An amplifiable and deletable locus of *Streptomyces ambofaciens* RP181110 contains a very large gene homologous to polyketide synthase genes

Bertrand Aigle, Dominique Schneider,† Cécile Morilhat, Dominique Vandewiele, Annie Dary, Anne-Catherine Holl, Jean-Marc Simonet‡ and Bernard Decaris

Author for correspondence: Bernard Decaris. Tel: +33 83 91 20 96. Fax: +33 83 91 25 00. e-mail: decaris@nancy.inra.fr

*Laboratoire de Génétique et Microbiologie, Associé INRA, Faculté des Sciences, Université Henri Poincaré, Nancy I, BP 239, 54506 Vandoeuvre-lès-Nancy, France*

*Streptomyces ambofaciens* RP181110 produces the macrolide polyketide spiramycin. Like many other *Streptomyces* species, the RP181110 strain is prone to genetic instability involving genomic rearrangements (deletions and/or amplifications) in the large unstable region of the genome. It has previously been demonstrated that the amplification of a particular locus (AUD205) affects spiramycin biosynthesis and, conversely, the loss of this amplification is correlated with the restoration of antibiotic production. This report focuses on a 0.93 kb reiterated fragment specific for the AUD205 locus. Sequencing of 3596 bp including this reiteration revealed the presence of an ORF (orfPS) whose potential product was highly homologous to the EryA and Raps proteins, responsible for the biosynthesis of erythromycin in *Saccharopolyspora erythraea* and rapamycin in *Streptomyces hygroscopicus*, respectively. orfPS encodes a protein with at least four successive domains: ketoacyl synthase, acyltransferase, ketoreductase and acyl carrier protein. This organization is very similar to most eryA and rap modules. The reiterated sequence corresponds to the acyltransferase domain. orfPS was transcribed during rapid growth and stationary phase in RP181110 and overtranscribed in the amplified mutant. Both these results suggest that the gene encodes a type I polyketide synthase and its reorganization is responsible for the loss of spiramycin production in the amplified strains.

Keywords: polyketide synthase (PKS) gene, unstable region, overtranscription, spiramycin production, *Streptomyces ambofaciens*

INTRODUCTION

*Streptomyces* spp. are notable for their ability to produce a wide range of secondary metabolites. These metabolites represent a class of natural products with diverse biological activities such as antibiotics, insecticides and chemotherapeutic agents (Demain et al., 1983).

Some secondary metabolic characteristics are genetically unstable and can mutate at a frequency greater than $10^{-3}$. In *Streptomyces ambofaciens*, genetic instability affects pigmentation, aerial mycelium formation and antibiotic production (Leblond et al., 1989; Volff et al., 1993). Genomic rearrangements such as deletions, often associated with large-scale DNA amplification, are correlated with genetic instability (Birch et al., 1990; Leblond et al., 1991). The *Streptomyces* chromosomal DNA is linear (Lin et al., 1993; Lezhava et al., 1995; Leblond et al., 1996), and both deletions and amplifications have been located in regions overlapping the chromosome ends (Leblond et al., 1996; Redenbach et al., 1993).
In *S. ambofaciens* DSM 40697, the 2000 kb unstable region contains two amplifiable loci, **AUD6** and **AUD90** (Demuyter et al., 1988), without well-defined functions. A chloramphenicol resistance determinant is located in an amplifiable and deletable region of *Streptomyces lividans* 1326 (Dittrich et al., 1991). The unstable region of the *Streptomyces glaucescens* chromosome contains the melC and strC genes (Birch et al., 1989). Antibiotic biosynthesis and resistance genes are located within amplifiable units of DNA (AUDs) in some *Streptomyces* spp. (Potekhin & Danilenko, 1985; Hornemann et al., 1987), as is a mercury resistance determinant (Sedlmeier & Altenbuchner, 1992). Schneider et al. (1993) demonstrated the presence in *S. ambofaciens* DSM 40697 AUD6 of two transcribed genes showing increased transcription in the amplified strain and homology with regulatory genes. The AUD locus also contains two genes encoding products with about 30% identity with different bacterial repressors (Aubert et al., 1993).

Mutants derived from *S. ambofaciens* RP181110 and amplified in the **AUD90** locus show loss of the capacity to produce the macrolide antibiotic spiramycin (Spi- phenotype; Dary et al., 1992). The NSA205 strain contains an 89 kb amplified DNA sequence (ADS). **(ADS205)**. Instability in its progeny is associated with amplifications resulting in the Spi- phenotype, in which the loss of **ADS205** is accompanied by restoration of spiramycin production. Nevertheless, the spiramycin biosynthetic genes are not located in the unstable region of NSA205 (Dary et al., 1993), suggesting that **ADS205** could contain a gene interfering with spiramycin biosynthesis.

Here, we demonstrate the presence of one very large ORF in **AUD205**. This ORF is transcribed and its predicted product shows homology with the products of genes for the synthesis of polyketide-derived portions of erythromycin in *Saccharopolyspora erythraea* and of rapamycin in *Streptomyces hygroscopicus* (Donadio et al., 1991; Schwecke et al., 1995).

**METHODS**

**Bacterial strains and plasmids.** *S. ambofaciens* RP181110 was derived from the ATCC 23877 strain (Pinnert-Sindico et al., 1955). NSA205, a mutant derived from RP181110, has an 89 kb amplification (**ADS205** in the **AUD90** locus and is Spi- (Dary et al., 1992). *S. ambofaciens* NSA97H, in which the **AUD90** locus is totally deleted, is a spontaneous mutant of the DSM 40697 strain (Leblond et al., 1991). *Escherichia coli* SURE (Stratagene) was the host for cloning experiments.

The plasmids pOS15 and pOS17 (Fig. 1) resulted from insertion into pl702 (Kaz et al., 1983) of 4.6 and 5.6 kb BamHI DNA fragments, respectively. These inserts come from adjacent segments of DNA in an ADS belonging to the **AUD90** locus; in **AUD205**, the two segments are amplified. pBluescript II KS (Short et al., 1988) was purchased from Stratagene. The pKS2/M5 plasmid was kindly provided by C. R. Hutchinson. pKS2/M5 is pUC19 with a 1.3 kb *Smal* fragment from module 5 of the erythromycin *eryAIII* polyketide synthase gene containing the second ketoacyl synthase (KS) domain (Donadio et al., 1991).

**Fig. 1.** Restriction maps of plasmids pOS15 (a) and pOS17 (b). The narrow lines represent the pl702 vector; the boxed regions represent the cloned **AUD205** BamHI fragments of 4.6 kb (pOS15) and 5.6 kb (pOS17). The black box in (a) identifies the 0.93 kb SacI-BamHI sequence containing the AT motif; in (b) it identifies the continuation of this sequence in pOS17. The shaded boxes in both (a) and (b) indicate regions homologous to the second KS domain of the *Sacch.* *erythraea* erythromycin *eryAIII* KS gene.

**Media and culture conditions.** *S. ambofaciens* RP181110 was grown at 30°C on Hickey Tresner (HT) agar (Pridham et al., 1957). Large-scale DNA isolation was carried out after 48 h growth of this strain at 30°C with shaking at 200 r.p.m. in YEME liquid medium supplemented with 0.25% glycerol (Hopwood et al., 1985). *E. coli* SURE was grown at 37°C in LB liquid medium (Maniatis et al., 1982) supplemented with tetracycline (12 μg ml⁻¹), and with ampicillin (50 μg ml⁻¹) when the strain contained pBluescript II KS.

For RNA isolations, pre-germination medium (Hopwood et al., 1985) was inoculated with about 10⁷ c.f.u. spores ml⁻¹ and incubated for 90 min at 37°C. The pre-germination spores were
then used to inoculate 200 ml MP5 medium (Pernodet et al., 1993) in 21 baffled shake flasks, which were incubated at 25 °C with shaking at 250 r.p.m.

**DNA preparation and restriction analysis.** Total DNA of *S. ambofaciens* was prepared for Southern blots as described previously (Demuyter et al., 1988). *S. ambofaciens* and *E. coli* plasmid extractions and purifications were performed as described by Hopwood et al. (1985) and Maniatis et al. (1982).

Restriction enzymes were purchased from Boehringer Mannheim and New England Biolabs and used as recommended by the suppliers. DNA fragments were electrophoresed on agarose gels as described by Maniatis et al. (1982).

**RNA isolation and manipulation.** RNA was extracted from liquid cultures during exponential and stationary phases. The isolations were performed essentially as described by Guillfoile & Hutchinson (1992) except for the centrifugation conditions. After lysis of the cells, CsCl was added to the lysates (24 g per 6 ml); the samples were layered over a 2:8 ml cushion of 5 M CsCl in 0.1 M EDTA, pH 7.5, and centrifuged in a Kontron TST41-14 rotor for 18 h at 29000 r.p.m. and 20 °C.

**Southern and Northern hybridizations.** RNA samples (40 μg) were denatured with glyoxal and fractionated by electrophoresis in non-denaturing agarose gels as described by Hopwood et al. (1985): an RNA ladder (0.24-9.5 kb; Gibco BRL) and HindIII-digested lambda phage DNA were used as size markers. After electrophoresis, the gel was soaked for 20 min in 50 mM NaOH to partially hydrolyze the RNA and improve the efficiency of transfer. The gel was then rinsed in diethyl-pyrocarbonate-treated water and soaked for 45 min in 20 x SSC (1 x SSC is 0.15 M sodium chloride, 0.015 sodium citrate). The gel was blotted on Hybond nylon membranes (Amersham) with the Vacugene system (LKB).

Probes were labelled with [α-32P]dCTP using the Megaprime kit (Amersham). Prehybridizations, hybridizations and washes were as described by Demuyter et al. (1988) except that SSPE (30 M NaCl, 0.2 M NaH2PO4, 0.02 M Na2EDTA, pH 7-4) replaced SSC in the Northern hybridizations and washings. After hybridization, blots were washed at 65 °C (50 °C for the probe pKS2/M5) and exposed at −70 °C to X-ray film using two intensifying screens.

**DNA sequencing and sequence analyses.** The sequences were determined using the chain-termination method (as previously described by Schneider et al., 1993) with the TaqTrack sequencing kit (Promega) using 7-deaza-dGTP instead of dGTP. Products were analysed on 5% (w/v) polyacrylamide/7 M urea gels. For sequencing, the fragments were cloned into pBluescript II KS. Nested deletions were carried out within the cloned fragment using a double-stranded nested deletion kit (Pharmacia).

Comparisons of nucleotide and amino acid sequences with the databases were performed by the BLAST network service (Altschul et al., 1990). Protein sequences were aligned with the CLUSTAL V program (Higgins & Sharp, 1988) and the dendrogram was constructed using the PAUP program (Swofford, 1991).

**RESULTS**

**Evidence of specific reiterations in AUD205**

Hybridization of a BamHI digest of *S. ambofaciens* genomic DNA under high stringency conditions with a 4.6 kb fragment containing the 4.6 kb and 5.6 kb BamHI fragments of AUD205 as a probe revealed bands additional to those corresponding to the probe.

We first focused on identifying these reiterations. Total DNA of the RP181110 strain digested with BamHI and hybridized with pOS15 showed, in addition to the 4.6 kb band, 12 other bands of various intensities (data not shown). Identical results were found with *S. ambofaciens* ATCC 23877, the ancestor of the RP181110 strain (Volff et al., 1993), as well as with *S. ambofaciens* DSM 40697 (data not shown).

With the 0.66 kb SacI-Smal fragment of pOS15 as a probe, only seven bands in addition to the 4.6 kb BamHI fragment were detected (Fig. 2). Probing with the 0.51 kb SacI-SmaI labelled 0.66 kb SacI-Smal fragment of pOS15, the asterisks indicate the eight signals detected in the RP181110 DNA. Gels were exposed for 3 d at −70 °C with two intensifying screens.

![Fig. 2.](image-url)
B. AIGLE and OTHERS

GQGSQRLGMGRELYGAFPVF

GCG GAC GCT TTC GAC GCG GTG TGT GCG CGG GTG GAT CTT GAG CGT CCG CTG CGT GAG GTG

GGT CAG GGT TCG CAG CGG CTG GGG ATG GGC CGG GAG TTG TAC GGG GCA TTT CCG GTG TTC

GCC GTG GCC CTG GGA CAC GAC CGC GAG GCC TTG CTC ACC GAA CTC TCC GAC CTC GCT GAC

AGHLTVRDSAESGRLALLFS

GGA ATC AGC GGC ACC AAC GCA CAC GTG GTG CTG GAG CCG ATC CCC GCG ACG ACU

TEARPWPEPDRPRRAGVSAF

AMLAVEVAEEALELPDGVDL

ALRPAAATAAPDTVQDGFRA

ADAFDAVCARVDLERPLREV

PQALREQAARLAERVTRDDT

VGHSIGELAAAHVAGVLSLE

VFGEDAGLrERTVYAQAGLF

LHPVDVAVSLADGRAHLEHR

CTG GAG GAG CGG CTC CGG GCG CAG GAG GTC CGC GTC AAG CGC CTC GCG GTC TCG CAC GCC

GTG GGG CAT TCG ATC GGT GAG CTG GCC GCC GCG CAT GTG GCA GGG GTG TTG TCG CTG GAG

GCG CTG GAA GTG GCG CTG TTC CGG CTG GTG GAG TCG TGG GGT GTG ACC CCG GAC GTG CTG

GTG TTC GGC GAG GAC GCG GGG TTG CTG GAG CGG ACC GTC TAC GCG CAG GCG GGG

CCG CAG GCG CTG CGG GAG CAG GCC GCC CGA CTG GCC GAG CGG GTC ACG CGG GAC GAC ACT

GAA AGC CCC GTG CGC CGG ACG ACC CCA CCC CTG GTC CCG TGG GTG CTC TCG GCC CGC ACG

GISGTNAHVVLEPIPATTSS

CTG CAC GTG ACC GAG CCC ACG CCC GAG GTG GAC TGG ACG ACG GGC GCG GTG GAA CTG CTC

AAVNGPTSVTVSGEAEAIAE

GAT GCC TGC GCG TTG GTG TCG GCT CGG GGG CGG TTG ATG GAT GCG TTG ccc

TYHTPRIPVVTTAPGDIATP

GAG TAC TGG GTG GGC CAG GTG CGC GAG CCG GTC CGC TTC GCC GAC GCG CTC GGC CGG TTC

ACC TAC CAC ACC CCC

TTC CAC TCC CAC TTG ATG GAG CCG ATG CTG GCC GAC TTC GCG GCC GTG GCG CGG TCG TTG

FHSHLMEPMLADFAAVARSL

AUD205

GhC CCG CGC CTC GCC GCG CGG CTC GCG GCG ACC CTG ACC CGC ACC GhC GGC GAG GAC CAG

DPRLAARLAATLTRTDGEDQ

SPASGRPSTRATSSPWPPPR

WPDPPAPDAAARRRPPHHRR

AGVRTALELGPAGVLSALAQQ

HG GTC GGC GGC GTC ATC AAG ACG GTC ATG GCA CTG CGG CAC GGT GTG CTG CCC CGT ACC

LGVLAPLADTGVADLAHLVG

LGARVTVVACDLQARGDVSR

TLLRVIARAHVRGAEVDWSA

EQVDSLVAVPALRPGHGEAE

VAVRAAATLGRTLRHADPVP

TYHTPRIPVVTTAPGDIATP

GAG TAC TGG GTG GGC CAG GTG CGC GAG CCG GTC CGC TTC GCC GAC GCG CTC GGC CGG TTC

ACC TAC CAC ACC CCC

TTC CAC TCC CAC TTG ATG GAG CCG ATG CTG GCC GAC TTC GCG GCC GTG GCG CGG TCG TTG

NNN

LAHPRHWGGLVDLPADGLGD

CGA CGG TCG ACG CCT GCC GCG CGG CGC TCG CCG AAC GGC AGG GCG AGG TCG ACC TGC TCG

CGA CGG TCG ACG CCT GCC GCG CGG CGC TCG CCG AAC GGC AGG GCG AGG TCG ACC TGC TCG

SRPRRPARDGGFWWLRTTPT

PASTRPSVPGGTGSPGSPWR

1861/621 1891/631

14411481 1471/491

1321/441 1351/451

121/41 151/51

181/61 211/71

1081/361 1111/371

1021/341 1051/351

781/261 811/271

721/241 751/251

601/201 631/211

541/181 571/191

1051/361 1081/361

1531/511 1501/501

1621/541 1651/551

1541/501 1571/511

121/41 151/51

181/61 211/71

1281/41 1311/51

1621/541 1651/551

1661/51 1691/52

1901/52 1931/53

2161/721 2191/731

2221/741 2251/751

2281/761 2311/771

2321/781 2351/791

2421/801 2451/811

2491/821 2521/831

2551/841 2581/851

2641/861 2671/871

2741/881 2771/891

2841/901 2871/911

2941/921 2971/931

3041/941 3071/951

3181/961 3211/971

3421/981 3451/991

3301/1001 3331/1011

3391/1021 3421/1031

3391/1041 3421/1051

3481/1061 3511/1071

3531/1081 3561/1091

3591/1101 3621/1111

Fig. 3. Nucleotide and deduced amino acid sequences of the 3596 bp segment of pOS15. The relevant restriction sites are identified and are in small bold type.

NSA205, a mutant showing an amplification in the AUD205 locus, was also probed with pOS15 (Fig. 2a). The reiterations hybridized with amplified fragments, but some faint signals, different from the amplified fragments, were also detected. Fig. 2(b) shows hybridization with the 0.66 kb SacI-SmaI fragment of pOS15. The faint signals detected could correspond to junction fragments after deletion. However, these faint signals could also result from a clonal heterogeneity. The largest fragment (> 23 kb) detected in RP181110 was absent in the NSA205 strain. This might have resulted from the deletion process preceding the amplification. The 9.4 kb signal was given by two similarly sized amplified fragments (9.4 and 9.2 kb). The 9.2 kb BamHI fragment was not detected in the RP181110 digest and probably arose from the fragment created by tandem reiteration. The reiterations present in pOS15 are thus specific to the AUD205 locus.

Potential type I polyketide synthase (PKS) gene in AUD205

A 3596 bp region representing 1156 bp of the 4.6 kb BamHI fragment and 2440 bp of the adjacent 5.6 kb BamHI fragment, and including the 0.93 kb SacI-BamHI fragment of pOS15 was sequenced (Fig. 3). Its high G+C
content (75.2%) is in accord with the G+C content of Streptomyces genomes. Consistent with this, one of the reading frames, orfPS, shows a strong G+C bias (90%). The deduced amino acid sequence of orfPS shows significant sequence similarity to type I PKSs. These enzymes are large multifunctional proteins similar to the mammalian fatty acid synthases (FASs) and contain sets of domains with catalytic functions. The alignment shown in Fig. 4 with the products of rap module 11 from S. amhofaciens and module 5 from S. hygroscopicus (Raps3/M11) and module 5 from Sacch. erythraea (EryA3/M5) was done with the CLUSTAL program. Amino acids boxed in black are conserved in at least two sequences. The asterisks identify the amino acid motifs assigned as the end of the KS domain. Underlining with +, = and - identifies the ACP, KR and AT domains, respectively. Underscoring with \* identifies the sequence defining the substrate specificity of the AT domains.
enzymes activities catalysing successive rounds of elongation of the polyketide product. The highest similarities were to the type I PKSs for erythromycin and rapamycin synthesis in *Saccharophagus erythraea* (Cortes et al., 1990; Donadio et al., 1991) and *S. hygroscopicus* (Schwecke et al., 1995), respectively. The rapamycin PKS genes (*rap*), like those involved in erythromycin biosynthesis (*ery*), are clustered and correspond to three ORFs, *rapA* (28.5 kb), *rapB* (34.5 kb) and *rapC* (20.8 kb), containing four, six and four modules, respectively, encoding FAS-like activities (Schwecke et al., 1995). The *ery* genes correspond to three 10 kb ORFs, *eryAIII*, *eryAII* and *eryAII*, each including two modules (Donadio et al., 1991).

Homologies involve the KS, acyltransferase (AT), ketoreductase (KR) and acyl carrier protein (ACP) domains of these PKSs (Fig. 4). The amino acid motif G_45/TNAHVVE_27 is highly similar to the motif (GTNAHVVE, upper case letters referring to invariant residues) considered to be the likely end of the KS domains (Donadio & Katz, 1992). This implies that the first 71 aa of OrfPS represent the carboxy terminus of a KS domain. Consistent with this, the sequence shows 63.4% and 74.6% identity with the C-terminal part of KS domains encoded by module 5 of *eryAIII* and module 11 of *rapC*, respectively.

Hybridization of the 1.19 kb *KpnI–SacI* fragment of pOS15 with the *eryAIII* fragment from pKS2/M5 (see Methods) as a probe detected homology in the region immediately upstream of the 0.93 kb *SacI–BamHI* fragment (Fig. 1a), confirming the presence of the KS domain. Partial sequencing of the pOS15 insert from the *KpnI* site (in the opposite direction to the 0.93 kb *SacI–BamHI* fragment) showed that the region encoding the KS domain began immediately upstream of the *KpnI* site. That it ended a few nucleotides before the *SacI* site was shown by analysis of the 3596 bp *SacI–BamHI* sequence (Figs 3, 4). The pKS2/M5 insert was used to probe pOS17, containing the cloned 5.6 kb *BamHI* fragment from *S. ambofaciens* RP181110 (Fig. 1b). Hybridization with the 1.9 kb *KpnI–SacI* fragment of pOS17 (data not shown) indicated that another KS domain was located downstream of the sequenced region. Hybridizations using this probe with *BamHI*-digested total DNA of RP181110 and NSA205 confirmed the reiteration of the KS domain (by detecting many signals in addition to those given by the 4.6 kb and 5.6 kb *BamHI* fragments), and showed that the reiterations were specific to the amplifiable region (data not shown). The results are consistent with the presence of type I polyketide genes. For comparison, the *rap* and *eryA* loci contain 14 and 6 KS domains, respectively (Schwecke et al., 1995; Donadio et al., 1991).

The adjacent AT domain would begin at the amino acid motif A_272/LLEFGQSQQ_288 and would end a few residues after the motif G_272/AEVWDSAV_489 (Donadio & Katz, 1992). This domain showed 36.8% and 47.7% identity with the AT domains encoded by module 5 of *eryAIII* and module 11 of *rapC*, respectively. A large degree of sequence conservation was observed around the sequence G_261/HSSG_265 (x = any amino acid). This motif corresponds to the AT active site, S being the serine involved in forming the acyl-enzyme intermediate. The sequence upstream of the active site allowed us to predict the nature of the incorporated extender unit (see below). The 0.93 kb *SacI–BamHI* fragment contained a large part of the sequence encoding the AT domain (Fig. 3); since hybridization experiments with the reiteration as a probe showed seven signals in addition to those given by the 4.6 kb *BamHI* fragment (see Fig. 2b), the **AUSD205** locus might contain at least eight AT domains.

The amino acid motifs **P_74**DGTNLV_74 and **G965**PW-DGXXX_47 might represent the start and the end of the KR domain, respectively, which showed 57.4% and 44.3% identity with the KR domains encoded by module 5 of *eryAIII* and module 11 of *rapC*, respectively. The extremities of the ACP domain were assigned, on the basis of data from Donadio & Katz (1992), to the conserved motifs L_1961/AALSPVE_1968 and G_1962/LTLPLAVL_1969, the active site being defined by L_1961/GADS_1969. Identities of 47.8 and 57.9% were found with the ACPs from module 5 of *eryAIII* and module 11 of *rapC*, respectively.

These results predict that *orfPS* contains at least one PKS module encoding domains in the order KS–AT–KR–ACP. This organization is consistent with that of most modules in the PKSs for biosynthesis of erythromycin and rapamycin, the exceptions being modules containing dehydratase (DH) and/or enoyl reductase (ER) domains. The presence of a KS domain downstream of the sequenced region suggests that a second PKS module is also present.

**Substrate specificity of the AT domain**

Haydock et al. (1995) compared the amino acid sequences of a large number of type I PKS AT domains catalysing the transfer of acetyl-CoA esters onto ACP. A 20 aa motif, located 11 residues upstream of the active site (GHSxG) is particularly valuable for predicting the AT specificity. This motif, present in the OrfPS AT sequence (Fig. 4), shows an excellent match to one of the seven AT domains containing dehydratase (DH) and/or enoyl reductase (ER) domains. The presence of a KS domain downstream of the sequenced region suggests that a second PKS module is also present.

**Transcriptional analysis**

To investigate *orfPS* transcription, total RNAs were extracted from *S. ambofaciens* RP181110 and NSA205 after 37 h (rapid growth phase) and 68 h (stationary phase) of culture. Dot-bLOTS of stationary-phase RNAs revealed transcription of *orfPS* in RP181110, and showed a higher
level of transcription in NSA205 compared with RP181110 and a NSA205-derived, de-amplified mutant (data not shown). RNAs were then subjected to Northern analysis with the pOS15 0.66 kb SacI–SmaI fragment as a probe (Fig. 6). A 25 kb signal, corresponding to a very large RNA for bacteria, was detected in both 37 and 68 h cells, suggesting transcription of orfS during rapid growth and stationary phases in RP181110 and NSA205. The signal persisted in DNase-treated samples, and at both sampling times was much stronger in NSA205 than in RP181110. Two other bands of about 2-9 kb and 1-9 kb were also detected. These might correspond to transcripts of a sequence homologous to the AT motif, but they might also represent two degradation products.

DISCUSSION

The above results provide evidence of at least two reiterated sequences specific to AUD205, an amplifiable and deletable locus of S. ambofaciens RP181110. Analysis of a 3596 bp sequence that included one reiteration showed the presence of an incomplete ORF (orfPS). Northern blot hybridization demonstrated that orfPS is transcribed as a large unit. Together with other studies (Schneider et al., 1993; Aubert et al., 1993), these results confirm the presence of many actively transcribed genes in the unstable region of the S. ambofaciens RP181110 genome.

Protein database comparisons suggest that orfPS might encode a type I PKS similar to the PKs involved in rapamycin biosynthesis in S. hygroscopicus (Schwecke et al., 1995) and erythromycin biosynthesis in Sacch. erythraea.
(Donadio et al., 1991). The sequenced region corresponds to almost the entire PKS gene module encoding KS-AT-KR-ACP activities (part of the fragment encoding the KS domain was not sequenced). This organization is identical to 4 of 6 modules of Sacc. erythraea eryA and 4 of 14 of S. hygroscopicus rap. The presence of another KS domain downstream of the ACP suggests that a second PKS module is located immediately downstream of the sequenced region. This second unit probably belongs to another ORF since in the Rap and EryA multifunctional proteins the intervening sequence between the ACP and KS domains extends over about 20 residues. The 52 aa downstream of the OrfPS ACP domain do not include the motif thought to be at the beginning of the KS domains (PiiVgmaCR; Donadio & Katz, 1992). The AUD205 locus thus may contain at least two ORFs that together encode a PKS complex. However, we cannot exclude the possibility that the potential ACP-KS intervening region is larger in OrfPS.

It is clear that the KS and AT domains are reiterated in the AUD205 locus, and in fact many PKS subunits of the type KS-AT-KR-ACP, as in the eryA or rap complexes, may be present.

Transcriptional analyses revealed an orfPS transcript of about 25 kb, which is compatible with the size of PKS genes involved in antibiotic biosynthesis. The eryA, eryAII and eryAIII genes of Sacc. erythraea are larger than 10 kb (Donadio et al., 1991), and the rapA, rapB and rapC genes of S. hygroscopicus extend over 28-5, 34-5 and 20-8 kb, respectively (Scheweke et al., 1995). Several other large, type I PKS genes have been described, including srmG (at least 11 kb), responsible for spiramycin biosynthesis in S. ambiguus (Geistlich et al., 1992). Furthermore, the transcripts might be polycistronic. Donadio et al. (1992) suggested that the three genes of the eryA complex could lie on the same transcriptional unit, estimated as 30 kb. Consequently, orfPS might be a very large gene similar in size to the rap genes, or it might be transcribed as a polycistronic RNA.

orfPS was transcribed in both RP181110 and NSA205 strains during the stationary phase and, more surprisingly, during the rapid growth phase under the conditions used. The KS involved in the synthesis of the polyketide portion of tetracenomycin C in S. glaucescens is detected only in stationary-phase cultures (Gramajo et al., 1991), but concern directly the protein and not the RNA. Nevertheless, the biosynthetic structural genes for the polyketide antibiotic actinorhodin are transcribed only during the transition from exponential to stationary phase and during early stationary phase in Streptomyces coelicolor A3(2) (Gramajo et al., 1993).

Polyketide-derived compounds include a wide variety of natural products including pigments as well as antibiotics (for a review, see Hopwood & Sherman, 1990), but although AUD205 might contain genes involved in pigmentation, the PKSs catalysing the synthesis of polyketide-derived streptomycete pigments (Davis & Chater, 1990; Blanco et al., 1992) belong to the type II family (multi-enzyme complexes formed by several poly-peptides with individual enzymic activities). Moreover, whilst the whiE locus hybridizes with the chromosome of S. ambifaciens ATCC23877, the homologous whiE sequence is not located in the AUD90 region, equivalent to the AUD205 locus (data not shown). A type I PKS is responsible for the Aspergillus nidulans conidial wall pigment (Mayorga & Timberlake, 1992). However, no clear correlation was found between pigmentation and genomic rearrangements in the unstable region of the RP181110-derived mutants. In the ATCC 23877 strain, some mutants are unpigmented but no rearrangements in the AUD90 region were detected by hybridization (Voll et al., 1993). Thus, orfPS probably does not encode a pigment PKS.

Donadio et al. (1991) suggested that the macrolide polyketide biosynthesis genes are organized in modules as in eryA, since all these genes contain repeated units. Spiramycin, produced by S. ambifaciens RP181110, belongs to this antibiotic family and the polyketide portion is synthesized by a PKS encoded by srmG (Richardson et al., 1990). Dary et al. (1993) demonstrated that the spiramycin biosynthesis genes are not located in the AUD205 region and srmG is located outside the unstable region (P. Leblond, personal communication). Thus, orfPS is not directly responsible for spiramycin biosynthesis and does not correspond to srmG.

Nevertheless, AUD205 plays a role in spiramycin production: amplification of this locus results in a loss of antibiotic synthesis and spontaneous deamplification restores spiramycin production (Dary et al., 1992). One explanation for this would be the involvement of at least one gene present on ADS203 in this non-producing phenotype. Since AUD205 amplification leads to orfPS overtranscription, the orfPS-encoded protein might be overproduced and compete for spiramycin pathway substrates. The polyketide portion of spiramycin is synthesized by condensation of short-chain carbon units, including acetate. The OrfPS AT domain closely resembles AT domains thought to incorporate acetate units (Haydock et al., 1995). This result strengthens our hypothesis: in liquid medium, antibiotic biosynthesis occurs either during stationary phase or at low growth rates (Demain et al., 1983). Since orfPS is overtranscribed during the rapid growth phase in the amplified NSA205, OrfPS might divert precursors necessary for spiramycin production. orfPS is not overtranscribed in a NSA205-derived, de-amplified, spiramycin-producing mutant. In this mutant, and in the parental RP181110 strain, OrfPS would be present in much lower quantities than in NSA205 and would not divert precursors from the spiramycin pathway.

However, we cannot exclude the possible involvement of OrfPS in fatty acid biosynthesis. Many similarities exist between type I PKS and type I FAS (Cortes et al., 1990; Donadio et al., 1991).

ACKNOWLEDGEMENTS

We are very grateful to C. Richard Hutchinson for the pKS2/M5 plasmid and Keith F. Chater for his critical reading of...
the manuscript. We thank the Ministère de la Recherche et de l’Enseignement Supérieur for supporting B.A., C.M. and D.V. We thank also the Ligue Nationale Française Contre Le Cancer for its financial support to D.S.

REFERENCES


B. AIGLE and OTHERS


Received 14 March 1996; revised 30 May 1996; accepted 4 June 1996.