Microbiology (1997), **143**, 3137–3147 Printed in Great Britain

Structural analysis of the 6 kb cryptic plasmid pFAJ2600 from *Rhodococcus erythropolis* NI86/21 and construction of *Escherichia coli–Rhodococcus* shuttle vectors

René De Mot, István Nagy, Adinda De Schrijver, Pranee Pattanapipitpaisal,† Geert Schoofs and Jos Vanderleyden

Author for correspondence: René De Mot. Tel: +32 16 32 23 98. Fax: +32 16 32 29 90. e-mail: rene.demot@agr.kuleuven.ac.be

F. A. Janssens Laboratory of Genetics, Department of Applied Plant Sciences, Catholic University of Leuven, Willem de Croylaan 42, B-3001 Heverlee, Belgium

The complete nucleotide sequence of the 5936 bp cryptic plasmid pFAJ2600 from *Rhodococcus erythropolis* NI86/21 was determined. Based on the characteristics of its putative replication genes, *repA* and *repB*, pFAJ2600 was assigned to the family of pAL5000-related small replicons identified in *Mycobacterium* (pAL5000), *Corynebacterium* (pXZ10142), *Brevibacterium* (pRBL1), *Bifidobacterium* (pMB1) and *Neisseria* (pJD1). The replication systems of these plasmids show striking similarities to the ones used by the ColE2 family of plasmids from Enterobacteria with respect to both *trans*-acting factors and *ori* sequences. Two possible plasmid stabilization systems are encoded on pFAJ2600: a site-specific recombinase (PmrA) related to the *Escherichia coli* Xer proteins for plasmid multimer resolution and an ATPase (ParA) related to the A-type of proteins in *sop*/*par* partitioning systems. The proposed replication termination region of pFAJ2600 has features in common with the *Ter* loci of *Bacillus subtilis*. Chimeras composed of a pUC18-Cm^r derivative inserted in the *parA–repA* intergenic region of vector pFAJ2600 produced vectors that could be shuttled between *Escherichia coli* and several *Rhodococcus* species (*R. erythropolis*, *R. fascians*, *R. rhodochrous*, *R. ruber*). The pFAJ2600-based shuttle vector pFAJ2574 was stably maintained in *R. erythropolis* and *R. fascians* growing under non-selective conditions.

Keywords: pAL5000-related plasmids, ColE2 plasmid family, site-specific recombinase, partitioning ATPase, *Bacillus* replication terminator

INTRODUCTION

Bacteria from the genus *Rhodococcus* display a remarkable versatility with respect to biodegradation or biotransformation of organic compounds (reviewed by Warhurst & Fewson, 1994). A major factor limiting the exploration of this potential for (environmental) biotechnology, undoubtedly is the lack of appropriate genetic tools (Finnerty, 1992). In recent years, a limited number of *Escherichia coli–Rhodococcus* shuttle vectors have been developed, by using replication regions contained on indigenous *Rhodococcus* plasmids.

A fragment of the cryptic 13 kb plasmid pMVS300 from *Rhodococcus* sp. strain H13-A was used to construct pMVS301 (Vogt Singer & Finnerty, 1988), which was further reduced in size by Shao *et al.* (1995) generating pBS305. These plasmids were stably maintained in *R. erythropolis*, *R. equi* and *R. globerulus* but not in *R. rhodochrous*. Plasmid pBS305 enabled heterologous expression of genes involved in herbicide degradation (Shao & Behki, 1995).

A similar approach was used by Hashimoto *et al.* (1992) to construct the shuttle vector pK4, based on the 2·6 kb cryptic plasmid pRC4 from *R. rhodochrous* IFO 3338. Plasmid pK4 has a host range limited to *R. rhodochrous* and was used to characterize the conversion of nitriles by this species (Komeda *et al.*, 1996a, b, 1997).

Another set of shuttle vectors, for replication in the phytopathogen *R. fascians*, rely on a fragment from the

†**Present address:** Department of Biological Science, Ubon Ratchathani University, Ubon Ratchathani 34190, Thailand.

The GenBank accession numbers for the sequences reported in this paper are AF015088 (pFAJ2600) and AF015087 (*cmrA*).

160 kb plasmid pRF2 which confers chloramphenicol and cadmium resistance on *R. fascians* NCPPB 1675 (Desomer *et al.*, 1990). These vectors can also be maintained in *R. erythropolis* IGTS8, a strain capable of dibenzothiophene desulfurization (Denome *et al.*, 1993; Piddington *et al.*, 1995; Li *et al.*, 1996).

Stably maintained chimeras consisting of an unstable rhodococcal element joined to a portion of nocardiophage Q4 have been developed into suitable shuttle vectors such as pDA71 (Dabbs *et al.*, 1990, 1995; Quan & Dabbs, 1993). Plasmids of this type have enabled cloning of genes involved in decolorization of sulfonated azo dyes (Heiss *et al.*, 1992), degradation of herbicides (Nagy *et al.*, 1995a, b) and isopropylbenzene (Kesseler *et al.*, 1996), and of a rifampicin inactivation gene (Andersen *et al.*, 1997).

The rather narrow host range of the currently known shuttle vectors limits their use to only a few *Rhodococcus* species. The molecular basis for this is unknown since no attempts have yet been made to characterize the respective rhodococcal DNA fragments responsible for stable maintenance of the various shuttle vectors. Here we describe the isolation of a cryptic 6 kb plasmid (pFAJ2600) from *R. erythropolis* NI86/21 and report its complete nucleotide sequence. Analysis of these data revealed putative regions involved in replication, partitioning and multimer resolution. The replicon was used to construct shuttle vectors that could be stably maintained in at least *R. erythropolis*, *R. fascians*, *R. rhodochrous* and *R. ruber*.

METHODS

Bacterial strains and growth conditions. *Escherichia coli* DH5α served as a host for plasmid DNA manipulations and was grown at 37 °C in LB medium. The same medium was used to grow the following actinomycete strains at 30 °C: *R. erythropolis* NI86/21 and SQ1, *R. fascians* D188-5, *R. rhodochrous* N54, *R. ruber* N361, *Dietzia* (*Rhodococcus*) *maris* N1015 (Nagy *et al.*, 1995b) and *Mycobacterium smegmatis* mc^{2}155 (Snapper *et al.*, 1990).

Isolation of plasmid DNA from *Rhodococcus* strains. Ampicillin (0·6 mg ml^{-1} final concentration) was added to cells grown to an OD_{600} of 0·6–1·0 in 5 ml LB medium and the cells were further incubated for 2–3 h. The cells were then collected by centrifugation, washed with 1 ml 50 mM Tris/HCl buffer, pH 7·2, and incubated for 30 min at 37 °C in 0·25 ml TE buffer (10 mM Tris/HCl, pH 8·0, 1 mM EDTA) supplemented with 2 mg lysozyme ml^{-1}. Complete lysis of the weakened cells was achieved by adding 25 µl 10 % (w/v) SDS and 60 µl 5 M NaCl. The lysate was kept at −20 °C for 1 h and then centrifuged (12000 ***g***) for 15 min at 4 °C. RNA and proteins were removed from the supernatant by sequential treatment with RNase A (200 µg ml^{-1}) and proteinase K (50 µg ml^{-1}) for 15 min at 37 °C. Further purification consisted of standard chloroform/isoamylalcohol extraction (twice) and ethanol precipitation steps. Plasmid DNA was dissolved in 50 µl TE buffer.

Sequence analysis of pFAJ2600. To determine the complete nucleotide sequence of plasmid pFAJ2600 on both strands, overlapping fragments were subcloned in pUC18 and subjected to double stranded DNA sequencing using the automated ALF sequencer (Pharmacia). Computer-assisted sequence assembly and analyses were performed with the PCGENE software package (IntelliGenetics). Potential coding regions were identified with the GCWIND program (Shields *et al.*, 1992). For homology searches, the FASTA and BLAST servers were used.

Construction of *E. coli–Rhodococcus* shuttle vectors. The chloramphenicol resistance marker of plasmid pDA71 (Dabbs *et al.*, 1995) was isolated as a 1·8 kb *Bbr*PI–*Stu*I fragment and cloned in the *Sma*I site of pUC18. This pUC derivative (pFAJ2576) and pFAJ2600 were both digested with either *Hin*dIII or *Ecl*136II (unique sites in both vectors), ligated and used to transform *E. coli* DH5α. For both types of chimeras, the two possible orientations were selected by PCR and restriction analysis: pFAJ2573 and pFAJ2574 for the *Hin*dIII-joined replicons, and pFAJ2577 and pFAJ2578 for the *Ecl*136II-joined replicons.

Segregational plasmid stability. *Rhodococcus* cells containing pFAJ2574 were first grown in selective liquid medium (LB containing 40 µg chloramphenicol ml^{-1}). The late-exponential growth phase cells were diluted 100-fold in LB medium without the antibiotic and grown at 30 °C. At 24 h intervals, cultures were diluted 100-fold. At each dilution step, the cells were plated on LB medium to estimate the number of generations. From these plates, 100 colonies each were transferred to non-selective plates. Following growth, each colony was put on selective medium to determine the frequency of plasmid loss based on the percentage of chloramphenicol-sensitive colonies.

RESULTS AND DISCUSSION

Plasmid pFAJ2600 of *R. erythropolis* NI86/21

The plasmid profile of *R. erythropolis* NI86/21 revealed the presence of a small and a very large plasmid (data not shown). The complete nucleotide sequence (5936 bp) was determined for the small replicon (pFAJ2600). Eight potential ORFs were identified (Fig. 1a) and their characteristics are compiled in Table 1. The overlap of the start codon of *orf4* with the stop codon of *parA* (GTGA) suggests their translational coupling. Such coupling is also inferred from the lack of intergenic region between *repA* and *repB* (TAAATG). Two potential stem–loop structures that may be involved in transcriptional termination were detected downstream of *orf4* (stem length of 12 bp) and between *pmrA* and *orf1* (stem length of 15 with 1 mismatch) (Fig. 1a). The replicon has a G+C content of 60·8 % but base composition is not evenly distributed along the plasmid. The segment carrying *parA–orf1* contains 64·9 % G+C whereas a value of 59·1 % was calculated for the region spanning *orf3–repB* (Fig. 1a). These segments are separated by short regions of particularly low G+C content: 49·2 % G+C between *orf3* and *parA* and 46·6 % between *repB* and *orf1*. The *orf2*-proximal segment has a G+C content of 71·1 % and contains 11 repeats of a CGCC core sequence (Fig. 1a).

Replication genes in pFAJ2600

The RepA and RepB deduced protein sequences are related to the products of similarly organized genes in a number of small cryptic plasmids from various actino-

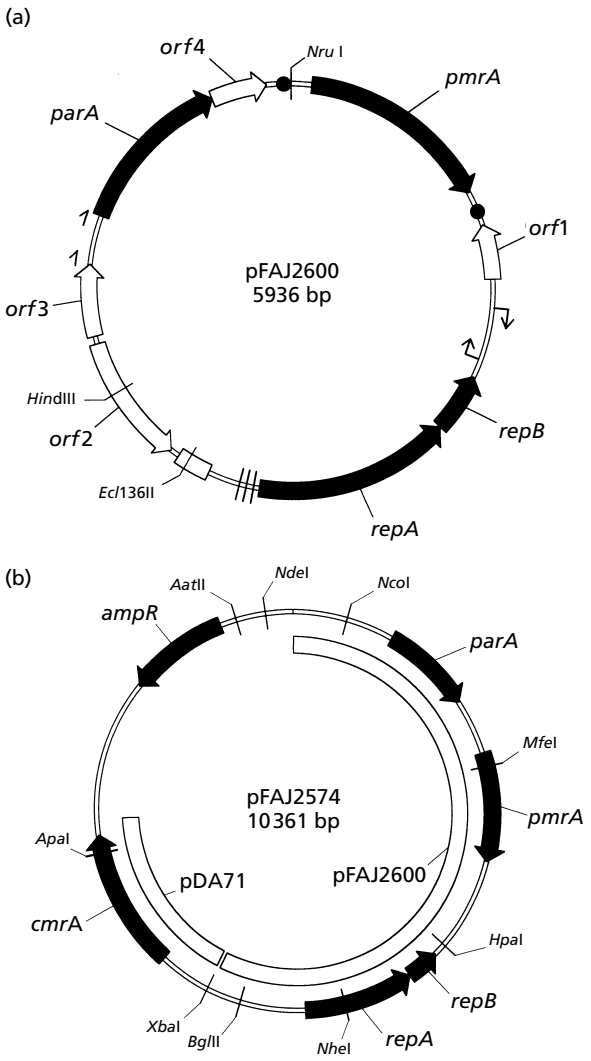

Fig. 1. (a) Gene organization and sequence motifs in pFAJ2600. The open and filled arrows represent the identified ORFs and putative genes, respectively. Potential stem–loop structures are shown as filled circles. The positions of the imperfect direct repeats are marked by half arrows. The opposing hooked arrows delineate the inverted repeats in the putative terminator region. The triplet of lines defines the conserved 15 bp sequence elements upstream of *repA* resembling the *ori* region of ColE2 plasmids. The DNA containing multiple repeats of the CGCC motif is shown as a box. The unique sites that were used for shuttle vector construction (*Hin*dIII and *Ecl*136II) are shown. The numbering starts from the *Nru*I site upstream of *pmrA*. (b) Gene organization in the *E. coli–Rhodococcus* shuttle vector pFAJ2574. The portions containing pFAJ2600 and the *cmrA* gene from pDA71 are indicated. The rest of the molecule (including the β-lactamase gene *ampR*) consists of pUC18 DNA. Unique restriction sites in pFAJ2574 are shown.

mycetes (Fig. 2). For pMB1 (Rossi *et al*., 1996) and pAL5000 (Stolt & Stoker, 1996a) it has been demonstrated that both *repA* and *repB* are required for replication. Our database searches also revealed significant homology with the products of two tandem genes located on the pJD1 plasmid from *Neisseria gonorrhoeae* (Korch *et al*., 1985). The alignment of the different RepA sequences (Fig. 2a) revealed the highest identity (32·5 %) of the pFAJ2600-encoded protein with the pAL5000 gene product. RepA from pFAJ2600 and its counterparts on the other actinomycete plasmids, as well as on pJD1, are all related to the Rep proteins encoded by the enterobacterial ColE2-type plasmids (Hiraga *et al*., 1994) and to the Rep homologue on the *Thiobacillus intermedius* plasmid pTiK12 (English *et al*., 1995). For instance, the alignment of the Rep protein encoded by ColE3-CA38 with RepA of pFAJ2600 revealed 24·1 % identical residues (alignment not shown).

When comparing pFAJ2600 RepB with the different known RepB proteins, the best alignment was obtained with the pXZ10142-encoded protein (31·0 % identity). Using the method of Dodd & Egan (1990), two regions with a helix–turn–helix motif were predicted in both of these RepB proteins (Fig. 2b). However, with this method only a single N-terminal DNA-binding motif was detected in RepB proteins from pMB1, pAL5000 and pJD1. Rauzier *et al*. (1988) proposed the C-terminal region of pAL5000 RepB as a putative DNA-binding domain. Actually, this region corresponds to the second proposed helix–turn–helix motif in the *Rhodococcus* and *Corynebacterium* RepB proteins. Furthermore, we noticed that the equivalent pJD1 RepB domain contains a sequence highly similar to a C-terminal motif in ColE2-related Rep proteins, which has been proposed as a DNA-binding site (Hiraga *et al*., 1994) (Fig. 2b). These observations suggest the existence of two DNA-binding domains in the RepB proteins. In this respect, RepB resembles the C-terminal domain of ColE2 Rep which carries two regions, about 30 residues apart, involved in binding to the origin of replication (Shinohara & Itoh, 1996). Unlike the first motif, the second motif does not display sequence conservation between the different RepB proteins and may therefore be important for specificity in binding.

The significant similarity of the RepA proteins with the ColE2-type of Rep proteins suggests that RepA may also act as a plasmid-specific primase which synthesizes a unique RNA primer (Takechi *et al*., 1995; Takechi & Itoh, 1995). The ColE2-type of Rep proteins contain two regions close to the C terminus that are determinants of specificity for the interaction of the initiator with the origin. We have not been able to identify a possible DNA-binding region in the different RepA proteins. This is consistent with the inability to demonstrate DNA binding for the pAL5000-encoded RepA protein (Stolt & Stoker, 1996b). The RepB protein of pAL5000, however, binds specifically to two DNA segments only 37 bp apart and located in the region upstream of *repA* (Stolt & Stoker, 1996b). Actually, we predict two putative DNA-binding motifs in the RepB proteins. These observations suggest that RepA (primase) and RepB (DNA-binding protein) act together when initiating DNA replication, in a way similar to Rep proteins of ColE2 plasmids which possess both activities.

Table 1. Characteristics of ORFs in pFAJ2600 and deduced protein sequences

ORF				Deduced protein sequence		
Name	% G+C	RBS*	Start/stop codon	Size (aa/Da)	pI	Homology†
parA	62·6	AGG [8]	ATG/TGA	258/26754	4·3	Soj/26·9
pmrA	69·1	GGAAGGA [6]	GTG/TGA	306/34198	11·0	YT23M/34·9
repA	59·1	GAGGG [8]	ATG/TAA	310/35451	10·0	RepA/32·5
repB	57·5	GGAGG [5]	ATG/TGA	93/10672	10·1	Rep B/31·0
orf1	65·1	GAA [10]	GTG/TAG	86/9140	11·6	–
orf2	58·3	AGGA [11]	TTG/TAG	202/22890	11·8	–
orf3	57·1	GGAA [9]	GTG/TGA	114/12906	9·8	–
orf4	64·1	–	GTG/TAG	89/9693	10·7	–

* Putative RBS [number of nucleotides from start codon].

† Most similar protein/percentage identical amino acids; Soj of *B. subtilis* (Ogasawara & Yoshikawa, 1992), extended ORF YT23M in Tn*4556* of *S. fradiae* (Siemieniak *et al*., 1990), RepA of *Mycobacterium fortuitum* plasmid pAL5000 (Rauzier *et al*., 1988), RepB homologue of *Corynebacterium glutamicum* plasmid pXZ10142 (GenBank X72691).

PmrA, a putative site-specific recombinase of the λ integrase family

Plasmid pFAJ2600 carries an ORF encoding a putative new member of the λ integrase family, tentatively named PmrA (plasmid multimer resolution protein A), reflecting its proposed function in pFAJ2600 maintenance. PmrA contains the characteristic RHRY tetrad involved in catalytic activity of this type of recombinase (Fig. 3). The extent of homology of PmrA with λ integrase-like proteins is low, but such low sequence conservation is a feature of this protein family (Stark *et al*., 1992). Among the well-characterized enzymes of this family, the closest relative is the chromosomally encoded XerD of *E. coli* (298 residues; 20% identity). XerD, in concert with its homologue XerC, is required for site-specific recombination at chromosomal (*dif*) and plasmid sites (like *cer* in ColE1) (Colloms *et al*., 1990; Blakely *et al*., 1993, 1997). Xer recombination on natural plasmids is preferentially intramolecular, converting multimers to monomers (Blake *et al*., 1997). Plasmid multimer resolution contributes to plasmid stability by maximizing the number of independently segregating units at division (Williams & Thomas, 1992). Several site-specific recombinases are encoded by plasmids of low-G+C Gram-positive bacteria, but these enzymes belong to the Tn*3* family of site-specific recombinases (reviewed by Alonso *et al*., 1996). To our knowledge, no plasmid stabilization system relying on the activity of a λ integrase-type of site-specific recombinase has been described previously in actinomycetes.

The most pronounced homology detected for PmrA was with ORF5 (=YT23) encoded by the *Streptomyces fradiae* transposon Tn*4556* (Siemieniak *et al*., 1990). We noticed that the homology extends into the upstream region of this ORF, disregarding a TAG stop codon (Fig. 3). The deduced polypeptide of 324 residues is very similar in length to PmrA (306 residues; 34·9% identity) and to the Xer proteins (298 residues). The transposition activity of Tn*4556* is not affected by interruption or deletion of YT23 (Siemieniak *et al*., 1990; Sohaskey *et al*., 1992). Indirect evidence suggesting that the PmrA-related protein may be involved in plasmid stabilization stems from the observations that insertion of this transposon into unstable plasmids generated stable plasmids (Chung, 1987).

Homology of ParA with partitioning proteins

Sequence comparisons with ParA revealed significant homology with a family of ATPases involved in active partitioning of various bacterial plasmids (Motalebbi-Veshareh *et al*., 1990), in which three motifs are particularly well conserved (Koonin, 1993). The motif close to the N termini of these A-type of Sop/Par partitioning proteins (Williams & Thomas, 1992) constitutes a modified NTP binding motif, KGGXXK[S,T] (Koonin, 1993), which is also present in ParA of pFAJ2600 (Fig. 4). Additional plasmid-encoded members of this ATPase family have been reported in *Rhizobium leguminosarum* (Turner & Young, 1995) and *Ralstonia eutropha* (*Alcaligenes eutrophus*) (Taghavi *et al*., 1996). Among actinomycetes, such plasmid-encoded ATPase has only been reported on the small linear plasmid pSCL1 from *Streptomyces clavuligerus* (Wu & Roy, 1993). The best alignment was obtained with the small IncC protein (21·6% identity; Fig. 4). These plasmid partitioning proteins are related to proteins that are encoded in the replication region of various bacterial chromosomes (Ogasawara & Yoshikawa, 1992; Fsihi *et al*., 1996). As shown in Fig. 4, the ParA from pFAJ2600 aligned somewhat better with the *Bacillus subtilis* chromosomal ATPase, Soj (26·9% identity), than with IncC from the broad-host-range plasmid RK2.

Active plasmid partitioning minimizes the frequency of plasmid-free cells arising (Williams & Thomas, 1992).

(a)

```
                                        ↓                    ↓         ↓          ↓ ↓   ↓
Rhod                 METMTPERWEAEHWYGRYPLASNNPRIDGSFRMPRSDALLKRFIQANPREYVSQIVVDIDHADAELRA  68
Myco                  MSHVADEFE-QLWLPYWPLASDDL-LEGIYRQSRASALGRRYIEANPTALANLLVVDVDHPDAALRA  65
Cory               MRPSTRPEGRFE-QLWLPLWPLASDNL-RDGIYRTSRDKALDKRYIETNPDAISNLLVVDIDKQDALLRS  68
Brev                  MSTASTETWG-QMWLPLWPLASDDL-LQGIYRTSRHNALELRYIEANPQSLSNLLVVDIDHPDALMRA  66
Bifi  MYAVAMSDEYSQPTLELSRTFEG-WWLPERPLCCDDD-YSRLHRRSRADALKCKHIEANPAALVNTIVVDIDDANAKAMA  79
                           -*   =**  ==***     *  = *=  ==  - =*-==    * -====-=  *=  *-
              ↓          ↓                                    ↓         ↓ ↓
Rhod  FSMHEVGLVPNMFAYSTRPGHGQATFLLRSPVSLSEASRRKPINLLARCQQGLTVALSGDPHYSGPLARNPTHPMANTRW 148
Myco  LSARGSHPLPNAIVGNRANGHAHAVWALNAPVPRTEYARRKPLAYMAACAEGLRRAVDGDRSYSGLMTKNPGHIAWETEW 145
Cory  MWDRDGWR-PNAVVENPXNGHAHAVWALAEPFTRTEYARRKPPXYAAAVTEGLRRSVDGDKGYSGLITKNPEHTAWDSHW 147
Brev  MWNRKAWQ-PNAVVENPANGHAHAVWALAEPVTRTEYARRKPLAYAAAVTEGLRRSVDGDKGYSGLITKNPTHDQWEASW 145
Bifi  LWEHEGMR-PNWIAENPANGHAHAGWVLTFPVPRTDLARLKPLKLLHATTEGLRRSCDGDMGYSGLLMKNPEHPAWASDI 157
                ==    *  *==**= - =  =* *-- -=*==    **  *==**- *==  ===*- *== =  * - *
           ↓ ↓                                  ↓
Rhod  CHPTPYELRDLARALGRDLPMPPTSRRVEASIDSA---LGRNCWMFDVTRQWAYRA-----WTRYPQRTDWDDAVF---- 216
Myco  LHSDLYTLSHIEAELGAN--MPPPRWRQQTTYKAAPTPLGRNCALFDSVRLWAYRPALMRIYLPTRNVDGLGRAIY---- 219
Cory  VTDKLYTLDELRFWLEETGFMPPESWKK--TRRKSPIGLGRNCALFESARSWAYR----EIRHHFGDPDGLGRSIQATAQ 222
Brev  LTDHLYNLDELTEHLTVSDFMPPDSWQR--TKRKNPVGLGRNCTLFETVRWDVYRVARTIRKRNEHPTPEDRHDLEAAIV 223
Bifi  IEWDTYDLEQLVQSLQEHGDMPPVSWKR--TKRARTQGLGRNCTLFDKARTLAYRYV----AAAADRSEASSEALRLYVR 231
           = =  -   =     === **    -       ===== -=-  =  *==                    -

Rhod  AYTWGRNPELVSHS--KGPLPDSELRTISRSVASFVW-NSEMRAKGAEQFDADFSKIQANRGHKGGKTMTEKRREVIAET 293
Myco  AECHARNAEFPCNDVCPGPLPDSEVRAIANSIWRWITTKSRIWADGIVVYEATLSARQSAISRKGAAARTAASTVARRAK 299
Cory  ----ALNQELFS-----EPLPVAEVDQIARSIHRWIITKSRMWTDGPAVYEATFTTIQATRGKRSAEHRWGTTNAERIER 292
Brev  NLCQGMN-STFS-----EALPASEIRATIRSFYKWITTRYTGWLDSRTTSQSKSAAYHRNTGRKGGLKGGVVSGQVRRQK 297
Bifi  RTCHELNVSLF-----PDPLHAREVEDIAKSIHKWIVTRSRMWRDGAIANAATFIAIQSARGHKHGENKYQQVMKEALEW 306
            =           *=*  =-     =   -- * *  * **     -     *   * -

Rhod  NRRRKVDRAAVLAELEA                                                              310
Myco  SASAMEALL                                                                      308
Cory  FIND                                                                          296
Brev  ALERVSMLEAGKL                                                                 310
```

(b)

```
Rhod                  MPAENQTRRN**KTAKEMAERLGVSERTIRNIVA**ESRDSYQARAA-ERRDTAVKLREQG**MKYREIAEE**  65
Cory                 MTKRTRIPRNG**KTIREVAEGTGLSTATIERWTS**APREDYLA-QANEKRVRVQELRAKG**LSMRAIAAE**  66
Bifi                 MVRTTLRKKRP**VSARELAEAYGVSTRTIQSWVA**MKREDWIDEQAAMREAVRSYHDDEGHTWPQTAEH  67
Myco  MSDGYSDGYSDGYNRQPTVRKKRR**VTAAEGARITGLSERHVVRLVA**QERSEWLAEQAARRERIRAYHDDEGHSWPQTAKH  80
Neis               MAYPKLKTTKRD**VTAKELAKRFGCSTRTVFRAWS**QSRADYLAENSISRD----------KPWEHFGIS  58
                           -*  -* = =   = = **-    -  =  - *  -  -                  *

Rhod  **MEISTGAVGRLLH**DAKKHDEKNIESRAS                                               93
Cory  **IGCSVGLVHRYVK**EVEEKKTA                                                      87
Bifi  FNMSQGAVRQRCYRARKEREDEAAEKSKHLPGEIPLFD                                         95
Myco  FGLHLDTVKRLGYRARKERAAEQEAAQKAHNEADNPPLF                                       110
Neis  RATWYRRGKPMPSETDNQSETA                                                         80
```

Fig. 2. Sequence alignments of the putative replication proteins encoded by pFAJ2600 from *R. erythropolis* (*Rhod*) with RepA (a) and RepB (b) homologues of the actinomycete plasmids pAL5000 from *Mycobacterium fortuitum* (*Myco*; Rauzier *et al.*, 1988), pXZ10142 from *Corynebacterium glutamicum* (*Cory*; GenBank X72691), pRBL1 from *Brevibacterium linens* (*Brev*; Ankri *et al.*, 1996) and pMB1 from *Bifidobacterium longum* (*Bifi*; Rossi *et al.*, 1996). Two positions with presumptive frameshifts in the pXZ10142 *repA* sequence are translated as 'X'. A *repB*-like gene is not contained on the partial sequence data available for pRBL1, but the RepB-like protein from the *N. gonorrhoeae* plasmid pJD1 (*Neis*) is included (Korch *et al.*, 1985). Positions with identical (=) or similar (−) residues in all sequences are indicated. Residues identical in four out of five sequences are marked (*). Arrows denote RepA positions with identical amino acids in eleven ColE2-type Rep proteins (Hiraga *et al.*, 1994), pTiK12 Rep (English *et al.*, 1995) and pJD1 RepA (Korch *et al.*, 1985). The helix–turn–helix motifs in the RepB proteins predicted by the method of Dodd & Egan (1990) are in bold. The DNA-binding region proposed by Rauzier *et al.* (1988) for pAL5000 RepB is underlined, as is the sequence of the *Neisseria* RepB-like protein with high similarity to the C-terminal consensus sequence KPWE[A,T,K]LGISRA[T,W]YY[Q,R][L,K]$_2$K of eight ColE2-type Rep proteins (Hiraga *et al.*, 1994).

```
YT23M  MPTVVQLPTGKALSVRAAADVFLDSLGNPNTVRNYGIGVGKTAERLGESRPLASVADDEIGEALELLWGTSAVNT*NSR  79
PmrA   MGQVRALPV-RGVRLKDAVDQYLSTIPSVNTRRGYAVALNQLVRDFGADSDVGLLEAERVGGWFTFKWGGSSAQTFNVR  78
                                                +                                  ↓
YT23M  RAAVLSWLGWCR--ERGYDGPAVPAWAKRLAVPDSETPARSRMAIDRLIARREVHLRE**KTLWRMLYETAGRSEEILGVN** 156
PmrA   LASLRGACEYWRAQEWLLGDPVVRLVPRR-VAPD-HSRAMTRDEVTGLLAQ-DVPLRE**RVLWTMLYETAARADELLLLD** 154

YT23M  **IED**LDFAGRRCPAKAKGAKSKSRRRGQVREDFVLETLYWDAGTARLLPRLLKGRTRGPVFVTHRRPGPGKVVSPRDICP 235
PmrA   **IPD**LDTA-NRCGTVTRKGGAR-------------DVVAWQTGTARLLPRLLAGRKKGPVFLTDRKAKPGVAV--DDIDP 217
                                  ↓  ↓                              ↓
YT23M  DTGFARLSYGQARALLDHHTAVRGVGTGWDL**HEYRHSALTHLGEQGASLLMLMAKSRHKKPENVRRYFK**PSPEAIAELT 314
PmrA   STLRGRLSYRRALELFETHTADFTRGP-FTL**HQLRHSRLTHAAEDGASTPMLMRMSGHTSVRSLARYAN**PSAESLMRWQ 295

YT23M  SLLAPGNSSR                                                                      324
PmrA   AQTDPAARRRR                                                                     306
```

Fig. 3. Alignment of the PmrA sequence encoded by pFAJ2600 with the amino acid sequence deduced from the *S. fradiae* Tn*4556* sequence (Siemieniak *et al.*, 1990). The reported N-terminal methionine of ORF YT23 is marked (+). The alignment is extended to the translated 5′ region of YT23 which is interrupted by a stop codon (UAG), indicated as '*' (modified ORF = YT23M). The upstream region of YT23 has a high coding probability (GCWIND analysis). Unlike YT23 (Siemieniak *et al.*, 1990), the putative start codon of YT23M is preceded by an appropriate potential RBS. Residues that are identical (=) or similar (–) in both sequences are marked. The two domains most conserved among integrases of the phage λ integrase family of site-specific recombinases are in bold (Colloms *et al.*, 1990). The amino acids within these domains constituting the catalytic RHRY tetrad are indicated with arrows (Kwon *et al.*, 1997).

```
IncC   MK**TLVTANQKGGVGKT**STLVHLAFDFFER**GLRVAVIDLDP**QGNASYTLK---DFATGLHASKLFGA---------VPAG  67
ParA   MT**VIAVVNQKGGSGKS**TTVLGLASAASAR**GIETLVIDLDP**QCNASEALGIV-YPVEGYTAAELLAADFPGTALDAVHAS  78
Soj    MGK**IIAITNQKGGVGKT**TTSVNLGACLAYI**GKRVLLVDIDP**QGNATSGLGIEKADVEQCVYDILVDDA---DVIDIIKAT  77

IncC   GWTETAPAAGDGQAARLALIESNPVLANAERLSLDDARELFGANIKALANQGF**DVCLID**TAPTLGVGLAAALFAADYVLS 147
ParA   SWDNVAVIPGDLDLADL---DAVAGLGVEQRLR--------AALDGEDWQGRF**PLILID**CPPSVGKLVSNALIAADTALV 147
Soj    TVENLDVIPATIQLAG-AEIELVPTISREVRLK--------RALEA--VKQNY**DYIIID**CPPSLGLLTINALTASDSVVI 146

IncC   PIELEAYSIQGIKKMVTTIANVRQ-KNAKLQFLGMVPSKVDARNPRHARHQAELLAAYPKMMIPATVGLRSSIADALASG 226
ParA   ATEPSFMASRGVSKILQAIETIQRYYNPALTVAGVLIGRVPAQGREAAHRTAEIREALGDQVLPLVVPQRAAVAEAAGDR 227
Soj    PVQCEYYALEGLSQLLNTVRLVQKHLNTDLMIEGVLLTMLDARTNLGIQVIEEVKKYFRDKVYKTVIPRNVRLSEAPSHG 226

IncC   VPVWKIKKTAARKASKEVRALADYVFTKMEISQ                                              259
ParA   RPIHQVRPVVTEVVDAFDAALDRVLAAVVVS                                               258
Soj    KPIILYDPRSRGAEVYLDLAKEVAANG                                                   253
```

Fig. 4. Sequence alignment of the putative partitioning protein ParA from pFAJ2600 with the chromosomal Soj protein from *B. subtilis* (Ogasawara & Yoshikawa, 1992) and the small IncC protein from the broad-host-range plasmid RK2 (Thomas & Smith, 1986). Residues that are identical (=) or similar (–) in equivalent positions of the three proteins are marked. In addition, residues that are shared by ParA with either Soj or IncC are indicated (*). The sequences covering the conserved ATPase motifs (Koonin, 1993) are in bold.

The putative rhodococcal partitioning system resembles the 'simple' systems, consisting of only an A-type of protein (such as ParA of the *Agrobacterium* pTAR plasmid), and a *cis*-acting site (upstream repeat region in the case of pTAR ParA) (Gallie & Kado, 1987). A putative ParA-only system appears also to be present on the linear plasmid pSCL1 from *S. clavuligerus* (Wu & Roy, 1993) and on the pAD1 plasmid from *Enterococcus faecalis* (Weaver *et al.*, 1993). The upstream region of *parA* in pFAJ2600 displays a peculiar structural organization with two long imperfect direct repeats (66 and 67 bp), one of which contains a 36 bp perfect palindromic sequence (Fig. 5a). This part of pFAJ2600 may provide the *cis*-acting partitioning site and/or be involved in negative *parA* autoregulation. This type of regulation involving the A component is frequently found in the *sop/par* partitioning systems (Williams & Thomas, 1992; Hiraga, 1992).

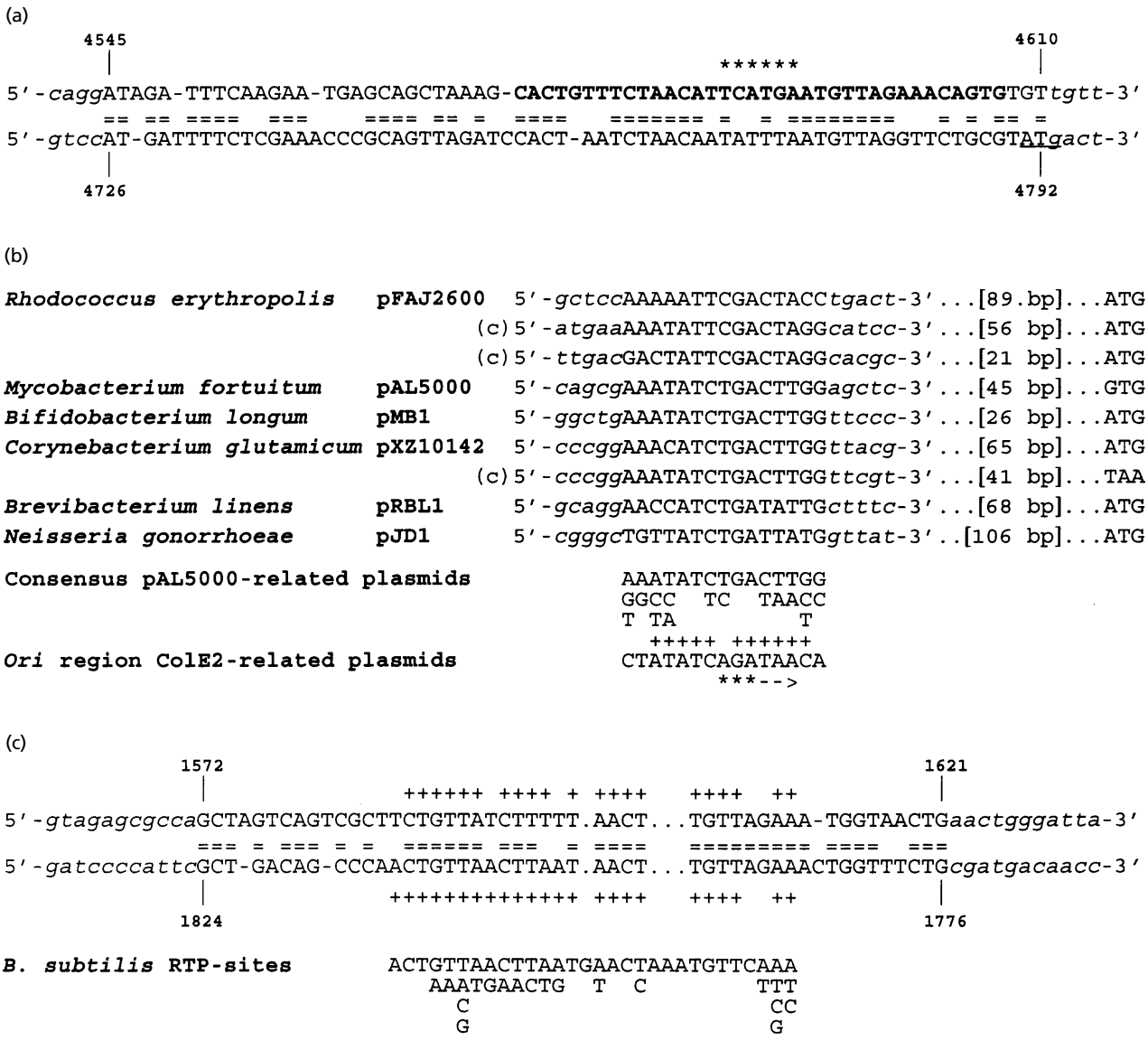

Fig. 5. DNA sequence motifs in pFAJ2600. (a) Alignment of the A+T-rich imperfect direct repeats located upstream of the *parA* gene in pFAJ2600. The nucleotides constituting a perfect 36 bp palindromic sequence are in bold. This palindrome is centred around the unique *Bsp*HI recognition site of pFAJ2600 (*). The proposed start codon of *parA* is underlined. (b) Alignment of the conserved 15 bp sequence identified in the promoter region of the *repA* gene in pAL5000-related plasmids and the deduced consensus sequence. The numbers in square brackets indicate the distance from the respective proposed *repA* start codons (ATG, GTG) or *repB* stop codon (TAA). The motifs with an inverse orientation relative to *repA* (on the complementary strand) are denoted (c). Nucleotides flanking the conserved sequences are in italicized lower-case characters. For comparison, the core sequence of the *ori* region in ColE2-related plasmids is included (Hiraga *et al.*, 1994) and the primer RNA (5′ ppApGpA) for leading strand synthesis (arrow) is marked (***) (Takechi *et al.*, 1995). Positions where the 15 bp consensus sequence matches the ColE2 *ori* sequence are marked (+). (c) Alignment of the two pFAJ2600 inverted repeat sequences resembling the replication terminator sequences of *B. subtilis*. The numbering is as in Fig. 1(a). Identical nucleotides are indicated (=). The consensus DNA sequence (RTP box) within the *B. subtilis* terminator regions, involved in binding to the replication terminator protein (RTP), was deduced from eight chromosomal *Ter* sites (Franks *et al.*, 1995) and one plasmid (pLS20) terminator (Meijer *et al.*, 1996). Nucleotides matching this consensus sequence are marked (+) in each of the two repeats. Dots correspond to gaps introduced in the pFAJ2600 sequence to optimize alignment with the *B. subtilis* terminator consensus sequence.

Putative *ori* in pAL5000-related plasmids

When comparing the putative promoter regions of the *repAB* genes from the pAL5000-related plasmids, the sequence conservation of a 15 bp element located within approximately 100 bp of the *repA* start was noticed (Fig. 5b). Whereas such a motif occurs only once in the plasmids from *Mycobacterium*, *Brevibacterium* and *Neisseria*, three such sequences are present in the 5′ region of *repA* from pFAJ2600, only one being in the

same orientation as in the previous plasmids (Fig. 1a). In the *Corynebacterium* plasmid a second copy of the motif was also detected in the 3′ region of *repB*, but located on the opposite strand (Fig. 5b). Remarkably, in pAL5000 this 15 bp sequence is positioned right between the RepB binding sites that were identified by Stolt & Stoker (1996b). We noticed that this sequence resembles the *ori* region of ColE2 plasmids (Hiraga *et al*., 1994). During replication of these plasmids, unidirectional leading strand synthesis is initiated at this site following synthesis of a specific primer by the DNA-bound Rep proteins (Takechi & Itoh, 1995; Takechi *et al*., 1995).

A *B. subtilis Ter*-like region in pFAJ2600

Between *repB* and *orf1*, a 50 bp sequence and a 49 bp sequence are organized as an imperfect inverted repeat (79·6% match), separated by 154 bp (Fig. 5c). The repeats are characterized by a higher than average proportion of A+T (64 and 61%, respectively). This structural feature is reminiscent of the organization of the chromosomal replication termination sites with IR-I and IR-II of *B. subtilis* (Baker, 1995). These inverted repeats of 47 and 48 bp, respectively, are 59 bp apart and share 77% identical nucleotides (Hill, 1992). In addition to this structural similarity, significant sequence similarity was found with the 30 bp sequence that constitutes the contact region for *B. subtilis* replication terminator protein, as outlined in Fig. 5(c). These striking similarities suggest that this pFAJ2600 region may be involved in termination of plasmid replication.

E. coli–Rhodococcus shuttle vectors based on pFAJ2600

It was inferred from the sequence analysis of pFAJ2600 that certain regions of the plasmid may be dispensable for stable maintenance in *Rhodococcus*. Unique sites located in *orf2* (*Hin*dIII) and in the high-G+C region downstream of *orf2* (*Ecl*136II) were therefore selected to construct potential shuttle vectors by generating chimeras of pFAJ2600 and pUC18. To provide a marker gene selectable in *Rhodococcus*, the chloramphenicol resistance gene of plasmid pDA71 was selected. This chloramphenicol resistance determinant, which originates from an unstable genetic element of *R. rhodochrous* ATCC 12674 (Quan & Dabbs, 1993), was further sequenced. Our sequence data revealed a gene, *cmrA*, with 77·0% identity to the *R. fascians cmr* gene that encodes a membrane-bound chloramphenicol efflux system (Desomer *et al*., 1992). This is consistent with the non-enzymic type of chloramphenicol resistance for the pDA71 determinant reported by Quan & Dabbs (1993). The deduced amino acid sequence of CmrA shares 85·7% identity with *R. fascians* Cmr. The higher sequence conservation at protein level compared to DNA level reflects the significantly different G+C content of *cmrA* (66·2%) and *cmr* (61·1%).

The *cmrA*-containing pDA71 fragment was first introduced into pUC18. Then, the resulting *E. coli* replicon was joined to pFAJ2600, using their unique *Hin*dIII or *Ecl*136II sites. For each type of fused replicon, two variants differing in relative orientation of the fused fragments were obtained: pFAJ2573/pFAJ2574 with *orf2* interrupted at the *Hin*dIII site and pFAJ2577/pFAJ2578 with the pUC18 derivative inserted in the G+C-rich region (*Ecl*136II site). The gene organization of pFAJ2574 is shown in Fig. 1(b). Introduction of the putative shuttle vector pFAJ2574 in *R. erythropolis* SQ1, *R. fascians* D188-5, *R. rhodochrous* N54, *R. ruber* N361, *D. maris* N1015, and *M. smegmatis* mc^{2}155 was attempted using electroporation (Desomer *et al*., 1990). By colony hybridization with the entire pFAJ2600 plasmid as probe, the absence of a closely related plasmid in these strains had been previously confirmed. Transformants were selected using chloramphenicol (40 μg ml^{-1}) and obtained with *R. erythropolis* SQ1 [10^4 (μg plasmid DNA)$^{-1}$], *R. fascians* (10^2 μg^{-1}), *R. rhodochrous* (< 10 μg^{-1}) and *R. ruber* (< 10 μg^{-1}). Similar efficiencies were obtained when the related constructs pFAJ2573, pFAJ2577 and pFAJ2578 were used to electrotransform *R. erythropolis* SQ1. Successful transformation was not achieved with *M. smegmatis* and *D. maris*, but no attempts were made to optimize electroporation conditions for individual strains. pFAJ2574 was reisolated from a number of randomly selected colonies for the four *Rhodococcus* species and reintroduced in *E. coli* DH5α. The integrity of the shuttle vector was confirmed by restriction analysis, indicating that no rearrangements or deletions had taken place during the sequential passages in *E. coli* and *Rhodococcus* cells. These data showed that the *parA–repA* intergenic region is not required for plasmid maintenance in *Rhodococcus*.

The segregational stability of pFAJ2574 under non-selective conditions differed depending on the host strain used (Fig. 6). The shuttle vector was quite stable in both *R. erythropolis* SQ1 and *R. fascians* D188-5. A moderate stability was observed in *R. rhodochrous* N54, with about 70% plasmid loss after 60 generations. Similar plasmid loss rates (1–1·5% per generation) have been reported for plasmid pMVS301 in *Rhodococcus* sp. strain AS-50-1 (Vogt Singer & Finnerty, 1988) and plasmid pK4 in *R. rhodochrous* ATCC 12674 (Hashimoto *et al*., 1992). In *R. ruber* N361 grown under non-selective conditions, pFAJ2574 was lost at a considerably higher rate (Fig. 6). Evidently, such plasmid stability data should be interpreted with care since *Rhodococcus* species display different morphological stages during their growth cycle. Depending on the species, the extent of differentiation ranges from elementary branching (as observed for *R. erythropolis* in this study) to production of hyphae-like structures. Later these structures fragment into rods and cocci. Since colony formation may thus occur from more than one cell, this could lead to overestimating plasmid stability. On the other hand it should be pointed out that our estimation of the plasmid-retaining subpopulation

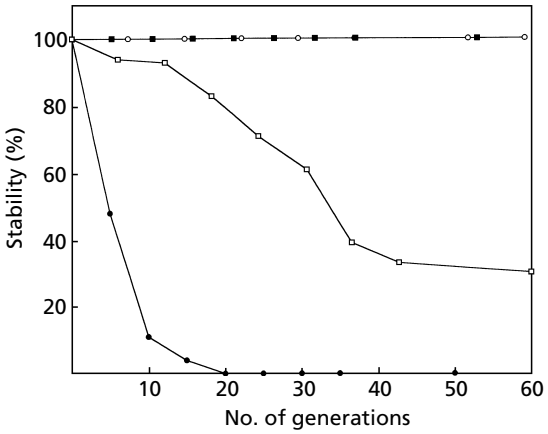

Fig. 6. Segregational stability of shuttle vector pFAJ2574 in *R. erythropolis* SQ1 (○), *R. fascians* D188-5 (■), *R. rhodochrous* N54 (□) and *R. ruber* N361 (●) in the absence of antibiotic resistance selection.

was carried out after two passages on non-selective LB plates, which tends to underestimate somewhat the actual plasmid stability. The stable maintenance of the pFAJ2600 derivatives in *R. erythropolis* (and *R. fascians*) is in line with our failure to cure pFAJ2600 from strain NI86/21. The reduced plasmid stability observed in two other species probably reflects in part the involvement of host-encoded proteins in plasmid replication (initiation, termination) and partitioning. For instance, for the pAL5000-based replicons there are indications of a requirement for mycobacterial DNA polymerase I to initiate DNA synthesis (Stolt & Stoker, 1996b).

ACKNOWLEDGEMENTS

R. De Mot is a Senior Research Associate with the Fund for Scientific Research (Flanders). Part of this research was supported by this foundation. Strains used in this study were kindly provided by E. Dabbs, J. Desomer and M. Goodfellow. The GCWIND program was obtained from D. Shields.

REFERENCES

Alonso, J. C., Ayora, S., Canosa, I., Weise, F. & Rojo, F. (1996). Site-specific recombination in Gram-positive theta-replication plasmids. *FEMS Microbiol Lett* **142**, 1–10.

Andersen, S. J., Quan, S., Gowan, B. & Dabbs, E. R. (1997). Monooxygenase-like sequence of a *Rhodococcus equi* gene conferring increased resistance to rifampin by inactivating this antibiotic. *Antimicrob Agents Chemother* **41**, 218–221.

Ankri, S., Bouvier, I., Reyes, O., Predali, F. & Leblon, G. (1996). A *Brevibacterium linens* pRBL1 replicon functional in *Corynebacterium glutamicum*. *Plasmid* **36**, 36–41.

Baker, T. A. (1995). Replication arrest. *Cell* **80**, 521–524.

Blake, J. A., Ganguly, N. & Sherratt, D. J. (1997). DNA sequence of recombinase-binding sites can determine Xer site-specific recombination outcome. *Mol Microbiol* **23**, 387–398.

Blakely, G., May, G., McCulloch, R., Arciszewska, L. K., Burke, M., Lovett, S. T. & Sherratt, D. J. (1993). Two related recombinases are required for site-specific recombination at *dif* and *cer* in *E. coli* K12. *Cell* **75**, 351–361.

Blakely, G. W., Davidson, A. O. & Sherratt, D. J. (1997). Binding and cleaving of nicked substrates by site-specific recombinases XerC and XerD. *J Mol Biol* **265**, 30–39.

Chung, S.-T. (1987). Tn*4556*, a 6·8-kilobase-pair transposable element of *Streptomyces fradiae*. *J Bacteriol* **169**, 4436–4441.

Colloms, S. D., Sykora, P., Szatmari, G. & Sherratt, D. J. (1990). Recombination at ColE1 *cer* requires the *Escherichia coli xerC* gene product, a member of the lambda integrase family of site-specific recombinases. *J Bacteriol* **172**, 6973–6980.

Dabbs, E. R., Gowan, B. & Andersen, S. J. (1990). Nocardioform arsenic resistance plasmids and construction of *Rhodococcus* cloning vectors. *Plasmid* **23**, 242–247.

Dabbs, E. R., Gowan, B., Quan, S. & Andersen, S. J. (1995). Development of improved *Rhodococcus* plasmid vectors and their use in cloning genes of potential commercial and medical importance. In *Proceedings of the Ninth Symposium on the Biology of Actinomycetes*, pp. 129–135. New York: Allerton Press.

Denome, S. A., Olson, E. S. & Young, K. D. (1993). Identification and cloning of genes involved in specific desulfurization of dibenzothiophene by *Rhodococcus* sp. strain IGTS8. *Appl Environ Microbiol* **59**, 2837–2843.

Desomer, J., Dhaese, P. & Van Montagu, M. (1990). Transformation of *Rhodococcus fascians* by high-voltage electroporation and development of *R. fascians* cloning vectors. *Appl Environ Microbiol* **56**, 2818–2825.

Desomer, J., Vereecke, D., Crespi, M. & Van Montagu, M. (1992). The plasmid-encoded chloramphenicol-resistance protein of *Rhodococcus fascians* is homologous to the transmembrane efflux proteins. *Mol Microbiol* **6**, 2377–2385.

Dodd, I. B. & Egan, J. B. (1990). Improved detection of helix-turn-helix DNA-binding motifs in protein sequences. *Nucleic Acids Res* **18**, 5019–5026.

English, R. S., Lorbach, S. C., Huffman, K. M. & Shively, J. M. (1995). Isolation and characterization of the replicon of a *Thiobacillus intermedius* plasmid. *Plasmid* **33**, 1–6.

Finnerty, W. R. (1992). The biology and genetics of the genus *Rhodococcus*. *Annu Rev Microbiol* **46**, 193–218.

Franks, A. H., Griffiths, A. A. & Wake, R. G. (1995). Identification and characterization of new DNA replication terminators in *Bacillus subtilis*. *Mol Microbiol* **17**, 13–23.

Fsihi, H., De Rossi, E., Salazar, L., Cantoni, R., Labò, M., Riccardi, G., Takiff, H. E., Eiglmeier, K., Bergh, S. & Cole, S. T. (1996). Gene arrangement and organization in a ~76 kb fragment encompassing the *oriC* region of the chromosome of *Mycobacterium leprae*. *Microbiology* **142**, 3147–3161.

Gallie, D. R. & Kado, C. I. (1987). *Agrobacterium tumefaciens* pTAR *parA* promoter region involved in autoregulation, incompatibility and plasmid partitioning. *J Mol Biol* **193**, 465–478.

Hashimoto, Y., Nishiyama, M., Yu, F., Watanabe, I., Horinouchi, S. & Beppu, T. (1992). Development of a host–vector system in a *Rhodococcus* strain and its use for expression of the cloned nitrile hydratase gene cluster. *J Gen Microbiol* **138**, 1003–1010.

Heiss, G. S., Gowan, B. & Dabbs, E. R. (1992). Cloning of DNA from a *Rhodococcus* strain conferring the ability to decolorize sulfonated azo dyes. *FEMS Microbiol Lett* **99**, 221–226.

Hill, T. M. (1992). Arrest of bacterial DNA replication. *Annu Rev Microbiol* **46**, 603–633.

Hiraga, S. (1992). Chromosome and plasmid partition in *Escherichia coli*. *Annu Rev Biochem* **61**, 283–306.

Hiraga, S.-I., Sugiyama, T. & Itoh, T. (1994). Comparative analysis of the replicon region of eleven ColE2-related plasmids. *J Bacteriol* **176**, 7233–7243.

Kesseler, M., Dabbs, E. R., Averhoff, B. & Gottschalk, G. (1996). Studies on the isopropylbenzene 2,3-dioxygenase genes encoded by the linear plasmid of *Rhodococcus erythropolis* BD2. *Microbiology* **142**, 3241–3251.

Komeda, H., Hori, Y., Kobayashi, M. & Shimizu, S. (1996a). Transcriptional regulation of the *Rhodococcus rhodochrous* J1 *nitA* gene encoding a nitrilase. *Proc Natl Acad Sci USA* **93**, 10572–10577.

Komeda, H., Kobayashi, M. & Shimizu, S. (1996b). Characterization of the gene cluster of high molecular mass nitrile hydratase (H-NHase) induced by its reaction product in *Rhodococcus rhodochrous* J1. *Proc Natl Acad Sci USA* **93**, 4267–4272.

Komeda, H., Kobayashi, M. & Shimizu, S. (1997). A novel transporter involved in cobalt uptake. *Proc Natl Acad Sci USA* **94**, 36–41.

Koonin, E. V. (1993). A superfamily of ATPases with diverse functions containing either classical or deviant ATP-binding motif. *J Mol Biol* **229**, 1165–1174.

Korch, C., Hagblom, P., Öhman, H., Göransson, M. & Normark, S. (1985). Cryptic plasmid of *Neisseria gonorrhoeae*: complete nucleotide sequence and genetic organization. *J Bacteriol* **163**, 430–438.

Kwon, H. J., Tirumalai, R., Landy, A. & Ellenberger, T. (1997). Flexibility in DNA recombination: structure of the lambda integrase catalytic core. *Science* **276**, 126–131.

Li, M. Z., Squires, C. H., Monticello, D. J. & Childs, J. D. (1996). Genetic analysis of the *dsz* promoter and associated regulatory regions of *Rhodococcus erythropolis* IGTS8. *J Bacteriol* **178**, 6409–6418.

Meijer, W. J. J., Smith, M., Wake, R. G., de Boer, A. L., Venema, G. & Bron, S. (1996). Identification and characterization of a novel type of replication terminator with bidirectional activity on the *Bacillus subtilis* plasmid pLS20. *Mol Microbiol* **19**, 1295–1306.

Motalebbi-Veshareh, M., Rouch, D. A. & Thomas, C. M. (1990). A family of ATPases involved in active partitioning of diverse bacterial plasmids. *Mol Microbiol* **4**, 1455–1463.

Nagy, I., Compernolle, F., Ghys, K., Vanderleyden, J. & De Mot, R. (1995a). A single cytochrome P-450 system is involved in degradation of the herbicides EPTC (*S*-ethyl dipropylthiocarbamate) and atrazine by *Rhodococcus* sp. NI86/21. *Appl Environ Microbiol* **61**, 2056–2060.

Nagy, I., Schoofs, G., Compernolle, F., Proost, P., Vanderleyden, J. & De Mot, R. (1995b). Degradation of the thiocarbamate herbicide EPTC (*S*-ethyl dipropylcarbamothioate) and biosafening by *Rhodococcus* sp. NI86/21 involve an inducible cytochrome P-450 system and aldehyde dehydrogenase. *J Bacteriol* **177**, 676–687.

Ogasawara, N. & Yoshikawa, H. (1992). Genes and their organization in the replication origin region of the bacterial chromosome. *Mol Microbiol* **6**, 629–634.

Piddington, C. S., Kovacevich, B. R. & Rambosek, J. (1995). Sequence and molecular characterization of a DNA region encoding the dibenzothiophene desulfurization operon of *Rhodococcus* sp. strain IGTS8. *Appl Environ Microbiol* **61**, 468–475.

Quan, S. & Dabbs, E. R. (1993). Nocardioform arsenic resistance plasmid characterization and improved *Rhodococcus* cloning vectors. *Plasmid* **29**, 74–79.

Rauzier, J., Moniz-Pereira, J. & Gicquel-Sanzey, B. (1988). Complete nucleotide sequence of pAL5000, a plasmid from *Mycobacterium fortuitum*. *Gene* **71**, 315–321.

Rossi, M., Brigidi, P., Gonzalez Vara y Rodriguez, A. & Matteuzzi, D. (1996). Characterization of the plasmid pMB1 from *Bifidobacterium longum* and its use for shuttle vector construction. *Res Microbiol* **147**, 133–143.

Shao, Z. Q. & Behki, R. (1995). Cloning of the genes for degradation of the herbicides EPTC (*S*-ethyl dipropylthiocarbamate) and atrazine from *Rhodococcus* sp. strain TE1. *Appl Environ Microbiol* **61**, 2061–2065.

Shao, Z., Dick, W. A. & Behki, R. M. (1995). An improved *Escherichia coli–Rhodococcus* shuttle vector and plasmid transformation in *Rhodococcus* spp. using electroporation. *Lett Appl Microbiol* **21**, 261–266.

Shields, D. C., Higgins, D. G. & Sharp, P. M. (1992). GCWIND: a microcomputer program for identifying open reading frames according to codon positional G+C content. *Comput Appl Biosci* **8**, 521–523.

Shinohara, M. & Itoh, T. (1996). Specificity determinants in interaction of the initiator (Rep) proteins with the origins in the plasmids ColE2-P9 and ColE3-CA38 identified by chimera analysis. *J Mol Biol* **257**, 290–300.

Siemieniak, D. R., Slightom, J. L. & Chung, S.-T. (1990). Nucleotide sequence of *Streptomyces fradiae* transposable element Tn*4556*: a class-II transposon related to Tn*3*. *Gene* **86**, 1–9.

Snapper, S. B., Melton, R. E., Mustafa, S., Kieser, T. & Jacobs, W. R., Jr (1990). Isolation and characterization of efficient plasmid transformation mutants of *Mycobacterium smegmatis*. *Mol Microbiol* **4**, 1911–1919.

Sohaskey, C. D., Im, H. & Schauer, A. T. (1992). Construction and application of plasmid- and transposon-based promoter-probe vectors for *Streptomyces* spp. that employ a *Vibrio harveyi* luciferase reporter cassette. *J Bacteriol* **174**, 367–376.

Stark, W. M., Boocock, M. R. & Sherratt, D. J. (1992). Catalysis by site-specific recombinases. *Trends Genet* **8**, 432–439.

Stolt, P. & Stoker, N. G. (1996a). Functional definition of regions necessary for replication and incompatibility in the *Mycobacterium fortuitum* plasmid pAL5000. *Microbiology* **142**, 2795–2802.

Stolt, P. & Stoker, N. G. (1996b). Protein–DNA interactions in the *ori* region of the *Mycobacterium fortuitum* plasmid pAL5000. *J Bacteriol* **178**, 6693–6700.

Taghavi, S., Provoost, A., Mergeay, M. & van der Lelie, D. (1996). Identification of a partition and replication region in the *Alcaligenes eutrophus* megaplasmid pMOL28. *Mol Gen Genet* **250**, 169–179.

Takechi, S. & Itoh, T. (1995). Initiation of unidirectional ColE2 DNA replication by a unique priming mechanism. *Nucleic Acids Res* **23**, 4196–4201.

Takechi, S., Matsui, H. & Itoh, T. (1995). Primer RNA synthesis by plasmid-specified Rep protein for initiation of ColE2 DNA replication. *EMBO J* **14**, 5141–5147.

Thomas, C. M. & Smith, C. A. (1986). The *trfB* region of broad host range plasmid RK2: the nucleotide sequence reveals *incC* and key regulatory gene *trfB*/*korA*/*korD* as overlapping genes. *Nucleic Acids Res* **14**, 4453–4469.

Turner, S. L. & Young, J. P. (1995). The replicator region of the *Rhizobium leguminosarum* cryptic plasmid pRL8JI. *FEMS Microbiol Lett* **133**, 53–58.

Vogt Singer, M. E. & Finnerty, W. R. (1988). Construction of an *Escherichia coli–Rhodococcus* shuttle vector and plasmid transformation in *Rhodococcus* spp. *J Bacteriol* **170**, 638–645.

Warhurst, A. W. & Fewson, C. A. (1994). Biotransformations catalyzed by the genus *Rhodococcus*. *Crit Rev Biotechnol* **14**, 29–73.

Weaver, K. E., Clewell, D. B. & An, F. (1993). Identification, characterization, and nucleotide sequence of a region of *Enterococcus faecalis* pheromone-responsive plasmid pAD1 capable of autonomous replication. *J Bacteriol* **175**, 1900–1909.

Williams, D. R. & Thomas, C. M. (1992). Active partitioning of bacterial plasmids. *J Gen Microbiol* **138**, 1–6.

Wu, X. & Roy, K. L. (1993). Complete nucleotide sequence of a linear plasmid from *Streptomyces clavuligerus* and characterization of its RNA transcripts. *J Bacteriol* **175**, 37–52.

Received 14 May 1997; revised 2 July 1997; accepted 8 July 1997.