Mol Gen Genet (1991) 230:401–412

MGG

Organisation and functions of the *act*VA region of the actinorhodin biosynthetic gene cluster of *Streptomyces coelicolor*

José L. Caballero[1], Eduardo Martinez[2], Francisco Malpartida[2], and David A. Hopwood[1]

[1] John Innes Institute, John Innes Centre, Norwich NR4 7UH, UK
[2] Centro Nacional de Biotecnologia, Serrano 115, E-28006, Madrid, Spain

Received May 8, 1991 / July 26, 1991

Summary. Sequence analysis of the *act*VA region of the actinorhodin biosynthetic gene cluster of *Streptomyces coelicolor* revealed a succession of six open reading frames (ORFs), all running in the same direction and extending over 5.32 kb. The protein product of *act*VA-ORF1 strongly resembles that of another gene, elsewhere in the *act* cluster (*act*II-ORF2), which codes for a trans-membrane protein previously implicated in actinorhodin export from the mycelium. This suggests that the two gene products may co-operate in actinorhodin export, perhaps being sufficient for self-protection of the organism against suicide. At least four of the other five ORFs are implicated in the control of the C-6 and C-8 ring-hydroxylation reactions, lacking in *act*VA mutants, that occur at middle to late stages in the actinorhodin biosynthetic pathway. This conclusion was reached by genetic mapping of *act*VA mutants to *act*VA-ORF3 and -ORF5 (and perhaps -ORF4), and by the finding of strong resemblances between the protein products of *act*-VA-ORF2 and -ORF6 and the products of genes of the oxytetracycline or tetracenomycin gene clusters that have been implicated in ring-hydroxylation reactions in the biosynthesis of these other aromatic polyketide antibiotics.

Key words: Antibiotic biosynthesis – Antibiotic resistance – Hydroxylases – *Streptomyces*

Introduction

Two of the most characteristic properties of the actinomycetes are their production of a huge range of secondary metabolites, including many antibiotics, and their manifestation of an unusually complex cycle of prokaryotic morphological differentiation. Recent genetic studies (Chater 1990; Seno and Baltz 1989; Martín and Liras 1989) have begun to elucidate the control of these two properties, and have revealed intimate regulatory interconnections between them. Much of this work has utilized *Streptomyces coelicolor* A3(2), which produces at least five secondary metabolites, and has a comparatively long history of in vivo and in vitro genetic analysis (Chater 1990; Hopwood 1988).

Among the secondary metabolites produced by *S. coelicolor* A3(2), the blue-pigmented polyketide actinorhodin has received most attention as a model for antibiotic production and export, and the entire cluster of biosynthetic genes has been cloned on a continuous segment of chromosomal DNA (Malpartida and Hopwood 1984). Recent sequence analysis and gene disruption studies have characterised the central and (conventionally) right-hand regions of the *act* cluster. The central 5.7 kb segment, corresponding to the *act*II class of mutants (Rudd and Hopwood 1979), contains two regulatory genes: a transcriptional activator for the *act* structural genes, and a repressor for the adjacent transcript that carries two genes involved in export of actinorhodin from the mycelium (Fernandez-Moreno et al. 1991; Caballero et al. 1991). To the right of the *act*II region is a series of genes, extending over at least 6 kb and corresponding to several mutant classes, which code for proteins of the polyketide synthase complex (the *act*I, III, VII products), and at least two enzymes (the products of *act*IV and VB) for later steps in the biosynthetic pathway (Hallam et al. 1988; M.A. Fernandez-Moreno et al., in preparation).

The organization and functions of the remaining part of the *act* cluster, to the left of the *act*II region, are less well understood. The region has been divided into two roughly equal sections, corresponding to the approximate locations of two classes of *act* gene mutations, *act*VA and *act*VI (Malpartida and Hopwood 1986). Here we report sequence analysis and mutant mapping studies in the *act*VA region. An analysis of the *act*VI region will be presented elsewhere (E. Martinez et al., in preparation).

Offprint requests to: D.A. Hopwood

Materials and methods

Bacterial strains, plasmid and cloning vectors. The *Escherichia coli* strain was JM101 (Yanisch-Perron et al. 1985). The *S. coelicolor actVA* mutants were derived from strain 1190 (*hisA1*, *uraA1*, *strA1* SCP1$^-$ SCP2$^+$) as previously described (Rudd and Hopwood 1979). *S. lividans* strain 1326 was used as recipient for the construction of recombinant ϕC31 phages (Hopwood et al. 1985a). pIJ2323 – pBR325 carrying *act* DNA between sites 2 and 11 of the restriction map (Malpartida and Hopwood 1986) – was used as the source of cloned *act* DNA. The *Streptomyces* phages used for the complementation analysis were derivatives of the *att*$^-$ *c*$^+$ ϕC31 phages KC857 and KC861 (Bruton et al. 1991).

Media, culture conditions and microbiological procedures. For *Streptomyces*, agar media, transformation and transfection were as in Hopwood et al. (1985a). For vector selection, thiostrepton (a gift of S.J. Lucania, Squibb Institute, Princeton, USA) was used at 50 μg/ml in agar and 10 μg/ml in broth cultures. *E. coli* was grown on L agar or in L broth with or without carbenicillin at 100 μg/ml (Maniatis et al. 1982).

DNA sequencing. DNA sequencing was by the dideoxy chain-termination method (Sanger et al. 1977), using the 7-deaza-dGTP reagent kit from US Biochemicals according to the manufacturer's recommendations. Convenient DNA fragments were previously cloned in Bluescript vectors (Stratagene) from suitable restriction fragments, or generated by ExoIII digestion (Henikoff 1984). Every part of the sequence was determined at least once in both directions, with every sequence junction for one strand being overlapped by a sequence for the other.

Computer analysis of sequences. The DNA sequence was analysed for open reading frames using the FRAME Program (Bibb et al. 1984). Derived amino acid sequences were analysed using programs from the UWCG package (version 6.2, 1990: Devereux et al. 1984): sequences were compared to the EMBL Nucleic Acid database (release 24.0, August 1990) and the PIR database (release 26.0, September 1990) using BESTFIT, COMPARE and DOTPLOT. Hydropathy plots (Fig. 5) used the program HYDROPUB, kindly provided by Dr. M. Boursnell (Poultry Virus Research Station, Houghton, UK) which scans amino acid sequences and calculates a mean hydropathy value for a given window, based on the amino acid hydropathy values of Kyte and Doolittle (1982).

Construction of recombinant phages, and complementation analysis. KC857 and KC861 derivatives carrying two different DNA fragments covering part of the *act*VA region (Fig. 1) were constructed. A 2.71 kb *Xho*I fragment from pIJ2323 was cloned into the *Xho*I site of KC861 to produce KC592 and a 1.41 kb *Aat*II-*Xho*I fragment was cloned to replace the *Xho*I-*Sca*I fragment of KC857 to produce KC591. The resulting recombinant phages carried open reading frames (ORFs) 4, 5 and 6 and the 3′ end of ORF3, or just ORF5 and ORF6, respectively (Fig. 7). Each of the phages retained the *tsr* gene to allow selection for thiostrepton-resistant recombinants. The phages were used to transfect *S. lividans* 1326 protoplasts and positive plaques were detected by plaque hybridisation (Hopwood et al. 1985a) using a fragment from *Bgl*II site 6 to a *Sph*I site to the right of the *act*VA region (Fig. 1) as a radioactive probe. Positive plaques were purified and a high titre lysate was obtained for each construct to infect a set of *act*VA *S. coelicolor* mutants. Spores from the resulting lysogens were harvested and plated on R2 medium containing thiostrepton and the proportions of blue lysogens were scored after 5–6 days of growth.

Results

Sequence and overall organization of the act*VA region*

Figure 1 shows a restriction map of the entire *act* region (Malpartida and Hopwood 1986). The sequence of the segment of DNA from *Bam*HI site 4 to a point overlapping the sequence of the *act*II region previously reported (Fernandez-Moreno et al. 1991) was determined (Fig. 2: deposited as EMBL Nucleic Acid Database number X58833). Analysis of the relevant part of this sequence for open reading frames (ORFs) by the FRAME program revealed six complete ORFs, all in the same orientation, as shown in Fig. 3 and summarised in Fig. 1.

The most likely start codons for the ORFs – named *act*VA-ORF1–6 – were assigned by choosing the most upstream ATG or GTG, which in most cases was preceded by a good potential ribosome binding site. In two cases there is potential translational coupling of pairs of ORFs, with overlaps of the extreme 3′ end of one ORF with the 5′ end of the next ORF downstream: this applies at the ORF4/5 (ATGA) and ORF5/6 (*ATGGCTGA*) boundaries. There is only one nucleotide between the stop codon of ORF1 and the start codon of ORF2; between ORF2 and ORF3 there are 94 nucleotides; and between ORF3 and ORF4 there are 24 nucleotides. Interestingly, the coding region of *act*VA-ORF6 would overlap for its last 11 codons with the last 11 codons of *act*II-ORF1 (the repressor gene), which converges with the *act*VA region (Fernandez-Moreno et al. 1991; Caballero et al. 1991).

The product of act*VA-ORF1 appears to be a second trans-membrane protein of the* act *cluster*

The deduced protein product of *act*VA-ORF1 shows a strong resemblance to the product of a gene in the *act*II region, *act*II-ORF2 (identity 30%, similarity 57% revealed by BESTFIT), as shown in Fig. 4a. The *act*II-ORF2 product was shown to resemble proteins that confer resistance to tetracycline on gram-negative and gram-positive bacteria, to aminotriazole on *Saccharomyces cerevisiae*, and to methylenomycin on *S. coelicolor*

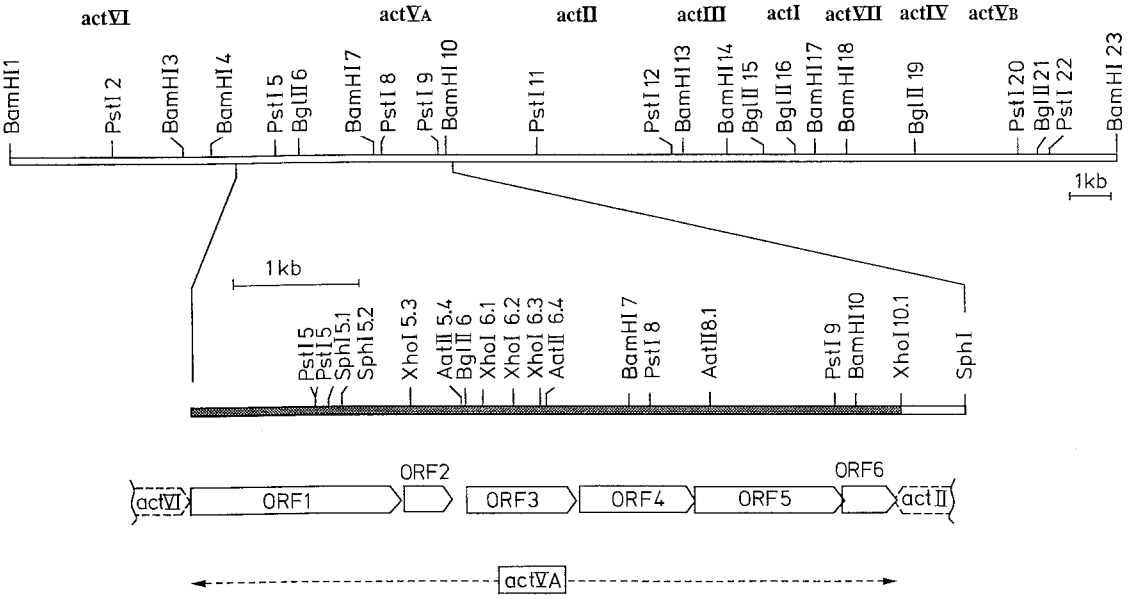


Fig. 1. Restriction map of the *act* cluster with approximate locations of *act* mutant classes I–VII, and organization of the six *act*VA open reading frames (ORFs) revealed and analysed in this paper. Integral restriction site numbers are those of Malpartida and Hopwood (1986) and decimal numbers were assigned in the present work

(Fernandez-Moreno et al. 1991). Fig. 4b shows the strong resemblance between the products of *act*V-ORF1 and the methylenomycin resistance gene *mmr*. The tetracycline (Klock et al. 1985) and aminotriazole (Kanazawa et al. 1988) resistance proteins (and presumably the methylenomycin resistance protein: Neal and Chater 1987) are trans-membrane proteins; they show a characteristic pattern of hydrophobic domains separated by less hydrophobic regions. The same is true of the *act*II-ORF2 product, which is postulated to act, together with the coupled *act*II-ORF3 product, to export endogenously produced actinorhodin from the mycelium (Fernandez-Moreno et al. 1991) and for the *act*VA-ORF1 product (Fig. 5a). Figure 5b compares the hydrophobicity profiles of the *act*VA-ORF1 and *act*II-ORF2 products. The two proteins have a remarkably similar arrangement of predicted trans-membrane domains: the only obvious difference is in the C-terminal regions. Both proteins have a 13th hydrophobic domain at the extreme C-terminus, but this is separated from the set of 12 hydrophobic domains of ActII-ORF2, but not of ActVA-ORF1, by a hydrophilic domain. Interestingly, comparison of the nucleotide sequences of the two genes by DOTPLOT showed no significant similarity, except for a segment of about 80 bp near the 3′ ends of the genes with an identity revealed by BESTFIT of 70%.

The products of act*VA-ORF2, -ORF3 and -ORF6 resemble those of biosynthetic genes for other aromatic polyketide antibiotics*

Comparisons of the amino acid sequences of the remaining *act*VA ORFs with the sequences of known proteins or the translation products of reported DNA sequences indicated resemblances of three of the *act*VA products each with a different gene product from another pathway of aromatic polyketide antibiotic biosynthesis.

*act*VA-ORF2. As shown in Fig. 6a, the deduced product of *act*VA-ORF2 shows a strong end-to-end alignment with the product of a recently sequenced ORF, *otc*-ORFX1, in the cluster of genes for oxytetracycline biosynthesis in *S. rimosus* (McDowall 1991), with an identity of 49% and a similarity of 75%.

*act*VA-ORF3. The product of *act*VA-ORF3 strongly resembles, over its C-terminal half, the available sequence (again for the C-terminal portion) of an incompletely sequenced ORF (*gra*-ORF7) from the granaticin biosynthetic gene cluster of *S. violaceoruber* (Fig. 6b). The sequence of *gra*-ORF7 was determined by D.H. Sherman (unpublished: EMBL Nucleic Acid Database number X16144); it lies to the right of the set of six genes coding for components of the granaticin polyketide synthase complex and converging with the right-hand members of this set of genes (Sherman et al. 1989) but its function is unknown.

*act*VA-ORF6. The *act*VA-ORF6 product aligns strongly with the sequence (H. Motamedi and C.R. Hutchinson, personal communication) of the product of one of the genes (*tcm*VI-ORF1) in the tetracenomycin biosynthetic gene cluster of *S. glaucescens* (Fig. 6c), with identity 39% and similarity 55%.

The act*VA-ORF4 product*

The product of *act*VA-ORF4 was found to show a resemblance over its N-terminal one-third to a sequence near the N-terminus of the dihydroflavonol-4-reductases

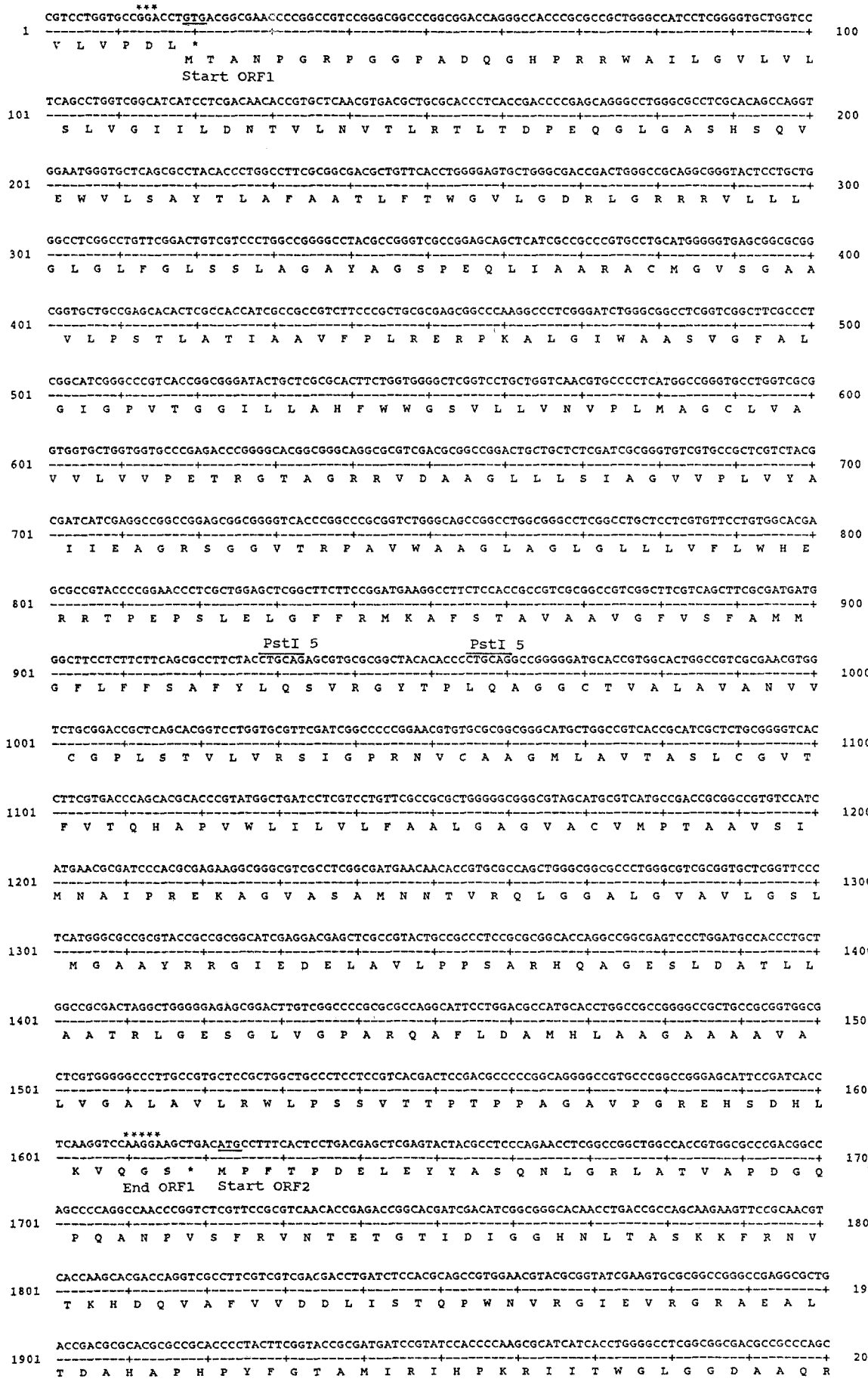

***
CGTCCTGGTGCCGGACCTGTGACGGCGAACCCCGGCCGTCCGGGCGGCCCGGCGGACCAGGGCCACCCGCGCCGCTGGGCCATCCTCGGGGTGCTGGTCC
V L V P D L *
M T A N P G R P G G P A D Q G H P R R W A I L G V L V L
Start ORF1
TCAGCCTGGTCGGCATCATCCTCGACAACACCGTGCTCAACGTGACGCTGCGCACCCTCACCGACCCCGAGCAGGGCCTGGGCGCCTCGCACAGCCAGGT
S L V G I I L D N T V L N V T L R T L T D P E Q G L G A S H S Q V
GGAATGGGTGCTCAGCGCCTACACCCTGGCCTTCGCGGCGACGCTGTTCACCTGGGGAGTGCTGGGCGACCGACTGGGCCGCAGGCGGGTACTCCTGCTG
E W V L S A Y T L A F A A T L F T W G V L G D R L G R R R V L L L
GGCCTCGGCCTGTTCGGACTGTCGTCCCTGGCCGGGGCCTACGCCGGGTCGCCGGAGCAGCTCATCGCCGCCCGTGCCTGCATGGGGGTGAGCGGCGCGG
G L G L F G L S S L A G A Y A G S P E Q L I A A R A C M G V S G A A
CGGTGCTGCCGAGCACACTCGCCACCATCGCCGCCGTCTTCCCGCTGCGCGAGCGGCCCAAGGCCCTCGGGATCTGGGCGGCCTCGGTCGGCTTCGCCCT
V L P S T L A T I A A V F P L R E R P K A L G I W A A S V G F A L
CGGCATCGGGCCCGTCACCGGCGGGATACTGCTCGCGCACTTCTGGTGGGGCTCGGTCCTGCTGGTCAACGTGCCCCTCATGGCCGGGTGCCTGGTCGCG
G I G P V T G G I L L A H F W W G S V L L V N V P L M A G C L V A
GTGGTGCTGGTGGTGCCCGAGACCCGGGGCACGGCGGGCAGGCGCGTCGACGCGGCCGGACTGCTGCTCTCGATCGCGGGTGTCGTGCCGCTCGTCTACG
V V L V V P E T R G T A G R R V D A A G L L L S I A G V V P L V Y A
CGATCATCGAGGCCGGCCGGAGCGGCGGGGTCACCCGGCCCGCGGTCTGGGCAGCCGGCCTGGCGGGCCTCGGCCTGCTCCTCGTGTTCCTGTGGCACGA
I I E A G R S G G V T R P A V W A A G L A G L G L L L V F L W H E
GCGCCGTACCCCGGAACCCTCGCTGGAGCTCGGCTTCTTCCGGATGAAGGCCTTCTCCACCGCCGTCGCGGCCGTCGGCTTCGTCAGCTTCGCGATGATG
R R T P E P S L E L G F F R M K A F S T A V A A V G F V S F A M M
PstI 5
PstI 5
GGCTTCCTCTTCTTCAGCGCCTTCTACCTGCAGAGCGTGCGCGGCTACACACCCCTGCAGGCCGGGGGATGCACCGTGGCACTGGCCGTCGCGAACGTGG
G F L F F S A F Y L Q S V R G Y T P L Q A G G C T V A L A V A N V V
TCTGCGGACCGCTCAGCACGGTCCTGGTGCGTTCGATCGGCCCCCGGAACGTGTGCGCGGCGGGCATGCTGGCCGTCACCGCATCGCTCTGCGGGGTCAC
C G P L S T V L V R S I G P R N V C A A G M L A V T A S L C G V T
CTTCGTGACCCAGCACGCACCCGTATGGCTGATCCTCGTCCTGTTCGCCGCGCTGGGGGCGGGCGTAGCATGCGTCATGCCGACCGCGGCCGTGTCCATC
F V T Q H A P V W L I L V L F A A L G A G V A C V M P T A A V S I
ATGAACGCGATCCCACGCGAGAAGGCGGGCGTCGCCTCGGCGATGAACAACACCGTGCGCCAGCTGGGCGGCGCCCTGGGCGTCGCGGTGCTCGGTTCCC
M N A I P R E K A G V A S A M N N T V R Q L G G A L G V A V L G S L
TCATGGGCGCCGCGTACCGCCGCGGCATCGAGGACGAGCTCGCCGTACTGCCGCCCTCCGCGCGGCACCAGGCCGGCGAGTCCCTGGATGCCACCCTGCT
M G A A Y R R G I E D E L A V L P P S A R H Q A G E S L D A T L L
GGCCGCGACTAGGCTGGGGGAGAGCGGACTTGTCGGCCCCGCGCGCCAGGCATTCCTGGACGCCATGCACCTGGCCGCCGGGGCCGCTGCCGCGGTGGCG
A A T R L G E S G L V G P A R Q A F L D A M H L A A G A A A A A V A
CTCGTGGGGGCCCTTGCCGTGCTCCGCTGGCTGCCCTCCTCCGTCACGACTCCGACGCCCCCGGCAGGGGCCGTGCCCGGCCGGGAGCATTCCGATCACC
L V G A L A V L R W L P S S V T T P T P P A G A V P G R E H S D H L
*****
TCAAGGTCCAAGGAAGCTGACATGCCTTTCACTCCTGACGAGCTCGAGTACTACGCCTCCCAGAACCTCGGCCGGCTGGCCACCGTGGCGCCCGACGGCC
K V Q G S * M P F T P D E L E Y Y A S Q N L G R L A T V A P D G Q
End ORF1 Start ORF2
AGCCCCAGGCCAACCCGGTCTCGTTCCGCGTCAACACCGAGACCGGCACGATCGACATCGGCGGGCACAACCTGACCGCCAGCAAGAAGTTCCGCAACGT
P Q A N P V S F R V N T E T G T I D I G G H N L T A S K K F R N V
CACCAAGCACGACCAGGTCGCCTTCGTCGTCGACGACCTGATCTCCACGCAGCCGTGGAACGTACGCGGTATCGAAGTGCGCGGCCGGGCCGAGGCGCTG
T K H D Q V A F V V D D L I S T Q P W N V R G I E V R G R A E A L
ACCGACGCGCACGCGCCGCACCCCTACTTCGGTACCGCGATGATCCGTATCCACCCCAAGCGCATCATCACCTGGGGCCTCGGCGGCGACGCCGCCCAGC
T D A H A P H P Y F G T A M I R I H P K R I I T W G L G G D A A Q R

Fig. 2.

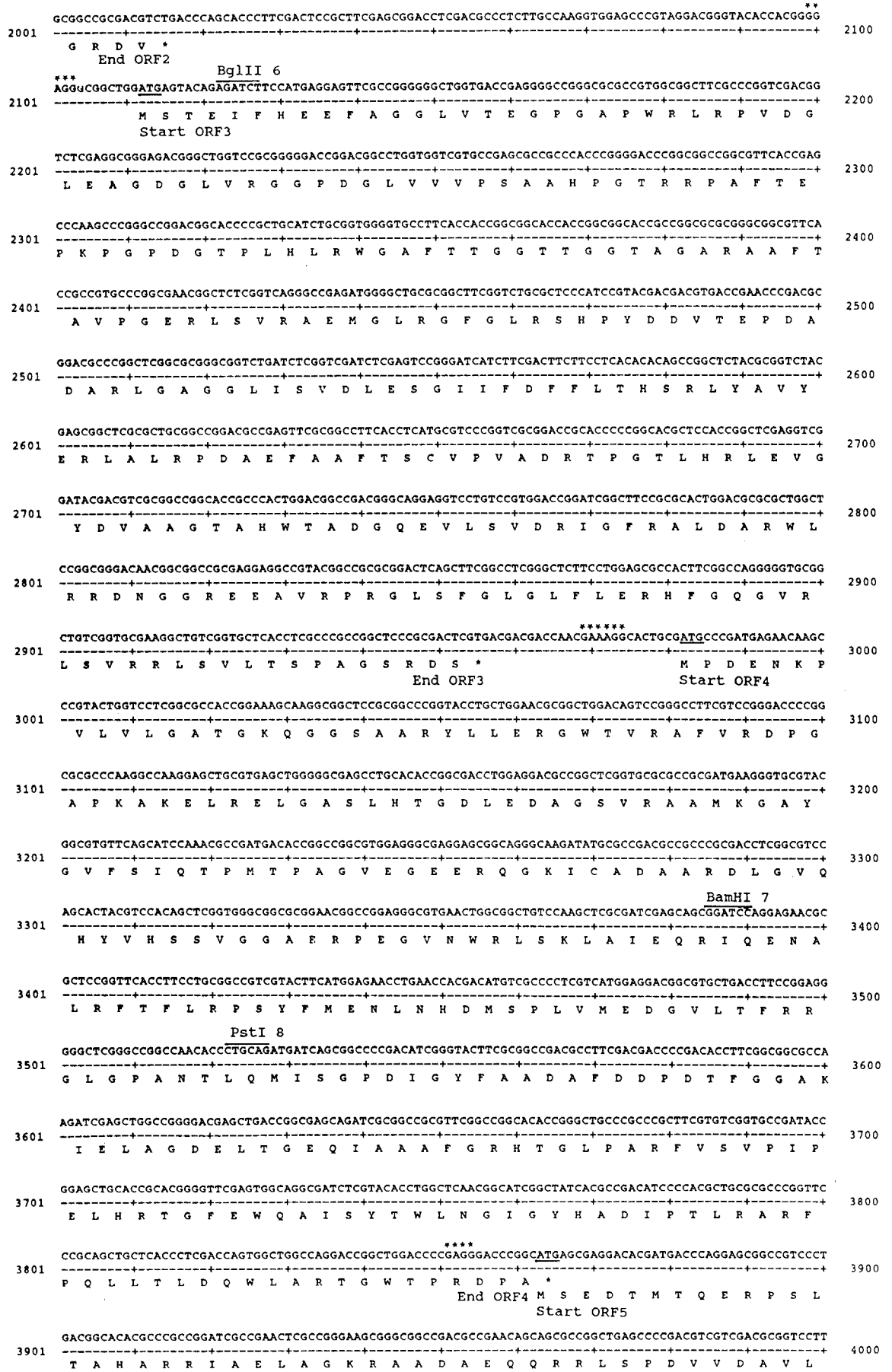

End ORF2
BglII 6
Start ORF3
End ORF3
Start ORF4
BamHI 7
PstI 8
End ORF4
Start ORF5

Fig. 2.

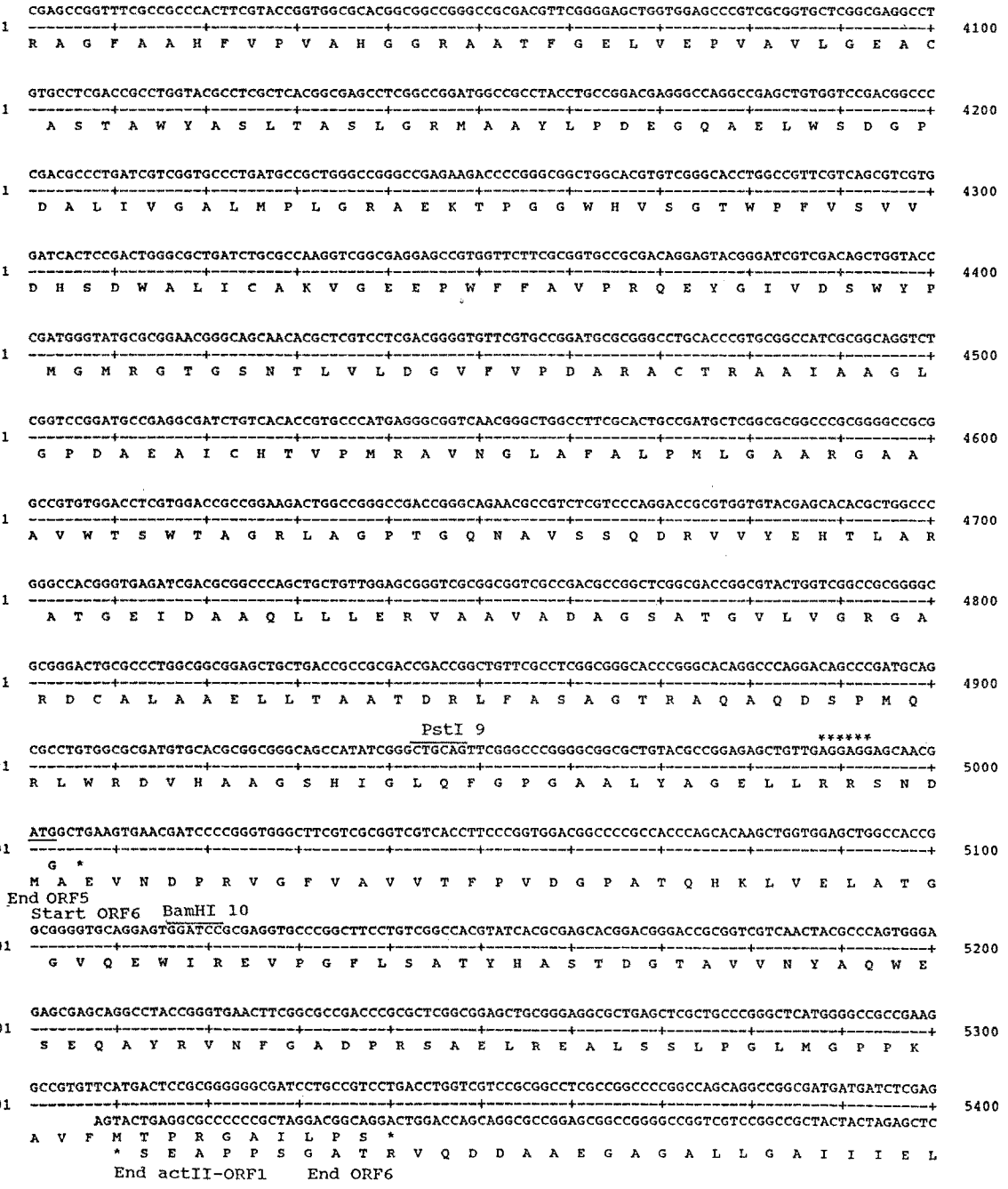


Fig. 2. Nucleotide sequence of the *act*VA region and deduced amino acid sequences (single letter code) of the six *act*VA ORFs. Key restriction sites are shown numbered as in Fig. 1. (*Pst*I site 5 turned out to be two very close *Pst*I sites.) Putative ribosomal binding sites are marked by *asterisks*. ATG or GTG start codons are *underlined*

of plants (Beld et al. 1989). The significance of this partial similarity remains to be explained.

Mapping of the act*VA mutants to various ORFs*

In order to assign the previously isolated mutations of class *act*VA (Rudd and Hopwood 1979) to a specific ORF or ORFs in the sequence, recombinant phages were constructed which carried two different segments of the *act*VA DNA. It was already known that several *act*V mutations must lie between *Bgl*II site 6 and *Bam*HI site 10 because two plasmids with overlapping inserts of *act* DNA – pIJ2315 carrying DNA between sites 3 and 10 and pIJ2305 carrying DNA between sites 6 and 15 (Fig. 1) – caused blue pigment to be produced when introduced into *act-101*, *act-109*, *act-163* and *act-218* (Malpartida and Hopwood 1986). (At least one other *act*V mutant, *act-235*, was not complemented by these two plasmids and was later found to be phenotypically different from the *act-101* type strain in accumulating a different metabolite; Cole et al. 1987). Mutants resem-

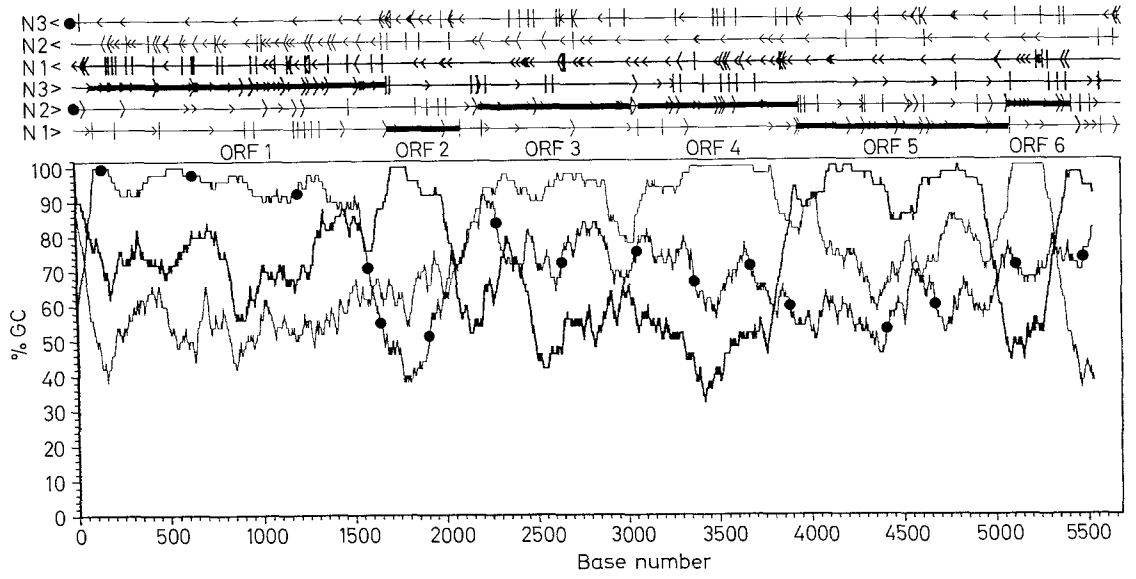


Fig. 3. FRAME analysis of the *act*VA sequence showing (below) the base composition in each of the three triplet positions over a 150 bp window and (above) the limits and direction of *act*VA ORFs 1–6 as *heavy lines* (*large arrowheads* are ATG codons, *small arrowheads* are GTG codons and *vertical lines* are stop codons). *Solid circles* serve to distinguish the line of intermediate density. The lengths of the ORFs (amino acid residues) are: ORF1, 533; ORF2, 130; ORF3, 281; ORF4, 294; ORF5, 381; ORF6, 112

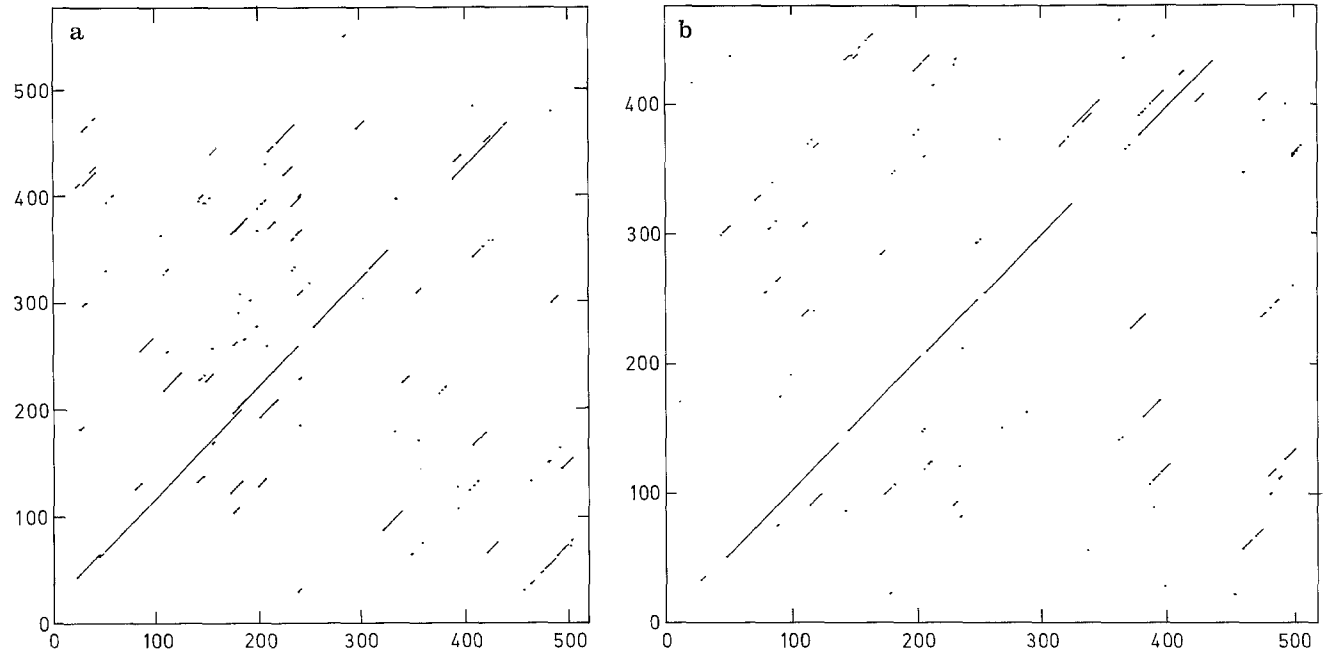


Fig. 4a and b. DOTPLOT comparisons of ActVA-ORF1 with two other proteins using a window of 40 and a stringency of 20. **a** ActVA-ORF1 (horizontal axis) vs ActII-ORF2 (Fernandez-Moreno et al. 1991); **b** ActVA-ORF1 (horizontal axis) vs Mmr (Neal and Chater 1987)

bling *act-101* were therefore re-classified as *act*VA and *act-235* represented a new class, called *act*VB, mapping near the right hand end of the *act* cluster: F. Malpartida, unpublished). All transformants of each of the four *act*-VA mutants with pIJ2315 produced abundant blue pigment, indicative of true *in trans* complementation. Therefore, since pIJ2315 is now seen to lack the C-terminal region of ORF6 (Fig. 7), we can deduce that the four *act*VA mutations cannot lie in this ORF. Transformants of *act-101*, *-109* and *-163* by pIJ2305 were also fully complemented, ruling out a location for these mutations in ORF3, because pIJ2305 is seen to lack the N-terminus of this ORF (or in ORFs 1 or 2, which are absent from the plasmid). Thus, these three *act*VA mutations were candidates for either ORF4 or ORF5. In contrast, pIJ2305 produced only 5–10% of blue transformants with *act-218* (Malpartida and Hopwood 1986), suggesting that this mutation might lie in ORF3, with

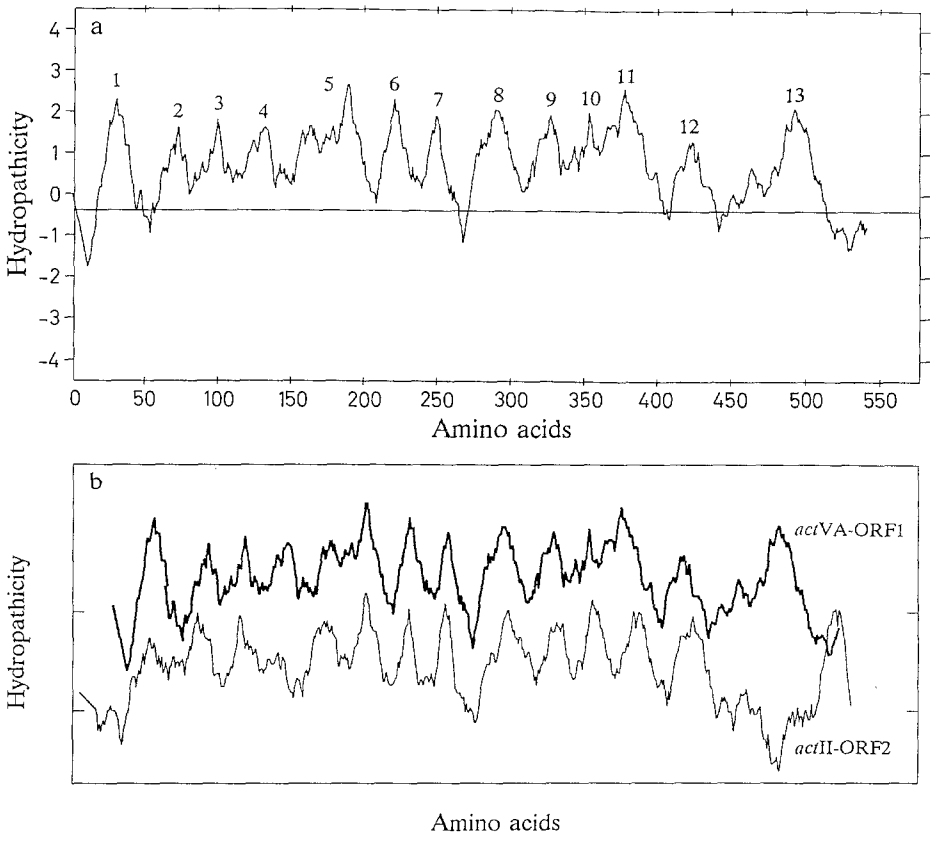


Fig. 5. a Hydropathy plot (HYDROPUB) of ActVA-ORF1, using a window of 19 amino acids. The horizontal line is at −0.4, the average overall hydropathy for a collection of soluble proteins (Kyte and Doolittle 1982). The numbers identify hydrophobic peaks above a threshold value of 1.6. **b** Comparison of hydropathy of ActVA-ORF1 with ActII-ORF2 (Fernandez-Moreno et al. 1991). The marks on the vertical axis indicate the −0.4 point for each of the plots

a requirement for crossing-over to generate *act*$^+$ recombinants.

In the light of these earlier results we constructed two *att*$^-$ *c*$^+$ recombinant phages, KC591 and KC592, carrying the DNA segments shown in Fig. 7. When used to lysogenise the four *act*VA mutants studied by Malpartida and Hopwood (1986), as well as others of the set isolated by Rudd and Hopwood (1979), by homologous recombination between the cloned fragment and the host *act* region, a proportion of blue colonies was observed in some phage/mutant combinations, but not in others (right side of Fig. 7). The mutations fall into three clear groups.

The first group, containing *act-218*, gave no blue colonies with either phage, ruling out a location in ORF4, 5 or 6. The production of a proportion of blue colonies when transformed with pIJ2305 (Malpartida and Hopwood 1986) confirmed its location in ORF3, rather than in some other part of the genome.

Two mutations, *act-156* and *-278*, gave blue colonies with KC592 but none (in the case of *act-278* out of more than 5500 lysogens) with KC591. Although a location at the extreme left-hand end of ORF5 is not excluded, one or both of these mutations are candidates for ORF4.

The remaining 14 mutations gave a proportion of blue colonies with both phages, indicative of a location in ORF5 or ORF6, and presumably lying progressively further towards the right of the cloned DNA fragments as the proportion of blue colonies increases (they are listed in this order in Fig. 7). The rightmost mutation (*act-249*) was transformed with pIJ2315: nearly all the transformants were blue, ruling out a location in ORF6. Thus, all of this group of mutations are likely to be in ORF5.

Discussion

Sequence analysis of the *act*VA region has revealed an unexpectedly large number of genes. Of these, one (*act*-VA-ORF1) is likely to be involved in the mechanism of self-protection of the organism against actinorhodin, while at least four of the other five would be structural genes for middle to late steps in the biosynthetic pathway. These two aspects are discussed in turn.

act*VA-ORF1*

The finding that cloning of the entire *act* cluster into the actinorhodin-sensitive host *S. parvulus* led to production of the antibiotic without suicide of the culture led to the conclusion that the cluster includes information for a self-protective mechanism (Malpartida and Hopwood 1984). When sequencing of the *act*II region revealed a gene (*act*II-ORF2) whose product strongly resembled, in its sequence and its hydrophobicity profile, genes for resistance by export of exogenously applied tetracycline in bacteria or aminotriazole in yeast (Fernandez-Moreno et al. 1991; Klock et al. 1985; Kanazawa et al. 1988), this gene was at first thought to be

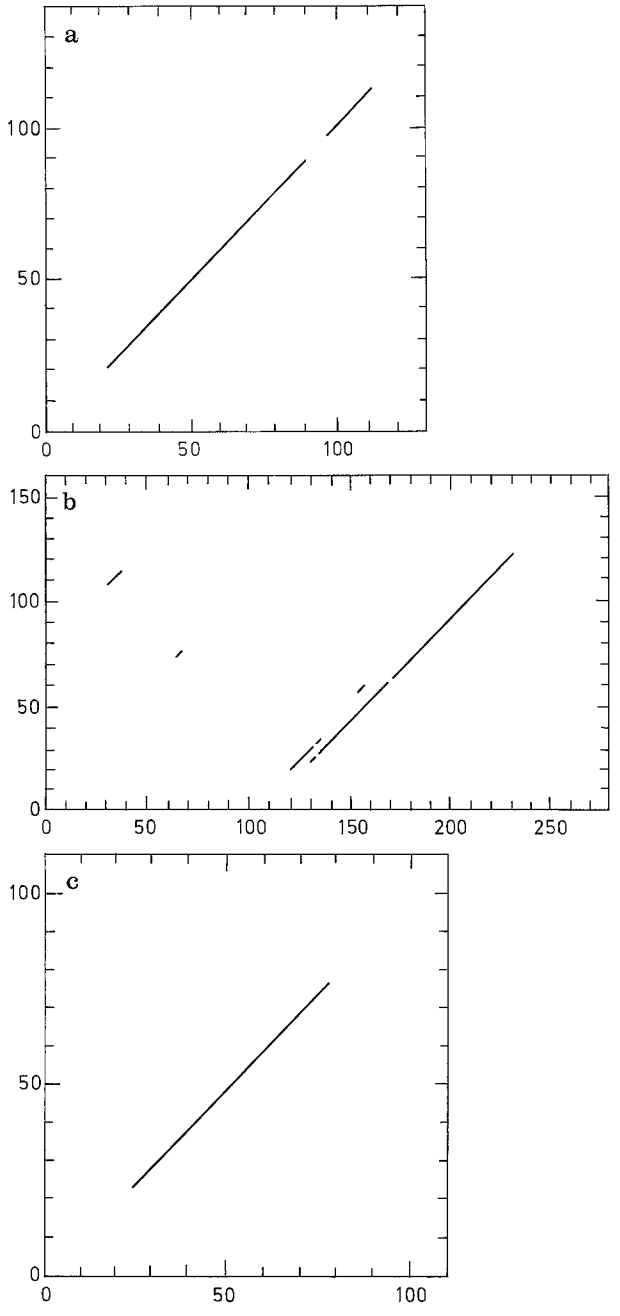
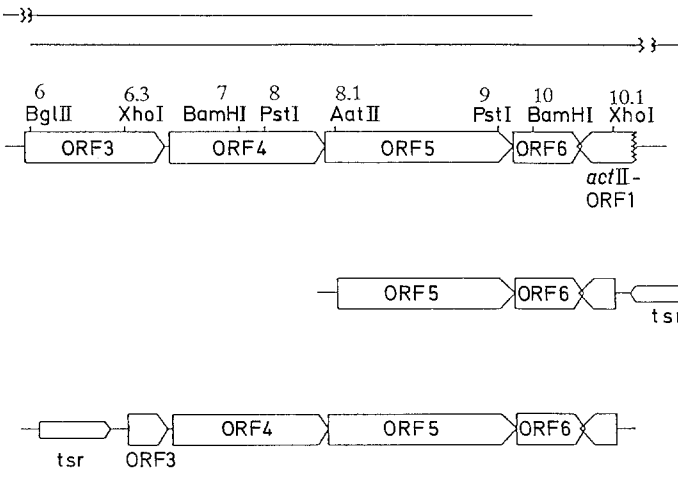


Fig. 6a–c. DOTPLOT comparisons of *act*VA sequences with various other proteins using a window of 40 amino acids and a stringency of 20. **a** ActVA-ORF2 (horizontal axis) vs Otc-ORFX1 (McDowall 1991). **b** ActVA-ORF3 (horizontal axis) vs Gra-ORF7 (D.H. Sherman, unpublished, EMBL No. X16144). **c** ActVA-ORF6 (horizontal axis) vs TcmVI-ORF1 (H. Motamedi and C.R. Hutchinson, personal communication)

the predicted actinorhodin resistance determinant. This view was made even more likely by the finding that *act*II-ORF1 and -ORF2 form a divergently transcribed pair of genes resembling the repressor/resistance pairs for tetracycline, and even that the promoter-operator signals in the intergenic region resemble those of *E. coli tet* determinants and are expressed in *E. coli* (Caballero et al. 1991). However, disruption of *act*II-ORF2 was found not to be lethal, nor to prevent actinorhodin biosynthesis; instead appearance of the antibiotic in the medium was delayed. It was therefore postulated that this gene indeed codes for a component of an export complex for actinorhodin (together with the product of the translationally coupled *act*II-ORF3 downstream of it), but that there must be a further gene involved in actinorhodin resistance to account for the survival of *act*II-ORF2 mutants (Fernandez-Moreno et al. 1991). Whether this gene could be *act*VA-ORF1 is now open to test. In view of the finding that the *act*VA-ORF1 product appears to be a trans-membrane protein like those of the products of *act*II-ORF2 and the tetracycline and aminotriazole resistance genes, it seems likely that the *act*VA-ORF1 product would also operate as an actinorhodin export system, either together with the *act*II-ORF2 protein or forming a parallel channel. If so, the delay of actinorhodin export in an *act*II-ORF2 mutant, rather than complete retention of the antibiotic in the mycelium (Fernandez-Moreno et al. 1991), could be explained. Conceivably, efficient export via the products of both *act*II-ORF2 and *act*VA-ORF1 in the wild-type strain could be sufficient for self-protection, without the need for a further mechanism of resistance to intracellular antibiotic. This possibility is currently being tested by gene disruption studies in the wild-type and in an *act*II-ORF2 mutant.

act*VA-ORFs 2–6*

*act*VA-ORF2 clearly codes for a protein very similar to the product of *otc*-ORFX1 in the cluster of genes for biosynthesis of oxytetracycline, OTC (McDowall 1991). *otc*-ORFX1 represents part of the DNA that complements a class of mutants, *otcX*, that are blocked at a step in OTC biosynthesis before 4-aminodedimethylaminoanhydrotetracline (4-amino-ATC). Of the other two classes of mutants blocked in 4-amino-ATC biosynthesis (Rhodes et al. 1981), *otcY* is now known to correspond to the polyketide synthase (PKS) for OTC because *otcY* DNA hybridises with the *actI* region, which is known to encode the actinorhodin PKS, and partial sequence information revealed a characteristic series of type II PKS ORFs in the *otcY* region (McDowall et al. 1991; Hopwood and Sherman 1990). *otcZ* codes for a protein that resembles a hydroxyindole-O-methyltransferase from bovine pineal gland (Ishida et al. 1987) and so is probably involved in the methylation that occurs at an early step in 4-amino-ATC biosynthesis (McDowall et al. 1991). Apart from assembly and cyclisation of the polyketide chain (catalysed by the PKS) and methylation, two hydroxylation reactions and two transaminations would be required in 4-amino-ATC biosynthesis (Rhodes et al. 1981). There is no requirement for amination reactions in actinorhodin biosynthesis, but two hydroxylations are required, both of which fail to occur in *act*VA mutants which accumulate product 1 in Fig. 8 (Cole et al. 1987; Bartel et al. 1990) instead of actinorhodin (2 in Fig. 8). It therefore seems highly likely that *act*VA-ORF2 (and *otc*-ORFX1) are involved in ring-hydroxylation reactions.



act mutations	Percentage of blue lysogens with each recombinant phage		Likely location of mutations
	KC591	KC592	
act-218	<0.2 (0/582)	<0.2 (0/219)	ORF3
act-156	<0.5 (0/275)	25	?ORF4
act-278	<0.02 (0/5528)	25	?ORF4
act-190	0.5	33	ORF5
act-101	2	38	ORF5
act-280	5	40	ORF5
act-243	6	43	ORF5
act-234	5	47	ORF5
act-109	4	55	ORF5
act-222	8	57	ORF5
act-211	31	59	ORF5
act-176	31	70	ORF5
act-157	36	63	ORF5
act-245	46	71	ORF5
act-163	64	74	ORF5
act-150	78	84	ORF5
act-249	99	95	ORF5

Fig. 7. Left. Structure of recombinant plasmids pIJ2315 and pIJ2305 and recombinant phages KC591 and KC592. Right. Results of "complementation" tests of KC591 and KC592 against *act*VA mutants

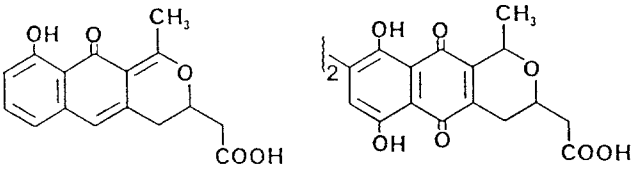


Fig. 8. Structures of (1) the *act*VA intermediate (Cole et al. 1987) and (2) actinorhodin

In a similar way, the finding that the product of *act*-VA-ORF6 resembles that of *tcm*VI-ORF1 implies a common mechanism in the reactions catalysed by these two gene products. *tcm*-VI mutants are blocked in one or more early reactions in the biosynthesis of the anthracycline antibiotic tetracenomycin, after polyketide chain assembly (Motamedi et al. 1986). There are three ORFs in the *tcm*VI region (H. Motamedi and C.R. Hutchinson, personal communication) and recent sub-cloning and complementation analysis suggests that *tcm*VI-ORF1 is needed for hydroxylation of C-5 of the tetracyclic ring system of tetracenomycin F1 to tetracenomycin D3 (C.R. Hutchinson, personal communication). Thus the *act*VA-ORF6 product, also, is a likely candidate for a ring hydroxylase.

Genetic evidence implicates at least two further members of the *act*VA set of ORFs in ring hydroxylation reactions. One representative of the *act*VA class of mutants maps to *act*VA-ORF3, two are possible candidates for ORF4, and 14 others map to ORF5, pointing to involvement of at least two, and possibly all three, of these ORFs in conversion of the *act*VA metabolite to the next intermediate in the pathway. The available evidence suggests that the hydroxyl at C-6 is introduced before that at C-8 because the *act*VB mutant, *act-235*, accumulates as a shunt product, kalafungin, which is hydroxylated at C-6 but not at C-8 (Cole et al. 1987). Thus, *act*VA-ORF3 and -ORF5 (and perhaps -ORF4) seem to be involved in hydroxylation at C-6. That DNA in the *act*VA region is also implicated in C-8 hydroxylation is indicated by the finding that pIJ2315 (Fig. 7), which carries all of the *act*VA ORFs except ORF6, caused the production of the hybrid compound, mederrhodin A (8-hydroxymedermycin), when introduced into the medermycin producer, *Streptomyces* sp. AM-7161 (Hopwood et al. 1985b). (The accumulation of kalafungin by *act*VB mutants does not imply that *act*VB DNA, instead of *act*VA DNA, controls C-8 hydroxylation: chemical arguments suggest that dimerisation of the two monomers to give the dimeric actinorhodin ring system would precede C-8 hydroxylation and so it is reasonable to suppose that *act*VB mutants are not deficient in C-8 hydroxylation ability itself, but perhaps are defective in the ability to carry out the dimerisation reaction.)

Taking the homology results and the genetic data together, at least four of the five ORFs 2–6 are implicated in the introduction of hydroxyl groups onto the ring carbons of actinorhodin. Hydroxylation reactions in other systems are often quite complex. The so-called cytochrome P450 systems require up to three dedicated proteins (O'Keefe and Harder 1991), while some other

systems, such as methane monooxygenase and phenol hydroxylase, consist of even more subunits (Cardy et al. 1991; Nordlund et al. 1990). Since there are no significant resemblances between the products of any of the *act*VA ORFs and cytochromes, ferredoxins or reductases, the C-6 and C-8 hydroxylations in actinorhodin biosynthesis are presumably not carried out by P450 monooxygenases; no similarities were found, either, with any of the other multicomponent systems. Thus the hydroxylases of the actinorhodin pathway must be of a different type which remains to be characterised.

The organization and roles of the genes in the *act*II (Fernandez-Moreno et al. 1991), *act*III (Hallam et al. 1988) and *act*I, VII, IV, VB regions (M.A. Fernandez-Moreno et al., in preparation) are now largely understood in general terms. The present work has revealed a set of six genes in the *act*VA region and has allowed a preliminary allocation of roles to at least five of them. In parallel, a set of genes in the adjacent *act*VI region has been characterised (E. Martinez, in preparation). Thus, we are almost at the point where the basic architecture of the whole gene cluster that controls production and export of actinorhodin is defined well enough to provide a powerful springboard for the next stages of more detailed functional understanding of the complete set of genes for this model antibiotic.

Acknowledgements. We are grateful to K.J. McDowall, and H. Motamedi and C.R. Hutchinson for unpublished sequence information for *otc*-ORFX1 and *tcm*VI-ORF1, respectively. This research was supported in part by grants (to F.M.) from the Spanish C.I.C.Y.T. (PBY86-0025), and EEC (BAP 0393-E) and Smith, Kline and French, S.A.E. and (to. D.A.H.) from the Agricultural and Food Research Council and the John Innes Foundation. J.L.C. acknowledges a M.E.C./British Council Fleming Fellowship and one from the John Innes Foundation.

References

Bartel PL, Zhu C-B, Lampel JS, Dosch DC, Connors NC, Strohl WR, Beale JM, Floss HG (1990) Biosynthesis of anthraquinones by interspecies cloning of actinorhodin biosynthesis genes in *Streptomyces*: clarification of actinorhodin gene functions. J Bacteriol 172:4816–4826

Beld M, Martin C, Huits H, Stuitje AR, Gerats AGM (1989) Flavonoid synthesis in *Petunia hybrida*: partial characterization of dihydroflavonol-4-reductase genes. Plant Mol Biol 13:491–502

Bibb MJ, Findlay PR, Johnson MW (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

Bruton CJ, Guthrie EP, Chater KF (1991) Phage vectors that allow monitoring of transcription of secondary metabolism genes in *Streptomyces*. Biotechnology 9:652–656

Caballero JL, Malpartida F, Hopwood DA (1991) Transcriptional organization and regulation of an antibiotic export complex in the producing *Streptomyces* culture. Mol Gen Genet 228:372–380

Cardy DNL, Laidler V, Salmond GPC, Murrell JC (1991) Molecular analysis of the methane monooxygenase (MMO) gene cluster of *Methylosinus trichosporium* OB3b. Mol Microbiol 5:335–342

Chater KF (1990) Multilevel regulation of *Streptomyces* differentiation. Trends Genet 5:372–377

Cole SP, Rudd BAM, Hopwood DA, Chang C-J, Floss GH (1987) Biosynthesis of the antibiotic actinorhodin: analysis of blocked mutants of *Streptomyces coelicolor*. J Antibiot (Tokyo) 40:340–347

Devereux J, Haeberli P, Smithies O (1984) A comprehensive set of sequence analysis programs for the Vax. Nucleic Acids Res 12:387–395

Fernandez-Moreno MA, Caballero JL, Hopwood DA, Malpartida F (1991) The actinorhodin gene cluster contains regulatory and antibiotic export genes that are direct targets for translational control by the *bldA* tRNA gene of *Streptomyces coelicolor*. Cell 66:769–780

Hallam SE, Malpartida F, Hopwood DA (1988) DNA sequence, transcription and deduced function of a gene involved in polyketide antibiotic synthesis in *Streptomyces coelicolor*. Gene 74:305–320

Henikoff S (1984) Unidirectional digestion with ExoIII creates targeted breakpoints for DNA sequencing. Gene 28:351–359

Hopwood DA (1988) Towards an understanding of gene switching in *Streptomyces*, the basis of sporulation and antibiotic production. Proc R Soc Lond [Biol] 235:121–138

Hopwood DA, Sherman DH (1990) Molecular genetics of polyketides and its comparison to fatty acid biosynthesis. Annu Rev Genet 24:37–66

Hopwood DA, Bibb MJ, Chater KF, Kieser T, Bruton CJ, Kieser HM, Lydiate DJ, Smith CP, Ward JM, Schrempf H (1985a) Genetic manipulation of *Streptomyces*. A laboratory manual. John Innes Foundation, Norwich

Hopwood DA, Malpartida F, Kieser HM, Ikeda H, Duncan J, Fujii I, Rudd BAM, Floss HG, Omura S (1985b) Production of 'hybrid' antibiotics by genetic engineering. Nature 314:642–644

Ishida I, Obinata M, Deguchi T (1987) Molecular cloning and nucleotide sequence of cDNA encoding hydroxyindole O-methyltransferase of bovine pineal glands. J Biol Chem 262:2895–2899

Kanazawa S, Driscoll M, Struhl K (1988) *ATR1*, a *Saccharomyces cerevisiae* gene encoding a transmembrane protein required for aminotriazole resistance. Mol Cell Biol 8:664–673

Klock G, Unger B, Gatz C, Hillen W, Altenbuchner J, Schmid K, Schmitt R (1985) Heterologous repressor-operator recognition among four classes of tetracycline resistance determinants. J Bacteriol 161:326–332

Kyte J, Doolittle RF (1982) A simple method for displaying the hydropathic character of a protein. J Mol Biol 157:105–132

Malpartida F, Hopwood DA (1984) Molecular cloning of the whole biosynthetic pathway of a *Streptomyces* antibiotic and its expression in a heterologous host. Nature 309:462–464

Malpartida F, Hopwood DA (1986) Physical and genetic characterisation of the gene cluster for the antibiotic actinorhodin in *Streptomyces coelicolor* A3(2). Mol Gen Genet 205:66–73

Maniatis T, Fritsch EF, Sambrook J (1982) Molecular cloning: A laboratory manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York

Martín JF, Liras P (1989) Organization and expression of genes involved in the biosynthesis of antibiotics and other secondary metabolites. Annu Rev Microbiol 43:173–206

McDowall KJ (1991) An investigation of the molecular genetics of oxytetracycline biosynthesis in *Streptomyces rimosus*. Ph D Thesis, University of Glasgow, UK

McDowall KJ, Doyle D, Butler MJ, Binnie C, Warren M, Hunter IS (1991) Molecular genetics of oxytetracycline production by *Streptomyces rimosus*. In: Baumberg S, Krügel H, Noack D (eds) Genetics and product formation in *Streptomyces*. Plenum Press, New York, pp 105–116

Motamedi H, Wendt-Pienkowski E, Hutchinson CR (1986) Isolation of tetracenomycin C nonproducing *Streptomyces glaucescens* mutants. J Bacteriol 167:575–580

Neal RJ, Chater KF (1987) Nucleotide sequence analysis reveals similarities between proteins determining methylenomycin A resistance in *Streptomyces* and tetracycline resistance in eubacteria. Gene 58:229–241

Nordlund I, Powlowski J, Shingler V (1990) Complete nucleotide sequence and polypeptide analysis of a multicomponent phenol hydroxylase from *Pseudomonas* sp. strain CG600. J Bacteriol 172:6826–6833

O'Keefe DP, Harder PA (1991) Occurrence and biological function of cytochrome P450 monooxygenases in the actinomycetes. Mol Microbiol 5:2099–2105

Rhodes PM, Winskill N, Friend EJ, Warren M (1981) Biochemical and genetic characterization of *Streptomyces rimosus* mutants impaired in oxytetracycline biosynthesis. J Gen Microbiol 124:329–338

Rudd BAM, Hopwood DA (1979) Genetics of actinorhodin biosynthesis by *Streptomyces coelicolor* A3(2). J Gen Microbiol 114:35–43

Sanger F, Nicklen S, Coulson AR (1977) DNA sequencing with chain-terminating inhibitors. Proc Natl Acad Sci USA 74:5463–5467

Seno ET, Baltz RH (1989) Structural organization and regulation of antibiotic biosynthesis and resistance genes in actinomycetes. In: Shapiro S (ed) Regulation of secondary metabolism in Actinomycetes. CRC, Boca Raton, Florida, pp 1–48

Sherman DH, Malpartida F, Bibb MJ, Kieser HM, Bibb MJ, Hopwood DA (1989) Structure and deduced function of the granaticin-producing polyketide synthase gene cluster of *Streptomyces violaceoruber* TÜ22. EMBO J 8:2717–2725

Yanisch-Perron CJ, Vieira J, Messing J (1985) Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. Gene 33:103–119

Communicated by H. Hennecke