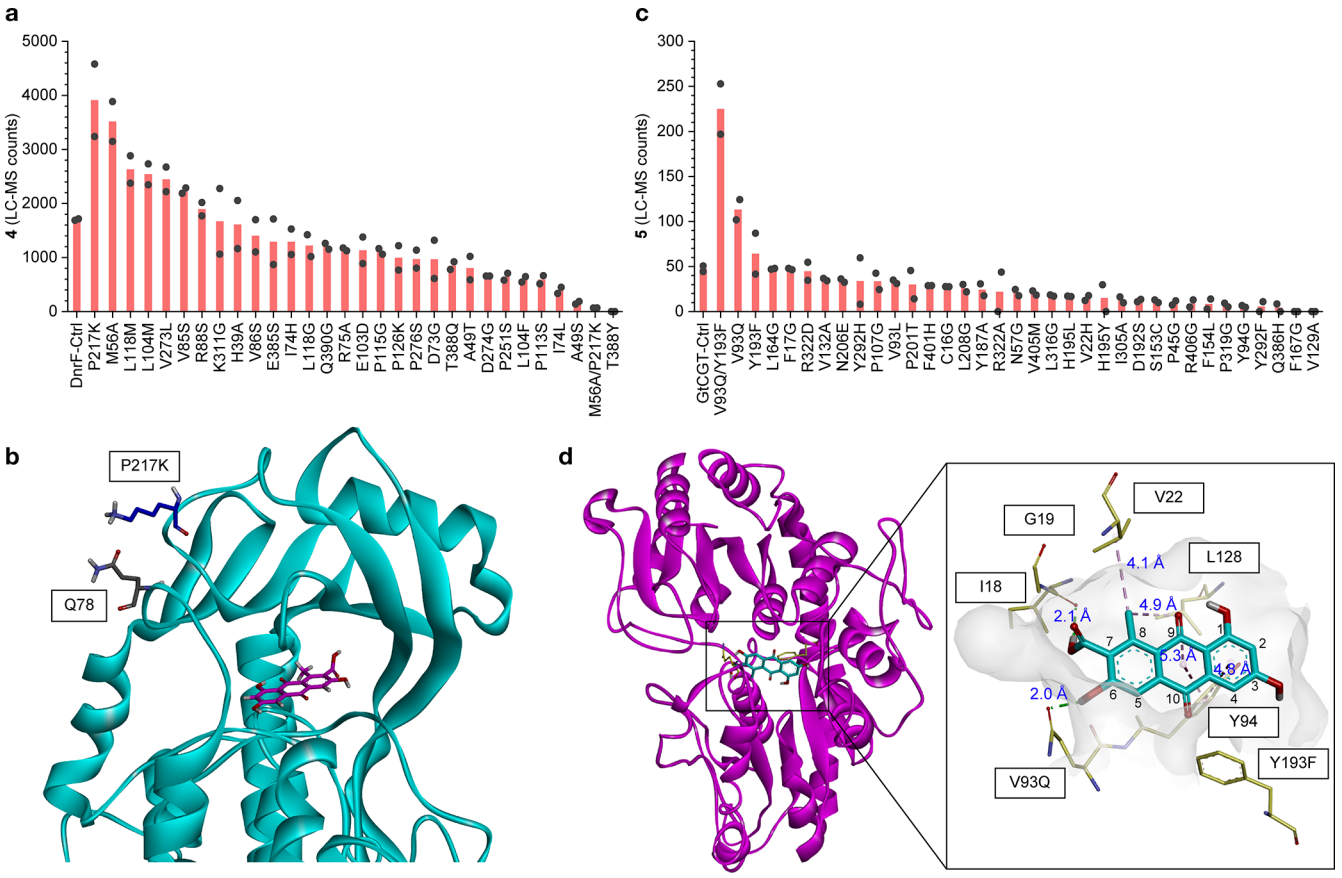


Figure 3. Homology modeling and docking simulations of DnrF and GtCGT for enhanced enzymatic activities. (a) Enhanced production of **4** by testing DnrF mutants predicted by *in silico* simulations. (b) The mutated amino acid residue (P217K) and the mode of the ligand (**3**) binding in the best DnrF mutant ($DnrF^{P217K}$). (c) Enhanced production of **5** by testing GtCGT mutants predicted by *in silico* simulations. (d) Major amino acid residues (including the mutated amino acid residues V93Q and Y193F) interacting with the ligand (**3**) in the best GtCGT mutant ($GtCGT^{V93Q/Y193F}$). (a, c) $n = 2$.

fragmentation pattern of **5′** with the predicted MS/MS spectra of possible **5** isomers (generated by CFM-ID 3.0[49]) indicated that **5′** is likely to be an *O*-glucosylated **3** due to the presence of characteristic ion peaks for *O*-glucosylation (Figure S7); whereas the stable *C*-glucoside bonds are well preserved among the predicted fragment ions generated from the *C*-glucoside isomers, the less stable *O*-glucoside bonds are predicted to be more easily cleaved to result in ion peaks such as *m*/*z* 269.04555 (decarboxylated **3** anion) and *m*/*z* 313.03538 (**3** anion) (Figure S7). As these characteristic peaks were observed in the MS/MS spectrum of **5′**, we identified **5′** as an *O*-glucoside of **3**. Another peak shown at a retention time of 14.09 min (between **5** and **5′**) in the EIC 475.08 does not correspond to an isomer of **5**, since the *m*/*z* value of the precursor ion (*m*/*z* 475.0702) was different from the expected *m*/*z* value of a **5** isomer (*m*/*z* 475.0882; error 37.9 ppm). When ZmCGT was employed, **5′′**, another isomer of **5**, was produced (Figure S8). In the MS/MS spectrum of **5′′**, one of the characteristic ion peaks for *O*-glucosylation (*m*/*z* 269.0476) was observed while the characteristic ion peaks for *C*-glucosylation (e.g. *m*/*z* 341.06668, 311.05611) were not observed (Figures S7 and S8). Also, **5′′** was identified as an *O*-glucosylated **3** by a previous report (Table S5).[25] Next, conversion of **3** into **5**, **5′**, and **5′′** by GtCGT (with **5** as the major product) was confirmed by an *in vitro* enzyme assay (Figure S5). The authenticity of **5** was confirmed by LC-MS/MS (Figure 2d and Table S5) and NMR (Figure S23) analyses and a comparson of the spectra with those of the previously published data.[25,26,45,59]

As *C*-glucosylation requires a high UDP-glucose pool, overexpressions of *galU* (encoding UTP-glucose-1-phosphate uridylyltransferase), *pgm* (encoding phosphoglucomutase), and *ndk* (encoding nucleoside-diphosphate kinase) were tested for the enhanced generation of UDP-glucose and enhanced UDP recycling,[60] further increasing **5** production by 2.3-fold (Figure 2c).

After verification of all the conversion steps required for the production of **6**, an *E. coli* BAP1 strain harboring pDS00-antDEFBG-zhuIJ and pCDF-dnrF-GtCGT, which thus express all the necessary genes (*antDEFBG, zhuIJ, dnrF,* and *GtCGT*) was constructed and cultured in baffled flasks. However, the production of **6** was not observed. On consideration that 180.3 mg/L of **3** could be produced from glucose, it was speculated that the downstream enzymes DnrF and GtCGT did not have sufficient activities to convert **3** to **6**.

Development of DnrF and GtCGT Mutants Having Enhanced Activities on the Basis of Homology Modeling and Docking Simulations. Structure-based *in silico* docking simulations of DnrF and GtCGT with **3** (as their common ligand) were performed to generate the mutant enzymes having enhanced activities. Since the crystal structures of DnrF and GtCGT were not available, homology modeling was first performed using MODELLER[61] to predict their structures. Next, docking simulations were performed using

https://doi.org/10.1021/jacs.0c12406
J. Am. Chem. Soc. 2021, 143, 5364−5377

PyRosetta[62] to predict beneficial mutants with enhanced binding affinities toward **3**. Beneficial mutants were also predicted by rationally examining the predicted protein structures, especially amino acid arrangements within the catalytic pockets.

First, homology modeling was performed on DnrF using RdmE (aklavinone 11-hydroxylase) from *Streptomyces purpurascens* (PDB ID 3IHG)[63] as a template (Figure S9a), which had the highest protein sequence identity with DnrF (60.4%). The generated DnrF structure model was used for *in silico* docking simulations with **3** (see Materials and Methods in the Supporting Information for the detailed simulation procedure). Among the 20^{234} theoretical mutants, 162 *in silico* mutants that were stable in terms of potential energy were first chosen, followed by the selection of 142 mutants having docking scores higher than that of the wild type enzyme (Table S7). The top 20 mutants (denoted as bold letters in Table S7) with the highest docking scores were selected for experimental analysis. An additional 9 mutants (denoted as blue letters in Table S7) generated by rational examination of the predicted DnrF structure were also selected. Each of these 29 mutants was introduced into the FK strain (BAP1 harboring pDS00-antDEFBG-zhuIJ) and was flask-cultured.

Seven mutants showed a higher level production of **4** in comparison with the wild type DnrF, among which the DnrF^{P217K} mutant showed the highest **4** production (Figure 3a). This corresponds to a 2.2-fold higher titer of **4** in comparison with that obtained with the strain harboring the wild type DnrF. Examining this mutant revealed that the new lysine residue (P217K) is close to the active site, putatively causing a conformational change in the overall DnrF structure including the active site (Figure 3b and Figure S9b). Also, as the Lys217 residue forms new hydrogen bonds with the Gln78 residue, it might have affected the flexibility or thermostability of the enzyme (Figure 3b). Since Lys217 is located at the surface of the enzyme and is positively charged near the neutral pH, it might also have contributed to the enhanced solubility of the enzyme. However, the expression of *dnrF*^{P217K} was unaffected by the P217K mutation (Figure S10), suggesting that the higher level production of **4** was not due to the enhanced enzyme (expression) level or higher enzyme solubility but rather was due to an unknown reason such as higher specific activity.

Employing the M56A mutant resulted in the second highest **4** titer (greater than 2-fold increase in **4** production in comparison with that of the wild type DnrF; Figure 3a). The Ala56 residue in the mutant is also located near the active site. Whereas the original Met56 residue is predicted to have hydrophobic interactions with Met114, Ala117, and Ile68, the new Ala56 residue is predicted to only interact with Met114, eliciting a conformational change in the active site (Figure S9c). To see the synergistic effects of the two mutations P217K and M56A, the double mutant DnrF$^{P217K/M56A}$ was generated and was introduced into the FK strain. Surprisingly, the resultant double mutant showed a significantly decreased **4** titer (Figure 3a). Thus, the DnrF^{P217K} mutant was used for further experiments.

To better understand the docking simulation results, docking of DnrF, DnrF^{P217K}, and DnrF^{M56A} with the desired substrate (**3**) and the native substrate (aklavinone) were compared using Autodock Vina (Figure S11 and Table S8). As predicted by PyRosetta (Table S8), docking of the DnrF mutants with **3** showed stronger binding affinities in comparison to that of the wild type DnrF. On the other hand, aklavinone is predicted to bind well to both the wild type DnrF and the DnrF mutants (Figure S11 and Table S8). The binding energy calculated for the DnrF^{P217K} mutant with **3** (−9.9 kcal/mol) was comparable with that (−10.4 kcal/mol) calculated for the wild type DnrF with the native substrate aklavinone, confirming that the docking simulations were successful. In addition, the volume of the binding pocket within the DnrF mutants (DnrF^{P217K} and DnrF^{M56A}) was predicted to be increased (Table S8), which is thought to be one of the probable reasons for the enhanced enzymatic activity toward **3**.

Next, the steady-state kinetic parameters of DnrF and DnrF^{P217K} were determined (concentration of enzymes set to 1 μM), showing a 23.7% reduction in the K_M value of DnrF^{P217K} (43.31 μM) in comparison with that (56.75 μM) of DnrF (P = 0.012; Table 1 and Figure S12). Although the V_{max} value of DnrF^{P217K} (1.01 μM min^{-1}) was similar to that (0.99 μM min^{-1}) of DnrF (P = 0.78), the V_{max}/K_M value of DnrF^{P217K} (2.34×10^{-2} min^{-1}) showed 33.4% enhancement in comparison with that (1.75×10^{-2} min^{-1}) of DnrF (Table 1), confirming the enhanced catalytic efficiency of the DnrF^{P217K} mutant.

Table 1. Steady-State Kinetic Parameters of DnrF, DnrF^{P217K}, GtCGT, and GtCGT$^{V93Q/Y193F}$, with 3 as the Common Substrate

	K_M (μM)[a]	V_{max} (μM min^{-1})[a,b]	V_{max}/K_M (min^{-1})[b]
DnrF	56.75 ± 4.68	0.99 ± 0.086	1.75×10^{-2}
DnrF^{P217K}	43.31 ± 2.64	1.01 ± 0.062	2.34×10^{-2}
GtCGT	6.30 ± 1.64	$1.60 \times 10^{-3} \pm 1.39 \times 10^{-4}$	2.54×10^{-4}
GtCGT$^{V93Q/Y193F}$	5.07 ± 0.74	$1.89 \times 10^{-3} \pm 1.13 \times 10^{-4}$	3.73×10^{-4}

[a]Data are means ± standard deviation (SD) of three independent experiments. [b]Since the enzyme preparations were not pure (Figure S12b), V_{max}/K_M was calculated instead of k_{cat}/K_M.

Next, homology modeling was performed on GtCGT using TcCGT (a *C*-glycosyltransferase) from *Trollius chinensis* (PDB ID 6JTD)[64] as a template (Figure S13a), which had the highest protein sequence identity with GtCGT (35.1%). The generated GtCGT structure model was used for *in silico* docking simulations with **3** (see Materials and Methods in the Supporting Information for the detailed simulation procedure). Among the 20^{249} theoretical mutants, 203 *in silico* mutants that are stable in terms of potential energy were first chosen followed by selection of 113 mutants having docking scores higher than that of the wild type enzyme (Table S9). The top 20 mutants (denoted as bold letters in Table S9) with the highest docking scores were selected for experimental analysis. An additional 14 mutants (denoted as blue letters in Table S9) generated by rational examination of the predicted GtCGT structure were also selected. Each of these 34 mutants was introduced into the FK strain and was flask-cultured.

Six mutants showed a higher level production of **5** in comparison with the wild type GtCGT, among which the GtCGT^{V93Q} mutant showed the highest **5** production (Figure 3c). This corresponds to a 2.8-fold higher titer of **5** in comparison to that obtained with the strain harboring the wild type GtCGT. In this mutant, the new Gln93 residue is located in the active site and is predicted to directly interact with the

https://doi.org/10.1021/jacs.0c12406
J. Am. Chem. Soc. 2021, 143, 5364−5377

ligand **3** (Figure S13b,c). The amino acid residues surrounding the active site properly orient the ligand so that it can be easily *C*-glucosylated at C2 (Figure S13c).

Employing the Y193F mutant resulted in the second highest **5** titer, showing a greater than 2-fold increase in **5** production in comparison with that obtained with the wild type GtCGT (Figure 3c). To see the synergistic effects of the two mutations V93Q and Y193F, the double mutant GtCGT$^{V93Q/Y193F}$ was generated and was introduced into the FK strain. Employing the double mutant resulted in a 5.3-fold increase in the titer of **5** in comparison with that obtained with the wild type GtCGT (Figure 3c). In the GtCGT^{V93Q} mutant, the Tyr193 residue forms a hydrogen bond with the carbonyl group at C10, blocking the Gln93 residue from forming a hydrogen bond with the hydroxyl group at C6 (Figure S13c). Thus, exchanging the Tyr193 residue with Phe193 (to construct GtCGT$^{V93Q/Y193F}$) is predicted to result in the formation of a robust hydrogen bond between Gln93 and the hydroxyl group at C6, orienting **3** to be more favorable for *C*-glucosylation at C2 (Figure 3d and Figure S13d). The expression of *GtCGT*$^{V93Q/Y193F}$ was not affected by the V93Q/Y193F double mutant (Figure S10), suggesting that the higher level production of **5** was not due to the enhanced enzyme (expression) level or higher enzyme solubility but was rather due to the better alignment of the ligand **3** in the active site.

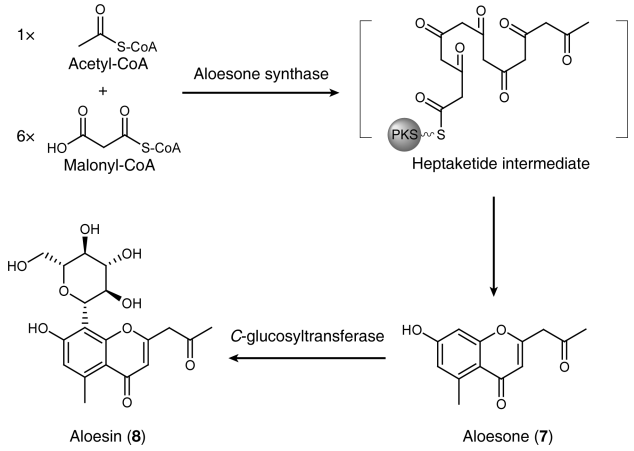
To better understand the docking simulation results, docking of the best mutant GtCGT$^{V93Q/Y193F}$ and the wild type GtCGT with the ground-state substrate (**3**), the transition-state analogue of **3** (**3′**), and the product (**4**) were compared using Autodock Vina (Figure S14 and Table S10). The chemical structure of **3′** was created on consideration of the reaction mechanism of *C*-glucosylation (Figure S14).[47] As high enzymatic activity is generally accompanied by a strong binding between the enzyme and the transition state of the substrate, the enhanced enzymatic activity of GtCGT$^{V93Q/Y193F}$ toward **3** can be reasoned by the lower predicted binding energy of GtCGT$^{V93Q/Y193F}$ with **3′** (−9.6 kcal/mol) in comparison with that (−7.9 kcal/mol) of GtCGT (Figure S14 and Table S10). Moreover, the predicted binding energy of GtCGT$^{V93Q/Y193F}$ with the product **4** (−8.1 kcal/mol) was higher in comparison with that (−8.4 kcal/mol) of GtCGT (Figure S14 and Table S10), allowing a more favorable release of **4** by GtCGT$^{V93Q/Y193F}$.

Next, steady-state kinetic parameters of GtCGT and GtCGT$^{V93Q/Y193F}$ were determined (concentration of enzymes set to 0.1 μM), showing 19.5% reduction in the K_M value of GtCGT$^{V93Q/Y193F}$ (5.07 μM) in comparison with that (6.30 μM) of GtCGT, despite an insignificant difference (P = 0.30; Table 1 and Figure S12). On the other hand, the V_{max} value of GtCGT$^{V93Q/Y193F}$ (1.89×10^{-3} μM min^{-1}) was increased by 18.2% in comparison with that (1.60×10^{-3} μM min^{-1}) of GtCGT (P = 0.048). When this is taken together, the V_{max}/K_M value of GtCGT$^{V93Q/Y193F}$ (3.73×10^{-4}) showed 46.8% enhancement in comparison with that (2.54×10^{-4}) of GtCGT (Table 1), confirming the enhanced catalytic efficiency of the GtCGT$^{V93Q/Y193F}$ mutant.

GtCGT$^{V93Q/Y193F}$ Can *C*-Glucosylate Aloesone to Produce Aloesin. As *C*-glucosylation is a relatively unexplored reaction in bacteria, including *E. coli*, we sought to test whether the CGT can also be applied to other similar *C*-glucosylation reactions; for this, GtCGT$^{V93Q/Y193F}$ having the highest activity toward converting **3** into **5** was employed for aloesin production. Aloesin (**8**) is an aromatic heptaketide (an aromatic polyketide synthesized from seven ketide units) extracted from the leaves of *Aloe vera*. Due to its antityrosinase and antimelanogenesis activities, it has been widely used as a skin-whitening agent in the cosmetic industry.[65,66] As **8** is also reported to exhibit anti-inflammatory and free-radical-scavenging activity, it is also considered as a potential therapeutic or cosmetic ingredient for a wide range of applications.[67]

We previously reported the construction of an engineered *E. coli* BL21(DE3) strain harboring pCDF-RpALS, pWAS-anti-pabA, and pBBR1-zwf, which expresses *RpALS* (encoding aloesone synthase from *R. palmatum*), anti-*pabA* synthetic sRNA, and *zwf* (encoding glucose 6-phosphate 1-dehydrogenase from *E. coli*).[5] This engineered strain was capable of producing 30.9 mg/L of **7** from glucose.[5] However, the CGT responsible for the conversion of **7** to **8** has not yet been identified (Scheme 2), which led us to test GtCGT$^{V93Q/Y193F}$

Scheme 2. Proposed Biosynthetic Pathway for the Production of Aloesin (8)



for the production of **8**. Before testing the *C*-glucosylation of **7**, enhanced production of **7** was pursued first by further increasing the expression level of *RpALS* through its expression from a compatible plasmid (Supplementary Note S4). The resultant strain *E. coli* BL21(DE3) harboring pCDF-RpALS, pRSF-RpALS, pWAS-anti-pabA, and pBBR1-zwf produced 102.1 mg/L of **7** by a flask culture (Figure 4a and Supplementary Note S4). Next, the *E. coli* BL21(DE3) strain harboring pWAS-anti-pabA, pRSF-RpALS, and pBBR1-zwf was transformed with pCDF-GtCGT or pCDF-GtCGTmut (harboring *GtCGT*$^{V93Q/Y193F}$) to convert **7** to **8**. Both strains successfully produced **8**, while the strain harboring *GtCGT*$^{V93Q/Y193F}$ produced **8** to a higher titer (0.060 μg/L; P = 0.046; Figure 4b). As in the production of **7** described above and in Supplementary Note S4, additional copies of *RpALS* were introduced by insertion of RpALS into pCDF-GtCGTmut (pCDF-RpALS-GtCGTmut) to construct the ALS strain, which is BL21(DE3) harboring pRSF-RpALS, pCDF-RpALS-GtCGTmut, pWAS-anti-pabA, and pBBR1-zwf. The ALS strain could produce 0.30 μg/L of **8** (P = 0.011; Figure 4b). This represents the first report on the microbial production of **8**. Also, this result suggests that GtCGT$^{V93Q/Y193F}$ is likely to be a generally applicable CGT for aromatic polyketides in *E. coli*. The titer of **8** can be further enhanced by carrying out additional metabolic engineering and performing docking simulations on GtCGT$^{V93Q/Y193F}$ with **7** as the ligand. The authenticity of **8** was confirmed by LC-MS/MS

https://doi.org/10.1021/jacs.0c12406
J. Am. Chem. Soc. 2021, 143, 5364−5377

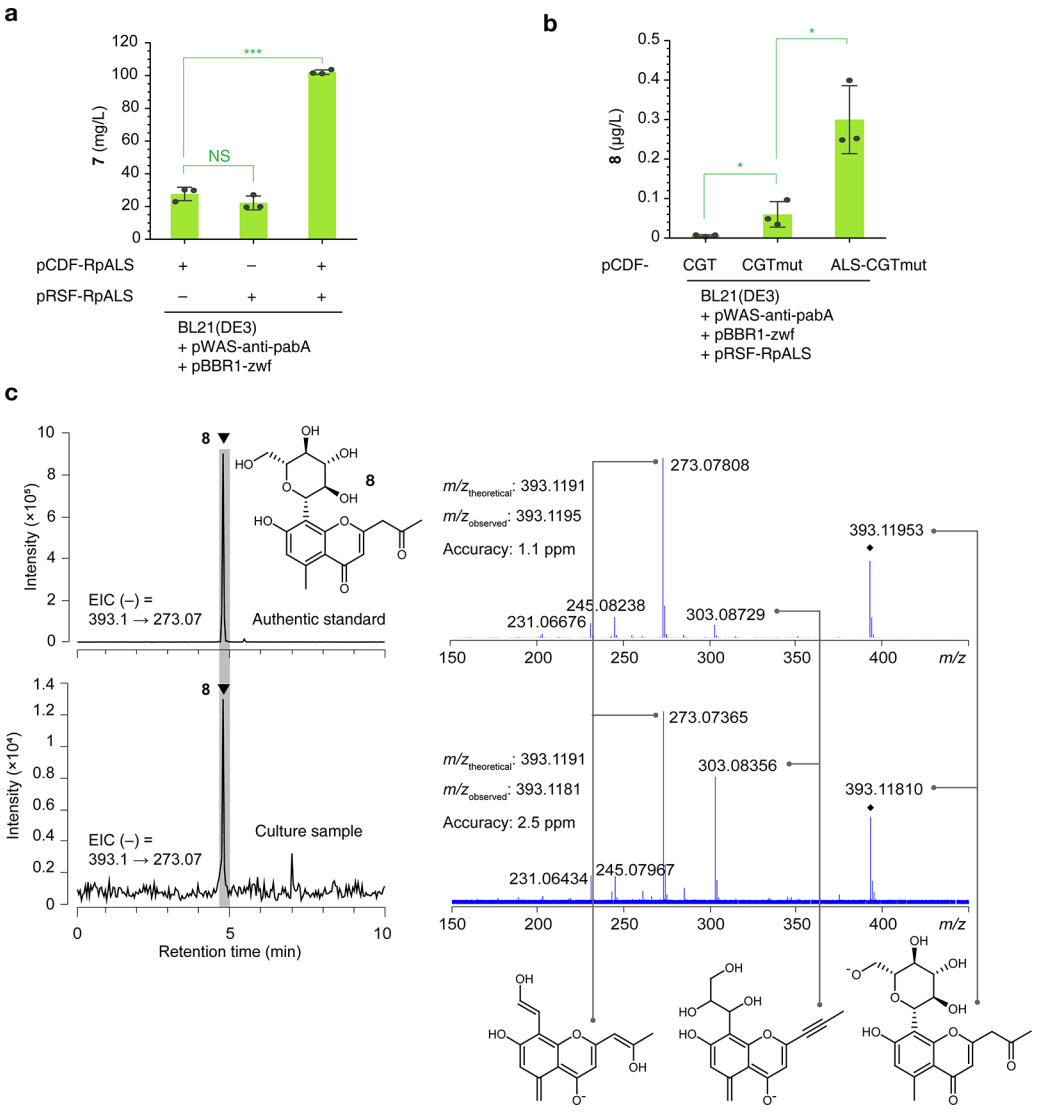


Figure 4. Production of aloesin (**8**) from glucose. (a) Production of aloesone (**7**) by testing different copies of aloesone synthase (ALS). The plasmids pCDF-RpALS and/or pRSF-RpALS were additionally introduced into *E. coli* BL21(DE3) harboring pWAS-anti-pabA and pBBR1-zwf for the flask culture. (b) Production of **8** by employing GtCGT (noted as CGT) or the GtCGT mutant (GtCGT$^{V93Q/Y193F}$, denoted as CGTmut). pCDF-GtCGT, pCDF-GtCGTmut, or pCDF-RpALS-GtCGTmut was additionally introduced into *E. coli* BL21(DE3) harboring pWAS-anti-pabA, pBBR1-zwf, and pRSF-RpALS for the flask culture. (c) EICs generated by LC-MS and fragmentation patterns of **8** generated by MS/MS from the authentic standard (upper panels) or from the culture sample of the ALS strain (lower panels). Data are representative of three independent replicates. Peak annotation and fragment structures were modeled using the CFM-ID tool.[49] Orientations of the fragment ions were adjusted to that of the precursor ion so that fragmentation could be easily identified. Parent ions are marked with filled diamonds. The $m/z_{theoretical}$ and $m/z_{observed}$ values noted are for the parent ions $[M - H]^-$. (a, b) Error bars are mean ± SD (n = 3). *P < 0.05, ***P < 0.001, determined by a two-tailed Student's t-test. NS denotes not significant.

analysis and a comparison of the MS/MS fragmentation pattern with that of the authentic standard chemical (Figure 4c).

Since the native substrate of GtCGT is luteolin,[68] which is a flavonoid, the efficiencies for the conversion of luteolin and other similar flavonoids into their *C*-glucosides by GtCGT and GtCGT$^{V93Q/Y193F}$ were compared (Supplementary Note S4). *In vivo* conversion assays were performed by exogenously feeding luteolin, apigenin, and naringenin during the flask culture of *E. coli* BL21(DE3) harboring either pCDF-GtCGT or pCDF-GtCGTmut. As a result, GtCGT$^{V93Q/Y193F}$ showed a generally decreased conversion efficiency for flavonoids (Figure S15 and Supplementary Note S4). This indicates that the increased activity of the GtCGT$^{V93Q/Y193F}$ mutant toward aromatic polyketides resulted in decreased activity toward its native substrates, as an outcome of enzyme engineering.

Production of Carminic Acid from Glucose. Having constructed the mutants of DnrF and GtCGT having enhanced activities, production of **6** from glucose was attempted by constructing the CA strain, which is the BAP1 strain harboring pFK and pCA. A flask culture of the CA strain successfully produced 22.2 μg/L of **6** from glucose (Figure 5a). To further enhance **6** production, *C. glutamicum accBCD1* overexpression, *pabA* knockdown, and/or *galU-pgm-ndk* overexpression were tested alone or in combinations. The CA-ACC strain, which overexpresses *accBCD1*, produced 74.9 μg/L of **6**, corresponding to a 3.4-fold increase in comparison with that (22.2 μg/L) obtained with the base strain (P = 2.3 × 10^{-3}; Figure 5a and

https://doi.org/10.1021/jacs.0c12406
J. Am. Chem. Soc. 2021, 143, 5364–5377

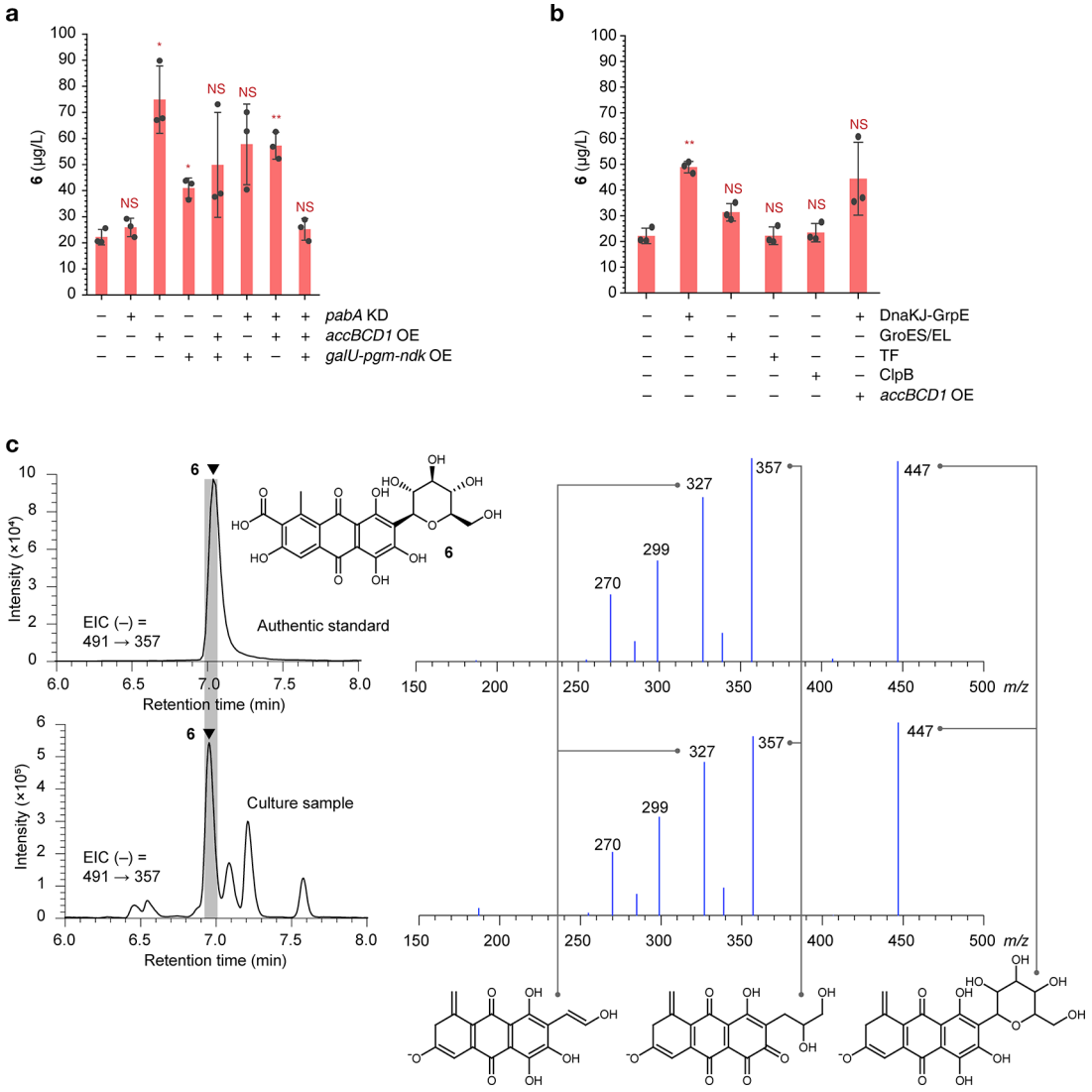


Figure 5. Production of carminic acid (**6**) from glucose. (a) Production of **6** by applying different metabolic engineering strategies. *P < 0.0071, **P < 0.0014, determined by a two-tailed Student's t-test. P value thresholds were adjusted using a Bonferroni correction ($P < \alpha/m$). (b) Production of **6** by coexpression of chaperones. Abbreviations: KD, knockdown; OE, overexpression. Error bars are the mean ± SD (standard deviation; $n = 3$). **P < 0.002, determined by a two-tailed Student's t-test. P value thresholds were adjusted using a Bonferroni correction ($P < \alpha/m$). NS, not significant. (c) EICs generated by LC-MS and fragmentation patterns of **6** generated by MS/MS in MRM mode from the authentic standard (upper panels) or from the culture sample of the CA-ACC strain (lower panels). Data are representative of three independent replicates (Figure S17). Peak annotation and fragment structures were modeled using the CFM-ID tool.[49] Orientations of the fragment ions were adjusted to that of the precursor ion so that fragmentation could be easily identified. The $m/z_{\text{theoretical}}$ and m/z_{observed} values noted are for the parent ions $[M - H]^-$.

Figure S16). Four different *E. coli* chaperones, DnaK-DnaJ-GrpE, GroES/EL, trigger factor (TF), and ClpB, were also employed to facilitate folding of insoluble heterologous enzymes.[69] Overexpression of *dnaKJ-grpE* and *groES/EL* gave enhanced productions of 48.9 μg/L ($P = 2.4 \times 10^{-4}$) and 31.3 μg/L ($P = 0.024$) of **6** (Figure 5b). It was reasoned that simultaneous overexpression of *accBCD1* and *dnaKJ-grpE* would further increase the titer of **6**. However, the titer of **6** (44.4 μg/L; Figure 5b) did not further increase, and thus the CA-ACC strain was selected as the final strain. The authenticity of **6** was confirmed by an LC-MS/MS analysis and a comparison of the MS/MS fragmentation pattern with that of the commercially available standard chemical (Figure 5c and Figure S17).

Finally, fed-batch cultures were carried out to assess the performance of the strain in a 6.6 L bioreactor. The pH-stat fed-batch fermentation of the CA-ACC strain (for details, see Materials and Methods in the Supporting Information) resulted in the production of 0.65 mg/L of **6** (Figure 6a–c and Figure S18). Another fed-batch fermentation performed independently produced 0.61 mg/L of **6** (Figure 6d–f), showing good reproducibility. However, a maximum of 3.66 g/L of **3** was also produced (Figure 6b,e), suggesting that a large portion of **3** was not converted to the final product **6**, leaving room for further improvement. When this was taken together, we were able to produce **6** for the first time by a metabolically engineered bacterium.

CONCLUSION

Aromatic polyketides encompass a diverse set of compounds essential for our daily lives. However, the production of these

https://doi.org/10.1021/jacs.0c12406
J. Am. Chem. Soc. 2021, 143, 5364–5377

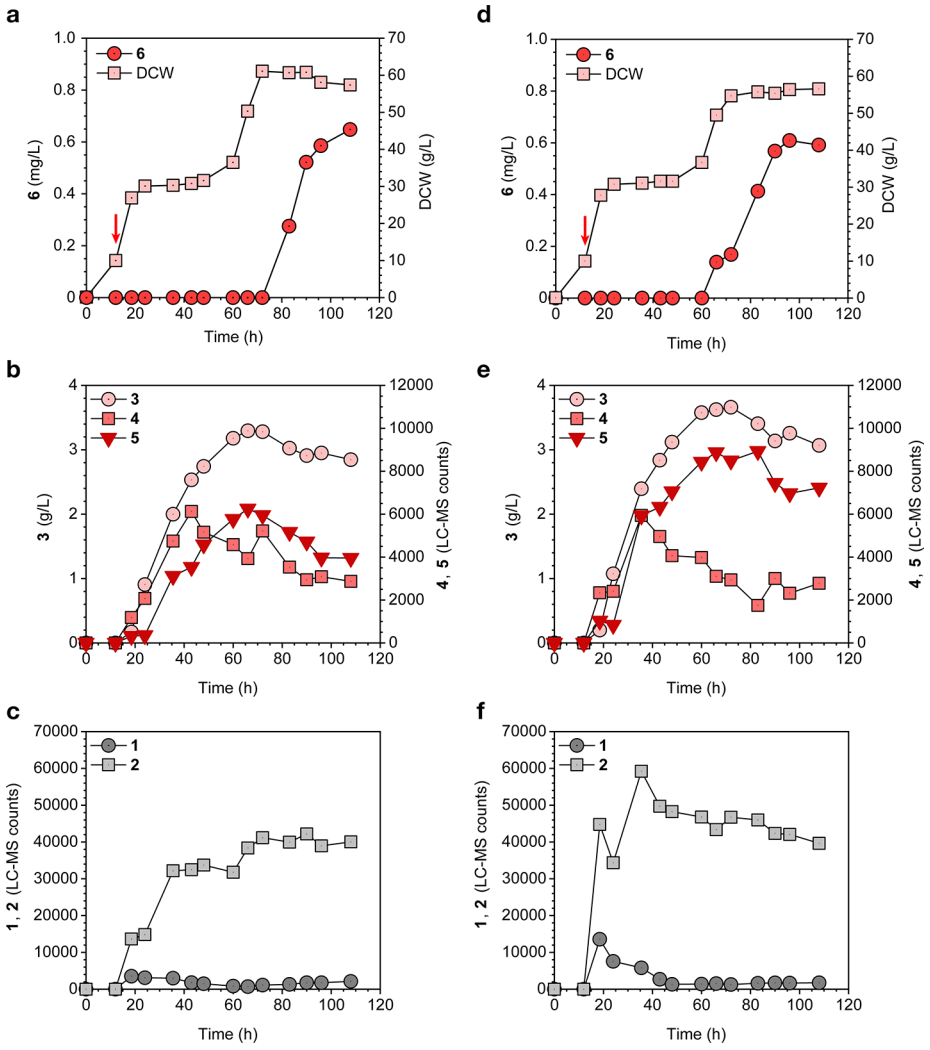


Figure 6. Time profiles of the fed-batch fermentation of the CA-ACC strain producing **6**. Fed-batch fermentation profiles of CA-ACC regarding (a) **6** and DCW, (b), intermediates **3**–**5**, and (c) shunt products **1** and **2**. The profiles of independently repeated fed-batch fermentation of CA-ACC are shown in (d–f), showing reproducibility. The red arrow denotes the isopropyl *β*-D-1-thiogalactopyranoside (IPTG) induction time point. DCW denotes dry cell weight.

important polyketides from their native producers has been limited due to a number of problems, including the slow growth of the host, the requirement of a large cultivation area and specific climate, the incapability of a high cell density culture, and the requirement for multistep purification process, necessitating the development of better microbial production systems. In this study, we constructed a novel biosynthetic pathway for the production of **6** in *E. coli* from glucose. After optimization of the type II minimal PKS construct along with the introduction of heterologous cyclases, the production of **3** was achieved. The biosynthetic pathway of **6** has been previously proposed in the filamentous fungus *A. nidulans*,[24] but the hydroxylase responsible for the conversion of **3** to **4** (or **5** to **6**) in *Aspergillus* remained unidentified and the CGT from *D. coccus* responsible for the conversion of **3** to **5** (or **4** to **6**) was not functional in *E. coli*. Therefore, to complete the remaining metabolic pathways for the biosynthesis of **6**, biochemical reaction analyses and *in silico* homology modeling coupled with docking simulations were performed to replace or complement the unidentified hydroxylase and DcUGT2. Integrating all these strategies led to the production of **6** from glucose.

This study not only represents a demonstration of microbial production of **6** and **8** but also suggests an important strategy for constructing novel biosynthetic pathways for polyketide production. Biochemical reaction analyses coupled with *in silico* docking simulations can be a powerful tool in constructing novel biosynthetic pathways containing unknown enzymes, thus widening the spectrum of natural products that can be produced in model microorganisms such as *E. coli*. We also showed that GtCGT$^{V93Q/Y193F}$ is likely to be a generally applicable CGT for aromatic polyketides in *E. coli*, as demonstrated by the production of **8** in addition to **6**; the CGT responsible for the conversion of **7** to **8** has not yet been identified, but we were able to use GtCGT$^{V93Q/Y193F}$ to convert **7** to **8**. This strategy will be particularly useful in producing pharmaceuticals and nutraceuticals (such as **6** or **8**), as *C*-glucosides generally improve pharmacokinetic properties such as stability, resistance to hydrolysis, and bioavailability.[70] Also, performing docking simulations using an alternative target ligand will further enhance the activity of the CGT toward the production of a desired *C*-glucoside. Since *C*-glucosylation reactions have not been well reported in bacteria, the strategies employed for the development of GtCGT$^{V93Q/Y193F}$ reported

https://doi.org/10.1021/jacs.0c12406
J. Am. Chem. Soc. 2021, 143, 5364–5377

here will advance the microbial production of diverse *C*-glucoside natural products with beneficial properties.

The promising result of producing **6** by metabolic engineering strategies described in this study suggests that further improvement in the strain performance would allow a transition from the traditional insect-based **6** production process into a microbial fermentation-based process. As 3.66 g/L of **3** was produced during the fermentation, it is reasonable to expect that the **6** titer can reach up to ∼5 g/L in the near future with further systems metabolic engineering effort.[19,71] Assuming this improved titer, a single run of fed-batch fermentation (∼5 days) in a 100000 L scale industrial fermenter can produce the same amount of **6** produced from the insect culture per hectare of land per annum (Supplementary Note S5). Switching to the microbial production system will also allow the whole process to be independent of geopolitical or climate issues, much less labor-intensive, and better standardized to reduce the current variations in product properties and compositions. When this is taken together, the strains and metabolic engineering strategies developed in this study for the production of **6** (and also **8**) will contribute to the sustainable production of diverse aromatic polyketides important for the food, pharmaceutical, and cosmetic industries.

EXPERIMENTAL SECTION

Full experimental details on the following can be found in Materials and Methods in the Supporting Information: materials; plasmid construction; media and culture conditions; homology modeling and docking simulations of DnrF and GtCGT; fed-batch fermentation; SDS-PAGE analysis; enzyme purification; *in vitro* enzyme assay; HPLC, LC-MS, and LC-MS/MS analysis; NMR analysis; statistical analysis; data availability.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/jacs.0c12406.

Supplementary notes, tables, figures, and materials and methods as described in the text (PDF)

AUTHOR INFORMATION

Corresponding Author

Sang Yup Lee − *Metabolic and Biomolecular Engineering National Research Laboratory, Department of Chemical and Biomolecular Engineering (BK21 plus program), Institute for the BioCentury, Korea Advanced Institute of Science and Technology (KAIST), Daejeon 34141, Republic of Korea; Systems Metabolic Engineering and Systems Healthcare Cross-Generation Collaborative Laboratory and BioProcess Engineering Research Center and BioInformatics Research Center, Korea Advanced Institute of Science and Technology (KAIST), Daejeon 34141, Republic of Korea; Novo Nordisk Foundation Center for Biosustainability, Technical University of Denmark, 2800 Kongens Lyngby, Denmark;* orcid.org/0000-0003-0599-3091; Email: leesy@kaist.ac.kr

Authors

Dongsoo Yang − *Metabolic and Biomolecular Engineering National Research Laboratory, Department of Chemical and Biomolecular Engineering (BK21 plus program), Institute for the BioCentury, Korea Advanced Institute of Science and Technology (KAIST), Daejeon 34141, Republic of Korea; Systems Metabolic Engineering and Systems Healthcare Cross-Generation Collaborative Laboratory and BioProcess Engineering Research Center and BioInformatics Research Center, Korea Advanced Institute of Science and Technology (KAIST), Daejeon 34141, Republic of Korea;* orcid.org/0000-0002-0299-3128

Woo Dae Jang − *Metabolic and Biomolecular Engineering National Research Laboratory, Department of Chemical and Biomolecular Engineering (BK21 plus program), Institute for the BioCentury, Korea Advanced Institute of Science and Technology (KAIST), Daejeon 34141, Republic of Korea; Systems Metabolic Engineering and Systems Healthcare Cross-Generation Collaborative Laboratory, Korea Advanced Institute of Science and Technology (KAIST), Daejeon 34141, Republic of Korea;* orcid.org/0000-0003-1649-0174

Complete contact information is available at: https://pubs.acs.org/10.1021/jacs.0c12406

Notes

The authors declare the following competing financial interest(s): S.Y.L. and D.Y. declare that the *C*-glycosyltransferase technology developed here is in a patent filed including, but not limited to, KR 10-2021-0011326.

ACKNOWLEDGMENTS

The authors thank Dr. Seon Young Park and Dr. Kyeong Rok Choi for helpful discussions, Eunju Lee at UC Davis Economics Department for advice on statistics, Yu Ri Lee at Chungnam National University Center for Research Facilities for LC-HRMS analyses, Dr. Yun-Cheol Na at Korea Basic Science Institute (KBSI), Western Seoul Center, for LC-MS/MS analyses, Ji Won Kim at KAIST Bio Core Center for LC-MS (MRM) analyses, and Jae Hyun Lee at KBSI, Western Seoul Center, for NMR analysis. This work was supported by the Technology Development Program to Solve Climate Changes on Systems Metabolic Engineering for Biorefineries (Grants NRF-2012M1A2A2026556 and NRF-2012M1A2A2026557) of the Ministry of Science and ICT (MSIT) through the National Research Foundation (NRF) of Korea and the KAIST Cross-Generation Collaborative Lab project; S.Y.L. and D.Y. were also supported by Novo Nordisk Foundation grant NNF16OC0021746.

REFERENCES

(1) Khosla, C.; Keasling, J. D. Metabolic engineering for drug discovery and development. *Nat. Rev. Drug Discovery* **2003**, *2* (12), 1019−1025.

(2) Staunton, J.; Weissman, K. J. Polyketide biosynthesis: a millennium review. *Nat. Prod. Rep.* **2001**, *18* (4), 380−416.

(3) Kim, E.; Moore, B. S.; Yoon, Y. J. Reinvigorating natural product combinatorial biosynthesis with synthetic biology. *Nat. Chem. Biol.* **2015**, *11* (9), 649−659.

(4) Zhang, H.; Wang, Y.; Wu, J.; Skalina, K.; Pfeifer, B. A. Complete biosynthesis of erythromycin A and designed analogs using *E. coli* as a heterologous host. *Chem. Biol.* **2010**, *17* (11), 1232−1240.

(5) Yang, D.; Kim, W. J.; Yoo, S. M.; Choi, J. H.; Ha, S. H.; Lee, M. H.; Lee, S. Y. Repurposing type III polyketide synthase as a malonyl-CoA biosensor for metabolic engineering in bacteria. *Proc. Natl. Acad. Sci. U. S. A.* **2018**, *115* (40), 9835−9844.

(6) Hertweck, C.; Luzhetskyy, A.; Rebets, Y.; Bechthold, A. Type II polyketide synthases: gaining a deeper insight into enzymatic teamwork. *Nat. Prod. Rep.* **2007**, *24* (1), 162−190.

(7) Zaya, P. L. B.; Oehlschlager, A. C. *Carmine extraction technology*; International Development Research Centre (IDRC) Digital Library 1999 (Accessed online February 9, 2021); https://idl-bnc-idrc.dspacedirect.org/handle/10625/21568.
(8) Lloyd, A. G. Extraction and chemistry of cochineal. *Food Chem.* **1980**, *5* (1), 91−107.
(9) Borges, M. E. T. R. L.; Díaz, L.; Esparza, P.; Ibáñez, E. Natural dyes extraction from cochineal (*Dactylopius coccus*). *Food Chem.* **2012**, *132* (4), 1855−1860.
(10) Soteriou, H. S. W. Why you may have been eating insects your whole life. BBC News (April 27, 2018; accessed online February 9, 2021); https://www.bbc.com/news/business-43786055.
(11) Takeo, N.; Nakamura, M.; Nakayama, S.; Okamoto, O.; Sugimoto, N.; Sugiura, S.; Sato, N.; Harada, S.; Yamaguchi, M.; Mitsui, N.; Kubota, Y.; Suzuki, K.; Terada, M.; Nagai, A.; Sowa-Osako, J.; Hatano, Y.; Akiyama, H.; Yagami, A.; Fujiwara, S.; Matsunaga, K. Cochineal dye-induced immediate allergy: Review of Japanese cases and proposed new diagnostic chart. *Allergol. Int.* **2018**, *67* (4), 496−505.
(12) Lizaso, M. T. M. I.; García, B. E.; Acero, S.; Quirce, S.; Tabar, A. I. Identification of allergens involved in occupational asthma due to carmine dye. *Ann. Allergy, Asthma, Immunol.* **2000**, *84* (5), 549−552.
(13) Dapson, R. W. The history, chemistry and modes of action of carmine and related dyes. *Biotech. Histochem.* **2007**, *82* (4−5), 173−187.
(14) Lee, S. Y.; Kim, H. U.; Chae, T. U.; Cho, J. S.; Kim, J. W.; Shin, J. H.; Kim, D. I.; Ko, Y. S.; Jang, W. D.; Jang, Y. S. A comprehensive metabolic map for production of bio-based chemicals. *Nat. Catal.* **2019**, *2* (1), 18−33.
(15) Nielsen, J.; Fussenegger, M.; Keasling, J.; Lee, S. Y.; Liao, J. C.; Prather, K.; Palsson, B. Engineering synergy in biotechnology. *Nat. Chem. Biol.* **2014**, *10* (5), 319−322.
(16) Pontrelli, S.; Chiu, T. Y.; Lan, E. I.; Chen, F. Y. H.; Chang, P. C.; Liao, J. C. *Escherichia coli* as a host for metabolic engineering. *Metab. Eng.* **2018**, *50*, 16−46.
(17) Ko, Y. S.; Kim, J. W.; Lee, J. A.; Han, T.; Kim, G. B.; Park, J. E.; Lee, S. Y. Tools and strategies of systems metabolic engineering for the development of microbial cell factories for chemical production. *Chem. Soc. Rev.* **2020**, *49* (14), 4615−4636.
(18) Li, M.; Schneider, K.; Kristensen, M.; Borodina, I.; Nielsen, J. Engineering yeast for high-level production of stilbenoid antioxidants. *Sci. Rep.* **2016**, *6*, 36827.
(19) Paddon, C. J.; Westfall, P. J.; Pitera, D. J.; Benjamin, K.; Fisher, K.; McPhee, D.; Leavell, M. D.; Tai, A.; Main, A.; Eng, D.; Polichuk, D. R.; Teoh, K. H.; Reed, D. W.; Treynor, T.; Lenihan, J.; Fleck, M.; Bajad, S.; Dang, G.; Dengrove, D.; Diola, D.; Dorin, G.; Ellens, K. W.; Fickes, S.; Galazzo, J.; Gaucher, S. P.; Geistlinger, T.; Henry, R.; Hepp, M.; Horning, T.; Iqbal, T.; Jiang, H.; Kizer, L.; Lieu, B.; Melis, D.; Moss, N.; Regentin, R.; Secrest, S.; Tsuruta, H.; Vazquez, R.; Westblade, L. F.; Xu, L.; Yu, M.; Zhang, Y.; Zhao, L.; Lievense, J.; Covello, P. S.; Keasling, J. D.; Reiling, K. K.; Renninger, N. S.; Newman, J. D. High-level semi-synthetic production of the potent antimalarial artemisinin. *Nature* **2013**, *496* (7446), 528−532.
(20) Ajikumar, P. K.; Xiao, W. H.; Tyo, K. E. J.; Wang, Y.; Simeon, F.; Leonard, E.; Mucha, O.; Phon, T. H.; Pfeifer, B.; Stephanopoulos, G. Isoprenoid pathway optimization for taxol precursor overproduction in *Escherichia coli*. *Science* **2010**, *330* (6000), 70−74.
(21) Galanie, S.; Thodey, K.; Trenchard, I. J.; Interrante, M. F.; Smolke, C. D. Complete biosynthesis of opioids in yeast. *Science* **2015**, *349* (6252), 1095−1100.
(22) Srinivasan, P.; Smolke, C. D. Biosynthesis of medicinal tropane alkaloids in yeast. *Nature* **2020**, *585* (7826), 614−619.
(23) Luo, X. Z.; Reiter, M. A.; d'Espaux, L.; Wong, J.; Denby, C. M.; Lechner, A.; Zhang, Y. F.; Grzybowski, A. T.; Harth, S.; Lin, W. Y.; Lee, H.; Yu, C. H.; Shin, J.; Deng, K.; Benites, V. T.; Wang, G.; Baidoo, E. E. K.; Chen, Y.; Dev, I.; Petzold, C. J.; Keasling, J. D. Complete biosynthesis of cannabinoids and their unnatural analogues in yeast. *Nature* **2019**, *567* (7746), 123−126.
(24) Frandsen, R. J. N.; Khorsand-Jamal, P.; Kongstad, K. T.; Nafisi, M.; Kannangara, R. M.; Staerk, D.; Okkels, F. T.; Binderup, K.; Madsen, B.; Møller, B. L.; Thrane, U.; Mortensen, U. H. Heterologous production of the widely used natural food colorant carminic acid in *Aspergillus nidulans*. *Sci. Rep.* **2018**, *8* (1), 12853.
(25) Andersen-Ranberg, J.; Kongstad, K. T.; Nafisi, M.; Staerk, D.; Okkels, F. T.; Mortensen, U. H.; Møller, B. L.; Frandsen, R. J. N.; Kannangara, R. Synthesis of C-glucosylated octaketide anthraquinones in *Nicotiana benthamiana* by using a multispecies-based biosynthetic pathway. *ChemBioChem* **2017**, *18* (19), 1893−1897.
(26) Rasmussen, S. A.; Kongstad, K. T.; Khorsand-Jamal, P.; Kannangara, R. M.; Nafisi, M.; Van Dam, A.; Bennedsen, M.; Madsen, B.; Okkels, F.; Gotfredsen, C. H.; Staerk, D.; Thrane, U.; Mortensen, U. H.; Larsen, T. O.; Frandsen, R. J. N. On the biosynthetic origin of carminic acid. *Insect Biochem. Mol. Biol.* **2018**, *96*, 51−61.
(27) McDaniel, R.; Ebert-Khosla, S.; Hopwood, D. A.; Khosla, C. Rational design of aromatic polyketide natural-products by recombinant assembly of enzymatic subunits. *Nature* **1995**, *375* (6532), 549−554.
(28) Gao, X.; Wang, P.; Tang, Y. Engineered polyketide biosynthesis and biocatalysis in *Escherichia coli*. *Appl. Microbiol. Biotechnol.* **2010**, *88* (6), 1233−1242.
(29) Zhang, W. J.; Li, Y. R.; Tang, Y. Engineered biosynthesis of bacterial aromatic polyketides in *Escherichia coli*. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105* (52), 20683−20688.
(30) Stevens, D. C.; Conway, K. R.; Pearce, N.; Villegas-Peñaranda, L. R.; Garza, A. G.; Boddy, C. N. Alternative sigma factor overexpression enables heterologous expression of a type II polyketide biosynthetic pathway in *Escherichia coli*. *PLoS One* **2013**, *8* (5), No. e64858.
(31) Cummings, M.; Peters, A. D.; Whitehead, G. F. S.; Menon, B. R. K.; Micklefield, J.; Webb, S. J.; Takano, E. Assembling a plug-and-play production line for combinatorial biosynthesis of aromatic polyketides in *Escherichia coli*. *PLoS Biol.* **2019**, *17* (7), No. e3000347.
(32) Liu, X.; Hua, K.; Liu, D.; Wu, Z. L.; Wang, Y.; Zhang, H.; Deng, Z.; Pfeifer, B. A.; Jiang, M. Heterologous biosynthesis of type II polyketide products using *E. coli*. *ACS Chem. Biol.* **2020**, *15* (5), 1177−1183.
(33) Jakociunas, T.; Klitgaard, A. K.; Kontou, E. E.; Nielsen, J. B.; Thomsen, E.; Romero-Suarez, D.; Blin, K.; Petzold, C. J.; Gin, J. W.; Tong, Y. J.; Gotfredsen, C. H.; Charusanti, P.; Frandsen, R. J. N.; Weber, T.; Lee, S. Y.; Jensen, M. K.; Keasling, J. D. Programmable polyketide biosynthesis platform for production of aromatic compounds in yeast. *Synth. Syst. Biotechnol.* **2020**, *5* (1), 11−18.
(34) Pfeifer, B. A.; Admiraal, S. J.; Gramajo, H.; Cane, D. E.; Khosla, C. Biosynthesis of complex polyketides in a metabolically engineered strain of *E. coli*. *Science* **2001**, *291* (5509), 1790−1792.
(35) Bräuer, A.; Zhou, Q.; Grammbitter, G. L. C.; Schmalhofer, M.; Rühl, M.; Kaila, V. R. I.; Bode, H. B.; Groll, M. Structural snapshots of the minimal PKS system responsible for octaketide biosynthesis. *Nat. Chem.* **2020**, *12* (8), 755−763.
(36) Fu, H.; Ebert-Khosla, S.; Hopwood, D. A.; Khosla, C. Engineered biosynthesis of novel polyketides: dissection of the catalytic specificity of the *act* ketoreductase. *J. Am. Chem. Soc.* **1994**, *116*, 4166−4170.
(37) Fu, H.; Hopwood, D. A.; Khosla, C. Engineered biosynthesis of novel polyketides: evidence for temporal, but not regiospecific, control of cyclization of an aromatic polyketide precursor. *Chem. Biol.* **1994**, *1* (4), 205−210.
(38) Ames, B. D.; Lee, M. Y.; Moody, C.; Zhang, W.; Tang, Y.; Tsai, S. C. Structural and biochemical characterization of ZhuI aromatase/cyclase from the R1128 polyketide pathway. *Biochemistry* **2011**, *50* (39), 8392−8406.
(39) Funa, N.; Funabashi, M.; Ohnishi, Y.; Horinouchi, S. Biosynthesis of hexahydroxyperylenequinone melanin via oxidative aryl coupling by cytochrome P-450 in *Streptomyces griseus*. *J. Bacteriol.* **2005**, *187* (23), 8149−8155.

https://doi.org/10.1021/jacs.0c12406
J. Am. Chem. Soc. 2021, 143, 5364−5377

(40) Hutchinson, C. R.; Colombo, A. L. Genetic engineering of doxorubicin production in *Streptomyces peucetius*: a review. *J. Ind. Microbiol. Biotechnol.* **1999**, *23* (1), 647−652.
(41) Taguchi, T.; Maruyama, T.; Sawa, R.; Igarashi, M.; Okamoto, S.; Ichinose, K. Structure and biosynthetic implication of 5*R*-(*N*-acetyl-L-cysteinyl)-14*S*-hydroxy-dihydrokalafungin from a mutant of the *act*VA-ORF4 gene for actinorhodin biosynthesis in *Streptomyces coelicolor* A3(2). *J. Antibiot.* **2015**, *68* (7), 481−483.
(42) Park, S. Y.; Yang, D.; Ha, S. H.; Lee, S. Y. Biosynthesis of dihydroquercetin in *Escherichia coli* from glycerol. **2020**-11−27. *bioRxiv* DOI: 10.1101/2020.11.27.401000 (accessed March 24, 2021).
(43) Jones, J. A.; Collins, S. M.; Vernacchio, V. R.; Lachance, D. M.; Koffas, M. A. Optimization of naringenin and *p*-coumaric acid hydroxylation using the native *E. coli* hydroxylase complex, HpaBC. *Biotechnol. Prog.* **2016**, *32* (1), 21−25.
(44) Eppink, M. H.; Schreuder, H. A.; van Berkel, W. J. Lys42 and Ser42 variants of *p*-hydroxybenzoate hydroxylase from *Pseudomonas fluorescens* reveal that Arg42 is essential for NADPH binding. *Eur. J. Biochem.* **1998**, *253* (1), 194−201.
(45) Kannangara, R.; Siukstaite, L.; Borch-Jensen, J.; Madsen, B.; Kongstad, K. T.; Staerk, D.; Bennedsen, M.; Okkels, F. T.; Rasmussen, S. A.; Larsen, T. O.; Frandsen, R. J. N.; Møller, B. L. Characterization of a membrane-bound C-glucosyltransferase responsible for carminic acid biosynthesis in *Dactylopius coccus* Costa. *Nat. Commun.* **2017**, *8*, 1987.
(46) Huang, G.; Lv, M.; Hu, J.; Huang, K.; Xu, H. Glycosylation and activities of natural products. *Mini-Rev. Med. Chem.* **2016**, *16* (12), 1013−1016.
(47) Putkaradze, N.; Teze, D.; Fredslund, F.; Welner, D. H. Natural product C-glycosyltransferases - a scarcely characterised enzymatic activity with biotechnological potential. *Nat. Prod. Rep.* **2021**, DOI: 10.1039/D0NP00040J.
(48) Lee, H. C.; Bernstein, H. D. The targeting pathway of *Escherichia coli* presecretory and integral membrane proteins is specified by the hydrophobicity of the targeting signal. *Proc. Natl. Acad. Sci. U. S. A.* **2001**, *98* (6), 3471−3476.
(49) Djoumbou-Feunang, Y.; Pon, A.; Karu, N.; Zheng, J.; Li, C.; Arndt, D.; Gautam, M.; Allen, F.; Wishart, D. S. CFM-ID 3.0: Significantly improved ESI-MS/MS prediction and compound identification. *Metabolites* **2019**, *9* (4), 72.
(50) Chen, D.; Chen, R.; Xie, K.; Duan, Y.; Dai, J. Production of acetophenone C-glucosides using an engineered C-glycosyltransferase in *Escherichia coli*. *Tetrahedron Lett.* **2018**, *59* (19), 1875−1878.
(51) Fischbach, M. A.; Lin, H.; Liu, D. R.; Walsh, C. T. *In vitro* characterization of IroB, a pathogen-associated C-glycosyltransferase. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102* (3), 571−576.
(52) Falcone Ferreyra, M. L.; Rodriguez, E.; Casas, M. I.; Labadie, G.; Grotewold, E.; Casati, P. Identification of a bifunctional maize C- and O-glucosyltransferase. *J. Biol. Chem.* **2013**, *288* (44), 31678−31688.
(53) Nagatomo, Y.; Usui, S.; Ito, T.; Kato, A.; Shimosaka, M.; Taguchi, G. Purification, molecular cloning and functional characterization of flavonoid C-glucosyltransferases from *Fagopyrum esculentum* M. (buckwheat) cotyledon. *Plant J.* **2014**, *80* (3), 437−448.
(54) Chen, D.; Chen, R.; Wang, R.; Li, J.; Xie, K.; Bian, C.; Sun, L.; Zhang, X.; Liu, J.; Yang, L.; Ye, F.; Yu, X.; Dai, J. Probing the catalytic promiscuity of a regio- and stereospecific C-glycosyltransferase from *Mangifera indica*. *Angew. Chem., Int. Ed.* **2015**, *54* (43), 12678−12682.
(55) Brazier-Hicks, M.; Edwards, R. Metabolic engineering of the flavone-C-glycoside pathway using polyprotein technology. *Metab. Eng.* **2013**, *16*, 11−20.
(56) Hirade, Y.; Kotoku, N.; Terasaka, K.; Saijo-Hamano, Y.; Fukumoto, A.; Mizukami, H. Identification and functional analysis of 2-hydroxyflavanone C-glucosyltransferase in soybean (*Glycine max*). *FEBS Lett.* **2015**, *589* (15), 1778−1786.
(57) Shrestha, A.; Pandey, R. P.; Dhakal, D.; Parajuli, P.; Sohng, J. K. Biosynthesis of flavone C-glucosides in engineered *Escherichia coli*. *Appl. Microbiol. Biotechnol.* **2018**, *102* (3), 1251−1267.
(58) Choudhri, P.; Rani, M.; Sangwan, R. S.; Kumar, R.; Kumar, A.; Chhokar, V. *De novo* sequencing, assembly and characterisation of *Aloe vera* transcriptome and analysis of expression profiles of genes related to saponin and anthraquinone metabolism. *BMC Genomics* **2018**, *19* (1), 427.
(59) Stathopoulou, K.; Valianou, L.; Skaltsounis, A. L.; Karapanagiotis, I.; Magiatis, P. Structure elucidation and chromatographic identification of anthraquinone components of cochineal (*Dactylopius coccus*) detected in historical objects. *Anal. Chim. Acta* **2013**, *804*, 264−272.
(60) Leonard, E.; Yan, Y.; Fowler, Z. L.; Li, Z.; Lim, C. G.; Lim, K. H.; Koffas, M. A. G. Strain improvement of recombinant *Escherichia coli* for efficient production of plant flavonoids. *Mol. Pharmaceutics* **2008**, *5* (2), 257−265.
(61) Webb, B.; Sali, A. Comparative protein structure modeling using MODELLER. *Curr. Protoc. Bioinf.* **2016**, *54*, 5.6.1−5.6.37.
(62) Chaudhury, S.; Lyskov, S.; Gray, J. J. PyRosetta: a script-based interface for implementing molecular modeling algorithms using Rosetta. *Bioinformatics* **2010**, *26* (5), 689−691.
(63) Lindqvist, Y.; Koskiniemi, H.; Jansson, A.; Sandalova, T.; Schnell, R.; Liu, Z.; Mantsala, P.; Niemi, J.; Schneider, G. Structural basis for substrate recognition and specificity in aklavinone-11-hydroxylase from rhodomycin biosynthesis. *J. Mol. Biol.* **2009**, *393* (4), 966−977.
(64) He, J. B.; Zhao, P.; Hu, Z. M.; Liu, S.; Kuang, Y.; Zhang, M.; Li, B.; Yun, C. H.; Qiao, X.; Ye, M. Molecular and structural characterization of a promiscuous C-glycosyltransferase from *Trollius chinensis*. *Angew. Chem., Int. Ed.* **2019**, *58* (33), 11513−11520.
(65) Jones, K.; Hughes, J.; Hong, M.; Jia, Q.; Orndorff, S. Modulation of melanogenesis by aloesin: a competitive inhibitor of tyrosinase. *Pigm. Cell Res.* **2002**, *15* (5), 335−340.
(66) Wang, Z.; Li, X.; Yang, Z.; He, X.; Tu, J.; Zhang, T. Effects of aloesin on melanogenesis in pigmented skin equivalents. *Int. J. Cosmet. Sci.* **2008**, *30* (2), 121−130.
(67) Yagi, A.; Kabash, A.; Okamura, N.; Haraguchi, H.; Moustafa, S. M.; Khalifa, T. I. Antioxidant, free radical scavenging and anti-inflammatory effects of aloesin derivatives in *Aloe vera*. *Planta Med.* **2002**, *68* (11), 957−960.
(68) Sasaki, N.; Nishizaki, Y.; Yamada, E.; Tatsuzawa, F.; Nakatsuka, T.; Takahashi, H.; Nishihara, M. Identification of the glucosyltransferase that mediates direct flavone C-glucosylation in *Gentiana triflora*. *FEBS Lett.* **2015**, *589* (1), 182−187.
(69) Baneyx, F.; Mujacic, M. Recombinant protein folding and misfolding in *Escherichia coli*. *Nat. Biotechnol.* **2004**, *22* (11), 1399−1408.
(70) Harle, J.; Gunther, S.; Lauinger, B.; Weber, M.; Kammerer, B.; Zechel, D. L.; Luzhetskyy, A.; Bechthold, A. Rational design of an aryl-C-glycoside catalyst from a natural product O-glycosyltransferase. *Chem. Biol.* **2011**, *18* (4), 520−530.
(71) Yang, D.; Park, S. Y.; Park, Y. S.; Eun, H.; Lee, S. Y. Metabolic engineering of *Escherichia coli* for natural product biosynthesis. *Trends Biotechnol.* **2020**, *38* (7), 745−765.

https://doi.org/10.1021/jacs.0c12406
J. Am. Chem. Soc. 2021, 143, 5364−5377