

Sequence Analysis and Functional Characterization of the Violacein Biosynthetic Pathway from *Chromobacterium violaceum*

P.R. August^{1*}, T.H. Grossman¹, C. Minor⁺, M.P. Draper¹, I. A. MacNeil¹, J.M. Pemberton², K.M. Call¹, D. Holt¹, and M. S. Osburne¹

¹Ariad Pharmaceuticals, Inc, 26 Landsdowne Street, Cambridge, Massachusetts 02139, USA

²Department of Microbiology, University of Queensland, St. Lucia, Australia 4072

+ Current address:

Millennium Pharmaceuticals
Cambridge, Massachusetts 02139

Abstract

Violacein is a purple-colored, broad-spectrum antibacterial pigment that has a dimeric structure composed of 5-hydroxyindole, oxindole and 2-pyrrolidone subunits formed by the condensation of two modified tryptophan molecules. The violacein biosynthetic gene cluster from *Chromobacterium violaceum* was characterized by DNA sequencing, transposon mutagenesis, and chemical analysis of the pathway intermediates produced heterologously in *Escherichia coli*. The violacein biosynthetic gene cluster spans eight kilobases and is comprised of the four genes, *vioABCD*, that are necessary for violacein production. Sequence analysis suggests that the products of *vioA*, *vioC* and *vioD* are nucleotide-dependent monooxygenases. Disruption of *vioA* or *vioB* completely abrogates the biosynthesis of violacein intermediates, while disruption of the *vioC* or *vioD* genes results in the production of violacein precursors.

Introduction

Many antibacterial, antifungal, and anticancer agents are secondary metabolites originally isolated from soil microorganisms. Over the past few decades, numerous efforts have focused on harnessing the biosynthetic capabilities of soil microbes to generate novel molecular structures for drug discovery efforts. These efforts have primarily utilized members of the Gram positive

Abbreviations:

Amp, Ampicillin
Kan, Kanamycin
Apra, Apramycin
Chlor, Chloramphenicol
ACPS, Acyl Carrier Protein Synthase

Received January 1, 2000; revised May 20, 2000; accepted May 20, 2000.
*For correspondence. Email Paul.august@aventis.com; Tel. (617)-768-4142; Fax. (617) 374-8811.

actinomycetale family since they are prolific producers of secondary metabolites. Molecular genetic analyses of these organisms have revealed that the genes encoding their secondary metabolite biosynthetic pathways are clustered and range in size from five to greater than 100 kilobases (Malpartida and Hopwood 1984; August *et al.*, 1998). As these natural product biosynthetic pathways become elucidated, more innovative efforts have been placed on producing these molecules in heterologous hosts such as *E. coli*, that are more amenable to physiological and genetic manipulations (Kealer *et al.*, 1998; Gehring *et al.*, 1997). However, a frequent problem with recombinant expression systems is that many biosynthetic enzymes may not be functional due to the absence of essential precursors, cofactors and/or accessory proteins in the host organism. For example, polyketide synthase enzymes from actinomycetes require co-expression of the native holo-ACP synthase (ACPS) since they are not effectively modified with 4'-phosphopantetheine by the *E. coli* ACPS (Cox *et al.*, 1997). Therefore it is important to identify biosynthetic pathways that are more amenable to heterologous expression in the *E. coli* physiological background in order to improve our ability to produce heterologous secondary metabolites in this host. Towards this effort, we have proceeded to examine the molecular genetics and biosynthesis of violacein, a natural product that can be produced heterologously in *E. coli* (Pemberton *et al.*, 1991).

Violacein is a purple-colored broad-spectrum antibacterial pigment produced by *Chromobacterium* (Duran *et al.*, 1994). The violacein molecule has a dimeric structure composed of 5-hydroxyindole, oxindole and 2-pyrrolidone subunits. The biosynthesis of violacein has been shown by radiolabel incorporation experiments to initiate with L-tryptophan (Hoshino *et al.*, 1987). However, the biosynthetic pathway and the type of enzymes involved remained unclear. The genes encoding the violacein biosynthetic pathway were previously isolated from *Chromobacterium violaceum* on a 14 kb fragment that was functionally expressed in *E. coli* and other Gram negative microorganisms (Pemberton *et al.*, 1991). In this earlier work, transposon mutagenesis of the 14 kb fragment resulted in the production of white and green colonies containing insertions that mapped to a 5 kb region. Here we report the complete sequence of the cloned violacein locus (*vio*), a functional assessment of the activity of the genes encoded therein, and infer a biosynthetic pathway for violacein production.

Results and Discussion

Mutant Phenotypes

Plasmid pVIO4, encoding the *vio* cluster, was mutagenized *in vitro* with the transposon GPS-Apramycin, and the resulting plasmids were transformed into *E. coli* strain DH10B.

Colony color and morphology could easily distinguish phenotypes of the transformant colonies from the purple phenotype conferred by the pVIO4 parent plasmid. Several distinct phenotypic classes were observed, in agreement with earlier work (Pemberton *et al.*, 1991). Of 569 clones, the majority (63.82%) were purple, reflecting the high probability of the transposon inserting into the vector without interrupting violacein biosynthetic genes. However, a small fraction of the clones (1.11%) formed intensely purple very small colonies, presumably violacein hyper-producers. Another small fraction (4.27%) formed very light purple colonies, potentially violacein hypo-producers. In addition, blue (1.48%), light-blue (0.56%), aqua (4.08%) and light-aqua (0.56%) clones were also obtained. A relatively large number of clones (24.11%) were white. The appearance of a large number of mutants (over 30%) encompassing a variety of mutant phenotypes suggested that violacein biosynthesis was encoded by multiple genes.

Sequence Analysis

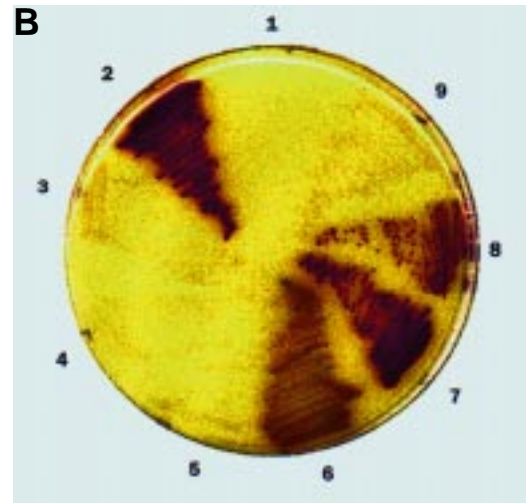
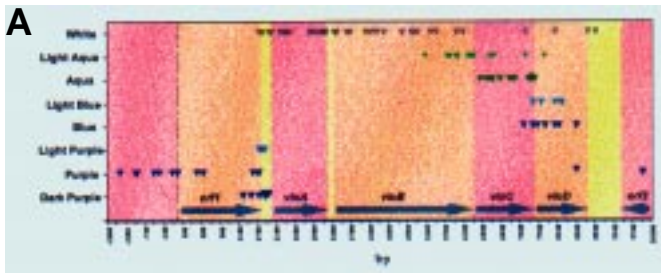
The sequence of the violacein gene cluster in plasmid pVIO4 was obtained from plasmids containing transposon insertion mutations, using primers that anneal to the ends of the integrated GPS-*Apra* transposon. Analysis of the assembled sequence revealed an insert of 10,094 bp encoding five open reading frames (ORFs) designated *orf1* (1629 bp), *vioA* (1257 bp), *vioB* (2997 bp), *vioC* (1290 bp), *vioD* (1122 bp) and *orf2* (partial). The locations of these genes and the phenotypes of representative mutants are shown in figures 1A and 1B, respectively. Initial homology determinations were performed using BLAST (Altschul *et al.*, 1990). The product of *orf1* was weakly similar (61%; 59 a.a.) to phospholipase C precursor protein (Swissprot accession number P09978, Projan *et al.*, 1989) by BLAST analysis. The predicted proteins VioA (418 a.a.), VioC (429 a.a.), and VioD (373 a.a.), encoded by genes *vioA*, *vioC* and *vioD*, respectively, failed to show significant homology to any known proteins using BLAST. Interestingly, ProfileScan analysis (Gribskov *et al.* 1987) revealed a conserved N-terminal NAD(P)H binding domain in each of these proteins (VioA, a. a. 7-39; VioC, a. a. 3-33; VioD, a. a. 4-32, Bellamacina, 1996). An alignment of these domains is shown in Figure 2. It is known that the biosynthesis of violacein requires NADPH for the conversion of L-tryptophan to violacein in cell-free extracts, in support of the idea that one or more of the NADPH binding sites in VioA, VioC and VioD may be responsible for this requirement (Hoshino and Yamamoto, 1997).

Proteins weakly similar to VioA, VioC and VioD were identified using the sensitive homology search tool PSI-BLAST. PSI-BLAST combines statistically significant alignments produced by the program BLAST into a position-specific score matrix, and subsequently utilizes that matrix to search a protein database (Altschul *et al.*, 1997). This search tool is much more sensitive than BLAST alone and excels at identifying weak but functionally relevant similarities. Results of this analysis suggested that the proteins VioA, VioC and VioD are related to the PheA (phenol hydroxylase) /TfdB (2,4-dichlorophenol hydroxylase) family of FAD monooxygenases. VioA exhibited 22% identity and 35% similarity to the aplysianin-A protein of the sea hare *Aplysia kurodai*, a putative L-amino acid oxidase (Takamatsu *et al.*, 1995). This family of monoamine oxidases catalyzes the oxidative

Figure 1. A.) Transposon insertion points into the pVIO4 insert overlaid onto the violacein biosynthetic gene cluster ORF map. The Y-axis position of the insertion points identifies the mutant phenotype. The region to the left of the insert start site is displayed to illustrate that transposon integration into the vector sequence does not generate any mutant phenotypes. B.) Phenotypes of selected transposon generated mutants. 1.) Control (pVIO4) No growth on apramycin. 2.) clone 00001452b_e08 (GPS-*Apra*::*orf1*) 3.) clone 00001452b_a10 (GPS-*Apra*::*vioABCD* promoter) Although the overall phenotype is white, small brilliant purple colonies can be seen in the background. 4.) clone 00001452b_d06 (GPS-*Apra*::*vioA*) 5.) clone 00001452b_f05 (GPS-*Apra*::*vioB*) 6.) clone 00001452b_c02 (GPS-*Apra*::*vioC*) 7.) clone 00001452b_a01 (GPS-*Apra*::*vioD* upstream) 8.) clone 00001452b_c01 (GPS-*Apra*::*vioD*) 9.) clone 00001452b_c12 (GPS-*Apra*::terminator).

Figure 2. Alignment of the VioA, VioB and VioC proteins revealing a conserved GXGXXG nucleotide binding motif.

Figure 5. This scheme illustrates a hypothetical route for violacein biosynthesis from tryptophan by evaluating the homology of the four genes required for violacein biosynthesis and the intermediates produced from mutants of these genes. The enzyme names are placed adjacent to the reactions that we postulate they respectively catalyze. The symbol (*) indicates incorporation patterns with stable ¹³C into violacein from tryptophan (Hoshino and Ogasawara, 1990). 1, Tryptophan; 2, 5-Hydroxytryptophan; 3, Indole Pyruvic Acid; 4, D-2-Amino-2-(3-indolyl)Propionic Acid; 5, Prodeoxyviolacein; 6, Violacein; 7, Oxyviolacein; 8, Deoxy-violacein.



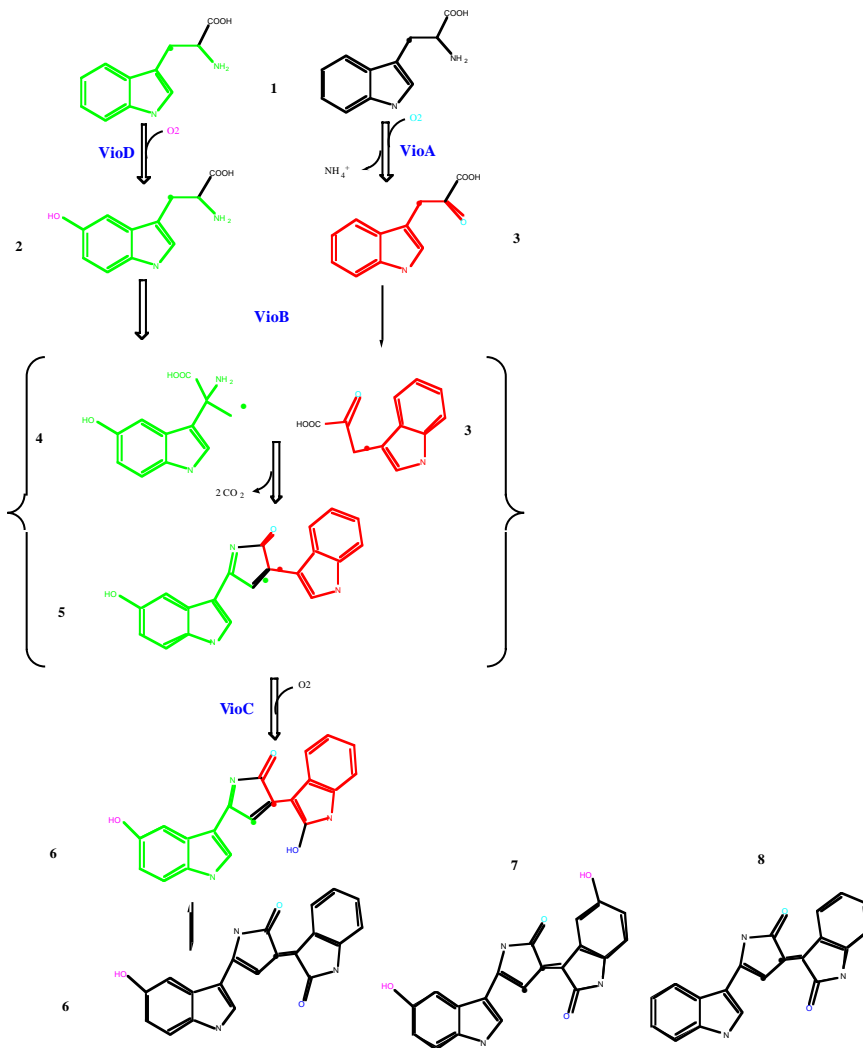
10 20 30 40 50

VioD/1-443 ---MKILVIGAGPAGLVFASQLKQARPLWAIDIIIEKN-D-EQEVLGWGVVLPGRPG--

VioC/1-443 ---MKRAIIVGGGLAGGLTAYLAKRG--YEVHVVEKRGDFLODLLSSYVDVVSSRAIG

VioA/1-443MKHSSDICIVGAGISGLTCASHLLDS PACRGLSLRIFDMQQEAGGRIRSKMLD GKAS--

GXGXXG



```

1331 GGTTCAGACCTGGTGGGGCGGGTGGGGGGCAACCAGTACGCGTACTACACCGCCCAGGGGGCCGCTCAATCGTTCGCCGAGCTGCGCCGTGATC
orf1 403▶ V L D L V G R R G G Q P V R V L H R P G A A Q S F R R A A P D
1426 AATCACAGCCGGCGGGCGGCTGCCTGCAGGACGGCGACGTGGTGTTCAGGATTGGGCGCGGGCCGACTACTACCTGCACCGCGTGGGT
435▶ Q S Q P A G R L P A G R R R G V V Q G L G A R G R L L P D R V G
1521 TGGCGGCGACCATGCCGACCGCTGTATCTGTGGCAGCCGGCTCTGGGGATGGGGACGATTCCGGCTGCGGATGGGCGCATCGCCGACGATATC
467▶ W R Q P C R P A V S V A A G V W G W G T I P A A D G R I A A V S
1616 TGGACTGGCGGGACGATTGCTTTTCGCCAAGCCGCTGAGGGATTGCTTGGCTTTTGTATCCAAATCTCCACGCTCATGCCCTGACCCCTTGGA
498▶ G L A R T I A F R Q T P L R D C L A F V S N S P R H A L T L G
1711 CAGTATCAGCCGATTGTGACACAATATGGTGTGATGATGAGTccggttgaacTaatgtggettgaaattAcaaaagcattttatgtttaaatt
530▶ T V S A A L S D N M V D E .
1805 aAaatGtcaaacgattttttgtgccCggattgaccggGcAgtgacAggaaAcgaaaggcagCatcccgatccaabgcegcagtgccgtAgg
1897 ttttcgccgcCAgtgcccgtccgcggttgccgcgcccgggGtgagttgActagtcaaaaggacattcgcGATGAAGCATTCTCCGATATC
vioA 1▶ M K H S S D I
1990 TGCATTGTGCGGCCCGCATCAGCGGCTGACCTGCGCCAGCCATCTGCTCGATTGCGCCGCTTGGCCGGCCCTGTGCTGCGCATCTTCGACA
8▶ C I V G A G I S G L T C A S H L L D S P A C R G L S L R I F D
2084 TGCAGCAGGAGGGCGGGCGCCGCATCCGCTCGAAGATGCTGGATGGCAGGGCTTCGATAGAGCTGGGCGCGGGCGATACCTCCCGCAGCTGCAC
39▶ M Q Q E A G G R I R S K M L D G K A S I E L G A G R Y S P Q L H
2179 CCGCATTTCCAGAGCGCGATGCGAGCTTACAGCCAGAGAGCGAGGTGTATCCGTTACCCAGTTGAATTCAGAGCCATGCTCCAGCAGAGCC
71▶ P H F Q S A M Q H Y S Q K S E V Y P F T Q L K F K S H V Q Q K

```

Figure 3. Region of the *vio* locus in which transposon insertions generate violacein hyperproducing mutants. Nucleotides in shadow font indicate bases where transposon insertions generate mutants that possess the white phenotype. Nucleotides in bold font indicate transposon insertion points that yield violacein hyperproducers (dark purple). Nucleotides in grey font indicate insertion points that yield violacein hypoproducers (light purple). Nucleotides in outline font indicate transposon insertion points that yield violacein nonproducers (white). The arrow on top of the bases reveals the direction of transcription from the transposon encoded apramycin resistance gene.

deamination of hydrophobic and aromatic L-amino acids (Raibekas and Massey, 1998). VioC exhibited 27% identity and 47% similarity to kynurenine 3-monooxygenase from *Homo sapiens*. Kynurenine 3-monooxygenase is a flavin-dependent monooxygenase that catalyzes the oxidation of L-kynurenine to 3-hydroxy-L-kynurenine in the kynurenine pathway of tryptophan metabolism (Alberati-Giani *et al.*, 1997). VioD exhibited 25% identity and 38% similarity to a putative hydroxylase from *Streptomyces coelicolor* A3(2) (Genbank accession number AL0949863) and a number salicylate hydroxylases.

PSI-BLAST analysis of VioB (998 aa) revealed limited similarity to any known proteins. The nucleotide sequence of the *orf2* gene is partially complete. However, BLAST analysis revealed that gene product of *orf2* is 45% similar and 25% identical to a hypothetical hydroxylase from the actinomycete *Amycolatopsis orientalis* (Genbank accession number AJ223998).

Functional Analysis

Transposon insertions in the *vio* gene cluster were localized using the sequence assembly program Phrap, and by visual inspection using the program Consed (Ewing *et al.*, 1998; Gordon *et al.*, 1998). In agreement with the earlier work of Pemberton *et al.*, we observed a correlation between colony phenotypes and transposon location in the gene cluster. These results indicated the limits of the violacein gene cluster and demonstrated that *vioA*, *vioB*, *vioC* and *vioD* were the only genes required on pVIO4 for violacein biosynthesis in *E. coli*. As illustrated in figure 1A, transposon integration into *vioA* and *vioB* yielded mutants with a white colony phenotype, while integration into *vioC* and *vioD* yielded aqua and blue colony phenotypes, respectively. Integration into *orf1* or *orf2* did not disrupt or alter the

production of violacein, demonstrating that these two genes are apparently not involved in the biosynthesis of violacein. However, due to the similarity of the predicted ORF1 protein to phospholipase C, we tested *E. coli* bearing pVIO4 for hemolytic activity (using blood agar plates). Strain DH10B (pVIO4) was in fact hemolytic, unlike the control host strain bearing the parent plasmid without the *vio* cluster. In addition, strains containing transposon insertions in or just upstream of the *orf1* gene lacked hemolytic activity (not shown), confirming that the hemolytic activity is encoded in the *orf1* region. The location of *orf1* adjacent to the *vio* cluster suggests the interesting possibility that the DNA insert in pVIO4 may be derived from a pathogenicity island, a discrete segment of DNA that encodes virulence traits and often appears to have a foreign origin. Often times, these DNA segments are inserted into the chromosome and confer upon the host bacterium a variety of virulence traits (*e.g.* hemolysins and toxins).

Sequence analysis indicates that *vioA*, *vioB*, *vioC* and *vioD* genes are transcribed in the same direction. The extent of the *vio* operon was examined in *E. coli* harboring pVioSub4 by Northern analysis using the *vioD* gene as a probe (data not shown). Although a clear transcript size was difficult to obtain, the results were consistent with the production of a large transcript (~8 kb) encoding the *vioABCD* genes. A putative transcriptional terminator was identified downstream of *vioD* (nt 913-9344) and upstream of *orf2*. The fact that transposon-generated hyperproducers arose only when the transposon was inserted proximal to *vioA* and distal to *orf1* (Figure 1A) also suggests that the *vioABCD* cluster is transcribed as an operon. Interestingly, violacein hyper-producing colonies were extremely small, and when restreaked on apramycin plates, white non-producing variants arose spontaneously at a high frequency (Figure 1B). This observation suggests that

hyperproduction of violacein is deleterious to *E. coli*. Although Gram negative bacteria are resistant to external violacein, it is possible that violacein is toxic to *E. coli* at high intracellular concentrations. Transposon integrations that resulted in the hyper-production phenotype were clustered directly upstream of *vioA*, with the transposon apramycin-resistance gene positioned in the opposite orientation with regard to the direction of transcription of the gene cluster (Figure 3). Conversely, the apramycin resistance gene was in the same orientation as the *vioABCD* genes for transposon insertions that resulted in the hypo-production phenotype. Violacein hyper-production did not result from transposon integrations elsewhere in the biosynthetic gene cluster, suggesting that either the stoichiometry of the biosynthetic proteins is critical for hyper-production, or that VioA is the rate-limiting enzyme in the violacein biosynthetic pathway.

Analysis of Pathway Intermediates

Despite efforts to find homologies to the individual *vio* genes, their biosynthetic roles remained unclear. In an effort to define their roles, we sought to obtain structural information on the intermediates produced by the *vio* mutants. Extracts of the various classes of *vio* mutants were generated and analyzed by HPLC. Biosynthetic intermediates from the *vioA* and *vioB* mutant extracts were not detected by this method. However, pigments from the *vioC* and *vioD* mutants were isolated and their molecular masses determined. A major product of the *vioC* mutant had an M+1 mass of 311.21 and a major product of the *vioD* mutant had an M+1 mass of 328.2 (Figure 4A and 4B). As the mass of violacein is 343 and that of deoxyviolacein is 327, the mass differences between the products isolated from the *vioD* and *vioC* mutants differ

from violacein by 16 and 32 mass units, respectively. These differences could be attributed to the absence of an oxygen atom: the mass of the product from the *vioD* mutant corresponds to that of deoxyviolacein, while the product isolated from the *vioC* mutant is similar to that of prodeoxyviolacein (Hoshino *et al.*, 1995).

Previous experiments by Hoshino *et al.* (1987) revealed that all of the oxygen atoms in violacein are derived from molecular oxygen. Based on our results and previous work of others, a hypothetical pathway for violacein production is presented in Figure 5. VioA, VioC and VioD likely catalyze the three reactions necessary to introduce these oxygen atoms into the violacein molecule. The chemical structures of the molecules produced from mutants in each of the corresponding genes permits for tentative assignment in the pathway.

Interestingly, VioA shows a high degree of overall similarity with toxic proteins such as snake venom, aplysianin and achacin, that appear to be L-amino acid oxidases (Raibekas *et al.*, 1998; Obara *et al.*, 1992; Yamazaki *et al.*, 1989). These enzymes are bactericidal and cause apoptosis in a general fashion by deaminating several types of amino acids, including tryptophan. Based upon similarities to these proteins, the function of the *vioA* gene product is proposed to be an L-amino acid oxidase, more specifically, tryptophan 2-monooxygenase. Since the disruption of *vioA* results in phenotypically white mutants that do not produce any discernable violacein intermediates, VioA must catalyze a reaction early in the violacein biosynthetic pathway. Therefore, we propose that VioA generates an indole pyruvate intermediate that feeds into the violacein biosynthetic pathway (Figure 5). The mass spectra which demonstrate that the *vioD* mutant produces deoxyviolacein, suggest that VioD is responsible for the

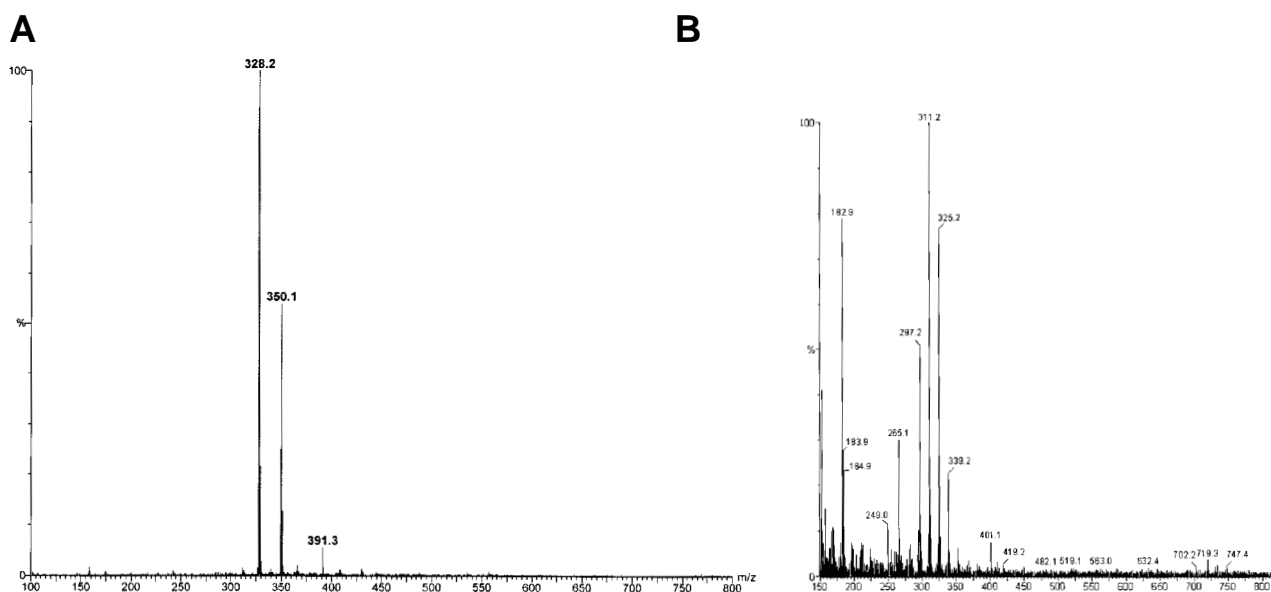


Figure 4. Daughter ion mass spectra for the A. *vioC* mutant (00001452b_c02) m/z 311.21 and B. *vioD* mutant (00001452b_c01) m/z 328.2.

additional oxygen atom incorporated on the violacein oxindole subunit. Similarly, the mass of the major product produced from the *vioC* mutant suggests that VioC acts on prodeoxyviolacein and hydroxylates the indole ring to generate the violacein oxindole subunit.

vioB mutants also failed to produce any detectable violacein intermediates, suggesting that VioB may be involved in generating the violacein chromophore. However, it is difficult to assign a clear biochemical role for VioB based on the lack of homology to any known protein. Within the violacein biosynthetic pathway there is a unique decarboxylation reaction in which a 1,2-shift of a tryptophan side chain on the indole ring occurs, followed by the condensation with another tryptophan metabolite to generate the violacein pyrrol-2-one nucleus (Hoshino and Ogasawara, 1990). This reaction is very unusual and may explain why other VioB homologs have not been previously characterized. This biosynthetic step does not involve the incorporation of oxygen, and therefore VioA, VioC and VioD are unlikely to be involved in this reaction. We therefore speculate that VioB, which is a large protein, is a multifunctional enzyme that catalyzes the 1,2-indole shift of tryptophan as well as the condensation reaction to generate the violacein pyrrole ring (Figure 5).

VioC shows amino acid sequence similarity to the kynurenine hydroxylases. Although these enzymes hydroxylate an aromatic ring, it is plausible that VioC in fact incorporates oxygen into the pyrrole ring of violacein. This hypothesis is supported by the fact that *vioC* mutants generate a violacein intermediate that appears to have the molecular weight of prodeoxyviolacein, which has been previously characterized as lacking oxygenation at the 2 position of the indole nucleus (Hoshino *et al.*, 1995). This result suggests that VioC is responsible for catalyzing the oxygenation reaction that converts prodeoxyviolacein to deoxyviolacein (Figure 5).

The VioD protein shows similarity to salicylate hydroxylases, which hydroxylate the aromatic ring of salicylate. The most likely substrate for VioD would be the aromatic ring of tryptophan, which is hydroxylated to generate one of the violacein subunits. Evidence for this role is provided by the fact that the molecular mass of the pigmented intermediate isolated from the *vioD* mutant was 327.2, which corresponds to the published mass for deoxyviolacein. Therefore, it is plausible that VioD converts tryptophan to 5-hydroxytryptophan (Figure 5).

Production of violacein in *E. coli* demonstrates that monooxygenases from heterologous microorganisms can be functionally expressed in this host to generate interesting secondary metabolites. Additional work in our laboratory has shown that similar monooxygenases, which generate secondary metabolites from tryptophan, are present in other biosynthetic pathways cloned from soil genomic DNA (MacNeil *et al.*, in preparation). It is likely that heterologous pathways encoding monooxygenases can be more readily identified in *E. coli* because their aromatic substrates are available in this host. Overall, the observation that monooxygenases can be functionally expressed in *E. coli* suggests that future efforts to produce heterologous secondary metabolites in this microorganism should include a focus on biosynthetic gene clusters that encode these enzymes.

Experimental Procedures

Plasmids

Plasmid p4969 (~23 kb, ampR, chlorR, kanR), encoding the *Chromobacterium violaceum* gene cluster, was a generous gift from John Pemberton at the University of Queensland, Australia (Pemberton *et al.*, 1991). To reduce its size, plasmid p4969 was partially digested with *HindIII*, religated, and the ligation was transformed into *E. coli* DH10B cells. One transformant, pVIO4 (~17 kb, ampR, chlorR, kanS), encoding the violacein gene cluster, was chosen for further study.

Assay for Hemolytic Activity

Strains were assayed for hemolytic activity by streaking on Bacto TSA Blood Agar Base containing amp (50 µg/ml) and 5% sterile defibrinated blood.

Transposon Mutagenesis and DNA Sequencing

Plasmid pVIO4 was subjected to transposon mediated sequencing. Transposition reactions were performed according to the manufacturer's instructions (New England BioLabs, Beverly MA) with the exception that a modified transposon vector, pGPS-*Apra*, was used. pGPS-*Apra* contained the apramycin resistance gene in place of the pGPS-1 kanamycin resistance gene (August *et al.*, in preparation). Apramycin-resistant transformants were grown for 20 hours in deep 96-well plates containing super broth supplemented with apramycin at 100 µg/µl (Botstein *et al.*, 1975). Plasmid DNA was isolated using the Qiagen Biorobot 9600 according to the manufacturer's instructions. Sequencing reactions were performed on purified plasmid DNA using ABI Big Dye at one quarter strength and run on an ABI 377 DNA sequencer.

Sequence Analysis

DNA sequence was determined with the UNIX program Phred. The sequence data were assembled using the program Phrap and edited using the program Consed (Gordon *et al.*, 1998; Ewing *et al.*, 1998; Ewing and Green, 1998). Open reading frames were identified and translated using the program ORF finder at the National Center for Biotechnology Information (NCBI). BLAST and PSI-BLAST analysis were performed at NCBI. ProfileScan analysis was performed using the pscan Swiss Institute for Experimental Cancer Research server at www.isrec.isb-sib.ch/software/PFSCAN_form.html. The sequence of the violacein biosynthetic gene cluster has been deposited in Genbank (accession number AF172851).

Extraction and Purification

E. coli cultures were grown overnight in 25 ml of LB supplemented with apramycin (100 µg/ml), except for cultures containing pVIO4, which were supplemented with ampicillin (50 µg/ml). 10 ml of butanol were added to each culture. The mixture was shaken vigorously and centrifuged at 10K for 7 minutes. The upper (butanol) phase containing the pigmented compound was removed. Extracts were concentrated using a rotovap under high vacuum. The concentrate was resuspended in distilled water and then lyophilized. Extracts were stored at 4°C. Cultures of the *vioC* mutant were extracted with ethyl acetate (1:1). Extracts were dried using a stream of dry nitrogen gas and subjected to TLC using an ethyl acetate/hexane solvent system (9:1). The major product derived from the *vioC* mutant was purified using preparative TLC and the above solvent system. A band with an R_f of 0.13 was excised from a preparative TLC plate and extracted with methanol. The sample was concentrated and subjected to mass spectral analysis.

MS Analysis

Mass spectra were obtained in positive and negative ion electrospray mode on a Micromass Platform II single quadrupole mass spectrometer. The samples were dissolved in 50:50 v/v acetonitrile:water solution and introduced into the mass spectrometer via loop injection using a Hewlett-Packard HP1050 liquid chromatograph.

Acknowledgements

We thank Jo Handelsman, Robert Goodman, and Michelle Rondon for acquainting us with the violacein cluster. Special thanks to Richard Newcombe, Ming Ding, Danielle Clarke and Lee Hendricks in the Hoehst-Ariad Genomic Center sequencing group for generating the raw DNA sequence data used in this study.

References

- August, P.R., Tang, L., Yoon, Y.J., Ning, S., Muller, R., Yu, T.W., Taylor, M., Hoffmann, D., Kim, C.G., Zhang, X., Hutchinson, C.R., and Floss, H.G. 1998. Biosynthesis of the ansamycin antibiotic rifamycin: deductions from the molecular analysis of the *rif* biosynthetic gene cluster of *Amycolatopsis mediterranei* S699. *Chem. Biol.* 5: 69-79.
- Alberati-Giani, D., Cesura, A.M., Broger, C., Warren, W.D., Rover, S., and

- Malherbe P. 1997. Cloning and functional expression of human kynurenine 3-monoxygenase. *FEBS Lett.* 410: 407-12.
- Altschul, S.F., Madden T.L., Schäffer, A.A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D.J. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucl. Acids Res.* 25: 3389-3402.
- Altschul, S.F., Gish, W., Miller, W., Myers, E.W. and Lipman, D.J. 1990. Basic local alignment search tool. *J. Mol. Biol.* 215: 403-410.
- Bellamacina, C.R. 1996. The nicotinamide dinucleotide binding motif: a comparison of nucleotide binding proteins. *FASEB J.* 10: 1257-69.
- Botstein, K., Lew, K.K., Jarvik, V., and Swanson, C.A. 1975. Role of antirepressor in the bipartite control of repression and immunity by bacteriophage P22. *J. Mol. Biol.* 91: 439-62.
- Cox, R.J., Hitchman, T.S., Byrom, K.J., Findlow, I.S., Tanner, J.A., Crosby, J., and Simpson, T.J. 1997. Post-translational modification of heterologously expressed *Streptomyces* type II polyketide synthase acyl carrier proteins. *FEBS Lett.* 405: 267-72.
- Duran, N., Antonio, R.V., Haun, M. and R.A. Pilli. 1994. Biosynthesis of a trypanocide by *Chromobacterium violaceum*. *World J. Micro. Biotech.* 10: 686-690.
- Ewing, B., Hillier, L., Wendl, M.C., and Green, P. 1998. Base-calling of automated sequencer traces using phred. I. Accuracy assessment. *Genome Res.* 8:175-85.
- Ewing, B., and Green, P. 1998. Base-calling of automated sequencer traces using phred. II. Error probabilities. *Genome Res.* 8:186-94.
- Gehring, A.M., Lambol, R.H., Vogel, K.W., Druceckhammer, D.G., and Walsh, C.T. 1997 Ability of *Streptomyces* spp. acyl carrier proteins and coenzyme A analogs to serve as substrates *in vitro* for *E. coli* holo-ACP synthase. *Chem. Biol.* 4:17-24.
- Gordon D., Abajian C., and Green, P. 1998. Consed: a graphical tool for sequence finishing. *Genome Res.* 8:195-202.
- Gribskov, M., McLachlan, A.D., and Eisenberg, D. 1987. Profile analysis: detection of distantly related proteins. *Proc. Natl. Acad. Sci. USA.* 84: 4355-4358.
- Hoshino, T., Hayashi, T., and Odajima, T. 1995. Biosynthesis of Violacein: Oxygenation at the 2-position of the indole ring and structures of proviolacein, prodeoxyviolacein and pseudoviolacein, the plausible biosynthetic intermediates of violacein and deoxyviolacein. *J. Chem. Soc. Perkin Trans. 1:* 1565-1571.
- Hoshino, T., Takano, T., Hori, S., and Ogasawara, N. 1987. Biosynthesis of Violacein: Origins of hydrogen, nitrogen, and oxygen atoms in the 2-pyrrolidone nucleus. *Agric. Biol. Chem.* 51: 2733-2741.
- Hoshino, T., and Yamamoto, M. 1997. Conversion from tryptophan precursor into violacein pigments by a cell-free system from *Chromobacterium violaceum*. *Biosci. Biotech. Biochem.* 61: 2134-2136.
- Kealer, J.T., Liu, L., Santi, D.V., Betlach, M.C., and Barr P.J. 1998. Production of a polyketide natural product in non-polyketide-producing prokaryotic and eukaryotic hosts. *Proc. Nat. Acad. Sci. USA.* 95: 505-509.
- Malpartida, F., and Hopwood, D.A. 1984. Molecular cloning of the whole biosynthetic pathway of a *Streptomyces* antibiotic and its expression in a heterologous host. *Nature.* 309: 462-464.
- Momen, A., Mizuoka, T., and Hoshino, T. 1998. Studies on the Biosynthesis of Violacein. Part 9. Green Pigments possessing tetraindole and dipyrromthene moieties, chromoviridans and deoxychromoviridans produced by a cell-free extract of *Chromobacterium violaceum* and their biosynthetic origins. *J. Chem. Soc., Perkin Trans. 1:* 3087-3092.
- Obara, K., Otsuka-Fuchino, H., Sattayasai, N., Nonomura, Y., Tsuchiya, T., and Tamiya, T. 1992. Molecular cloning of the antibacterial protein of the giant African snail, *Achatina fulica* Ferussac. *Eur. J. Biochem.* 209:1-6.
- Pemberton, J.M., Vincent, K.M., and R.J. Penfold. 1991. Cloning and heterologous expression of the violacein biosynthesis gene cluster from *Chromobacterium violaceum*. *Current Microbiol.* 22: 355-358.
- Projan, S.J., Kornblum, J., Kreiswirth, B., Moghazeh, S.L., Eisner, W., and Novick, R.P. 1989. Nucleotide sequence: the beta-hemolysin gene of *Staphylococcus aureus*. *Nucleic Acids Res.* 17: 3305.
- Raibekas, A.A., and Massey, V. 1998. Primary structure of the snake venom L-amino acid oxidase shows high homology with the mouse B cell interleukin 4-induced Fig1 protein. *Biochem. Biophys. Res. Commun.* 248: 476-478.
- Takamatsu, N., Shiba, T., Muramoto, K., and Kamiya, H. 1995. Molecular cloning of the defense factor in the albumen gland of the sea hare *Aplysia kurodai*. *FEBS Lett.* 377: 373-6.
- Yamazaki, M., Kimura, K., Kisugi, J., Muramoto, K., and Kamiya, H. 1989. Isolation and characterization of a novel cytolytic factor in purple fluid of the sea hare *Aplysia*. *Cancer Res.* 49: 3834-8.

