Limited regions of homology between linear and circular plasmids encoding methylenomycin biosynthesis in two independently isolated streptomycetes

Masayuki Yamasaki, Yasuhiro Ikuto, Akiyo Ohira, Keith Chater and Haruyasu Kinashi

Correspondence
Haruyasu Kinashi
kinashi@hiroshima-u.ac.jp

1Department of Molecular Biotechnology, Graduate School of Advanced Sciences of Matter, Hiroshima University, Higashi-Hiroshima 739-8530, Japan
2John Innes Centre, Norwich Research Park, Colney, Norwich NR4 7UH, UK

INTRODUCTION

Biosynthetic genes for any particular antibiotic in Streptomyces usually form a gene cluster located on the chromosome. However, in a few cases the gene cluster is located on a large plasmid. Three examples have been studied in some detail, all involving linear plasmids: SCP1 for methylenomycin in Streptomyces coelicolor A3(2), pSLA2-L for lankacidin, lankamycin and an unknown type-II polyketide in Streptomyces rochei (Kinashi et al., 1994, 1998; Suwa et al., 2000; Mochizuki et al., 2003) and pPZG103 for oxytetracycline in Streptomyces rimosus (Pandza et al., 1997, 1998). Among these, SCP1 has been studied most extensively. Genetic evidence showed that the mmy gene cluster was located on a then unisolated plasmid, SCP1 (Kirby & Hopwood, 1977; Chater & Bruton, 1985). Physical analysis revealed that SCP1 is a 350 kb linear plasmid (Kinashi et al., 1987; Kinashi & Shimaji-Murayama, 1991) and located the mmy gene cluster on its EcoRI map (Redenbach et al., 1998). SCP1 has been found to interact with the linear chromosome of S. coelicolor A3(2) and to integrate into it at several different positions (Vivian & Hopwood, 1970, 1973; Hopwood & Wright, 1976; Hanafusa & Kinashi, 1992; Kinashi et al., 1993; Yamasaki et al., 2001). Recently, the complete nucleotide sequence of SCP1 was determined (GenBank accession numbers AL590463 and AL590464).

Haneishi et al. (1974) reported the original isolation and structure of methylenomycins from Streptomyces violaceoruber SANK95570. Okanishi et al. (1980) detected pSV1, a large circular plasmid, in SANK95570 by electron microscopy. Aguilar et al. (1982) detected the same plasmid by conventional agarose gel electrophoresis and implicated it in methylenomycin production by its hybridization to the methylenomycin resistance (mmr) gene. Although the size of pSV1 was indicated to be about 170 kb, based on its mobility in a conventional agarose gel, its physical properties have not been studied well. A 70 kb linear plasmid, pSV2, in SANK95570 appeared not to be involved in methylenomycin production because mmr did not hybridize to it (Kinashi et al., 1993).

The location of the mmy gene cluster on three different DNA elements, linear and circular plasmids and linear chromosome, is an unusual phenomenon, indicating an interesting degree of mobility of at least this set of antibiotic production genes without any precedence. To investigate this further, we constructed a precise physical map of pSV1 and compared homology between pSV1 and SCP1.

METHODS

Bacterial strains, plasmids and cosmids. S. violaceoruber SANK95570, a producer of methylenomycins (Haneishi et al., 1974), was
provided by T. Haneishi and used for construction of a cosmid library of pSV1 in the Supercos-1 vector (Evans et al., 1989; Stratagene). The ordered cosmid library of SCP1 (Redenbach et al., 1998) was used for physical comparison of pSV1 and SCP1.

**DNA manipulations and PFGE conditions.** DNA samples for PFGE were prepared as described by Kinashi (1994) and Lezhava et al. (1995). We used contour-clamped homogeneous electric fields (CHEF) (Chu et al., 1986) for PFGE, which was conducted at 14 °C in 0.5 x TBE using 1-0 % agarose gel. SCP1 DNA was extracted from gels by electroelution. Digestion of DNA in gels with restriction endonucleases was carried out in reaction buffer containing 100 μg bovine serum albumin ml⁻¹. MidRange I PFGE marker (New England Biolabs) was used as a size marker for CHEF electrophoresis.

**Construction of cosmid library and cosmid walking.** A cosmid library of total DNA of *S. violaceoruber* SANK95570, which carries pSV1, was constructed. Total DNA was partially digested with Sau3AI to a mean size of 40–60 kb, ligated to Supercos-1, packaged into λ phage using a packaging extract (Stratagene) and transfected into *Escherichia coli* SURE2. Supercos-1 possesses T3 and T7 promoters flanking a BamHI cloning site, which can be used for preparation of RNA transcripts from both ends of the insert. DNAs of cosmid clones were isolated, restricted with *Pvu*II, non-radioactively labelled with digoxigenin-11-dUTP (Roche Diagnostics) using T3 and T7 RNA polymerases, separately, and used for cosmid walking.

**Southern hybridization.** DNA fragments were separated by CHEF or conventional agarose electrophoresis and transferred to nylon membrane filters by the capillary method. Hybridization was carried out using the DIG system (Roche Diagnostics) overnight at 70 °C in standard buffer according to the supplier’s protocol. After hybridization, washing was done twice for 5 min each in 2 x wash solution at room temperature, and then twice for 15 min each in 0-1 x wash solution at 70 °C.

**Nucleotide sequencing.** Nucleotide sequencing was performed by the dyeoxy termination method using a dye terminator cycle sequencing kit (Amersham Pharmacia Biotech) and the ABI373 sequencing system (PE Biosystems). Genetyx-Mac 10.1 (Software Development, Tokyo, Japan) was used for analysis of the sequence data.

**RESULTS AND DISCUSSION**

**Physical map of pSV1**

Using conventional agarose gel electrophoresis, we could not obtain enough pSV1 DNA for analysis, due to its low copy number. We therefore constructed a cosmid library from total DNA of SANK95570, using the vector Supercos-1. One thousand two hundred cosmid clones were obtained and then screened with the 10-0 kb *Eco*RI fragment of SCP1, which is located in the centre of the *mmy* gene cluster (Chater & Bruton, 1985; Redenbach et al., 1998). Sixteen clones showed positive hybridization, among which cosmid 7C2 was selected as a starting clone for walking, because it carried the intact 10-0 kb *Eco*RI fragment.

The left and right end probes of cosmid 7C2 were prepared as RNA transcripts and used for hybridization against colony blots of all 1200 clones to identify neighbouring cosmids. By successive rounds of hybridization, six cosmids, 7C2, 12D7, 12A4, 6C11, 7A12 and 9B1, were unambiguously aligned on a circular map (Fig. 1; shown here in a linear form, for ease of representation and comparison). To confirm the circularity of pSV1 further, total DNA of SANK95570 in gels was digested with *Spe*I and *Nde*I, which cut pSV1 once and twice, respectively, separated by CHEF and analysed by Southern hybridization. When probed by a mixture of the six ordered cosmids, undigested total DNA gave a signal at the origin (Fig. 2, lane 2), because large circular DNA molecules are retarded during PFGE (Levene & Zimm, 1987). On the other hand, total DNA digested with *Spe*I, *Nde*I and *Spe*I plus *Nde*I gave signals at 163 kb (lane 3), 148 and 15 kb (lane 4), and 133 and 15 kb (doublet) (lane 5), respectively. This result confirmed that pSV1 is a 163 kb circular plasmid.

To construct a more detailed physical map of pSV1, we used the partial digestion method previously applied to SCP1 (Redenbach et al., 1998). Each of the six cosmid clones was
first completely digested with AseI and then partially with EcoRI. The partial digest was separated by CHEF electrophoresis and hybridized with the end probes of Supercos-1 (0.9 and 4.3 kb AseI–EcoRI fragments). The size differences of the ladder of hybridizing signals determined the order of EcoRI fragments. As a result, all of the EcoRI recognition sites were aligned on a circular map together with the SpeI and NdeI sites (Fig. 1). The total size of EcoRI fragments was calculated to be 162.5 kb, which agrees with the size determined by the SpeI and NdeI digestion.

**Homology between pSV1 and SCP1**

The circularity of pSV1 raised the possibility that pSV1 was originally generated by circularization of a progenitor of SCP1 or vice versa. To test this possibility, the homology between pSV1 and SCP1 was studied. Labelled SCP1 DNA was used to probe EcoRI digests of each of the six pSV1 cosmids. As shown in Fig. 3(a), cosmids 7C2 and 12D7 gave strong hybridizing signals, cosmids 6C11 and 7A12 moderate signals and cosmid 12A4 weak signals. By referring to the EcoRI map of pSV1, the hybridization was located on three separate regions of pSV1 as indicated by shading in Fig. 1; strong hybridization to three central EcoRI fragments (8.8, 10.0 and 10.0 kb; region I), moderate hybridization to the junction of two left fragments (13.7 and 41.0 kb; region II) and weak hybridization to the junction of two right fragments (11.9 and 19.4 kb; region III).

Next, in the opposite direction, the mixture of the six ordered cosmids of pSV1 was hybridized to each of the 11 ordered SCP1 cosmids (Redenbach et al., 1998) digested with EcoRI. As shown in Fig. 3b, SCP1 cosmid 35 gave signals at 9.0 and 8.5 kb, and cosmid 73 gave signals at 11.0, 10.0, 9.0 and 2.7 kb in addition to a common vector signal at 6.8 kb. Cross-hybridized regions were analysed in more detail. Region I, which carries the mmy gene cluster and its flanking regions, hybridized to the 11.0, 10.0 and 9.0 kb EcoRI fragments of SCP1 cosmid 73 (Fig. 1). Region II hybridized to the 9.0 and 8.5 kb fragments of SCP1 cosmid 35, both of which are located at the left side of the mmy gene cluster in SCP1. It should be noted that region II is located far from region I in pSV1. Region III weakly hybridized only to the 2.7 kb fragment in SCP1 cosmid 73, which is located to the right of the mmy cluster. These results ruled out the possibility that pSV1 was generated by circularization of SCP1, or alternatively that SCP1 was formed by linearization of pSV1.

**Analysis of homology region I**

The complete nucleotide sequence of SCP1 has been determined recently (GenBank accession numbers...
AL590463 and AL590464), revealing the plasmid size (356–0 kb) and further information on the organization of the mmy gene cluster (accession number AJ276673) in relation to the flanking sequences. To define the homology between pSV1 and SCP1 more precisely, regions I–III in pSV1 were analysed by finer restriction mapping and finally by nucleotide sequencing.

As shown in Fig. 4(a), the correspondence of the relative positions of restriction sites for EcoRI, PstI, SacI and Sall in pSV1 and SCP1 suggest that homologous region I extends for 25 kb around the mmy gene cluster. The left (7–3 kb) and right (2–7 kb) sides of the mmy gene cluster in pSV1 were subcloned and sequenced (DDJB/EMBL/GenBank accession numbers AB099705 and AB099706). The homologous region extended from nt 228 040 to nt 254 223 in SCP1. Thus, in addition to the genes considered to belong to the mmy gene cluster (ORF228c–ORF246), ORF225–ORF227c to the left of the cluster and ORF247c–ORF252c to the right are conserved in both plasmids, although two additional genes, SCP1.250 and SCP1.251, are present in SCP1 (we have used the same ORF names for pSV1 and SCP1, because the homologous regions showed almost identical sequences except for the regions of insertion or deletion). Small insertions were also found in SCP1.226A in SCP1 and on the left side of ORF225 in pSV1. The 5’-terminal sequences of SCP1.252c in SCP1 and ORF252c in pSV1 are different, because the homology ends inside the 5’ terminus of SCP1.252c. All of the ORFs in the mmy flanking regions encode hypothetical proteins of unknown function.

**Analysis of homology regions II and III**

Homology to region II of pSV1 extends from nt 217 448 to nt 228 591 in SCP1 (Fig. 4b; accession numbers AB099848 and AB099849). Thus, in SCP1 the regions of homology to pSV1 regions I and II actually abut. The rightmost 552 bp of region II of pSV1 is a direct repeat of the leftmost end of region I. This non-coding DNA (named DR1) is duplicated with a spacing of 42 kb in pSV1. This result suggests that the intervening 42 kb DNA segment flanked by dnaE (ORF224c) and ORF225 might have been inserted into pSV1 or excised from SCP1. The 10–5 kb homologous region (region II) of SCP1 contains four ORFs with predicted functions, transposase (SCP1.220c, pseudogene), parA2 (SCP1.221), parB2 (SCP1.222) and dnaE (SCP1.224c), and five hypothetical ORFs (SCP1.217, SCP1.217Ac, SCP1.218c, SCP1.219 and SCP1.223c). The juxtaposition of a transposase gene (SCP1.220c) and a 17-0 kDa protein gene (SCP1.219) reminded us of a similar alignment of a transposase gene and a 13-3 kDa protein gene, which has a subsidiary function in transposition, on the mini-circle

---

**Fig. 4.** Comparison of gene organization and restriction sites in homology region I (a) and homology region II (b) in pSV1 and SCP1. ORFs in SCP1 are depicted based on their complete nucleotide sequence. The positions of ORFs in pSV1 were based on nucleotide sequencing and restriction analysis. The mmy genes are indicated by black arrows and other genes by grey arrows. The central regions of the mmy gene cluster are omitted, while the inserted and deleted regions are shown by grey bars and dotted lines. E, EcoRI; P, PstI; S, SacI; Sc, Sall; Sp, Sphi.
(IS117) in *S. coelicolor* A3(2) (Lydiate et al., 1986; Henderson et al., 1989). The 17-0 kDa protein shows 32% identity to the 13-3 kDa protein of IS117 and the transposase shows 84-9% identity over 73 aa to another transposase in *S. coelicolor* A3(2) (AL557524, SCO4344).

The size of the hybridizing area in region III is at most 0.5 kb in length and the corresponding region in SCP1 lies in a hypothetical coding sequence (SCP1.254). We did not analyse this region in more detail.

### Implications of the homology analysis

Our analysis suggests that the *mmy* gene cluster (region I) and the *parAB* region on its left side (region II) were co-transferred as a set between progenitors of pSV1 and SCP1. At first, we considered that in the transfer event, the transposase (ORF220c) and the 17-0 kDa protein (ORF219) played some part as in the case of IS117 in *S. coelicolor* A3(2). Unlike most other transposable elements, IS117 lacks extensive terminal inverted repeats and does not give rise to duplications of target site sequences on integration (Henderson et al., 1990; Smokvina & Hopwood, 1993), which agrees with the calculated retrospectively to be about seven per chromosome based on the data presented by Kinashi & Shimaji-Murayama (1991) and the determined chromosomal size (8-7 Mb) of *S. coelicolor* A3(2) (Bentley et al., 2002). Therefore, *parAB* in SCP1 may be the partition genes for a linear plasmid, and *parA2B2* in SCP1 and *parAB* in pSV1 for a circular plasmid.

Taking these data together, a very probable sequence of transfer events may be as follows. The *mmy* gene cluster and the *parAB* region were co-transferred as a set from pro-pSV1 to pro-SCP1, which already carried the *parAB* genes for a linear plasmid. After this, a circular DNA species containing a 552 bp DR1 sequence was integrated into pro-pSV1 by a single crossover to generate a 42 kb insert flanked by the 552 bp direct repeats in pSV1. The transfer direction of the *mmy-parAB* segment from pro-pSV1 to pro-SCP1 was further supported by the following data. The protein product (610 aa) of SCP1.216Ac at the left side of the *mmy-parAB* segment in SCP1 shows 93-0% identity with SCP1.136 helicase (879 aa), but its possible N-terminal 283 aa are truncated just at the left junction of homology region II (Fig. 5a). On the right side of the junction, 17 bp of DNA, that forms a fusion gene in SCP1.216Ac, and SCP1.217 are located, and their nucleotide sequences are identical with those of pSV1 (Fig. 5b). Thus, the transfer of the *mmy-parAB* segment from pro-pSV1 to pro-SCP1 might have truncated the original SCP1.216Ac. We expected the truncated DNA at the opposite side of the *mmy-parAB* segment in SCP1, but could not find it at the 5’ terminus of SCP1.252c.

### Note

During the preparation of this manuscript, Spatz et al. (2002) published a paper describing a circular restriction map of pSV1.
ACKNOWLEDGEMENTS

We are indebted to T. Haneishi for S. violaceoruber SANK95570. This work was supported by Grant-in-Aid for Scientific Research on Priority Areas (C) ‘Genome Biology’ from the Ministry of Education, Culture, Sports, Science and Technology of Japan. M. Yamasaki has been supported by a scholarship from the Japan Society for the Promotion of Sciences.

REFERENCES


