JOURNAL OF BACTERIOLOGY, May 1995, p. 2942–2945
0021-9193/95/$04.00+0

Vol. 177, No. 10

Nucleotide Sequences and Expression of Genes from *Streptomyces purpurascens* That Cause the Production of New Anthracyclines in *Streptomyces galilaeus*

JARMO NIEMI* AND PEKKA MÄNTSÄLÄ

Department of Biochemistry and Food Chemistry, University of Turku, FIN-20500 Turku, Finland

Received 22 July 1994/Accepted 10 March 1995

Six open reading frames, *rdmA* to *rdmF*, in a 6,077-bp segment of *Streptomyces purpurascens* DNA which caused the production of hybrid anthracyclines were identified. The minimal fragment that produced anthracyclines modified at the 10th position contained *rdmB* to *rdmD*; *rdmE* is the gene for aklavinone-11-hydroxylase. RdmC is similar to a putative open reading frame in the daunorubicin biosynthetic cluster of *Streptomyces peucetius* and is likely to participate in the removal of the side chain at the 10th position.

Anthracyclines are a group of substances that contain the 7,8,9,10-tetrahydro-5,12-naphthacene quinone structure (31); they are usually antibiotics produced by actinomycetes or their synthetic analogs. Some of them are important pharmaceutical agents which are used in cancer therapy. The anthracycline chromophore, aglycone, is synthesized by the polyketide pathway in producing organisms (6). The polyketide synthases of several streptomycetes have been found to have a high degree of conservation (17, 18), which has enabled the biosynthetic genes of additional polyketide antibiotics to be cloned (22).

Hybrid antibiotics are molecules which combine structural features of two or more antibiotics that are not normally produced by the same organism and are often produced by transferring genes from one antibiotic producer to another strain that produces a structurally related antibiotic by recombinant DNA techniques (14, 16).

Previously, Niemi et al. (25, 27) described hybridizing a segment of *Streptomyces purpurascens* DNA near the locus with an *actI* (polyketide synthase) probe from *Streptomyces coelicolor* (22), which caused the production of glycosides of several modified aglycones (Fig. 1) in *Streptomyces galilaeus*, a producer of aklavinone glycosides. The genes that caused hybrid production were located on a 6.1-kb *Bam*HI-*Sau*3AI fragment, EB3 (including 9 bp from the polylinker of λEMBL4 [8] with an *Eco*RI site).

We have sequenced this fragment and expressed the aglycone-modifying enzyme activities it encodes by using a strong, constitutive promoter. Expression and sequence data indicate interesting similarities with daunorubicin and doxorubicin biosynthesis in *Streptomyces peucetius.*

Sequence analysis. A random subcloning strategy was used (28). M13mp18 (37) clones were sequenced with *Taq* polymerase (Promega Corporation, Madison, Wis.); with most clones, 7-deazaguanidine nucleotide mixes were used. Clone preselection by single-nucleotide tracking (26), clone inversion, and subcloning of selected restriction fragments were also used to completely sequence both strands of DNA. Assembly and analysis of the nucleotide sequence were performed with the Genetics Computer Group package (10). ATG and GTG start codons (30) were accepted, and the codon usage table for CODONPREFERENCE was prepared by using previously published data (36).

On the basis of codon bias (1), six presumed protein-coding open reading frames (ORFs), *rdmA* to *rdmF*, were observed (Fig. 2).

The beginning of *rdmA* is uncertain because although an apparent ribosome binding site precedes the GTG codon at 226, a relatively high third-base bias with no stop codons is observed from the beginning of this sequence. Our preliminary sequencing results for the homologous gene in *S. galilaeus* lead us to expect *rdmA* to begin a few amino acids before the start of this sequence. *rdmB* corresponds to a 374-amino-acid peptide (M_r, 39,764; isoelectric point, 4.5). The product of *rdmC* would be 298 amino acids (M_r, 31,792; isoelectric point, 5.0), that of *rdmD* would be 237 amino acids (M_r, 25,851; isoelectric point, 6.7), and that of *rdmE* would be 535 amino acids (M_r, 57,436; isoelectric point, 4.8). *rdmF* appears to be incomplete.

The peptide sequences of these ORFs were compared with the PIR (release 40.0), GenBank (release 82.0), and EMBL (release 38.0) databases. RdmA, RdmB, and RdmC resemble peptide sequences that have been translated from a nucleotide sequence that contains the carminomycin 4-*O*-methyltransferase (COMT) gene of *S. peucetius* (21). RdmB shows 51.7% identity with COMT. RdmA and RdmC resemble peptide sequences that correspond to the regions that are 5′ and 3′, respectively, of COMT, with the last 51 amino acids of RdmA 96.1% identical to the translated reading frame that begins at nucleotide 3 of the sequence and with the first 110 amino acids of RdmC 64.5% identical to the same reading frame downstream of COMT. RdmB also shows similarities to several other *O*-methyltransferases, and the conserved dinucleotide binding site DvGGG-G (21) is present.

RdmC shows relatively weak similarities to carboxyl esterase from *Pseudomonas* sp. strain KWI-56 (29), 2-hydroxy-6-oxo-6-phenylhexa-2,4-dienoic acid hydrolases from *Pseudomonas putida* KF715 (12) and *Pseudomonas* sp. strain KKS102 (20), and 2-hydroxy-6-oxohepta-2,4-dienoate hydrolase from *P. putida* F1 (23) (S34609, C35124, Espssk, and Jh0245, respectively [Fig. 3]). When RdmC, the corresponding peptide sequence from *S. peucetius* (Dnrorfc; translated from the previously published nucleotide sequence [21]), and the sequences of these hydrolases were aligned (Fig. 3), several conserved regions were observed.

* Corresponding author. Mailing address: Department of Biochemistry and Food Chemistry, University of Turku, Vatselankatu 2, FIN-20500, Turku, Finland. Phone: 358 21 633 6856. Fax: 358 21 633 6860. Electronic mail address: jarnie@utu.fi.

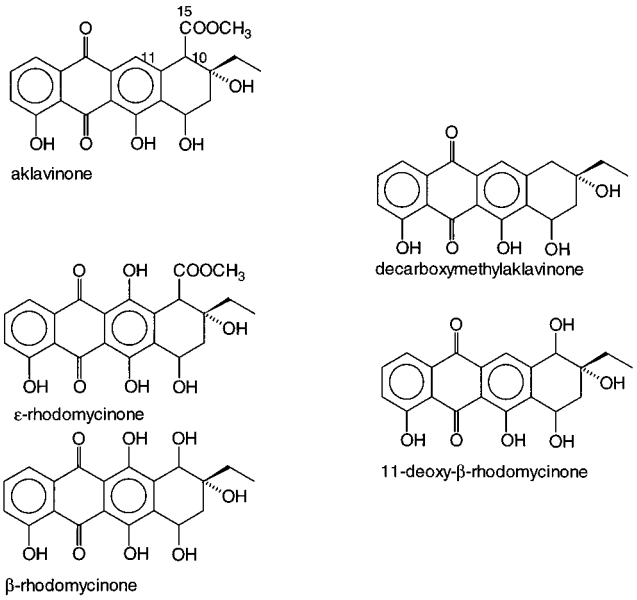

FIG. 1. Identified aglycones of anthracycline glycosides produced by *S. galilaeus* ATCC 31615 (pEB3). The host alone produces glycosides of aklavinone; the gene donor, *S. purpurascens*, produces glycosides of ε-rhodomycinone and β-rhodomycinone.

RdmD exhibits 45.6% identity with the product of *srmX* from *Streptomyces ambofaciens* (9). The function of the latter is unknown, but it probably participates in biosynthesis of spiramycin, a macrolide antibiotic.

RdmE shows significant similarities to several aromatic hydroxylases, including the *tcmG* hydroxylase (32) that participates in tetracenomycin biosynthesis and the fragment that is listed as aklavinone-11-hydroxylase in GenBank sequence file M73758 (11). Over this 78-amino-acid fragment, RdmE and the latter sequence are 68.8% identical.

RdmF resembles the product of an ORF from *Daucus carota* AX110P mRNA (24) (28.3% identity), which is expressed in carrot embryogenic cells and whose function is unknown.

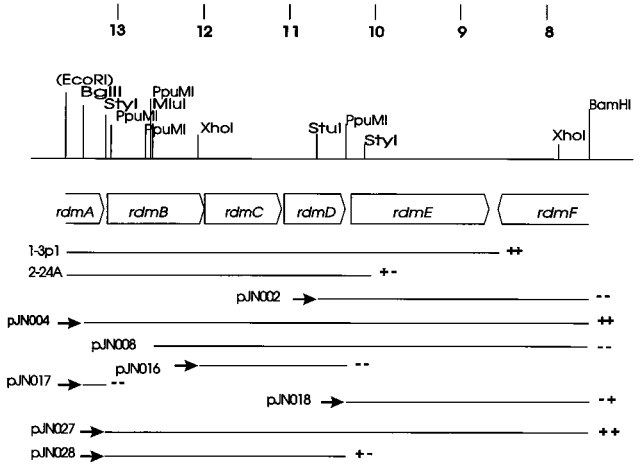

FIG. 2. Schematic representation of the proposed ORFs in the EB3 sequence and expression constructions tested. The kilobase scale above the plot shows the same coordinates as those in previously published diagrams (27). Only the sites of restriction enzymes used to make these constructions are shown (the *Eco*RI site in parentheses is a vector site). The portion of sequence in each construction is indicated by a solid line; an arrow indicates the *ermE* promoter. Constructions without *ermE* are in pIJ486 in the expressing orientation. The first + or − after each construction indicates production or lack of production, respectively, of aglycones with 10-modifications; the second indicates production or lack of production of aglycones with 11-hydroxylation.

Expression studies. The bacterial strains used in this study have previously been described (27). Expression constructions, shown schematically in Fig. 2, which contained the wild-type *ermE* promoter (2) were assembled in *Escherichia coli* between the *Eco*RI and *Hin*dIII sites in the polylinker of pUC18 (37) or pBS(−) (Stratagene, Inc., La Jolla, Calif.), subcloned into *Streptomyces lividans* in the polylinker of pIJ487 (15, 34), and subsequently transferred into *S. galilaeus*. Plasmid pJN008 was constructed by digesting pEB3 (27) (EB3 in pIJ486, prepared from *S. galilaeus*) with *Eco*RI and *Mlu*I, isolating the large fragment, producing blunt ends with Klenow polymerase, ligating, and introducing by transformation into *S. galilaeus*. Culture conditions for *S. galilaeus* and the extraction and thin-layer chromatography methods for anthracycline products have previously been described (27). The production of products that were modified at the 10th and 11th positions (10- and 11-modified products, respectively) by each clone is indicated in Fig. 2.

Interpreted in terms of observed sequence features, previous results (27) indicated that deletions that extended into *rdmE* caused the loss of 11-hydroxylase activity and deletions into *rdmD* caused the loss of 10-modifications (deletion constructions 1-3p1 and 2-24A [Fig. 2], respectively).

All *ermE* constructions conferred kanamycin resistance at a level of at least 5 μg/ml in *S. lividans*, suggesting readthrough transcripts into the *aphI* gene of promoter-probe vector pIJ487.

Conclusions. Since *rdmA* and *rdmF* are not needed in constructions for hybrid production, no function can be assigned to them. The fact that the sequence of *rdmA* is highly conserved among anthracycline producers *S. purpurascens*, *S. peucetius*, and *S. galilaeus* but, according to Southern hybridization, is not present in *S. lividans* (data not shown) suggests that it is an important component in the anthracycline biosynthetic apparatus.

It is surprising that no compounds with 10-modifications were produced by constructions that lacked *rdmB*, which by sequence comparison appears to code for an *O*-methyltransferase whose closest relative is the COMT of *S. peucetius*. The nucleotide sequence homology that extends from *rdmA* to *rdmC* is equally intriguing and suggests that these three genes form an evolutionally conserved combination. No 4-O-methylated products from *S. purpurascens* have been described, and none were found in hybrid-producing *S. galilaeus* transformants (27).

In 10-demethoxycarbonylaklavinone, the side chain has been removed in the same manner as in daunorubicin. RdmC may participate in this reaction; a highly similar ORF is apparently present in *S. peucetius* in close association with the gene of another late modification, COMT. Similarities with enoate hydrolases (Fig. 3) (12, 20, 23) suggest that this reaction proceeds by hydrolysis of the C_{10}—C_{15} bond, but similarity with the *Pseudomonas* esterase (29) indicates possible hydrolysis of the methyl group. It is possible that RdmC catalyzes both reactions, as decarboxylation of aklavinone carboxylic acid is known to occur spontaneously in dimethylformamide (33, 35). The conserved active-serine sequence GXSXG (4) is present around Ser-102, in addition to a conserved putative active-site histidine at 276. The stop codon of *rdmB* overlaps the putative ribosome binding site of *rdmC*, an arrangement that is associ-

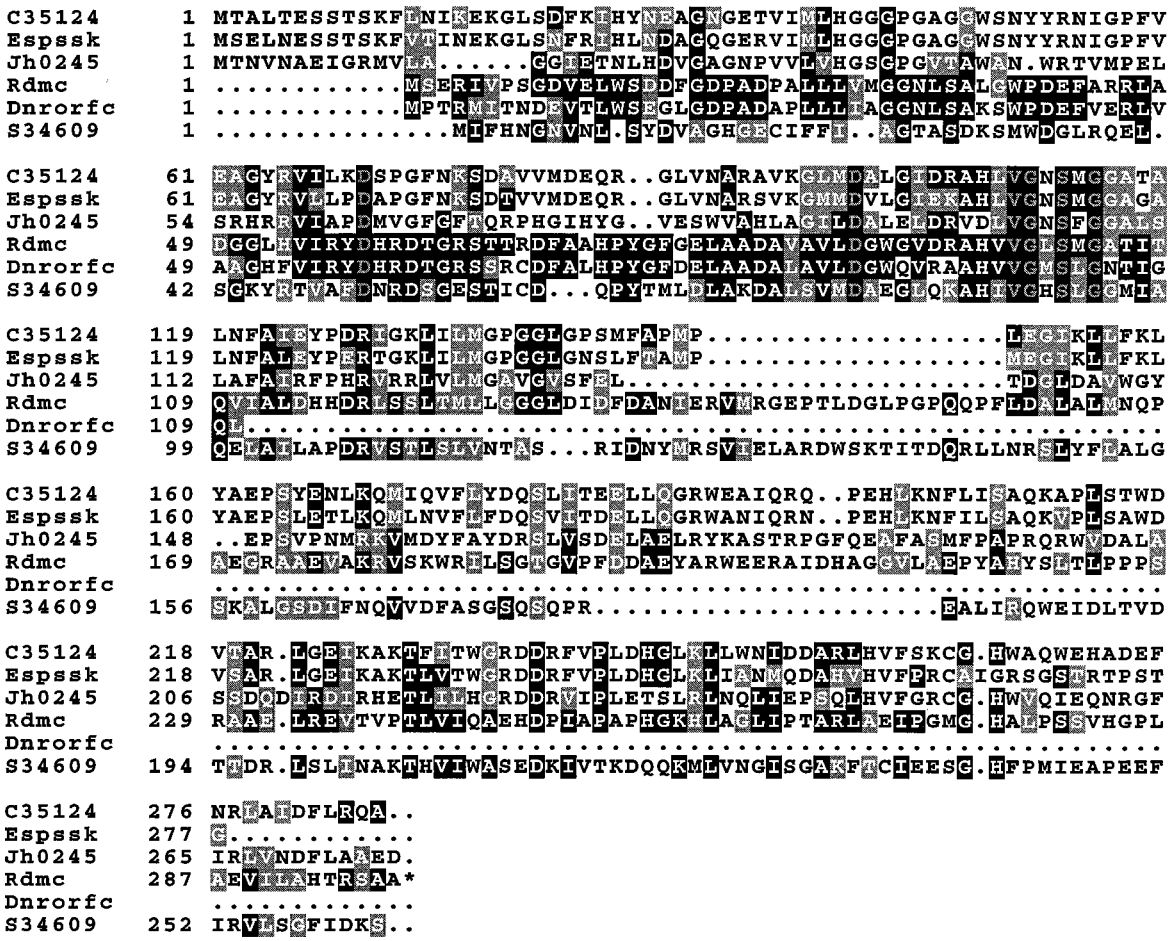

FIG. 3. Multiple sequence alignment (PILEUP [10]) of RdmC, the sequenced portion of the homologous ORF from *S. peucetius* (Dnrorfc), three *Pseudomonas* enoate hydrolases (C35124, Espssk, and Jh0245), and a *Pseudomonas* esterase (S34609). Amino acids that are similar (shaded background) or identical (black background) to those in RdmC have been marked with the BOXSHADE program (13). The sequences have been ordered according to pairwise similarity.

ated with translational coupling (38) and is often observed in *Streptomyces* operons (3, 7).

rdmD appears to be necessary for the production of 10-modified hybrids, but because the function of the homologous *srmX* gene is unknown, it is impossible to deduce its function from available data.

Both experimental data and sequence comparisons indicate that *rdmE* is the aklavinone-11-hydroxylase gene. However, two constructions in which *rdmE* was complete but other, upstream ORFs were truncated didn't produce 11-hydroxylated products.

Why were no 10-modified anthracyclines produced with constructions that lacked *rdmB*? The following possibilities may explain their absence.

(i) The enzymatic activity of RdmB is required. This is improbable, because no 4-O-methylated products have been observed and because another known *S*-adenosylmethionine-dependent methyl transferase in anthracycline biosynthesis, aklanonic acid methyltransferase (5), is already expected to be present in the *S. galilaeus* host.

(ii) The RdmB protein is required. If RdmB, RdmC, and RdmD form a multienzyme complex which is inactive in the absence of RdmB, the necessity of RdmB has been explained. The translational coupling of RdmB and RdmC also suggests that these two peptides may be subunits of a protein; in *S. peucetius*, however, the homologous ORFs do not appear to be translationally coupled (21).

(iii) Translation of RdmB is required for translation of the ORFs that follow it. Disruption of translational coupling between *actI*-ORF3 and *actVII* in *S. coelicolor* caused reduced expression of the downstream genes (19); presumably, lack of translation of *rdmB* could have a similar effect.

The observation (27) that in the opposite orientation, only 11-hydroxylation was observed suggests that *rdmE* is transcribed independently from the other ORFs. Unfortunately, we have not been able to characterize RNA transcripts produced from these constructions because of extensive degradation of RNAs isolated from antibiotic-producing cultures.

Our results demonstrate that anthracycline biosynthetic genes can be used to generate novel anthracyclines, and we are proceeding to further dissect the late steps in rhodomycin biosynthesis, which are apparently very relevant to understanding the biosynthesis of daunorubicin and doxorubicin also. Eventually, expression constructions of biosynthetic genes may allow the construction of artificial biosynthetic pathways in a building block-like fashion.

Nucleotide sequence accession number. The nucleotide sequence reported in this paper has been submitted to the GenBank data bank with accession number U10405.

We thank K. Ylihonko, J. Hakala, L. Cong, and I. Suominen for helpful discussions; H. Myllyharju for some early sequencing work; M. J. Bibb for kindly analyzing the preliminary sequence with FRAME; R. Losick and R. Nagata for information about *srmX* and AX110P, respectively; and D. A. Hopwood for advice and critical reading of the manuscript.

J.N. gratefully acknowledges grants from the Varsinais-Suomi Fund of the Finnish Cultural Foundation and the Jenny and Antti Wihuri Foundation.

REFERENCES

1. **Bibb, M. J., P. R. Findlay, and M. W. Johnson.** 1984. The relationship between base composition and codon usage in bacterial genes and its use for the simple and reliable identification of protein-coding sequences. Gene **30:**157–166.
2. **Bibb, M. J., G. R. Janssen, and J. M. Ward.** 1985. Cloning and analysis of the promoter region of the erythromycin resistance gene (*ermE*) of *Streptomyces erythraeus*. Gene **38:**215–226.
3. **Birch, A., A. Leiser, and J. A. Robinson.** 1993. Cloning, sequencing, and expression of the gene encoding methylmalonyl-coenzyme A mutase from *Streptomyces cinnamonensis*. J. Bacteriol. **175:**3511–3519.
4. **Brenner, S.** 1988. The molecular evolution of genes and proteins: a tale of two serines. Nature (London) **334:**528–530.
5. **Connors, N. C., P. L. Bartel, and W. R. Strohl.** 1990. Biosynthesis of anthracyclines: enzymic conversion of aklanonic acid to aklavinone and ε-rhodomycinone by anthracycline-producing streptomycetes. J. Gen. Microbiol. **136:**1887–1894.
6. **Eckardt, K., and C. Wagner.** 1988. Biosynthesis of anthracyclinones. J. Basic Microbiol. **28:**137–144.
7. **Fernández-Moreno, M. A., E. Martínez, L. Boto, D. A. Hopwood, and F. Malpartida.** 1992. Nucleotide sequence and deduced functions of a set of cotranscribed genes of *Streptomyces coelicolor* A3(2) including the polyketide synthase for the antibiotic actinorhodin. J. Biol. Chem. **267:**19278–19290.
8. **Frischauf, A.-M., H. Lehrach, A. Poustka, and N. Murray.** 1983. Lambda replacement vectors carrying polylinker sequences. J. Mol. Biol. **170:**827–842.
9. **Geistlich, M., R. Losick, J. R. Turner, and R. N. Rao.** 1992. Characterization of a novel regulatory gene governing the expression of a polyketide synthase gene in *Streptomyces ambofaciens*. Mol. Microbiol. **6:**2019–2029.
10. **Genetics Computer Group.** 1991. Program manual for the GCG package, version 7. Genetics Computer Group, Madison, Wis.
11. **Guilfoile, P. G., and C. R. Hutchinson.** 1991. A bacterial analog of the *mdr* gene of mammalian tumor cells is present in *Streptomyces peucetius*, the producer of daunorubicin and doxorubicin. Proc. Natl. Acad. Sci. USA **88:**8553–8557.
12. **Hayase, N., K. Taira, and K. Furukawa.** 1990. *Pseudomonas putida* KF715 *bphABCD* operon encoding biphenyl and polychlorinated biphenyl degradation: cloning, analysis, and expression in soil bacteria. J. Bacteriol. **172:**1160–1164.
13. **Hoffmann, K.** Unpublished computer program.
14. **Hopwood, D. A.** 1981. Future possibilities for the discovery of new antibiotics by genetic engineering, p. 585–598. *In* M. R. J. Salton and G. D. Shockman (ed.), Beta-lactam antibiotics. Academic Press, New York.
15. **Hopwood, D. A., M. J. Bibb, K. F. Chater, T. Kieser, C. J. Bruton, H. M. Kieser, D. J. Lydiate, C. P. Smith, J. M. Ward, and H. Schrempf.** 1985. Genetic manipulation of Streptomyces: a laboratory manual. The John Innes Foundation, Norwich, England.
16. **Hopwood, D. A., F. Malpartida, H. M. Kieser, H. Ikeda, J. Duncan, I. Fujii, B. A. M. Rudd, H. G. Floss, and S. Ōmura.** 1985. Production of "hybrid" antibiotics by genetic engineering. Nature (London) **314:**642–644.
17. **Hopwood, D. A., and D. H. Sherman.** 1990. Molecular genetics of polyketides and its comparison to fatty acid biosynthesis. Annu. Rev. Genet. **24:**37–66.
18. **Katz, L., and S. Donadio.** 1993. Polyketide synthesis: prospects for hybrid antibiotics. Annu. Rev. Microbiol. **47:**875–912.
19. **Khosla, C., S. Ebert-Khosla, and D. A. Hopwood.** 1992. Targeted gene replacements in a *Streptomyces* polyketide synthase gene cluster: role for the acyl carrier protein. Mol. Microbiol. **6:**3237–3249.
20. **Kimbara, K., T. Hashimoto, M. Fukuda, T. Koana, M. Takagi, M. Oishi, and K. Yano.** 1989. Cloning and sequencing of two tandem genes involved in degradation of 2,3-dihydroxybiphenyl to benzoic acid in the polychlorinated biphenyl-degrading soil bacterium *Pseudomonas* sp. strain KKS102. J. Bacteriol. **171:**2740–2747.
21. **Madduri, K., F. Torti, A. L. Colombo, and C. R. Hutchinson.** 1993. Cloning and sequencing of a gene encoding carminomycin 4-*O*-methyltransferase from *Streptomyces peucetius* and its expression in *Escherichia coli*. J. Bacteriol. **175:**3900–3904.
22. **Malpartida, F., S. E. Hallam, H. M. Kieser, H. Motamedi, C. R. Hutchinson, M. J. Butler, D. A. Sugden, M. Warren, C. McKillop, C. R. Bailey, G. O. Humphreys, and D. A. Hopwood.** 1987. Homology between Streptomyces genes coding for synthesis of different polyketides and its use to clone antibiotic biosynthetic genes. Nature (London) **325:**818–821.
23. **Menn, F.-M., G. J. Zylstra, and D. T. Gibson.** 1991. Location and sequence of the *todF* gene encoding 2-hydroxy-6-oxohepta-2,4-dienoate hydrolase in *Pseudomonas putida* F1. Gene **104:**91–94.
24. **Nagata, R., M. Wada, S. Satoh, H. Kamada, and H. Harada.** 1993. A cDNA differentially expressed between embryogenic and nonembryogenic carrot (*Daucus carota* L.) cells. Plant Physiol. (Bethesda) **103:**1021.
25. **Niemi, J., J. Hakala, A. Kopio, R. Pärssinen, K. Ylihonko, and K. Lampi.** March 1992. PCT patent application PCT/FI92/00084.
26. **Niemi, J., and P. Mäntsälä.** 1994. TRACTOR, a program to locate subclones in a nucleotide sequence using only one sequencing reaction. Comput. Appl. Biosci. **10:**563–566.
27. **Niemi, J., K. Ylihonko, J. Hakala, R. Pärssinen, A. Kopio, and P. Mäntsälä.** 1994. Hybrid anthracycline antibiotics: production of new anthracyclines by cloned genes from *Streptomyces purpurascens* in *Streptomyces galilaeus*. Microbiology **140:**1351–1358.
28. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
29. **Shimada, Y., T. Nagao, A. Sugihara, T. Iizumi, T. Yui, K. Nakamura, T. Fukase, and Y. Tominaga.** 1993. Cloning and analysis of an esterase gene from *Pseudomonas* sp. KWI-56. Biochim. Biophys. Acta **1174:**79–82.
30. **Strohl, W. R.** 1992. Compilation and analysis of DNA sequences associated with apparent streptomycete promoters. Nucleic Acids Res. **20:**961–974.
31. **Strohl, W. R., P. L. Bartel, N. C. Connors, C.-B. Zhu, D. C. Dosch, J. M. Beale, Jr., H. G. Floss, K. Stutzman-Engwall, S. L. Otten, and C. R. Hutchinson.** 1989. Biosynthesis of natural and hybrid polyketides by anthracycline-producing streptomycetes, p. 68–84. *In* C. L. Hershberger, S. W. Queener, and G. Hegeman (ed.), Genetics and molecular biology of industrial microorganisms. American Society for Microbiology, Washington, D.C.
32. **Summers, R. G., E. Wendt-Pienkowski, H. Motamedi, and C. R. Hutchinson.** 1992. Nucleotide sequence of the *tcmII-tcmIV* region of the tetracenomycin C biosynthetic gene cluster of *Streptomyces glaucescens* and evidence that the *tcmN* gene encodes a multifunctional cyclase–dehydratase–*O*-methyl transferase. J. Bacteriol. **174:**1810–1820.
33. **Tanaka, H., T. Yoshika, Y. Shimauchi, Y. Matsuzawa, T. Oki, and T. Inui.** 1980. Chemical modification of anthracycline antibiotics. I. Demethoxycarbonylation, 10-epimerization and 4-O-methylation of aclacinomycin A. J. Antibiot. **33:**1323–1330.
34. **Ward, J. M., G. R. Janssen, T. Kieser, M. J. Bibb, M. J. Buttner, and M. J. Bibb.** 1986. Construction and characterisation of a series of multi-copy promoter-probe plasmid vectors for Streptomyces using the aminoglycoside phosphotransferase gene from Tn5 as indicator. Mol. Gen. Genet. **203:**468–478.
35. **Wiley, P. F., D. W. Elrod, D. H. Houser, J. L. Johnson, L. M. Pschigoda, W. C. Krueger, and A. Moscowitz.** 1979. Nogalamycin. Stereochemistry and chemical modification. J. Org. Chem. **44:**4030–4038.
36. **Wright, F., and M. J. Bibb.** 1992. Codon usage in the G+C-rich *Streptomyces* genome. Gene **113:**55–65.
37. **Yanisch-Perron, C., J. Vieira, and J. Messing.** 1985. Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. Gene **33:**103–119.
38. **Zalkin, H., and D. J. Ebbole.** 1988. Organization and regulation of genes encoding biosynthetic enzymes in *Bacillus subtilis*. J. Biol. Chem. **263:**1595–1598.