Cloning, Sequencing, and Heterologous Expression of the *elmGHIJ* Genes Involved in the Biosynthesis of the Polyketide Antibiotic Elloramycin from *Streptomyces olivaceus* Tü2353

Elpidio R. Rafanan, Jr.,[†] Linda Le,[†] Lili Zhao,[†] Heinrich Decker,[‡] and Ben Shen*,[†]

Department of Chemistry, University of California, One Shields Avenue, Davis, California 95616, and Hoechst Marion Roussel Deutschland GmbH, Process Development, D-65926 Frankfurt, Germany

Received January 12, 2001

Elloramycin A (**1**) belongs to a small family of naphthacenequinones characterized by a unique highly hydroxylated cyclohexenone moiety. A cosmid clone 16F4, harboring genes for the production of **1** from *Streptomyces olivaceus* Tü2353, has been previously isolated. DNA sequence analysis of a 3.2-kb fragment from 16F4 revealed four open reading frames—the *elmGHIJ* genes. Heterologous expressions of the *elmGHI* genes in either *Escherichia coli* or *Streptomyces lividans*, followed by biochemical characterizations of the ElmGHI proteins, established ElmG as tetracenomycin B2 oxygenase, ElmH as tetracenomycin F1 monooxygenase, and ElmI as tetracenomycin F2 cyclase. These results provide direct biochemical evidence for the hypothesis that the biosynthesis of **1** in *S. olivaceus* parallels that of tetracenomycin C (**2**) in *Streptomyces glaucescens* and support the notion that the biosynthesis of the highly hydroxylated cyclohexenone moiety in other polyketides most likely follows the same paradigm as the tetracenomycin B2 or A2 oxygenase.

Elloramycin (Elm) A (**1**), an antitumor and antibacterial antibiotic produced by *Streptomyces olivaceus* Tü2353,[1] belongs to a small group of naphthacenequinones characterized by a unique structural feature of the highly hydroxylated cyclohexenone moiety (boxed in Figure 1). Other examples of this family include tetracenomycin (Tcm) C[2] (**2**), Tcm X,[3] dutomycin,[4] and viridicatumtoxin.[5] The biosynthesis of **2** in *Streptomyces glaucescens* GLA.0 has been studied extensively with the isolation of all biosynthetic intermediates, cloning and characterization of the entire *tcm* gene cluster, and purification and biochemical investigation of individual enzymes.[6] The striking structural similarities between **1** and **2** have inspired us to propose that the biosynthesis of **1** in *S. olivaceus* parallels that of **2** in *S. glaucescens*. This hypothesis is consistent with (1) the accumulation of several known biosynthetic intermediates of **2**, including Tcm D3 (**3**) and Tcm B3 (**4**), by blocked mutants of *S. olivaceus*[7] and (2) the isolation of cosmid clone 16F4 from the *S. olivaceus* genomic library, the introduction of which into *Streptomyces lividans* resulted in the production of **4** and 8-demethyl-Tcm C[8] (**5**) (Figure 2). Here we now report the sequencing and characterization of a 3.2-kb fragment from the 16F4 cosmid, revealing four complete open reading frames, designated the *elmGHIJ* genes. Heterologous expressions of the *elmGHI* genes in either *Escherichia coli* or *S. lividans* and biochemical examination of the ElmGHI proteins provided direct evidence supporting the proposed biosynthetic pathway for **1**.

Results and Discussion

Isolation from Cosmid 16F4 and DNA Sequences of the *elmGHIJ* Genes. We have previously reported[8] that introduction of cosmid 16F4 into *S. lividans* resulted in the production of **5**, suggesting that 16F4 likely harbors all genes for the production of **1** except those for the biosynthesis of the permethylated L-rhamnose moiety. Since the biosynthetic gene cluster of **1** appears to resemble that of **2**,[8] we proceeded to isolate the *elmGHIJ* locus from 16F4 using homologous genes from the *tcm* cluster as probes. Cosmid 16F4 was analyzed by Southern hybridization with *tcmG*[9] as a probe. A single 5.0-kb *Bgl*II fragment that hybridizes to the *tcmG* probe was isolated and cloned (pBS4001), the nucleotide (nt) sequence of a 3.2-kb fragment of which was determined.

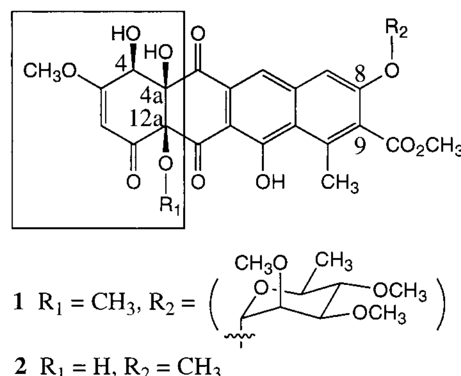

Figure 1. Structures of elloramycin A (**1**) and tetracenomycin C (**2**).

Four complete open reading frames (ORFs), transcribed in the same direction, were identified within the 3.2-kb fragment of pBS4001 (Figure 3), the 3164-nt sequence of which has been deposited in the GenBank database under accession number AF263463. The first ORF, designated *elmG*, most likely begins with an ATG at nt 339, preceded by a probable ribosome binding site (RBS), GGAGA, and ends with a TGA stop codon at nt 1958. *elmG* should therefore encode a 539-amino acid protein with a molecular weight of 58 085 and an isoelectric point of 6.04. The deduced product of *elmG* is highly homologous to a family of known and putative FAD-containing, NADPH-dependent oxygenases, including TcmG[9] from *S. glaucescens* (69.8% similarity, 64.2% identity), DntB[10] from *Burkholderia* sp. strain DNT (50.6% similarity, 43.5% identity), MmcT[11] from *Streptomyces lavendulae* (50.0% similarity, 44.0%

* To whom correspondence should be addressed. Tel: (530) 754-9382. Fax: (530) 752-8995. E-mail: shen@chem.ucdavis.edu.
[†] University of California, Davis.
[‡] Hoechst Marion Roussel Deutschland GmbH.

10.1021/np010007+ CCC: $20.00 © 2001 American Chemical Society and American Society of Pharmacognosy
Published on Web 03/22/2001

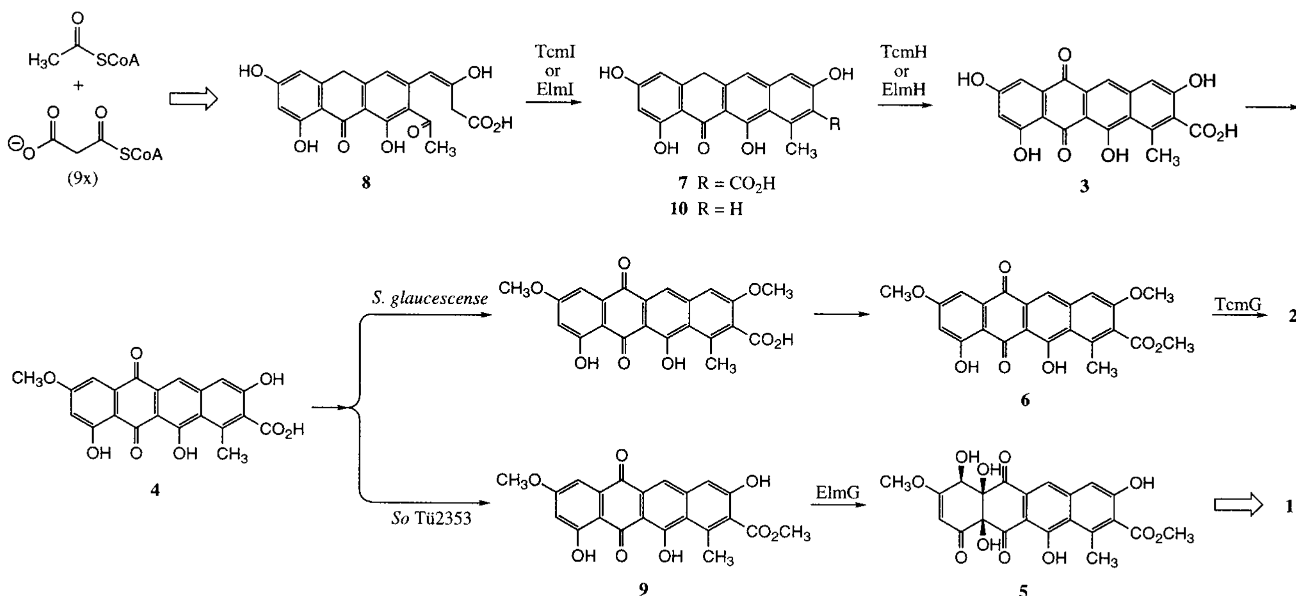

Figure 2. Biosynthetic pathway of tetracenomycin C (**1**) in *S. glaucescens* and of elloramycin A (**2**) in *S. olivaceus*.

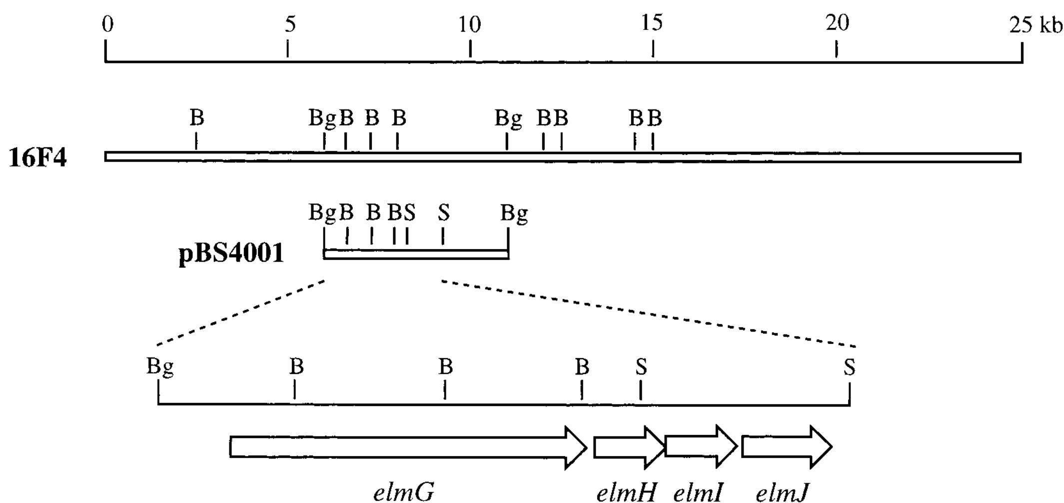

Figure 3. Restriction map of the 25-kb *elm* biosynthetic gene cluster from *S. olivaceus* harbored by cosmid 16F4 and genetic organization of the *elmGHIJ* genes. B, *Bam*HI; Bg, *Bgl*II; S, *Sac*I.

identity), SchC[12] from *Streptomyces halstedii* (49.7% similarity, 43.8% identity), RdmE[13] from *Streptomyces puspurascens* (47.5% similarity, 39.2% identity), and DnrF[14] from *Streptomyces peucetius* (46.4% similarity, 36.8% identity). The functions of TcmG to catalyze the triple hydroxylation of Tcm A2 (**6**) to **2**[15-17] and DnrF to catalyze the hydroxylation of aklavinone to ϵ-rhodomycinone[14] have been confirmed by in vitro enzyme assays following expressions of the *tcmG* gene in *S. lividans* and of the *dnrF* gene in either *S. lividans* or *E. coli*. The function of DntB to catalyze the oxidation of 4-methyl-5-nitrocatechol to 2-hydroxy-5-methylquinone was proven directly by following the oxygenase activity, leading to the purification of the DntB protein to homogeneity and the cloning of the *dntB* gene.[10]

Located 31 nt downstream of *elmG* are the second and third ORFs, designated *elmH* and *elmI*, which are translationally coupled. *elmH* most likely begins with an ATG at nt 1990, preceded by a probable RBS, GGAGG, and ends with a TGA stop codon at nt 2319. *elmH* should therefore encode a 109-amino acid protein with a molecular weight of 12 304 and an isoelectric point of 5.91. *elmI* most likely starts with an ATG at nt 2316, preceded by a probable RBS, GGAGG, and ends with a TGA stop codon at nt 2636. *elmI* should therefore encode a 106-amino acid protein with a molecular weight of 12 421 and an isoelectric point of 4.62. The deduced product of *elmH* is closely related to a family of known or putative oxygenases, including TcmH[18] from *S. glaucescens* (59.0% similarity, 52.7% identity), ActVA-ORF6[19] from *Streptomyces coelicolor* (45.2% similarity, 39.4% identity), and ORF7 of the jadomycin biosynthetic gene cluster[20] from *Streptomyces venezuelae* (37.4% similarity, 27.3% identity). While direct evidence for jadomycin ORF7 as an oxygenase, presumably catalyzing the oxidation of a benz[*a*]anthranone intermediate to rebelomycin,[20] is lacking, both TcmH[21] and ActVA-ORF6[22] have been characterized as monooxygenases, catalyzing the oxidation of the naphthacenone Tcm F1 (**7**) into naphthacenequinone **3**. The *elmI* gene product shows strong homology to a family of known and putative polyketide cyclases, including TcmI[18] from *S. glaucescens* (76.6% similarity, 71.0% identity), UrdF[23] from *Streptomyces fradiae* (49.1% similarity, 35.8% identity), LanF[24] from *Streptomyces cyanogenuse* (48.1% similarity, 34.0% identity), JadI[25] from *S. venezuelae* (46.2% similarity, 34.0% identity), and WhiE-ORFVII[26] from *S. coelicolor* (43.4% similarity, 37.7% identity). Although direct characterization of polyketide cyclases in general has been hindered by the lack of the presumed

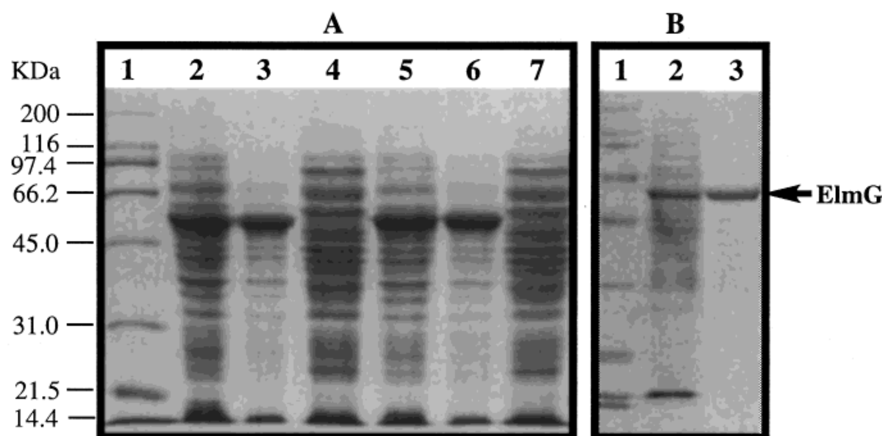

Figure 4. Heterologous expression and purification of ElmG. (A) ElmG overproduced in *E. coli* BL-21(DE-3)(pBS4004) (lanes 2–4) and *N*-His$_6$-tagged ElmG overproduced in *E. coli* BL-21(DE-3)(pBS4003) (lanes 5–7), cultured at 30 °C with 100 μM IPTG induction. Lane 1, MW Std; lanes 2 and 5, total proteins; lanes 3 and 6, insoluble proteins; lane 4 and 7, soluble proteins. (B) ElmG overproduced in *S. lividans* (pBS4006). Lane 1, MW Std; lane 2, total proteins; lane 3, purified ElmG.

substrates, the function of JadI to facilitate cyclization of decapolyketides into angucyclinones has been demonstrated by expression in *S. lividans* of a jadomycin gene cassette consisting of the *jadABCDEI* genes.[27] TcmI has been previously purified and characterized as Tcm F2 cyclase, catalyzing the intramolecular aldol condensation of Tcm F2 (**8**) to **7**.[28]

The fourth ORF, designated *elmJ*, is located 11 nt downstream of *elmI*. It most likely starts with an ATG at nt 2648, preceded by a probable RBS, GGAGA, and ends with a TGA stop codon at nt 3079. *elmJ* should therefore encode a 143-amino acid protein with a molecular weight of 15 251 and an isoelectric point of 6.90. The deduced *elmJ* product appears to be closely related to TcmJ[18] from *S. glaucescens* (70.0% similarity, 64.2% identity), SchB[29] from *S. halstedii* (61.6% similarity, 53.6% identity), WhiE-ORFII[26] from *S. coelicolor* (58.2% similarity, 49.6% identity), and CurC[30] from *Streptomyces cyaneus* (57.0% similarity, 46.5% identity). The function of this family of proteins remains a mystery. It has been shown in *S. glaucescens* that TcmJ is not essential but can influence the production of **2** profoundly.[18,31] The latter result led to the speculation that TcmJ could provide an optimal activity of the Tcm polyketide synthase multiprotein complex by facilitating the interaction among the constitutive proteins.[31] On the basis of the high degree of amino acid sequence conservation, an equivalent role could be proposed for ElmJ, which could enhance the production of **1** in *S. olivaceus* by interaction with the Elm polyketide synthase.

Heterologous Expressions of *elmG* in *E. coli* and *S. lividans*. *elmG* was efficiently expressed in *E. coli* BL-21(DE-3)(pBS4004) as the native protein or in *E. coli* BL-21(DE-3)(pBS4003) as an *N*-His$_6$-tagged fusion protein (Figure 4A). The overproduced ElmG and its *N*-His$_6$-tagged fusion protein migrated as distinctive bands with a M_r of 58 000 (Figure 4A, lane 2) or 60 000 (Figure 4A, lane 5), respectively, consistent with the predicted M_r of 58 085 and 60 266 from the *elmG* gene and that plus the *N*-His$_6$-tag. However, the recombinant proteins were completely insoluble. The solubility of both ElmG (Figure 4A, lanes 3 vs 4) and its *N*-His$_6$-tagged fusion proteins (Figure 4A, lanes 6 vs 7) could not be improved under all conditions investigated, including the lowering of IPTG concentration or incubation temperature. Since it is known that genes of *Streptomyces* origin are more likely expressed in soluble form in a *Streptomyces* host, *elmG* was next expressed in *S. lividans*. pBS4006 was constructed in pWHM3 so as to place the expression of *elmG* under the control of the *ermE** promoter. Analysis of cell-free extract of *S. lividans* (pBS4006) showed a distinctive band that migrated with a M_r of 58 000 (Figure 4B, lane 2), consistent with the predicted M_r of 58 085 from the *elmG* gene. ElmG overproduced in *S. lividans* was completely soluble. Purification of the ElmG protein to homogeneity (Figure 4B, lane 3) was accomplished by a procedure described previously.[15] Since Tcm B2 (**9**), the most likely substrate for ElmG, was not available, **6** was used as a substrate to test the enzyme activity of the purified ElmG. The purified ElmG catalyzes quantitative conversion of **6** into **2**, as determined by HPLC analysis with authentic standards **6** and **2** as references. The observed t_R (min) for **6** (20.3) and **2** (11.5) are consistent with those reported previously.[15] ElmG utilizes either NADH or NADPH as an electron donor and requires molecular O_2.[16]

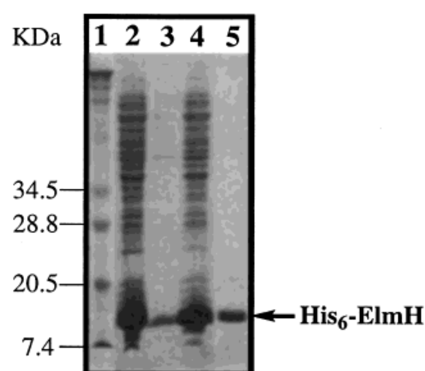

Figure 5. Heterologous expression and purification of *N*-His$_6$-tagged ElmH from *E. coli* BL-21(DE-3)(pBS4009) cultured at 25 °C with 100 μM IPTG induction. Lane 1, MW Std; lane 2, total proteins; lane 3, insoluble proteins; lane 4, soluble proteins; lane 5, purified *N*-His$_6$-tagged ElmH.

Heterologous Expression of *elmH* and *elmI* in *E. coli*. *elmH* was efficiently expressed in *E. coli* as an *N*-His$_6$-tagged fusion protein. While incubation of *E. coli* BL-21-(DE-3)(pBS4009) at 37 °C upon induction with 400 μM IPTG resulted in the overproduction of ElmH predominantly in an insoluble form, lowering the incubation temperature to 25 °C and the IPTG concentration to 100 μM dramatically improved the solubility of ElmH. Under the latter conditions, ElmH was overproduced predominantly in the soluble form (Figure 5, lanes 3 vs 4), from which it was purified to homogeneity (Figure 5, lane 5) by affinity chromatography on Ni–NTA resin. The purified ElmH protein migrated with a M_r of 15 000, consistent with

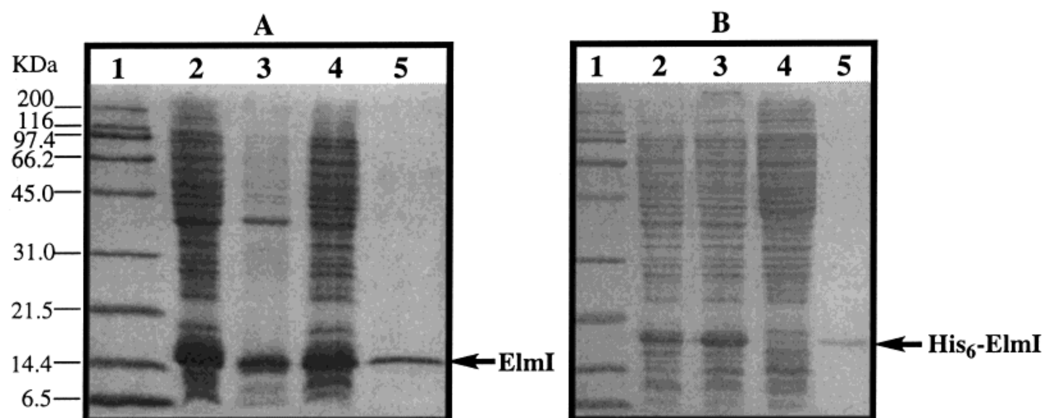

Figure 6. Heterologous expression and purification of ElmI. (A) ElmI overproduced in *E. coli* BL-21(DE-3)(pBS4011) and (B) *N*-His$_6$-tagged ElmI overproduced in *E. coli* BL-21(DE-3)(pBS4010), cultured at 25 °C with 100 μM IPTG induction. Lane 1, MW Std; lane 2, total proteins; lane 3, insoluble proteins; lane 4, soluble proteins; lane 5, purified ElmI or *N*-His$_6$-tagged ElmI.

the predicted M_r of 14 485 from the *elmH* gene plus the *N*-His$_6$-tag. The purified ElmH catalyzed the quantitative oxidation of **7** to **3**, and this reaction was confirmed by TLC analysis with authentic standards of **7** and **3** as references. The observed R_f values for **7** (0.33) and **3** (0.26) are consistent with those reported previously.[21,22,28] The oxidation of **7** by ElmH depends solely on molecular O_2, and ElmH does not require any exogenous cofactors or metal ions.

elmI was efficiently expressed in *E. coli* BL-21(DE-3)-(pBS4011) as the native protein or in *E. coli* BL-21(DE-3)-(pBS4010) as an *N*-His$_6$-tagged fusion protein (Figure 6). Like *elmH*, incubation of either *E. coli* BL-21(DE-3)-(pBS4011) or *E. coli* BL-21(DE-3)(pBS4010) at 37 °C upon induction with 400 μM IPTG resulted in the overproduction of ElmI predominantly in an insoluble form, and lowering the incubation temperature to 25 °C and the IPTG concentration to 100 μM significantly improved the solubility of ElmI. Under the latter conditions, ElmI was overproduced predominantly in the soluble form (Figure 6A, lanes 3 vs 4), while the *N*-His$_6$-tagged ElmI fusion protein was overproduced partially in the soluble form (Figure 6B, lanes 3 vs 4). ElmI was purified to homogeneity (Figure 6A, lane 5) by a procedure reported previously,[28] and the *N*-His$_6$-tagged ElmI fusion protein was purified to homogeneity (Figure 6B, lane 5) by affinity chromatography on Ni–NTA resin. The purified ElmI and its *N*-His$_6$-tagged fusion protein migrated with a M_r of 15 000 and 17 000, respectively, consistent with the predicted M_r of 12 421 and 14 602 from the *elmI* gene and that plus the *N*-His$_6$-tag. Both ElmI and its *N*-His$_6$-tagged fusion protein were assayed for the cyclase activity, and both proteins are active in catalyzing the intramolecular aldol condensation of **8** to yield **7** with 9-decarboxyl-Tcm F1 (**10**) as a minor product when the enzymatic reaction was carried out at pH 8.5. The formation of **7** and **10** from **8** was confirmed by both TLC and HPLC analyses with authentic standards **8** and **7** as references. The observed R_f for **8** (0.52), **7** (0.33), and **10** (0.83) and t_R (min) for **8** (9.7), **7** (11.9), and **10** (14.0) are consistent with those reported previously.[28]

Except for the additional C-12a-*O*-methyl group and the C-8-*O*-permethylated L-rhamnose moiety, **1** shares the identical chromophore with **2**.[1,2] This structural resemblance has promoted the hypothesis that the biosynthesis of **1** in *S. olivaceus* parallels that of **2** in *S. glaucescens*. The latter is consistent with (1) the isolation of known biosynthetic intermediates of **2**, such as **3** and **4**, from blocked mutants of *S. olivaceus*;[7] (2) the complementation with *elm* genes to *tcm* mutants defective in the *tcmKLM* polyketide synthase genes;[8] and (3) the production of **4** and **5** in *S. lividans* upon expression of the cosmid clone 16F4 harboring a part of the *elm* gene cluster.[8] The fact that **4** is the most advanced intermediate identified from both pathways suggests that **1** and **2** might share a similar pathway in the early stage of their biosynthesis but branch off at **4**. Thus, in *S. glaucescens* **4** is doubly *O*-methylated at the C-8-OH and C-9-CO_2H groups to yield **6**, which undergoes the Tcm A2 oxygenase-catalyzed triple hydroxylation to **2**. In *S. olivaceus,* however, **4** is selectively *O*-methylated at the C-9-CO_2H group to **9**. The latter is then triply hydroxylated into **5**, presumably catalyzed by an enzyme analogous to the Tcm A2 oxygenase, followed by *O*-methylation at the C-12a-OH and glycosylation at the C-8-OH groups of **5** to give **1** (Figure 2). This hypothesis is now supported directly by (1) the cloning and sequencing of *elmGHIJ* genes from *S. olivaceus* and (2) the heterologous expressions and biochemical characterizations of the ElmGHI proteins. The deduced gene products of *elmGHIJ* are highly homologous to TcmGHIJ,[9,18] suggesting that they could play similar roles as TcmGHIJ in the biosynthesis of **1** in *S. olivaceus*. The latter predictions were confirmed by direct characterizations of the purified ElmG, ElmH, and ElmI proteins, which displayed biochemical and kinetic (data not shown) properties very similar to those of TcmG,[15,16] TcmH,[21] and TcmI,[28] respectively. Hence, in analogy with TcmGHI, we name ElmG as Tcm B2 oxygenase, presumably catalyzing the triple hydroxylation of **9** to **5** in vivo,[8] ElmH as Tcm F1 monooxygenase, catalyzing the oxidation of **7** to **3**, and ElmI as Tcm F2 cyclase, catalyzing the intramolecular aldol condensation of **8** to **7**. Although, in the absence of **9**, we used **6** as a substrate analogue to demonstrate the oxygenase activity of ElmG, we prefer to name ElmG as Tcm B2 oxygenase on the basis that **5** has been demonstrated previously as an intermediate for **1**, presumably resulting from direct triple hydroxylation of **9**.[8]

ElmH[21,22] belongs to an unusual class of monooxygenases that oxidize anthranones or naphthacenones to the corresponding anthraquinones or naphthacenequinones. The mechanism of this oxidation is very intriguing because these enzymes appear to require no cofactor or metal ions. Although a number of putative monooxygenases of this family are known,[18–22] ElmH, TcmH,[21] and ActVA-ORF6[22] are the only three members of this family that have been purified and characterized biochemically. We have previously shown that histidine residues were essential for TcmH,[21] and Marsh and co-workers proposed that histidine-52 acts as a general base during catalysis for ActVA-

ORF6.[22] Sequence analysis indeed showed that the latter histidine residue is absolutely conserved among ElmH and its homologues. We envisage that the cloning and heterologous expression of *elmH* and the purification of ElmH reported here shall facilitate the mapping of the active sites of this family of novel monooxygenases.

Polyketide cyclases catalyze the intramolecular aldol condensations in the synthesis of aromatic polyketides. Although genetic studies have identified a number of putative polyketide cyclases,[18,23−26] little is known about the mechanism of these enzymatic aldol condensations. DnrD,[32] TcmI,[28] and TcmN[33] are the only other polyketide cyclases that have been purified and biochemically characterized. However, no sequence homology is observed among DnrD, TcmN, and TcmI or ElmI, consistent with the fact that they utilize different substrates. In contrast, ElmI and TcmI are highly homologous and act on the same substrate. Of special interest are the three histidine residues that are conserved between ElmI and TcmI. Preliminary study has already shown that ElmI can be completely inactivated by diethylpyrocarbonate, suggesting that histidine may play a critical role in enzyme catalysis. Therefore, comparative studies of ElmI and TcmI could provide an understanding of the mechanism and active sites of these proteins, and the cloning and heterologous expression of *elmI* and the purification of ElmI reported here have set the stage for such investigations.

The characterizations of ElmG[16,17] for **1** in *S. olivaceus*, together with TcmG[9,15−17] for **2** in *S. glaucescens*, support the notion that the biosynthesis of the highly hydroxylated cyclohexenone moiety in other polyketides most likely follows the same paradigm as the Tcm B2 or Tcm A2 oxygenase. Although ElmG and TcmG are the only two examples known to date that catalyze this unprecedented triple hydroxylation of an aromatic compound, it is interesting to note that sequence analysis has revealed that both ElmG and TcmG are related to other FAD-containing, NADPH-dependent hydroxylases,[11−15] such as DnrF, a known monooxygenase catalyzing the hydroxylation of aklavinone to ϵ-rhodomycinone.[15] The latter is in fact consistent with the proposed mechanism for TcmG, in which **6** is first hydroxylated into 4-hydroxyl-Tcm A2, presumably by a monooxygenase-like activity.[15−17] The cloning and heterologous expression of *elmG* and the purification of ElmG reported here should greatly facilitate future investigations to clarify the exact mechanism of this intriguing hydroxylation reaction.

Experimental Section

Bacterial Strains, Plasmids, and Other Materials. *E. coli* DH5(α)[34] was used as a general host for routine subcloning. *E. coli* BL-21 (DE-3) (Novagen, Madison, WI) and *S. lividans* 1326[35] were used as expression hosts. pGEM-7zf(+), pGEM-9zf(+), and pGEM-T were from Promega (Madison, WI), pUC19 was from New England Biolabs (Beverly, MA), and pET-28a(+) and pET-29a(+) were from Novagen. pWHM3,[36] pWHM63,[15] and 16F4[8] were described previously. Ampicillin and kanamycin were from Fisher Scientific, Inc. (Fair Lawn, NJ), and thiostrepton was a gift from Sal Luciana at the Squibb Institute for Medical Research (Princeton, NJ). Unless specified otherwise, restriction enzymes and other molecular biology reagents, biochemicals, and chemicals were from standard commercial sources.

DNA Manipulation, Sequencing, and Analysis. Plasmid preparation and DNA extraction were carried out using commercial kits (Qiagen, Santa Clarita, CA). Restriction endonuclease digestion, dephosphorylation, ligation, and other molecular biology procedures were done according to standard methods.[34] For Southern analysis, digoxigenin labeling of the *tcmG* probe,[9] hybridization, and detection were performed by the protocols provided by the manufacturer (Boehringer-Mannheim, Mannheim, Germany), and Hybond N nylon membranes (Amersham, Braunschweig, Germany) were used. Automated DNA sequencing was carried out on an ABI Prism 377 DNA sequencer using the ABI Prism dye terminator cycle sequencing ready reaction kit and Ampli*Taq* DNA polymerase FS (Perkin-Elmer/ABI, Foster City, CA). Both strands were sequenced, and the sequencing service was provided either by DBS Automated DNA Sequencing Facility, University of California, Davis, or by Davis Sequencing, Inc. (Davis, CA). Data were analyzed by ABI Prism Sequencing 2.1.1 and Genetics Computer Group (GCG) Wisconsin Package (Version 9.1, Madison, WI). The deduced gene products were compared with proteins in the GenBank through the NCBI Internet site using BLASTP 2.0.12.[37]

Polymerase Chain Reaction (PCR). Primers were synthesized at the Protein Structure Laboratory, University of California, Davis. The 5.0-kb *Bgl*II fragment of cosmid 16F4, harboring the *elmGHIJ* genes, was cloned into the *Bam*HI site of pUC19 to yield pBS4001 and was used as a template for amplification of the desired genes. PCR was carried out on a GeneAmp PCR System 2400 (Perkin-Elmer/ABI) with *Vent* DNA polymerase from New England Biolabs. A typical PCR mixture consisted of 44 ng of the template DNA, 50 pmol of each primer, 250 μM of dNTP, 5% (v/v) DMSO, 1 unit of *Vent* DNA polymerase, and 1 × buffer in a final volume of 50 μL. The PCR temperature program was as follows: initial denaturing at 96 °C for 5 min, 30−35 cycles of 45 s at 94 °C, 1 min at 55 °C, and 2 min at 72 °C, followed by an additional 7 min at 72 °C. For *elmG*, the following pair of primers was used: 5′-TGC CGA CCA TAT GGA CCG AAT AGA AAT TC-3′ (forward, the *Nde*I site is underlined) and 5′-GCG AAG CTT CAG GTG CGG GCC AGG ATC-3′ (reverse, the *Hind*III site is underlined). A distinctive product with the predicted size of 1.62 kb was amplified and cloned into pGEM-T according to the protocol provided by the manufacturer (Promega) to yield pBS4002. For *elmH*, the following pair of primers was used: 5′-AGT ATT CAT ATG CCC CGA CTG TCC CCC GAC-3′ (forward, the *Nde*I site is underlined) and 5′-TAT GGA TCC TCA TCC GGC GGC TTC CTC GAA-3′ (reverse, the *Bam*HI site is underlined). A distinctive product with a predicted size of 330 bp was amplified. For *elmI*, the following pair of primers was used: 5′-AGT ATT CAT ATG ACC TAC CGT GCC CTG-3′ (forward, the *Nde*I site is underlined) and 5′-ATG AAT TCT CAG GCT TCC CAG GTG-3′ (reverse, the *Eco*RI site is underlined). A distinctive product with a predicted size of 320 bp was amplified.

Heterologous Expressions of the *elmGHI* Genes in *E. coli* or *S. lividans* and Purification of the ElmGHI Proteins. To express *elmG* in *E. coli*, the *elmG* gene was cloned as a 1.6-kb *Nde*I-*Hind*III fragment from pBS4002 into the same sites of pET-28a(+) or pET29a(+) to yield pBS4003 or pBS4004, respectively, in the former of which *elmG* will be expressed as an *N*-His$_6$-tagged fusion protein. *E. coli* BL-21 (DE-3)(pBS4003) and *E. coli* BL-21 (DE-3)(pBS4004) were grown in LB[34] containing kanamycin (30 μg/mL) under conditions recommended by Novagen. To improve the solubility of the recombinant protein, a lower concentration of IPTG (100 μM) and lower incubation temperatures (30 °C and 25 °C) were also tried. To express *elmG* in *S. lividans*, pBS4001 was digested with *Xba*I and *Sac*I, and the resulting 2.2-kb *Xba*I-*Sac*I fragment containing *elmG* was first cloned into the same sites of pGEM-9zf(+) to yield pBS4005. To clone the *ermE** promoter,[38] pWHM63 was digested with *Kpn*I and *Sac*I, and the resulting 0.45-kb *Kpn*I-*Sac*I fragment was inserted into the same sites of pGEM-7zf(+) to yield pBS4006. The *elmG* gene was then moved as a 2.2-kb *Hind*III-*Nsi*I fragment from pBS4005 into the same sites of pBS4006 to yield pBS4007. The latter was finally digested with *Eco*RI and *Nsi*I, and the resulting 2.6-kb *Eco*RI-*Nsi*I fragment of *ermE*::elmG* was cloned into the *Eco*RI and *Pst*I sites of pWHM3 to yield pBS4008, in which the expression of *elmG* is under the control of the *ermE** promoter. pBS4008 was introduced into *S. lividans* by ethylene glycol-mediated protoplast transforma-

tion,[35] selected with thiostrepton (50 μg/mL). The resulting *S. lividans* (pBS4008) transformants were grown in R2YE with thiostrepton (10 μg/mL) and incubated in a rotary shaker (Series 25, New Brunswick Scientific, Inc., Edison, NJ) at 300 rpm and 30 °C for 2 days. Purification of ElmG was carried out by following essentially the procedure reported previously for TcmG.[15]

To express *elmH* in *E. coli*, the PCR-amplified 330-bp fragment of the *elmH* gene was digested with *Nde*I and *Bam*HI and cloned into the same sites of pET-28a(+) to yield pBS4009, in which *elmH* will be expressed as an *N*-His_6-tagged fusion protein. *E. coli* BL-21 (DE-3)(pBS4009) was grown in LB[34] containing kanamycin (30 μg/mL) under conditions recommended by Novagen. Purification of the *N*-His_6-tagged ElmH fusion protein was carried out by affinity chromatography on Ni−NTA resin under the recommended conditions by Qiagen.

To express *elmI* in *E. coli*, the PCR-amplified 320-bp fragment of the *elmI* gene was digested with *Nde*I and *Eco*RI and cloned into the same sites of both pET-28a(+) and pET-29a(+) to yield pBS4010 and pBS4011, respectively, in the former of which *elmI* will be expressed as an *N*-His_6-tagged fusion protein. *E. coli* BL-21 (DE-3)(pBS4010) and *E. coli* BL-21 (DE-3)(pBS4011) were grown in LB[34] containing kanamycin (30 μg/mL) under conditions recommended by Novagen. Purification of ElmI was accomplished by following essentially the procedure described previously for TcmI,[28] and purification of the *N*-His_6-tagged ElmI fusion protein was carried out by affinity chromatography on Ni−NTA resin under the recommended conditions by Qiagen.

Enzyme Assays. The authentic substrates **8**,[28] **7**,[21] and **6**[15] and products **10**,[28] **3**,[21] and **2**[15] were isolated and characterized as described previously. Enzyme assays for ElmI using **8** as a substrate,[28] for ElmH using **7** as a substrate,[21] and for ElmG using **6** as a substrate[15] were carried out by the procedures described previously. Briefly, a typical assay solution (250 μL) for ElmI consisted of 100 μM **8** and 4 μg of ElmI in 100 mM Tris-HCl, pH 8.0, that for ElmH consists of 80 μM **7**, 20 μg of ElmH, and 1 mM DTT in 100 mM sodium phosphate, pH 7.5, and that for ElmG consisted of 100 μM **6**, 250 μM NADPH, 30 μg of ElmG, 1 mM DTT, and 10% (v/v) DMSO in 50 mM Tris-HCl, pH 9.0. The assays were initiated by addition of substrates **8**, **7**, or **6**, respectively, incubated at 30 °C for 0.5 to 2 h, and terminated by addition of NaH_2PO_4 to saturation and extraction with EtOAc (3 × 250 μL). The EtOAc extracts were concentrated in vacuo to dryness. The residues were dissolved in 50−100 μL of MeOH and analyzed by HPLC (for ElmG and ElmI assays) or TLC (for ElmH and ElmI assays). HPLC analysis was performed on a C_{18} column (Nova-Pak, 4 μm, 8 mm × 10 cm, Waters, Milford, MA), developed with $CH_3CN/H_2O/AcOH$ (20:80:0.1, v/v) for 4 min, followed by a linear gradient to $CH_3CN/AcOH$ (99.9:0.1, v/v) for 10 min, and held at $CH_3CN/AcOH$ (99.9:0.1, v/v) for an additional 7 min, at flow rate of 0.8 mL/min and with UV detection at 280 nm, on a Dynamax gradient HPLC system (Rainin, Walnut Creek, CA). The column was calibrated with authentic **2**, **6**, **7**, and **8**, and under these conditions, **2**, **6**, **7**, **8**, and **10** have retention times (t_R) of 11.5, 20.3, 11.9, 9.7, and 14.0 min, respectively. TLC analysis was performed on Keiselgel 60 F_{254} glass plates (0.25 mm, Merck KGaA, Germany) with authentic standards **3**, **7**, and **8** as references, developed with $CHCl_3/MeOH/AcOH$ (85:15:0.25, v/v). Under these conditions **3**, **7**, **8**, and **10** have R_f values of 0.26, 0.33, 0.52, and 0.83, respectively.

Nucleotide Sequence Accession Number. The nucleotide sequence reported here has been deposited in the GenBank database under accession number AF263463.

Acknowledgment. We thank Prof. C. Richard Hutchinson, University of Wisconsin, Madison, for pWHM3 and pWHM63. This work was supported in part by the Committee on Research, University of California, Davis, the donors of the Petroleum Research Fund administered by the American Chemical Society, and the Searle Scholars Program/The Chicago Community Trust.

References and Notes

(1) Drautz, H.; Reuschenbach, P.; Zähner, H.; Rohr, J.; Zeeck, A. *J. Antibiot.* **1985**, *38*, 1291−1301.
(2) Egert, E.; Noltemeyer, M.; Siebers, J.; Rohr, J.; Zeeck, A. *J. Antibiot.* **1992**, *45*, 1190−1192.
(3) Anderson, M. G.; Khoo, C. L.; Rickards, R. W. *J. Antibiot.* **1989**, *42*, 640−643.
(4) Xuan, L. J.; Xu, S. H.; Zhang, H. L.; Xu, Y. M. *J. Antibiot.* **1992**, *45*, 1974−1976.
(5) De Jesus, A. E.; Hull, W. E.; Steyn, P. S.; Van Heerden, F. R.; Vleggar, R. *J. Chem. Soc., Chem. Commun.* **1982**, 902−904.
(6) Hutchinson, C. R. *Chem. Rev.* **1997**, *97*, 2525−2535.
(7) Rohr, J.; Eick, S.; Zeeck, A.; Peuschenbach, P.; Zähner, H.; Fiedler, H. P. *J. Antibiot.* **1988**, *41*, 1066−1073.
(8) Decker, H.; Rohr, J.; Motamedi, H.; Zähner, H.; Hutchinson, C. R. *Gene* **1995**, *166*, 121−126.
(9) Decker, H.; Motamedi, H.; Hutchinson, C. R. *J. Bacteriol.* **1993**, *175*, 3876−3886.
(10) Haigler, B. E.; Suen, W. C.; Apain, J. C. *J. Bacteriol.* **1996**, *178*, 6019−6024.
(11) Mao, Y.; Varoglu, M.; Sherman, D. H. *Chem. Biol.* **1999**, *6*, 251−263.
(12) Blanco, G.; Pereda, A.; Brian, P.; Mendez, C.; Chater, K. F.; Salas, J. A. *J. Bacteriol.* **1993**, *175*, 8043−8048.
(13) Niemi, J.; Mantsala, P. *J. Bacteriol.* **1995**, *177*, 2942−2945.
(14) Filippini, S.; Solinas, M. M.; Breme, U.; Schluter, M. B.; Gabellini, D.; Biamonti, G.; Colombo, A. L.; Garofano, L. *Microbiology* **1995**, *141*, 1007−1016.
(15) Shen, B.; Hutchinson, C. R. *J. Biol. Chem.* **1994**, *269*, 30726−30733.
(16) Rafanan, E. R., Jr.; Hutchinson, C. R.; Shen, B. *Org. Lett.* **2000**, *2*, 3225−3227.
(17) Beynon, J.; Rafanan, E. R., Jr.; Shen, B.; Fisher, A. J. *Acta Crystallogr.* **2000**, *D56*, 1647−1651.
(18) Summers, R. G.; Wendt-Pienkowski, E.; Motamedi, H.; Hutchinson, C. R. *J. Bacteriol.* **1992**, *174*, 1810−1820.
(19) Fernandez-Moreno, M. A.; Martinez, E.; Caballero, J. L.; Ichinose, K.; Hopwood, D. A.; Malpartida, F. *J. Biol. Chem.* **1994**, *269*, 24854−24863.
(20) Yang, K.; Han, L.; Ayer, S. W.; Vining, L. C. *Microbiology* **1996**, *142*, 123−132.
(21) Shen, B.; Hutchinson, C. R. *Biochemistry* **1993**, *32*, 6656−6663.
(22) Kendrew, S. G.; Hopwood, D. A.; Marsh, E. N. *J. Bacteriol.* **1997**, *179*, 4305−4310.
(23) Decker, H.; Haag, S. *J. Bacteriol.* **1995**, *177*, 6126−6136.
(24) Westrich, L.; Domann, S.; Faust, B.; Bedford, D.; Hopwood, D. A.; Bechthold, A. *FEMS Microbiol. Lett.* **1999**, *170*, 381−387.
(25) Han, L.; Yang, K.; Ramalingam, E.; Mosher, R. H.; Vining, L. C. *Microbiology* **1994**, *140*, 3379−3389.
(26) Davis, N. K.; Chater, K. F. *Mol. Microbiol.* **1990**, *4*, 1679−1691.
(27) Kulowski, K.; Wendt-Pienkowski, E.; Han, L.; Yang, K.; Vining, L. C.; Hutchinson, C. R. *J. Am. Chem. Soc.* **1999**, *121* 1786−1794.
(28) Shen, B.; Hutchinson, C. R. *Biochemistry* **1993**, *32*, 11149−11154.
(29) Blanco, G.; Brian, P.; Pereda, A.; Mendez, C. Salas, J. A.; Chater, K. F. *Gene* **1993**, *130*, 107−116.
(30) Bergh, S.; Uhlen, M. *Gene* **1992**, *117*, 131−136.
(31) Shen, B.; Summers, R. G.; Wendt-Pienkowski, E.; Hutchinson, C. R. *J. Am. Chem. Soc.* **1995**, *117*, 6811−6821.
(32) Kendrew, S. G.; Katayama, K.; Deutch, E.; Madduri, K.; Hutchinson, C. R. *Biochemistry* **1999**, *38*, 4794−4799.
(33) Shen, B.; Hutchinson, C. R. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 6600−6604.
(34) Sambrook, J. E.; Maniatis, T. *Molecular Cloning, a Laboratory Manual,* 2nd Ed.; Cold Spring Harbor Laboratory: Cold Spring Harbor, NY, 1989.
(35) Kieser, T.; Bibb, M. J.; Buttner, M. J.; Chater, K. F.; Hopwood, D. A. *Prectical Streptomyces Genetics*; The John Innes Foundation: Norwich, UK, 2000.
(36) Vara, J.; Skarbek, M. L.; Wang, Y. G.; Donadio, S.; Hutchinson, C. R. *J. Bacteriol.* **1989**, *171*, 5872−5881.
(37) Altschul, S. F.; Madden, T. L.; Schäffer, A. A.; Zhang, J.; Zhang, Z.; Miller, W.; Lipman, D. J. *Nucleic Acids Res.* **1997**, *25*, 3389−3402.
(38) Bibb, M. J.; White, J.; Ward, J. M.; Janssen, G. R. *Mol. Microbiol.* **1994**, *14*, 533−545.

NP010007+