Engineered Biosynthesis of a Novel Amidated Polyketide, Using the Malonamyl-Specific Initiation Module from the Oxytetracycline Polyketide Synthase

Wenjun Zhang, Brian D. Ames, Shiou-Chuan Tsai, and Yi Tang

Department of Chemical and Biomolecular Engineering, University of California, Los Angeles, California 90095, and Department of Molecular Biology and Biochemistry, University of California, Irvine, California 92697

Received 11 November 2005/Accepted 6 February 2006

Tetracyclines are aromatic polyketides biosynthesized by bacterial type II polyketide synthases (PKSs). Understanding the biochemistry of tetracycline PKSs is an important step toward the rational and combinatorial manipulation of tetracycline biosynthesis. To this end, we have sequenced the gene cluster of oxytetracycline (oxy and otc genes) PKS genes from Streptomyces rimosus. Sequence analysis revealed a total of 21 genes between the otrA and otrB resistance genes. We hypothesized that an amidotransferase, OxyD, synthesizes the malonanate starter unit that is a universal building block for tetracycline compounds. In vivo reconstitution using strain CH999 revealed that the minimal PKS and OxyD are necessary and sufficient for the biosynthesis of amidated polyketides. A novel alkaloid (WJ35, or compound 2) was synthesized as the major product when the oxy-encoded minimal PKS, the C-9 ketoreductase (OxyJ), and OxyD were coexpressed in CH999. WJ35 is an isoquinolone compound derived from an amidated decaketide backbone and cyclized with novel regioselectivity. The expression of OxyD with a heterologous minimal PKS did not afford similarly amidated polyketides, suggesting that the oxy-encoded minimal PKS possesses novel starter unit specificity.

Tetracyclines are among the most important antibiotics known to mankind in the last half century (15). The broad-spectrum antimicrobial activities of tetracyclines have resulted in their widespread clinical use against infectious gram-positive and gram-negative bacteria (21). The emergence of microbial resistance to tetracyclines has severely limited their effectiveness and has prompted the search for analogs that can overcome the known negative bacteria (21). The emergence of microbial resistance to tetracyclines has severely limited their effectiveness and has prompted the search for analogs that can overcome the known negative bacteria (21). The emergence of microbial resistance to tetracyclines has severely limited their effectiveness and has prompted the search for analogs that can overcome the known negative bacteria (21). The emergence of microbial resistance to tetracyclines has severely limited their effectiveness and has prompted the search for analogs that can overcome the known negative bacteria (21). The emergence of microbial resistance to tetracyclines has severely limited their effectiveness and has prompted the search for analogs that can overcome the known negative bacteria (21). The emergence of microbial resistance to tetracyclines has severely limited their effectiveness and has prompted the search for analogs that can overcome the known negative bacteria (21). The emergence of microbial resistance to tetracyclines has severely limited their effectiveness and has prompted the search for analogs that can overcome the known negative bacteria (21). The emergence of microbial resistance to tetracyclines has severely limited their effectiveness and has prompted the search for analogs that can overcome the known negative bacteria (21). The emergence of microbial resistance to tetracyclines has severely limited their effectiveness and has prompted the search for analogs that can overcome the known negative bacteria (21). The emergence of microbial resistance to tetracyclines has severely limited their effectiveness and has prompted the search for analogs that can overcome the known negative bacteria (21). The emergence of microbial resistance to tetracyclines has severely limited their effectiveness and has prompted the search for analogs that can overcome the known negative bacteria (21). The emergence of microbial resistance to tetracyclines has severely limited their effectiveness and has prompted the search for analogs that can overcome the known negative bacteria (21). The emergence of microbial resistance to tetracyclines has severely limited their effectiveness and has prompted the search for analogs that can overcome the known negative bacteria (21). The emergence of microbial resistance to tetracyclines has severely limited their effectiveness and has prompted the search for analogs that can overcome the known negative bacteria (21). The emergence of microbial resistance to tetracyclines has severely limited their effectiveness and has prompted the search for analogs that can overcome the known negative bacteria (21). The emergence of microbial resistance to tetracyclines has severely limited their effectiveness and has prompted the search for analogs that can overcome the known negative bacteria (21). The emergence of microbial resistance to tetracyclines has severely limited their effectiveness and has prompted the search for analogs that can overcome the known negative bacteria (21). The emergence of microbial resistance to tetracyclines has severely limited their effectiveness and has prompted the search for analogs that can overcome the known negative bacteria (21). The emergence of microbial resistance to tetracyclines has severely limited their effectiveness and has prompted the search for analogs that can overcome the known negative bacteria (21). The emergence of microbial resistance to tetracyclines has severely limited their effectiveness and has prompted the search for analogs that can overcome the known negative bacteria (21). The emergence of microbial resistance to tetracyclines has severely limited their effectiveness and has prompted the search for analogs that can overcome the known negative bacteria (21). The emergence of microbial resistance to tetracyclines has severely limited their effectiveness and has prompted the search for analogs that can overcome the known negative bacteria (21). The emergence of microbial resistance to tetracyclines has severely limited their effectiveness and has prompted the search for analogs that can overcome the known negative bacteria (21). The emergence of microbial resistance to tetracyclines has severely limited their effectiveness and has prompted the search for analogs that can overcome the known negative bacteria (21). The emergence of microbial resistance to tetracyclines has severely limited their effectiveness and has prompted the search for analogs that can overcome the known negative bacteria (21). The emergence of microbial resistance to tetracyclines has severely limited their effectiveness and has prompted the search for analogs that can overcome the known negative bacteria (21). The emergence of microbial resistance to tetracyclines has severely limited their effectiveness and has prompted the search for analogs that can overcome the known negative bacteria (21). The emergence of microbial resistance to tetracyclines has severely limited their effectiveness and has prompted the search for analogs that can overcome the known negative bacteria (21). The emergence of microbial resistance to tetracyclines has severely limited their effectiveness and has prompted the search for analogs that can overcome the known negative bacteria (21). The emergence of microbial resistance to tetracyclines has severely limited their effectiveness and has prompted the search for analogs that can overcome the known negative bacteria (21). The emergence of microbial resistance to tetracyclines has severely limited their effectiveness and has prompted the search for analogs that can overcome the known negative bacteria (21). The emergence of microbial resistance to tetracyclines has severely limited their effectiveness and has prompted the search for analogs that can overcome the known negative bacteria (21). The emergence of microbial resistance to tetracyclines has severely limited their effectiveness and has prompted the search for analogs that can overcome the known negative bacteria (21). The emergence of microbial resistance to tetracyclines has severely limited their effectiveness and has prompted the search for analogs that can overcome the known negative bacteria (21). The emergence of microbial resistance to tetracyclines has severely limited their effectiveness and has prompted the search for analogs that can overcome the known negative bacteria (21). The emergence of microbial resistance to tetracyclines has severely limited their effectiveness and has prompted the search for analogs that can overcome the known negative bacteria (21). The emergence of microbial resistance to tetracyclines has severely limited their effectiveness and has prompted the search for analogs that can overcome the known negative bacteria (21). The emergence of microbial resistance to tetracyclines has severely limited their effectiveness and has prompted the search for analogs that can overcome the known negative bacteria (21). The emergence of microbial resistance to tetracyclines has severely limited their effectiveness and has prompted the search for analogs that can overcome the known negative bacteria (21). The emergence of microbial resistance to tetracyclines has severely limited their effectiveness and has prompted the search for analogs that can overcome the known negative bacteria (21). The emergence of microbial resistance to tetracyclines has severely limited their effectiveness and has prompted the search for analogs that can overcome the known negative bacteria (21). The emergence of microbial resistance to tetracyclines has severely limited their effectiveness and has prompted the search for analogs that can overcome the known negative bacteria (21). The emergence of microbial resistance to tetracyclines has severely limiting...
Sequencing of oxy cluster. The complete genomic DNA library of *S. rimosus* was constructed using a pWEB cosmid cloning kit (Epigenics). The cosmid clone pYT264, which harbors the 21.2-kb oxy gene cluster flanked by *otc1* (18) and *otc2*, was identified by PCR screening. A combination of shotgun and primer walking techniques was used to obtain the sequence information. Open reading frames (ORFs) were detected and analyzed using Frameplot software (http://www.nih.go.jp/~jun/cgi-bin/frameplot.pl), and the putative roles of the proteins were assigned using protein-protein BLAST and Pfam analysis. To simplify naming and analysis of the cluster for subsequent studies, the genes were renamed oxy to otc in a linear fashion. The five previously sequenced genes of the cluster are cross-referenced in Table 1.

Table 1. Genes in the oxytetracycline biosynthetic cluster and deduced roles based on sequence homology

<table>
<thead>
<tr>
<th>Gene</th>
<th>Start position</th>
<th>Stop position</th>
<th>Predicted size (kDa)</th>
<th>ORFs</th>
<th>Homologs</th>
<th>Protein (reference)</th>
<th>Deduced role</th>
<th>% Identity</th>
</tr>
</thead>
<tbody>
<tr>
<td>otcA</td>
<td>134</td>
<td>625</td>
<td>18</td>
<td>CAD15553 (52)</td>
<td>Transcription regulation</td>
<td>40</td>
<td></td>
<td></td>
</tr>
<tr>
<td>otcE</td>
<td>5521</td>
<td>6777</td>
<td>46</td>
<td>AsnB (8)</td>
<td>Asparagine synthetase</td>
<td>60</td>
<td></td>
<td></td>
</tr>
<tr>
<td>otcF</td>
<td>7846</td>
<td>8609</td>
<td>37</td>
<td>BAB09170 (44)</td>
<td>Methylenetetrahydromethanopterin synthase</td>
<td>61</td>
<td></td>
<td></td>
</tr>
<tr>
<td>otcG</td>
<td>8186</td>
<td>7899</td>
<td>11</td>
<td>MtmOII (32)</td>
<td>Oxygense</td>
<td>45</td>
<td></td>
<td></td>
</tr>
<tr>
<td>otcH</td>
<td>9853</td>
<td>8241</td>
<td>59</td>
<td>MtmL (32)</td>
<td>Acyl-CoA ligase</td>
<td>47</td>
<td></td>
<td></td>
</tr>
<tr>
<td>otcI</td>
<td>10324</td>
<td>9872</td>
<td>17</td>
<td>MtmX (32)</td>
<td>Cyclase</td>
<td>49</td>
<td></td>
<td></td>
</tr>
<tr>
<td>otcJ</td>
<td>10593</td>
<td>11384</td>
<td>27</td>
<td>Sim5 (20)</td>
<td>C-9 ketoreductase</td>
<td>69</td>
<td></td>
<td></td>
</tr>
<tr>
<td>otcK</td>
<td>11443</td>
<td>12396</td>
<td>35</td>
<td>ChaF (61)</td>
<td>Aromatase</td>
<td>54</td>
<td></td>
<td></td>
</tr>
<tr>
<td>otcL</td>
<td>12386</td>
<td>14059</td>
<td>60</td>
<td>MtmXI (32)</td>
<td>Ketoreductase</td>
<td>54</td>
<td></td>
<td></td>
</tr>
<tr>
<td>otcM</td>
<td>14077</td>
<td>14835</td>
<td>25</td>
<td>MtmOIII (32)</td>
<td>Ketoreductase</td>
<td>31</td>
<td></td>
<td></td>
</tr>
<tr>
<td>otcN</td>
<td>14855</td>
<td>15628</td>
<td>28</td>
<td>MtmOIV (32)</td>
<td>Ketoreductase</td>
<td>31</td>
<td></td>
<td></td>
</tr>
<tr>
<td>otcQ</td>
<td>15734</td>
<td>16795</td>
<td>38</td>
<td>ZP00569712</td>
<td>Unknown</td>
<td>31</td>
<td></td>
<td></td>
</tr>
<tr>
<td>otcP</td>
<td>17385</td>
<td>18613</td>
<td>35</td>
<td>AknF (16)</td>
<td>Acyltransferase</td>
<td>61</td>
<td></td>
<td></td>
</tr>
<tr>
<td>otcQ</td>
<td>18911</td>
<td>19763</td>
<td>38</td>
<td>AaspB (44)</td>
<td>Aminotransferase</td>
<td>56</td>
<td></td>
<td></td>
</tr>
<tr>
<td>otcR</td>
<td>19336</td>
<td>19808</td>
<td>16</td>
<td>ActVA ORF2 (11)</td>
<td>PNP-oxidase</td>
<td>50</td>
<td></td>
<td></td>
</tr>
<tr>
<td>otcS</td>
<td>19492</td>
<td>21003</td>
<td>54</td>
<td>MtmOVI (32)</td>
<td>Oxygense</td>
<td>46</td>
<td></td>
<td></td>
</tr>
<tr>
<td>otcT</td>
<td>21064</td>
<td>22098</td>
<td>37</td>
<td>MtmXII (32)</td>
<td>N-Methyltransferase</td>
<td>45</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

a We compared the partially annotated chlorotetracycline PKS cluster (43, 51) from *Streptomyces aureofaciens* with the oxy c-encoded PKS cluster.

b Gene names in parentheses are loci of the oxy cluster first identified through mutant complementation studies. Sequence information for these genes is not available publicly, except for that shown in bold. The cluster was renamed in this study to facilitate sequence analysis. References for previously sequenced genes are listed in the text.

c No stop codon was detected. The putative protein size was based on alignment with closely related acyl-CoA ligases.

d A 60-bp deletion at the 3' end was found in the sequenced oxy cluster and has been corrected here.

FIG. 1. Proposed biosynthetic pathway of oxytetracycline (1) in *S. rimosus*. The biosynthetic pathway was established previously by Hunter and coworkers based on an exhaustive analysis of shunt products. The assignments of the tailoring enzymes are based on their functions obtained through protein-protein BLAST analysis. We confirmed in this report the role of OxyD and the minimal PKS in the biosynthesis of an amidated decaketide. We also confirmed the role of OxyJ as the C-9 ketoreductase.
GTACGGAATCCGGTACCCTC-3'. The introduced restriction sites are shown in italics, and the restriction enzymes are indicated in parentheses. The optimal ribosome binding site was introduced at the 5' end of each gene and is underlined. All oxy genes were amplified from pYT264, and multicistronic cassettes were constructed using the compatible XbaI/SpeI cohesive ends for most of the genes, except for oxyA and oxyB, which were cloned as a single PacI/XbaI cassette. Different combinations of genes were introduced into pYT315 (a pRM5-derived vector) to yield the constructs shown in Table 2.

**Culture conditions and purification of polyketides.** Strains were grown on solid R5 plates with 25 mg/liter thiotrepton at 30°C for 7 to 10 days. For analytical high-performance liquid chromatography (HPLC) analysis, a well-pigmented plate was chopped into fine pieces and extracted with 50 ml of ethyl acetate-methanol-acetic acid (89%-9.8%-1.2%). Extracts were dried over anhydrous Na2SO4. The solvent was removed in vacuo, and the residue was dissolved in 2 ml of acetone to be loaded onto two preparative thin-layer chromatography (TLC) plates (silica gel 60F-254) at 30°C for 7 to 10 days. The plates were chopped into fine pieces and extracted with 2 liters of ethyl acetate-methanol-acetic acid (89%-9.8%-1.2%). The solvent was removed in vacuo, and the residue was dissolved in 2 ml of dimethyl sulfoxide (DMSO). The polyketide products were separated by reverse-phase HPLC and detected at 254 and 280 nm using an analytical C18 column (Varian Pursuit 5u; 250 mm × 4.6 mm) with a linear gradient of 5% acetonitrile (ACN) in water (0.1% trifluoroacetic acid [TFA]) to 95% ACN in water (0.1% TFA) over 30 min with a flow rate of 1 ml/min. HPLC retention times were as follows: for Rm20b, 19.3 min; and for WJ35 (compound 2), 15.0 min. For large-scale production and isolation of compound 2, 60 R5 plates (2 liters) streaked with the transformed CH999 strains were incubated at 30°C for 7 to 10 days. The plates were chopped into fine pieces and extracted with 2 liters of ethyl acetate-methanol-acetic acid (89%-9.8%-1.2%). The solvent was removed in vacuo, and the residue was dissolved in 10 ml of H2O-ACN-DMSO (50%-25%-25%) and filtered for injection into a semipreparative reverse-phase HPLC column (Alltech Alltima 5u C18 column; 250 mm × 10 mm). A 10% to 50% acetonitrile and water (0.1% TFA) gradient was used over 45 min with a flow rate of 3 ml/min. The solvent was removed in vacuo from the collected fractions containing the expected biosynthetic product. The residue was dissolved in 2 ml of acetone to be loaded onto two preparative thin-layer chromatography (TLC) plates (20-cm × 20-cm × 0.25-mm silica gel [60F-254]). The preparative TLC plates were developed in ethyl acetate-methanol-acetic acid (94%-5%-1%), and the desired band (Rf = 0.3) was excised from the TLC plate and eluted from silica with ethyl acetate-methanol (90%-10%).

**Spectroscopic analysis.** High-resolution mass spectrometry (HRMS) was performed at the UCLA Pasarow Mass Spectrometry Laboratory with an IonSpec Ultima 7.0 Telsa electrospray ionization and matrix-assisted laser desorption ionization–Fourier transform mass spectrometry. The HRMS result for compound 2 was m/z = 388.1043 (C18H19NO7), calculated [M + H]+, 388.1027. Nuclear magnetic resonance (NMR) spectra were obtained on Bruker DRX-500 spectrometers at the NMR facility of the Department of Chemistry and Biochemistry at UCLA. The 1H and 13C chemical shifts were referenced to the solvent peak (acetone-d6) and were 6.205 and 29.9 ppm, respectively. Standard parameters were used for one-dimensional (1D) and 2D NMR experiments, which included 1H, 13C heteronuclear multiple quantum correlation (HMQC) (1H-13C), 1H, 15N heteronuclear multiple-bond correlation (HMBC) (1H-15N), and heteronuclear field-symmetry (HSQC) (1H-13C, 1H-15N) analyses. NMR experiments were performed on a DRX-600 instrument, and formamide (6 × 95; DMSO-d6) was used as an internal reference. The observed 15N NMR signal for compound 2 was 6 = 115. For detailed NMR data, see Table 3.

**Nucleotide sequence accession number.** The sequence of the oxy gene cluster was deposited in GenBank under accession number DQ143963.

**RESULTS**

**Sequencing of oxytetracycline gene cluster.** To study the origin of the amide moiety observed among all tetracyclines and to gain additional insight into the biosynthesis of tetracyclines, we sequenced the oxytetracycline (oxy and otc) gene cluster using a combination of shotgun sequencing and cosmid walking techniques. The gene cluster was previously mapped to between the two resistance genes otrA (18) and otrB (2, 10, 38). A total of 21 ORFs were clustered between otrA and otrB (Fig. 2 and Table 1), including the previously sequenced minimal PKS (otrY1-1, otrY1-2, and otrY1-3) (28), cyclase (otcD1) (46), and ATC oxygenase (otcC) (2, 45) genes. Sequencing information for the remainder of the cluster, which has been studied through mutant complementation experiments (24), was not available from GenBank prior to this study. To simplify designation of the gene cluster, we named the oxy genes oxyA to oxyT from end to end, as graphically represented in Fig. 2. The functions of the proteins were assigned based on sequence similarities to known aromatic PKS enzymes and are listed in Table 1.

A putative oxytetracycline biosynthetic pathway was previously established, largely aided by the identification of metabolic shunt products (Fig. 1) (9, 24). The minimal PKS (OxyA, OxyB, and OxyC) accepts a malonamyl starter unit and condenses eight equivalents of malonate to afford an amidated decaketide backbone. Two distinct nitrogen-inserting enzymes (OxyD and OxyQ) were found to be encoded in the oxy gene cluster. OxyD is an amidotransferase and participates in the biosynthesis of the amide starter unit (see below). OxyQ is homologous to aspartate/tyrosine/aromatic aminotransferases involved in amino acid metabolism (22) and is likely the enzyme that transaminates C-4 of 4-keto-ATC, using pyridoxal 5'-phosphate as a cofactor. OxyR, the putative pyridoxamine 5-phosphate oxidase, is translationally coupled to OxyQ with overlapping stop/start codons. A small set of proteins homologous to OxyR are found in the literature, with the most notable being ActVA-ORF2 (11), an enzyme of unknown function involved in actinorhodin biosynthesis.

Two ORFs (oxyJ and oxyM) encoding NADPH-dependent ketoreductases (KR) are present in the gene cluster. OxyJ catalyzes regiospecific C-9 reduction of the oxytetracycline backbone (see below). The specific role of OxyM in the oxy cluster is unresolved (24) and perhaps reflects redundancy in function. Only one KR gene is present in the recently sequenced chlorotetracycline gene cluster (51). The bifunctional cyclase/dehydratase OxyK (OtcD1) was previously identified by Petkovic and coworkers and was assigned to catalyze formation of the D ring (46). OxyN shows strong sequence similarity to second-ring cyclases (such as DpsY from the daunorubicin cluster [33]) and presumably catalyzes the aldol reaction between the C-5 carbonyl and the acidic C-14 meth-
TABLE 2. Plasmid constructs and resulting polyketide products

<table>
<thead>
<tr>
<th>Plasmid</th>
<th>Genes</th>
<th>Major product</th>
<th>Approx. yield (mg/liter)</th>
</tr>
</thead>
<tbody>
<tr>
<td>pYT319</td>
<td>oxyAB</td>
<td>SEK15</td>
<td>30</td>
</tr>
<tr>
<td>pYT318</td>
<td>oxyABCI</td>
<td>RM20b</td>
<td>30</td>
</tr>
<tr>
<td>pWJ35</td>
<td>oxyABCDJ</td>
<td>Compound 2</td>
<td>20</td>
</tr>
<tr>
<td>pWJ35a</td>
<td>oxyABCD actIII</td>
<td>Compound 2</td>
<td>20</td>
</tr>
<tr>
<td>pWJ40</td>
<td>act ORPs I-II oxyCDJ</td>
<td>Mutacin</td>
<td>20</td>
</tr>
<tr>
<td>pWJ48</td>
<td>tcmKL oxyCDJ</td>
<td>RM20b</td>
<td>&lt;5</td>
</tr>
</tbody>
</table>

* Streptomyces coelicolor strain CH999 was used as the host for polyketide biosynthesis. Each plasmid is derived from pRM5.

ylene to form the C ring. Subsequent formation of the B ring should be spontaneous, as seen for other aromatic polyketides such as aklanonic acid (33). Formation of the A ring is presumably catalyzed by OxyI, a small protein that is homologous to MtmX. MtmX has been implicated in catalyzing the formation of the final ring during mithramycin biosynthesis (32).

Two S-adenosylmethionine (SAM)-dependent methyltransferases (OxyF and OxyT) are encoded in the gene cluster. OxyT shows high sequence similarity to O-methyltransferases and is cotranscribed with OxyQ, OxyR, and OxyS (OtcC) (37), and it may therefore be associated with these downstream tailoring steps. We putatively assigned OxyT to dimethylation of the free amine present in 4-amino-ATC to yield ATC. OxyF is putatively assigned to methylate C-6 of pretetramid that yields 6-methyl-pretetramid. Four oxygenases are present in the oxy gene cluster, including that for the previously identified ATC oxygenase OxyS (OtcC) (2, 45). Interestingly, all four oxygenases are homologous to enzymes encoded in the mithramycin gene cluster in Streptomyces argillaceus (32). OxyE, OxyL, OxyG, and OxyS (OtcC) show strong sequence similarities to MtmOII, MtmOIII, MtmOIV, and MtmOIV, respectively, indicating that the oxy and mtm gene clusters are evolutionarily closely related. Prado and coworkers showed that MtmOII is involved in an early hydroxylation step to yield the mithramycin precursor 4-demethylpremithramycinone (47). Thus, OxyL is assigned to catalyze an analogous reaction in the biosynthesis of compound 1, using 6-methyl-pretetramid as a substrate. OxyG is a small (11-kDa) quinone-forming oxygenase and is also homologous to ElmH from the tetacenomyein gene cluster (48). OxyG is therefore possibly involved in the quinone formation of ring A in 4-keto-ATC. OxyE is a flavin adenine dinucleotide-dependent monooxygenase and is putatively assigned to catalyze the C-5 oxidation of 5a,11a-dehydroxetacycline to yield 5a,11a-dehydroxytetacycline.

The last step in the enzymatic cascade is the reduction of 5a,11a-dehydroxetacycline to compound 1. The gene encoding the reductase (TchA) responsible for the same step in chlorotetracycline biosynthesis was mapped outside the chlorotetracycline gene cluster (42). We have used degenerate primers and isolated a homologous ORF in the S. rimosus genome (result not shown).

**Enzymes putatively involved in malonamyl starter unit biosynthesis.** We previously showed that nonacetate-primed aromatic PKSs contain discrete initiation modules that typically consist of (i) a second ACP (ACP<sub>p</sub>) that primes the minimal PKS with the nonacetate unit (34), (ii) a ketosynthase III that synthesizes the ACP<sub>p</sub>-bound primer unit (39), and (iii) a potent acetyl-ACP thiolase that hydrolyzes the competing acetyl-ACP species that may otherwise initiate polyketide assembly with acetate (56). An ORF (oxyP) 15 kb downstream of the minimal PKS genes encodes an enzyme of high sequence homology to acetyl-ACP thiolases, such as ZhuC from the R1128 PKS (34) (49% identity) and FrnK from the frenolicin PKS (1) (43% identity). Based on sequence homology, OxyP may play an analogous role in the biosynthesis of compound 1 to ensure correct chain initiation by the malonamyl starter unit.

Surprisingly, no homologs of ketosynthase III and ACP<sub>p</sub> are present in the oxy cluster, suggesting that the oxy-encoded PKS employs a distinct set of enzymes to initiate polyketide biosynthesis. OxyD, encoded immediately downstream of the minimal PKS genes (oxyAB), is a 613-amino-acid protein that shows high sequence identity (60%) to the type II (also known as the Ntn family) asparaginyl synthetase (31). Asparagine synthetase converts aspartic acid to asparagine in an ATP-dependent two-step reaction, using glutamine or ammonia as the amine donor (3–5). OxyD contains the conserved N-terminal nucleophilic cysteine (Cys2) which is involved in the hydrolysis of glutamine (64) as well as the conserved adenylation domain that activates the acid moiety of the amine acceptor (6). We hypothesize that OxyD may therefore amidate either malonyl-coenzyme A (malonyl-CoA) or malonyl-ACP to yield malonamyl-CoA or malonamyl-ACP (see Fig. 4), which then primes the oxy-encoded KS-CLF for chain elongation.

**Biosynthesis of an amidated polyketide by extended minimal oxy-encoded PKS.** To examine the roles of OxyD in the biosynthesis of an amidated polyketide in vivo, a series of Streptomyces coelicolor shuttle vectors derived from pRM5 were constructed (35). The plasmids were transformed into S. coelicolor strain CH999 (35) and were analyzed for polyketide biosynthesis (Table 2).

The minimal oxy-encoded PKS (oxyABC, in pYT319) produced ample amounts of SEK15 (see Fig. 4) (53) (30 mg/liter), while the addition of OxyJ (pYT318) yielded reduced decaketide RM20b as the major product (30 mg/liter), consistent with the data reported by Fu et al. (19). OxyJ is therefore the regiospecific C-9 ketoreductase in the oxy-encoded PKS. Therefore, to reconstitute the steps of oxytetracycline biosynthesis, OxyJ was coexpressed in all subsequent studies.

Significantly, a new polyketide product was identified when OxyD was coexpressed with the minimal oxy-encoded PKS and OxyJ. The oxyD gene encoding the amidotransferase was cloned into pYT318 to yield pWJ35. CH999 transformed with pWJ35 produced an intense yellow pigmentation on solid R5 medium that was not observed before. HPLC analysis of the CH999/pWJ35 extract revealed that a new metabolite (WJ35, or compound 2) was synthesized as the major product, with an excellent yield (>20 mg/liter) (Fig. 3). The acetate-primed RM20b product was also synthesized by this strain, although in much lower quantities (<5 mg).

The mass ([M + H]+) of compound 2 was detected to be 388. LC-MS analysis of the CH999/pYT318 extract with positive ion extraction for m/z 388 showed that no trace of compound 2 was produced in the absence of OxyD, confirming the essential role of OxyD in the biosynthesis of compound 2. A sufficient amount of compound 2 was extracted from 2 liters of CH999/pWJ35 fermentation and was purified to homogeneity.
using reverse-phase preparative HPLC and normal-phase thin- 
layer chromatography. HRMS indicated a molecular formula 
of C_{19}H_{17}NO_{8} (\text{m/z} = 388.1027 \text{ [M + H]^+}; difference of 0.0016), consistent with the molecular composition of an am-
dated decaketide that has been subjected to one ketoreduction 
at C-9 and two cyclization/dehydration events. When OxyJ was 
replaced by the heterologous ActIII KR (pWJ35a; Table 2), no 
change in yield and selectivity of compound 2 was observed.

The structure of compound 2 was elucidated using a com-
bination of 1D and 2D NMR experiments (Table 3) and is 
shown in Fig. 4. Data obtained from $^{13}$C, distortionless en-
hancement by polarization transfer (90, 135), and HMQC ex-
periments identified 19 carbon signals, supporting the number 
of carbon atoms predicted by HRMS. No methyl carbon atom 
is present, consistent with the lack of an acetate starter unit. 
The presence of the $\alpha$-pyrone moiety as a result of O-1/C-5 
cyclization is clearly evident from both $^1$H (\delta_{H} = 5.3 and \delta_{H} = 6.1) and $^{13}$C (63) data. The C-6 methylene was assigned 
using $^1$H-$^{13}$C HMBC (Table 3). The linear connectivities 
among C-6 to C-9 were readily established using correlation 
spectroscopy and $^1$H-$^{13}$C HMBC. The $^1$H multiplet at \delta_{H} = 4.5 is 
assigned to H-9 and is nearly identical to the H-9 protons 
present in mutactin (66) and YT46 (57), with each containing a 
9-OH as a result of C-9 ketoreduction (Fig. 4). The assign-
ments of the C-8/C-10 methylenes were readily achieved using 
correlation spectroscopy, HMBC, and HMQC experiments.

The key assignment of the C-7 aliphatic ketone enabled us to 
eliminate the common C-7/C-12 intramolecular cyclization ob-
served among aromatic polyketides (36). The remaining car-
bon backbone cyclizes and dehydrates through a C-13/C-18 
intramolecular aldol condensation, which is confirmed by the 
coupling between H-14 and H-16 (\text{J}_{HH} = 2.1 \text{ Hz}). The nucleo-
philic amide group is thus favorably positioned to attack the 
electrophilic C-11 to yield, after dehydration, an isoquinolone 
heterocycle. The sharp proton singlet at \delta 12.9, which shows 
long-range coupling to C-16, C-17, and C-18, is assigned to 
OH-17. The broad singlet at \delta 10.2 is assigned to the NH 
proton. These assignments are consistent with the isoqui-
olonone protons observed in the natural product fredericamycin 
(7). The presence of a lactam ring was unambiguously sup-
ported by the $^1$H-$^{15}$N HMBC experiment, in which both H-10 
and H-12 showed long-range coupling to the nitrogen (Table 
3). Therefore, WJ35 is a shunt product of the oxa-encoded PKS 
in the absence of the first-ring cyclase.

\begin{table}
\centering
\caption{Proton and carbon NMR data for WJ35 (compound 2)\textsuperscript{a}}
\begin{tabular}{cccc}
\hline
Position\textsuperscript{b} & $^{13}$C \text{ (ppm)} & $^1$H \text{ (ppm)} & HMBC\textsuperscript{c} \\
\hline
1 & 165.9 & & C1, C3, C4 \\
2 & 90.9 & 5.35 (d, 1H, 2.1) & C1, C3, C4 \\
3 & 172.7 & 11.68 (s, 1H, OH) & C2, C5 \\
4 & 105.0 & 6.09 (d, 1H, 2.0) & C2, C5 \\
5 & 160.9 & & C2, C5 \\
6 & 49.2 & 3.73 (s, 2H) & C4, C5, C7 \\
7 & 204.7 & & C4, C5, C7 \\
8 & 50.7 & 2.82 (d, 2H, 5.3) & C7, C9, C10 \\
9 & 68.2 & 4.50 (m, 1H) & C8, C9, C11, C12, N \\
10 & 41.5 & 2.69 (dd, 1H, 8.0, 14.5) & C8, C9, C11, C12, N \\
11 & 140.7 & & C8, C9, C11, C12, N \\
12 & 107.8 & 6.28 (s, 1H) & C10, C11, C13, C14, C18, N \\
13 & 142.8 & & C10, C11, C13, C14, C18, N \\
14 & 102.5 & 6.40 (d, 1H, 2.1) & C12, C15, C16, C18 \\
15 & 164.9 & 11.16 (s, 1H, OH) & C12, C15, C16, C18 \\
16 & 101.7 & 6.25 (d, 1H, 2.1) & C14, C15, C17, C18 \\
17 & 165.2 & 12.91 (s, 1H, OH) & C16, C17, C18 \\
18 & 106.2 & & C16, C17, C18 \\
19 & 167.9 & & C16, C17, C18 \\
NH & 10.18 (s, 1H) & & C16, C17, C18 \\
\hline
\end{tabular}
\textsuperscript{a}Spectra were obtained at 500 MHz for protons and 125 MHz for carbon and were recorded in acetone-$d_6$ for compound 2.
\textsuperscript{b}Numbering of the carbon backbone is shown in Fig. 4.
\textsuperscript{c}Observed $^1$H-$^{13}$C and $^1$H-$^{15}$N HMBC signals.
\end{table}

OxyD does not interact with heterologous minimal PKSs.

We assayed whether heterologous minimal PKSs that are nor-
mally primed by acetate can interact productively with OxyD 
and OxyC to yield amidate polyketides. The act-encoded KS-
CLF (65) and the tcm-encoded KS-CLF (54) were each coex-
pressed in the presence of Act KR, OxyD, and OxyC in CH999. 
Heterologous KS-CLFs have been shown to function with the 
OxyC ACP to produce acetate-primed compounds (26). To 
our surprise, no amidated polyketides were recovered from 
these strains (Table 2). We were able to detect only acetate-
primed mutactin and RM20b from CH999/pWJ40 and CH999/
pWJ48, respectively. This is in sharp contrast to the broad 
compatibility between the R1128 initiation module and heter-
ologous minimal PKSs (57). Hence, fundamental differences 
exist between oxa-encoded and other minimal PKSs that allow 
the former to accept both amide and acetate starter units.

\section*{DISCUSSION}

The tetracyclines are universally primed with a unique ma-
lonamate starter unit not found in any other polyketides. Past 
work has cloned five oxa (oxt) genes and shed light on the 
minimal PKS. However, the initiation mechanism that leads to 
the production of the unique malonamate starter unit is not 
well understood. The lack of this information has hampered 
efforts to engineer the initiation of tetracycline biosynthesis.

Understanding the biosynthesis and incorporation of the novel 
amide functionality is therefore a top priority in studying tet-
racycline biosynthesis.

In this work, we sequenced the previously mapped oxa (oxt) 
gene cluster from S. rimosus, which allowed us to propose that 
OxyD is involved in the formation of the malonamate starter unit.

We reconstituted the minimal oxa-encoded PKS in CH999,
as well as an extended minimal PKS including OxyD. Significantly, the coexpression of minimal PKS with the C-9 KR and OxyD produced a unique, amidated polyketide 2 at a high yield (Table 2). Establishing the origin of the amide unit and constructing the complete backbone required for tetracycline biosynthesis are significant steps towards rational bioengineering of this family of compounds. Isolation of this novel isoquinolone 2 revealed several important biochemical properties of the oxy-encoded PKS, as follows.

(i) OxyD is the only enzyme required to biosynthesize and insert an amide starter unit into the polyketide backbone in the heterologous host. CH999/pWJ35 produced predominantly the amidated polyketide, indicating preferential incorporation of a malonamyl unit over acetyl-OxyC by the oxy-encoded minimal PKS, even in the absence of an acetyl-ACP editing enzyme. A homologous enzyme (TcsG) from the chlorotetracycline PKS (43) likely performs the same catalytic function as OxyD during chlorotetracycline biosynthesis.

(ii) The amidated compound 2 is derived from an intact decaketide (10-carbonyl) backbone, indicating that in the heterologous host CH999, the oxy-encoded minimal PKS is able to maintain correct chain length control. We did not find any truncated polyketide products in the extract of CH999/pWJ35, presumably because these truncated products are present at much lower levels in this strain and may have escaped our purification and detection protocols. We did detect a truncated polyketide in a pWJ35-derived construct coexpressing OxyK (OtcD1) (unpublished data), consistent with the conclusion by Petkovic and coworkers that the oxy-encoded PKS also synthesizes truncated polyketides (46).

(iii) The most surprising structural feature of compound 2 is perhaps the lack of C-7/C-12 cyclization. The decaketide backbone must undergo two unique cyclization steps (C-13/C-18 and N-19/C-11) to yield the isoquinolone backbone. This is apparently a dominant mode of cyclization, since no alternatively cyclized amidated polyketides were detected in the fermentation extract. The cyclization regioselectivity of compound 2 was completely unanticipated, considering that C-9-reduced, acetate-primed decaketides cyclize solely between C-7 and C-12 (as observed in RM20b/c) (49). From the crystal structures of act-encoded KR (29) and act-encoded KS-CLF (25), it has been tempting to propose that the C-7/C-12 connectivity is formed within the active site of KS, prior to C-9 reduction by the KR. It is evident from the novel structural features of compound 2 that C-9 keto-reduction must take place independently of C-7/C-12 cyclization. OxyJ must therefore recognize an uncyclized polyketide backbone to yield compound 2.

(iv) Furthermore, formation of the isoquinolone must take place immediately after the complete assembly and release of the polyketide product. Premature C-13/C-18 cyclization will prevent the correct C-7/C-12 cyclization catalyzed by OxyK during the tailoring steps of tetracycline biosynthesis (Fig. 1). Malonate is a very rare metabolite under normal physiological conditions and is a toxic compound due to its potent inhibition of succinate dehydrogenase in the tricarboxylic acid cycle (30), and hence it is unlikely to serve as a substrate for.
OxyD. Therefore, we hypothesize that biologically plausible malonyl substrates for OxyD are malonyl-CoA and malonyl-OxyC. Malonyl-OxyC may be preferred over malonyl-CoA for the following reasons. (i) Malonyl-CoA is involved in other essential cellular processes, including fatty acid biosynthesis. Amidation of a significant amount of the intracellular malonyl-CoA pool may therefore be detrimental to host strain viability (we did not observe any difference in the growth characteristics of CH999/pWJ35). Malonyl-OxyC, on the other hand, is dedicated to the oxy-encoded PKS and is not involved in the primary metabolism of the host. The robust growth of CH999/pWJ35 supports the above argument that the malonyl-CoA pool is not depleted. (ii) It is known that a cognate ACP-bound acyl substrate has a micromolar $K_m$ to the KS-CLF, while the same acyl-CoA has $K_m$ values exceeding 100 $\mu$M (56). The lower $K_m$ of acyl-ACP is a result of extensive protein-protein interactions between the acyl-ACP and the KS-CLF. Therefore, under in vivo conditions, a significantly lower concentration of malonamyl-ACP than of malonamyl-CoA may be required to prime the KS-CLF. We are currently investigating the substrate specificities of OxyD in vitro.

It is unknown why heterologous pairing of minimal PKSs with OxyD failed to yield amidated polyketides. One likely possibility is that the exclusively acetate-primed KS-CLFs, including the act- and tem-encoded PKSs, do not tolerate the presence of a polar starter unit in their active sites. Alternatively, OxyD may interact with the KS-CLF heterodimer to shuttle the malonamyl starter unit to the active site of KS and may require specific residues present on the surfaces of OxyA-OxyB. We have demonstrated that the chain length specificities of KS-CLFs can be drastically altered through rational mutagenesis (58). Therefore, the specificity of oxy-encoded KS-CLF may be similarly engineered to synthesize amidated polyketides of various lengths in coordination with OxyD.

While our in vivo results show that OxyD is a key determinant in the formation of the amide unit in W335 and, most logically, oxytetracycline, our results do not completely rule out a late role of OxyD during tetracycline assembly. In this model, the oxy-encoded PKS is primed directly by a malonyl group and proceeds with chain elongation. The acid-primed polyketide can then be amidated by OxyD to yield an amidated polyketide (W335) or can undergo spontaneous decarboxylation to yield an acetate-primed compound (RM20b, which will require the substrate specificities of OxyD in vitro.

ACKNOWLEDGMENTS

This work was supported by a UCLA faculty research grant and University of California Cancer Research Coordinating Committee funds. The UCLA Pasarow Mass Spectrometry IonSpec 7/OT Ultima FTMS-MALDI and ESI sources are supported by NSF grant CHE0092036. Our NMR instrumentation is partially supported by the NSF under equipment grants CHE9974928 and CHE0116853. We thank Chaitan Khosla, Christopher Bodd, and Iain Hunter for helpful discussions.

REFERENCES


