ChemBioChem

Research Article
doi.org/10.1002/cbic.202400357

Chemistry Europe
European Chemical Societies Publishing

www.chembiochem.org

Biosynthesis of Arcyriaflavin F from *Streptomyces venezuelae* ATCC 10712

Hung-En Lai[+],[a] Agata Kennedy,[c] Lewis Tanner,[b] Emma A. Bartram,[b] Soo Mei Chee,[a] Paul S. Freemont,*[a, d, e] and Simon J. Moore*[b]

Indolocarbazoles are natural products with a broad spectrum of bioactivity. A distinct feature of indolocarbazole biosynthesis is the modification of the indole and maleimide rings by regioselective tailoring enzymes. Here, we study a new indolocarbazole variant, which is encoded by the *acfXODCP* genes from *Streptomyces venezuelae* ATCC 10712. We characterise the pathway by expressing the *acfXODCP* genes in *Streptomyces coelicolor*, which led to the production of a C-5/C-5′-dihydroxylated indolocarbazole, which we assign as arcyriaflavin F. We also show that a flavin-dependent monooxygenase AcfX catalyses the C-5/C-5′ dihydroxylation of the unsubstituted arcyriaflavin A into arcyriaflavin F. Interestingly, AcfX shares homology to EspX from erdasporine A biosynthesis, which instead catalyses a single C-6 indolocarbazole hydroxylation. In summary, we report a new indolocarbazole biosynthetic pathway and a regioselective C-5 indole ring tailoring enzyme AcfX.

Introduction

Bisindoles and indolocarbazoles are a diverse natural product family with potential therapeutic applications, including Gram-positive antibacterials,[1–3] antitumor,[4–6] antifungal and antimalaria bioactivity.[7] Midostaurin (Rydapt®, Novartis) is an approved glycosylated indolocarbazole for clinical treatment of acute myeloid leukemia.[8] Two natural product indolocarbazoles, staurosporine and rebeccamycin, specifically inhibit protein kinases[9] and DNA topoisomerase I[5], respectively. As an indolocarbazole sub-class, the arcyriaflavins were originally isolated from the slime mold *Arcyria denudate* and thus are a prominent pigment within the red fruiting bodies of this organism. Arcyriaflavin A (**1**) has a core indolocarbazole scaffold fused to a maleimide ring. Although evolutionarily distinct, these natural products are also made by bacteria. Bacterial biosynthesis of indolocarbazoles is either low or variable, therefore either synthetic chemistry[10–14] or recombinant approaches are used to make these molecules. A number of indolocarbazole variants were recombinantly expressed from metagenomic sources by the Brady group,[15–18] while indolocarbazoles have also been overproduced in *Pseudomonas putida*[19] and *Streptomyces albus* J1074.[20]

The biosynthesis of indolocarbazoles was first characterised for the rebeccamycin pathway.[6,21,22] The indolocarbazole scaffold forming enzymes are RebO, -D, -C and -P, with the last letter of the protein name shared in all indolocarbazole pathways – i.e., for staurosporine biosynthesis: StaO, -D, -C and -P (Figure 1A). Indolocarbazole biosynthesis involves the oxidative dimerization of two L-tryptophan (L-Trp) derived indole rings. First, the flavin adenine dinucleotide (FAD)-dependent L-Trp oxidase (RebO) forms an indole-3-pyruvate imine. Then a cytochrome P450 oxygenase (RebD) catalyses the transformation of an imine and enamine L-Trp intermediate[23,24] to form the product chromopyrrolic acid (CPA).[21,23] Last, another cytochrome P450 enzyme RebP, catalyses the formation of a C–C bond forming an unstable indolocarbazole intermediate.[25] RebP works together with a FAD-dependent enzyme (RebC), to catalyse the oxidation of two carboxylic acids to a maleimide ring,[22] creating the product arcyriaflavin A (**1**) (Figure 1B). Variant indolocarbazoles stem from modifications to the indole ring or the maleimide ring.[15,26,27] Examples include 2-pyrrolidinone indolocarbazoles like K-252c (**2**) and carboxypyrrole indolocarbazole erdasporine A, which were isolated from soil-derived sources.[15] Other hydroxylated variants include arcyriaflavin E (**3**),[28] BE-13793 C (**4**),[5] and hydroxysporine.[16] These findings emphasise the role of late stage regioselective tailoring enzymes that decorate the indole ring. This includes the flavin-dependent C-6 hydroxylase EspX[15] and a two-component

[a] *H.-E. Lai,[+] S. Mei Chee, P. S. Freemont*
Section of Structural and Synthetic Biology, Centre for Synthetic Biology, Department of Infectious Disease, Imperial College London, London SW7 2AZ, UK
E-mail: p.freemont@imperial.ac.uk

[b] *L. Tanner, E. A. Bartram, S. J. Moore*
School of Biological and Behavioural Science, Queen Mary University of London, London E1 4NS, UK
E-mail: simon.moore@qmul.ac.uk

[c] *A. Kennedy*
School of Biosciences, University of Kent, Canterbury CT7 2NJ, UK

[d] *P. S. Freemont*
UK Dementia Research Institute Care Research and Technology Centre, Imperial College London, Hammersmith Campus, Du Cane Road, London W12 0NN, UK

[e] *P. S. Freemont*
UK Innovation and Knowledge Centre for Synthetic Biology (SynbiCITE) and the London Biofoundry, Imperial College Translation & Innovation Hub, Imperial College London, White City Campus 80 Wood Lane, London W12 0BZ, UK

[+] *These authors contributed equally to this work.*

Supporting information for this article is available on the WWW under https://doi.org/10.1002/cbic.202400357

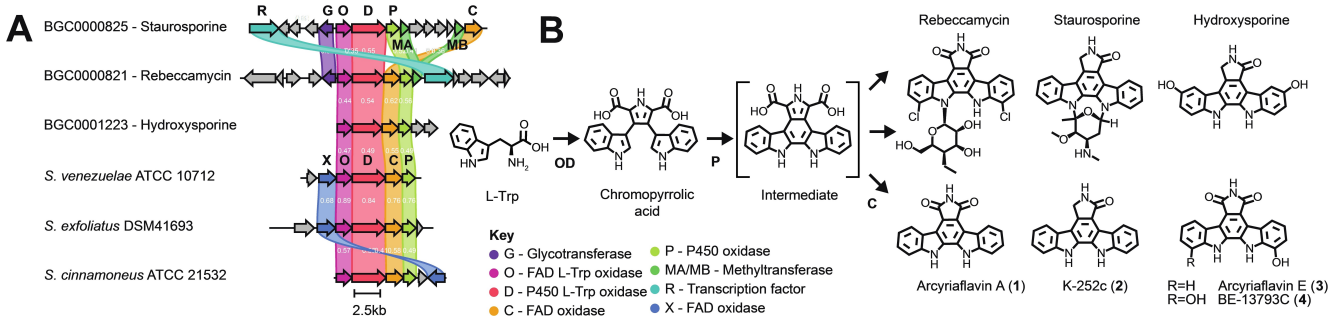


Figure 1. Overview of indolocarbazole biosynthesis and genetic organisation of the *S. venezuelae* ATCC 10712 arcyriaflavin F BGC.
A: Alignment of selected known (with MIBiG IDs) and putative indolocarbazole biosynthetic gene clusters using Clinker software.[34] Genes present in two or more BGCs are coloured and represented as single letters in the key. Cluster specific co-localised genes are coloured in grey. **B**: A general scheme for indolocarbazole biosynthesis and a selection of natural products. As an exception rebeccamycin biosynthesis begins with C-7 chlorination of L-Trp, whereas typically other indole ring tailoring steps occur late in the biosynthetic pathway.

cytochrome P450 C-5 hydroxylase HysX1/X2.[16] There are also a range of indolocarbazole variants isolated from environmental Actinomycetota.[28–31]

Building upon our interest in *Streptomyces venezuelae* ATCC 10712,[32,33] we identified a five-gene indolocarbazole biosynthetic gene cluster (BGC) encoded within its chromosome, which shares similarity to rebeccamycin (MIBiG: BGC0000821) and staurosporine (MIBiG: BGC0000825) BGCs[6,21] (Figure 1A). Distinctly, the *S. venezuelae* indolocarbazole BGC carries an uncharacterised flavoprotein (Locus tag: VNZ<?_>RS03650) with similarity to EspX from erdasporine A biosynthesis. We report the characterisation of the *S. venezuelae* indolocarbazole BGC, the flavoprotein enzyme and the pathway's product arcyriaflavin F.

Results and Discussion

Characterisation of the *S. venezuelae* Arcyriaflavin BGC

S. venezuelae ATCC 10712 encodes 30 biosynthetic gene clusters (BGCs) predicted by AntiSMASH bioinformatics analysis.[35] Five of these BGCs are known to encode chloramphenicol,[36,37] jadomycin,[36] watasemycin,[38] venepeptide[39] and pikromycin[40] biosynthesis. The other 25 BGCs remain uncharacterised, with one encoding five genes homologous to *rebO*, *-D*, *C* and *-P* from rebeccamycin biosynthesis, as well as *espX* from the erdasporine A pathway (MIBiG: BGC0001336). Interestingly, *S. venezuelae* was reported to make a form of arcyriaflavin,[41] however, the experimental details were not subsequently published (Professor Mervyn Bibb, personal communication). To investigate whether a laboratory condition could trigger the *S. venezuelae* ATCC 10712 type strain to make arcyriaflavin A or variants, we initially screened a range of routine *Streptomyces* growth media,[42] using both liquid and solid agar cultures. Analysis of organic extracts separated by C18 reverse-phase high-performance liquid chromatography (HPLC) failed to detect any arcyriaflavin-like metabolites by either ultraviolet-visible (UV-Vis) absorbance or mass spectrometry (MS) detection approaches, when compared to an arcyriaflavin A commercial standard. This suggested that cluster activation requires either an unknown effector(s), potentially present in the environment, or through genetic modification.

To investigate the proposed *S. venezuelae* arcyriaflavin (*acfXODCP*) BGC, we switched to heterologous gene expression. Using Gibson DNA assembly, we inserted the *acfXODCP* or *acfODCP* genes within the integrative pAV-gapdh vector,[43,44] which carries a strong $P_{gapdh(EL)}$ constitutive promoter upstream of the operon and a *phiC31* integrase for integration into the *S. coelicolor* M1152 chromosome.[44] An empty vector (pAV-gapdh) lacking the *acf* BGC was also included as a control. After conjugation with *S. coelicolor* M1152, exconjugants from the recombinant *acfXODCP* or *acfODCP* strains showed production of a light brown or orange pigment, respectively, after approximately 7–10 days of growth. The control colonies were a typical white-grey colouration, suggesting that the recombinant arcyriaflavin BGCs was active and responsible for pigment production. Next, we analysed ethyl acetate extracts of solid MS agar growth of the *S. coelicolor* M1152 recombinant *acfODCP* and *acfXODCP* strains by C-18 HPLC-MS in positive ion mode. First, no apparent indolocarbazoles or intermediates were detected from extracts of the M1152 control strain. For the *acfODCP* strain, known indolocarbazole biosynthetic intermediates and side-products were detected in low quantities, with the exception that CPA (Expected $m/z = 384.10$) was absent. CPA was previously detected as an intermediate during heterologous production of rebeccamycin in *Streptomyces albus*.[45] An arcyriaflavin A standard had a retention time (RT) of 39.9 min with an $[M+H]^+$ $m/z = 326.09$ (**1**). In the *acfODCP* strain, detected indolocarbazoles included arcyriaflavin A (**1**) at $[M+H]^+$ $m/z = 326.09$ (RT = 39.9 min), K-252c (also known as staurosporinone) (**2**) at $[M+H]^+$ $m/z = 312.11$ (RT = 35.5 min) and trace levels of an unknown species with $[M+H]^+$ $m/z = 328.11$ (RT = 33.75 min). In the *acfODCP* pathway variant, the major product was arcyriaflavin A (**1**), constituting over 95% of the total extracted ion chromatogram (EIC) for all indolocarbazole metabolite species detected (Figure 2). The putative species corresponding to K-252c (**2**) and $m/z = 328.11$ accounted for <5% of the total EIC. For the *acfXODCP* strain,

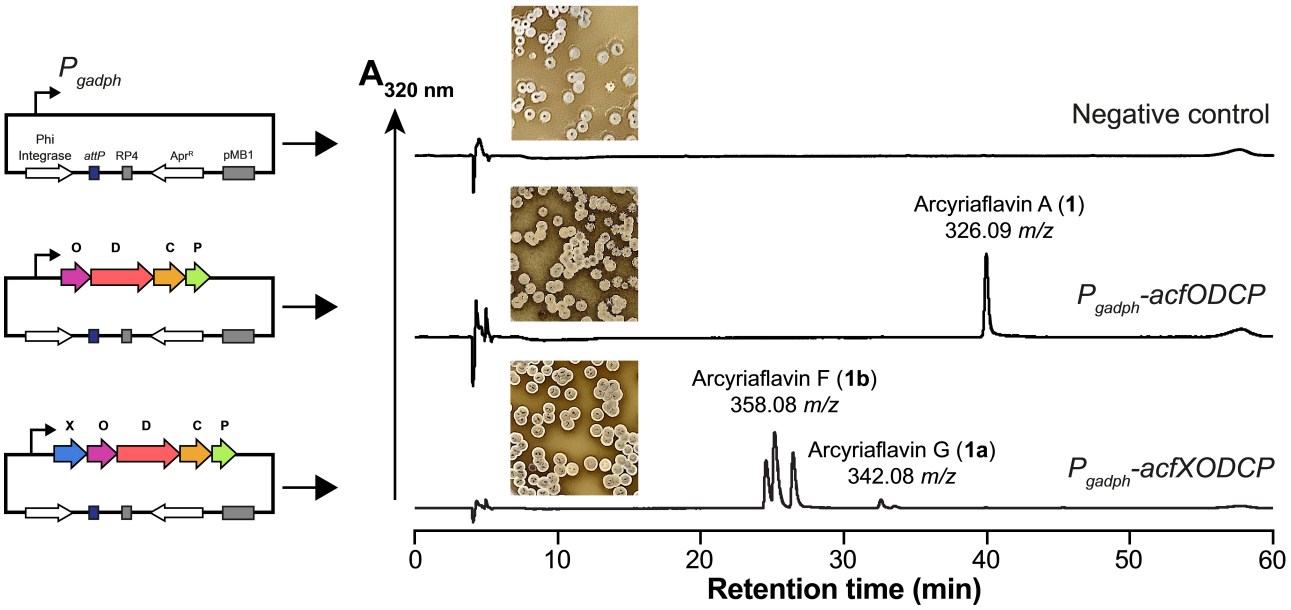


Figure 2. HPLC-UV chromatograms of organic extracts of *S. coelicolor* M1152 overexpressing the *S. venezuelae* ATCC 10712 *acfODCP* and *acfXODCP* operons. Cells were grown for 7 days at 30 °C on MS agar plates and prepared for HPLC-UV-MS analysis as described in the methods. Chromatograms represent relative absorbance at 320 nm, with the dominant mass for each peak displayed. Extracted ion chromatograms are provided in Figure S1

arcyriaflavin A was absent. Instead, new peaks for a minor (~2%) and a major (~98%) product with $[M+H]^+$ *m/z* = 342.08 (RT = 32.8 min) and 358.08 (RT = 24.7 min) was observed, respectively (Figures S1, S2). These products corresponded to the addition of 15.99 and 31.99 mass units, or one and two oxygen atoms respectively, suggesting potential hydroxylation. The decrease in RT for both these molecules in comparison to arcyriaflavin A (compound **1**), also suggested reduced hydrophobicity. For this, we initially assigned these peaks as potential mono- and dihydroxylated analogues of arcyriaflavin A (**1 a** and **1 b**), respectively. Interestingly, the dihydroxylated species at 24.7 min, consistently split into three peaks both on UV-Vis absorbance and MS (EIC of 358.08) detection. Interestingly, the chemical synthesis of arcyriaflavin analogues, including arcyriaflavin F, was previously reported, which also noted general instability for the hydroxylated species.[4] Overall, the HPLC-MS data supports that arcyriaflavin A is the major product for the recombinant strains expressing the *acfODCP* genes, whereas a dihydroxylated analogue of arcyriaflavin A, is the major product for the *acfXODCP* genes.

Structural Characterisation of Arcyriaflavin F and C-5/C-5′ Hydroxylation

The product of the *acfXODCP* recombinant pathway is consistent with mono- (minor) and dihydroxylated (major) variants of arcyriaflavin A. To confirm the structure, we scaled-up the growth of the *S. coelicolor* M1152 *acfXODCP* strain (in 4 L of MR5 media) for compound isolation and NMR analysis. Due to low quantities of the presumed mono-hydroxylated analogue arcyriaflavin G (**1 a**), we did not isolate this species for structural characterisation.

Next, we purified the major product of the *S. coelicolor* M1152 *acfXODCP* strain as outlined in the methods. From 4 litres of culture ferment, 13.8 mg of **1 b** was purified as a brown solid. Based on the EIC for the mass 358.08 *m/z* and related pathway metabolites, the purity of **1 b** was estimated to be greater than 99%. **1 b** was dissolved in 600 μL d_6-DMSO and the ^{1}H, ^{13}C, HMBC, HMQC, HSQC and ROESY NMR spectra were obtained (NMR data is provided in Table S2 and Figure S3). The ^{1}H NMR spectrum showed 11 protons, including three aromatic protons between δ 7.02 (dd), 7.59 (d) to 8.40 (d). In comparison to previous NMR spectra for arcyriaflavin A, four aromatic protons (between δ 7.35 and 8.98) were previously observed for the C-4 and C-7 positions,[28] suggesting one of the positions is connected to a new group. Further, there were three downfield-shifted protons, which were assigned as a single OH (δ 9.2) and two NH (δ 10.85 and 11.4) groups. For the ^{13}C NMR spectrum, 20 carbons were present, consistent with the arcyriaflavin A structure.[28] Based on the combined MS and NMR spectra, the data supports the structure of **1 b** with two hydroxyl groups at the C-5 and C-5′ indole ring positions (atomic position 10/22 in Figure 3 and Table S2). The C-5/C-5′ hydroxyl group assignment is supported by the following observations: ROESY through-space interactions between the indole NH group and the C-7 proton, and ROESY connections for the C-5/C-5′ hydroxyl proton with the C-4 and C-6 protons. In addition, C-5 is downfield shifted (151.48 ppm) in comparison to C-4 (108.82 ppm) and C-6 (116.32 ppm), which favours C-5/C-5′ as the hydroxylation position. In addition, the HSQC and HMQC spectra for arcyriaflavin F also revealed three sets of C–H couplings each, indicating mixed species. This is consistent with the split HPLC-MS peak and previous literature for chemical synthesis of arcyriaflavin F.[4] Based on the major species, we

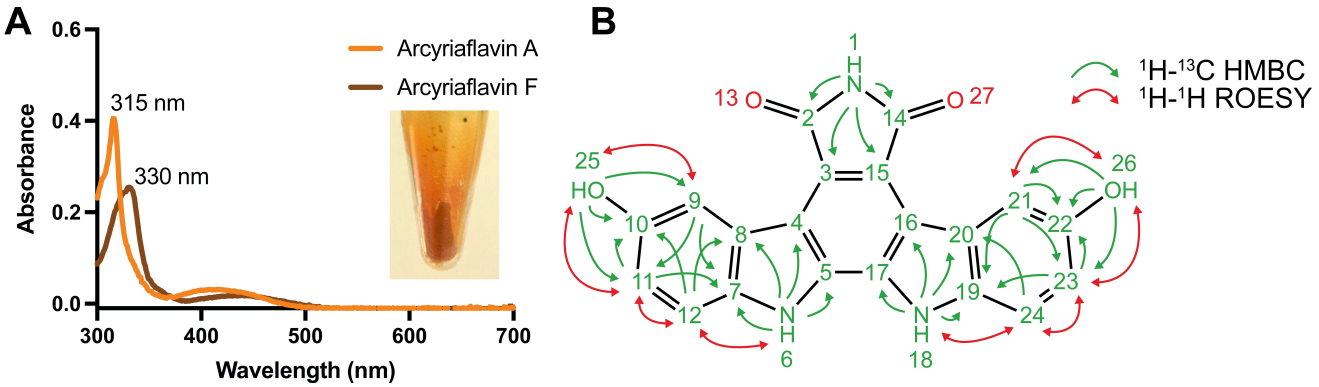


Figure 3. Characterisation of arcyriaflavin F.
A: UV-Vis absorbance spectrum of purified arcyriaflavin A and arcyriaflavin F in ethanol (inset image – dried arcyriaflavin F extract). **B**: Structure of arcyriaflavin F and ^{1}H-^{13}C HMBC and ^{1}H-^{1}H ROESY correlations. One- and two-dimensional NMR spectra are provided in Figure S2. Atom positions are labelled according to Table S2.

assign **1b** as a C-5/C-5′ dihydroxylated form of arcyriaflavin, hereafter referred to as arcyriaflavin F.

AcfX Catalyses C-5 Hydroxylation of Arcyriaflavin A

C-5 and C-5′ indole hydroxylation is observed in hydroxysporine biosynthesis, but in contrast, this reaction is distinctly catalysed by two HysX1 and HysX2 cytochrome P_{450} enzymes.[16] AcfX shares 54.7% amino acid (aa) identity to EspX, a FAD hydroxylase that catalyses a single hydroxylation at the C-6 indole ring position to form erdasporine A biosynthesis.[15] AcfX is annotated as a FAD-dependent monooxygenase (GenBank ID: WP_145953692) with 21% aa identity to the *Pseudomonas fluorescens* kynurenine 3-monooxygenase (PDB: 6FOX_A), as well as several other aromatic hydroxylases. These group A FAD monooxygenases collectively belong to the *p*-hydroxybenzoate-3-hydroxylase prototype family.[46] AcfX also has a typical Rossman fold GXGXXG motif at the N-terminus for FAD binding and two predicted disordered regions at the N- (1-26 aa) and C-terminus (551-566 aa). First, we suspected that the N-terminal disordered region was misannotated on the NCBI database (WP_145953692), due to an alternative start codon (GTG) and an unknown ribosome binding site. A BLASTP search confirmed that the N-terminus (1-26 aa) is absent on AcfX homologues present on the Protein Data Bank or more closely related *Streptomyces* genus protein sequences. Due to this uncertainty, we cloned three variants of *acfX* – termed *acfX1*, *acfX2* and *acfX3* (See Supporting information). The full-length *acfX1* variant encoded both the N- and C-terminal disordered extensions, which was cloned with a C-terminal His_6-tag into pSF1 C-SP44, a plasmid driven by a constitutive promoter. For *acfX2* we removed 1–26 aa but kept the extended C-terminus, which was cloned into pET15b with a T7 promoter, N-terminal His_6-tag and thrombin protease site. *acfX3* lacking the disordered C-terminus was similarly cloned into pET15b. We began by characterising the *acfX1* variant and found it was soluble when produced in *E. coli* BL21 Star (DE3) grown in autoinduction media at 28 °C. AcfX1 was purified by immobilized metal ion affinity (IMAC) chromatography (see methods) with yields of approximately 2–5 mg/L of culture. AcfX1 was observed as a single band at approximately 62 kDa by denaturing polyacrylamide gel electrophoresis (Figure 4A and Figure S4). In comparison, the yields of AcfX2 were much lower (<0.5 mg/L), with most of the protein located in the insoluble fraction. The N- and C-terminally truncated AcfX3 variant was poorly expressed and insoluble. Therefore, we continued with AcfX1 for further experiments.

During the purification, AcfX1 purified as pale-yellow coloured protein. Unfortunately, AcfX1 was unstable in a range of typical pH/salt conditions and gradually precipitated during elution with high concentrations (~200–400 mM) of imidazole. To stabilise the protein, we eluted AcfX1 at 200 mM imidazole and rapidly buffer exchanged the protein into S30 buffer (see methods), with the addition of 15% (v/v) glycerol to help stabilise the protein. Despite careful preparation, AcfX1 slowly aggregated at high concentrations (1–5 mg/mL) at either room

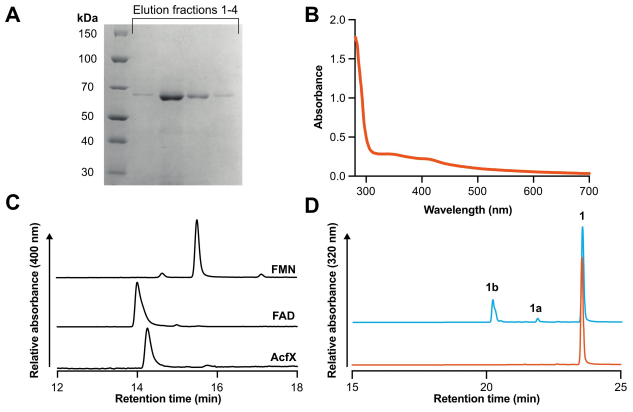


Figure 4. Characterisation of the AcfX C-5/C-5′ indolocarbazole hydroxylase.
A: Denaturing polyacrylamide gel electrophoresis (PAGE) of purified AcfX. Please see Figure S4 for additional information on expression and purification of AcfX1 and AcfX2. **B**: UV-Visible absorbance of AcfX1. **C**: HPLC analysis of AcfX flavin cofactor with analytical standards of FMN and FAD. **D**: HPLC-MS analysis of an AcfX1 reaction after 16 hours of incubation. Negative control and additional AcfX1 and AcfX2 HPLC enzyme activity experiments are shown in Figures S5–S10.

temperature or 4 °C. For the yellow pigment, and prediction that AcfX is a flavoprotein, we measured the UV-Vis light absorption spectrum and observed major absorbance peaks at 350 nm and 420 nm (Figure 4B), typical of flavin bound to the protein. This was further confirmed using FAD and FMN standards by HPLC-MS analysis (Figure 4C). To study AcfX1 enzymatic activity, we used C-18 HPLC analysis (Method B) to monitor conversion of **1** into **1 b**. 10 μM of freshly purified AcfX was incubated in the presence of 50 μM **1** at 21 °C. To provide a source of electrons, we supplemented the reaction with 5 mM NADPH. After 16 hours of incubation, approximately 25–40 % of **1** (20.7 min) was converted into **1 b** (20.50 min) across four independent repeats using only freshly purified AcfX1 (Figure 4D and Figures S5, S6). Interestingly, trace levels of the mono-hydroxylated **1 a** were also detected (21.87 min), which complements the cell-based production data (Figure 2). For the controls, in the absence of AcfX or NADPH, no product conversion was observed (Figure S7, S8), while the enzyme was also active with 5 mM NADH. Surprisingly, AcfX1 activity was equivalent at either 4 °C or 21 °C (Figure S9) but was less active at 30 °C (Figure S5), suggesting the main limitation with AcfX enzymatic activity is its instability. Further, after storage at 4 °C for 48 hrs in S30 buffer with 15 % (v/v) glycerol, AcfX1 only showed traced levels of activity in comparison to freshly purified enzyme (Figure S10). Finally, we also tested the AcfX2 variant for hydroxylation activity, but it was weaker than AcfX1 at either 4 °C or 21 °C. While we also tried to increase **1** to **1 b** conversion by increasing AcfX1 concentration, this change only increased aggregation. Additional optimisation of the pH (HEPES, Tris-HCl, phosphate buffers) and salt (0.1–1 M NaCl or KCl) conditions also did not improve activity, with the most active condition obtained in S30 buffer. We suggest the limited *in vitro* enzymatic activity of AcfX1 is likely due to the N- and C-terminal disordered regions – which while serendipitously enabling expression and purification – could promote aggregation over prolonged incubations. In contrast, AcfX fully converts **1** into **1 b** in cells (Figure 2). It is known that disordered regions, particularly at the C-terminus can play a role in protein function such as protein-protein interactions,[47] which may be important for its efficient activity and stability in cells.

Overall, while the *in vitro* enzymatic activity of AcfX was relatively modest, our data supports the role of AcfX as a FAD-dependent C-5/C-5′ arcyriaflavin A hydroxylase. The general instability of AcfX also limited further kinetic or mechanistic studies. Therefore, we suggest AcfX, like other group A FAD monooxygenases,[46] uses an external NAD(P)H donor to reduce the AcfX-bound FAD cofactor to $FADH_2$ before oxygen splitting and FAD-4a-OOH formation. FAD monooxygenases are known to readily hydroxylate kinetically accessible carbanions, which are found in indole, phenol, and pyrrole rings. Another example is OxyS from oxytetracycline biosynthesis, which shares 33 % identity to AcfX. OxyS modifies the enol and keto scaffold in the final steps of oxytetracycline biosynthesis.[48]

Conclusions

Herein, we study the arcyriaflavin F biosynthetic pathway from *S. venezuelae*. We show MS and NMR data supporting C-5/C-5′ hydroxylation at the indole ring positions. This tailoring reaction is distinct to the related erdasporine A, which has only a single hydroxyl group at the C-6 position catalysed by the EspX hydroxylase.[15] The biosynthetic rationale of AcfX also differs to hydroxysporine biosynthesis, which instead installs C-5/C-5′ hydroxyl groups through the action of a dual cytochrome P_{450} (HysX1 and HysX2) system. Overall our work reveals a new insight into regioselective tailoring of indolocarbazole biosynthesis and contributes to the growing area of synthetic biology for natural products.

Experimental Section

Bacterial Strains and Growth Media

For general cloning and plasmid purification, *E. coli* JM109 or DH10β strains were used. *E. coli* strains were grown in Luria-Bertani (LB) autoclaved media with appropriate antibiotic concentrations (carbenicillin 100 μg/mL, apramycin 25 μg/mL and nalidixic acid 30 μg/mL) at 30 °C or 37 °C. *S. coelicolor* M1152 was provided by Prof. Mervyn Bibb, John Innes Centre, UK, and was grown in mannitol soya agar (20 g mannitol, 20 g soya flour, 20 g agar) or liquid media (i.e., LB or GYM – 4 g/L glucose, 4 g/L yeast extract and 10 g/L malt extract pH 7.2) at 30 °C. Scaled-up cultures were grown in modified R5 (MR5) media[49] (0.25 g/L K_2SO_4, 10.12 g/L $MgCl_2.6H_2O$, 10.0 g/L D-glucose, 0.1 g/L casamino acids, 5.0 g/L yeast extract, 2.0 g/L $CaCO_3$, 0.2 % (w/v) trace elements, adjusted to pH 6.8, and grown at 30 °C with shaking (150–200 rpm).

Assembly and Conjugation of the acf BGC

The *acf* BGC was amplified from *S. venezuelae* ATCC 10712 genomic DNA as ~3 kb fragments and assembled into the pAV-gapdh vector[43,44] (a gift from Professor Jay Keasling) by Gibson DNA assembly. Each PCR fragment was generated using Q5® High-Fidelity DNA Polymerase and ensuring ~30 bp overlap between the homologous regions. 5 μL of NEBuilder® HiFi DNA Assembly Master Mix (2X) was added to 5 μL of DNA mixture containing 0.05 pmol of each PCR fragment and incubated at 50 °C for 2 h before transforming into competent *E. coli* DH10β cells. The *acf* BGC was then verified by commercial long-read sequencing (by Full Circle Labs, UK). Oligonucleotides and plasmids are listed in Table S1. Conjugation was performed using *E. coli* ET12567::pUZ002 into *S. coelicolor* M1152 as described previously.[42]

HPLC-MS Analysis

A single *S. coelicolor* M1152 colony was picked and inoculated into 10 mL autoclaved MR5 media with nalidixic acid and apramycin in 50 mL sterile tubes. Culture was incubated at 30 °C for 12 days and extracted with 1 volume of ethyl acetate. The mixture was further incubated at 30 °C, 200 rpm shaking for 1 h, and the organic layer was transferred to a fresh tube, dried *in vacuo* and reconstituted in 300 μL of MeOH. Pathway metabolites were analysed by HPLC-MS (Method A) using an Agilent 1100 LC system connected to a Bruker micrOTOF II MS and detected with positive electrospray ionization (ESI), using an ACE 5 AQ C-18 column (2.1×250 mm, 5 μm particle

size). For HPLC analysis, buffer A was water with 0.1% (v/v) trifluoroacetic acid and buffer B was acetonitrile with 0.1% (v/v) trifluoroacetic acid, for a total running time of 60 min including 10 min post-run time. Metabolites were detected at 320 nm, close to the maximal extinction coefficient for arcyriaflavins. 1–10 μL samples were injected, with a flow rate of 0.2 mL/min, oven temperature of 25 °C, and the following gradient: 0 min 5% B, 50 min 100% B, 55 min 100% B, 60 min 5% B (Method A). Absorbance spectra was monitored by Diode Array UV-Vis detection (Agilent Technologies).

Structural Characterisation of Arcyriaflavin F

S. coelicolor M1152 exconjugant with the *acf* BGC integrated was inoculated into 4 L of MR5 media and grown for 12 d at 30 °C and 200 rpm. Culture was extracted with an equal volume of ethyl acetate, and the organic layer was separated and dried *in vacuo*. The crude extract was reconstituted in 4 mL of methanol and separated on a pre-equilibrated Biotage SNAP Ultra C-18 column 12 g attached to a Biotage Isolera Spektra system. Buffer A was ultrapure water (HPLC grade) and buffer B was acetonitrile (HPLC grade), for a total running time of 16.5 min including 2.4 min equilibration time. The column was run at 36 mL/min with the following gradient: 0 min 10% B, 2.35 min 15% B, sample loading, 6.60 min 45% B, 7.10 min 95% B with A as water and B as acetonitrile. Fractions that absorbed strongly at 290 nm and 335 nm were collected and pooled, dried *in vacuo* and the remaining solid was weighed. Approximately 7.8 mg of solid was dissolved in 600 μL of DMSO-d_6 and transferred to an NMR tube 507 grade (Norell). The NMR sample was run on a Bruker Avance 400 MHz (DRX400) NMR spectrometer and data was collected for ^{1}H, ^{13}C, HMBC, HMQC, and HSQC experiments. NMR data and peak assignment were carried out using the MestReNova software 12.0.4-22023 release.

Purification of AcfX and Activity Assay

Recombinant AcfX was over-produced in *E. coli* BL21 Star (DE3) as either a C-terminal (AcfX1) or N-terminal His_6-tagged (AcfX2-3) variant – see supporting information for cloning details and sequences. Cells were grown at 28 °C, 200 rpm for 16 hours in autoinduction medium with 25 μg mL^{-1} apramycin for pSF1 C-SP44-*acfX1* or 100 μg mL^{-1} carbenicillin (pET15b-*acfX2-3*). The morning after, cells were collected by centrifugation at 6,000×*g*, 4 °C for 20 min, and re-suspended in binding buffer (20 mM Tris-HCl pH 8, 500 mM NaCl, 5 mM imidazole) and lysed by sonication. Cell-lysates were clarified with centrifugation at 45,000×*g*, 4 °C for 20 min and purified by gravity flow using nickel affinity chromatography (Cytiva). His_6-tagged proteins were washed with increasing concentrations of imidazole (5, 30 and 70 mM) in 20 mM Tris-HCl pH 8, 500 mM NaCl and 15% (v/v) glycerol, before elution with 20 mM Tris-HCl pH 8, 500 mM NaCl, 15% (v/v) glycerol and 200–400 mM imidazole. Protein fractions were analysed by denaturing gel electrophoresis and Coomassie staining as described previously.[50] Purified proteins were buffer-exchanged and concentrated using a centrifugation filter (Amicon MWCO 10,000 Da) to 5 mg/mL in 50 mM Tris-HCl pH 7.7, 14 mM Mg-glutamate, 60 mM K-glutamate and 15% (v/v) glycerol (S30 buffer). Enzyme incubations were set up in S30 buffer lacking 15% (v/v) glycerol as specified in the results and supporting information. Incubations were performed overnight at 4 °C, 21 °C and 30 °C in a glass vial. All reactions included 10% ethanol (v/v) to solubilise the substrate arcyriaflavin A. Reactions were terminated with 1% (v/v) acetic acid and centrifuged at 18,000×*g* for 10 min. 200 μL supernatant was extracted twice with an equal volume of ethyl acetate, which was dried and then re-suspended in 200 μL ethanol. For HPLC analysis, buffer A was water with 0.1% (v/v) trifluoroacetic acid and buffer B was acetonitrile with 0.1% (v/v) trifluoroacetic acid, 2–5 μL of sample was injected onto an ACE 5 AQ C-18 column (4.8 mm×150 mm, 5 μm particle size) with a flow rate of 1 mL/min, oven temperature of 25 °C, and the following gradient (Method B): 0 min 5% B, 5 min 5% B, 20 min 100% B, 25 min 100% B, 26 min 5% B and 31 min 5% B. An analytical standard (dissolved in ethanol) of arcyriaflavin A (Santa-Cruz) was used for comparison. An arcyriaflavin F HPLC standard was prepared from the recombinant cultures and verified by LC–MS.

Supporting Information

Table S1. Oligonucleotides used for cloning of *acfXODCP* or *acfODCP*, and *acfX* expression plasmids.

Table S2. NMR characterisation data for Arcyriaflavin F. Positions are numbered according to Figure 3B.

Figure S1. Absorbance at 320 nm and EIC of extracts from *S. coelicolor* M1152 recombinant strains grown on MS agar.

Figure S2. UV-Vis and MS spectra of arcyriaflavin A and F.

Figure S3. NMR spectra of arcyriaflavin F.

Figure S4. Summary of AcfX1-2 purification.

Figures S5–S10. AcfX reaction controls and time-course reactions.

acfX1-3 and AcfX1-3 sequences.

Acknowledgements

We would like to thank Professor Mervyn Bibb (John Innes Centre) for personal communication on the *S. venezuelae* arcyriaflavin BGC, which is referenced within the manuscript. We would also like to thank Dr Guy Hanke for providing ferredoxin and ferredoxin reductase for enzyme assays (Queen Mary University of London) and Dr Andrew Lawrence (University of Southampton) for discussions on the manuscript. PSF acknowledges funding support from EPSRC (EP/K038648/1, EP/L011573/1). HEL was funded by an Imperial College President's PhD Scholarship. AK was funded as a graduate teaching assistant (University of Kent). SJM acknowledges support from the Leverhulme Trust (RPG-2021-018) and Biotechnology and Biological Sciences Research Council (BB/V007939/1) for funding LT and EB, respectively.

Conflict of Interests

The authors declare no conflict of interest.

Data Availability Statement

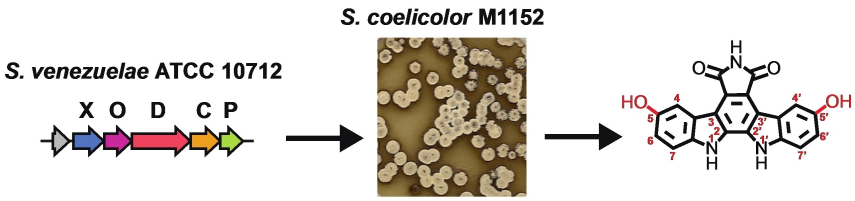
The data that support the findings of this study are available from the corresponding authors upon reasonable request.

Keywords: Arcyriaflavin · Streptomyces · synthetic biology · biosynthetic gene cluster · indolocarbazole

[1] H. Nakano, S. Omura, *J. Antibiot. (Tokyo)* **2009**, *62*, 17–26.
[2] M. Sancelme, S. Fabre, M. Prudhomme, *J. Antibiot. (Tokyo)* **1994**, *47*, 792–798.
[3] K. Kamata, T. Suetsugu, Y. Yamamoto, M. Hayashi, K. Komiyama, M. Ishibashi, *J. Nat. Prod.* **2006**, *69*, 1252–1254.
[4] M. J. Slater, S. Cockerill, R. Baxter, R. W. Bonser, K. Gohil, C. Gowrie, J. E. Robinson, E. Littler, N. Parry, R. Randall, W. Snowden, *Bioorg. Med. Chem.* **1999**, *7*, 1067–1074.
[5] C. Bailly, J. F. Riou, P. Colson, C. Houssier, E. Rodrigues-Pereira, M. Prudhomme, *Biochemistry* **1997**, *36*, 3917–3929.
[6] C. Sánchez, I. A. Butovich, A. F. Braña, J. Rohr, C. Méndez, J. A. Salas, *Chem. Biol.* **2002**, *9*, 519–531.
[7] M. D. Wilkinson, H.-E. Lai, P. S. Freemont, J. Baum, *Antimicrob. Agents Chemother.* **2020**, *64*, DOI: 10.1128/AAC.02129-19.
[8] R. M. Stone, S. J. Mandrekar, B. L. Sanford, K. Laumann, S. Geyer, C. D. Bloomfield, C. Thiede, T. W. Prior, K. Döhner, G. Marcucci, F. Lo-Coco, R. B. Klisovic, A. Wei, J. Sierra, M. A. Sanz, J. M. Brandwein, T. de Witte, D. Niederwieser, F. R. Appelbaum, B. C. Medeiros, M. S. Tallman, J. Krauter, R. F. Schlenk, A. Ganser, H. Serve, G. Ehninger, S. Amadori, R. A. Larson, H. Döhner, *N. Engl. J. Med.* **2017**, *377*, 454–464.
[9] L. Prade, R. A. Engh, A. Girod, V. Kinzel, R. Huber, D. Bossemeyer, *Structure* **1997**, *5*, 1627–1637.
[10] S. Lee, K.-H. Kim, C.-H. Cheon, *Org. Lett.* **2017**, *19*, 2785–2788.
[11] K. Wang, Z. Liu, *Synth. Commun.* **2009**, *40*, 144–150.
[12] S. Saha, A. Banerjee, M. S. Maji, *Org. Lett.* **2018**, *20*, 6920–6924.
[13] M. Ohkubo, T. Nishimura, H. Jona, T. Honma, H. Morishima, *Tetrahedron* **1996**, *52*, 8099–8112.
[14] Y. Men, Z. Hu, J. Dong, X. Xu, B. Tang, *Org. Lett.* **2018**, *20*, 5348–5352.
[15] F.-Y. Chang, M. A. Ternei, P. Y. Calle, S. F. Brady, *J. Am. Chem. Soc.* **2013**, *135*, 17906–17912.
[16] F.-Y. Chang, M. A. Ternei, P. Y. Calle, S. F. Brady, *J. Am. Chem. Soc.* **2015**, *137*, 6044–6052.
[17] L. Li, L. W. Maclntyre, S. F. Brady, *Curr. Opin. Biotechnol.* **2021**, *69*, 145–152.
[18] F.-Y. Chang, S. F. Brady, *Chembiochem.* **2014**, *15*, 815–821.
[19] N. L. Bitzenhofer, T. Classen, K.-E. Jaeger, A. Loeschcke, *Chembiochem.* **2023**, *24*, e202300576.
[20] Z. Zhang, S. Yang, Z. Li, Y. Wu, J. Tang, M. Feng, S. Chen, *Appl. Microbiol. Biotechnol.* **2023**, *107*, 5701–5714.
[21] A. R. Howard-Jones, C. T. Walsh, *J. Am. Chem. Soc.* **2006**, *128*, 12289–12298.
[22] K. S. Ryan, A. R. Howard-Jones, M. J. Hamill, S. J. Elliott, C. T. Walsh, C. L. Drennan, *Proc. Natl. Acad. Sci. U. S. A.* **2007**, *104*, 15311–15316.
[23] A. R. Howard-Jones, C. T. Walsh, *Biochemistry* **2005**, *44*, 15652–15663.
[24] T. Spolitak, D. P. Ballou, *Arch. Biochem. Biophys.* **2015**, *573*, 111–119.
[25] A. R. Howard-Jones, C. T. Walsh, *J. Am. Chem. Soc.* **2007**, *129*, 11016–11017.
[26] L. M. Alkhalaf, K. S. Ryan, *Chem. Biol.* **2015**, *22*, 317–328.
[27] Y. L. Du, K. S. Ryan, *ACS Synth. Biol.* **2015**, *4*, 682–688.
[28] S. Hoshino, L. Zhang, T. Awakawa, T. Wakimoto, H. Onaka, I. Abe, *J. Antibiot.* **2015**, *68*, 342–344.
[29] C. L. Yang, B. Zhang, W. W. Xue, W. Li, Z. F. Xu, J. Shi, Y. Shen, R. H. Jiao, R. X. Tan, H. M. Ge, *Org. Lett.* **2020**, *22*, 4665–4669.
[30] Y. I. Shamikh, A. A. El Shamy, Y. Gaber, U. R. Abdelmohsen, H. A. Madkour, H. Horn, H. M. Hassan, A. H. Elmaidomy, D. H. M. Alkhalifah, W. N. Hozzein, *Microorganisms* **2020**, *8*, 783.
[31] X. Cheng, B. Zhou, H. Liu, C. Huo, W. Ding, *Nat. Prod. Res.* **2018**, *32*, 1–6.
[32] S. J. Moore, H.-E. Lai, S.-M. Chee, M. Toh, S. Coode, K. Chengan, P. Capel, C. Corre, E. L. de Los Santos, P. S. Freemont, *ACS Synth. Biol.* **2021**, *10*, 402–411.
[33] S. J. Moore, H.-E. Lai, H. Needham, K. M. Polizzi, P. S. Freemont, *Biotechnol. J.* **2017**, *12*, 1600678.
[34] C. L. M. Gilchrist, Y.-H. Chooi, *Bioinformatics* **2021**, *37*, 2473–2475.
[35] K. Blin, S. Shaw, K. Steinke, R. Villebro, N. Ziemert, S. Y. Lee, M. H. Medema, T. Weber, *Nucleic Acids Res.* **2019**, *47*, W81–W87.
[36] J. L. Doull, A. K. Singh, M. Hoare, S. W. Ayer, *J. Ind. Microbiol.* **1994**, *13*, 120–125.
[37] J. Doull, Z. Ahmed, C. Stuttard, L. C. Vining, *J. Gen. Microbiol.* **1985**, *131*, 97–104.
[38] Y. Inahashi, S. Zhou, M. J. Bibb, L. Song, M. M. Al-Bassam, M. J. Bibb, G. L. Challis, *Chem. Sci.* **2017**, *8*, 2823–2831.
[39] S. Kodani, K. Sato, H. Hemmi, M. Ohnish-Kameyama, *J. Antibiot.* **2014**, *67*, 839–842.
[40] J.-E. Kim, J.-S. Choi, J.-H. Roe, *J. Microbiol.* **2019**, *57*, 388–395.
[41] J. P. Gomez-Escribano, N. A. Holmes, S. Schlimpert, M. J. Bibb, G. Chandra, B. Wilkinson, M. J. Buttner, M. J. Bibb, *J. Ind. Microbiol. Biotechnol.* **2021**, *48*, DOI 10.1093/jimb/kuab035.
[42] T. Kieser, M. J. Bibb, M. J. Buttner, *Practical Streptomyces Genetics*, John Innes Foundation, **2000**.
[43] L. Van Mellaert, L. Mei, E. Lammertyn, S. Schacht, J. Ann, *Microbiology* **1998**, *144(Pt 12)*, 3351–3358.
[44] R. M. Phelan, D. Sachs, S. J. Petkiewicz, J. F. Barajas, J. M. Blake-Hedges, M. G. Thompson, A. Reider Apel, B. J. Rasor, L. Katz, J. D. Keasling, *ACS Synth. Biol.* **2017**, *6*, 159–166.
[45] C. Sánchez, L. Zhu, A. F. Braña, A. P. Salas, J. Rohr, C. Méndez, J. A. Salas, *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102*, 461–466.
[46] M. M. E. Huijbers, S. Montersino, A. H. Westphal, D. Tischler, W. J. H. van Berkel, *Arch. Biochem. Biophys.* **2014**, *544*, 2–17.
[47] S. Sharma, M. R. Schiller, *Crit. Rev. Biochem. Mol. Biol.* **2019**, *54*, 85–102.
[48] P. Wang, G. Bashiri, X. Gao, M. R. Sawaya, Y. Tang, *J. Am. Chem. Soc.* **2013**, *135*, 7138–7141.
[49] Y.-L. Du, T. Ding, K. S. Ryan, *Org. Lett.* **2013**, *15*, 2538–2541.
[50] S. J. Moore, H.-E. Lai, R. J. R. Kelwick, S. M. Chee, D. J. Bell, K. M. Polizzi, P. S. Freemont, *ACS Synth. Biol.* **2016**, *5*, 1059–1069.

Manuscript received: April 17, 2024
Revised manuscript received: July 12, 2024
Accepted manuscript online: July 22, 2024
Version of record online: ■■, ■■

RESEARCH ARTICLE



We report the characterisation of the *Streptomyces venezuelae* ATCC 10712 arcyriaflavin F biosynthetic gene cluster. We express the gene cluster using *Streptomyces coelicolor* M1152 as a heterologous host, isolate the product and confirm distinct C-5/C-5′ indole ring hydroxylation catalysed by a regioselective flavoprotein AcfX.

H.-E. Lai, A. Kennedy, L. Tanner, E. A. Bartram, S. Mei Chee, P. S. Freemont, S. J. Moore**

1 – 8

Biosynthesis of Arcyriaflavin F from *Streptomyces venezuelae* ATCC 10712