Targets for pSAM2 integrase-mediated site-specific integration in the *Mycobacterium smegmatis* chromosome

Asunción Seoane, Jesús Navas and Juan M. García Lobo

Author for correspondence: Juan M. Garcia Lobo. Tel: +34 42 201948. Fax: +34 42 201945. e-mail: jmglobo@medi.unican.es

An improved integrative cassette from plasmid pSAM2 has been constructed containing plasmid int and attP genes but excluding the xis gene, which should result in increased stability by suppression of the excision reaction. This cassette was included in both suicide and thermosensitive plasmids and used for integration in *Mycobacterium smegmatis*. Suicide plasmids containing this cassette integrated at a single site (attB1) in the *M. smegmatis* chromosome. The sequence of the attB1 site has been determined and was identified as a putative tRNA*<sup>α</sup>* gene. Thermosensitive plasmids containing the cassette integrated both at the same attB1 site and at other different sites, often giving rise to simultaneous integration at two sites. A second integration site (attB2) has been sequenced, which was located in the region encoding 16S rRNA of one of the two rrn operons of *M. smegmatis*.

**Keywords**: *Mycobacterium smegmatis*, plasmid pSAM2, integrase, site-specific recombination

**INTRODUCTION**

Genetic manipulation is a key procedure for the study of mycobacterial biology and for the engineering of the genus to produce either new vaccines or other attenuated strains. Homologous recombination in mycobacteria is a process which is not fully understood, and seems to work better in fast-growing than in slow-growing pathogenic species. Its use for genetic analysis of this genus is also hampered by the high rates of illegitimate recombination (McFadden, 1996). Site-specific recombination in mycobacteria has not been well studied either. However, this kind of RecA-independent process is simpler than homologous recombination and does not usually require extensive sequence homology between the recombinating partners. A single recombinase is sufficient with the assistance of some accessory proteins that are ubiquitous in bacteria. Site-specific recombination may therefore provide a useful approach to the genetic manipulation of mycobacteria.

Mycobacteriophages such as L5 have integrases similar to λ-integrase which have been successfully used for site-specific gene integration in mycobacteria (Lee *et al.*, 1991). Integrative plasmids from *Streptomyces* are capable of autonomous replication but they can also integrate into the bacterial chromosome through a site-specific recombination event (Pernodet *et al.*, 1984). The integrative functions of plasmid pSAM2 have been well characterized (Boccard *et al.*, 1989a; Kuhstoss *et al.*, 1991; Smokvina *et al.*, 1991). A gene that encodes an integrase belonging to the λ-integrase family has been located in pSAM2 (Boccard *et al.*, 1989b). Two identical 58 bp *att* sequences have been found in the plasmid (*attP*) and in the *Streptomyces ambofaciens* chromosome (*attB*). The *attB* site overlaps a putative gene encoding a tRNA*<sup>Pro</sup>* conserved in many actinomyces, including mycobacteria (Mazodier *et al.*, 1990). It is not known whether the complete tRNA gene is needed for integration or whether a part of it could be sufficient. The use of the integrative functions of plasmid pSAM2 has already been exploited for *Streptomyces* vector construction (Smokvina *et al.*, 1990; Kuhstoss *et al.*, 1991). In addition, the presence of a sequence similar to pSAM2 *attB* in the chromosome of mycobacteria suggests the use of this plasmid for the construction of mycobacterial integrative vectors (Eiglmeier *et al.*, 1991). Integration mediated by the pSAM2 integrase has been demonstrated in *Mycobacterium smegmatis* (Martin *et al.*, 1991) and shown to be site specific. Here, with the aim of improving both the stability and frequency of the integration of pSAM2-based plasmids...
into the *M. smegmatis* chromosome, we have developed new integrative plasmids and have analysed their integration sites in the *M. smegmatis* chromosome.

**METHODS**

**Plasmid constructions.** A 1.4 kb MscI–BamHI DNA fragment from the plasmid pTSN39 (Martín et al., 1991) containing the *int* gene and the *attP* site of pSAM2 was ligated to the EcoRI (Klenow filled)/BamHI sites of the vector pUC18. The resulting plasmid was called pLAS3. The kanamycin-resistance gene from Tn903 obtained from Pharmacia (Kanblock) was inserted as a *PstI* fragment to the *PstI* site of pLAS3. This 5.4 kb plasmid was called pLAS4 and was used as a suicide integrative vector for *M. smegmatis*. An Int- derivative of pLAS4 was obtained by introducing a frameshift mutation into its *int* gene by filling the BglII site as described by Boccard et al. (1989b). This Int- plasmid was called pLAS4b.

A thermosensitive *Mycobacterium–Escherichia coli* shuttle plasmid, carrying the *int–attP* integrative cassette from pLAS4, was also constructed. The thermosensitive plasmid pCG63 (Guilhot et al., 1992) was linearized with BamHI (Klenow filled)/SphI and then ligated to a 1.6 kb *PvuII–SphI* fragment from pLAS3 containing the pSAM2 integrative cassette. The resulting plasmid was called pLAS633. An Int- derivative of pLAS633, pLAS633b, was constructed as described above for pLAS4. A restriction and genetic map of these plasmids is shown in Fig. 1.

**Isolation of integration events.** Purified DNA (0.1 µg) of either pLAS4 or pLAS4b was used to electrotransform competent *M. smegmatis* mc²155 cells. Transformants growing at 37 °C on plates containing kanamycin at 50 µg ml⁻¹ should carry the plasmid integrated in the chromosome. Integration of the thermosensitive plasmid pLAS633 was done as described by Guilhot et al. (1994) with slight modifications. Briefly, pLAS633 was electrotransformed into competent *M. smegmatis* mc²155 cells and stable transformants were selected on plates containing kanamycin at 30 °C. Independent transformants were analysed for the presence of the plasmid and grown in 5 ml kanamycin-supplemented 7H9 medium for 48 h at 30 °C. Dilutions were plated on kanamycin-containing plates, which were incubated for 24 h at 30 °C and then shifted to 42 °C until colonies developed. Colonies obtained from these plates should carry pLAS633 integrated in the chromosome.

**Isolation of chromosomal DNA and Southern blot hybridization.** Total DNA from mycobacteria was purified basically as described by Ausubel et al. (1990). Chromosomal DNA to be used as a sequencing template was further purified in CsCl gradients. Total DNA was separated in agarose gels after treatment with restriction endonucleases and blotted on nylon membranes. Probes were labelled with digoxigenin and hybridization detected by a luminescent assay using the DIG DNA labelling and detection kit (Boehringer Mannheim).

**DNA sequencing.** Plasmid and chromosomal DNA sequences were determined with the *Taq* polymerase based *fmol* cycle-sequencing system (Promega), using the primers indicated in the text. Samples of 1.5 pmol primer end-labelled with [*γ-³²P*]ATP and polynucleotide kinase were mixed with either 6 µg chromosomal DNA or 100 ng plasmid DNA. The thermocycling programme included a denaturing step at 95 °C for 2 min, and then 30 cycles each at 95 °C for 30 s, 50 °C for 30 s and 70 °C for 1 min. The sequences obtained were compared to those in GenBank using the BLAST server at the National Centre for Biotechnology Information.

**RESULTS**

**Plasmid integration in the chromosome of *M. smegmatis***

The suicide plasmid pLAS4 was introduced by electroporation into *M. smegmatis* mc²155. The number of colonies obtained was around 150 per µg plasmid DNA. Since this plasmid was unable to replicate in mycobacteria, we assumed that kanamycin-resistant trans-
formants arose as a result of IntpSAM2-mediated integration in the chromosome. This was confirmed by the lack of kanamycin-resistant transformants when the same technique was repeated with the plasmid pLAS4b, an Int- derivative of pLAS4.

To improve the integration frequency we performed a similar experiment with the plasmid pLAS633, containing a thermosensitive mycobacterial replicon. The plasmid was introduced into M. smegmatis mc²155, and kanamycin-resistant transformants were selected at 30 °C, a temperature allowing pLAS633 replication. Colonies able to grow in the presence of kanamycin at 42 °C were selected as described in Methods. With this procedure we routinely obtained around 2000 kanamycin-resistant colonies per ml culture. Experiments similar to those described above were performed with plasmid pLAS633b, an Int- derivative of pLAS633. In this case, however, no kanamycin-resistant colonies could be obtained after the shift to 42 °C.

Hybridization analysis of the integration sites

The colonies obtained in the above experiments were checked for the absence of free replicating plasmids by electrophoretic analysis of plasmid lysates and by hybridization of total DNA, with negative results. Total DNA from six colonies obtained from independent experiments that should carry pLAS4 integrated into the chromosome was digested with PvuII and XbaI, blotted and hybridized with pLAS4. The results, shown in Fig. 2, indicated that the six colonies analysed were identical.

In addition, we found that all colonies contained two of the pLAS4 restriction fragments but the third fragment, containing the attP site, was not present and was replaced by two new fragments formed as a result of the integration event. These results demonstrated that the plasmid was integrated into the chromosome and that integration probably occurred as a result of IntpSAM2-mediated site-specific recombination.

In a similar way, total DNA from nine colonies obtained in the integration experiment with pLAS633 was digested with PstI and hybridized against a 1.3 kb EcoRI–BamHI fragment from pLAS3, containing the int–attP region (identical to pLAS633). The probe hybridized to a 2.1 kb PstI fragment of pLAS633. This fragment did not hybridize in any of the recombinant clones, suggesting that the plasmid was indeed integrated into the chromosome. Nevertheless, the nine colonies analysed...
were not identical. At least four different hybridization patterns could be observed, indicating that more than one integration site was being used in the chromosome. In some cases only two hybridization bands were present, suggesting a single integrative event, while in other cases there were four hybridizing bands, suggesting the presence of two copies of the plasmid at different chromosomal positions (Fig. 3).

To show that the complete plasmid was integrated in the chromosome of these colonies, chromosomal DNA from two recombinants was digested with Smal and probed with pLAS633. The results (Fig. 4) indicated that the whole plasmid was integrated into the chromosome. Analysis of fragment sizes indicated again that the plasmid was probably integrated through the attP site.

DNA sequence of the integration sites

The nucleotide sequence of pLAS4 around the attP site was determined using as sequencing primer the oligonucleotide TAGTCACGCAGATAGACAC, mapping at the 3' side of attP, according to the sequence reported for pSAM2 attP by Boccard et al. (1989b). The second chain was sequenced with the oligonucleotides AGTGTTAACA CTTGCAGGCAAA and AGAAGCACACTGGGCCC AA, deduced from the sequence obtained with the first primer (Fig. 5). The same primers were used to determine the sequence of the attL and attR sequences in the clones carrying the integrated plasmids.

Chromosomal fragments containing the integrated plasmid pLAS4 were cloned taking advantage of the absence of KpnI restriction sites in this plasmid. Chromosomal DNA was digested with KpnI, and the fragments were self-ligated and used to transform E. coli cells to kanamycin resistance. With this method a 7 kb plasmid was obtained, and was used to sequence the pLAS4-
chromosome junctions using the primers described above. The sequences obtained corresponding to the
attR1 and attL1 sites are shown in Fig. 5.

The sequence at the attB1 site was deduced from the
attL1 and attR1 sequences on the basis of the previously
described integration mechanism of pSAM2. The se-
quence was compared to DNA sequences in GenBank. A
region of attB1 was found to be identical to the rRNAPro
from S. ambofaciens and similar to other bacterial
tRNAPro genes. In addition, the analysis of attB1 with
the program tRNAsearch from the PCgene package
revealed that the region between position 87 and
position 160 (74 bases) of attB1 could be a tRNAPro with
the anticodon CGG.

In contrast, integration sites of pLAS633 were deter-
mined by direct sequencing of chromosomal DNA, due
to the lack of restriction enzymes appropriate for the
self-ligation approach. Chromosomal DNA was purified
from the colonies analysed in Fig. 3. An examination of
the sequences revealed that integration in colony 8 had
occurred in the previously described attB1 site. How-
ever, integration in colonies 1, 2 and 3 occurred at a new
site which was identical in all three cases. The attR2 and
attL2 sequences obtained from this site, and the sequence
attB2 deduced from them, are shown in Fig. 5. Com-
parison with sequences in GenBank revealed that the sequence
attB2 was identical to a region of the M.
smegmatis gene encoding the 16S rRNA (Rogall et al.,
1990). Sequencing of DNA from colonies 4, 5, 6, 7 and 9
produced mixed sequencing ladders, confirming that
they carried at least two plasmid copies in the chro-
mosome and precluding the identification of other in-
tegration sites.

**Stability of the Integrated Plasmids**

Loss of the integrated plasmids was investigated through
the loss of the kanamycin-resistance marker after growth
in a nonselective medium. Three clones carrying each of
the integrated plasmids were grown for 56 generations
at either 37°C (pLAS54) or 42°C (pLAS633) and then
plated on an antibiotic-free medium. Two hundred
colonies from each clone were replicated onto kanamy-
cin plates. All the replicas grew in the presence of
kanamycin, indicating that integration was stable, and
plasmid loss was lower than 0.5% after serial passage in
a drug-free medium.

**DISCUSSION**

In this study we have demonstrated that IntpSAM2-
mediated site-specific recombination was active in M.
smegmatis if we introduced a plasmid carrying an
integrative cassette containing only the integrase gene
and the attP site; attB sequences were present in the M.
smegmatis chromosome, and any accessory protein
necessary for the recombination process was provided
by the host.

A different integrative behaviour was observed depend-
ing on the type of plasmid used to harbour the integrative
cassette. Single insertions at an unique chromosomal
location were observed for suicide integrating plasmids.
This integration site coincided with the previously
described pSAM2 attB. Double insertions involving at
least three different integration sites were found when
the integrative cassette was placed in a temperature-
sensitive plasmid. This difference probably resulted
from the time that integrase had to react in the
ciobacterial cells. Non-replicating plasmids are ex-
pected to be diluted within the population, if not lost, as
a result of degradation by the restriction systems of the
host. As a result, the level of expression attained by the
integrase will be high enough to produce only single,
highly specific integrative events at the attB1 site. How-
ever, if thermosensitive plasmids carrying the
integrative cassette were allowed to grow for some time
at the permissive temperature, the integrase had a longer
time for expression and to act in the cells, producing
multiple integrative events involving more than one
integration site. We have sequenced one of the additional
sites and found it in one of the two rRNA operons
present in the M. smegmatis chromosome. This site was
located precisely in the middle of the 16S rRNA gene,
and was called attB2. Integration at the canonical attB1
site was presumably favoured by the existence of 43 bp
of sequence in common between attB1 and attP sites.
Although only three common base pairs were observed
between the sites attP and attB2, these could be enough
to facilitate recombination at the attB2 site as previously
demonstrated for integration of the bacteriophage
φC31 (Kuhstoss & Rao, 1991).

The existence of 43 bp of homology between the attB1
and attP sites prevented determination of the exact
location of the crossing-over in the IntpSAM2 recom-
bination reaction at attB1. However, the observation of
recombination at attB2, with only 3 bp of sequence
homology with attP, allowed us to localize the crossing-
over site to these 3 bp, located in the S' end of attP. If is
assumed that the same site was used in the recom-
bination at attB1, then the crossing-over site in attB1
could also be precisely located. This position was
coincident with the first nucleotide of the anticodon
loop of the tRNA as described for other site-specific
recombination systems that use tRNA genes as targets.
The mycobacteriophage L5 integrates in the chro-
mosome of M. smegmatis through a tRNAPro gene adjacent
to a tRNAPro gene (Lee et al., 1991). Interestingly, this
gene was slightly different from the tRNAPro gene used
for pSAM2 integration.

Plasmid integration at the 16S RNA gene of M.
smegmatis should result in some inactivation of one of
the two rRNA operons present in this species (Bercovier
et al., 1986). However, no evident differences in growth
rate were observed between clones carrying plasmid
pLAS633 integrated at attB2 and clones with insertions
at attB1.

Integrations of suicide and thermosensitive plasmids
were both stable. Excision of pSAM2 has been observed
in Streptomyces lividans (Simonet et al., 1987) and S.
ambofaciens (Kuhstoss et al., 1989). The increased
stability of both pLAS4 and pLAS633 integrations in the M. smegmatis chromosome can probably be explained by the absence of the pSAM2 xis gene in these plasmids. The results reported here represent a new useful tool for the integration of genes into the mycobacterial chromosome and provide further insight into the understanding of the IntpSAM2 mechanism of site-specific recombination.

ACKNOWLEDGEMENTS

We acknowledge that A. Seoane and J. Navas made equal contributions to this work. We are grateful to C. Martin and B. Gicquel for generously providing plasmid pTSN39 and valuable methodological information, and to Ian Williams for the correction of the manuscript. The work described herein was performed under contract no. BIO2 CT92 0520 with the European Union and cofinanced with grant PB93-0386 from the Spanish DGICYT to J.M.G.L.

REFERENCES


Received 21 February 1997; revised 13 June 1997; accepted 18 June 1997.