A semi-synthetic strategy for derivatization of the violacein natural product scaffold

bioRxiv manuscript

Hung-En Lai[1,3], Soo Mei Chee[1,3], Marc Morgan[2], Simon Moore[1,3], Karen Polizzi[2,3], Paul Freemont*[1,3]

[1]Department of Medicine, Imperial College London, London SW7 2AZ, UK.

[2]Department of Life Sciences, Imperial College London, London SW7 2AZ, UK.

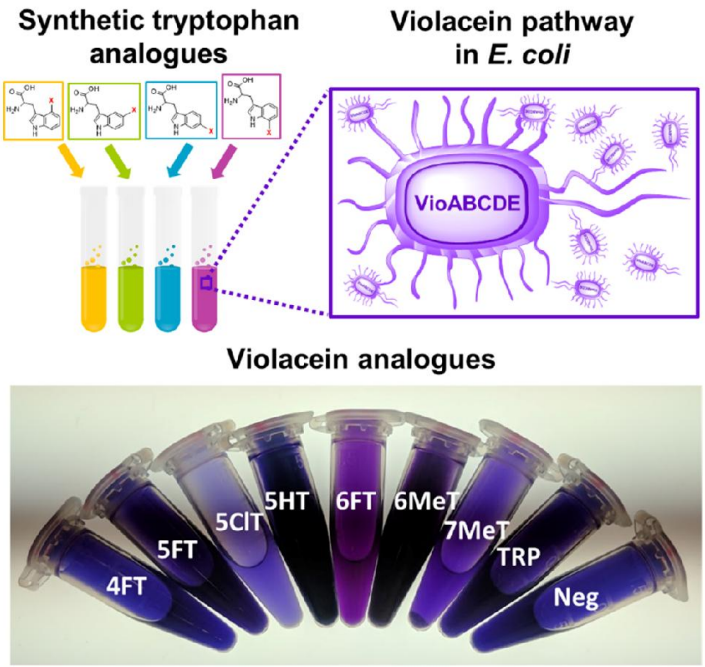
[3]Centre for Synthetic Biology and Innovation, Imperial College London, London SW7 2AZ, UK.

* To whom correspondence should be addressed (p.freemont@imperial.ac.uk)

Keywords

Violacein, antibiotics biosynthesis, natural product analogue, rebeccamycin, bisindole, indolocarbazole

Graphical Abstract



Abstract

The next frontier in drug discovery could be the semi-synthesis of non-natural, xenobiotic compounds combining both natural product biosynthesis and synthetic chemistry. However, the required tools and underlying engineering principles are yet to be fully understood. One way to investigate non-natural product biosynthesis is to probe the substrate promiscuity of a clinically relevant biosynthesis pathway. Violacein is a bisindole compound produced by the VioABCDE biosynthesis pathway using L-tryptophan as the starting substrate. Previous studies have shown that violacein exhibits antimicrobial properties, and synthetic analogues of violacein might give rise to new targets for therapeutic development to combat antimicrobial resistance. By adding seven types of tryptophan analogues available commercially, 62 new violacein or deoxyviolacein analogues were generated with a synthetic violacein biosynthesis pathway expressed in *Escherichia coli*, demonstrating the promiscuity of violacein biosynthesis enzymes. Growth inhibition assays against *Bacillus subtilis*, a Gram-positive bacterium, were carried out to measure growth inhibitory activity of violacein analogues compared to violacein. In addition, we show that four new 7-chloro analogues of violacein or deoxyviolacein can be generated *in vivo* by combining the rebeccamycin and violacein biosynthesis pathways and purified 7-chloro violacein was found to have similar growth inhibitory activity compared to violacein. Structural studies of VioA revealed active site residues that are important for catalytic activity, and further pathway recombination with VioA homologues in related bisindole pathways may lead to more efficient enzymes that would accept tryptophan analogues more readily.

Introduction

Violacein is a violet pigment first isolated from the bacterium, *Chromobacterium violaceum* and is part of the bisindole biosynthetic family that utilises L-tryptophan as the starting substrate. The violacein pathway is encoded within a conserved operon of five genes (*vioABCDE*), whose gene-products catalyse a 14-electron oxidative biosynthesis pathway[1] (see **Figure 1A** for pathway details). The bisindole biosynthetic pathways have attracted considerable interest because of their therapeutic potential for medical applications, including antimicrobial, antiviral, trypanocidal and anti-tumorigenic properties[2]. The violacein biosynthesis pathway also produces deoxyviolacein (DV) as a by-product, and the coloured properties of violacein and deoxyviolacein make them interesting targets for natural product pathway engineering. For example, high-throughput combinatorial promoter library screening[3], CRISPR-based multiplex transcriptional regulation[4] or diverting pathway flux from deoxyviolacein to violacein via RBS engineering[5].

Total synthesis of natural products and their analogues, especially for structure-activity relationship (SAR) studies, has often been difficult and cost-ineffective due to their structural complexities[6]. However, given the larger chemical space accessible by total synthesis compared to enzyme-catalysed biosynthesis, a combination of both synthetic routes (also known as semi-synthesis) to generate drug screening libraries has become increasingly common[7]. More recently, analogues of a 'last-resort' antibiotic vancomycin had been found to have >200-fold improvement of potency compared to vancomycin in resistant *Enterococci* strains[8]. In another study, the addition of a sterically unhindered primary amine group to a Gram-positive antibiotic deoxynybomycin expanded its antimicrobial activity to multi-drug-resistant Gram-negative pathogenic strains[9]. Therefore, we hypothesised that violacein analogues with similar features could also have the potential to have improved activity against resistant strains.

A study on oxyviolacein, a hydroxylated analogue of violacein generated by feeding exogenous 5-hydroxy-tryptophan[10], and its bioactivity against a collection of pathogenic

bacteria and fungi strains[11] revealed that violacein analogues may be a good starting point to develop more potent antibiotics. Although total chemical synthesis of violacein has been reported[12,13], there is no literature to date on generating other violacein analogues or derivatives either through chemical synthesis or biosynthesis. In addition, VioA, the first enzyme in the violacein biosynthesis pathway, has been shown to react with different tryptophan analogues, with 5-methyl-tryptophan having a slightly higher activity (33%) compared to the cognate substrate L-tryptophan[14], demonstrating potential for analogue generation using the violacein biosynthesis pathway.

Based on these previous studies, we hypothesize that specific violacein analogues can be generated by corresponding chemically synthesised tryptophan analogues with a heteroatom or chemical group (such as chloro, fluoro, methyl or hydroxyl) attached at the C4-, C5-, C6- or C7- positions of the indole ring of tryptophan. These sites are selected as they are far away from the reaction active site[15] (alpha-carbon and its adjacent amine), therefore they should not impede enzymatic conversion of tryptophan to violacein.

There is also increasing interest in the closely related bisindole biosynthesis pathways[16–18] due to the prospect of biosynthesizing clinically relevant metabolites[19], with rebeccamycin (Reb)[20–22] and staurosporine (Sta)[23–25] as the most characterized pathways. Depending on the chemical structure of their core scaffold, bisindole compounds can be classified into indolocarbazoles (including rebeccamycin and staurosporine), bisindolylmaleimide (methylarcyriarubin) and indolotryptolines (such as cladoniamide and borregomycin)[17] (**Figure S1A**). All bisindole pathways use chromopyrrolic acid (CPA) as a common biosynthetic precursor, with the exception of violacein pathway which employs a unique VioE enzyme that forms the violacein 2-pyrrolinone core[26,27]. Although most bisindole pathways share the conserved ODCP genes that construct the indolocarbazole scaffold (**Figure S1B**), some pathways such as the rebeccamycin pathway also contain a halogenase (RebH) and a flavin reductase (RebF) that are responsible for generating 7-chloro-L-tryptophan from L-tryptophan via a chloramine-lysine intermediate[28,29]. In addition, due to the functional similarity of

tryptophan oxidase and CPA synthase in these pathways, RebO and RebD were shown to reconstitute violacein production by substituting VioA and VioB respectively in the violacein pathway[1]. Other studies have demonstrated the potential of generating novel bisindole compounds by recombining enzyme homologues from various bisindole pathways[30,31], therefore we anticipated that new violacein analogues could also be generated via a similar strategy (**Figure 1B**).

Result and Discussion

Tryptophan analogues as starting substrates to generate various violacein analogues

Violacein analogues were generated from commercially available tryptophan analogues as starting substrates, including 4-fluoro-, 5-fluoro-, 6-fluoro-, 5-chloro-, 5-hydroxyl-, 6-methyl- and 7-methyl-tryptophan analogues. By adding the tryptophan analogues to *E. coli* cells that were expressing a synthetic violacein biosynthesis pathway *vioABCDE*, different proportions of violacein analogues were generated depending on the type of tryptophan analogue added (**Figure 2A**). Hybrid analogues where more than one type of heteroatom is incorporated into the product (for example chloro-fluoro-violacein or fluoro-methyl-violacein) were also observed by adding two different tryptophan analogues to the same culture (**Figure S2**). A total of 62 new compounds were observed including structural isomers due to the asymmetry of the violacein core scaffold, demonstrating promiscuity of violacein biosynthesis enzymes (full structures of all analogues are available in **Table S1**).

In some cases, structural isomers of mono-substituted analogues are distinguished by a specific fragment in MS/MS spectra (**Table S2**). For example, 6”-fluoro-violacein (fluorine attached to indole) was identified by fragment with *m/z* = 183, while 6’-fluoro-violacein (fluorine attached to keto-indole) was identified by fragment with *m/z* = 201, which is the same fragment with fluorine added. This is due to the subtle difference in fragmentation pattern between analogues with substituted group attached to the indole or keto-indole moiety. The same dichotomous fragmentation pattern was observed for other monosubstituted analogues, such as chlorine (*m/z* = 217), methyl (*m/z* = 197) and hydroxyl (*m/z* = 199). For di-substituted isomers with different groups, the group attached to the keto-indole moiety was used in identifying the fragment, for example 5”-fluoro-6’-methyl-deoxyviolacein would be identified by fragment with *m/z* = 197 whereas 5’-fluoro-6”-methyl-deoxyviolacein by fragment with *m/z* = 201.

The percentage compositions of the various violacein and deoxyviolacein analogues were calculated based on the peak area obtained from LC-MS analyses (**Figure 2B**). Among the seven tryptophan analogues tested, 6-fluoro-tryptophan (6FT) has the highest total analogue percentage composition (73.6 ± 25.9%). While 5-chloro-tryptophan (5ClT), 6-methyl-tryptophan (6MeT) and 7-methyl-tryptophan (7MeT) generated similar total analogue percentage composition (between 44.1% and 47.2%), 4-fluoro-tryptophan (4FT) was by far the least proportion of total analogues (11.8 ± 5.4%). These data suggested that 6-fluoro tryptophan analogues can be utilized as effective substrates by the violacein pathway compared to 4-fluoro analogues. This could be due to either differences in binding specificities between the different analogues and violacein enzymes (see below) or stability of the modified pathway intermediates. There were also more variants observed for deoxyviolacein than violacein (**Table S1**), and the percentage composition of the deoxyviolacein analogues were generally higher than that of violacein analogues (**Figure 2B**). This could be because VioD, the key enzyme that directs the pathway towards violacein being more selective than VioC, which produces deoxyviolacein, although this would need further study. However, our data clearly show that a wide variety of different tryptophan analogues can be used as substrates to generate modified versions of violacein and deoxyviolacein.

Growth inhibition assays of *B. subtilis* using extracts containing violacein analogues

The activity of the crude violacein extracts containing the various analogues and a control containing only violacein and deoxyviolacein (Vio/DV) were assayed in a 96-well plate using *B. subtilis* cultures adjusted to the same OD at 600 nm. We found that all extracts have similar minimal inhibitory concentration (MIC) of 0.5 ng/μL compared to the positive control tetracycline at 5 ng/μL (not shown). Further screening of concentrations between 0.05 ng/μL to 0.5 ng/μL of crude extracts of violacein analogues revealed that none of the crude extracts have higher activity (or lower IC50) against *B. subtilis* compared to the extract containing

unmodified crude violacein (Vio/DV), as shown in **Figures 2C and 2D**. However, because of the presence of mixture of analogues in each crude extract, it was not possible to assign growth inhibitory activity to a specific analogue in the extracts. Furthermore, most analogues from any single extract had very close retention times (**Table S2**), making separation of these analogues difficult by reverse-phase liquid chromatography. Interestingly, the 6MeT extract has about 4-fold higher IC50 compared to a 7MeT extract, where the only difference between these two samples is the position of methyl group attached to the indole ring, suggesting that the C6 position might be important for violacein activity. In addition, cell extracts containing only deoxyviolacein (DV) did not show any significant inhibition towards *B. subtilis*, suggesting that the 5-hydroxyl group of violacein has an important function in relation to antimicrobial activity.

Hybrid pathway of Rebeccamycin-Violacein generates 7-chloro violacein analogues

Due to the promiscuity of violacein biosynthesis pathway enzymes, we rationalized that a combination of enzymes from closely related pathways could generate compatible tryptophan analogues leading to other interesting non-natural analogues of violacein that might be otherwise difficult to synthesise. As a proof-of-principle, by combining the rebeccamycin (Reb) and violacein (Vio) biosynthesis pathways using the Golden Gate cloning kit EcoFlex[32], 7-chlorinated violacein analogues were successfully produced in *E. coli* cells (**Figure 3A**) and identified via LC-MS/MS (**Figure 3B**). A kanamycin resistance gene was also included to improve plasmid stability with dual antibiotic selection. Unsurprisingly, we found that although the RebOD+VioCDE strain (lacking the RebH halogenase and RebF reductase) did not produce any chlorinated analogues of violacein, it was able to functionally substitute VioAB to produce violacein and deoxyviolacein, as reported previously[1]. Next, by comparing RebFH+VioABCDE and RebFHOD+VioCDE (with RebOD replacing VioAB) strains, we observed that RebOD increased the proportion of chlorinated analogues from 6.1 ± 0.5% to 32.9 ± 1.2% when compared with VioAB (**Figure 3C**). This is not surprising as RebO has 57-

times higher k_{cat}/K_m value for 7-Cl-L-tryptophan than that of L-tryptophan[33], thus RebO and RebD preferentially accept the 7-chloro analogue of tryptophan and IPA imine respectively, passing on the 7-chlorinated intermediate to downstream enzymes in the violacein biosynthesis pathway.

Next, we rationalized that higher amounts of chlorinated tryptophan will increase the proportion of chlorinated analogues, so we changed the strength of the individual promoters controlling expression of RebF, RebH and VioD, to either a weak (BBa_J23114) or medium (BBa_J23108)[34] strength promoter. Interestingly, this adjustment led to an increase in the proportion of chlorinated analogues from 28.6 ± 6.4% for weak strength promoters (annotated WWW strain in **Figure 3D**) to 56.5 ± 4.3% for medium strength promoters (annotated MMM strain in **Figure 3D**), although the total peak area of chlorinated analogues decreased by approximately 13.2% (**Figure 3D**). The large error in peak area could be due to variations in pathway expression profile between different biological replicates, possibly as a result of the presence of medium-strength promoters causing metabolic burden to cells[35]. However, by expressing VioD using the weak promoter, but RebF and RebH with medium promoters (annotated MMW strain in **Figure 3D**), the proportion of chlorinated analogues further increased to 60.2 ± 7.7%, although the total peak area of chlorinated analogues decreased by 22.1% compared to the WWW strain. This shows that it is possible to fine tune the proportion of chlorinated analogue by changing the promoter strength of pathway enzymes, and optimization of fermentation protocols such as length of incubation, temperature, media conditions might further increase both proportion and the total amount of chlorinated analogues[3].

Purification of 7-chloro analogues of violacein and deoxyviolacein

To probe the activity of 7-chloro analogues of violacein, we needed to produce sufficient amounts of the compounds to allow purification of the analogues. Due to their coloured nature, we observed visible separation of the compounds on a reverse-phase C18 column (**Figure S4A**). Violacein is the most polar molecule among the four, therefore it eluted the quickest,

followed by chloro-violacein, chloro-deoxyviolacein and dichloro-deoxyviolacein. The elution fractions were pooled as necessary, and the purity of each fraction was then assayed by LC-MS/MS. We found that a second purification step (by pooling all relevant fractions and repeating the purification) was necessary to obtain purified chloro-violacein and chloro-deoxyviolacein. The two chloro-deoxyviolacein isomers, each having the chlorine atom attached to either side of the molecule, did not separate well enough even after the second purification step due to close retention times (**Figure S4B**). A preliminary assay of purified 7-chloro analogue of violacein did not show any increased activity compared to purified violacein (**Figure S3C**), suggesting that the 7-chloro group does not enhance the growth inhibitory activity of violacein.

VioA crystal structure and multiple sequence alignment with VioA homologues reveal conserved residues important for catalytic activity

The promiscuity of violacein biosynthesis pathway enables the generation of various analogues of violacein from non-natural tryptophan substrates. Compared to RebO, VioA has a 825-fold higher catalytic efficiency for L-tryptophan (k_{cat}/K_M = 6548 min^{-1} mM^{-1})[1] than RebO (k_{cat}/K_M = 7.94 min^{-1} mM^{-1})[33], therefore a structural insight of the VioA active site in the context of catalytic activity would be interesting to determine the basis of these differences. To this end, we purified VioA with an N-terminal hexahistidine tag to homogeneity (**Figure S5**) and determined the crystal structure of VioA at 2.6 Å in complex with L-tryptophan (PDB Accession **6ESE, Figure 4A**). L-tryptophan, the VioA substrate, was found to be close to the isoalloxazine rings of FAD, with the nearest interatomic distance at 4.0 Å, similar to other flavoenzymes[36] (**Figure S6**). At the substrate-FAD interface, two key residues, Arg64 and Tyr309, were close to the charged carboxylate group of the substrate, positioning the Cα atom next to the FAD isoalloxazine plane (**Figure 4B**). However, a previous study found that while Arg64 and Tyr309 are essential for catalytic activity, His163 is unnecessary as a VioA H163A mutant retained 85% of WT activity, suggesting that the hydride transfer from Cα and amine groups of L-tryptophan may not require acid-base catalysis of the H163 side chain[14].

The active site of VioA also reveals the residues surrounding the C4- to C7-positions of tryptophan substrate indole ring. Of the four positions, C4 and C7 has the most space around the binding site pocket compared to C5 and C6 due to Asp311 and Ala145 side chains (**Figure 4C**), suggesting that tryptophan analogues with substituent groups at the C5- and C6-positions would have the lowest affinity for VioA. Surprisingly, C5- and C6-substituted tryptophan analogues tested in this study, such as 5-fluoro-tryptophan and 6-fluoro-tryptophan, were shown to have higher total analogues percentage composition than 4-fluoro-tryptophan (**Figure 2B**). In addition, 5-methyl-tryptophan was shown to have higher affinity than L-tryptophan in a previous study[14], suggesting that C5-substituted tryptophan analogues may bind to the VioA active site with alternative conformation which confers higher catalytic turnover, although further studies would be needed to validate this hypothesis.

Multiple sequence alignment of 14 VioA homologues from various bisindole pathways that were characterized previously, shows that while Arg64 and Tyr309 are conserved among the VioA homologues, His163 is replaced by Asn in HysO and AbeO (**Figure S7**), supporting the hypothesis that His163 is not required for catalytic activity of VioA in a previous study[14]. It would be interesting to compare the reaction kinetics of VioA homologues from other bisindole pathways, such as StaO (staurosporine pathway)[23], EspO (erdasporine pathway)[37] and HysO (hydroxysporine pathway)[38], against several tryptophan analogues in order to identify key active site residues that either accept or inhibit analogue binding. In addition, combining the most efficient VioA and VioB homologues with the rest of violacein biosynthesis pathway enzymes may also result in improved activity or flux of substrate analogues that are not readily accepted by VioA and VioB. For tryptophan analogues with C5- and C6-substituted groups, combining VioA homologues that natively accepts 5-chloro-tryptophan (such as ClaO, cladoniamide pathway[39,40]) and 6-chloro-tryptophan (BorO, borregomycin pathway[41]) and their corresponding CPA synthases (ClaD, BorD), flavin reductases (ClaF, BorF) and halogenases (ClaH, BorH) in a hybrid pathway could generate respective chlorinated analogues of violacein, further expanding the diversity of non-natural violacein analogues.

A structure of VioA in complex with a competitive inhibitor 2-[(1H-indol-3-yl)methyl]prop-2-enoic acid (IEA) has been published independently (PDB accession code **5G3U**)[14], with slightly different crystallization conditions and crystallographic space group. In this study, all VioA active site mutants exhibited decreased activity with most tryptophan analogues tested compared to wild-type VioA, demonstrating the difficulty in rationally engineering VioA towards increased substrate promiscuity[14]. Given that biosynthesis pathways often consist of multiple enzymatic cascades to generate the final products, rational engineering of all the enzymes involved in the pathway may not be the most efficient strategy for non-natural product biosynthesis. Therefore, alternative methods such as directed evolution[42], mutagenesis[25,43], pathway recombination[30,44] or semi-synthesis[6,45,46] strategies for generating natural product analogues should be systematically explored in greater detail.

Conclusion

We have shown that the violacein biosynthesis pathway is promiscuous, accepting various tryptophan analogues to generate 62 new violacein or deoxyviolacein analogues that have not been described thus far (except oxyviolacein[10,11]). Although an initial screen of crude extracts containing various analogues showed no higher growth inhibitory activity against *B. subtilis* than extract containing only violacein and deoxyviolacein, the diverse range of analogues could be used as biosynthetic precursors for semi-synthesis in drug discovery pipelines. In addition, we also combined enzymes from the rebeccamycin and violacein biosynthesis pathways to construct a hybrid pathway that first generates 7-chloro-tryptophan in *E. coli*, and subsequently 7-chloro analogues of violacein and deoxyviolacein. We demonstrated that it is possible to increase the proportion of 7-chloro analogues by increasing the promoter strength of key enzymes in the rebeccamycin-violacein hybrid pathway, and showed that purified 7-chloro violacein has similar growth inhibitory activity against *B. subtillis* compared to purified violacein. We also determined the crystal structure of VioA to identify residues that might be important for VioA catalytic activity and binding of tryptophan analogues. We speculate that

further pathway recombination using enzyme homologues from related bisindole pathways could lead to further biosynthesis of diverse violacein analogues.

Materials and Methods

Molecular Biology

Restriction enzymes (NEB), T4 DNA ligase (Promega), Phusion (Agilent) and Q5 DNA polymerase (NEB) were used for routine cloning. QIAprep Miniprep Kit (Qiagen) and Zymoclean Gel DNA Recovery Kit (Zymogen) were used for plasmid extraction and purification. Primers were synthesized by Integrated DNA Technologies and sequencing was performed by Eurofins Genomics. Chemically competent DH10B (NEB) or JM109 (Promega) *Escherichia coli* strains were used for plasmid maintenance, cloning and expression of pathway enzymes. Depending on the plasmid antibiotic resistance marker, LB broth supplemented by carbenicillin (50 µg/mL), chloramphenicol (35 µg/mL) and kanamycin (50 µg/mL) was used unless stated otherwise.

Culture growth and violacein extraction protocol

For test cultures to check for presence of novel analogues, typically 10 mL of autoclaved LB media in 50 mL sterile tubes were used. Overnight culture was diluted 1:100 with fresh media and grown for 4 to 6 hours at 37 ºC until reaching OD 0.5, and 0.04% (w/v) tryptophan analogue (dissolved in 0.5 M HCl) and equal volume of 0.5 M NaOH were added. Cultures were then incubated at 25 °C for up to 65 hours. Cells were harvested by centrifugation (4000 x *g*, 4 °C, 10 min) and supernatant was removed. For each 10 mL of culture grown, 1 mL of ethanol was added to the cell pellet, and the mixture was transferred into smaller tubes and incubated at 75 °C for about 30 min, or until cell pellet turns greyish-white. The mixture was then centrifuged (15,000 x *g*, RT, 1 min) and the supernatant transferred to fresh tube for storage at -20 °C, or diluted 1:50 with ethanol for injection into LC-MS/MS for detection of analogues. For long-term storage, the ethanol extract was dried down under vacuum and resuspended in DMSO, and then stored at -20 °C. Concentration was measured by comparing absorbance at 575 nm using purified violacein (Sigma) as calibration standards.

LC-MS/MS detection of violacein analogues

Violacein analogues were detected by Agilent 1290 LC and 6550 Q-ToF mass spectrometer with electrospray ionization, using Agilent Zorbax Extend C-18 column (2.1 x 50 mm, 1.8 µm particle size). For LC, buffer A was water with 0.1% (v/v) formic acid and buffer B was acetonitrile with 0.1% (v/v) formic acid, for total running time of 13 minutes including 3 min post-run time. 2 µL diluted samples were injected, and the column was washed at 0.35 ml/min with the following gradient: 0 min 5% B, 1 min 5% B, 9 min 70% B, 10 min 70% B. Quantification was by peak area of integrated chromatograms at targeted $[M+H]^+$ mass with ±20 ppm window. The protonated molecules of each analyte, $[M+H]^+$, were targeted and subjected to collision-induced dissociation (collision energy 30 eV), with product ions accumulated throughout the analysis. In calculating percentage composition of analogues, all compounds were assumed to have the same mass response as there is no mass spectrometry standard available for every compound detected.

Growth inhibition bioassays against *B. subtilis*

100 µL overnight culture of *B. subtilis* is inoculated into 6 mL of fresh LB and grown for ~3 h until OD600 reaches around 0.3 to 0.5. In the meantime, serial dilutions of crude violacein extracts were prepared with DMSO to obtain a range of concentrations between 1 to 10 ng/µL. Concentration of crude violacein extract stocks were calibrated against Violacein standard (Sigma) using absorbance at 575 nm. 16.5 µL of crude violacein extract at each concentration is added to 313.5 µL of *B. subtilis* culture diluted with fresh LB to OD 0.1 to adjust final concentration of drugs to 0.05 to 0.5 ng/µL while keeping DMSO concentration constant at 5% (w/v). 100 µL of mixture was then transferred in triplicate wells in a 96-well flat-bottom plate (Greiner), and Breathe-easy sealing membrane (Sigma) was applied to minimize liquid evaporation and allow gaseous exchange. *B. subtilis* culture with 5% DMSO (without drug) was also prepared as negative control. The OD was measured at 600 nm at 10 min intervals over 18 hours, and the incubation chamber temperature was set at 30 °C throughout.

Experiments were performed on three separate days to obtain biological triplicates. Endpoint (18 h) OD600 values were normalised against negative control and plotted against drug concentration. Non-linear regression fit analysis was performed using GraphPad Prism version 7.03 for Windows, GraphPad Software, La Jolla California USA, www.graphpad.com. Specifically, the inhibitor concentration vs response - variable slope (four parameters) with robust fitting method was chosen, and IC50 and their associated RSDR (Robust Standard Deviation of the Residuals) values were obtained from the analysis and plotted for each crude violacein extract tested.

Constructing Rebeccamycin/Violacein hybrid pathway

The Rebeccamycin and Violacein pathway genes are available from UniProt databases. Violacein genes were PCR from Bba_K274002 plasmid (iGEM Parts Registry). RebF (Q8KI76), RebH (Q8KHZ8), RebO (Q8KHS0) and RebD (Q8KHV6) were synthesised by a commercial company (GeneArt, Thermo Fisher) and cloned into appropriate holding vectors. Assembly of complete pathway consisting of all genes (RebF, RebH, RebO, RebD, VioA, VioB, VioC, VioD, VioE) was done using a modular cloning toolkit EcoFlex[32], and is available on Addgene (Kit #1000000080). Briefly, all genes, including a kanamycin resistance marker KanR, were cloned into holding vector pBP-ORF, and then assembled into individual transcription units with standardised promoter (BBa_J23114 or BBa_J23108), pET-RBS and BBa_B0015 terminator. These transcription units (Level 1 plasmids) were then assembled into two Level 2 sub-pathway plasmids (*e.g.* RebaFHOD and VioCDE+KanR), and subsequently the final complete pathway into a Level 3 plasmid (plasmid and protein sequences available in **SI Data**). For full assembly protocol please refer to **SI Protocol**.

Purification of violacein analogues

Crude extract of violacein analogues was concentrated *in vacuo* with a rotary evaporator (Heidolph), and loaded on a pre-equilibrated Biotage SNAP Ultra C18 column 12 g attached to Biotage Isolera Spektra system. Buffer A was ultrapure water (HPLC grade) and buffer B was acetonitrile (HPLC grade), for total running time of 16.5 minutes including 2.4 min equilibration time. The column was washed with acetonitrile and methanol before each run, and run at 36 mL/min with the following gradient: 0 min 10% B, 2.4 min 25% B, sample loading, 14.2 min 65% B, 14.7 min 95% B. About 0.5 mL concentrated ethanol extract was loaded directly onto the column using a 1 mL syringe at 2.4 min (after equilibration). Relevant fractions containing coloured compounds were pooled and concentrated *in vacuo*, dried under nitrogen gas and assayed using the same LC-MS/MS method as described above.

Agar inhibition assay

Overnight culture of *B. subtilis* was inoculated into fresh LB and grown at 37 °C for about 3 hours until OD600 reaches 0.5, and then diluted with molten LB agar cooled to ~45 °C to a final OD600 of 0.1, poured onto plates and let air-dried under fume hood for 2 h. 5 µL droplets of violacein or 7-chloro violacein at 0, 1.25, 2.5, 5, 10 and 20 ng/µL in DMSO was added onto the agar, and air-dried under fume hood for 30 min before incubating at 37 °C overnight. Activity of compound was assayed based on diameter of growth inhibition zone.

VioA purification and crystallisation

The *vioA* gene was amplified and cloned between NdeI and BamHI sites of pET-15b expression vector (Novagen) which contained an N-terminal hexahistidine tag. The vector was transformed into BL21 DE3 pLysS *Escherichia coli* (NEB) strain, and grown in 1 L LB broth supplemented with 50 µg/mL ampicillin at 37 °C until OD_{600} reached 0.6, and then induced with 0.2 mM IPTG and further incubation at 20 °C for 18 hours. Cells were then pelleted by centrifugation (4k rpm, 4 °C, 20 min) and resuspended in Binding Buffer (20 mM Tris pH 8, 500 mM NaCl, 10 mM Imidazole). Cells were sonicated (2 s on 2 s off at 70% amplitude for 4 min) and cell debris pelleted by centrifugation (15k rpm, 4 °C, 30 min). The supernatant was subjected to affinity purification by injecting into HisTrap HP column (GE Healthcare), with Buffer A as Binding Buffer and Buffer B as Elution Buffer (20 mM Tris pH 8, 500 mM NaCl, 500 mM Imidazole). Elution fractions were checked with SDS-PAGE, and fractions containing most VioA (49 kD) were pooled and concentrated with Amicon Ultra-15 10K MWCO concentrator (EMD Milipore) and desalted with PD-10 Desalting Column (GE Healthcare) into GF Buffer (20 mM CHES pH 9.0, 25 mM NaCl). Protein solution was then subjected to SEC by injecting into HiLoad 16/60 Superdex 200 column (GE Healthcare) equilibrated with GF Buffer. Peak fractions were pooled and concentrated to 5 mg mL^{-1} measured on spectrophotometer with molar extinction coefficient from ProtParam (ExPASy). Concentrated proteins were then screened initially with sparse matrix (Crystal Screen 1 & 2 from Hampton Research), and hit condition was optimized in 24-well sitting drop plates with protein-to-buffer ratio 2:1. Best crystals were obtained with 100 mM HEPES pH 7.5, 9% PEG 8000 and 9% ethylene glycol incubated at 20 ˚C.

VioA and L-tryptophan-bound VioA structure determination

VioA crystal was fished and cryoprotected with 20% ethylene glycol, and then snap frozen in liquid nitrogen before sending to Diamond Light Source for X-ray diffraction. L-tryptophan-soaked crystals were prepared by immersing crystal in mother liquor with added 1 mM L-tryptophan for 10 to 60 minutes at room temperature. Datasets collected from DLS were processed with in-house DIALS/Xia2 algorithm[47], and phase information was solved by

PHASER[48] molecular replacement with a model structure with PDB ID **3X0V** (L-lysine oxidase, 17% sequence identity). Further rounds of refinement and model building was done in Phenix[49] and *Coot*[50] respectively (see **Table S3** for data collection and refinement statistics). L-tryptophan-bound VioA structure was solved using molecular replacement with the apoenzyme structure as model. Protein structures are visualised using PyMOL Molecular Graphics System, Version 1.8 Schrödinger, LLC. Sequence alignment was carried out using ESPript (http://espript.ibcp.fr)[51].

References

1. Balibar, C. J. & Walsh, C. T. In Vitro Biosynthesis of Violacein from L-Tryptophan by the Enzymes VioA−E from *Chromobacterium violaceum*. *Biochemistry* **45,** 15444–15457 (2006).

2. Durán, M. *et al.* Potential applications of violacein: A microbial pigment. *Med. Chem. Res.* **21,** 1524–1532 (2012).

3. Jones, J. A. *et al.* ePathOptimize: A Combinatorial Approach for Transcriptional Balancing of Metabolic Pathways. *Sci. Rep.* **5,** 11301 (2015).

4. Zalatan, J. G. *et al.* Engineering complex synthetic transcriptional programs with CRISPR RNA scaffolds. *Cell* **160,** 339–350 (2015).

5. Jeschek, M., Gerngross, D. & Panke, S. Rationally reduced libraries for combinatorial pathway optimization minimizing experimental effort. *Nat. Commun.* **7,** 11163 (2016).

6. Kirschning, A. & Hahn, F. Merging chemical synthesis and biosynthesis: A new chapter in the total synthesis of natural products and natural product libraries. *Angew. Chemie - Int. Ed.* **51,** 4012–4022 (2012).

7. Eichner, S. *et al.* The interplay between mutasynthesis and semisynthesis: Generation and evaluation of an ansamitocin library. *Angew. Chemie - Int. Ed.* **51,** 752–757 (2012).

8. Okano, A., Isley, N. A. & Boger, D. L. Peripheral modifications of [Ψ[CH2NH]Tpg4]vancomycin with added synergistic mechanisms of action provide durable and potent antibiotics. *Proc. Natl. Acad. Sci.* 201704125 (2017). doi:10.1073/pnas.1704125114

9. Richter, M. F. *et al.* Predictive compound accumulation rules yield a broad-spectrum antibiotic. *Nature* **545,** 299–304 (2017).

10. Hoshino, T. & Ogasawara, N. Biosynthesis of violacein: Evidence for the intermediacy of 5-hydroxy-L-tryptophan and the structure of a new pigment, oxyviolacein, produced by the metabolism of 5-hydroxytryptophan. *Agric. Biol. Chem.* **54,** 2339–2346 (1990).

11. Wang, H. *et al.* Biosynthesis and characterization of violacein, deoxyviolacein and oxyviolacein in heterologous host, and their antimicrobial activities. *Biochem. Eng. J.* **67,** 148–155 (2012).

12. Petersen, M. T. & Nielsen, T. E. Supporting Information Tandem Ring-Closing Metathesis / Isomerization Reactions for the Total Synthesis of Violacein. *Org. Lett.* **15,** 1986–1989 (2013).

13. Wille, G. & Steglich, W. A Short Synthesis of the Bacterial Pigments Violacein and Deoxyviolacein. *Synthesis (Stuttg).* **2001,** 0759–0762 (2001).

14. Füller, J. J. *et al.* Biosynthesis of violacein, structure and function of L-tryptophan oxidase VioA from chromobacterium violaceum. *J. Biol. Chem.* **291,** 20068–20084 (2016).

15. Hoshino, T. Violacein and related tryptophan metabolites produced by Chromobacterium violaceum: Biosynthetic mechanism and pathway for construction of violacein core. *Appl. Microbiol. Biotechnol.* **91,** 1463–1475 (2011).

16. Sánchez, C., Méndez, C. & Salas, J. A. Indolocarbazole natural products: occurrence, biosynthesis, and biological activity. *Nat. Prod. Rep.* **23,** 1007–1045 (2006).

17. Du, Y. L. & Ryan, K. S. Catalytic repertoire of bacterial bisindole formation. *Curr. Opin. Chem. Biol.* **31,** 74–81 (2016).

18. Alkhalaf, L. M., Du, Y.-L. & Ryan, K. S. in *Methods in Enzymology* (ed. Elsevier Inc) 21–37 (2016). doi:10.1016/bs.mie.2016.02.017

19. Sánchez, C., Méndez, C. & Salas, J. A. Engineering biosynthetic pathways to generate antitumor indolocarbazole derivatives. *J. Ind. Microbiol. Biotechnol.* **33,** 560–

568 (2006).

20. Moreau, P. *et al.* Synthesis, mode of action, and biological activities of rebeccamycin bromo derivatives. *J. Med. Chem.* **42,** 1816–1822 (1999).

21. Sánchez, C. *et al.* The Biosynthetic Gene Cluster for the Antitumor Rebeccamycin. *Chem. Biol.* **9,** 519–531 (2002).

22. Salas, J. A. & Méndez, C. Indolocarbazole antitumour compounds by combinatorial biosynthesis. *Curr. Opin. Chem. Biol.* **13,** 152–160 (2009).

23. Onaka, H., Taniguchi, S., Igarashi, Y. & Furumai, T. Cloning of the staurosporine biosynthetic gene cluster from Streptomyces sp. TP-A0274 and its heterologous expression in Streptomyces lividans. *J. Antibiot. (Tokyo).* **55,** 1063–71 (2002).

24. Gani, O. A. B. S. M. & Engh, R. A. Protein kinase inhibition of clinically important staurosporine analogues. *Nat. Prod. Rep.* **27,** 489 (2010).

25. Fu, P. *et al.* New indolocarbazoles from a mutant strain of the marine-derived actinomycete Streptomyces fradiae 007M135. *Org. Lett.* **14,** 6194–6197 (2012).

26. Ryan, K. S., Balibar, C. J., Turo, K. E., Walsh, C. T. & Drennan, C. L. The violacein biosynthetic enzyme VioE shares a fold with lipoprotein transporter proteins. *J. Biol. Chem.* **283,** 6467–6475 (2008).

27. Hirano, S., Asamizu, S., Onaka, H., Shiro, Y. & Nagano, S. Crystal structure of VioE, a key player in the construction of the molecular skeleton of violacein. *J. Biol. Chem.* **283,** 6459–6466 (2008).

28. Yeh, E., Garneau, S. & Walsh, C. T. Robust in vitro activity of RebF and RebH, a two-component reductase/halogenase, generating 7-chlorotryptophan during rebeccamycin biosynthesis. *Proc. Natl. Acad. Sci.* **102,** 3960–3965 (2005).

29. Yeh, E., Blasiak, L. C., Koglin, A., Drennan, C. L. & Walsh, C. T. Chlorination by a

long-lived intermediate in the mechanism of flavin-dependent halogenases. *Biochemistry* **46,** 1284–1292 (2007).

30. Sánchez, C. *et al.* Combinatorial biosynthesis of antitumor indolocarbazole compounds. *Proc. Natl. Acad. Sci. U. S. A.* **102,** 461–466 (2005).

31. Du, Y. L. & Ryan, K. S. Expansion of Bisindole Biosynthetic Pathways by Combinatorial Construction. *ACS Synth. Biol.* **4,** 682–688 (2015).

32. Moore, S. J. *et al.* EcoFlex: A Multifunctional MoClo Kit for E. coli Synthetic Biology. *ACS Synth. Biol.* **5,** 1059–1069 (2016).

33. Nishizawa, T., Aldrich, C. C., David, H. & Sherman, D. H. Molecular Analysis of the Rebeccamycin L-Amino Acid Oxidase from *Lechevalieria aerocolonigenes* ATCC 39243 Molecular Analysis of the Rebeccamycin L -Amino Acid Oxidase from *Lechevalieria aerocolonigenes* ATCC 39243. *J. Bacteriol.* **187,** 2084–2092 (2005).

34. DeCelle, K. Registry of Standard Biological Parts: Part:BBa_J04500. *Registry of Standard Biological Parts* (2005). doi:iGEM07_Example (2008-09-08)

35. Ceroni, F., Algar, R., Stan, G.-B. & Ellis, T. Quantifying cellular capacity identifies gene expression designs with reduced burden. *Nat. Methods* **12,** 415–418 (2015).

36. Fraaije, M. W. & Mattevi, A. Flavoenzymes: Diverse catalysts with recurrent features. *Trends Biochem. Sci.* **25,** 126–132 (2000).

37. Chang, F. Y., Ternei, M. A., Calle, P. Y. & Brady, S. F. Discovery and synthetic refactoring of tryptophan dimer gene clusters from the environment. *J. Am. Chem. Soc.* **135,** 17906–17912 (2013).

38. Chang, F. Y., Ternei, M. A., Calle, P. Y. & Brady, S. F. Targeted Metagenomics: Finding Rare Tryptophan Dimer Natural Products in the Environment. *J. Am. Chem. Soc.* **137,** 6044–6052 (2015).

39. Ryan, K. S. Biosynthetic gene cluster for the cladoniamides, bis-indoles with a rearranged scaffold. *PLoS One* **6,** (2011).

40. Du, Y. L., Williams, D. E., Patrick, B. O., Andersen, R. J. & Ryan, K. S. Reconstruction of cladoniamide biosynthesis reveals nonenzymatic routes to bisindole diversity. *ACS Chem. Biol.* **9,** 2748–2754 (2014).

41. Chang, F.-Y. & Brady, S. F. Discovery of indolotryptoline antiproliferative agents by homology-guided metagenomic screening. *Proc. Natl. Acad. Sci. U. S. A.* **110,** 2478–83 (2013).

42. Smanski, M. J. *et al.* Functional optimization of gene clusters by combinatorial design and assembly. *Nat. Biotechnol.* **32,** 1241–1249 (2014).

43. Ladner, C. C. & Williams, G. J. Harnessing natural product assembly lines: structure, promiscuity, and engineering. *J. Ind. Microbiol. Biotechnol.* **43,** 371–387 (2016).

44. Kim, E., Moore, B. S. & Yoon, Y. J. Reinvigorating natural product combinatorial biosynthesis with synthetic biology. *Nat. Chem. Biol.* **11,** 887–887 (2015).

45. Smanski, M. J. *et al.* Synthetic biology to access and expand nature's chemical diversity. *Nat. Rev. Microbiol.* **14,** 135–149 (2016).

46. Crane, E. A. & Gademann, K. Capturing Biological Activity in Natural Product Fragments by Chemical Synthesis. *Angew. Chemie - Int. Ed.* **55,** 3882–3902 (2016).

47. Winter, G. Xia2: An expert system for macromolecular crystallography data reduction. *J. Appl. Crystallogr.* **43,** 186–190 (2010).

48. McCoy, A. J. *et al.* Phaser crystallographic software. *J. Appl. Crystallogr.* **40,** 658–674 (2007).

49. Adams, P. D. *et al.* PHENIX: A comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **66,**

213–221 (2010).

50. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **66,** 486–501 (2010).

51. Robert, X. & Gouet, P. Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Res.* **42,** 320–324 (2014).

Acknowledgement

HEL receives funding from the Imperial College President's PhD Scholarships. The authors thank all staff involved in Diamond Light Source Beamline I04-1 for assistance in VioA crystal data collection.

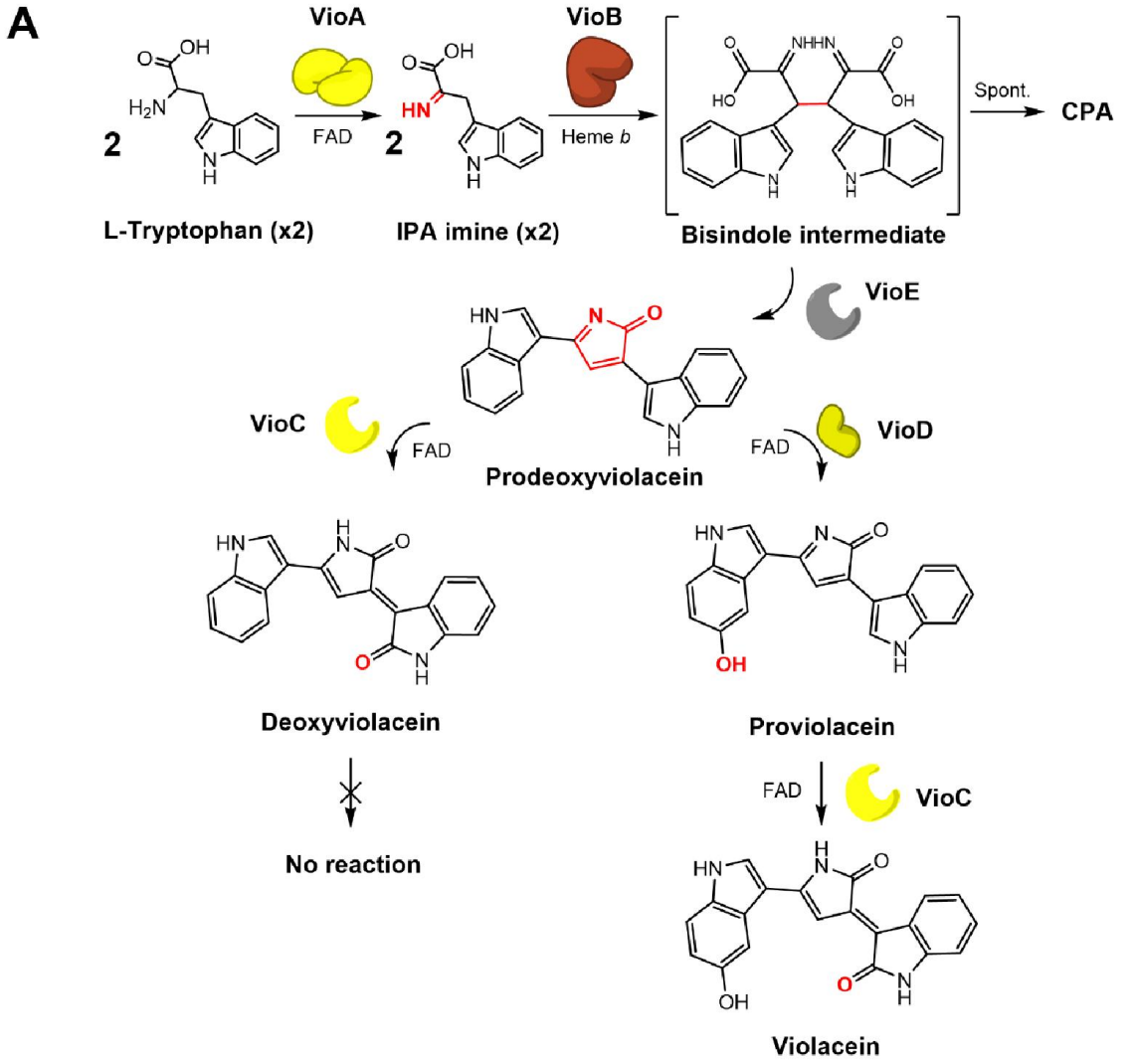
Author Contribution

HEL, KP and PF designed the study, analysed the data, and wrote the manuscript. SM provided feedback and technical assistance for experiments. HEL, SMC and MM performed the experiments.

Notes

The authors declare no competing financial interests.

Figures



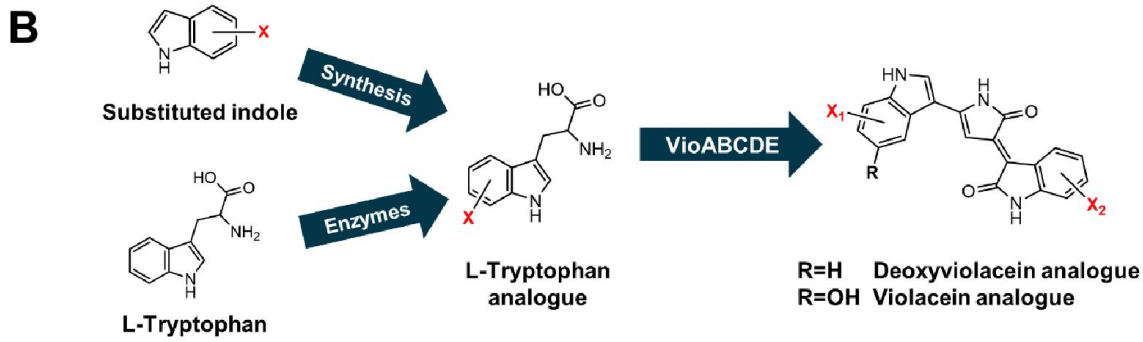
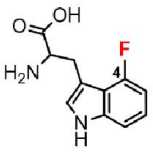
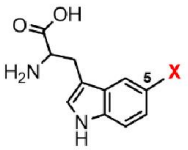


Figure 1. Violacein biosynthesis pathway. **A**, VioA, VioC and VioD (yellow) are FAD-dependent oxidases, VioB (red) is a heme-protein, while VioE (grey) is a colourless protein without any bound cofactors or prosthetic groups. For each molecule of violacein or deoxyviolacein, two molecules of L-tryptophan are required. Chemical groups marked red in structures and reaction arrows indicate atoms incorporated or removed as part of the reaction. CPA is produced as a side-product in the absence of VioE which converts the unstable bisindole intermediate to PDV. IPA, indole-3-pyruvic acid; CPA, chromopyrrolic acid; PDV, prodeoxyviolacein. **B**, L-tryptophan analogues can be generated by two different routes, either via chemical synthesis from substituted indole or enzyme-catalysed modification of L-tryptophan. Tryptophan analogues can be used as precursors for the violacein biosynthesis pathway to generate violacein or deoxyviolacein analogues.

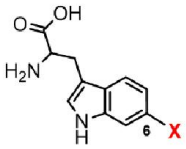
A



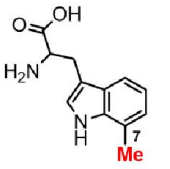
4-F-Trp (**4FT**)



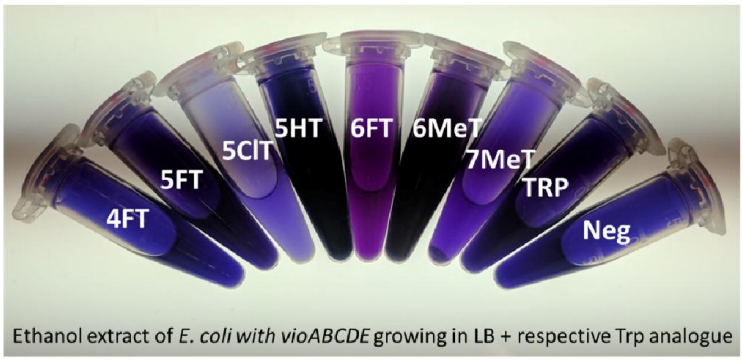
X = F 5-F-Trp (**5FT**)
X = Cl 5-Cl-Trp (**5CIT**)
X = OH 5-OH-Trp (**5HT**)



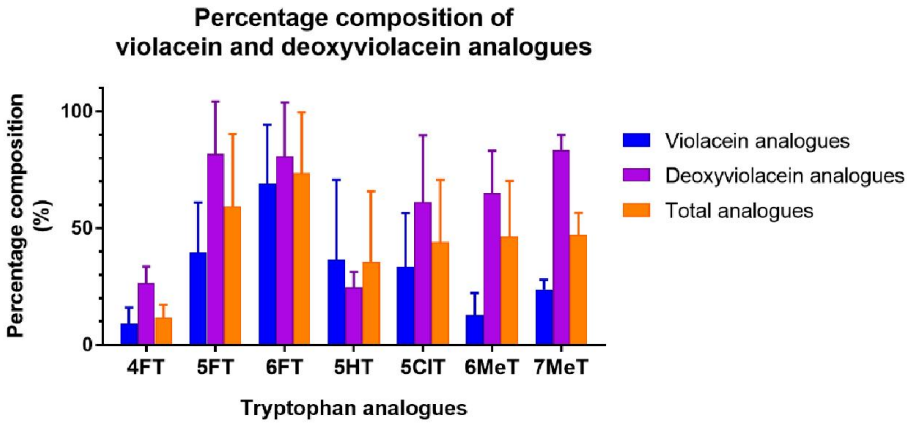
X = F 6-F-Trp (**6FT**)
X = Me 6-Me-Trp (**6MeT**)



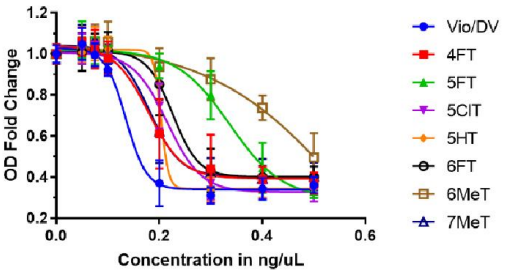
7-Me-Trp (**7MeT**)



B



C



D

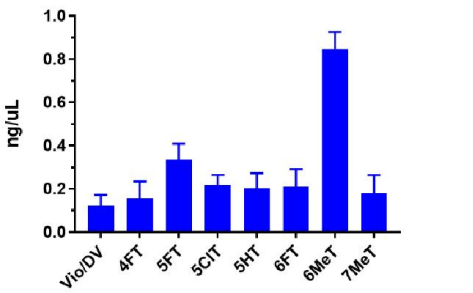
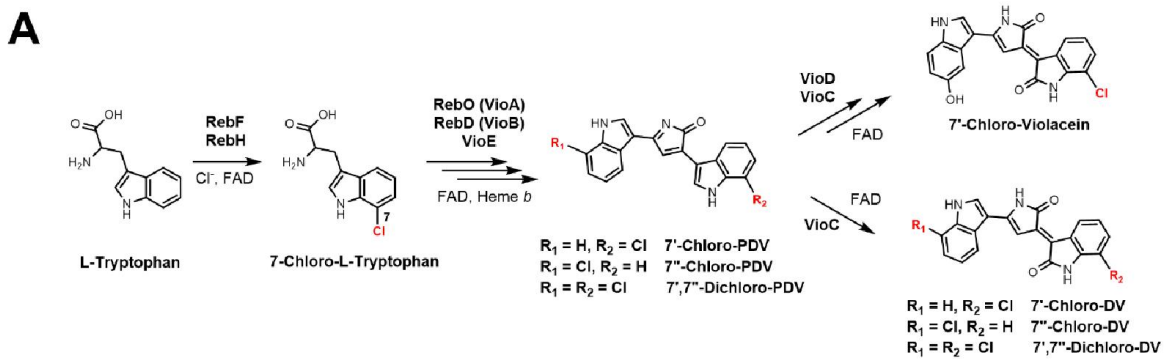
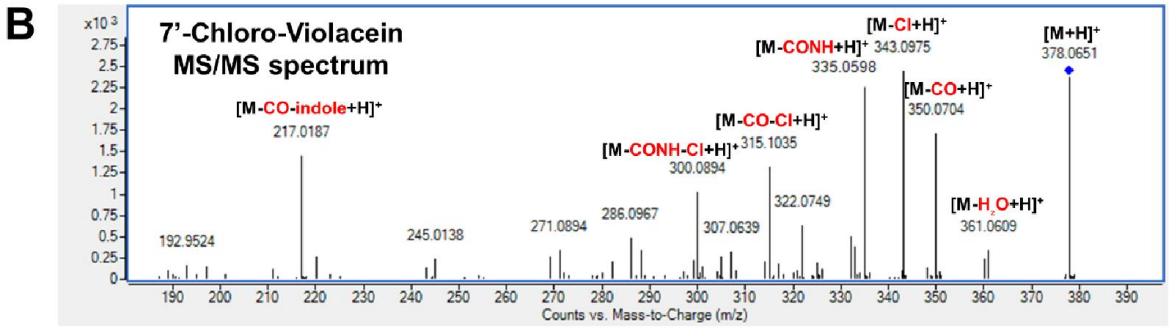
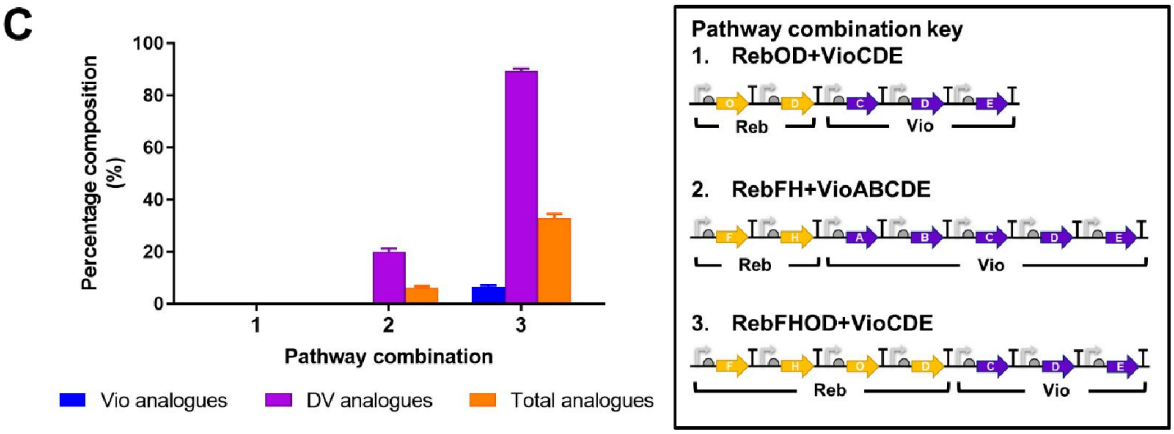


Figure 2. Generation of violacein analogues via tryptophan analogues. **A**, Tryptophan analogues used in generating violacein analogues from *E. coli* strains expressing the VioABCDE pathway. Abbreviations: 4FT, 4-fluoro-tryptophan; 5FT, 5-fluoro-tryptophan; 5ClT, 5-chloro-tryptophan; 5HT, 5-hydroxy-tryptophan; 6FT, 6-fluoro-tryptophan; 6MeT, 6-methyl-tryptophan; 7MeT, 7-methyl-tryptophan; TRP, L-tryptophan; NEG, negative control (no added tryptophan or its analogue). Cells were grown in LB which contains tryptophan, so negative control would still produce violacein and deoxyviolacein. **B**, Percentage compositions of the different individual violacein and deoxyviolacein analogues compared to the total pool of analogues. The mean and SD from biological triplicates is indicated. **C**, Dose response curves of violacein analogues against *B. subtilis*, showing range of OD fold change (set to 1 at 0 ng/μL) at 0, 0.05, 0.075, 0.1, 0.2, 0.3, 0.4 and 0.5 ng/μL concentrations of analogues respectively. The final concentration of DMSO is kept at 5% (w/v) for all samples. Vio/DV contains 76% violacein and 24% deoxyviolacein as assayed by LC-MS/MS. Data is shown from three independent biological replicates carried out on three different days. **D**, Inhibitor concentration at 50% response (IC50) of violacein analogues activity against *B. subtilis.* Error bars represent the RSDR of IC50 as calculated by Prism GraphPad 7.0 (see Methods for details).







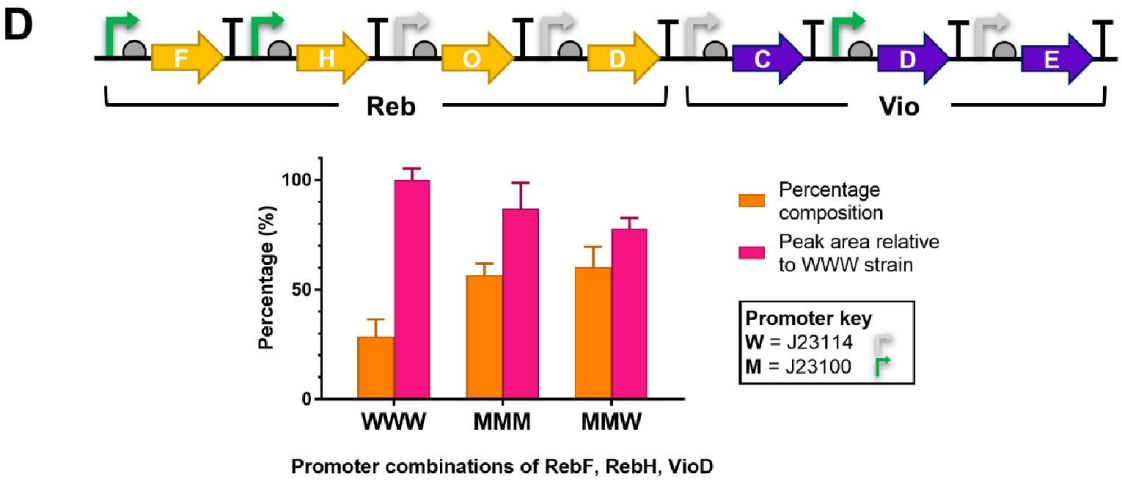


Figure 3. Hybrid Reb-Vio biosynthesis pathway generates chlorinated violacein and deoxyviolacein (DV) analogues. **A**, Schematic of the hybrid rebeccamycin-violacein pathway enzymes that gave the final 7-chloro analogues of violacein and deoxyviolacein. **B**, Typical LC-MS/MS fragmentation spectrum of 7-chloro-violacein with the major fragments labelled. Mass spectra for all other violacein/deoxyviolacein analogues is given in **Table S2**. **C**, Percentage compositions of 7-chloro analogues of violacein and deoxyviolacein extracted from different *E. coli* strains expressing RebOD+VioCDE, RebFH+VioABCDE and RebFHOD+VioCDE. The data shows the mean and SD from biological duplicates. Pathway symbols are drawn using SBOL visual standard for biological parts (bent arrow for promoter, half-circle for RBS, straight arrow (yellow – rebeccamycin; purple – violacein) for genes, and T-shape for terminator. **D**, Percentage compositions of 7-chloro analogues of violacein and deoxyviolacein extracted from *E. coli* expressing RebFHOD+VioCDE with different combinations of medium (J23108, green) and weak (J23114, grey) promoters controlling the expression of RebF, RebH and VioD. The data is shown as the mean and SD from biological triplicates.

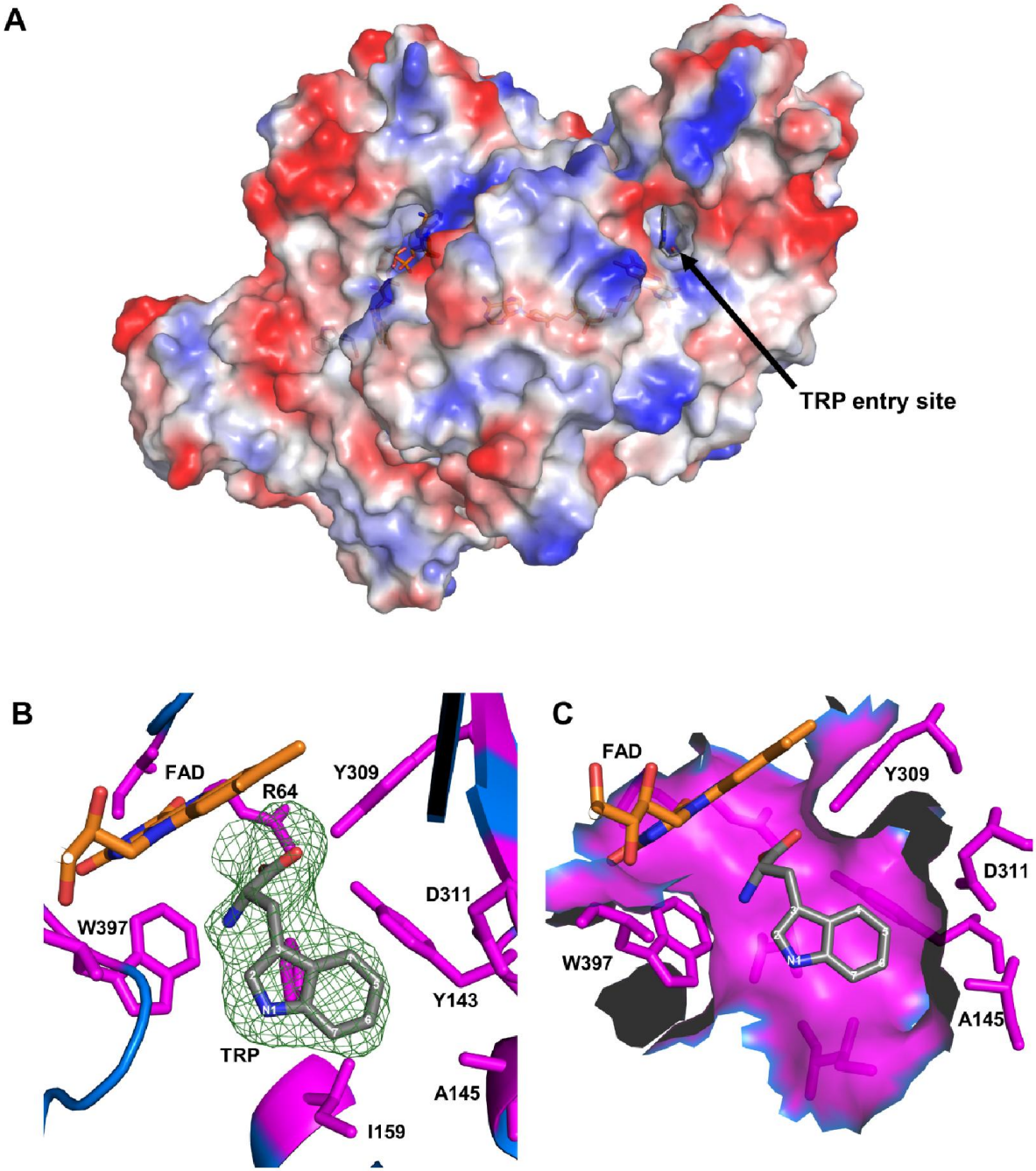


Figure 4. VioA crystal structure in complex with L-tryptophan at 2.6 Å resolution (PDB **6ESE**). **A**, Electrostatic surface representation of VioA crystal structure with bound L-tryptophan and FAD (shown in stick representation). Tryptophan (TRP) entry site is labelled. **B**, VioA active site residues (magenta) with L-tryptophan (TRP grey with carbon positions of indole ring labelled) and FAD cofactor (orange with elemental colour). In this view, H163 is located behind TRP. Green mesh is the mF_o-DF_c omit map of TRP contoured at σ=1.5. **C**, Binding site pocket showing space enclosed by residues (magenta) around L-tryptophan (grey with elemental colour). The carbon positions of tryptophan indole ring are labelled.