INSIGHTS INTO RADICICOL BIOSYNTHESIS VIA HETEROLOGOUS SYNTHESIS OF INTERMEDIATES AND ANALOGS

Hui Zhou (周卉)1,*, Kangjian Qiao (乔康健)1,*, Zhizeng Gao (高志增)3,*, John. C. Vederas3 and Yi Tang (唐奕)1,2

From the 1Department of Chemical and Biomolecular Engineering, University of California, Los Angeles, USA, 2Department of Chemistry and Biochemistry, University of California, Los Angeles, USA and 3Department of Chemistry, University of Alberta, Canada

Running head: Radicicol Biosynthesis

Address Correspondence to: Yi Tang, Department University of California Los Angeles, Los Angeles, CA 90095 (USA), E-mail: yitang@ucla.edu, john.vederas@ualberta.ca

* Equal contribution authors.

Resorcylic acid lactones (RALs) are fungal polyketides that display diverse biological activities, with the potent Hsp90 inhibitor radicicol being an important representative member. Two fungal iterative polyketide synthases (IPKSs), Rdc5 the highly-reducing IPKS (HRPKS) and Rdc1 the non-reducing IPKS (NRPKS), are required for the biosynthesis of radicicol in Pochonia chlamydosporia. In this work, the complete reconstitution of Rdc5 and Rdc1 activities both in vitro and in Saccharomyces cerevisiae uncovered the earliest RAL intermediate of the radicicol biosynthetic pathway, (R)-monocillin II. The enzymatic synthesis of (R)-monocillin II confirmed the exquisite timing of the Rdc5 ER domain. Using precursor directed biosynthesis, the chemical modularity of the dual IPKS system was determined. Rdc1 readily accepted a SNAC mimic of the reduced pentaketide product of Rdc5 to synthesize (R)-monocillin II with four additional iterations of polyketide elongation, indicating the C2’ ketone group found in (R)-monocillin II is incorporated via the functions of Rdc1 instead of Rdc5. The involvement of the TE domain in Rdc1 in macrolactonization was confirmed through both site-directed mutagenesis and domain deletion. The Rdc1 TE domain was also shown to be tolerant of the opposite stereochemistry of the terminal hydroxyl nucleophile, demonstrated in the precursor directed synthesis of the enantiomeric (S)-monocillin II. Finally, reconstitution of the halogenase Rdc2 was demonstrated both in vivo and in vitro in the synthesis of pochonin D and a new halogenated analog Cl-DHZ.
Like other RALs studied to date, such as hypothemycin (2) (16,17) and zearalenone (4) (18,19) (Fig. 1A), the carbon scaffold of radicicol is synthesized by the collaborative functions of two type I iterative polyketide synthases (IPKSs). IPKSs are megasynthases in which linearly juxtaposed catalytic domains function in an iterative and a highly programmed fashion (20). Genetic knockout experiments with two different radicicol producing fungi, Pochonia chlamydospora and Chaetomium chiversii (16,21), confirmed that a highly reducing PKS (HRPKS) and a non-reducing PKS (NRPKS) are required for the biosynthesis of radicicol. The domain structures and putative functions of the two P. chlamydospora PKSs are shown in Fig. 1B. The HRPKS Rdc5 contains the following domains: ketosynthase (KS) that performs the decarboxylative condensation; malonyl-CoA:ACP transacylase (MAT) that selects the building block malonyl-CoA; and acyl carrier protein (ACP) that serves as the tether of the growing polyketide via its phosphopantetheinyl arm. It also has the complete ensemble of β-keto reductive domains, which include ketoreductase (KR); dehydratase (DH); and enoyl reductase (ER). A core domain is located between DH and ER domains, which is proposed to be composed of a pseudo methyltransferase (MT) and a structural KR domain (16,20). Via iterative condensation and different combinations of β-keto reduction, Rdc5 is proposed to synthesize the reduced portion of the radicicol scaffold. In place of the reductive step, the NRPKS Rdc1 contains a N-terminal starter-unit:ACP transacylase (SAT) (22) that transfers the completed reduced polyketide from Rdc5 to Rdc1; a product template (PT) domain (23) putatively involved in the cyclization of the completed nonaketide to yield the resorcylate core; and a C-terminal thioesterase (TE) domain (24) that performs the macrolactonization to release the RAL product (25) (Fig. 1B).

Although the assignment of Rdc5 and Rdc1 to synthesize the two chemically distinct portions of radicicol parallels the “bi-module” strategy employed by the hypothemycin (2) (16) and zearalenone (4) (18,19) biosynthetic pathways, the proposed rdc pathway shown in Fig. 1B contains several unique features not present in other RAL pathways. Four of these intriguing mechanisms, which are highlighted in Fig. 1B in shaded boxes, in turn lead to unique structural features of radicicol:

I) The C2′-C6′ dienone. The three-dimensional structure of radicicol is largely determined by the trans-cis diene moiety and the adjacent C7′-C8′ epoxide (26). Since the epoxide is probably produced by oxidation of a double bond, it is possible that the reduced portion of the radicicol frame may be derived from a trienone. This then would suggest that the ER domain in Rdc5 may not be active during the iterative process and may instead serve as a C5′-C6′ isomerase. Alternatively, the ER may function once at the tetraketide stage (Fig. 1B, Box I) to afford the C5′-C6′ single bond and lead to the synthesis of (R)-monocillin II (6) as a possible intermediate to radicicol, which is then desaturated during post-PKS modifications to afford the trienone.

II) The C2′ ketone. This is a unique structural feature of radicicol compared to other RALs, which has been explored as a reactive handle for oxime formation (11,27). Two possible pathways leading to the installation of the C2′ ketone of radicicol have been proposed (Fig. 1B, Box II). In the first case, the ketone is synthesized by Rdc5 in the form of a β-keto group in a hexaketide intermediate, which is elongated by Rdc1 by three additional ketides (6+3 combination). This distribution of ketides between the two IPKSs is identical to that of some other RALs (17,19,28). Alternatively, Rdc5 could synthesize a pentaketide, which is transferred by the SAT domain to Rdc1 followed by chain extension of four additional ketides (5+4 combination) to complete the nonaketide. The C2′ ketone would be synthesized by Rdc1 in this case.

III) The C10′ R-OH. The R-stereochemistry of the ester-forming, terminal hydroxyl group in radicicol differs from the S configuration present in hypothemycin (2) or zearalenone (4). While the difference in stereochemistry between these structurally related RALs is most likely the result of the stereospecific β-keto reduction by the KR domain, it remains unknown if the macrolactonizing TE domain of Rdc1 has evolved to be stereoselective towards the terminal R-hydroxyl group (Fig. 1B, Box III).

IV) C6′-chlorination. One of the key tailoring steps in radicicol biosynthesis is chlorination of the resorcylate core by a putative chlorinase Rdc2 (Fig. 1B, Box IV). Knockout of the Rdc2
homolog (RadH) in C. chiversii resulted in the accumulation of the non-halogenated version of radicicol, monocillin I (21). Rdc2 shares sequence homology to a variety of FAD-dependent halogenases found in bacteria (29), fungi (21,30) and protozoa (31). Structure-activity relationship studies have revealed that the chloride contributes to the potent activity of radicicol (32). The presence of halide may attenuate the electron density on the aromatic ring and stabilize the compound. The chlorination step has been proposed to take place immediately following the carbon scaffold synthesis; however the function of Rdc2 has not been verified. Understanding the substrate specificity of Rdc2 may therefore lead to the synthesis of other chlorinated RALs. During the review of this manuscript, an independent study verified the role of Rdc2 as the C6-chlorinase, which is also discussed here (33).

Comprehensive understanding of these highlighted steps of the rdc pathway is therefore an important goal towards the synthesis of radicicol analogs. In this work, we express Rdc5 and Rdc1 using an engineered Saccharomyces cerevisiae strain and completely reconstitute the activities of the two IPKSs to demonstrate that (R)-monocillin II (6) is the initial IPKS product. Using a combination of heterologous pathway reconstitution, precursor-directed biosynthesis and domain dissection, we provide insights into the unique functions of enzymes in the rdc pathway.

**EXPERIMENTAL PROCEDURES**

Strains and general techniques for DNA manipulation. *Pochonia chlamydosporia* ATCC 16683 (previously *Verticillum chlamydosporium* var. *catenulatum*) was obtained from the American Type Culture Collection. *P. chlamydosporia* genomic DNA was prepared using the ZYMO (Orange, CA) ZR fungal/bacterial DNA kit according to the supplied protocols. Total RNA was isolated from *P. chlamydosporia* using ZYMO ZR Fungal/Bacterial RNA MiniPrep™ kit. The ImProm-II™ Reverse Transcription System kit was used for reverse transcription-PCR (RT-PCR). The gene-specific primers are listed in Table S1. *Escherichia coli* (E. coli) XL1-Blue (Stratagene) and *E. coli* TOP10 (Invitrogen) were used for cloning. DNA restriction enzymes were used as recommended by the manufacturer (New England Biolabs). PCR was performed using Platinum® Pfx DNA polymerase (Invitrogen). Sequences of PCR products were confirmed by DNA sequencing (Laragen, CA). *E. coli* BL21(DE3) (Novagen) was used for protein expression. *Saccharomyces cerevisiae* strain BJ5464-NpgA (MATα ura3-52 his3-A200 leu2-A1 trp1 pep4::HIS3 prb1 Δ1.6R can1 GAL) was used as expression host (34,35).

**Synthetic procedures.** The synthetic strategies and procedures for (2E,6E,9R)-9-hydroxydeca-2,6-dienoyl-N-acetyl-cysteamine (7) and (2E,6E,9E)-9-hydroxydeca-2,6-dienoyl-N-acetyl-cysteamine (9) and standards are described in detail in supplemental data.

**Spectroscopic analyses.** Nuclear magnetic resonance (NMR) spectra were obtained on Varian Inova 500 MHz, 500 MHz (cold probe), 600 MHz, 700 MHz (cryoprobe) instruments and a Bruker 500 MHz spectrometer. 1H NMR chemical shifts are reported in parts per million (ppm) using the residual proton resonance of solvents as reference: CDCl₃ δ 7.26, CD₂OD δ 3.30 and CD₃COCD₃ δ 2.04. 13C NMR chemical shifts are reported relative to CDCl₃ δ 77.0, CD₂OD δ 49.0 and CD₃COCD₃ δ 29.9. LC-MS was conducted with a Shimadzu 2010 EV liquid chromatography mass spectrometer by using both positive and negative electrospray ionization, and a Phenomenex Luna 5 μm 2.0 × 100 mm C18 reverse-phase column. Samples were separated on a linear gradient of 5 to 95% CH₃CN (v/v) in H₂O supplemented with 0.05% (v/v) formic acid at a flow rate of 0.1 mL/min. Chiral HPLC analysis was performed using a Lux 3 μm Cellulose-1 column (4.60 × 250 mm, Phenomenex) with an isocratic condition of 25% iso-propanol in n-hexane (v/v) at a flow rate of 0.75 mL/min.

**Construction of plasmids.** Genomic DNA from *P. chlamydosporia* was used as the template for PCR amplifications. The two exons of rdc5 gene were amplified and ligated into the 2μ-based yeast-*E. coli* shuttle vector YEpADH2p with *URA3* marker to give pKJ61. The gene encoding Rdc1 was amplified and ligated into YEpADH2p with *TRP1* marker to give pKJ91. An engineered FLAG-tag expression vector was constructed by introducing DNA sequence 5’-CATATGGCTAGCGATTATAAGGATGATGA-3’ into YEpADH2p with *URA3* marker between *Nde*I and *Spe*I sites. Then,
both rdc5 and rdc1 genes were amplified using primers pairs P7/P8 and P9/P10. The genes were digested by SpeI and Swal, and ligated into the FLAG-tag expression vector to yield pZH223 and pZH232. Rdc1ΔTE gene was amplified by primer pair P1 and P11. The PCR product was ligated into YEplac139p with TRP1 marker to give pZH228. The Ser1889 to Ala mutation was introduced into rdc2 gene from RT-PCR showed variation compared to the predicted rdc2 CDS (HQ149729) (supplemental Fig. S4). The intron-free rdc2 was ligated into pET24a vector to give E. coli expression plasmid pZH200 or into YEplac139p with LEU2 marker to yield pZH208. All the plasmids in this work are listed in supplemental data Table S2.

Protein purification and in vitro assays. C-terminal hexahistidine tagged Rdc1ΔTE and Rdc1_S1889A were expressed and purified from S. cerevisiae strain BJ5464-NpgA harboring plasmid pZH228 and pZH252, respectively. C-terminal hexahistidine tagged Rdc2 was expressed and purified from E. coli BL21(DE3) strain harboring pZH200 (see supplemental data for details). N-terminal FLAG tagged Rdc1 and Rdc5 were expressed and purified from S. cerevisiae strain BJ5464-NpgA harboring plasmid pZH232 and pZH223, respectively. For 1 L of yeast culture, the cells were grown at 25°C in YPD media with 1% dextrose for 72 hours. The cells were harvested by centrifugation (3500 rpm, 15 minutes, 4°C), resuspended in 20 mL lysis buffer (50 mM NaH2PO4, 0.15 M NaCl, 10 mM imidazole, pH=8.0) and lysed by sonication on ice. Cellular debris was removed by centrifugation (15000 g, 1 hour, 4°C). The supernatant was loaded onto a gravity flow column containing 3 mL anti-FLAG M2 affinity resin and washed with 10 column volumes of TBS Buffer (150 mM Tris-HCl, 50 mM NaCl, pH=7.4). Purified proteins were eluted with 5 column volumes of TBS Buffer with 100 mg/mL FLAG peptide. Purified proteins were concentrated and buffer exchanged into Buffer A (50 mM Tris-HCl, 50 mM NaCl, pH=8.0) +10% glycerol, concentrated, aliquoted and flash frozen.

Protein concentrations were determined using the Bradford dye-binding assay (Bio-Rad).

For in vitro assays, 10 µM of Rdc1 and 10 µM of Rdc5 were incubated with 2 mM of NADPH and malonyl-CoA. In the assays for Rdc1, the chemically synthesized starter units were added to 2 mM final concentration. For halogenation assays with Rdc2, 1 mM RAL substrate was incubated with 50 µM Rdc2, 10 µM SsuE, 100 µM FAD, 2 mM NADPH and 50 mM NaCl. SsuE was used as a FAD reductase, which is cloned and expressed as reported (36). All the reactions were quenched and extracted twice with 99% ethyl acetate (EtOAc)/1% acetic acid (AcOH). The resultant organic extracts were evaporated to dryness, redissolved in methanol, and then analyzed by LC-MS.

Isolation of (R)-monocillin II (6), isocoumarin (8) and (S)-monocillin II (10). To purify 6 for structural analysis, S. cerevisiae harboring pKJ61 and pKJ91 plasmids was cultured in YPD liquid media (4 L) for 4 days. The pH of culture supernatant was adjusted to 5.0 and then extracted three times with equal volume of EtOAc. The resultant organic extracts were combined and evaporated to dryness, redissolved in methanol, and purified by reverse-phase HPLC (XTerra Prep MS C18 5 µm, 19 mm × 50 mm) on a linear gradient of 5 to 95% CH3CN (v/v) over 15 min and 95% CH3CN (v/v) further for 15 min in H2O supplemented with 0.1% (v/v) trifluoroacetic acid at a flow rate of 2.5 mL/min. The eluent was extracted with EtOAc, and dried in vacuo to give pure solid 6 (approximate yield of 15 mg/L). To purify 8, S. cerevisiae/pKJ61+pZH228 was cultured in YPD (20 L) media for 4 days followed by the same purification steps as above. To produce 10, S. cerevisiae strain harboring pKJ91 were cultured in YPD media (400 mL) for 2 days. The cell culture was supplemented with 100 mg/L (2E,6E,9S)-9-hydroxydeca-2,6-dienoyl-N-acetyl cysteamine (9) and continued growing for another day. Compound 10 was purified as described above for 6 with an approximate yield of 1.3 mg/L.

Isolation of pochonin D (11) and 6-chloro, 7', 8'-dehydrozearalenol (Cl-DHZ) (14). S. cerevisiae/pKJ61+pKJ91+pZH208 was cultured in YPD media (4 L) for 4 days and 11 was purified to a final yield of 14.3 mg/L. S. cerevisiae strain harboring pZH78+pZH74+pZH208 plasmids were
cultured in YPD media (4 L) for 7 days and 14 was purified to a final yield of 9 mg/L.

RESULTS

Heterologous reconstitution of the enzymatic activities of Rdc5 and Rdc1. Although the rdc gene cluster has been genetically verified in two different fungal hosts, the biochemical characterization of the key IPKSs has not been performed and the earliest RAL intermediate has not been conclusively identified. As a starting point to study the unique features of the rdc pathway, we aimed to reconstitute the functions of Rdc5 and Rdc1 in vitro and in the heterologous host S. cerevisiae. To do so, yeast 2μ expression plasmids encoding Rdc1 and Rdc5 under the ADH2 promoter were separately transformed into S. cerevisiae BJ5464-NpgA. The use of this vacuolar protease-deficient yeast strain with genome-integrated NpgA (37) enables heterologous expression of intact and phosphopantetheinylated PKSs (34). Both megasynthases were solubly expressed and were each purified to single-band purity by using anti-FLAG affinity chromatography (supplemental Fig. S1).

When the purified Rdc5 and Rdc1 were incubated with 2 mM malonyl-CoA and NADPH, a major compound emerged with a mass of 316, which is identical to that of 6. The UV absorbance of the compound is characteristic of resorcylate chromophore with λmax at 217, 260 and 301 nm. In order to determine the structure of this RAL product, two plasmids harboring rdc1 and rdc5 with different auxotrophic selection markers were co-transformed into S. cerevisiae BJ5464-NpgA. After three days of culturing in YPD media, both the culture broth and cell pellet were extracted with organic solvent and analyzed with LC-MS. Compared to BJ5464-NpgA expressing either Rdc1 or Rdc5 individually, the co-expression strain synthesized the identical m/z 316 compound as a predominant product at a titer of 15 mg/L (Fig. 2). The new compound was then purified and analyzed by extensive NMR characterization. The 13C spectrum showed 18 carbon signals, indicative of a nonaketide backbone. One carbon signal at δ 196.5 ppm suggested the existence of an aliphatic ketone. Combining all spectroscopic information, we assigned the structure of this RAL as that of 6 (supplemental Table S3), which has been isolated from different radicicol producing strains (21,38,39). To confirm the structure of 6 and especially to verify the R stereochemistry of the lactone at C1, an authentic standard of 6 was chemically synthesized (supplemental Table S4). The UV pattern, retention time on chiral column, and all NMR data of 6 are all identical to that of the standard.

The enzymatic synthesis of 6 by purified Rdc1 and Rdc5 conclusively indicates 6 is the first RAL compound in the rdc pathway, and is the product directly offloaded from Rdc1 as shown in Fig. 1B. Interestingly, a fully reduced ketide unit at C5’-C6’ is found in 6 instead of the possible trienone structure, which indicates that the ER domain in Rdc5 is indeed functional and can completely reduce the β-keto of the tetraketide to a methylene following the actions of KR and DH. However, the ER domain is not active during the other reductive tailoring iterations, thereby giving rise to the enone at C2’-C4’ and the trans double bond at C7’-C8’ that eventually becomes epoxidized. The additional C5’-C6’ cis double bond present in radicicol must therefore be introduced by post-PKS processing. This is consistent with the genetic studies with C. chiversii, which suggested that C7’-C8’ epoxidation and C5’ hydroxylation may be catalyzed by a single P450 Rdc4 (RadP), although this has not been biochemically confirmed (21).

Polyketide chain length control of Rdc5 and Rdc1. We next probed whether the starter unit of the NRPKS Rdc1 is a β-keto hexaketide or a pentaketide as shown in Fig. 1B. This would reveal the “distribution of labor” in the synthesis of 6 by the two rdc IPKSs. To verify the feasibility of the “5+4” pathway, we tested whether chemically synthesized N-acetyl-cysteamine thioester (SNAC) of (2E, 6E, 9R)-9-hydroxydeca-2,6-dienoic acid (7) (Fig. 3A and supplemental data) can mimic the proposed pentaketide product of Rdc5 as shown in Fig. 1B. We have previously shown that KS domains of certain NRPKSs can accept specific acyl-SNAC substrates as the starter unit to prime the functions of the remaining steps (17,28,40). We therefore reasoned that if the pentaketide is indeed the starter unit for Rdc1, the RAL product 6 should be observed in the reaction mixture. However, if Rdc1 is only programmed to perform three additional iterations as in the “6+3”
biosynthesis of...further demonstrate the precursor-directed absorbance spectrum when purified from the yeast culture, the identity was seen in the organic extracts (Fig. 3A, trace iii). After one additional day of culturing, the production of 

6 at a concentration of 100 mg/L. After one...intermediates could be found. However, if other RAL compounds or biosynthetic intermediates are responsible for the final macrolactonization using the secondary hydroxyl nucleophile can affect the recognition and macrocyclization of the resorcylate ester by the TE domain (Fig. 1B, Box III). To probe the stereoselectivity of the TE domain, the active site Ser1889 of the catalytic triad identified from sequence alignment with PKS13TE was mutated to Ala. This Rdc1_S1889A mutant was solubly expressed and purified from BJ5464-NpgA (supplemental Fig. S1). When Rdc1_S1889A was incubated with thioester 7 and malonyl-CoA, we could no longer detect the synthesis of 6. Instead, a new compound 8 with the same mass of 6 (m/z of 316) was observed (Fig. 4 and supplemental Fig. S2). Compound 8 was also present in the culture extract of BJ5464-NpgA co-expressing Rdc1_S1889A and Rdc5. However, the titer of 8 is >100 fold less than that of 6 synthesized in Fig. 2. Since the starter unit 7 or Rdc5 is required for the synthesis of 8, we inferred that the pentaketide acyl chain is part of 8. The UV spectrum of 8 does not display the characteristic peaks observed from resorcylate-containing compounds, but is more complex and similar to isocoumarin compounds SMA76b and SMA76c produced by the NRPKS PKS4 from Gibberella fujikuroi (40). To solve the structure of 8, we scaled up the yeast culture and purified 8 for NMR characterization (supplemental Table S5). The NMR data is consistent with that of a pentaketide-primed isocoumarin 8 shown in Fig. 4.

To exclude any possible involvement of Rdc1 TE in the formation of 8, a truncated Rdc1 without the TE domain (Rdc1ΔTE) was constructed. Rdc1ΔTE was similarly expressed and purified from yeast (supplemental Fig. S1). Both in vitro assay of Rdc1ΔTE supplemented with starter unit 7 and in vivo product of Rdc1ΔTE co-expressed with Rdc5 showed that 8 was synthesized (data not shown). Our results confirm the role of the Rdc1 TE domain in catalyzing the macrolactonization reaction during the biosynthesis of 6. When the TE activity is compromised, 8 is released as a shunt product of Rdc1 through enolization of the C2’ ketone and nucleophilic attack on the thioester carbonyl to form the benzopyrone (Fig. 4). Interestingly, we did not find any pentaketide-primed resorcyclic acids in the reaction mixture, suggesting releasing by hydrolysis of the thioester linkage is limited for Rdc1. The mechanism shown in Fig. 4 may also account for the biosynthesis of structurally similar isocoumarin co-metabolites from other RAL producing fungi strains, such as paraphaeosphaerins and chaetochiversins (41), which can accumulate as a result of derailment of NRPKS TE functions.

Probing the stereoselectivity of the TE domain. Having confirmed the involvement of the Rdc1 TE domain in macrolactone formation, we then examined whether the stereochemistry of the C10’ hydroxyl nucleophile can affect the recognition and macrocyclization of the resorcylate ester by the TE domain (Fig. 1B, Box III). To probe the
stereochemical requirement of TE domain, we chemically synthesized (2E, 6E, 9S)-9-hydroxydeca-2,6-dienoyl-S-NAC (9), which is the enantiomer of the usual starter unit 7. After supplementing 9 into the Rdc1-expressing BJ5464-NpgA culture, the organic extract contained a compound that is indistinguishable from 6 based on retention time, UV absorbance and mass spectrum on reverse phase C18 HPLC column (Fig. 3A, trace ii). This compound was purified to homogeneity and all NMR spectra perfectly matched to that of 6 (supplemental Table S6). However, the compound can be clearly separated from 6 when co-injected onto a chiral HPLC column as shown in Fig. 3B, suggesting that it is 10, which contains an S-stereocenter at C10’. Indeed, the circular dichroism spectra of 6 and 10 were mirror images of each other (supplemental Fig. S3), which is expected as 6 and 10 are enantiomers.

The titer of 10 from the in vivo precursor feeding study is nearly the same as that of 6 when supplemented with 9. Furthermore, in vitro experiments also gave similar levels of the enantiomers. Therefore, the Rdc1 TE domain appears insensitive towards the stereochemical configuration of the terminal nucleophile in the macrocyclization reaction. It is possible that the linear portion of the unlaconotized molecule is flexible enough to allow the hydroxyl in either configuration to be deprotonated by the histidine that serves as the general base, which can then lead to formation of the S or R configured RAL compounds with equal efficiency. This observation thus expands the already broad substrate specificities of the RAL TE domains characterized to date and further hints at a structurally very loose cyclization chamber for this unique class of TE domains. Lastly, the biosynthesis of 10 also reveals the KS and PT domains of the Rdc1 are not specific towards the chirality of the hydroxyl group.

In vivo reconstitution of pochonin D synthesis. Establishment of 6 as the earliest RAL compound in the radicicol biosynthetic pathway represents the starting point to investigate the downstream post-PKS tailoring enzymes. Pochonin D (11), which is the chlorinated version of 6, appears to be biosynthetically accessible through the action of the FAD-dependent halogenase Rdc2 (Fig. 1B, Box IV). Compound 6 displays good affinity towards Hsp90 and its C2’ oxime derivatives are considerably more stable than radicicol. Therefore, reconstituting the biosynthesis of 11 in S. cerevisiae may provide a convenient route of affording this potential drug lead.

The cDNA of Rdc2 was obtained by RT-PCR and inserted into a yeast expression plasmid that contains a LEU2 selection marker, as well as into an E. coli pET24a expression vector. When the BJ5464-NpgA host is transformed with three expression plasmids separately encoding Rdc1, Rdc2 and Rdc5, we observed the emergence of a major RAL product from the third day culture extract (Fig. 5A, trace i). While a small amount of 6 remained, the titer was less than 5% of the overall RAL pool at the end of four days of continued culturing. The mass spectrum of the new compound matched precisely to that of the estimated isotope mass pattern of 11. To verify the identity of the compound and the regioselectivity of Rdc2, we scaled up the yeast culture and purified the compound at a final titer of 14.3 mg/L. One and two dimensional NMR was performed to verify that the new compound is indeed 11 (supplemental Table S7). The 1H NMR displayed a single aromatic proton signal at δ 6.62 ppm instead of the typical spin system observed for the resorcylate core of 6, indicative of loss of one proton. Further comparison of the 2D NMR data to 6 confirmed that no proton is linked to C6 in the HSQC spectrum; while correlation between C1’ proton and C6 remained in HMBC spectrum. Therefore, co-expression of the three rdc genes in yeast led to robust synthesis of 11 and thereby confirmed the role of Rdc2 as a C6-specific chlorinase. This high in vivo conversion of 6 to 11 also demonstrated the robust activity of the heterologously expressed Rdc2.

We also investigated the in vitro conversion of 6 to 11 using purified Rdc2 from E. coli and the E. coli SsuE as a flavin reductase accessory enzyme. We were able to observe near complete conversion of 6 to 11 in the presence of NADPH and NaCl in four hours (Fig. 5A, trace ii). By LC-MS analysis, we also detected a compound 12 that appeared to be the dichlorinated 6 from both in vitro assay and in vivo extract at a low yield. The structure of 12, however, has not been confirmed by NMR analysis.

Engineered biosynthesis of a chlorinated RAL analog. Since RALs share the similar structural
scaffold and the core resorcylicate, we tested whether Rdc2 may display substrate tolerance towards other RAL compounds. Dehydrozearalenol (13) (Fig. 5), the earliest RAL product from the hypothemycin (2) pathway (17), serves as an ideal candidate for the assay. As shown in Fig. 5B (trace ii), the in vitro mixture of 13, Rdc2, SsuE and NADPH led to the emergence of a new peak of which the isotopic mass pattern is indicative of a chlorinated compound. To characterize 14, we introduced the yeast Rdc2 expression plasmid into the DHZ-producing strain, which co-expresses the HRPKS Hpm8 and the NRPKS Hpm3. Following organic extraction of the culture, we also observed the presence of 14 that accumulated to approximately the same level as 13 after seven days of culturing. Compound 14 was purified from the yeast culture to a final titer of 9 mg/L and was confirmed by NMR to be 6-chloro, 7',8'-dehydrozearalenol (Cl-DHZ, 14) (supplemental Table S8).

DISCUSSION

In this work, we fully reconstituted the enzymatic activities of Rdc5 and Rdc1 from radicicol biosynthetic pathway with the production of (R)-monocillin II (6) as an early intermediate. Following the same biosynthetic model as that of other RAL compounds (16-19), the assembly of 6 from malonyl-CoA requires one HRPKS (Rdc5) and one NRPKS (Rdc1). The tandem activities of the two IPKSs represent the fungal version of the well-studied modular PKS assembly lines found in bacteria. However, the fungal strategy is significantly more compact and efficient, utilizing only two “modules” and a total of 11 active sites to complete 37 steps of catalytic events. The workload is divided based on the capabilities of each IPKS: the upstream module Rdc5 is solely responsible for assembling the reducing portion of 6, while the downstream module Rdc1 is responsible for the aromatic portion and cyclization of 6. The iterative nature of the fungal IPKSs may initially appear to make them less amendable to combinatorial biosynthesis approaches that have been successfully implemented on modular type I PKSs. However, the relaxed substrate specificity of the NRPKS enzymes and their ability to accept smaller thioesters as starter units should permit formation of many interesting compounds. This highly efficient manner of synthesizing and tailoring the polyketide backbone is evolutionarily more advantageous, especially for the metabolically more complex eukaryotic organisms. However, the IPKSs, especially the HRPKS, require considerably more complex programming rules as discussed below. The use of a “bi-modular” strategy to synthesize polyketides containing structurally distinct subunits is also found in other fungal biosynthetic pathways. For example in the aflatoxin biosynthetic pathway, a FAS-like module synthesizes the hexanoyl starter unit, which is then transferred to the NRPKS PksA to synthesize the anthraquinone portion of the precursor norsolorinic acid (42). Similarly, in the recently uncovered biosynthetic pathway of asperfuranone (43), an HRPKS and a NRPKS synthesize the linear tetraketide and the β-resorcylic aldehyde subunits, respectively.

Using the pentaketide SNAC 7 and precursor directed biosynthesis, we showed that the “5+4 distribution of effort” between Rdc5 and Rdc1 is a highly plausible mechanism of assembling 6. This programming rule is apparently different from the “6+3” split employed by the zearalenone (PKS4 and PKS13) and hypothemycin (Hpm8 and Hpm3) pathways (17,19,28). The “5+4” mechanism maintains the aforementioned chemical modularity of the two IPKSs. Each β-keto position synthesized by Rdc5 is reduced, while those synthesized by Rdc1 remain completely unreduced. With this split, the KR of Rdc5 reduces the β-carbon in a nondiscriminatory fashion after each chain extension step, and therefore requires no additional programming rule to distinguish a β-keto hexaketide. Indeed, this feature of KR can be found in all RAL HRPKSs. We previously suggested that PKS4 KR is inactive at the tetraketide step in the synthesis of zearalenone scaffold, thereby giving rise to the C6’ ketone. However, further studies with the highly parallel hypothemycin pathway (17) suggest that the KR is indeed active at this step and affords the secondary alcohol at C6’ of compound 13 that is eventually oxidized to the ketone by an alcohol oxidase encoded in both pathways (16,18,19). To maintain this orthogonal chemical modularity between the HRPKS and NRPKS in the rdc pathway, the substrate specificity of the Rdc1 SAT domain
must therefore be different from that of Hpm3 and PKS13. The SAT domain is the key enzyme that facilitates acyl transfer of the completed, reduced precursor from the ACP domain of the HRPKS to the ACP domain of the NRPKS, thereby ensuring the correct substrate is passed between the two IPKSs (17,22). The Rdc1 SAT domain therefore must display specificity for the pentaketide acyl chain instead of the hexaketide chain. Following chain transfer, the KS domain of Rdc1 maintains strict chain length control to synthesize the complete nonaketide.

While the KR domain appears to function during each cycle, the DH and the ER domains are programmed to be different, thereby leading to the different reduced chains among zearalenone (4), hypothemycin (2) and radicicol (1). All three DH domains are programmed to be inactive towards the β-hydroxybutyryl diketide intermediate, which is different from LNKS and LDKS in the lovastatin pathway (44,45). While the Rdc5 DH does not discriminate against the remaining β-hydroxyl intermediates, the PKS4 and Hpm8 DH domains do not dehydrate the β-hydroxyl position of the tetraketide, which is a required process to install the C6′ hydroxyl in zearalenol and 13. Similar analysis of the substrate specificities of the ER domains also reveals subtle, yet important differences between the three RAL HRPKSs. For example, our work here demonstrates that Rdc5 ER can only reduce the double bond of the tetraketide during biosynthesis of 6. In contrast, the Hpm8 ER only functions on a α,β-unsaturated pentaketide whereas the PKS4 ER acts on both the diketide and pentaketide. These differences in substrate specificity of a single DH/ER domain towards different intermediates of different chain lengths, as well as the differences between homologous DH/ER domains towards substrates of the same chain lengths are faithfully maintained to produce the different RAL compounds. Understanding the structural basis of these catalytic differences will be a key to decode HRPKS programming rules and allow precise prediction of product structures from protein sequences.

Although the three RAL KR domains can reduce β-keto functionality in all intermediates, the programming rules dictating stereochemical outcomes are different. In contrast to the S-

stereochemistry found in 13 and zearalenol, the first β-keto reduction of acetoacetyl diketide by Rdc5 takes place to give an R-configuration. Compared to the KR domains of bacterial type I modular PKSs, of which the biochemical and structural basis of stereochemical control has been well established (46-49), the stereospecificity of KR in fungal PKSs remains enigmatic. For the type I modular KRs, it has been proposed that β-keto polyketides gain access to the catalytic site via different paths, which are oriented and controlled by characteristic residues in the active site. From structural and mutational analyses, B-type KR contains the LDD tripeptide fingerprint and reduces with R stereochemistry (49,50). Sequence alignment between Rdc5 KR with B-type KR revealed a similarly placed patch of LRD in Rdc5 KR instead of the LDD sequence. However, this LRD motif is also conserved in Hpm8 and PKS4, both of which reduce the same acetoacetyl diketide with S stereochemistry. Therefore, the classification of type I modular KR stereochemistry based on sequence analysis does not apply to fungal IPKS KRs. Another degree of complexity associated with HRPKS KR domains may be their abilities to reduce substrates of different chain lengths with differing stereochemistry. This is not apparent in our current study with Rdc5 since the other four β-hydroxyl products of the KR are all dehydrated by the DH domain, thereby masking the stereochemistry of these keto reductions. However, we previously observed that during the synthesis of 13 by Hpm8 and Hpm3, the two nondehydrated hydroxyl groups at C6′ and C10′ were apparently the results of β-keto reduction by Hpm8 KR with opposite stereochemistry. Therefore, it is possible that the masked reduction steps in Rdc5 may also proceed via different stereochemistries, which may result from orientation of acyl chains of different sizes divergently in the cavity. We are currently in the process of establishing the correlation between chain length and stereochemical control of the various RAL KR domains using synthetic β-keto substrate mimics. It may also be possible that the differences in β-hydroxyl stereochemistry may influence the selectivity of the subsequent DH-catalyzed dehydrations as discussed previously (51,52).
Regardless of the stereochemistry of the terminal hydroxyl nucleophile, the Rdc1 TE domain can complete the macrocyclization to yield either (R)- or (S)-monocillin II. Substrate tolerance was also observed in TEs from type I modular PKS, such as that of DEBS. Engineered bi-modules fused with cognate TE domain can produce different 6-membered triketide lactones with inversion in hydroxyl group stereochemistry (53,54). However, the ability to form a same-sized macro lactone using nucleophiles of opposite stereochemistry has not been demonstrated for a bacterial TE domain. Our results here, combined with our previous studies with RAL TEs from Hpm3 and PKS13 (17,25,28), further illustrate the remarkable substrate tolerance of this class of fungal macro lactonizing TE domains.

In addition to reconstituting the combined 37 individual reactions catalyzed by Rdc5 and Rdc1, we also confirmed the function of Rdc2 as a flavin-dependent halogenase in the biosynthesis of 11, in agreement with the previous genetic study (21) and a recently published, independent study (33). Although the catalytic mechanisms of halogenases in this family have been well-studied for homologs in bacteria, this represents the first heterologous reconstitution of a fungal aromatic chlorinase. It has been shown that certain aromatic chlorinases generate hypochlorous acid through the flavin cofactor and molecular oxygen (29). The nucleophilic hypochlorous acid is then covalently linked to a lysine residue in the active pocket to form a reactive chloramidine, which can attack the electron rich aromatic substrate (55,56). Sequence alignment of Rdc2 with selected bacterial halogenases revealed the presence of a likely active site containing the lysine residue K74. Interestingly, Rdc2 appears to be a very efficient chlorinase from in vitro analysis, which is in contrast to most bacterial chlorinases reconstituted to date (29). Rdc2 also functioned efficiently in the heterologous host as evidenced by the nearly complete conversion of 6 to 11. Furthermore, Rdc2 was readily combined with heterologous RAL IPKSs to generate a different chlorinated RAL 14, demonstrating its potential utility in combinatorial biosynthesis.

This work on reconstitution of the rdc enzymes further demonstrates the usefulness of the BJ5464-NpgA heterologous host. In addition to being a host capable of expressing the fungal megasynthases at preparative quantities, it can be considered as a versatile system for engineered biosynthesis of fungal polyketides. We first demonstrated that precursor directed biosynthesis can be performed with the host through the synthesis of (R)-monocillin II (6) and its enantiomer 10. The large acyl-SNAC substrates 7 and 9 were able to penetrate the yeast cell membrane and were available intracellularly to prime Rdc1. This can be highly useful in the probing of fungal IPKSs programming rules, as well as generation of polyketide analogs. The syntheses of 11 and 14 demonstrate that the yeast host is efficient in coexpression of multiple fungal biosynthetic enzymes through the multi-plasmid approach. This feature can be particularly useful in the heterologous reconstitution of a multistep fungal biosynthetic pathway and the study of fungal-specific tailoring enzymes.

REFERENCES


**FOOTNOTES**
This work was supported in part by NIH grants 1R01GM085128 and 1R01GM092217; and a David and Lucile Packard Fellowship in Science and Engineering to Y.T. Research in the lab of J. C. V. is supported by the Natural Sciences and Engineering Research Council of Canada (NSERC) and the Canada Research Chair in Bioorganic and Medicinal Chemistry.
FIGURE LEGENDS

Figure 1. Resorcylic acid lactones (RALs). (A) The structures of RAL natural products; (B) The proposed biosynthetic pathway of (R)-monocillin II (6) by Rdc5 (HRPKS) and Rdc1 (NRPKS) as well as pochonin D (11). 6 is an early intermediate in the biosynthetic pathway of the natural product radicicol (1). The four featured biosynthetic steps are highlighted in shaded box.

Figure 2. In vivo reconstitution of biosynthesis of (R)-monocillin II (6). LC-MS profiles of organic extract from (i) S. cerevisiae strain BJ5464-NpgA/pKJ91 expressing Rdc1; (ii) S. cerevisiae strain BJ5464-NpgA/pKJ61 expressing Rdc5; and (iii) S. cerevisiae strain BJ5464-NpgA/pKJ91+pKJ61 co-expressing Rdc5 and Rdc1. All traces are monitored at 300 nm.

Figure 3. Precursor directed feeding using SNAC substrates. (A) LC-MS profiles of organic extracts from S. cerevisiae strain BJ5464-NpgA/pKJ91 expressing Rdc1 and i) without substrate feeding; ii) supplementation with pentaketide-SNAC 9; iii) supplementation with pentaketide-SNAC 7. (B) Chiral HPLC separation of the enantiomers 6 and 10 on a Lux 3 µm Cellulose-1 column (Phenomenex): i) purified 10; ii) purified 6; iii) co-injection of a mixture containing purified 10 and 6. All traces are monitored at 300 nm.

Figure 4. Proposed biosynthetic pathway of the isocoumarin 8 upon inactivation or deletion of the Rdc1 TE domain.

Figure 5. Biosynthesis of chlorinated RAL compounds using the halogenase Rdc2. (A) LC-MS profiles for production of 11 from: i) BJ5464-NpgA expressing Rdc5, Rdc1 and Rdc2; ii) in vitro assay of 50 µM Rdc2 incubated with 1 mM 6, 100 µM FAD, 2 mM NADPH, 50 mM NaCl and 10 µM SsuE. Inset in trace i shows observed [M-H] - MS spectrum of 11, and inset in trace ii shows the observed [M-H] MS spectrum of 12; (B) LC-MS profiles for production of 14: i) BJ5464-NpgA expressing Hpm8, Hpm3 and Rdc2, ii) in vitro assay of 50 µM Rdc2 incubated with 1 mM 13, 100 µM FAD, 2 mM NADPH, 50 mM NaCl and 10 µM SsuE. Inset in trace ii shows observed [M-H] - MS spectrum of 14. All traces are monitored at 300 nm.
Fig. 1
Fig. 2
Fig. 3
Spontaneous lactonization

Nonaketide

Spontaneous lactonization

Fig. 4
Fig. 5
Insights into radicicol biosynthesis via heterologous synthesis of intermediates and analogs
Hui Zhou, Kangjian Qiao, Zhizeng Gao, John C. Vederas and Yi Tang

J. Biol. Chem. published online October 20, 2010

Access the most updated version of this article at doi: 10.1074/jbc.M110.183574

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts