THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 274, No. 46, Issue of November 12, pp. 32829–32834, 1999

Printed in U.S.A.

Disruption of an Aromatase/Cyclase from the Oxytetracycline Gene Cluster of *Streptomyces rimosus* Results in Production of Novel Polyketides with Shorter Chain Lengths*

(Received for publication, September 1, 1999)

Hrvoje Petkovič‡§, Arinthip Thamchaipenet¶, Li-Hong Zhou‡, Daslav Hranueli‖, Peter Raspor§, Peter G. Waterman‡, and Iain S. Hunter‡**

From the ‡Department of Pharmaceutical Sciences, University of Strathclyde, Glasgow, Scotland G1 1XW, United Kingdom, §Biotechnical Faculty, University of Ljubljana, Ljubljana, 61000 Slovenia, ¶Department of Genetics, Kasetsart University, Bangkok, 10900 Thailand, and ‖PLIVA d.d., Research Institute, Anti-infective Research, 10000 Zagreb, Croatia

Oxytetracycline is a polyketide antibiotic made by *Streptomyces rimosus*. From DNA sequencing, the gene product of *otcD1* is deduced to function as a bifunctional cyclase/aromatase involved in ring closure of the polyketide backbone. Although *otcD1* is contiguous with the ketoreductase gene, they are located an unusually large distance from the genes encoding the "minimal polyketide synthase" of the oxytetracycline gene cluster. A recombinant, disrupted in the genomic copy of *otcD1*, made four novel polyketides, all of shorter chain length (by up to 10 carbons) than oxytetracycline. All four novel structures contained the unusual carboxamido group, typical of oxytetracycline. This implies that the carboxamido group is present at the start of biosynthesis of oxytetracycline, a topic that has been debated in the literature. Loss of the cyclase protein has a profound influence on the length of polyketide chain assembled, implying that OtcD1 plays a greater role in the overall integrity of the quaternary structure of the polyketide complex than hitherto imagined.

Oxytetracycline (OTC)[1] is a broad spectrum antibiotic made by *Streptomyces rimosus* (1). It is a member of the "polyketide" class of secondary metabolites biosynthesized by condensation of coenzyme A derivatives of metabolic precursors (2). The backbone of the antibiotic, consisting of 19 carbon atoms, is thought to be derived from an aminated starter unit (most likely malonamyl-CoA), to which eight acetyl (malonyl-CoA) extender units are added sequentially (Fig. 1). Molecular genetic analysis of OTC biosynthesis revealed evidence that the pathway genes are clustered together on the chromosome of *S. rimosus* (3, 4), and that the tetracyclic backbone is assembled by a (so-called) "Type II" polyketide synthase (PKS) (5), in which this multienzyme complex carries out the iterative assembly of the backbone from its precursors and subsequently folds and cyclizes the nascent polyketide. Type II PKSs that catalyze the biosynthesis of different chemical structures have a common genetic architecture: their genes share substantial homology (6).

The minimal Type II PKS has been defined (7) as the three genes which, together, encode the minimal number of gene products that can direct biosynthesis of a polyketide chain. The minimal PKS consists of two translationally coupled "ketosynthase" genes, KS_α and KS_β, and an acyl carrier protein to which the growing polyketide chain is attached. KS_α contains the essential amino acid residues involved in catalytic condensation of the co-acylated derivative of the extender units, whereas the KS_β partner (which is highly similar but non-identical to KS_α) is thought to influence the number of iterations of the condensation cycle and hence the length of the polyketide chain, justifying its original designation as the chain length factor (8). Matrix experiments, in which the KS and acyl carrier protein elements have been exchanged among different PKSs, have reinforced this view. When expressed alone in *Streptomyces coelicolor* using the elegant trancriptionally activated pRM5/CH999 system, the two KS subunits and acyl carrier protein from the *otc* gene cluster directed the biosynthesis of an incorrectly cyclized 20-carbon backbone, without the terminal carboxamido group that is typical of oxytetracycline (9). This implied that the minimal PKS for OTC biosynthesis could not direct synthesis of the natural 19-carbon backbone on its own.

After the nascent polyketide chain has been formed, it must be folded and cyclized to take up its final structure. Two different types of gene products have been associated with this process: "ketoreductases" (KR), which are responsible for reduction of one of the keto groups within the chain; and "cyclases" (CYC), which are responsible for carbon-carbon bond formation resulting in the closed rings of the final products (6). A KR step is not obligatory for formation of some polyketides, *e.g.* the unreduced polyketide, tetracemonycin, folds and is cyclized without a KR (10). However, cyclases are ubiquitous in polyketide gene clusters. They fall into two classes: (*a*) those with "didomain" architecture (typified by frenolicin; see Ref. 11) in which the N-terminal segment of the gene has likely been duplicated to form the C-terminal second domain, and (*b*) those with "monodomain" architecture (typified by *whi*E-ORFVI; see Ref. 12) containing a single cyclase domain. In some instances a single cyclase domain may be linked to a second domain of unrelated function *e.g. tcmN*, which has an

* This work was supported by ALIS LINK Grant CRO007 (to I. S. H. and D. H.) from the British Council and the Ministry of Science and Technology, Croatia and by Grant 058407 from the Ministry of Science and Technology, Croatia (to D. H.). We thank the Governments of Slovenia (Ministry of Science and Technology Award P38-0259-0490-007/13542/94 (to P. R.)) and Thailand for the awards of a studentship and fellowship (to H. P. and A.T., respectively). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

The nucleotide sequence(s) reported in this paper has been submitted to the GenBank™/EBI Data Bank with accession number(s) AF087137.

** To whom correspondence should be addressed: Dept. of Pharmaceutical Sciences, University of Strathclyde, Royal College Building, Glasgow, Scotland G1 1XW, UK. Tel.: 44-141-548-4111; Fax: 44-141-548-4124; E-mail: i.s.hunter@strath.ac.uk.

[1] The abbreviations used are: OTC, oxytetracycline; HMBC, heteronuclear multiple bond connectivity; PKS, polyketide synthase; KR, ketoreductase; CYC, cyclase; ARO, aromatase; kb, kilobase(s).

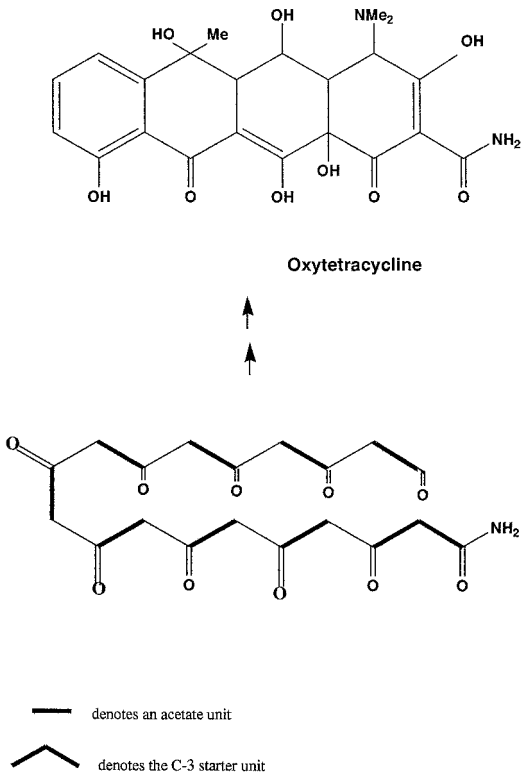

FIG. 1. **The structure of oxytetracycline and derivation of its backbone.**

O-methyltransferase activity at its C terminus (13). Some polyketide structures may require more than one cyclase to fold them correctly (14).

Similarity among DNA sequences of different Type II PKSs has been used to identify (*e.g.* see Ref. 15) and clone homologous sequences from other streptomycete species. Biosynthesis of actinorhodin has become the paradigm for study of Type II PKSs. The genes responsible for the assembly of the backbone of actinorhodin are contiguous on the chromosome of *S. coelicolor*, an architecture that is present in most other homologous gene clusters. When *actI* (the minimal PKS) and *actIII* (KR) of the actinorhodin cluster were used as probes with the cloned *otc* cluster, two non-contiguous regions of homology were identified (16): *otcY1* hybridized with *actI* (and DNA sequencing has shown it to constitute the *otc* minimal PKS (17)), whereas *otcY2* lay 10 kb away and hybridized to *actIII*. This paper is concerned with the analysis of the DNA adjacent to the *otcY2* locus. It describes the characterization of a bifunctional cyclase gene, *otcD1*, its subsequent manipulation to produce a recombinant blocked in that function, and characterization of the polyketide metabolites made by the recombinant.

EXPERIMENTAL PROCEDURES

Strain Maintenance and Cultivation—*S. rimosus* R6 is used for the commercial production of OTC. It was maintained and grown in fermentation medium as described previously (18).

General DNA Techniques—The methods used were described by Sambrook *et al.* (19) for *Escherichia coli* and by Hopwood *et al.* (20) for *Streptomyces*. DNA was introduced into *S. rimosus* by an electrotransformation technique (21). DNA was sequenced by the Sanger method (22). Templates were prepared by subcloning the DNA into vectors that contained universal and reverse primer binding sites. If necessary, the region of sequence determination was extended with custom primers.

Isolation of Metabolites—1.5 liters of fermentation medium was saturated with NaH_2PO_4 and extracted with ethyl acetate (3 × 1 liter). After evaporation to dryness, the extract was applied to a silica gel flash chromatography column and eluted with chloroform containing increasing amounts of methanol. Fractions containing the desired compounds were purified further on a Sephadex LH-20 column and eluted in a chloroform/methanol gradient.

Structural Elucidation of Metabolites—This was undertaken primarily by NMR (400 MHz) using, in addition to standard one-dimensional experiments, the two-dimensional homonuclear correlated spectroscopy-45 technique for direct ^{1}H-^{1}H coupling (23), the hydrogen-carbon correlation with a bilinear rotation decoupling pulse technique for 1-bond (^1J) heteronuclear (^{1}H-^{13}C interactions) (24), and the HMBC[4] pulse sequence for 2-bond (^2J) and 3-bond (^3J) interactions (23).

RESULTS

DNA Sequence of otcD1 and Analysis of the Deduced Gene Product—When the entire *otc* cluster was digested with various restriction enzymes and subjected to Southern blots using *actIII* (KR; Ref. 25) as a probe, a *Sac*I fragment of approximately 700 base pairs gave the strongest signal (HB in Fig. 2; Sac_{29}–Sac_{31} in the restriction map of the *otc* cluster described in Ref. 26), indicating that it likely contained the major part of the gene encoding the functional equivalent the KR for OTC biosynthesis. Analysis of this DNA sequence and the segment adjacent (Sac_{31} to Sph_{28}) revealed two open reading frames: the incomplete C terminus of the ActIII homologue, OtcY2-ORF2, and the complete open reading frame of *otcD1* (Fig. 2). The DNA sequence of *otcD1* and its flanks has been deposited at GenBank™.

The deduced polypeptide, OtcD1, was predicted to contain 317 amino acids with M_r of 34635. Using the TBLASTN software (27), OtcD1 showed substantial end-to-end similarity with other deduced cyclases from polyketide biosynthetic clusters. The N- and C-terminal halves of OtcD1 have reasonable similarity to each other (Fig. 3) indicating a didomain architecture, and aromatase/cyclase (ARO/CYC) function.

Construction of a S. rimosus Recombinant Disrupted in otcD1—The restriction site Sac_{29} lies within the open reading frame of *otcD1* (Fig. 2). *ErmE*, encoding erythromycin resistance (28), was subcloned into the Sac_{29} site of an *E. coli* plasmid-based copy of the gene and this disrupted copy of *otcD1* (*otcD1*::*ermE*) was homogenotized into the chromosome of *S. rimosus*, using an unstable bifunctional construct, pGLW121, based on the thiostrepton-resistant streptomycete vector pIJ487 (29). Transformation of pGLW121 into *S. rimosus* R6 was selected by resistance to thiostrepton and erythromycin. After two subcultures under non-selective conditions, transformants were screened to identify isolates (erythromycin-resistant, thiostrepton-sensitive) that had the disrupted copy of *otcD1* introduced into the chromosome by a double crossover and lost the vector sequences from the cell. The genotypes of such isolates were confirmed by Southern blotting (30), and strain ZGL3 was chosen for further work.

Characterization of Metabolites Made by the otcD1::ermE Recombinant, ZGL3—Three compounds were separated from the Sephadex column with 6% (v/v) methanol in chloroform, LH-1 (15.2 mg), followed by LH-2 (8.5 mg), and finally LH-3 (6.7 mg). A further compound, LH-4 (7.3 mg), was obtained by direct flash chromatography of the fermentation broth by elution with 8% (v/v) methanol in chloroform.

LH-1 failed to yield a molecular ion by electron impact mass spectroscopy but the ^{13}C NMR spectrum clearly resolved the presence of 9 carbons: two sp^3 (both methylenes) and seven sp^2 (two methines and five quaternary). The ^{1}H NMR spectrum confirmed the presence of two methylenes and two olefinic methines and in addition revealed broad resonances for OH and NH_2. The NMR chemical shift data are listed in Table I. The IR spectrum revealed a number of absorption bands between 1,716 and 1,639 cm^{-1}, which suggested a polyketide.

The structure was resolved largely on the basis of two-dimensional homonuclear and heteronuclear NMR. The two ole-

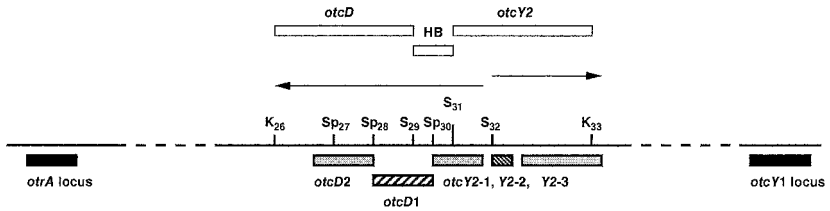

FIG. 2. **The genetic architecture of the *otc* cluster around the locus *otcY2*.** *Open boxes* denote the DNA fragments identified previously in complementation studies (*otcD*, *otcY2*; Ref. 3) or hybridization experiments (see text, HB; Ref. 16). *Arrows* denote the direction of transcription of the genes, which are shown in *shaded boxes*, with their designations below. The restriction sites around *otcY2* that are relevant to this work are shown (*K*, *Kpn*I; *Sp*, *Sph*I; *S*, *Sac*I). Numbering of restriction sites is from Hunter and Hill (26). The figure is not drawn to scale. The *broken lines* illustrate >8 kb gaps between the *otcY2* locus and the *otcY1* locus containing the genes for the minimal PKS (see text) on one flank and the *otrA* resistance gene on the other flank.

```
                                                          50
otcN    mpAptshr avhrteidAP Adrvyalird AaeWPrhftp tvhveraEld
otcC  agsdAellfs fedsetvhAP Aeavyaflae AgkWPdrlph vsrldltEps

        51                                                100
otcN  arserlriwa tANgevkHwH -ShRaldPeg qsvrfrQevc sPpvaamsGE
otcC  dgvqvmtmvt rANdgseHtH eSvRvcfPde lrivykQigt pPlmtlhtGE

      101                                               150
otcN  WvlRDlpgGr celtlHhafa avddrPEdve witTAtDrns RtelAnikal
otcC  WsiRDtgdGl lvtsqHtiri nesaiPEilg adaTAaDara RvraAvggns

        151
otcN  Aea
otcC  Aatlalanaf aeapha
                  166
```

FIG. 3. **Comparison of the N-terminal and C-terminal peptide sequences of OtcD1.** The PILEUP program (43) was used. Identical residues are shown in *capital letters*.

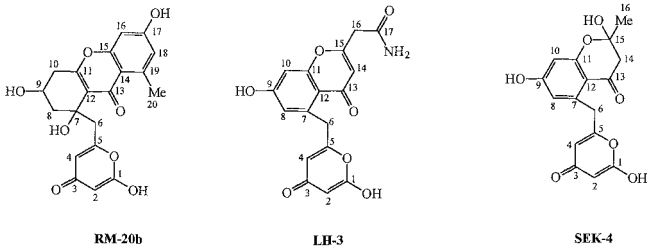

FIG. 4. **Comparison of the structures of the compounds Rm20b (9), LH-3, and SEK4 (33).**

TABLE I

NMR chemical shift data for LH-1 and comparison with published data for SEK4 (33)

All spectra were recorded in Me_2SO-d_6. Data for a SEK4 was modified to take into account the change in structure from a 4-oxo-2-hydroxypyran to a 2-oxo-4-hydroxypyran (31).

H/C	δ_H		δ_C	
	LH-1	SEK4	LH-1	SEK4
1			163.7	165.4
2	5.26 d (2.1)	5.19 d (2.2)	89.0	88.2
3	11.80 s	11.60 s	170.2	170.5
4	6.05 d (2.1)	5.66 d (2.2)	102.8	102.9
5			159.4	163.8
6	3.83 s		46.8	
7			199.6	
8	3.38 s		50.1	
9			167.5	
NH_2	7.52 s, 7.11 s			

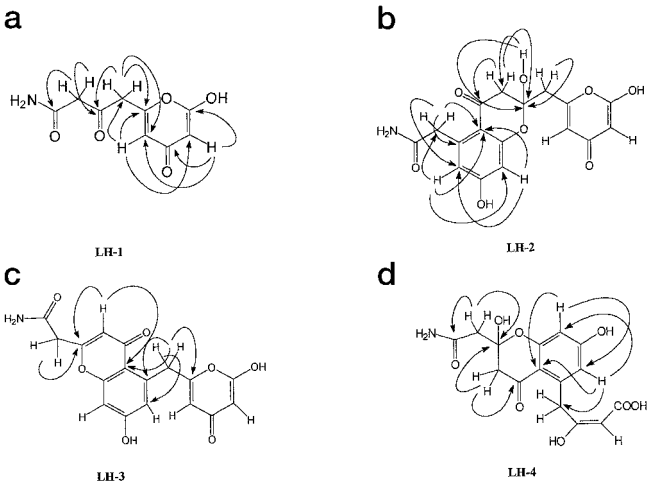

FIG. 5. **The long-range ^1H-^{13}C coupling observed in the NMR spectra of LH-1, LH-2, LH-3, and LH-4, which were useful in characterizing the structures.**

finic protons exhibited *meta*-coupling to each other with one (δ 5.26; δ_c 89.0) in an unusually electrophilic environment reminiscent of H-3 in a 2,4-oxygenated pyran as found in SEK4 (Fig. 4, reported in Ref. 33) (see Table I). This was confirmed by the HMBC experiment, the correlations observed being shown in Fig. 5 and which allowed LH-1 (Fig. 4) to be identified as 6-(4-amino-2,4-butadionyl)-4-hydroxypyran-2-one (Fig. 6).[2] The original report of SEK4 (33) proposed a 2-hydroxypyran-4-one system, but this was later revised to 4-hydroxypyran-2-one (31) on the basis of NMR studies. We support this revision as the C-4 resonance at δ_c 170.2 is at variance with that found in the pyran-4-one system of flavones and would be anticipated at ~δ_c 176.0 (32).

The second compound to be eluted again failed to show a molecular ion by electron impact mass spectroscopy but the ^{13}C NMR spectrum (Table II) revealed 17 atoms as three methylenes, four olefinic methines, and 10 quaternary carbons. The ^1H NMR revealed resonances for amide protons (δ 6.73, 7.09) and hydroxyl protons (δ 7.13, 10.40, 11.50). The presence of the 6-substituted 4-hydroxypyran-2-one moeity was identified by comparison with LH-1. The 6-substituent was established as a methylene (δ 3.07, 2.99; δc 44.0) by observation in an HMBC experiment (Fig. 5*b*) of the long range coupling of the methylene protons to C-6 and C-5 of the pyran ring. Likewise, further resonances were compatible with the presence of the methylenecarboxamide starter group of this compound.

The intervening structure was established by further analysis of the HMBC experiment (Fig. 5*b*). Thus, the methylene of the methylenecarboxamide was coupled with two *meta*-coupled protons of an aromatic ring, one of which (δ 6.21, δ_c 103.2) must be placed between the two oxygens. From the 6-methylene substituent of the pyran-4-one substituent there was a coupling with a carbon resonating at 100.2 ppm. This quaternary carbon was also correlated with the protons of another methylene and with the proton of the hydroxyl resonating at δ 7.13, and it was concluded that it must be an sp^3 carbon carrying two oxygen substituents. This additional methylene revealed further couplings to a carbonyl at δ_c 190.3, which is indicative of a flavanone-like dihydropyran system (31). On the basis of these

[2] The numbering system used in the figures and tables in this paper is non-IUPAC but follows that of Fu *et al.* (33) in reflecting the biosynthesis of cyclic compounds from the non-cyclic precursor.

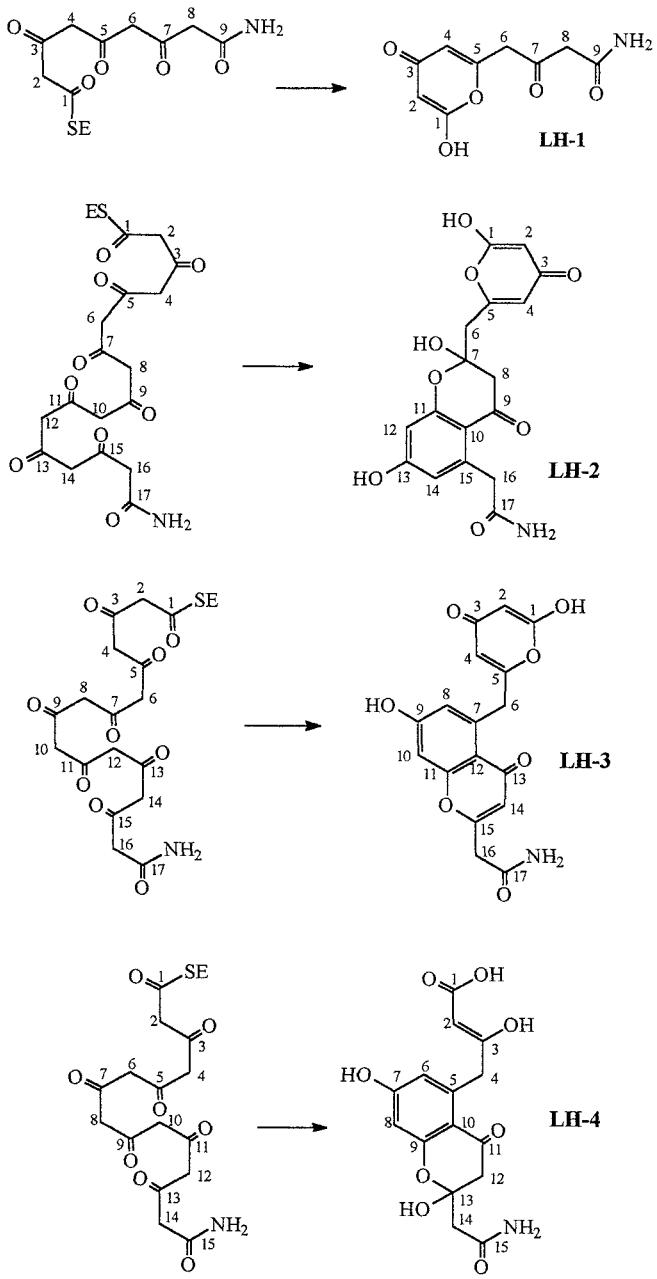

FIG. 6. **Structures of the isolated polyketides and their relationship to putative acyclic precursors.**

observations, it was concluded that the bicyclic was a 2,7-dihydroxy-2,3 dihydropyran-4-one substituted at C-5 by the methylenecarboxamide and at C-2 with the 6-methylene-2-hydroxypyran-4-one to give structure LH-2 (Figs. 4 and 6), 2-[2,7-dihydroxy-2-(4-hydroxy-6-oxo-4H-pyran-2-methyl-yl)-4-oxochroman-5-yl]-acetamide.

LH-3 exhibited many features in common with LH-2 and the presence of the 6-methylene-4-hydroxypyran-2-one and methylenecarboxamide groups could clearly be observed from the NMR spectra (Table II). The ^{13}C spectrum again revealed 17 carbon atoms but in this case there were three sp^2 methines and no methylenes associated with the intervening substructure. A carbonyl resonance at δ_c 177.7 was indicative of a benzopyran-4-one (31) and there were no signals associated with a benzodihydropyran-4-one, which suggested conversion of the latter to the former through dehydration to give 7-hydroxybenzopyran-4-one. Somewhat surprisingly, however, an HMBC experiment (Fig. 5*c*) showed that this compound was not the dehydration product of LH-2. The methylene protons of the methylenecarboxamide group correlated with the oxygen-bearing C-2 carbon of the benzopyran, whereas the methylene linked to the 4-hydroxypyran-2-one correlated with a methine (C-6) and two quaternary carbons (C-4a, C-5) in the benzenoid ring of the benzopyran. The structure of LH-3 (Fig. 6) must, therefore, be 2-[7-hydroxy-5-(4-hydroxy-6-oxo-4H-pyran-2-methyl-yl)-4-oxo-4H-chromen-2-yl]-acetamide.

The ^{13}C NMR spectrum of LH-4 revealed only 15 carbons, 9 of which gave chemical shift values comparable to the 2-hydroxy-2,3-dihydrobenzopyran ring system that had already been identified in LH-2, whereas two others were in agreement with expected values for a methylenecarboxamide substituent. However, unlike LH-2, the long range correlations observed in the HMBC spectrum of this compound (Fig. 5*d*) revealed that the methylenecarboxamide occurred at the 2-position of the 2,3-dihydropyran-4-one. This left a 4-carbon unit to be placed at C-5 of the benzopyran. This moeity consisted of a methylene, a highly shielded sp^2 methine (δ 4.65; δ_c 86.8) and two quaternary sp^2 carbons. The absence of direct H-H coupling between methylene and methine protons required that they were not adjacent, whereas long range H-C coupling between H-6 of the benzopyran and the methylene carbon meant the latter was linked to C-5 of the pyran. NMR spectra failed to provide conclusive data to support the structure of this side chain but given the polyketide origin of LH-4 a 3-hydroxybut-2-enoic acid substituent is plausible and is not contradicted by the chemical shift values (Table II). The tentative structure proposed for LH-4, as depicted in Fig. 6, is 4-(2-carbamoyl-2,7-dihydroxy-4-oxo-chroman-5-yl)-3-hydroxy-but-2-enoic acid.

DISCUSSION

otcD1 Encodes a Didomain Cyclase/Aromatase—The initial observation made by Southern blotting, that the minimal PKS and KR for OTC biosynthesis were non-contiguous (16), has been confirmed by DNA sequencing and extended now to show that the cyclase gene, *otcD1*, lies just downstream of the KR (Fig. 2). For most polyketide clusters (*e.g.* actinorhodin; see Ref. 34), the genes responsible for the assembly, folding, and cyclization of the backbone are contiguous, although there are some exceptions (*e.g.* for frenolicin (11) and nogalacin (35)). However, it is notable that the KR and cyclase genes for OTC biosynthesis are located a considerable distance (10 kb) from the minimal PKS (Fig. 2,) with genes encoding a variety of disparate functions (*e.g.* methylase, CoA ligase, and aminotransferase; see Ref. 26) between the two locations. Thus in the *otc* cluster, the KR and cyclase genes are expressed on a different mRNA from the PKS genes, which implies that expression of both mRNAs will need to be tightly co-ordinated.

The gene product, OtcD1, shows end-to-end homology with putative gene products from other aromatic polyketide gene clusters when comparisons are made using the TBLASTN software (27). Most notable (56% identity, 75% similarity) is the similarity with the putative aromatase/cyclase of the chlortetracycline biosynthetic cluster (36). There is also good homology with gene products from other polyketide clusters, *e.g.* SnoE of the nogalamycin cluster (53%, 64%) (35), Gris-Orf4 from the griseucin producer (49%, 61%) (37), DpsF from the daunorubicin producer (48%, 60%) (38), and ActVII (45%, 59%) (34) from *S. coelicolor*, the producer of actinorhodin.

The N- and C-terminal segments of the deduced OtcD1 polypeptide share homology with each other (26% identity, 35% similarity (Fig. 3)). It is therefore reasonable to deduce that *otcD1* encodes a didomain ARO/CYC, which is able to catalyze formation of the carbon-carbon bond that closes a ring and subsequently dehydrates the ring to aromatize it. Set against the existing literature (for review, see Ref. 14), the gene prod-

TABLE II

^{1}H NMR and ^{13}C NMR chemical shift data for compounds LH-2, LH-3, and LH-4

All spectra were run in $Me_2SO \cdot d_6$.

H/C	δ_H			δ_C		
	LH-2	LH-3	LH-4	LH-2	LH-3	LH-4
1				162.7	163.9	162.5
2	5.27 d (2.4)	5.16 d (2.1)	4.65 s	89.1	88.2	86.8
3(OH)	11.50 s	11.40 s	5.35 s	170.1	170.6	163.0[a]
4	6.11 d (2.4)	5.55 d (2.1)	4.04 d (15.9)	103.2	99.6	37.4
5				160.6[a]	165.7	139.7
6	3.07 d (14.4)	4.36 s	6.29 d (2.1)	44.0	37.4	113.0
	2.99 d (14.4)					
7(OH)	7.13 s		7.49 s	100.2	137.8	162.8[a]
8	2.97 d (16.0)	6.74 d (2.3)	6.18 d (2.1)	47.4	117.8	102.7
	2.55 d (16.0)					
9(OH)		10.86 s		190.3	161.4[a]	161.3
10		6.76 d (2.3)		111.8	102.2	111.7
11				160.8[a]	159.3[a]	190.6
12	6.21 d (2.4)		3.15 d (16.0)	102.2	113.8	48.0
13(OH)	10.40 s			163.7	177.7	100.6
14	6.31 d (2.4)	6.07 s	2.66 d (14.5)	113.4	112.3	43.8
15				140.0	162.1	171.3
16	3.82 d (15.2)	3.47 s		41.0	40.3	
	3.70 d (15.2)					
17				171.8	168.4	
NH2	6.73 s	7.17 s	7.20 s			
	7.09 s	7.63 s	7.61 s			

[a] Values in the same column are interchangeable.

uct would be responsible for the closing and aromatizing of the "left" ring (Fig. 1) of OTC.

ZGL3, a Strain Disrupted in otcD1, Makes Novel Polyketides of Shorter Chain Length—The four compounds isolated from the *S. rimosus* ZGL3 strain are all classical polyketides containing 9, 15, or 17 carbon atoms. They share a common biosynthetic origin, in which three (LH-1), six (LH-4), or seven (LH-2 and LH-3) acetate units have been added to a methylenecarboxamide precursor (Fig. 6). Thus, they have been made from a common starter unit through different iterative cycles of the PKS complex.

Disruption of *otcD1* has thus resulted in loss of specificity of chain length: normally the 19-carbon backbone of OTC is made. The "genetic surgery" performed on the *otc* cluster to generate the ZGL3 recombinant was some 10 kb distant from the minimal PKS (Fig. 2), which is implicated in specificity of chain length. As *otcD1* is transcribed on a different mRNA, expression of the minimal PKS should not have been altered through any polar effect.

Our current understanding of chain length elongation in polyketide biosynthesis (*e.g.* see Ref. 39) stems from studies with the pRM5/CH999 system in *S. coelicolor* (7). The host strain (CH999) is deleted for the entire actinorhodin cluster and the plasmid (pRM5) includes transcriptional activation machinery to ensure high expression of polyketide genes that can be added back to the null strain in different combinations and contexts. By design, this system investigates the molecular interactions among a small number of different gene products. Using hybrid PKS gene sets, the pRM5/CH999 system has shown (7, 8) that the KS_β subunit is a major determinant of chain length. However, specificity is not as absolute as, for example, the frenolicin KS duo direct biosynthesis of a mixture of octa- and nonaketides.

The data reported here adopts a different strategy to investigate gene function, to ablate selectively a gene within a chromosomally located cluster and investigate the result. It is presumed that all of the other gene products within the cluster will be expressed normally with the possible exception of those whose genes lie downstream on the same transcript and may be subject to polarity. In the ZGL3 recombinant, disruption of *otcD1* has completely "derailed" the capability of the PKS to synthesize polyketides of the usual chain length. This is most likely due to a change in quaternary interactions among the subunits of the PKS complex in the absence of the cyclase/aromatase polypeptide to act as a structural partner. It is not unprecedented that interactions between a cyclase/aromatase and minimal PKS can influence the fidelity of determination of chain length: (*a*) when the frenolicin minimal PKS was coexpressed with TcmN, the tetracenomycin monodomain cyclase, a nonaketide was produced whereas the minimal PKS alone produced an octaketide (39); (*b*) when the C-terminal domain of the ARO/CYC of the griseolin polyketide cluster was co-expressed with the minimal PKS from the tetracenomycin (*tcm*) cluster, a polyketide of 18 carbons was made, in addition to the usual 20-carbon structure that the *tcm* gene products specify (40); (*c*) when the *whiE* minimal PKS was expressed alone it made an 18-carbon product, but a 20-carbon polyketide was made when the *whiE* CYC/ARO was co-expressed (37). However, the remarkable feature of the products made by the ZGL3 recombinant is the series of polyketide molecules with a wide spectrum of chain lengths, which are shorter, in the case of LH-1 much shorter, than the wild-type strain. These products form a "quasi-homologous" series of chain length. Expression of the minimal PKS from the *otc* cluster in the pRM5/CH999 system resulted in a polyketide (RM20b, Fig. 4) with a backbone of 20 carbon atoms (9). The structural similarity between RM20b and LH-3 and SEK4, whose synthesis is directed by the minimal PKS from the actinorhodin pathway, is immediately apparent. The *act* system directs synthesis of a 16-carbon structure. On its own, the *otc* minimal PKS directs synthesis of a 20-carbon polyketide, whereas in wild-type *S. rimosus* the 19-carbon backbone of OTC is made. In the absence of *otcD1,* the backbone is reduced to 17 carbons or less. Thus, OtcD1 (as a representative cyclase/aromatase) may have a greater role to play in the determination of chain length than hitherto imagined.

The Novel Polyketides Have Aminated Termini, Implying That They Are All Derived from the Same Aminated Precursor—Each of the four new structures made by strain ZGL3 has an aminated

group, consistent with malonamyl-CoA being used as the starter unit for biosynthesis of these polyketides. In the literature (*e.g.* see Ref. 26), it has been hypothesized that malonamyl-CoA is the starter of OTC but, equally, malonyl-CoA could also act as starter with the amination taking place subsequent to completion of the backbone. The elegant work of Thomas and Williams (41, 42) used stable isotopic labeling to elucidate the nature of extender units and folding pathway of the backbone of OTC but did not address the issue of the starter unit. The truncated compounds made by ZGL3 are different from OTC in three-dimensional structure and also different from each other. It is unlikely that they would all be substrates for a single enzyme that performs an amination after the backbone is formed. This implies that the carboxamido group is present at the start of biosynthetic process and that malonamyl-CoA is very likely the natural starter unit for the OTC PKS.

The presence of the carboxamido group in each of the four structures helps also to resolve the issue of whether truncated polyketide products made by recombinant strains are the result of premature termination of the iterative process of chain extension or degradation products. All polyketide structures reported previously have been derived solely from acetate units, so it was impossible to come to a decision on this issue. Indeed, previous reports (*e.g.* see Ref. 31) have included the caveat that the structures could, formally, be the result of degradation of "full-length" chains. The carboxamido group in each of the four novel structures reported here provides a unique "tag" to indicate the point from which chain inititation is made, and affords the conclusion that the structures are the result of different numbers of iterations of chain extension.

Novel Polyketides Are the Result of Different Folding Patterns—Although LH-2 and LH-3 are derived from the same progenitor compound, their final structures are the result of different folding/cyclization patterns. LH-3 is the result of a cyclization pattern (Fig. 6) for the first ring (C12/C7) that is typical for cyclization of most aromatic polyketides, including actinorhodin (7). LH-2 has the unusual C10/C15 cyclization pattern (Fig. 6). Formation of the LH-2 and LH-3 pair thus parallels the situation with SEK4 and SEK4b, which are similarly derived from a single progenitor through different folding routes. The folding pattern of LH-3 is replicated in the 15 carbon compound, LH-4. It is likely that each of the 4 compounds has cyclized spontaneously in the absence of the *otcD1* gene product.

The Role of KR in the otcD1::ermE Strain, ZGL3—Each of the novel polyketides, LH-1–LH-4, has a structure that implies that no reduction of a keto group by the KR of the *otc* cluster has taken place in the ZGL3 recombinant. This is a curious result because the gene encoding the KR is upstream of *otcD1.* The genetic construction to make the recombinant should have left the KR intact. Three possibilities could plausibly explain this outcome: (i) the KR protein cannot integrate into the PKS assembly when the *otcD1* gene product is missing (as is the case in the ZGL3 recombinant); (ii) although the gene encoding the KR is proximal to *otcD1* on the polycistronic mRNA, its level of expression may be reduced by instability of the mutant mRNA containing the disrupted ARO/CYC (*otcD1::ermE*); or (iii) the nascent polyketide backbones of the progenitor structures of LH-1–LH-4 are not substrates for the *otcKR*. Interestingly, when the *otc* minimal PKS was co-expressed in *S. coelicolor* using the pRM5/CH999 system along with the KR (*actIII*) from the actinorhodin pathway, then the resulting polyketide had no reduced keto group (9). In this case, ActIII is either incapable of forming a catalytic partnership with the *otc* minimal PKS or the 20-carbon product formed by this minimal PKS is not a substrate for ActIII.

Addendum—After submission of this manuscript, a paper was published (Shen *et al.* (1999) *Proc. Natl. Acad. Sci. U. S. A.* **96,** 3622–3627), which shows that ectopic expression of the minimal PKS for WhiE, the polyketide-derived spore pigment of *S. coelicolor* results in a large number of compounds with different chain length. The authors deduce that the minimal PKS enzyme complex must rely on the stabilizing effects of additional subunits (*i.e.* the cyclase) to ensure that the chain reaches its full length.

REFERENCES

1. Chopra, I., Hawkey, P. M., and Hinton, M. (1992) *J. Antimicrobial. Chemother.* **29,** 245–277
2. Gatenbeck, S. (1962) *Biochim. Biophys. Res. Commun.* **6,** 4–6
3. Butler, M. J., Friend, E. J., Hunter, I. S., Kaczmarek, F. S., Sugden, D. A., and Warren, M. (1989) *Mol. Gen. Genet.* **215,** 231–238
4. Binnie, C., Warren, M., and Butler, M. J. (1989) *J. Bacteriol.* **171,** 887–895
5. Hopwood, D. A., and Sherman, D. H. (1990) *Annu. Rev. Genet.* **24,** 37–66
6. Hopwood, D. A., and Khosla, C. (1992) in *Secondary Metabolites: Their Function and Evolution* (Chadwick, D. J., and Whelan, J., eds) Vol. 171, pp. 88–112, Ciba Foundation Symposium, John Wiley and Sons, Chichester, UK
7. McDaniel, R., Ebert-Khosla, S., Fu, H., Hopwood, D. A., and Khosla, C. (1994) *Proc. Natl. Acad. Sci. U. S. A.* **91,** 11542–11546
8. McDaniel, R., Ebert-Khosla, S., Hopwood, D. A., and Khosla, C. (1993) *J. Am. Chem. Soc.* **115,** 11671–11675
9. Fu, H., Ebert-Khosla, S., Hopwood, D. A., and Khosla, C. (1994) *J. Am. Chem. Soc.* **116,** 6443–6444
10. Hutchinson, C. R., and Fujii, I. (1995) *Annu. Rev. Microbiol.* **49,** 201–238
11. Bibb, M. J., Sherman, D. H., Omura, S., and Hopwood, D. A. (1994) *Gene* **142,** 31–39
12. Davis, N. K., and Chater, K. F. (1990) *Mol. Microbiol.* **4,** 1679–1691
13. Summers, R. G., Wendt-Pienkowski, E., Motamedi, H., and Hutchinson, C. R. (1992) *J. Bacteriol.* **174,** 1810–1820
14. Hopwood, D. A. (1997) *Chem. Rev.* **97,** 2465–2497
15. Malpartida, F. M., Hallam, S. E., Kieser, H. M., Motamedi, H., Hutchinson, C. R., Butler, M. J., Sugden, D. A., Warren, M., McKillop, C., Bailey, C. R., Humphreys, G. O., and Hopwood, D. A. (1987) *Nature* **325,** 818–821
16. Butler, M. J., Binnie, C., Hunter, I. S., Sugden, D. A., and Warren, M. (1990) *Dev. Ind. Microbiol.* **31,** 41–50
17. Kim, E. S., Bibb, M. J., Butler, M. J., Hopwood, D. A., and Sherman, D. H. (1994) *Gene* **141,** 141–142
18. Hranueli, D., Pigac, J., and Vešligaj, M. (1979) *J. Gen. Microbiol.* **114,** 295–303
19. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual,* 2nd Ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
20. Hopwood, D. A., Bibb, M. J., Chater, K. F., Kieser, T., Bruton, C. J., Kieser, H. M., Lydiate, D. J., Smith, C. P., Ward, J. M., and Schrempf, H. (1985) *Genetic Manipulation of Streptomyces; A Laboratory Manual,* The John Innes Foundation, Norwich, UK
21. Pigac, J., and Schrempf, H. (1995) *Appl. Environ. Microbiol.* **61,** 352–356
22. Sanger, F., Nicklen, S., and Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. U. S. A.* **74,** 5463–5467
23. Williams, D. H., and Fleming, I. (1997) *Methods in Organic Chemistry,* 5th Ed., McGraw-Hill Publishing Co., London, UK
24. Sanders, J. K. M., and Hunter, B. K. (1987) *Modern NMR Spectrospcopy: A Guide for Chemists,* Oxford University Press, Oxford, UK
25. Hallam, S. E., Malpartida, F., and Hopwood, D. A. (1988) *Gene* **74,** 305–320
26. Hunter, I. S., and Hill, R. A. (1997) in *Bio/Technology of Antibiotics* (Strohl, W. R., ed), 2nd Ed., pp. 659–682, Marcel Decker, Inc., New York
27. Altschul S. F., Madden T. L., Schaffer A. A., Zhang J., Zhang Z., Miller W., and Lipman D. J. (1997) *Nucleic Acids Res.* **25,** 3389–3402
28. Bibb, M. J., Janssen, G. R., and Ward, J. M. (1985) *Gene* **38,** 215–226
29. Ward, J. M., Janssen, G. R., Kieser, T., Bibb, M. J., Buttner, M. J., and Bibb, M. J. (1986) *Mol. Gen. Genet.* **203,** 468–478
30. Southern, E. M. (1975) *J. Mol. Biol.* **98,** 503–517
31. Yu, T. W., Shen, Y., McDaniel, R., Floss, H. G., Khosla, C., Hopwood, D. A, and Moore, B. S. (1998) *J. Am. Chem. Soc.* **120,** 7749–7759
32. Agarwal, P. K. (1989) *Carbon-13 NMR of Flavonoids*, Elsevier, Amsterdam
33. Fu, H., Ebert-Khosla, S., Hopwood, D. A., and Khosla, C. (1994) *J. Am. Chem. Soc.* **116,** 4166–4168
34. Fernandez-Moreno, M. A., Martinez, E., Boto, L., Hopwood, D. A., and Malpartida, F. J. (1992) *J. Biol. Chem.* **267,** 19278–19290
35. Ylihonko, K., Tuikkanen, J., Jussila, S., Cong, L., and Mantsala, P. (1996) *Mol. Gen. Genet.* **251,** 113–120
36. Ryan, M., Lotvin, J., Strathy, N., and Fantini, S. (1996) U. S. Patent 5589385-A
37. Yu, T. W., Bibb, M. J., Revill, W. P., and Hopwood, D. A. (1994) *J. Bacteriol.* **176,** 2627–2634
38. Ye, J., Dickens, M. L., Plater, R., Li, Y., Lawrence, J., and Strohl, W. R. (1994) *J. Bacteriol.* **176,** 6270–6280
39. Kramer, P. J., Zawada, R. J., McDaniel, R., Hutchinson, C. R., Hopwood, D. A., and Khosla, C. (1997) *J. Am. Chem. Soc.* **119,** 635–639
40. Zawada, R. J. X., and Khosla, C. (1997) *J. Biol. Chem.* **272,** 16184–16188
41. Thomas, R., and Williams, D. J. (1983) *J. Chem. Soc. Chem. Commun.* **7,** 443–444
42. Thomas, R., and Williams, D. J. (1983) *J. Chem. Soc. Chem. Commun.* **12,** 128–130
43. Devereux, J., Haeberli, P., and Smithies, O. (1984) *Nucleic Acids Res.* **12,** 387–395