Polypeptides are a vast class of metabolites made by living cells, including prokaryotes and eukaryotes (1, 2); at least some of them have a possible role in the physiology of the producers, while others are commercially valuable antibiotics or pharmacologically active materials, pigments, or flavoring agents (3). Despite their chemical heterogeneity, all of them share a common pattern of biosynthesis via the successive condensation of simple carboxylic acid metabolites. Their synthesis is conceptually similar to that of long-chain fatty acids, as first suggested by Collie (4). However, while fatty acid and polypeptide biosynthesis resemble each other in enzymology, there are substantial differences in detail: the lack of some reductive steps in polypeptide chain assembly, the heterogeneity of choices of starter and chain extender units among the different pathways, and the need for aromatization in some of them. These differences account for the great variety of chemical structures found among the polypetides.

Recently, several sets of genes for fatty acid synthases (FAS) of animals (5–7), fungi (8–12), and Escherichia coli (13) and polypeptide synthases (PKS) from actinomycetes, fungi, and plants (reviewed in Ref. 3) have been cloned and sequenced. Two structural types of complex enzymes (type I: multifunctional proteins, and type II: multienzyme complexes) can be found among the different PKSs (see Ref. 3 for review). This basic knowledge of the physical structure, organization, and functions of genes encoding polypeptide synthases will constitute a powerful tool to understand the mechanisms by which the primary polypeptide carbon chain is assembled and channelled toward the final product in the different pathways.

Actinorhodin is a polypeptide antibiotic produced by Streptomyces coelicolor A3(2), which is genetically the most characterized actinomycete (14). Blocked mutants were isolated and grouped into seven phenotypic classes (actI–VII) (15); the ability of different classes to convert intermediates secreted by the other mutants, blocked later in the pathway, into actinorhodin, placed the mutants in a biosynthetic sequence. By genetic complementation of one of the late-blocked mutants, the whole pathway was cloned in a single DNA fragment that conferred the ability to produce actinorhodin on a nonproducer, Streptomyces parvulus, when it was introduced into it by transformation (16).

By a combination of genetic complementation of the blocked act mutants and insertional inactivation of the cloned genes, the physical localization and organization of the act genes was determined on a 25-kb region of DNA within the cloned fragment (17). The act genes were grouped into three

---

**Nucleotide Sequence and Deduced Functions of a Set of Cotranscribed Genes of *Streptomyces coelicolor* A3(2) Including the Polypeptide Synthase for the Antibiotic Actinorhodin**

(Received for publication, March 18, 1992)

Miguel A. Fernández-Moreno‡§, Eduardo Martínez‡, Luis Boto‡¶, David A. Hopwood¶, and Francisco Malpartida‡**

From the 1Centro Nacional de Biotecnología, Serrano 115 bis, 28006 Madrid, Spain and Facultad de Medicina de la Universidad Autónoma de Madrid, Departamento de Bioquímica, Arzobispo Morcillo 4, 28029 Madrid, Spain and the 2John Innes Institute, John Innes Centre, Norwich NR4 7UH, United Kingdom

A 5.3-kb region of the *Streptomyces coelicolor* actinorhodin gene cluster, including the genes for polypeptide biosynthesis, was sequenced. Six identified open reading frames (ORF1–6) were related to genetically characterized mutations of classes actI, VII, IV, and VB by complementation analysis. ORF1–6 run divergently from the adjacent actII gene, which encodes the polypeptide synthase (PKS) ketoeductase, and appear to form an operon. The deduced gene products of ORF1–3 are similar to fatty acid synthases (FAS) of different organisms and PKS genes from other polypeptide producers. The predicted ORF5 gene product is similar to type II β-lactamases of *Bacillus cereus* and *Bacteroides fragilis*. The ORF6 product does not resemble other known proteins. Combining the genetic, biochemical, and similarity data, the potential activities of the products of the six genes can be postulated as: 1) condensing enzyme/acyl transferase (ORF1 + ORF2); 2) acyl carrier protein (ORF3); 3) putative cyclase/dehydrase (ORF4); 4) dehydrase (ORF5); and 5) “dimerase” (ORF6). The data show that the actinorhodin PKS consists of discrete monofunctional components, like that of the *Escherichia coli* (Type II) FAS, rather than the multifunctional polypeptides for the macrolide PKSs and vertebrate FASs (Type I).

---

* This research was supported in part by grants (to F. M.) from the Spanish Comisión Interministerial de Ciencia y Tecnología (Grants PB96-0025 and BIO96-0760), from the Economic European Community (Grant BAP 0393-E), and Smith, Kline & French Laboratories; by grants-in-aid to the John Innes Institute from the Agricultural and Food Research Council and the John Innes Foundation; and by North Atlantic Treaty Organization Collaborative Research Grant 870118 (to D. A. H., and F. M.) and National Institutes of Health Grant GM 39784 (to D. A. H.). 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.

‡ Supported by a fellowship from the Spanish Ministerio de Educación.

§ Supported by a postdoctoral fellowship from the Spanish Consejo Superior de Investigaciones Científicas.

The nucleotide sequence(s) reported in this paper has been submitted to the GenBank™/EMBL Data Bank with accession number(s) X63449.

** To whom correspondence should be addressed: Facultad de Medicina de la Universidad Autónoma de Madrid, Arzobispo Morcillo 4, 28029 Madrid, Spain. Fax: 34-1-5854013; Tel: 34-1-3975452.
Genes for Actinorhodin Biosynthesis in S. coelicolor A3(2)

**MATERIALS AND METHODS**

**Bacterial Strains, Plasmids, and Bacteriophages**—The E. coli strains were JM101 (21) and XL1-Blue (22). The Streptomyces strains and E. coli and Streptomyces vectors are summarized in Tables I and II, respectively.

**Media, Culture Conditions, and Microbiological Procedures**—Streptomyces manipulations were as in Hopwood et al. (23). Thiostrepton (a gift of S. J. Lucania, Bristol-Myers-Squibb Research Institute, Princeton, NJ) was used at a concentration of 50 µg/ml for solid media and 10 µg/ml for liquid media. Hygromycin (Sigma, Cat. number H2638) was used at 200 and 50 µg/ml in solid and liquid media, respectively. E. coli strains were grown on L agar or L broth (24).

**DNA Sequencing**—DNA sequencing was carried out by the deoxy-chain termination method (25); we used the 7-deaza-dGTP reagent kit from United States Biochemical (Cat. number 70750), following the manufacturer’s recommendations. Convenient DNA fragments were previously cloned on either M13 mp18 or M13 mp19 vectors from suitable restriction fragments or generated by ExoIII digestion (26).

**Computer Analysis of Sequences**—The DNA sequence was analyzed for open reading frames using CODONPREFERENCE (from the UWGCG package (27)). Amino acid sequences were analyzed using various programs from the UWGCG package (version 7.0, April 1991); comparisons of sequences were made against the EMBL gene data base (daily updated, October 1991), and Swissprot Data Base, (Release 19.0, updated August 1991), using FASTA, TFASTA, BESTFIT, COMPARE, and DOTPLOT. Protein alignments were made using LINEUP (from the UWGCG package).

**Gene Disruption and Mutant Assignments**—For gene disruption and phage “complementation” experiments, we used insert-directed recombination as described previously (28), using S. coelicolor strain J1501 or act mutants as hosts. The trans-complementation tests were carried out by transformation as described elsewhere (23) with either low or high copy number Streptomyces plasmids.

**DNA and RNA Manipulations**—For isolation, cloning, and manipulation of nucleic acids, the methods used were those in Hopwood et al. (23) for Streptomyces and Maniatis et al. (24) for E. coli.

**RESULTS AND DISCUSSION**

**DNA Sequence of the act early and Some intermediate Genes**

It was shown previously (17) that DNA between sites 14 and 21, in the right-hand region of the act cluster (Fig. 1),

---

*Footnote:

2E. Martinez, M. A. Fernandez-Moreno, D. A. Hopwood, and F. Malpartida, manuscript in preparation.
complements all mutations belonging to the actI, VII, IV, and VB classes, which block the actinorhodin pathway at early steps (actI, VII, and IV) and an intermediate step (actVB: the original set of actV mutants (15) was shown to consist of two phenotypically different subgroups (29), with actVA in the left-hand part of the act cluster and actVB in the right-hand part), and that at least the early genes might be expressed as a polycistronic mRNA. To determine the organization and primary structure of the act genes, and as a basis for further study of the biosynthetic steps, the DNA sequence of the act gene cluster; 21 nucleotides further downstream of ORF6 is 5'-end of ORF6 seems to represent the limit of the act region of the cluster was determined (Fig. 2).

DNA sequencing began at BamHI site 14 (Fig. 1), adjacent to the actIII gene (20), and extended rightward 5.3 kb to include the actVB region (to a BglII site close to position 19.2). Computer analysis of the DNA sequence, using CODONPREFERENCE, revealed a set of six ORFs (Fig. 3) which were named (from left to right) ORF1–6, respectively. All of them are oriented from left to right. The translation start point for each ORF was tentatively located (Fig. 2) using the following criteria: (a) distribution of GC content in the third position of the codons (30); (b) codon usage (for this purpose a table of codon usage from 64 different Streptomyces sequenced genes was used); and (c) location of a potential ribosome binding site (RBS), based on reasonable complementarity to the 3'-end of the 16 S ribosomal RNA sequence, immediately upstream of the region where the two previous criteria were satisfied. In nearly all cases the choice of start codon was unambiguous, and four ORFs (ORF1, ORF3, ORF4, and ORF6) are preceded by a good RBS. The choice of start codon for ORF1 was given special consideration because a TTG (nucleotides 107–109) appeared to be the most likely candidate. Although unusual as a start codon, TTG is used in a low proportion of genes in some microorganisms (31), including one putative example in Streptomyces (32) in spite of its high GC content. The translation product starting at this TTG (preceded by a good RBS) would be a polypeptide of about the same size as, and aligning very closely with, two highly similar PKS gene products in the gra (33) and tem (34) clusters (see below). The nearest in-frame potential start codon upstream of this TTG, a GTG, would lie within the adjacent BamHI fragment (sites 12 and 14 in Fig. 1) coding for the actII gene (nucleotides 8–6 of the complementary strand) (20) and would give a polypeptide 36 amino acids longer than the corresponding gra and tem gene products. Moreover, it would lack a good RBS and have a long N-terminal run of unusual codons. Even further upstream (nucleotides 27–29), another GTG is preceded by a relatively good RBS but also suffers from the other objections. Both these potential GTG starts seem highly improbable. The lengths of the six ORFs in amino acid residues, and the corresponding molecular weights, would be: ORF1, 424 amino acids (M, 45,034); ORF2, 407 amino acids (M, 42,523); ORF3, 86 amino acids (M, 9,242); ORF4, 316 amino acids (M, 34,643); ORF5, 297 (M, 31,967); ORF6, 177 amino acids (M, 18,381). Three pairs of ORFs (ORF1/2, ORF3/4, and ORF4/5) appear to be "translationally coupled" (35, 36), overlapping by 4 bp in each case. There is a 25-bp untranslated segment between ORF2 and ORF3, and an 11-bp gap between ORF5 and ORF6.

An interesting polymorphism, at nucleotide 4665, was found in ORF5. It was previously shown (17) that the S. coelicolor strain from which the act DNA was derived, M145 (16), lacks BglII site 19, which is present in another S. coelicolor A3(2) derivative, J1501 (as well as in Streptomyces lividans). Site 19 was picked up by homologous recombination on various subclones of the act DNA which had been passed through J1501. In strain J1501 the hexanucleotide starting at this position is AGATCT, giving BglII site 19, whereas in strain M145 the sequence is CGATCT. The ORF5 product in J1501 would have a glutamic acid residue in this region of the protein, while M145 would have alanine instead.

The 3'-end of ORF6 seems to represent the limit of the act gene cluster; 21 nucleotides further downstream of ORF6 is the end of another ORF, as deduced from computer analysis, which appears to run in the opposite direction to that of ORF6.

**Characterization of the act Early and Intermediate Region**

**Definition of the Limits of the act Early and Intermediate Region: An Operon Carrying Six act Genes**—In order to find out if DNA beyond ORF6 was involved in actinorhodin biosynthesis, pMM1 was constructed (Fig. 1). No blue colonies (indicating restitution of actinorhodin production) were obtained when strains B135 and B185 (actVB mutants), carrying the rightmost of the known act mutations, were transformed by 4 bp in each case. There is a 25-bp untranslated segment between ORF2 and ORF3, and an 11-bp gap between ORF5 and ORF6.
Genes for Actinorhodin Biosynthesis in S. coelicolor A3(2)

| TABLE II |
| Vectors and recombinant clones |

| Plasmids | Relevant characteristics | Ref. |
|----------|--------------------------|------|
| pUC18/19 | pBR322-derived E. coli vectors. bla. | (21) |
| pSU19/20 | pACYC184-derived E. coli vectors. cat. | (56) |
| pIJ2921  | pUC-derived E. coli vector with a modified polylinker flanked by BglII sites. bla. | G. Janssen (personal communication) |
| pIJ2925  | pUC-derived E. coli vector with a modified polylinker flanked by BglII and EcoRI + HindIII sites. bla. | G. Janssen (personal communication) |
| pMV222   | SphII (19.2)-BglII (21) fragment cloned on pUC18 between SphII-BamHI sites. | This work |
| pIJ702   | Streptomyces high copy number plasmid. mel, tsr. | (57) |
| pIJ486   | Streptomyces high copy number plasmid. tsr. | (58) |
| pIJ941   | Streptomyces low copy number plasmid; tsr, hyg. | (59) |
| pMM1     | SphI fragment from pMV222 cloned in pIJ941 (EcoRV site). | This work |
| pMM4     | SphI (19.2)-BglII (21) fragment cloned in pIJ702 (SphI-BglII sites). | This work |
| pMM5     | BglII (19)-BglII (21) fragment cloned on pIJ702 (BglII site). | This work |
| pIJ2350  | Same as pMM5 but with insert in opposite orientation. | This work |
| pMM11   | SphI (13.4)-KpnI (15.1) fragment cloned in pIJ486 (EcoRI-HindIII sites). | This work |
| pMM12   | SphI (13.4)-BclI (17.2) fragment cloned in pIJ486 (EcoRI-HindIII sites). | This work |
| pMM13   | SphI (13.4)-DdeI (17.3) fragment cloned in pIJ486 (EcoRI-HindIII sites). | This work |
| pMM14   | EcoRI-PstI fragment of pMM11 cloned in pIJ941 (EcoRI-PstI sites). | This work |
| pMM15   | EcoRI-PstI fragment of pMM12 cloned in pIJ941 (EcoRI-PstI sites). | This work |
| pMM16   | EcoRI-PstI fragment of pMM13 cloned in pIJ941 (EcoRI-PstI sites). | This work |

| Phages | Relevant characteristics | Ref. |
|--------|--------------------------|------|
| MMmp18/19 | E. coli phages vectors for DNA sequencing. | (21) |
| PM1    | Streptomyces OC31-derived vector; att*, tsr, hyg. | (17) |
| KC516  | Streptomyces phage vector. att*, tsr, uph. | (60) |
| ØME3a  | SacII (nt. 4157–4812) fragment cloned in PM1 (BamHI-PstI) sites. tsr. | This work |
| ØME15a | BglII (16)-SphI (19.2) fragment cloned in KC516 (BglII site). tsr. | This work |
| ØME17a | BstEII (2.1 kb from nt. 3668) fragment cloned in KC516 BglII. tsr. | This work |
| ØAB19  | BclI fragment, sites 17.2 to 19.3 cloned in PM1 (BamHI-BglII sites). hyg. | This work |
| ØG7    | BglII (19)-BglII (21) fragment cloned in KC516 (BglII site). tsr. | This work |
| ØAB23b | SphI (13.4)-NcoI (nt. 2624) fragment cloned in PM1 (PvuII site). hyg. | This work |
| ØL22a  | DdeI (within actIII gene)-DdeI (17.3) fragment cloned in PM1 (PvuII site). hyg. | This work |

* The abbreviations used are: bla, ampicillin resistance gene; cat, chloramphenicol acetyltransferase gene; hyg, hygromycin resistance gene; mel, tyrosinase gene; tsr, thiostrepton resistance gene; uph, viomycin phosphotransferase genes; nt, nucleotide.

** These plasmids or phages were constructed by previously cloning either directly or as blunt ended fragments in an intermediate E. coli vector as follows, ** pIJ2925, ** pSU20, ** pUC19, then rescued and ligated to Streptomyces vectors as indicated.

With pMM1. To confirm that this result was really due to pMM1 not including the actVB gene, rather than a low level of gene products, the same fragment was cloned on the high copy number vector pIJ702 to yield pMM4, which gave the same result. We therefore conclude that the region to the right of SphI site 19.2 does not code for any act early and intermediate genes. In addition, we set up a set of overlapping DNA fragments, which covered the whole sequenced region from BamHI site 14 to BglII site 21, as probe in S1 nuclease protection experiments, a transcript of nearly 5.5 kb was identified, which covers the region coding for the actI, VII, IV, and VB genes (data not shown). The limits of the S1 protected fragment suggest that transcription starts upstream of BamHI site 14 and ends near SphI site 19.2 at the 3' end of ORF6. Two divergent promoters (37) were recently mapped within the BamHI fragment (sites 13 to 14 ~203 nucleotides upstream of the putative translation start codon for ORF1, one driving transcription toward the actI region and the second promoter being responsible for actIII transcription (20). The "actI promoter" described by Parro et al. (37) would be responsible for transcription of the actI, IV, VII, and VB genes leading to a polycistronic mRNA. Another piece of
Fig. 2. Nucleotide sequence of the 5.3-kb actI, VII, IV, and VB region of the act cluster. Below the DNA sequence are the deduced translated products in single letter code. The putative ribosomal binding sites are shown by multiple asterisks and start codons are underlined. The deduced directions of expression for the open reading frames are indicated by thick arrows and named accordingly. KS identifies the presumed active site Cys of the ketosynthase domain of ORF1 and the presumed active site Ser of its acyltransferase domain. PPB identifies the presumptive 4'-phosphopantetheine-binding Ser residue of the ACP (ORF3).

Evidence for transcription of ORF's 1–6 as an operon was obtained by insertional inactivation caused by phage OME3. This recombinant phage carries a 655-bp SacII fragment (Fig. 1) overlapping ORF5 and ORF6; the nonproducer phenotype of lysogens confirmed that the fragment is internal to a single transcript and, therefore, placed ORF5 and ORF6 in the same
Genes for Actinorhodin Biosynthesis in S. coelicolor A3(2)

---

**Figure 2—continued**
Genes for Actinorhodin Biosynthesis in S. coelicolor A3(2)

mRNA. Insertional inactivation with a BglII fragment (sites 16 to 19), which covers ORFs 2–5, was previously reported (17), in good agreement with our conclusions. (Note also that there are untranslated nucleotides only between ORFs 2 and 3 and ORF’s 5 and 6; see above.)

Identification of the act Early and Intermediate Loci—To try to assign act mutant classes to specific ORFs, we cloned a set of overlapping fragments (Fig. 1 and Table II), either in att C31 derivatives or in pJ702, pJ486, and pJ941 plasmids and introduced them by transduction or transformation into several act mutants. For the phages, restoration of actinorhodin production (blue color) needed a single crossover to introduce the wild-type allele carried on the phage vector into the chromosome; this occurs when either the end or the beginning of the transcript, including the promoter, is present in the cloned fragment and when the crossover takes place between the 5’-end of the mRNA and the mutation point. Thus, the percentage of wild type (blue) lysogens would reflect the distance between the mutation and the 5’-end of the cloned DNA. (Recombination downstream of the mutation, between it and the 3’-end of mRNA, would restore functional genes only if an efficient promoter from the phage can transcribe stably the genes located downstream of it; an orientation-dependent phenotype would indicate this situation (17).

In order to associate specific ORFs with the act mutations, phages ØME15, ØME17, ØAB19, and ØG7 were made (Fig. 1 and Table II). ØME15 carries 65% of ORF2, complete ORF3, ORF4, and ORF5, and an almost complete ORF6 (lacking the three C-terminal codons). ØME17 contains 16% of ORF4 and complete ORF5 and ORF6, plus 500 bp to the right. ØAB19 harbors less than 10% of ORF3, the whole of ORF4, ORF5, and ORF6, and 1.5 kb to the right. ØG7 carries less than 6% of the C terminus of ORF5, a complete ORF6, and approximately 2.5 kb to its right.

To locate unambiguously the actVB mutations plasmids pJL2350, pMM4, and pMM5 were constructed carrying the 3-kb BglII fragment (sites 19–21 in Fig. 1) in both orientations: in pJL2350, ORF6 is oriented in the same direction as the me1 gene, while in pMM5 ORF6 and me1 are divergent. pMM4 is the same as pJL2350 but lacks the 600-bp SphI fragment which harbors the whole of ORF6.

The above recombinant plasmids and phages were introduced by transformation or transduction into different act mutants. The results (summarized in Table III) enabled us to locate act mutations representing each of the phenotypic classes as follows.

The actIV mutations (act-112 and act-131) seem to lie in ORF5, in accord with the frequency of complementation observed with ØME15, ØME17, ØAB19, and ØG7: the closer the mutation is to the N terminus of the ORF5 protein (or conceivably in the C terminus of ORF4), while in TK16 the complementation frequency is higher, the act mutation (act-117) must be located nearer the C terminus of ORF5.

The actVB mutations were localized within ORF6 because of trans-complementation using pJL2350, but not by pMM1, pMM4, and pMM5, and by the frequencies of blue colonies obtained using the set of four phages. Complementation of the 5.3-kb actI, VII, IV, and VB region. A codon usage table, constructed from 64 different Streptomyces genes (see "Materials and Methods"), was used with the UWGCG program to scan the DNA sequence. The possible ORFs are indicated by open rectangles; the presumptive ORFs which fit the criteria for coding regions (see text) are indicated by shaded arrows showing the direction of transcription. Only the overall distribution of GC content in the third position of the codons is shown (jagged lines); the distributions of rare codons are included in each reading frame (short vertical lines).

Identification of the act Early and Intermediate Loci—To try to assign act mutant classes to specific ORFs, we cloned a set of overlapping fragments (Fig. 1 and Table II), either in att C31 derivatives or in pJ702, pJ486, and pJ941 plasmids and introduced them by transduction or transformation into several act mutants. For the phages, restoration of actinorhodin production (blue color) needed a single crossover to introduce the wild-type allele carried on the phage vector into the chromosome; this occurs when either the end or the beginning of the transcript, including the promoter, is present in the cloned fragment and when the crossover takes place between the 5’-end of the mRNA and the mutation point. Thus, the percentage of wild type (blue) lysogens would reflect the distance between the mutation and the 5’-end of the cloned DNA. (Recombination downstream of the mutation, between it and the 3’-end of mRNA, would restore functional genes only if an efficient promoter from the phage can transcribe stably the genes located downstream of it; an orientation-dependent phenotype would indicate this situation (17).

In order to associate specific ORFs with the act mutations, phages ØME15, ØME17, ØAB19, and ØG7 were made (Fig. 1 and Table II). ØME15 carries 65% of ORF2, complete ORF3, ORF4, and ORF5, and an almost complete ORF6 (lacking the three C-terminal codons). ØME17 contains 16% of ORF4 and complete ORF5 and ORF6, plus 500 bp to the right. ØAB19 harbors less than 10% of ORF3, the whole of ORF4, ORF5, and ORF6, and 1.5 kb to the right. ØG7 carries less than 6% of the C terminus of ORF5, a complete ORF6, and approximately 2.5 kb to its right.

To locate unambiguously the actVB mutations plasmids pJL2350, pMM4, and pMM5 were constructed carrying the 3-kb BglII fragment (sites 19–21 in Fig. 1) in both orientations: in pJL2350, ORF6 is oriented in the same direction as the me1 gene, while in pMM5 ORF6 and me1 are divergent. pMM4 is the same as pJL2350 but lacks the 600-bp SphI fragment which harbors the whole of ORF6.

The above recombinant plasmids and phages were introduced by transformation or transduction into different act mutants. The results (summarized in Table III) enabled us to locate act mutations representing each of the phenotypic classes as follows.

The actIV mutations (act-112 and act-131) seem to lie in ORF5, in accord with the frequency of complementation observed with ØME15, ØME17, ØAB19, and ØG7: the closer the mutation is to the N terminus of the ORF5 protein (or conceivably in the C terminus of ORF4), while in TK16 the complementation frequency is higher, the act mutation (act-117) must be located nearer the C terminus of ORF5.

The actVB mutations were localized within ORF6 because of trans-complementation using pJL2350, but not by pMM1, pMM4, and pMM5, and by the frequencies of blue colonies obtained using the set of four phages. Complementation of
Plasmids and phages carrying different combinations of ORFs and an Act- phenotype was observed. These results might well be due to a titration by the cloned thiostrepton-sensitive phage carrying actIIORF4, and thereand B185 (actVB mutants) when lysogenized with @ME17 that ORF6 might be transcribed from a vector promoter, while phenotypes because they lack ORF6. Moreover, the only overconfirming that ORF6 is responsible of actVB complementation.

The experiments leading to the identification of ORF4 as act/VB mutants with pIJ2350 but not with pMM5 suggests that ORF6 might be transcribed from a vector promoter, while neither pMM4 nor pMM1 are able to restore the wild type phenotype since they lack ORF6. Moreover, the only overlapping region between ORF7 and ORME15 is precisely ORF6, confirming that ORF6 is responsible of act/VB complementation. The frequencies of blue lysogens observed with B135 and B185 (act/VB mutants) when lysogenized with ORM17 are the same as for TK16, suggesting that act235 and act285 might be located close to the 5'-end of the ORF6 protein.

In order to localize the positions of the actI mutations, the available actI mutants were transformed with recombinant plasmids and phages carrying different combinations of ORFs 1-3. pMM11 (containing only ORF1 complete) and pMM13 (ORF1, ORF2, and ORF3) were constructed on the high copy number vector pIJ486. None of the actI mutants transformed with them yielded an Act- phenotype. Moreover, when the wild type S. coelicolor J1501 was transformed with pMM11 or pMM13 an Act- phenotype was observed. These results strongly suggest that the Act- phenotype observed with these constructions might well be due to a titration by the cloned fragments of some trans-acting element needed for expression of the act genes and thus presumably preventing expression of other act transcripts; a good candidate for such a trans-acting element is the actII-ORF4 gene product whose correct translation is essential for transcription of the actI gene (18). Furthermore, when pMM11 was introduced into S. coelicolor J1501, previously lysogenized with a hygromycin-resistant/thiostrepton-sensitive phage carrying actI ORF4, and therefore with one extra copy of the activator gene, the colonies selected (tar/hyg) overcame the mutagenic effect. In addition, the failure of complementation with high copy number plasmids contrasted with the success in using the low copy plasmid pIJ491 or “complementation” using att phages. The low copy number recombinant plasmids were pMM14 (carrying ORF1 complete), pMM15 (ORF1 and 2 complete), and pMM16 (ORF1, 2, and 3 complete); and the phages were ØL22 (containing complete ORFs 1, 2, and 3) and ØAB23 (carrying complete ORFs 1 and 2, without any ORF3 sequences), both of them starting at SphI site 13.4 (Fig. 1).

Table III (a) shows the results obtained with the different constructions in the complementation of the actI mutants. Four of the 12 actI mutations were complemented by ORF1 alone, and therefore lie in this ORF. The other eight were not complemented by ORF1 and are therefore candidates for ORF2 mutations (although we cannot exclude the possibility that some of them might be deletions involving ORF1 and ORF2 or polar mutations in ORF1). The absence of any within ORF3 may be due to the fact that ORF3 represented a small DNA target in relation to the whole act region (less than 9% of ORFs 1, 2, and 3) when the wild type S. coelicolor was mutagenized (15).

In view of the data reported here, based on DNA sequencing and complementation of the act mutants, the conclusion previously reported by Bartel et al. (39) must be corrected as follows: the minimal act DNA fragment which caused allo-saponarin I1 production by Streptomyces galilaeus strains (a XhoI fragment extending from a site to the left of the act region to XhoI site 17.1, cloned in pANT43) would have contained only ORF1 and ORF2 complete; it would have lacked the 3'-end of ORF3 (a component of the actI locus) and the whole of ORF4 (the actVII locus). Consequently the observed phenotype must have been the result of heterologous interaction between the ORF1 and/or ORF2 gene products from the act DNA and some endogenous components of S. galilaeus in the recombinant cultures.

Deduced Biochemical Functions of act Early and Intermediate Genes—To better understand the functions of each of the ORFs revealed by DNA sequencing, we searched the available data bases with their translated products. ORF1, ORF2, and ORF3 show strong similarities with other genes whose products are known to be components of either fatty
Genes for Actinorhodin Biosynthesis in S. coelicolor A3(2)

acid or polyketide synthases of different organisms. The resemblances are strongest (and are end-to-end) with the gra and tcm ORF1, 2, and 3 gene products (33, 34) (between 55 and 80% identity) and less extensive, although significant, with other gene products like 6-methylsalicylic acid synthase catalytic domains can be postulated within ORF1. These are: a \( \beta \)-ketoacyl-synthase (condensing enzyme) (Fig. 4) with significant alignments around the Cys residue, the presumptive active site of the synthase, with those in other Type I or Type II fatty acid and polyketide synthases. A second feature is around a Ser residue in a GHS motif which could represent a putative acyl transferase domain. This would imply that ORF1 codes for a bifunctional ketosynthase/acyl transferase and could reflect the lack of any identifiable separate acyl transferase(s) within the cluster. As in the gra and tcm clusters (33, 34), actI-ORF1 and actII-ORF2 strongly resemble each other (29% identity and 49% similarity), and their stop and start codons overlap as in many other bacterial operons (35, 36). This suggests a translational coupling which probably ensures equimolar production of the two gene products. There is no sign of the corresponding active site domains in ORF2. Therefore, as previously suggested for the corresponding genes in the granaticin and tetracenomycin PKS (33, 34), we suggest that the ORF2 product may perhaps function, together with the ORF1 product, as a heterodimeric protein.

The ORF3 product is a small polypeptide which, from its resemblance to other known proteins, would function as an acyl-carrier protein (ACP); a typical phosphopantetheine binding domain can be identified centered on the so-called "active-site" Ser, which is well conserved in other FAS and PKSs from different origins; a derived consensus sequence is given below. The conserved residues are in boldface. Ratfas, rat FAS (5); Scfas2, S. cerevisiae FAS2 (11); Scfas1, S. cerevisiae FAS1 (8-10); Ppfas2, P. patulum FAS2 (12); Ppmsa, P. patulum MSA synthase (40); DEBS 3 (C, N), S. erythraea actI-ORF3, S. glaucescens tcm-ORF1, 2, 3 gene products (33, 34); actIorf1, 2, 3, S. coelicolor actI-ORF1, 2, 3 (this work). The multiple alignments were done using GAP and PRETTY from the GCG package (see "Material and Methods").

\[ \beta \text{-keto synthase motif.} \]

\[
\text{Active site thiol}
\]

![Diagram of the \( \beta \)-keto synthase motif.]

\[
\text{Acyl transferase motif.}
\]

\[
\text{Active site hydroxyl}
\]

![Diagram of the acyl transferase motif.]

\[4'\text{-phosphopantetheine binding site}
\]

![Diagram of the 4' phosphopantetheine binding site.]

(a): malonyl/palmitoyl transferase domain
(b): acetyl transferase domain
PKS complexes (Fig. 4). Thus, altogether, the ORF1, ORF2, and ORF3 products would form the enzymatic complex needed for the condensation of one acetate starter unit with seven malonyl extender units to produce the poly-β-ketone carbon backbone of the actinorhodin half-molecule (Fig. 5). Searching the available data bases with the translated product of ORF4 failed to reveal similarities with other known proteins, apart from gra-ORF4 (33), tcmI-ORF4 (34), and whiE-ORFVI (43) products. By complementation experiments actVII mutations were mapped to ORF4 (38), while chemical evidence (44) has revealed that the product secreted by the actVII mutant is mutactin, a shunt product of the actinorhodin pathway, thought to be derived from an aberrant cyclization of the polyketide chain between C-15 and C-6 instead of C-14 and C-5 (Fig. 5). This led to the conclusion that the ORF4 product is needed for correct cyclization of the oligoketide leading to isochromanequinone formation. Since the hydroxyl group at C-9 is still present in mutactin, it was postulated that the ORF4 product also carries a dehydrase domain, in its C-terminal half (38).

Overlapping the TGA stop codon of ORF4 is an ATG (nucleotides 3830-3832); this start codon is not preceded by a typical RBS but is the first one after a typical change in the distribution of rare codons between reading frames (Fig. 3). We postulate that the beginning of ORF5 is this ATG, rather than the more upstream GTG (nucleotides 3419-3421) within ORF4. Searching the available data bases with the ORF5-translated product revealed strong similarities with the products of cphA (45) (55% similarity and 27% identity over 232 overlapping amino acids), cflA (46) (49% similarity and 26% identity over 239 overlapping amino acids) and ccrA (47) (49% similarity and 26% identity over 239 overlapping amino acids) (see Fig. 6 for Dotplots and alignments). These three proteins are Type II β-lactamases. There are less extensive similarities with other β-lactamases belonging to the same metallothioprotein group such as those from Bacillus cereus 569/H (48) and B. cereus 5/B/6 (49). The corresponding regions are in a domain thought to be important for Zn²⁺ binding: 2 closely located His residues, 1 Cys, and a further His near the C terminus of the proteins (48, 49). The two regions around the His residues can be aligned together, while the region next to the Cys residue of the Type II β-lactamases, although very similar to the region in the ORF5 product, lies immediately upstream of the Cys rather than being centered on it. Nevertheless, a Cys residue is located within this domain of ORF5 (amino acids 143, Fig. 6), whose counterpart in the Type II β-lactamases is a Tyr residue.

It is not at all obvious why the actIV gene product, which has been postulated (39) to act as a dehydrase that would remove a hydroxyl group from C-5 after the two carbon rings of the isochromanequinone are established (Fig. 5), should resemble β-lactamases, which are believed to be specific for hydrolysis of the β-lactam ring system; such a reaction cannot rationally be implicated in a biosynthetic scheme for actinorhodin. However, perhaps it is significant that the segments of the β-lactamases that resemble the ActIV protein are in regions thought to be important for Zn²⁺ binding, which include 2 closely located His residues (only 1 in CphA), a Cys, and a further His near the C terminus of the proteins (48). Corresponding residues are seen in the ActIV protein, except that the Cys is replaced by an aspartic acid residue. In some other Zn²⁺ proteins, 1 acidic residue acts as one of the three required ligands for the metal (50–52). Thus, perhaps the significance of the resemblance with the β-lactamases is a common evolutionary origin as Zn²⁺ enzyme, rather than a relationship between the present day enzymatic activities of the proteins, with the ActIV protein having acquired a role as a dehydrase. Interestingly, the homologies between the ORF5 product and the β-lactamases lie immediately downstream of the processing site of the pre-β-lactamases (46, 48, 49); this is consistent with the presumed intracellular location of the ActIV protein (as a biosynthetic enzyme), in contrast to the extracellular location of the β-lactamases.

Searching the available data bases with the sequence of the putative ORF6 product failed to show any significant similarity with other proteins. Cole et al. (29) argued that the
biosynthetic block caused by the actVB mutations, here shown to lie in ORF6, is probably in the dimerization reaction which would join two molecules of a late precursor of actinorhodin to produce the final antibiotic structure. Little is known about such phenolic oxidative coupling reactions in biological systems and so the further study of the ORF6 gene product may provide important insights into this class of bioorganic reactions. The sequencing strategy is outlined in Fig. 7.

CONCLUSION

This paper has described a set of five cotranscribed genes. Four of them (actI-ORF1, 2, 3, and actVII) code for components of the polyketide synthase that assembles and cyclizes the carbon backbone of the actinorhodin half-molecule. A fifth gene, actIII (20), which encodes the ketoreductase of the PKS, is adjacent to the actI/VII genes but, perhaps surprisingly, is transcribed independently from them, in the opposite direction, and from a promoter that appears to represent a different class from the promoter for the actI/VII transcript. This transcript also carries the gene (actIV) that appears to catalyze the next step in biosynthesis after the PKS but, again unexpectedly, continues on to include the actVB gene whose putative role, as a dimerase for the immediate precursors of the actinorhodin half molecule, acts considerable later in the pathway. Clearly, many intriguing features remain to be elucidated about the organization of the act cluster, which may be the first actinomycete antibiotic gene cluster to have been completely sequenced.

Acknowledgments—We thank M. J. Bibb for making available the Streptomyces coelicolor A3(2) codon usage table; S. J. Lucania, for the gift of thiostrepton; K. F. Chater, M. J. Bibb, C. Khosla, P. J. Revill, and J. Dubart for helpful comments on the manuscript.

REFERENCES

1. Birch, A. J., and Donovan, F. W. (1955) Aust. J. Chem. 6, 300–308
2. Mann, J. (1987) Secondary Metabolism, Clarendon, Oxford
3. Hopwood, D. A., and Sherman, D. H. (1990) Annu. Rev. Genet. 24, 37–66
4. Collie, J. N. (1977) J. Chem. Soc. 91, 1806–1813
5. Amy, C. M., Witkowski, A., Naggett, J., Williams, B., Randhawa, Z., and Smith, S. (1989) Proc. Natl. Acad. Sci. U. S. A. 86, 3114–3118
6. Holzer, K. P., Liu, W., and Hammes, G. G. (1989) Proc. Natl. Acad. Sci. U. S. A. 86, 4387–4391
7. Yuan, Z., Liu, W., and Hammes, G. G. (1988) Proc. Natl. Acad. Sci. U. S. A. 85, 6328–6332
8. Chirala, S. S., Kuziora, M. A., Spector, D. M., and WakiI, S. J. (1987) J. Biol. Chem. 262, 4231–4240
9. Schweizer, M., Roberts, L. M., Hölleke, H., Takabayashi, K., Höllere, E.,
Leskiw, B. K., Mevarech, M., Barritt, L. S., Jensen, D. J., Devereux, J., Haeberli, P., and Smithies, O. (1988) Mol. Gen. Genet. 205, 479-486

Mohamed, A. H., Chirala, S. S., Mody, N. H., Huang, W., and Wakil, S. J. (1988) J. Biol. Chem. 263, 12131-12135

Wiesner, P., Beck, J., Beck, K., Ripka, S., Müller, G., Lücke, S., and Schweizer, E. (1988) Eur. J. Biochem. 177, 69-79

Kauppinen, S., Siggaard-Andersen, M., and von Wettstein-Knowles, P. (1988) Carlsberg Res. Commun. 53, 357-370

Hopwood, D. A., and Kieser, T. (1990) in The Bacterial Chromosome (Drlica, K., and Riede, M., eds) pp. 147-162, American Society for Microbiology, Wash. D.C.

Rudd, B. A., and Hopwood, D. A. (1979) J. Gen. Microbiol. 114, 35-43

Malpartida, F., and Hopwood, D. A. (1984) Nature 309, 462-464

Malpartida, F., and Hopwood, D. A. (1986) Mol. Gen. Genet. 205, 66-73

Fernández-Moreno, M. A., Caballero, J. L., and Malpartida, F. (1991) Cell 66, 769-780

Caballero, J. L., Martinez, E., Malpartida, F., and Hopwood, D. A. (1991) Mol. Gen. Genet. 229, 401-412

Hallam, S. E., Malpartida, F., and Hopwood, D. A. (1988) Gene (Amst.) 74, 305-320

Yanisch-Perron, C., Vieira, J., and Messing, J. (1985) Gene (Amst.) 33, 103-119

Bullock, W., O., Fernández, J. M., and Short, J. M. (1987) BioTechniques 5, 376

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. (1984) Genetic Manipulation, A Laboratory Manual, The John Innes Foundation, Norwich, United Kingdom

Maniatis, T., Fritsch, E. F., and Sambrook, J. (1982) Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY

Sanger, F., Nicklen, S., and Coulson, A. R. (1977) Proc. Natl. Acad. Sci. U.S.A. 74, 5463-5467

Hoeft, J. R., and Chater, K. F. (1980) J. Bacteriol. 140, 287-295

Devereux, J., Haebeler, P., and Smithies, O. (1984) Nucleic Acids Res. 12, 887-896

Chater, K. F., and Bruton, C. J. (1988) Gene (Amst.) 74, 67-78

Cole, S. P., Rudd, B. A., Hopwood, D. A., Chang, C. J., and Floss, H. G. (1985) J. Bacteriol. (Tokyo) 40, 340-347

Bibb, M. J., Findlay, P. R., and Johnson, M. W. (1984) Gene (Amst.) 30, 157-166

Gold, L., and Stormo, G. (1987) in Escherichia coli and Salmonella typhimurium Cellular and Molecular Biology (Neidhardt, F. C., Ingraham, J. L., Low, K. B., Magasanik, B., Schaechter, M., and Umbarger, H. E., eds) pp. 1302-1307, American Society for Microbiology, Wash. D.C.

Leskiv, B. K., Mevarech, M., Barratt, L. S., Jensen, S. E., Henderson, D. J., Hopwood, D. A., Bruton, C. J., and Chater, K. F. (1989) J. Gen. Microbiol. 135, 1313-1322

Sherman, D. H., Malpartida, F., Bibb, M. J., Kieser, H. M., Bibb, M. J., and Hopwood, D. A. (1989) EMBO J. 8, 2717-2725

Bibb, M. J., Biro, S., Motamedi, H., Collins, J. F., and Hutchinson, C. R. (1989) EMBO J. 8, 2727-2736

Binas, E., Hugias, E., Chen, J. W., and Umbarger, H. E. (1988) J. Bacteriol. 170, 4785-4807

Zalkin, H., and Ebbole, D. J. (1988) J. Biol. Chem. 263, 1585-1588

Parro, Y., Hopwood, D. A., Malpartida, F., and Mellado, R. P. (1991) J. Bacteriol. 173, 263-267

Sherman, D. H., Bibb, M. J., Simpson, T. J., Johnson, D., Malpartida, F., Fernández-Moreno, M., Martínez, E., Hutchinson, C. R., and Hopwood, D. A. (1991) Tetrahedron 47, 9029-9043

Bartel, P. L., Zhu, C. B., Lampel, J. S., Doeh, D. C., Consnor, N. C., Strohl, W. R., Beal, J. J., and Floss, H. G. (1990) J. Bacteriol. 172, 4815-4826

Beck, J., Ripka, S., Signer, A., Schütz, E., and Schweizer, K. (1990) Eur. J. Biochem. 192, 487-498

Cortés, J., Haydock, S. F., Roberts, G. A., Bevitt, D. J., and Leadlay, P. F. (1990) Nature 346, 176-178

Bevitt, D. J., Cortés, J., Haydock, S. F., and Leadlay, P. F. (1992) Eur. J. Biochem. 204, 39-49

Davis, N. R., and Chater, K. F. (1990) Mol. Microbiol. 4, 1679-1691

Zhang, H., He, X. G., Adefarati, A., Gallici, J., Cole, S. P., Beale, J. M., Keller, P. J., Chang, C., and Floss, H. G. (1990) J. Org. Chem. 55, 1682-1684

Massidda, O., Rossolini, G. M., and Satta, G. (1991) J. Bacteriol. 173, 4611-4617

Thompson, J. S., and Malamy, M. H. (1990) J. Bacteriol. 172, 2584-2593

Rasmussen, B. A., Gluzman, Y., and Tallj, F. P. (1990) Antimicrob Agents Chemother. 34, 1690-1692

Hussain, M., Carlino, A., Madonna, M. J., and Lampen, J. O. (1985) J. Bacteriol. 164, 223-229

Lim, H. M., Pene, J. J., and Shaw, R. W. (1988) J. Bacteriol. 170, 2873-2876

Ekblud, H., Horjale, E., Jonnwall, H., Brandén, C.-I., and Jeffery, J. (1985) Biochemistry 24, 802-806

Venter, S. L., and Auld, D. S. (1990) Proc. Natl. Acad. Sci. U. S. A. 87, 220-224

Chater, K. F., Bruton, C. J., King, A. A., and Suarez, J. E. (1982) Gene (Amst.) 19, 21-32

Feiselson, J. S., and Hopwood, D. A. (1985) Mol. Gen. Genet. 190, 394-398

Hopwood, D. A., Kieser, T., Wright, H. M., and Bibb, M. J. (1983) J. Gen. Microbiol. 129, 2257-2268

Mylén, T., and Rostock, A. (1991) Gene (Amst.) 102, 75-78

Katz, E., Thompson, C. J., and Hopwood, D. A. (1983) Gen. Microbiol. 119, 2703-2714

Ward, J. M., Jensen, G. R., Kieser, T., Bibb, M. J., Buttner, M. J., and Bibb, M. J. (1986) Mol. Gen. Genet. 203, 468-478

Lydiat, D. J., Malpartida, F., and Hopwood, D. A. (1985) Gene (Amst.) 39, 223-235

Rodicio, M. R., Bruton, C. J., and Chater, K. F. (1985) Gene (Amst.) 34, 283-292