ACS chemical biology

Articles

pubs.acs.org/acschemicalbiology

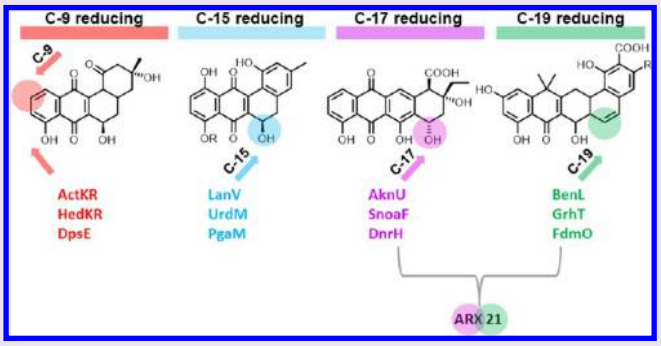
Comprehensive Analysis of a Novel Ketoreductase for Pentangular Polyphenol Biosynthesis

Timothy R. Valentic,[†] David R. Jackson,[†] Sean F. Brady,[‡] and Shiou-Chuan Tsai*[,†]

[†]Department of Molecular Biology and Biochemistry, Chemistry, and Pharmaceutical Sciences, University of California, Irvine, Irvine, California 92697, United States

[‡]Laboratory of Genetically Encoded Small Molecules, The Rockefeller University, 1230 York Avenue, New York, New York 10065, United States

ABSTRACT: Arixanthomycins are pentangular polyphenols (PP) with potent antiproliferative activities that were discovered through the heterologous expression of environmental DNA-derived gene clusters. The biosynthesis of arixanthomycin and other PPs is unusual because it requires several novel type II polyketide synthase (PKS) enzymes for its complete maturation. Most type II PKSs contain a ketoreductase (KR) that mediates the C7−C12 first ring cyclization and C-9 reduction. In contrast, based on previous studies of product analysis and genome mining, the arixanthomycin (ARX) gene cluster harbors a C-11 reducing KR (ARX 27), a C9−C14 first-ring aromatase/cyclase (ARX 19), and an unprecedented C-17 and C-19 reducing KR (ARX 21). While bioinformatics is useful for predicting novel enzymes, the functions of ARX 19, ARX 21, and ARX 27 have yet to be confirmed. Further, the structural features that predispose the ARX biosynthetic enzymes to process atypical poly-β-ketone scaffolds remain unknown. We report the crystal structure of ARX 21, the first structure of an enzyme involved in PP biosynthesis and likely a C-17 and C-19 reducing-KR, which is structurally similar to C-15 reducing KRs. Structural comparison of ARX 21 and other C-9 reducing KRs revealed a difference in the enzyme active site that may enlighten the molecular basis of KR substrate specificity. In addition, we report the successful *in vitro* reconstitution of ARX 19. The structural characterization of ARX 21 in conjunction with the *in vitro* results of ARX 19 lays the groundwork toward a complete *in vitro* and structural characterization of type II PKS enzymes involved in PP biogenesis.



Pentangular polyphenols (PP) are an underrepresented subfamily of bacterial aromatic polyketides that exhibit diverse biological activities.[1] Pradimicin, benastatin, and fredericamycin A are examples of PPs that serve as antifungals, chemotherapeutics, and antibiotics, respectively.[2] In *Actinomyces* and other bacteria, aromatic polyketides are generated by type II polyketide synthases (PKSs).[3−5] A prominent feature of type II PKSs is the utilization of a minimal PKS (Min PKS) that consists of a ketosynthase-chain length factor heterodimer (KS-CLF) and an acyl-carrier protein (ACP) for the generation of linear poly-β-keto thioester linked poly-β-ketones through repetitive decarboxylative condensations of malonyl-CoA.[6] Once a poly-β-ketone chain has been biosynthesized, it is enzymatically processed in a conserved order.[7,8] For example, in type II PKS reducing pathways, the mature poly-β-ketone chain appended to the ACP is first reduced at position C-9 by a ketoreductase (KR) followed by cyclization and aromatization by an aromatase/cyclase (ARO/CYC) to establish a set C7−C12 ring pattern.[9] Additional downstream cyclases, ketoreductase, and oxygenases further tailor the polyketide intermediate product to maturity.[10] While the biosyntheses of anthracycline and angucycline natural products by type II PKSs have been shown to follow a conserved sequence of enzyme catalyzed transformations, little is known about the order of processing and tailoring events that occur during PP biogenesis.[11,12] In addition, the gene clusters that encode the type II PKS enzymes involved in PP biosynthesis harbor many unusual features.[1,13] The PP scaffolds, for example, are derived from some of the longest polyketide chains, whereas most bacterial Min PKSs produce polyketide chains that range from 16 to 24 carbons in length. These chains are 26−30 carbons in length.[14] Second, based on product analysis, PP formation in certain cases follows noncanonical carbonyl reduction at positions C-11, C-17, and C-19, whereas most type II PKS KRs reduce the C-9 and C-15 carbonyl groups.[1,2,4,15,16] Finally, PP biogenesis requires an additional 15−20 tailoring enzymes for complete maturation, including glycosyl transferases, transamidases, and multiple oxygenases, which are found in lower abundance in typical type II PKSs.[4]

Advances in genome mining over the past decade have led to the discovery of unusual biosynthetic gene clusters from environmental samples.[17,18] Heterologous expression of these gene clusters has led to the identification of novel aromatic polyketides from unidentified or uncultured hosts.[14] In particular,

Received: August 2, 2016
Accepted: October 25, 2016
Published: October 25, 2016

DOI: 10.1021/acschembio.6b00658
ACS Chem. Biol. 2016, 11, 3421−3430

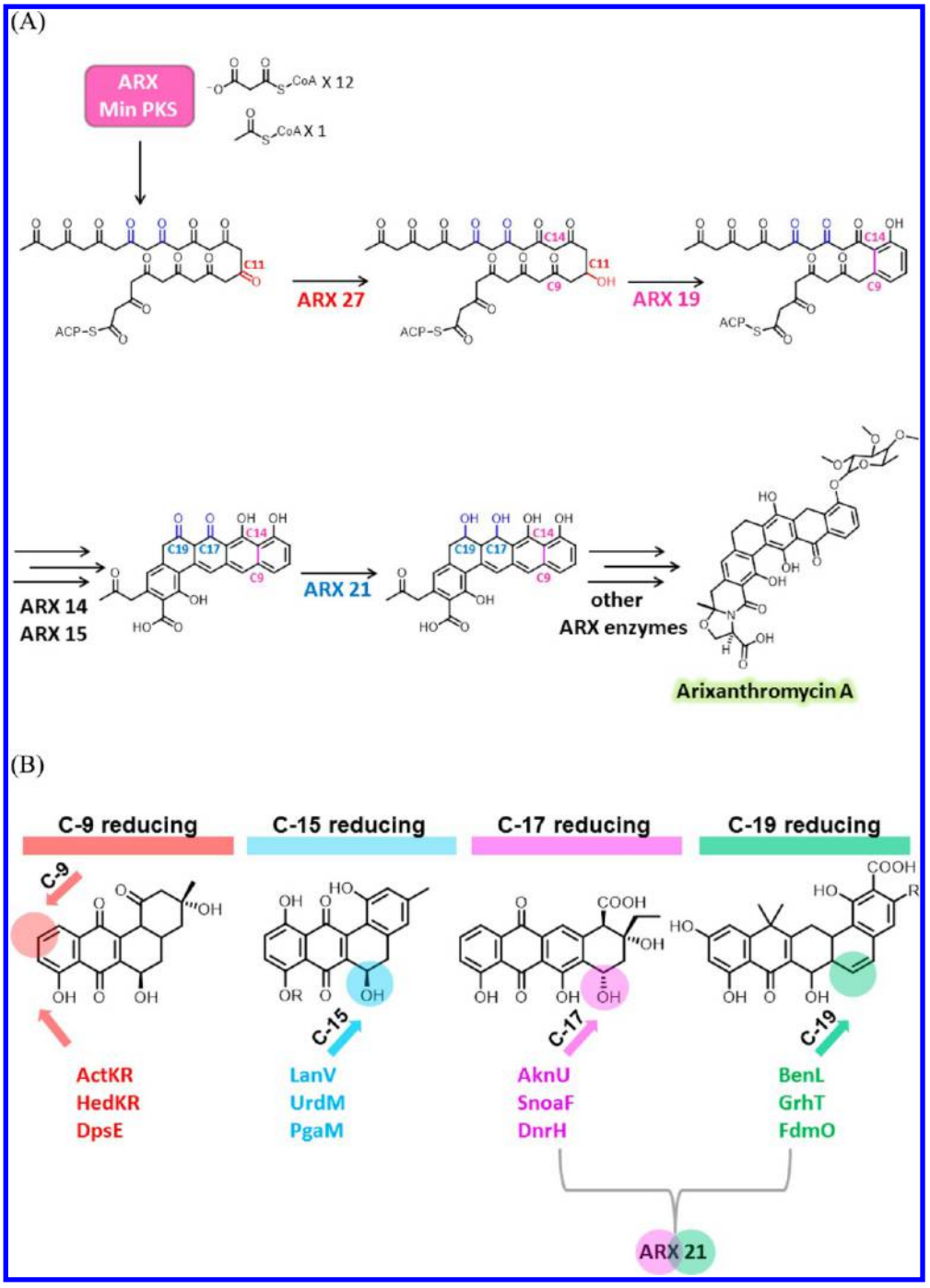


Figure 1. (A) The proposed biosynthesis of arixanthromycin A emphasizing key type II PKS processing enzymes. (B) A schematic of the differing KR clades and their native substrates. Note that ARX 21 has been predicted to behave as a C-17 and C-19 ketoreductase.

the arixanthomycins are a class of recently discovered PPs that exhibit attractive *in vitro* antiproliferative activities against bacteria and cancer cells.[1] Further, based on product analysis, the arixanthomycins gene cluster (ARX) may contain a treasure trove of unusual PKS biosynthetic enzymes, including an atypical ARO/CYC (ARX 19) that cyclizes a reduced polyketide chain at C9−C14, a novel C-11 reducing KR, and a tailoring C-17 and C-19 reducing KR (ARX 21; Figure 1A).[1] How ARX 19 and ARX 21 achieve such unusual feats of polyketide modification remains a mystery.

In this work, we present the crystal structure of ARX 21, the proposed C-17 and C-19 reducing KR, and elucidated the molecular basis for its unusual regio-specificity. We also report results from the *in vitro* reconstitution of ARX 19, a reducing monodomain ARO/CYC, in the presence and absence of ARX 21. We found that the ARO/CYCs from PP biosynthetic pathways have a more stringent chain length specificity in comparison to typical ARO/CYCs.[13,19] To the best of our knowledge, the crystal structure of ARX 21 represents the first structure of any enzyme domain responsible for PP biosynthesis, and this also likely represents the first structure of a bifunctional C-19 and C-17 reducing KR (Figure 1B). Interestingly, ARX 21 shares a high degree of structural similarity to C-15 KRs, which leads to our proposal about the molecular

DOI: 10.1021/acschembio.6b00658
ACS Chem. Biol. 2016, 11, 3421−3430

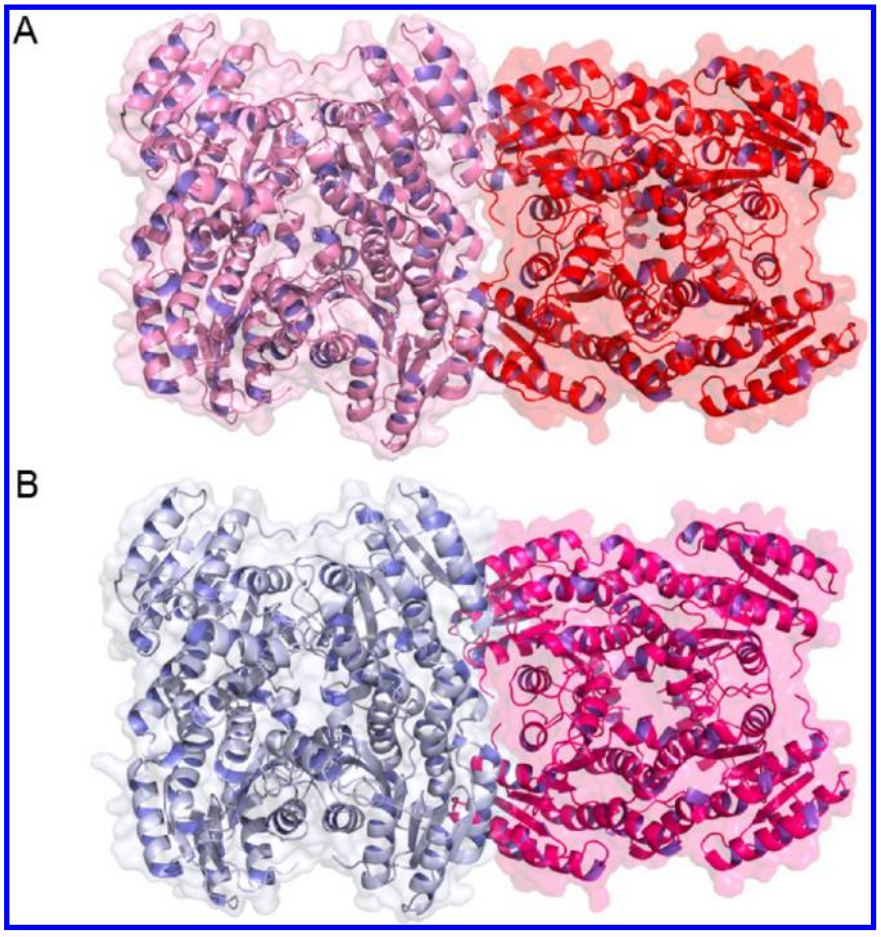


Figure 2. Overall architecture of ARX 21. (A) The structure of Apo ARX 21 with dimeric tetramers colored in pink and red. (B) The structure of NADPH bound ARX 21 with dimeric tetramers colored in gray and magenta.

mechanism on the ability of ARX 21 to reduce C-19 and C-17 carbonyl groups.[20] Taken together, the structural and functional studies here will further aid future efforts to implement regiospecific reduction and cyclization to diversify products generated by type II PKSs.

RESULTS AND DISCUSSION

Structure of *apo* ARX 21. ARX 21 crystallized in the absence of NADPH in the $P2_1$ space group with eight monomers in the asymmetric unit that form a dimer of tetramers (Figure 2A). The structure was determined to 2.7 Å via molecular replacement using the crystal structure of gluconate 5-dehydrogenase from *Thermotoga maritima* as a search model (PDB code: 1VL8). High quality electron density for the entire polypeptide chain could be resolved in the crystal structure with the exception of poor density corresponding to the α8–α9 helix–loop–helix lid region encompassing residues 188–203, which, depending on the monomer, was found to be partially or completely disordered. Surprisingly, upon inspection of the active site of ARX 21, partial electron density corresponding to NADPH was present for three of the eight ARX 21 monomers, suggesting that the NADPH cofactor was acquired intracellularly from the *E. coli* expression host. Previous structural studies with KRs have demonstrated that cofactor binding is crucial for the stabilization of the KR loop–helix–loop region through favorable structural perturbations that may improve crystal packing.[21] To assess if NADPH binding could aid in the modeling of a complete ARX 21 structure, the ARX 21 crystals were soaked in excess NADPH, and intensity statistics were collected. Clear electron density for NADPH and the missing residues 188–203 were present in the NADPH soaked crystal structure, providing an intact ARX 21 model that was refined to 2.4 Å (Figures 2B and 4C). Because the complete backbone of ARX 21 could only be successfully traced when bound to NADPH, the resulting structural analysis will refer to the NADPH-bound structure.

Overall Fold. Similar to other enzymes that belong to the short chain dehydrogenase/reductase (SDR) superfamily, ARX 21 contains a conserved Rossman fold that consists of six parallel α helices bundled together in two groups of three, flanking a central seven-stranded β-sheet core. In addition, ARX 21 harbors the canonical SDR TGXXXGXG NADPH binding consensus sequence and the conserved Ser–Tyr–Lys active site triad. The front of ARX 21's active site is demarcated by a groove that is formed between the α8–α9 flap and the loop region that spans from β4 to η1 (Figure 3). The back of the substrate-binding pocket is formed by the C-terminal end of the central seven-stranded β sheet core and a loop located between β1 and α1. The structural topology provides a large cavity that can accommodate both NADPH and a bulky pentacylic aromatic substrate. In both the apo- and NADPH-bound structures, the active site of ARX 21 is in an open conformation that leaves both the NADPH and substrate binding sites largely exposed to the solvent.

ARX 21 shares a high degree of structural homology but low sequence similarity (less than 34%) to previously determined

DOI: 10.1021/acschembio.6b00658
ACS Chem. Biol. 2016, 11, 3421–3430

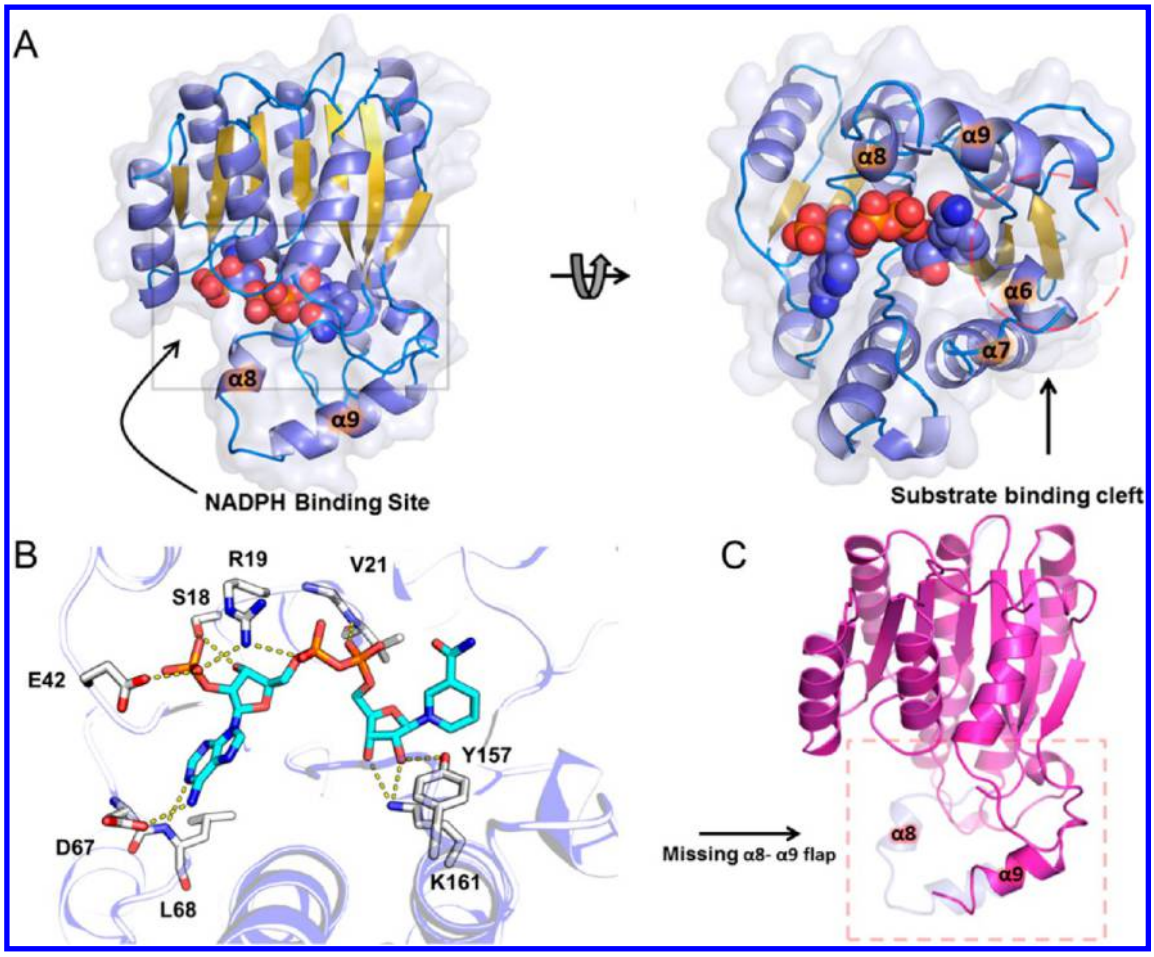


Figure 3. Cofactor and proposed substrate binding site of ARX 21. (A) The tertiary structure of one ARX 21 monomer bound to NADPH, with highlighting both the substrate and cofactor binding sites. The active site is located at a cleft formed by the $\alpha 8-\alpha 9$ flap and the loop region that spans $\beta 4$ and $\eta 1$. (B) Inspection of the residues that encompass the NADPH binding site of ARX 21 with the NADPH cofactor colored as cyan. (C) A structural overlay of an *apo* and NADPH bound ARX 21 monomer. Clear electron density defining the $\alpha 8-\alpha 9$ flap is absent in the *apo* form of ARX 21.

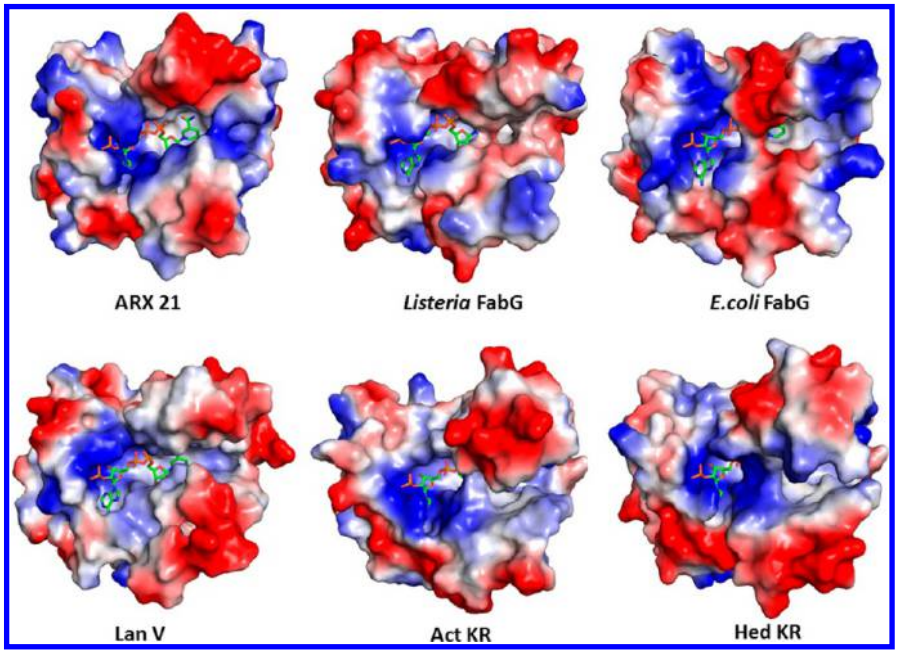


Figure 4. Electrostatic potential surfaces of ARX 21 and other KR homologues emphasizing the variance of charge distributions and active site accessibility across differing KR clades.

DOI: 10.1021/acschembio.6b00658
ACS Chem. Biol. 2016, 11, 3421−3430

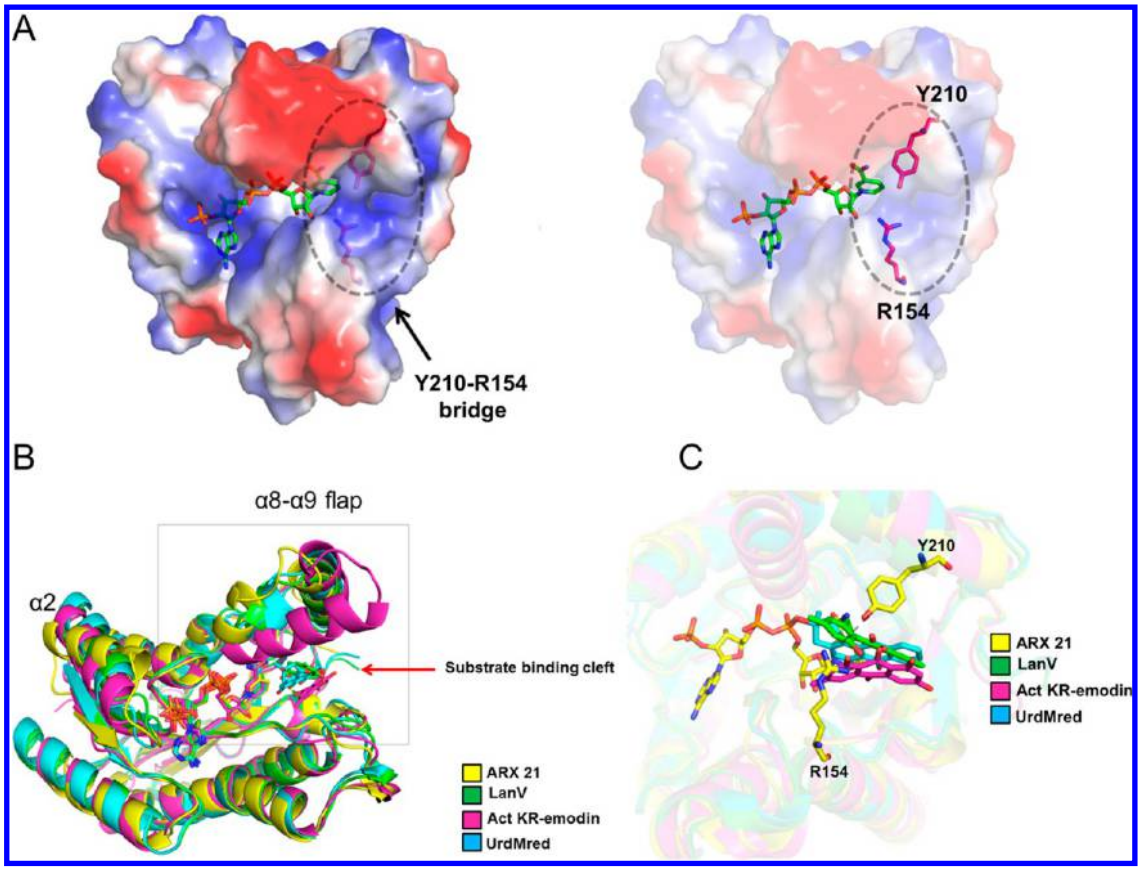


Figure 5. Residues proposed to be important for ARX 21 substrate binding and recognition. (A) An electrostatic surface representation of an NADPH bound ARX 21 monomer. Y210 and R154 block the entrance to the substrate binding cleft. (B) Superimposition of NADPH bound ARX 21 to the structure of ActKR (bound with emodin and NADP), LanV-11-deoxylandomycinone-NADP$^+$, and rebelomycin-NADP$^+$-UrdMred ternary complexes. The models are colored as gold, green, magenta, and cyan, respectively. (C) Structural overlay of substrate bound KRs, demonstrating a conserved substrate binding location highlighting Y204 and R154 of ARX 21.

KR structures involved in type II PKS and FAS biosyntheses. These similar KR structures include the actinorhodin KR (ActKR), the hedamycin KR (HedKR), LanV, and multiple FabG homologues. Dali structural alignment search revealed that ARX 21 shares the highest structural similarity to *Listeria* FabG (LFabG, PDB code 4JRO) with a *Z* score of 35.8, followed by an SDR from *Serratia marcescens* BCRC 10948 (SM_SDR, PDB code 4ZGW) with a Z-score of 35.5.[22] Superimposition of the LFabG and SM_SDR structures onto that of ARX 21 revealed that all three have an exposed active site cavity, and of these, ARX 21's pocket is the most solvent accessible (Figure 4). The low sequence similarity of ARX 21 to other structurally characterized type II PKS KRs merited additional structural comparisons of ARX 21 (a C-17 and C-19 reducing KR) to C-15 reducing KRs from different type II PKSs. ARX 21 shows the highest degree of structural similarity to the crystal structures of two C-15 reducing KRs, LanV and UrdMred, both bound to NADP+ and their putative substrates 11-deoxylandomycinone and rebelomycin, respectively, with Cα backbone RMSDs of 1.7 and 1.72 Å, respectively.[11,20,22,23] Like LanV and UrdMred, ARX 21 contains a deep active site cavity; however, it is slightly less accessible to the solvent in comparison to the crystal structure of the two C-15 reducing KRs. Given the roles C-15 and C-19 KRs play in the downstream reductive tailoring of angucyclines and pentangular polyphenol intermediates, the crystal structures of the LanV and UrdMred ternary complexes served as a solid framework to compare possible substrate binding modes between KRs with differing substrate reduction sites. In the LanV and UrdMred ternary complex crystal structures, 11-deoxylandomycinone and rebelomycin bind between the α6−α7 loop and the α8−α9 flap where the C-15 carbonyl oxygen of their respective substrates is placed within 3.8 Å of the C-4 atom of the nicotinamide ring of NADPH to achieve regiospecific reduction at position C-15. In ARX 21, residues that define the α6−α7 loop and α9 helix are bulkier in comparison to LanV and UrdMred, and consequently, the substrate binding cleft of ARX 21 is narrower (Figure 5A). Specifically, a weak salt bridge is formed between R154 and Y210 that constricts the substrate binding cleft of ARX 21, which would lead to protein−substrate steric clashes if an angucycline intermediate were bound in this position. For UrdMred and LanV, R154 and Y210 of ARX 21 correspond to E/Q156 and M/L206, respectively (Figure 5). The presence of smaller residues in these positions provides ample space between the α6−α7 loop and α8−α9 lid to accommodate the tetracyclic aromatic substrates for the C-15 reducing KRs. It is interesting to note that R154 and Y210 are highly conserved among annotated C-19 reducing KRs and may be important for substrate recognition and specificity (Figure 5). Because ARX 21 acts on large pentacyclic aromatic compound, it is within reason to suggest that substrate binding could cause Y210 and/or R154 to move in a manner that could either allow for entry to the substrate binding region of ARX 21 or make favorable contacts through π−π stacking interactions with an incoming

DOI: 10.1021/acschembio.6b00658
ACS Chem. Biol. 2016, 11, 3421−3430

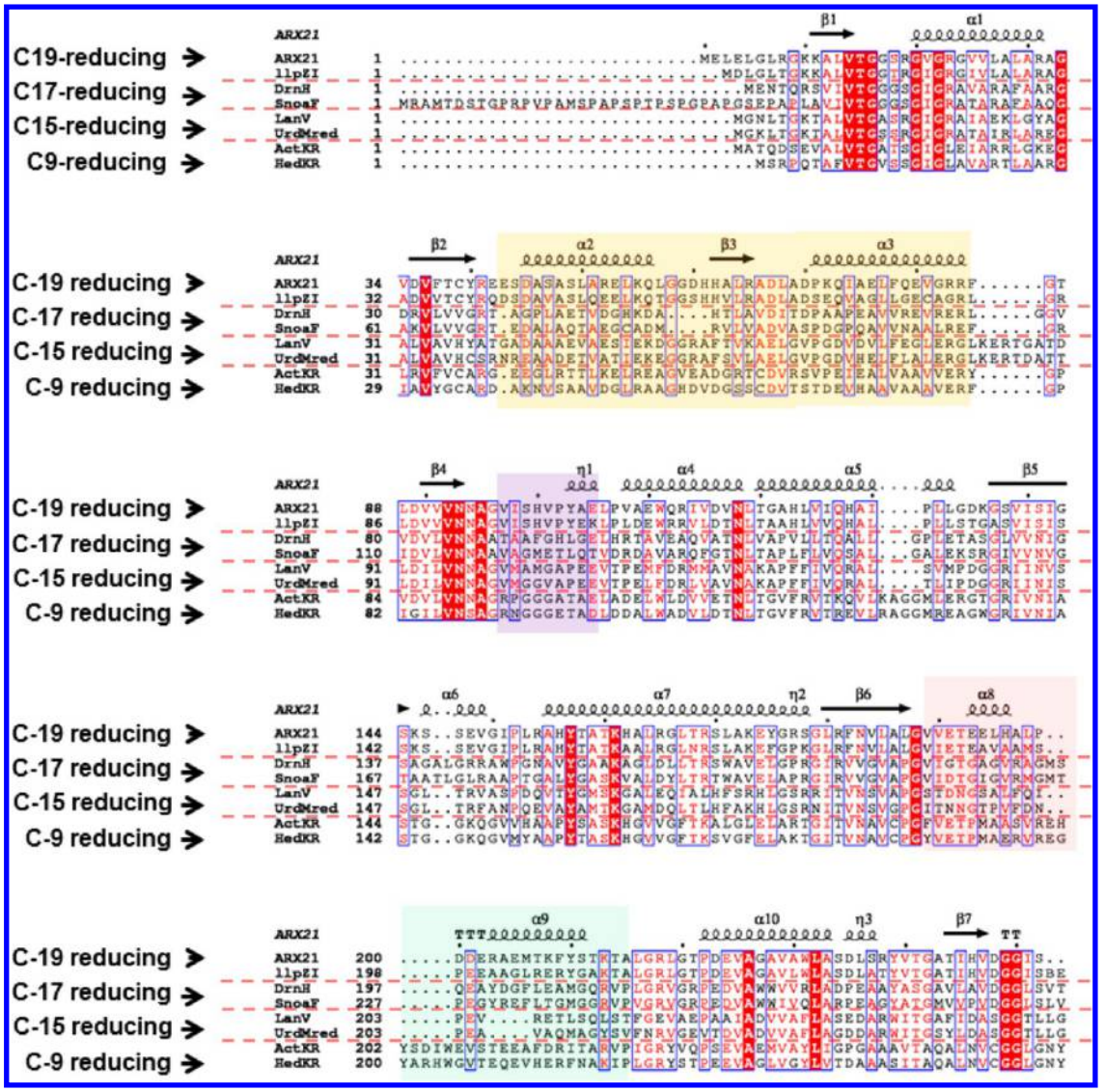


Figure 6. Sequence alignment of ARX 21 to functionally characterized C-19, C-17, C-15, and C-9 reducing KRs. Unconserved amino acid sequences are colored.

substrate (Figure 5A). Alternatively, it is also possible that ARX 21, a C-19 reducing KR, exhibits unique substrate binding modes in comparison to LanV and UrdMred. However, the cocrystal structure of the C-9 reducing ActKR bound to emodin suggests that the substrate binding regions are conserved across all KRs, as exemplified by the emodin binding site that is roughly in the same location as substrates 11-deoxylandomycinone and rebelomycin for LanV and UrdMred, respectively. Therefore, the crystal structure of ARX 21 further shows how KRs can acquire novel substrate and regio-specificities via the simple modification of several amino acids.

Structure-Sequence-Function Relationship. Due to the high degree of structural similarity of previously determined type II PKS KRs with diverse functions, a sequence alignment was performed for ARX 21 against functionally and/or structurally characterized C-19-, C-17-, C-15-, and C-9-reducing KRs. The majority of the structural deviations correlate with unconserved amino sequences that are localized in the following regions across all of the KR clades: helices α2, α3 α8, and α9 and loop regions that span between α2–β3, β4–α4, α6–α7, and α8–α9 (Figure 6). The amino acid sequences of the α9 helix are the most variable among all KRs and, consequently, the helix itself shows the greatest levels of structural deviations among deposited KR crystal structures with respect to ARX 21.

Assessing the Substrate Tolerance of ARX 21. To determine the substrate tolerance of ARX 21, we evaluated its activity *in vitro* using a well-established assay that utilizes the Pks4 minimal PKS (Pks4 Min PKS) and actinorhodin minimal PKS (Act Min PKS), which supply downstream enzymes such as KR with a linear octaketide and nonaketide intermediates, respectively (see Figure 7).[9,24] These two assays were previously used for C-9 reducing KRs such as Act KR, which typically have a high substrate tolerance, resulting in the biosynthesis of new polyketides.[15,19] We wished to test if ARX 21 may also exhibit similar substrate tolerance. However, when either Pks4 or Act Min PKS was used, the inclusion of ARX 21 did not lead to the formation of new products. This demonstrated that ARX 21 is not a promiscuous enzyme as are other C-9 reducing KRs. To assess whether the enzymatic activity of ARX 21 was dependent on a predetermined C9–C14 ring cyclization pattern, ARX 19 was incubated with ARX 21 in the presence of the Pks4 min

DOI: 10.1021/acschembio.6b00658
ACS Chem. Biol. 2016, 11, 3421–3430

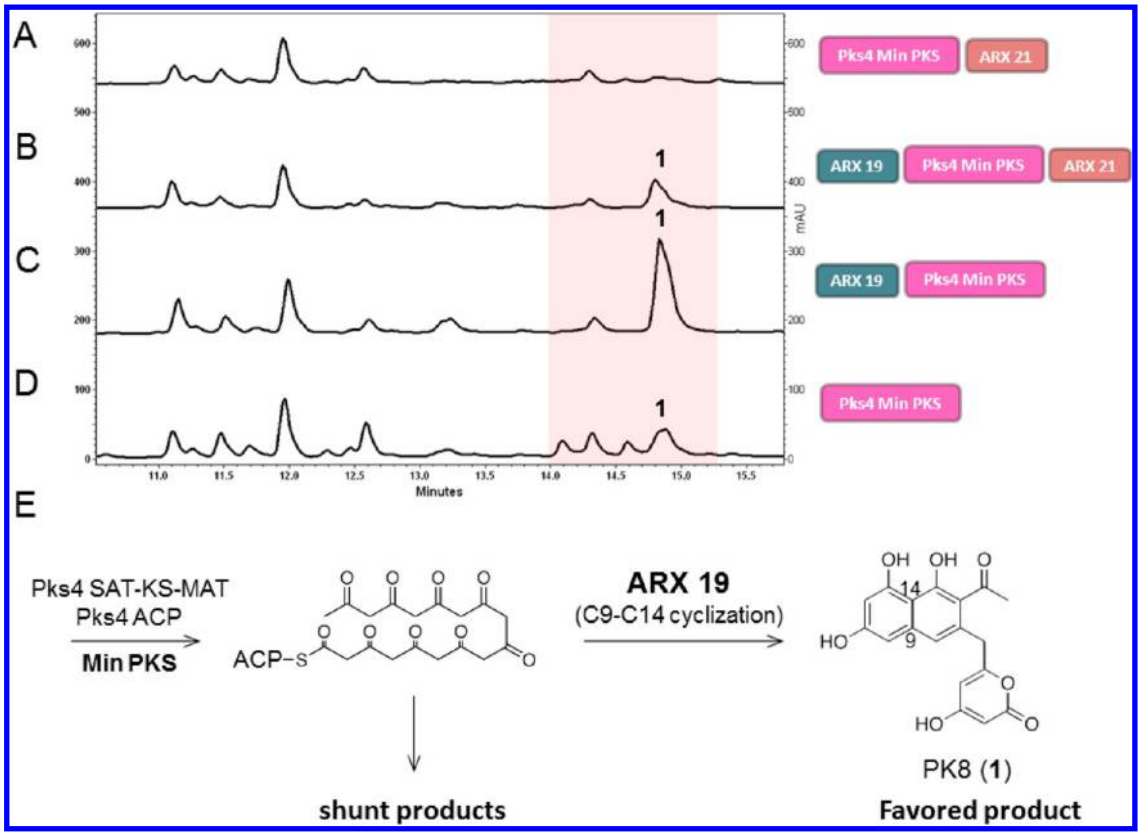


Figure 7. HPLC analysis of ARX 21 activities at 280 nm. (A) ARX 21 in the presence of the Pks4 Min PKS. (B) ARX 19 and ARX 21 in the presence of the Pks4 Min PKS. (C) ARX 19 in the presence of the Pks4 Min PKS. (D) Pks4 Min PKS alone. (E) A schematic of the Pks4 Min PKS based cyclization assay. In the absence of a C9−C14 ARO/CYC, the Pks4 Min PKS generates a nonaketide intermediate that affords spontaneous shunt products with myriad cyclization patterns. The inclusion of a C9−C14 ARO/CYC leads to a product profile shift favoring shunt products with C9−C14 cyclization patterns that can be detected using HPLC.

PKS; however, no new products were observed when ARX 19 and ARX 21 were included in the reaction. Interestingly, in the absence of ARX 21, ARX 19 was capable of acting upon a fully unreduced polyketide chain *in vitro*. This is consistent with a previous report that C9−C14 ARO/CYCs such as ARX 19 have high substrate tolerance when assayed with Pks4 Min PKS.[9,24] These findings suggest that the lack of ARX 21 activity toward octaketide and nonaketide is a consequence of its inability to process substrates that do not contain a pentangular polyphenol skeleton. The result of the *in vitro* assay is consistent with the observation that the structure of ARX 21 shares a degree of similarity with C-15 KRs, which do not accept ACP tethered linear polyketide substrates and only act upon substrates that harbor specific carbon scaffolds.

CONCLUSIONS

The rational combination of different PKSs for the generation of novel therapeutics is a major goal for natural products researchers and synthetic biologists. However, efforts toward controlled combinatorial synthesis have been hampered by a lack of structural and biochemical knowledge that clearly define which sets of protein−protein and protein−substrate interactions allow productive biogenesis. An understanding of compatible enzymes pairs from otherwise orthogonal biosynthetic pathways could allow for the introduction of foreign biosynthetic machinery into nonendogenous metabolic pathways for the formation of new natural products.

In this study, we present the crystal structure of ARX 21, a novel KR proposed to reduce C-17 and C-19 of the polyketide framework. The crystal structure of ARX 21 is the first reported for enzymes involved in PP biosynthesis, and likely the first structure for C-17 and C-19 reducing KRs. The exposed substrate pocket of ARX 21 bears a stronger resemblance to the C-15 reducing KRs, LanV and UrdMred. Superimposition of ARX 21 with the ternary complex structures of LanV, UrdMred, and ActKR revealed a conserved substrate binding location that is positioned between the α6−α7 loop and the α8−α9 flap. Although the structure of ARX 21 reported in this work is in its substrate free form, the conservation of the substrate binding site across multiple KR clades strongly suggested that the ARX 21 substrate is positioned in a similar location during the proposed C-17 and C-19 ketoreduction. Further, *in vitro* reconstitution demonstrated that ARX 21 was incapable of acting upon linear octaketide or nonaketide substrates produced by the Pks4 and Act Min PKSs, respectively. These findings support that the enzymatic activity of ARX 21 may be restricted to substrates that bear a PP scaffold.

In summary, this work has laid the framework for the *in vitro* reconstitution of the entire ARX pathway. Because the ARX pathway harbors unusual type II PKS biosynthetic machinery, a structural and biochemical investigation of the individual enzyme components involved in PP biogenesis will greatly aid in rational combinatorial biosynthesis of different type II polyketides.

MATERIALS AND METHODS

Cloning of ARX 19 and ARX 21. The genes encoding ARX 19 and ARX 21 were amplified from BAC-AZ1076/33/378 (BAC containing

DOI: 10.1021/acschembio.6b00658
ACS Chem. Biol. 2016, 11, 3421−3430

the arixanthomycin gene cluster) using primers listed below. The amplified PCR products were double-digested with NdeI and NotI (New England Biolabs), and then ligated onto pET28c (Novagen) for the generation of ARX 21 and ARX 19 constructs with N-terminal His_{6x} tags. All cloning steps were carried out with NovaBlue competent *E. coli* cells (Millipore). DNA sequences were verified using Genewiz automated sequencing.

ARX21F 5′-TCGCTGCATATGGAACTGGAACTC GGC-3′
ARX21R 5′-GCTATA-GCGGCC-GCTCAG-GAAATT-CCTCC-3′
ARX19F 5′-GCTATA-CATATG-GTGGCG-CACACC-GA-3′
ARX19R 5′-CATATA-GCGGCC-GCCTAG-CCGACC-GC-3′

Expression and Purification of ARX 21. The resulting N-terminal His_{6x} ARX 21 construct was transformed in BL21(DE3) *E. coli* competent cells (Novagen) for protein overexpression. Cell cultures harboring the His_{6x} ARX 21 plasmid were grown in terrific broth (2 × 1 L) supplemented with 50 μg/mL kanamycin until A_{600} reached 0.6−0.8. The temperature was then lowered to 18 °C, and protein expression was induced upon the addition of 0.5 mM IPTG. After incubation at 18 °C for 14−18 h, cells were harvested via centrifugation at 5000 rpm, pelleted, flash frozen, and stored at −80 °C. Cell pellets were thawed on ice and resuspended in lysis buffer (50 mM Tris·HCl at pH 8.0, 20 mM imidazole at pH 8.0, 300 mM NaCl, and 10% glycerol). The cell suspension was sonicated (6 × 30 s cycles), and the cell debris was separated by centrifugation at 14 000 rpm for 1 h at 4 °C. The clarified cell lysate was mixed with 5 mL of nickel NTA resin (Thermo Fischer Scientific) for 50 min at 4 °C. The nickel slurry was then transferred to a gravity flow column. The flow through volume was collected, and the resin was washed with 100 mL of lysis buffer (2 × 50 mL washes). Elution of ARX 21 was accomplished via the addition of 5 mL fractions containing increasing concentrations of imidazole in lysis buffer (40−450 mM). Protein purity was assessed by SDS-PAGE. Fractions containing ARX 21 were pooled and dialyzed against storage buffer (50 mM Tris·HCl at pH 8.0, 50 mM NaCl, 1 mM DTT, and 15% glycerol). Dialyzed protein was concentrated to 15−16 mg mL^{-1} using a 10 000 MWCO centrifugal filter device (Millipore) and buffer exchanged into crystallization buffer (50 mM Tris·HCl at pH 8.0 and 1 mM DTT) using a PD-10 desalting column (GE Healthcare) prior to crystal screens.

Expression and Purification of ARX 19. The ARX-19-pET28 construct was used to transform BL21(DE3) cells for protein overproduction. Four × 1 L cultures were grown in TB containing 50 μg/mL kanamycin until A_{600} reached 0.8−1.0 at 37 °C and 180 rpm. Protein expression was induced by the addition of 0.5 mM IPTG; the cells were then cooled and incubated at 18 °C for approximately 16−18 h. Cells were harvested via centrifugation at 5100 rpm for 20 min and resuspended in 50 mL of lysis buffer (50 mM Tris·HCl at pH 8.0, 20 mM imidazole at pH 8.0, 300 mM NaCl, and 10% glycerol). The cells were lysed through two rounds of microfluidization, and the cell debris was removed through centrifugation at 14 000 rpm at 4 °C for 1 h. The clarified cell lysate was incubated with 10 mL of nickel NTA resin (Bio-Rad) while being gently stirred at 4 °C for 1 h. The nickel−lysate slurry was then transferred to a gravity flow column where it was washed with 2 × 50 mL of lysis buffer followed by the addition of 10 mL of 40 mM imidazole in lysis buffer. Elution of ARX 19 was achieved through the inclusion of 3 × 5 mL fractions containing 100 mM imidazole in lysis buffer followed by the subsequent addition of 2 × 5 mL fractions containing 150, 250, and 450 mM imidazole in lysis buffer. The resulting fractions were subject to SDS-PAGE analysis prior to being pooled and dialyzed against Tris·HCl at pH 8.0, 50 mM NaCl, and 10% glycerol. ARX 19 was then concentrated to 5 mg mL^{-1} using a 10 000 MWCO centrifugal filter device (Millipore) and injected onto a Superdex 200 column (GE Healthcare) pre-equilibrated with 50 mM Tris·HCl at pH 8.0, 50 mM NaCl, and 1 mM DTT. Pure ARX 19 was concentrated to 13 mg mL^{-1}, flash frozen in liquid nitrogen, and stored at −80 °C for later use.

Crystallization and Data Collection. Small dagger-shaped ARX 21 crystals were obtained from combining 4 μL of protein solution at 7−10 mg mL^{-1} with 4 μL of well solution containing 50 mM Tris·HCl at pH 8.5, 0.2 M sodium acetate, and 30−32% PEG 4000 using the vapor diffusion method. ARX 21 crystals formed after 2−5 days at 20 °C. ARX 21 crystals were harvested and flash frozen in liquid nitrogen. For the NADPH-bound ARX 21 structure, *apo* crystals were soaked in the crystallization solution containing 20 mM NADPH (Sigma) for 20−30 s, then flash frozen in liquid nitrogen. Diffraction data for the *apo* ARX 21 crystals were collected at the Advanced Light Source (ALS) beamline 821 at the Lawrence Berkeley National Laboratory. Diffraction data for the NADPH-soaked ARX 21 crystals were collected at Stanford Synchrotron Radiation Lightsource (SSRL) beamline 12−2. Data collection statistics are listed under Table 1.

Phasing, Model Building, and Refinement. Data sets for both the *apo* and NADPH-bound forms of ARX 21 were indexed, integrated, and scaled using Mosflm.[25] Initial phases were obtained through molecular replacement (Phaser) using the crystal structure of gluconate 5-dehydrogenase from *Thermotoga maritima* as a search model (PDB code: 1VL8).[26] PHENIX AutoBuild was used to construct an initial model of ARX 21.[27] The final ARX 21 models were obtained through multiple rounds of model building and refinement using COOT and Phenix.REFINE.[28,29]

Table 1. Data Collection and Refinement Statistics

structure	apo ARX 21	NADPH-ARX 21
wavelength	0.9794	1.000
resolution range	52.67−2.6 (2.693−2.6)	86.3−2.3 (2.382−2.3)
space group	P1211	P1211
unit cell	105.38, 86.6499, 107.32, 90, 108.97, 90	105.34, 86.54, 106.97, 90, 108.8, 90
total reflections	104 658 (10172)	157 573 (15055)
unique reflections	54 082 (5326)	79 685 (7695)
multiplicity	1.9 (1.9)	6.7 (6.7)
completeness (%)	0.96 (0.95)	0.98 (0.96)
mean I/sigma(I)	6.66 (2.94)	10.70 (3.98)
Wilson B-factor	26.52	28.22
R-merge	0.03741 (0.1316)	0.03323 (0.1514)
R-meas	0.05291 (0.1861)	0.04699 (0.2141)
CC1/2	0.997 (0.944)	0.998 (0.954)
CC*	0.999 (0.985)	0.999 (0.988)
reflections used in refinement	52666 (4935)	79631 (7692)
reflections used for R-free	1959 (191)	1990 (198)
R-work	0.177 (0.194)	0.169 (0.1933)
R-free	0.236 (0.299)	0.217 (0.2632)
CC (work)	0.958 (0.932)	0.966 (0.938)
CC (free)	0.910 (0.852)	0.923 (0.841)
number of non-hydrogen atoms	14678	15662
macromolecules	14167	14844
ligands	96	384
protein residues	1919	1984
RMS (bonds)	0.008	0.017
RMS (angles)	1.11	1.46
Ramachandran outliers (%)	0.3	0.2
Rotamer outliers (%)	0.7	1.1
Clashscore	8.0	7.87
average B-factor	30.74	34.30
macromolecules	30.64	34.15
ligands	46.51	37.87
solvent	30.65	36.16

DOI: 10.1021/acschembio.6b00658
ACS Chem. Biol. 2016, 11, 3421−3430

For the NADPH bound ARX 21 structure, LigandFit was used to identify electron density corresponding to NADPH, and Elbow was used to generate ligand restraints.[29] The final NADPH-bound ARX 21 models were obtained through multiple rounds of model building and refinement using COOT and Phenix.REFINE.[28,29] The refinement statistics are listed under Table 1.

ARX 21/19 Assays. The fungal Pks4 minimal PKS was used to assay ARX 19 and ARX 21 activity *in vitro*. Then, 250 μL *in vitro* reactions with the following final concentrations with or without 100 μM ARX 19 and or 160 μM ARX 21, 50 μM *holo*-Pks4 ACP, 37 μM Pks4 SKM (SAT-KS-MAT tridomain), 2.5 mM malonyl-CoA, and NADPH 2.5 mM in 50 mM potassium phosphate at pH 7.5 were used to assess the activities of ARX 19 and ARX 21. The reactions were incubated overnight for approximately 15−17 h at 20 °C. The following day, the reactions were extracted three times with 300 μL of ethyl acetate/methanol/acetic acid (94:5:1). The organic layer was collected and dried using a speed vacuum concentrator. The resulting concentrate was then resuspended in 25 μL of DMSO and subject to reverse phase HPLC analysis as previously described.[9]

Expression and Purification of Pks4 SKM, Pks4 ACP, and SFP. The purification of Pks4 SKM, Pks4 ACP, and Sfp was performed as previously described. In all cases, purified protein was dialyzed against 50 mM Tris·HCl at pH 7.5, 50 mM NaCl, 10% glycerol, and 1 mM DTT at 4 °C. The dialyzed protein was concentrated, flash frozen in liquid nitrogen, and stored at −80 °C for later use.

Generation of Holo PKs4 ACP. *Holo* Pks4 ACP was generated via *in vitro* phosphopanetheinlyation of apo-Pks4 ACP. A 400 μL ACP loading reaction was set up using the following final concentrations: 375 μM apo pks4 ACP, 10 μM of Sfp, 1 mM CoA, and 12.5 mM $MgCl_2$ in 50 mM potassium phosphate at pH 7.0. The reaction was incubated at 37 °C for 2.5 h. The formation of holo-Pks4 ACP was verified using MALDI TOF mass spectrometry.

■ ASSOCIATED CONTENT

Accession Codes

The atomic coordinates of the apo and NADPH-bound ARX 21 structures have been deposited in the Protein Data Bank (PDB IDs: 5TII and 5THQ).

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: sctsai@uci.edu. ORCID: orcid.org/0000-0002-7082-9178.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

T.R.V. and S.-C.T. were funded by the National Institutes of Health through NIH R01 GM100305 and R01 GM076330. This work was part of the DOE Joint BioEnergy Institute (http://www.jbei.org) supported by the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research, through contract DE-AC02-05CH11231 between Lawrence Berkeley National Laboratory and the U.S. Department of Energy. A portion of this research was carried out at the Stanford Synchrotron Radiation Laboratory, a national user facility operated by Stanford University on behalf of the U.S. Department of Energy Office of Basic Energy Sciences. We would like to thank Amanda J. H. Gilliam for her help with editing the manuscript and thoughtful discussions surrounding experimental design.

■ REFERENCES

(1) Kang, H. S., and Brady, S. F. (2014) Arixanthomycins A-C: Phylogeny-guided discovery of biologically active eDNA-derived pentangular polyphenols. *ACS Chem. Biol. 9*, 1267−1272.

(2) Lackner, G., Schenk, A., Xu, Z., Reinhardt, K., Yunt, Z. S., Piel, J., and Hertweck, C. (2007) Biosynthesis of pentangular polyphenols: deductions from the benastatin and griseorhodin pathways. *J. Am. Chem. Soc. 129*, 9306−9312.

(3) Hertweck, C. (2015) Decoding and reprogramming complex polyketide assembly lines: prospects for synthetic biology. *Trends Biochem. Sci. 40*, 189−199.

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

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

(6) Javidpour, P., Bruegger, J., Srithahan, S., Korman, T. P., Crump, M. P., Crosby, J., Burkart, M. D., and Tsai, S. C. (2013) The determinants of activity and specificity in actinorhodin type II polyketide ketoreductase. *Chem. Biol. 20*, 1225−1234.

(7) Ames, B. D., Korman, T. P., Zhang, W., Smith, P., Vu, T., Tang, Y., and Tsai, S. C. (2008) Crystal structure and functional analysis of tetracenomycin ARO/CYC: implications for cyclization specificity of aromatic polyketides. *Proc. Natl. Acad. Sci. U. S. A. 105*, 5349−5354.

(8) Ames, B. D., Lee, M. Y., Moody, C., Zhang, W., Tang, Y., and Tsai, S. C. (2011) Structural and biochemical characterization of ZhuI aromatase/cyclase from the R1128 polyketide pathway. *Biochemistry 50*, 8392−8406.

(9) Caldara-Festin, G., Jackson, D. R., Barajas, J. F., Valentic, T. R., Patel, A. B., Aguilar, S., Nguyen, M., Vo, M., Khanna, A., Sasaki, E., Liu, H. W., and Tsai, S. C. (2015) Structural and functional analysis of two di-domain aromatase/cyclases from type II polyketide synthases. *Proc. Natl. Acad. Sci. U. S. A. 112*, E6844−6851.

(10) Meurer, G., Gerlitz, M., Wendt-Pienkowski, E., Vining, L. C., Rohr, J., and Hutchinson, C. R. (1997) Iterative type II polyketide synthases, cyclases and ketoreductases exhibit context-dependent behavior in the biosynthesis of linear and angular decapolyketides. *Chem. Biol. 4*, 433−443.

(11) Patrikainen, P., Kallio, P., Fan, K., Klika, K. D., Shaaban, K. A., Mantsala, P., Rohr, J., Yang, K., Niemi, J., and Metsa-Ketela, M. (2012) Tailoring enzymes involved in the biosynthesis of angucyclines contain latent context-dependent catalytic activities. *Chem. Biol. 19*, 647−655.

(12) Tian, J., Chen, H., Guo, Z., Liu, N., Li, J., Huang, Y., Xiang, W., and Chen, Y. (2016) Discovery of pentangular polyphenols hexaricins A-C from marine Streptosporangium sp. CGMCC 4.7309 by genome mining. *Appl. Microbiol. Biotechnol. 100*, 4189−4199.

(13) Zhou, H., Li, Y., and Tang, Y. (2010) Cyclization of aromatic polyketides from bacteria and fungi. *Nat. Prod. Rep. 27*, 839−868.

(14) Kang, H. S., and Brady, S. F. (2014) Mining soil metagenomes to better understand the evolution of natural product structural diversity: pentangular polyphenols as a case study. *J. Am. Chem. Soc. 136*, 18111−18119.

(15) Javidpour, P., Das, A., Khosla, C., and Tsai, S. C. (2011) Structural and biochemical studies of the hedamycin type II polyketide ketoreductase (HedKR): molecular basis of stereo- and regiospecificities. *Biochemistry 50*, 7426−7439.

(16) Javidpour, P., Korman, T. P., Shakya, G., and Tsai, S. C. (2011) Structural and biochemical analyses of regio- and stereospecificities observed in a type II polyketide ketoreductase. *Biochemistry 50*, 4638−4649.

(17) Reddy, B. V., Kallifidas, D., Kim, J. H., Charlop-Powers, Z., Feng, Z., and Brady, S. F. (2012) Natural product biosynthetic gene diversity in geographically distinct soil microbiomes. *Appl. Environ. Microbiol. 78*, 3744−3752.

(18) Handelsman, J., Rondon, M. R., Brady, S. F., Clardy, J., and Goodman, R. M. (1998) Molecular biological access to the chemistry of unknown soil microbes: a new frontier for natural products. *Chem. Biol. 5*, R245−249.

(19) Zhang, W., Li, Y., and Tang, Y. (2008) Engineered biosynthesis of bacterial aromatic polyketides in Escherichia coli. *Proc. Natl. Acad. Sci. U. S. A. 105*, 20683−20688.

(20) Patrikainen, P., Niiranen, L., Thapa, K., Paananen, P., Tahtinen, P., Mantsala, P., Niemi, J., and Metsa-Ketela, M. (2014) Structure-

DOI: 10.1021/acschembio.6b00658
ACS Chem. Biol. 2016, 11, 3421−3430

based engineering of angucyclinone 6-ketoreductases. *Chem. Biol. 21*, 1381−1391.
(21) Price, A. C., Zhang, Y. M., Rock, C. O., and White, S. W. (2001) Structure of beta-ketoacyl-[acyl carrier protein] reductase from Escherichia coli: negative cooperativity and its structural basis. *Biochemistry 40*, 12772−12781.
(22) Holm, L., and Rosenstrom, P. (2010) Dali server: conservation mapping in 3D. *Nucleic Acids Res. 38*, W545−549.
(23) Paananen, P., Patrikainen, P., Kallio, P., Mantsala, P., Niemi, J., Niiranen, L., and Metsa-Ketela, M. (2013) Structural and functional analysis of angucycline C-6 ketoreductase LanV involved in landomycin biosynthesis. *Biochemistry 52*, 5304−5314.
(24) Ma, S. M., Zhan, J., Xie, X., Watanabe, K., Tang, Y., and Zhang, W. (2008) Redirecting the cyclization steps of fungal polyketide synthase. *J. Am. Chem. Soc. 130*, 38−39.
(25) Battye, T. G., Kontogiannis, L., Johnson, O., Powell, H. R., and Leslie, A. G. (2011) iMOSFLM: a new graphical interface for diffraction-image processing with MOSFLM. *Acta Crystallogr., Sect. D: Biol. Crystallogr. 67*, 271−281.
(26) McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C., and Read, R. J. (2007) Phaser crystallographic software. *J. Appl. Crystallogr. 40*, 658−674.
(27) Terwilliger, T. C., Grosse-Kunstleve, R. W., Afonine, P. V., Moriarty, N. W., Zwart, P. H., Hung, L. W., Read, R. J., and Adams, P. D. (2008) Iterative model building, structure refinement and density modification with the PHENIX AutoBuild wizard. *Acta Crystallogr., Sect. D: Biol. Crystallogr. 64*, 61−69.
(28) Emsley, P., Lohkamp, B., Scott, W. G., and Cowtan, K. (2010) Features and development of Coot. *Acta Crystallogr., Sect. D: Biol. Crystallogr. 66*, 486−501.
(29) Adams, P. D., Afonine, P. V., Bunkoczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L. W., Kapral, G. J., Grosse-Kunstleve, R. W., McCoy, A. J., Moriarty, N. W., Oeffner, R., Read, R. J., Richardson, D. C., Richardson, J. S., Terwilliger, T. C., and Zwart, P. H. (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr., Sect. D: Biol. Crystallogr. 66*, 213−221.

DOI: 10.1021/acschembio.6b00658
ACS Chem. Biol. 2016, 11, 3421−3430