Transposition of IS117 of *Streptomyces coelicolor* A3(2) in *Mycobacterium smegmatis*

Apoorva Bhatt and Tobias Kieser

Author for correspondence: Apoorva Bhatt. Tel: +44 1603 452571. Fax: +44 1603 456844. e-mail: bhatta@bbser.ac.uk

Department of Genetics, John Innes Centre, Norwich Research Park, Colney, Norwich NR4 7UH, UK

Derivatives of IS117, the *Streptomyces coelicolor* A3(2) 2.6 kb minicircle, transpose efficiently in *Mycobacterium smegmatis*, targeting chromosomal sites resembling translation start signals. Two IS117 derivatives, plJ4696 and plJ4697, containing a *Streptomyces* hygromycin-resistance gene in opposite orientations were introduced into *M. smegmatis* by electroporation and found to integrate into one of three specific sites. Integrations at sites A and B were frequent while integration at site C was observed only once. Only one site was occupied in each transformant. Sites A and B had either single or tandem integrations. PFGE analysis located these sites on different genomic *Asel* fragments. The sequences of the chromosome-IS117 junctions confirmed that integration was via the same IS117 attachment site as in *Streptomyces*, that there was no target site duplication, and that the orientation of IS117 at each site was fixed. In contrast to the situation in *Streptomyces lividans*, no deletions were created by the transposition and no circular forms could be detected. Comparison of the three *M. smegmatis* chromosomal IS117 target sites (att<sub>s</sub>) with known primary and secondary *S. lividans* att<sub>p</sub> sites showed that only a 2 bp ‘AG’ sequence at the crossover point was conserved. Dividing the att<sub>s</sub> sites into two groups produced two longer consensus target sites, GtcAAGg and gCCGATAGg. Most of the IS117 target sites resemble translational start sites, and site C resembles strongly the amino-terminal sequence of a *Mycobacterium tuberculosis* aminopeptidase. The level of hygromycin resistance in the transformants was high and independent of the site of integration, the number of copies integrated, or the orientation of the *hyg* gene. plJ4696 at all three sites was stable in *M. smegmatis* in the absence of selection for at least 60 cell divisions. plJ4696, plJ4697 and other IS117 derivatives are promising vectors for the stable, integrative cloning of genes in *M. smegmatis*.

**Keywords:** 2.6 kb minicircle, hygromycin resistance, site-specific integration, consensus target sites, PFGE

INTRODUCTION

The 2527 bp element IS117 of *Streptomyces coelicolor* A3(2) forms a covalently closed circular transposition intermediate, the ‘2.6 kb minicircle’ (Henderson et al., 1989; Lydiate et al., 1986). It encodes a transposase gene which is similar to those of IS110 (Bruton & Chater, 1987) and IS116 (Leskiw et al., 1990) of *Streptomyces*, IS900 of *Mycobacterium paratuberculosis* (Green et al., 1989), IS492 of *Pseudomonas atlantica* (Bartlett & Silverman, 1989), IS1000 of *Thermus thermophilus* (Ashby & Bergquist, 1990) and several putative transposases of mycobacteria (Kunze et al., 1991; Moss et al., 1992), *Rhodococcus* (Kulakova et al., 1997), *Coxiella* (Hoover et al., 1992), *Moraxella* (Fulks et al., 1990), *Agrobacterium* (Zhu & Winans, 1998), *Rhizobium* (Freiberg et al., 1997) and *Yersinia* (Rakin & Heesemann, 1995). IS117 and its relatives lack terminal inverted repeats and do not create target site duplications. *Streptomyces lividans*, a close relative of *S. coelicolor*, lacks IS117 but contains a single preferred attachment site, att<sub>B</sub>, which is identical to one of the

**Abbreviations:** att<sub>p</sub>, bacterial (chromosomal) IS117 attachment site; att<sub>s</sub>, IS117 (minicircle) attachment site.
IS117 attachment sites of *S. coelicolor* (Henderson et al., 1989). The only sequence identity between the IS117 attachment site, attM (‘M’ stands for minicircle), and the attB sites of *S. lividans* and *S. coelicolor* is TAG at the crossover point. Most *S. lividans* transformants contained two adjacent tandem copies of IS117 which originated from independent elements that were introduced into the same cells by co-transformation (Smokvina et al., 1994). Transformation of an *S. lividans* strain lacking attB gave integration into secondary attachment sites at a reduced frequency (Henderson et al., 1990). Interestingly, integration into secondary IS117 attachment sites produced large chromosomal deletions extending from the sites of insertion (Smokvina & Hopwood, 1993).

Streptomyces are taxonomically related to the mycobacteria: both are high G+C, Gram-positive actinomycetes. *Streptomyces* transposable elements are thus likely to function in mycobacteria. Here we show that IS117 transposes into at least three chromosomal attachment sites in *Mycobacterium smegmatis* mc²155, two of which are used preferentially. The DNA sequences of the sites resemble those of the primary and secondary *Streptomyces* attachment sites. Integration did not cause deletions in *M. smegmatis*. The integrated elements were maintained stably without selection, suggesting that IS117 derivatives may be useful as shuttle cloning vectors for *Mycobacterium* and *Streptomyces*.

**METHODS**

**Bacterial strains, transformation and media.** The *Mycobacterium smegmatis* strain mc²155 (Snapper et al., 1990) was grown in 7H9 broth (Difco) containing 0.05% Tween 80 and 5% ODAC supplement (Difco; ODAC contains oleic acid, dextrose, albumin and catalase). Electroporation of mc²155 was done as described by Snapper et al. (1988, 1990). After the 2500 V electric pulse, the cells were diluted in 5 ml supplemented 7H9 broth and incubated with shaking for 2 h at 37 °C to allow expression of the resistance gene. Antibiotic-resistant transformants were selected on 7H11 agar (Difco) containing 5% ODAC, 0.5% glycerol and 10 μg kanamycin sulphate ml⁻¹ (Sigma) or 100 μg hygromycin ml⁻¹ (a gift from E. T. Seno, Eli Lilly, Indianapolis). For cloning in *Escherichia coli*, the strain DH5α (Woodcock et al., 1989) was used as host. Transformants were selected on L agar containing 50 μg carbenicillin ml⁻¹.

**Extraction of genomic DNA.** Genomic DNA was extracted from *S. smegmatis* mc²155 by the salting-out method (Pospiech & Neumann, 1995).

**PFGE.** Agarose plugs containing *M. smegmatis* cells were prepared, digested with Asel, and fractionated by PFGE as described for *S. coelicolor* (Kieser et al., 1992). The 0.5 × Tris/borate/EDTA (TBE) electrophoresis buffer needed to be supplemented with 50 μM thiourea (Sigma) to obtain sharp, high-molecular-mass bands (K. Dharmalingam, personal communication). The gels were subjected to electrophoresis for 24 h at 6 V cm⁻¹ and 14 °C. The pulse times were 5–25 s (linear ramp) at an angle of 120°.

**Southern hybridization.** Digested chromosomal DNA was fractionated on 0.8% agarose gels in TBE buffer and transferred onto Hybond-N membrane (Amersham) and hybridized as described in Sambrook et al. (1989). The final wash was in 0.1 × SSC at 67 °C.

**PCR amplification.** IS117 integration (attM) sites were PCR-amplified from non-transformed *M. smegmatis* genomic DNA in a Stratagene Robocycler using the following primers: A1 (5'-CCGGATCTCTTCTGCCGTGCGG-3') and A2 (5'-GGATCCGTTGAAGGTACTCGGCAG-3') for site A; B1 (5'-GAATTCGGTACGCGTACGCA-3') and B2 (5'-GGATCCGGTGCTACGCGTCTG-3') for site B; C1 (5'-GGATATGTAGATGATGCTCCGGAAATCG-3') and C2 (5'-GAAGTCTTGTGGTCGACCCGGAAAAGT-3') for site C. Boehringer Mannheim *Tag* DNA polymerase and buffer were used in the reaction with the following conditions: denaturation at 96 °C for 1 min, annealing at 55 °C (for site C) or 60 °C (for sites A and B) for 2 min, and extension at 72 °C for 1 min; total cycles = 30. The reaction was preceded by a hot start (96 °C for 5 min) and was finished with a final extension reaction for 10 min.

**DNA sequencing.** Plasmid DNA for double-strand sequencing was purified using the QIAamp plasmid purification kit. PCR products for sequencing were extracted from agarose gels using the QIAquick gel extraction kit. Sequencing reactions were done using the Amersham Thermo Sequenase dye terminator sequencing kit and the oligonucleotide primers M1 (5'-CAGCCGGCCGCTTCAGAGAGACAG-3') and M2 (5'-TCATGAGCACACCCTGGCAGCAGC-3').

**RESULTS AND DISCUSSION**

**Introduction of IS117 derivatives into *M. smegmatis* by electroporation**

Two IS117 derivatives, plJ4696 and plJ4697, were used (details in Fig. 1). They differ only in the orientation of the hyg gene, and they were useful to detect a possible dependence of the hygromycin resistance on external promoters. No hygromycin resistance was detected in E. coli because the *Streptomyces* promoter is not expressed. The two constructs were introduced into *M. smegmatis* mc²155 by electroporation. Similar numbers of hygromycin-resistant colonies were obtained with plJ4696 and plJ4697, and no hygromycin-resistant colonies appeared in the transformation control without DNA. The autonomously replicating pAL5000 derivative pYUB12 (Snapper et al., 1988) was used as an internal control for the transformation frequency. In four independent experiments using 1 μg plJ4696 and 0.5 μg pYUB12, the transformation frequency of plJ4696 was between 4 and 15% of that of pYUB12 [0.5 × 10⁴ transformants (μg DNA)⁻¹]. In *S. lividans* the transformation frequency of a minicircle derivative was only 0.2% of that of an autonomously replicating plasmid (Henderson et al., 1989).

**IS117 integrated into three specific sites in *M. smegmatis***

Total DNA from 10 plJ4696 and 10 plJ4697 transformants was digested with EcoRI, fractionated on an agarose gel and probed with random prime labelled plJ4696 or plJ4697 DNA. Each plasmid has two EcoRI sites, one of which is in the hyg fragment (Fig. 1). Digestion of the circular plasmids gave two EcoRI...
of pBR327 (at nucleotide 2873 there is neither a site of IS1 of the fragment containing the Bglll attachment site). plIJ4696 and plIJ4697 contain two Asel sites, making it easy to localize the insertion points. Results from the ethidium-bromide-stained gels were confirmed by probing with the labelled plasmids (data not shown). Site A is in an approximately 225 kb Asel fragment of S. lividans, producing new flanking fragments of approximately 75 kb and 150 kb. Sites B and C are in two different approximately 340–350 kb Asel fragments. New flanking bands of approximately 90 and 250 kb were generated for insertions in sites B and C, respectively.

Levels of hygromycin resistance

The strains were tested for their levels of antibiotic resistance by spot-inoculating 10-fold dilutions of cultures onto 7H11 plates containing 0, 5, 10, 20, 40, 80, 100, 200, 300 or 400 µg hygromycin ml⁻¹. Single colonies of untransformed mc2155 failed to grow on 40 µg hygromycin ml⁻¹ and a dense patch inoculated with approximately 10⁵ cells failed to grow on 40 µg hygromycin ml⁻¹, irrespective of the site of integration, the orientation of the hyg gene or the copy number. Therefore, it is unlikely that the observed target-site specificity was influenced by the selection of transformants on 100 µg hygromycin ml⁻¹.

plIJ4696 is stably maintained in M. smegmatis

Starting with a single hygromycin-resistant colony, one plIJ4696 transformant of each type (single integrations in each site and tandem integration in site A) was propagated nonselectively for 60 doublings in 7H9 broth. Each culture was then plated for single colonies on 7H11 agar. All 200 tested colonies (50 from each transformant) were hygromycin resistant. By comparison, the autonomously replicating pAL5000 derivative was lost from >70% of the culture under the same conditions (England et al., 1991).

Circular forms of IS117 were not detected in M. smegmatis containing single insertions

Circular transposition intermediates occur at a low but readily detectable frequency of one CCC molecule per 10–50 chromosomes in Streptomyces (Henderson et al., 1989; Lydiate et al., 1986). In M. smegmatis trans-
formants, they could not be detected in Southern blots of digested (Fig. 2) or undigested (data not shown) total DNA. Oligonucleotide primers MC1 and MC2, reading from both sides towards attM, were used for PCR amplification. This produces a 265 bp fragment both from circular forms and also from tandem copies. DNA from a transformant with tandem integrations which was used as a positive control gave a band of expected size. Low yields of products of different sizes were obtained from the transformants with single insertions at sites A, B and C. One of these faint bands co-migrated with the 265 bp positive control but after purification from the gel and reamplification, it gave a different HaeIII restriction pattern from the control. Therefore, circular IS117 derivatives could not be detected in M. smegmatis containing single insertions. This suggests that once IS117 is integrated excision is much less frequent than in Streptomyces. This may also explain why the elements did not transpose from the original insertion sites to the other sites to give multiple insertions.

Cloning of the M. smegmatis IS117 integration sites

pIJ4696 has the pBR327 replication origin and the bla gene for selection in E. coli, and no Asp718 site. It was therefore possible to excise the plasmid together with flanking sequences by digestion of total DNA with Asp718. The digested samples were re-ligated at low DNA concentration (200 ng ml⁻¹) and used to transform E. coli DH5α. Plasmids isolated from these transformants contained, as expected, a single Asp718 site. The sizes of the plasmids containing tandem copies confirmed that two copies were present. No trans-
IS117 transposition in Mycobacterium smegmatis

(a)

S62247  ATGCGGTAAGCCACATGAGG
S62248  CCGCTCTGCGACGCCCAGG
S62250  GGGCCGCTCCCGCAGG
S62242  GACGGGCATGGG
S62243  CGGCTGTCCGCGAGCC
S62245  CTGTCGAGGTCGCTCC
S62249  GCCACGTGAGCTCAGG

Site A  GCCACGTGAGCTCAGG  AAGGGG

Consensus  ggc--tgtgctggccccccc

(b)

Slp site  GCCCGGTACGGCCAGAGA
S62244  GCCACAGATTTTCTGAAG
S62241  GTTCAGGATCCGCTTGAG
Site B  GCGGCGCTGGCCAGG
Site C  GCGGCGCTGGCCAGG

Consensus  g--g--g--g--g--g

**Fig. 3.** Comparison of the *S. lividans* and *M. smegmatis* IS117 integration sites. Alignment of sequences of: (a) *M. smegmatis* site A and *S. lividans* secondary integration sites; (b) *M. smegmatis* sites B and C and *S. lividans* preferred (Slp) and secondary integration sites. The sequences with EMBL accession numbers S62241–50 are *S. lividans* secondary integration sites. The sequences in boxes are the crossover points; putative RBSs (sequences with at least 3 bp homology to the 3' end of 16S rRNA) are underlined; and all ATG and GTG start codons are double-underlined. Bases identical in 50% or more of the sequences are indicated in bold.

formants were obtained from the single insertion in site C. It was, however, possible to clone the right and left border sequences separately using BamHI and MluI, respectively. It is noteworthy that none of the transformants, not even those from strains containing tandem insertions, contained the original plasmid that might have excised naturally from the *M. smegmatis* chromosome.

**Sequences of the IS117 integration sites**

The above plasmids and two 24-mer primers, MCI and MC2, reading from each side towards att _λ_ were used for sequencing. The sequences confirmed that, as in *S. lividans*, integration was via the known att _λ_ site, that there was no target site duplication, and that the orientation of IS117 at each site was fixed. Fig. 3 shows an alignment at the crossover point of the three *M. smegmatis* target sites with those already known from *S. lividans* (Smokvina & Hopwood, 1993). The sequence ATG to the right of the crossover point is conserved in all 14 sequences. The sequence similarity extends to the left if the sequences are subdivided into two groups. Group A, which includes site A, has the consensus GtcAAGg and Group B, which includes sites B and C, has the consensus gCCGATAGg. These sequences probably occur randomly several times in the *M. smegmatis* genome. Either most of such sites are in essential regions or these short consensus sequences do not alone specify target sites. Compare-Dotblot sequence analysis (UWGCG package; Devereux et al., 1984) to detect direct and inverted repeats did not reveal any features that were similar in all or most of these sequences.

The IS117 target sequences (att _λ_ ) were compared with the sequence databases. The sequence downstream of the crossover point at site C matches the translation initiation region and amino-terminus of the *Mycobacterium tuberculosis* aminopeptidase gene (pepD), which also has a homologue in *S. lividans*. A putative ribosome-binding site (RBS) is present immediately downstream of the crossover point (underlined in Fig. 3). The *S. lividans* att _λ_ sites also have putative RBSs and GTG start codons. No perfect database matches were found for sites A and B, but both have a possible GTG start codon to the right of the crossover point. The G + C content of the three target sites (100 nucleotides) is 71–72%, only slightly higher than the 67% average for *M. smegmatis*. High G + C DNA contains few out-of-frame stop codons but protein coding regions can be recognized reliably because most amino acid codons containing G or C in positions one and three are used preferentially (Bibb et al., 1984). All the sites seem to be in untranslated regions. It is therefore likely that IS117 targets, at least partly, translation initiation signals, as was described by Doran et al. (1997) for IS900, which has a similar transposase.

**IS117 integration into *M. smegmatis* did not cause deletions**

Insertion of IS117 into secondary att _λ_ sites in *S. lividans* frequently caused deletions at the target site (Smokvina & Hopwood, 1993). The deletions may have arisen from two IS117 derivatives transposing into the chromosome a few kilobases apart but in the same orientation. Efficient recombination between the two elements would delete the intervening (non-essential) DNA and produce the observed strains containing IS117 flanked by non-contiguous DNA. To test for deletions in *M. smegmatis* transformants, the distance between the sequences that flank inserted IS117 was estimated in the non-transformed strain using PCR amplification with primers homologous to the left and right flanking sequences. The primers used are described in Methods. The sizes of the observed amplification products were as predicted from the sequence, and sequencing of the PCR products proved that they did contain the att _λ_ sites A, B or C. Therefore, IS117 did not cause deletions in *M. smegmatis*.

**Concluding remarks**

IS117 always integrated in the same orientation into one of three sites in *M. smegmatis*, and the hygromycin resistance was independent of the orientation of the gene or the site of insertion. This, and the high stability in the absence of antibiotic selection, make pIJ4696 and
plJ4697, which contain unique BamHI cloning sites, promising vectors for cloning foreign genes in M. smege,
mas, and possibly other mycobacteria. Other potentially useful IS117 vector derivatives have also been described (Motamedi et al., 1995). IS900 from M. paratuberculosis, which integrates into several genomic sites in M. smearm, has also been used as a stable integrative vector (Dellagostin et al., 1993; England et al., 1991). The relative transformation frequencies obtained with plJ4696 were 4–15% of that of the autonomously replicating pYUB12, which compare favourably with the values of 0.02–0.05% reported for the IS900 derivatives.

Surprisingly, there are two different IS117 attB consensus sequences of only 7–9 bp in length. Since these sites would be expected to occur several times in the M. smearm genome, it is likely that there are additional sequence features that are recognized by the IS117 transposase. No perfect direct or inverted repeats are present in all the attB sequences. It seems, however, that IS117 has a partial preference for translational start sites as was observed previously for IS900 (Doran et al., 1997). IS900 has the sequence AAGGA at its 3’ end which could function as a RBS while in IS117 the sequence GGA is present 1 bp upstream of the crossover point.

ACKNOWLEDGEMENTS
We thank Mervyn Bibb, Mark Buttner, Keith Chater, David Hopwood and Mike Merrick for their helpful comments on this manuscript. A.B. is supported by a PhD studentship from the John Innes Foundation, Norwich, UK.

REFERENCES
IS117 transposition in Mycobacterium smegmatis


Received 26 October 1998; revised 27 January 1999; accepted 2 February 1999.