Characterization of the replicator region of megaplasmid pTAV3 of Paracoccus versutus and search for plasmid-encoded traits

Dariusz Bartosik, Jadwiga Baj, Aneta A. Bartosik† and Mirosława Włodarczyk

Author for correspondence: Dariusz Bartosik. Tel: +48 22 554 13 44. Fax: +48 22 554 14 04. e-mail: bartosik@biol.uw.edu.pl

The replicon of the pTAV3 megaplasmid (approx. 400 kb) of Paracoccus versutus has been localized to a 4.3 kb EcoRI restriction fragment and its entire nucleotide sequence determined. The G+C content of the entire sequence is 66 mol%, which is within the range (62–66 mol%) previously determined for P. versutus total DNA. ORF1 encodes a replication initiation protein Rep (47-2 kDa), which shares substantial similarity with putative proteins of the Coxiella burnetii plasmids QpH1 and QpDV, and the replication protein of Pseudomonas syringae plasmid pPS10. ORF2, located in the opposite transcriptional orientation to ORF1, encodes a putative protein that shares similarity to a subfamily of ATPases involved in plasmid partitioning. The highest similarity was observed with homologous proteins (RepA) encoded by the repABC family of replicons found in several plasmids of Agrobacterium, Rhizobium and Paracoccus spp. The predicted product of ORF3 was similar to AcoR, Nif and NtrC transcriptional activators. A strong incompatibility determinant (inc) was localized between ORF1 (rep) and ORF2 (parA). The origin of replication of pTAV400 contains a short A+T-rich region and several imperfect palindromic sequences. Curing experiments demonstrated that the megaplasmid bears genes required for growth in minimal media and can therefore be referred to as a mini-chromosome. Megaplasmids pTAV3 of P. versutus UW1 and pKLW2 of Paracoccus pantotrophus DSM 11073 were found to carry closely related, incompatible replicons. It has been shown that plasmid pORI6 (containing oriV of pTAV3 cloned into plasmid pABW1, which does not replicate in Paracoccus spp.) can be trans activated not only by pTAV3, but also by pKLW2. Using pORI6, it was demonstrated that replication systems related to pTAV3 are also present in the replicons of Paracoccus alcaliphilus JCM 7364, Paracoccus thiocyanatus IAM 12816 and Paracoccus methylutens DM 12.

Keywords: Paracoccus versutus, pTAV3 megaplasmid, mini-replicon

INTRODUCTION

Megaplasmids are a group of large plasmids ranging in size from over 100 to 1700 kb. It seems that plasmids of this type are the most abundant in the α-Proteobacteria (Moreno, 1998; Jumas-Bilak et al., 1998), although they are also commonly identified in Pseudomonas strains (β-Proteobacteria), in which they determine the ability to degrade complex, toxic organic compounds (Boronin, 1992). The largest megaplasmids have been identified in rhizobia. Many of them carry genes related to the symbiotic properties of their hosts and ability to fix atmospheric nitrogen. Moreover, the plant pathogenicity of some Rhizobium species (until recently included in the genus Agrobacterium; Young et al., 2001) that invade the crown, roots and stems of many dicoty-
ledonous and some gymnospermy plants via wounds is determined by the presence of large tumour-inducing (Ti) or hairy-root-inducing (Ri) plasmids (Binns et al., 1992; Zhu et al., 2000). Megaplasmids can encode many other phenotypic traits. For example, the megaplasmid localization of genes involved in lithoautotrophic metabolism (Kortluke et al., 1987) and resistance to heavy metals (Mergay et al., 1985) has been reported for strains of *Ralstoniaeutropha* HG1 or CH34, respectively (formerly *Alcaligenes eutrophus*). In recent years, complete sequences of several bacterial megaplasmids have been reported. Particular interest is evoked by the largest megaplasmids, sometimes referred to as ‘minichromosomes’ (Krawiec & Riley, 1990; Friedrich & Schwartz, 1993; Winterstein & Ludwig, 1998). It prompts consideration of the choice of criteria that would allow the distinction between megaplasmids and chromosomes, since size alone cannot be used for this distinction.

The presence of megaplasmids in bacteria is probably more common than indicated by current literature. Plasmid patterns of newly analysed strains usually involve standard alkaline lysis procedures as well as conditions favouring the identification of small and medium-sized plasmids in the form of CCC DNA. Under such conditions, megaplasmid DNA is fragmented, thus precluding the identification of these molecules and their physical and chemical characterization. During the last decade, more than ten new species have been classified to the genus *Paracoccus* (which, like the *Rhizobiales*, is included among the *z*-Proteobacteria) and these are frequently capable of degrading compounds that are toxic to other organisms. A preliminary analysis of plasmid patterns in several strains, representing 11 out of 14 species of the genus *Paracoccus*, enabled the identification of over 20 plasmids with sizes of up to approx. 100 kb (Baj et al., 2000). However, it was not until the employment of the Eckhardt procedure (in gel lysis) modified by Wheatcroft et al. (1990) that we were able to show that all the examined strains (except for *Paracoccusalkenifer*) carry megaplasmids of a size roughly determined at 400–500 kb (Baj et al., 2000). These megaplasmids were never subject to any genetic studies. It may thus be possible that they carry genes involved in diverse catabolic pathways determining the bioremediation potential of their hosts.

In this paper we present the characteristics of the replication system of the megaplasmid pTAV3 (approx. 400 kb) that has been identified in *Paracoccus versutus* and the results of an analysis aimed at determining plasmid-encoded phenotypic traits. *P. versutus*, besides pTAV3, carries also a composite plasmid pTAV1 (107 kb), which is, so far, the most thoroughly studied plasmid in the genus *Paracoccus* (Bartosik et al., 1997b, 1998).

**METHODS**

**Bacterial strains, plasmids and growth conditions.** Bacterial strains and plasmids used in this study are listed in Table 1. Cultures were grown in Luria broth (LB) (Sambrook et al., 1989) at 30 °C (*P. versutus*, *Paracoccus pantotrophus*) or 37 °C (*Escherichia coli*). The synthetic medium used for *P. versutus* was as described by Wood & Kelly (1977) with sodium succinate (10 mM) as carbon source. For chemolithoautotrophic growth, succinate was replaced by Na2S2O3 (20 mM). Antibiotics included in the media were used at the following final concentrations (µg ml−1): kanamycin (Km), 50; rifampicin (Rif), 50; tetracycline (Tc), 3 for *P. versutus* and 20 for *E. coli*. For growth curves experiments, the overnight cultures grown in LB were spun down, washed and resuspended in appropriate medium before incubation. OD490 was measured at 1 h intervals using a Pharmacia Biotech Nova- spec II spectrophotometer.

**DNA manipulations.** Plasmid DNA was isolated according to the method of Birnboim & Dol in (1979) and when necessary purified by CsCl/ethidium bromide gradient centrifugation. Total DNA of *P. versutus* was isolated by phenol extraction (Williams et al., 1998). Megaplasmid visualization was done using the Eckhardt (1978) procedure or its simplified version (Wheatcroft et al., 1990). Cloning experiments, digestion with restriction enzymes, ligationa and agarose gel electrophoresis were conducted in accordance with standard procedures as described by Sambrook et al. (1989). All enzymes were purchased from either Promega or Roche. DNA restriction fragments were purified from agarose with DNA-Gel-Out kit (DNA Gdansk). For Southern hybridization (Sambrook et al., 1989), DNA probes were labelled with digoxigenin (Roche). Hybridization and visualization of hybridization products was carried out as recommended by the supplier.

**DNA sequencing and analysis.** The nucleotide sequence was determined in the DNA Sequencing and Oligonucleotides Synthesis Laboratory, Institute of Biochemistry and Biophysics, Polish Academy of Science, using a terminator sequencing kit and an automatic sequencer (ABI 377; Perkin Elmer). Four restriction fragments obtained after digestion with enzyme *Pst*I of a 4.3 kb fragment of EcoRI pTAV400 (0.2 kb, 1.2 kb, 1.3 kb and 1.6 kb), covering the entire genome of the minireplicon, were cloned into *E. coli* pBluescript KSII vector and the inserts were sequenced starting with universal and reverse primers and then with primers complementary to the previously determined sequence. Sequence analysis was done with programs included in the UWGCG Package (Devereux et al., 1984). Comparison searches through the databases were performed with the BLAST program provided by the National Center for Biotechnology Information (http://www.ncbi.nlm.nih.gov/BLAST). The alignment of the sequences was performed with thePILEUP program (GCC software).

**Electroporation and transformation.** Electroporation was carried out at 2500 V, 25 µF and 200 Ω (for *E. coli*) or 400 Ω (for *P. versutus* and *P. pantotrophus*), in a gene pulser apparatus (Bio-Rad) according to a modified Bio-Rad procedure (Wlodarczyk et al., 1994). Electrottransformants were selected on solidified LB medium supplemented with appropriate antibiotics. Competent cells of *E. coli* TG1 were prepared and transformed as described by Kushner (1978).

**Triparental mating.** The overnight cultures of the donor strain *E. coli* TG1 carrying the mobilizable recombinant vector based on pABW1 (Bartosik et al., 1997a) or pRK415 (Keen et al., 1988) were grown, spun down and washed twice to remove antibiotics. These were mixed at a ratio of 1:2:1 with a suitable recipient strain and *E. coli* DH5α carrying the helper plasmid pRK2013. An aliquot of 100 µl of such a mixture was spread on a plate with solidified LB medium. After overnight
Table 1. Bacterial strains and plasmids used in this work

<table>
<thead>
<tr>
<th>Strain or plasmid</th>
<th>Relevant characteristics</th>
<th>Reference or source</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Strains</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>P. versutus</em></td>
<td></td>
<td></td>
</tr>
<tr>
<td>UW1</td>
<td>Wild-type (first described as <em>Thiobacillus</em> sp. strain A2)</td>
<td>Taylor &amp; Hoare (1969)</td>
</tr>
<tr>
<td>UW1R</td>
<td>Rif(^r) derivative of UW1</td>
<td>This work</td>
</tr>
<tr>
<td>UW312</td>
<td>Rif(^r), pTAV3-less derivative of UW1R</td>
<td>This work</td>
</tr>
<tr>
<td><em>P. alcalophilus</em> JCM 7364R</td>
<td>Rif(^r) derivative of the wild-type strain JCM 7364</td>
<td>This work</td>
</tr>
<tr>
<td><em>P. aminophilus</em> JCM 7686R</td>
<td>Rif(^r) derivative of the wild-type strain JCM 7686</td>
<td>This work</td>
</tr>
<tr>
<td><em>P. aminovorans</em> JCM 7685R</td>
<td>Rif(^r) derivative of the wild-type strain JCM 7685</td>
<td>This work</td>
</tr>
<tr>
<td><em>P. methylutens</em> DM 12R</td>
<td>Rif(^r) derivative of the wild-type strain DM 12</td>
<td>This work</td>
</tr>
<tr>
<td><em>P. thiocyhanus</em> IAM 12816R</td>
<td>Rif(^r) derivative of the wild-type strain IAM 12816</td>
<td>This work</td>
</tr>
<tr>
<td><strong>Strains</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>P. pantotrophus</em> KL100</td>
<td>Rif(^r) derivative of the wild-type (DSM 11073) deprived of its natural plasmid pKLW1</td>
<td>This work</td>
</tr>
<tr>
<td>KL101</td>
<td>Derivative of strain KL100 deprived of megaplasmid pKLW2</td>
<td>This work</td>
</tr>
<tr>
<td><strong>Plasmids</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pTAV3</td>
<td>Megaplasmid (approx. 400 kb) of <em>P. versutus</em> UW1</td>
<td>This work</td>
</tr>
<tr>
<td>pTAV400</td>
<td>Mini-replicon carrying 4-3 EcoRI fragment of pTAV3 plus Km(^r) cassette (1-3 kb) derived from pUC4-K</td>
<td>This work</td>
</tr>
<tr>
<td>pUC4-K</td>
<td>Source of Km(^r) cassette</td>
<td>Vieira &amp; Messing (1982)</td>
</tr>
<tr>
<td>pBGS18</td>
<td>Km(^r), cloning vector, ColE1 origin of replication</td>
<td>Spratt et al. (1986)</td>
</tr>
<tr>
<td>pBBW100</td>
<td>Km(^r), derivative of pBGS18 containing 3.2 kb Dral–EcoRI fragment of pTAV400</td>
<td>This work</td>
</tr>
<tr>
<td>pBBW101</td>
<td>Km(^r), derivative of pBGS18 containing 2.9 kb KpnI–EcoRI fragment of pTAV400</td>
<td>This work</td>
</tr>
<tr>
<td>pBBW102</td>
<td>Km(^r), derivative of pBGS18 containing 2.6 kb EcoRI–SphI fragment of pTAV400</td>
<td>This work</td>
</tr>
<tr>
<td>pBBW103</td>
<td>Km(^r), derivative of pBGS18 containing 3 kb EcoRI–ApaI fragment of pTAV400</td>
<td>This work</td>
</tr>
<tr>
<td>pBBW104</td>
<td>Km(^r), derivative of pBBW105 with deletion of 0.3 kb BclI fragment</td>
<td>This work</td>
</tr>
<tr>
<td>pBBW105</td>
<td>Km(^r), derivative of pBGS18 containing 4.3 kb fragment EcoRI of pTAV400</td>
<td>This work</td>
</tr>
<tr>
<td>pABW1</td>
<td>Km(^r), mobilizable cloning vector based on pBGS18, oriT of RK2</td>
<td>Bartosik et al. (1997a)</td>
</tr>
<tr>
<td>pORI6</td>
<td>Km(^r), derivative of pABW1 containing two clusters of Sau3AI fragments (1 kb) carrying oriV of pTAV400</td>
<td>This work</td>
</tr>
<tr>
<td>pABW1/S1</td>
<td>Km(^r), derivative of pABW1 containing S1 cluster of Sau3AI fragments (451 bp)</td>
<td>This work</td>
</tr>
<tr>
<td>pABW1/S2</td>
<td>Km(^r), derivative of pABW1 containing S2 cluster of Sau3AI fragments (576 bp)</td>
<td>This work</td>
</tr>
<tr>
<td>pRK415</td>
<td>Te(^c), mobilizable broad-host-range cloning vector, RK2 oriT and origin of replication</td>
<td>Keen et al. (1988)</td>
</tr>
<tr>
<td>pRK415/10</td>
<td>Te(^c), pRK415 derivative containing 1–6 kb fragment EcoRI–PstI of pTAV400</td>
<td>This work</td>
</tr>
<tr>
<td>pRK415/11</td>
<td>Te(^c), pRK415 derivative containing 1–2 kb fragment PstI of pTAV400</td>
<td>This work</td>
</tr>
<tr>
<td>pRK415/12</td>
<td>Te(^c), pRK415 derivative containing 1–3 kb fragment PstI of pTAV400</td>
<td>This work</td>
</tr>
<tr>
<td>pRK415/13</td>
<td>Te(^c), pRK415 derivative containing 0–2 kb fragment PstI–EcoRI of pTAV400</td>
<td>This work</td>
</tr>
<tr>
<td>pRK415/NRU</td>
<td>Te(^c), pRK415 derivative containing 0–57 kb fragment NruI of pTAV400</td>
<td>This work</td>
</tr>
<tr>
<td>pRK2013</td>
<td>Km(^r), helper plasmid carrying RK2 tra genes</td>
<td>Ditta et al. (1980)</td>
</tr>
</tbody>
</table>

Incubation at 30 °C, the bacteria were washed off the plate and suitable dilutions were plated on selective media containing rifampicin (selective marker of the recipient strain) and other appropriate antibiotic (Tc or Km) to select transconjugants. Spontaneous resistance of the recipient strains to Km and Tc was undetectable under these experimental conditions.
**Plasmid stability.** *P. versutus* or *P. pantotrophus* strains containing tested plasmids were grown in LB medium supplemented with appropriate antibiotic. The late-exponential-phase cells were diluted 10^-3 in LB medium without the antibiotic and grown at 30°C. Cultures were diluted 10^-3 at 24 h intervals. Samples taken at each dilution step were plated on solidified LB medium for estimation of the number of generations. From these plates, 200 colonies were tested for the presence of Km\(^r\) or Tc\(^r\) marker by replica plating. The retention of plasmids was determined as the percentage of kanamycin-resistant colonies.

**Incompatibility testing.** The incompatibility characteristics of two plasmids were examined by conjugal transfer of the plasmid to be tested into recipient *P. versutus* UW1R (carrying pTAV3). Transconjugants were selected for the incoming plasmid. The plasmid pattern of the transconjugants was verified by screening 10 colonies by a rapid alkaline lysis and agarose gel electrophoresis and by the modified Eckhardt procedure (Wheatcroft et al., 1990).

**RESULTS**

**Construction and characterization of mini-replicon pTAV400**

A mini-replicon was constructed using total DNA from wild-type *P. versutus* strain UW1 that was digested with PstI, EcoRI, BamHI and then ligated with an appropriately restricted kanamycin-resistance cassette derived from plasmid pUC4-K. The ligation mixture was introduced by electroporation into *P. pantotrophus* KL100 (phylogenetically closely related to *P. versutus*) in order to avoid incompatibility between pTAV3 and the constructed mini-replicon. Km\(^r\) transformants were obtained only when a DNA preparation digested with EcoRI was used. They contained a replicon 5.8 kb in size (designated pTAV400), composed of two EcoRI fragments of 4.5 kb and 1.3 kb (Km\(^r\) cassette). The restriction map of the replication system pTAV400 is presented in Fig. 1.

The origin of the cloned mini-replicon was confirmed by hybridization analysis. Using the constructed mini-replicon pTAV400 as a probe, a positive result was obtained for DNA of megaplasmid pTAV3, whilst the result of hybridization with co-residing plasmid pTAV1 of *P. versutus* (Fig. 2a, lane 1) as well as with another, larger megaplasmid reported earlier by Wlodarczyk & Piechucka (1995) was negative. The latter, designated pTAV4 (Fig. 2a, lane 1), could be visualized only with prolonged electrophoretic separation of DNA.

An interesting observation was that the strains of *P. pantotrophus* KL100 carrying pTAV400 were devoid of megaplasmid pKLW2, which occurred in the original strain (Fig. 2a, lane 3 and 4). This result points to the incompatibility of the replication systems of megaplasmids pTAV3 and pKLW2 and indicates their close relatedness. This conclusion was supported by the positive result of hybridization between pTAV400 and pKLW2 (Fig. 2b, lane 3).

The mini-replicon pTAV400 is very unstable in culture under lack of selective pressure. After 24 h growth (approx. 10 generations) only 5% of the cells still carried the plasmid, whereas after 48 h none of the tested colonies contained pTAV400. One of the ‘cured’ clones of *P. pantotrophus*, missing both pKLW2 and

---

**Fig. 1.** Mutational analysis of the replicator region of pTAV400. Only the restriction sites used for plasmid construction are given on the pTAV400 restriction map. The three ORFs described in the text and their transcriptional orientation are shown by arrows (open arrows represent incomplete ORFs). Two regions of the A+T sequence enrichment are boxed and designated A1 and A2. The DNA region carrying the cis-acting incompatibility determinant has been marked as a solid box and indicated *inc*. The S1 (containing origin of replication) and S2 DNA fragments of pTAV400 are boxed and correspondingly indicated (see text for details). The open boxes represent DNA restriction fragments of pTAV400 cloned into pBG518 vector and tested for replication ability. The thin line joining the boxes indicates the region of pTAV400 removed by deletion. The plasmid derivative names are given on the right. In the column labelled Replication, + indicates that the plasmid derivative has the ability to replicate in *P. pantotrophus* KL101; – indicates that the plasmid is unable to replicate within this strain.
pTAV400 (Fig. 2, lane 4), was designated KL101 and used in further studies. Since the observed low stability could have resulted from the use of *P. pantotrophus* as a host for pTAV400, the experiment was repeated with strain UW312 of *P. versutus*, which was devoid of megaplasmid pTAV3 (construction of this strain is described below). Similar results with respect to stability were obtained, which indicates that the restriction fragment of pTAV3 cloned into pTAV400 does not carry stabilizing systems that are commonly coded by low-copy-number bacterial plasmids (Gerdes et al., 2000).

Attempts to introduce pTAV400 into *E. coli* TG1 by electroporation met with failure, which indicates that the cloned replication system does not function in this host.

**Nucleotide sequence analysis of pTAV400**

The complete nucleotide sequence of pTAV400 was determined (4371 bp). The overall G+C content of the sequence was 66 mol%, which is within the range previously determined for *P. versutus* total DNA (from 62 to 66 mol%). The sequence does not show any significant similarity to sequences deposited in the databases. Ten ORFs encoding putative proteins of the size above 100 aa were identified. Their amino acid sequences were compared (BLASTP) with the databases. Only three of them (embracing approx. 70% of the sequence of the replicon) showed significant sequence similarities to proteins encoded by other bacterial plasmids. We found ORF1 encoding a putative peptide of 419 aa, with predicted molecular mass 47.2 kDa, and pI of 9.6, showing similarities to several plasmid replication initiation proteins. The highest similarities were observed with ORF311 of QpH1 (40% identity and 49% similarity) (Thiele et al., 1994), ORF434 of QpDV of Coxiella burnetii (34% identity and 45% similarity) (accession no. NP 052842), and with Rep proteins of several pseudomonad plasmids, e.g. pPS10 of *Pseudomonas syringae* (28% identity and 37% similarity) (Nieto et al., 1992). The similarities are shared mainly between the N-terminal and central parts of the proteins whereas the C-terminal part is much less conserved (Fig. 3). Local similarities were also observed with several other plasmid-encoded Rep proteins, e.g. of pFA3 of Neisseria gonorrhoeae (Gilbride & Brunton, 1990), pSC101 of *E. coli* (Armstrong et al., 1984), pOM1 of Francisella tularensis (accession no. NP 052243) (data not shown). Based on this analysis, we speculated that ORF1 (*rep*) encodes the megaplasmid replication initiation protein.

The incomplete ORF2 (lacking its terminal 3’ end) which is located in the opposite transcriptional orientation to *rep* (Fig. 1), encodes a putative protein (343 aa) with significant similarity to the family of SopA/ParA, Walker-type ATPases involved in active partitioning of bacterial plasmids (Williams & Thomas, 1992; Koonin, 1993) (data not shown). The highest similarity was observed with RepA-type proteins encoded by *repABC* family of replicons found in many plasmids of *Agrobacterium*, *Rhizobium* and *Paracoccus* spp. (Turner et al., 1996; Bartosik et al., 1998; Palmer et al., 2000). For instance, the putative ParA of pTAV400 showed 34% identity and 45% similarity with Y4Ck of rhizobial plasmid pNGR234a (536 kb; Freiberg et al., 1997) and 30% identity, 47% similarity with RepA of paracoccal composite plasmid pTAV1 (107 kb; Bartosik et al., 1998). Local similarities were also observed with ParA of megaplasmid pMOL28 (180 kb) of *Ralstonia eutrophora* (Taghavi et al., 1996).

Within pTAV400, a second incomplete ORF lacking its 5’ end was localized (ORF3; Fig. 3). It coded for a fragment (254 aa) of a potential polypeptide whose amino acid sequence shows considerable similarity to the C-terminal parts of the transcription activators AcoR, NtrC and NifA. The greatest similarity was observed with the AcoR proteins of *Ralstonia eutropha* (Krüger & Steinbüchel, 1992) and *Pseudomonas aeruginosa* (Stover et al., 2000) (approx. 51% identity and 63% similarity). The terminal part of the putative ORF3-encoded protein also contained the HTH motif (helix–turn–helix, GCG software) that was located similarly as in the above-mentioned regulatory proteins (data not shown).

Analysis of structural features of the sequence did not reveal the presence of iteron-like repeats characteristic for the origin of replication (*oriV*) of many theta-replicating plasmids. The G+C profile of the sequence revealed the presence of two short A+T-rich regions (termed A1 and A2; Fig. 1). Analogous sequences frequently occur within *oriV*, where they constitute a site at which the DNA duplex unwinds and initiation of replication occurs. Region A1 (163 bp; 57 mol% A+T, in contrast to 44 mol% overall) occurs within a short intergenic region between *parA* and *rep* (Fig. 1; the
Fig. 3. Alignment of deduced amino acid sequences of Rep of pTAV400 (bold type), ORF434 of Coxiella burnetii QpDV (accession no. NP 052842), ORF311 of C. burnetii QpH1 (accession no. NP 052334) and RepA of Pseudomonas syringae pPS10 (accession no. S20615). Dots indicate gaps introduced to maximize the alignment. Amino acids identical to Rep of pTAV400 are shown against a dark background. Asterisk indicates the stop codon. Amino acid numbering is shown at the right of the sequence lines.

nucleotide sequence of A1 is shown in Fig. 4), whereas A2 (92 bp; 60% A + T) is located 354 bp downstream of ORF1 (rep) (Fig. 1).

Deletion analysis of pTAV400

To determine the region of the minimal replicon of pTAV400 and to demonstrate the role of ORF1 in replication of the plasmid, derivatives of pTAV400 were constructed. To this end, selected restriction fragments from pTAV400 (Fig. 1) were cloned into the multiple cloning site (MCS) vector pBGS18 (specific for E. coli and unable to replicate in Paracoccus species) and the plasmids obtained were introduced by means of electroporation into P. pantotrophus strain KL101(pKLW2−).

It was observed that the deletion of fragment ApaI–EcoRI (pBBW103; Fig. 1), embracing the terminal part of the mini-replicon, had no effect on plasmid replication. It appeared, however, that plasmids in which deletions were made in the terminal 5′ part of ORF1 (pBBW101; Fig. 1) or its 3′ end (pBBW102; Fig. 1) did not replicate in P. pantotrophus. It was also shown that a deletion of a 264 bp BclI fragment within ORF1 prevented the replication of plasmid pBBW104 (Fig. 1), indicating the crucial role of ORF1 (rep) in replication of pTAV400. Interestingly, the removal of the proximal part of the replicon (1.13 kb EcoRI–DraI fragment), embracing ORF2 (parA) and part of the intergenic region between parA and rep (pBBW100; Fig. 1), together with the part of the A+T-rich sequence A1 (Fig. 4), completely abolished the replication. This result suggested that the deleted fragment could have contained the origin of replication of the plasmid or, more likely, the promoter of the rep gene.

Localization of origin of replication

To locate the origin of replication of pTAV400, the DNA of this mini-replicon was partially digested with Sau3AI and the restriction fragments were cloned into E. coli-specific mobilizable vector pBGS18 (specific for E. coli and unable to replicate in Paracoccus species) and the plasmids obtained were introduced by means of electroporation into P. pantotrophus strain KL101(pKLW2−).

It was observed that the deletion of fragment ApaI–EcoRI (pBBW103; Fig. 1), embracing the terminal part of the mini-replicon, had no effect on plasmid replication. It appeared, however, that plasmids in which deletions were made in the terminal 5′ part of ORF1 (pBBW101; Fig. 1) or its 3′ end (pBBW102; Fig. 1) did not replicate in P. pantotrophus. It was also shown that a deletion of a 264 bp BclI fragment within ORF1 prevented the replication of plasmid pBBW104 (Fig. 1), indicating the crucial role of ORF1 (rep) in replication of pTAV400. Interestingly, the removal of the proximal part of the replicon (1.13 kb EcoRI–DraI fragment), embracing ORF2 (parA) and part of the intergenic region between parA and rep (pBBW100; Fig. 1), together with the part of the A+T-rich sequence A1 (Fig. 4), completely abolished the replication. This result suggested that the deleted fragment could have contained the origin of replication of the plasmid or, more likely, the promoter of the rep gene.
Fig. 4. Nucleotide sequence of the pTAV400 DNA fragment containing the intergenic region located between the *parA* (ORF2) and *rep* (ORF1) genes. The deduced ParA and Rep protein products are given below the sequence in the single-letter amino acid code in the first position of each codon. The putative start codon of *parA* (ATG), located within ORF2, is boxed. The −10 and −35 boxes of putative promoters are in bold type and underlined. Presumed RBSs are in bold type and double underlined. Relevant restriction sites are indicated. The numbers on the right refer to the deposited nucleotide sequence (accession no. AF390867). The S1 region, containing the origin of replication, is boxed and correspondingly indicated. The sequence of A + T enrichment is shown on a black background. The imperfect direct (DR) and inverted (IR) repeated sequences are indicated with arrows.

The S1 and S2 clusters were cloned independently into pABW1 (pABW1/S1 and pABW1/S2, respectively). Only the presence of S1 allowed the replication of the plasmid in strain UW1R. S1 thus contains the origin of replication of megaplasmid pTAV3. Analysis of the nucleotide sequence of this region revealed the presence of several imperfect inverted (IR) and direct repeats (DR), which are indicated in Fig. 4. No sequences were detected that would match the consensus sequence of the DnaA box (TTATCCACA), commonly found within oriV of theta-replicating plasmids (Helinski *et al*., 1996).

Incompatibility analysis

Replication systems of low-copy-number plasmids contain regulatory elements overseeing the process of replication initiation. These are frequently incompatibility determinants (*inc*) precluding the co-occurrence of two related plasmids in a cell. To localize the determinant(s) of incompatibility, several restriction frag-
Characterization of strains lacking pTAV3

We obtained (by incompatibility with pRK415/NRU) clones of *P. versutus* deprived of pTAV3. By growing one of them in the absence of selective pressure, strain UW312 (which spontaneously lost pRK415/NRU) was obtained. The absence of an autonomous form of pTAV3 was confirmed by electrophoretic analysis (Fig. 2a, lane 2). The negative result of hybridization of pTAV400 to total DNA of UW312 ruled out the possibility of the integration of pTAV3 into the chromosome (data not shown).

The comparison of several physiological functions of the megaplasmid-free strains with the parental UW1 revealed that these strains were unable to grow in minimal liquid media (Fig. 5) and failed to form normal colonies on agar of similar composition. Colonies appearing after prolonged (72–96 h) incubation of plates were very distinct (flat, transparent, hardly visible) from the wild-type ones grown after 48 h incubation. The substitution of succinate (commonly used as a carbon source in minimal media for the heterotrophic growth of *P. versutus*) with other organic compounds (glucose, saccharose, p-hydroxybenzoate or methylamine) or thiosulphate (for autotrophic growth) did not enable growth to occur in liquid media. Moreover, it had no influence on the morphology of colonies. There was no requirement for a defined growth factor. For all the megaplasmid-free strains, a rich medium resulted in better growth. The influence of two different concentrations of Casamino acids or of yeast extract on growth of strain UW312 is shown in Fig. 5. All the megaplasmid-free strains behaved similarly. Their growth in rich LB medium was identical to that of the parental strain UW1 (data not shown).

We also proved that pTAV3 is not involved in encoding gentamicin resistance, thiosulphate oxidation or ability to denitrify (all features characteristic for wild-type *P. versutus*). The role of pTAV3 in *P. versutus* metabolism remains to be elucidated; it seems, however, clear that the megaplasmid encodes genes required for growth in minimal media.

Identification of type pTAV3 replicons in the genus *Paracoccus*

The screening for the occurrence of extrachromosomal DNA by the modified Eckhardt gel electrophoresis procedure in several strains representing 11 species of *Paracoccus* revealed the presence of megaplasmids in the majority of them (Baj et al., 2000). We were interested to find if they carried replicons related to those of pTAV3.

As demonstrated in this study (using hybridization and incompatibility analysis), one of them, megaplasmid pKLW2 of *P. pantotrophus* DSM 11073, carried a replication system that is related to the pTAV3 system. Interestingly, it appeared that plasmid pORI6 (which contains oriV of pTAV3) replicates in strain KL100 (pKLW2^+), but is unable to replicate in strain KL101, which is devoid of this megaplasmid. This indicates that both pKLW2 and pTAV3 provide trans-acting factors, which are indispensable for replication of the pORI6. This result encouraged us to use pORI6 as a sensitive probe enabling the identification of replication systems closely related to pTAV3.

Plasmid pORI6 was thus introduced, via triparental mating, into Rif^r^ derivatives of available Paracoccus species: *P. alcalophilus* JCM 7364, *P. aminophilus* JCM 7686, *P. aminovorans* JCM 7685, *P. methylutens* DM 12, *P. thiocyranatus* IAM 12816. Km^r^ transconjugants were obtained in the case of *P. alcalophilus*, *P. thiocyranthus* and *P. methylutens* (which contain the megaplasmids pALC2, pTHI3 and pMTH3, respectively) but the presence of the autonomous form of pORI6 was determined in only the first two species. In the case of *P. methylutens*, pORI6 most likely formed co-integrates
with the megaplasmid as a result of a recombination event. This hypothesis is favoured by the results of hybridization of pTAV400 (probe) with P. methylutens pMTH3 DNA (data not shown). This result shows that pMTH3 carries sequences homologous with those of the replicator region of pTAV3, present in pORI6.

In conclusion, it can be said that replication systems related to pTAV3 are present in the replicons of P. pantotrophus DSM 11073, P. alcaliphilus JCM 7364, P. thiocyanatus IAM 12816 (in these cases activation of DSM 11073, initiation of replication (Nieto rrons) that play a crucial role in the regulation of plasmid pPS10 contains long repeated sequences (ite- 

DISCUSSION

This paper presents the results of the molecular character- ization of the replication system of the natural megaplasmid pTAV3 present in P. versutus. This is the first such analysis of one of the megaplasmids from Paracoccus that are ubiquitous in this group of bacteria. As a result of deletion analysis of the replicator region of pTAV3, designed on the basis of a nucleotide sequence of mini-replicon pTAV400, a gene encoding the replication protein Rep was identified. The amino acid sequence of this protein shows a striking similarity to the protein products of conserved ORFs present in plasmids QpH1 and QpDV of Coxiiella burnetii (whose replication systems have not yet been identified), as well as to the replication protein of the Pseudomonas syringae plasmid pPS10. The origin of replication of plasmid pPS10 contains long repeated sequences (ite- rons) that play a crucial role in the regulation of initiation of replication (Nieto et al., 1992). The replication origin of pTAV400, which is located up- stream of the rep gene, within a short intergenic stretch, has a different structure. Its sequence lacks iteron-like repeats, but contains several imperfect inverted repeats (Fig. 4), some of which are located within a short A +T- rich sequence. As revealed in the course of analysis, deletion of part of the stretch enriched in A +T pairs completely inhibits replication, which suggests that the sequence plays a key role in initiation of replication. Analogous sequences have frequently been identified within the origin region of other theta-replicating plasmids (Helsinki et al., 1996). The functionality of this sequence as the origin of pTAV3 replication has been confirmed in experiments in which the region cloned into the vector pABW1, which is inactive in members of the genus Paracoccus (resulting in plasmid pORI6), could be trans activated by a pTAV3-encoded factor (Rep).

We have demonstrated that the replication system of pTAV3 is not unique among Paracoccus megaplasmids. As a rule, the identification of related groups of replicons makes use of replicon probes and involves hybridization or PCR techniques. However, such analysis allows the determination of the degree of homology between the nucleotide sequences of selected fragments of the studied replicons. In our studies we employed plasmid pORI6 carrying a sensitive probe (in the form of oriV), which enables the identification of the functional similarity of the replicons studied. It appeared that the origin of replication of pTAV3 (present in pORI6) was trans activated by other mega-replicons present in different Paracoccus species (P. pantotrophus, P. alcaliphilus, P. thiocyanatus). This points to the close phylogenetic relationship between these replicons and suggests that the model of the replication system of pTAV3 suggested may also be valid for other megaplasmids found in the genus Paracoccus.

The constructed mini-replicon pTAV400 is very un- stable since it does not contain a complete partitioning sys- tem (par), which ensures that plasmid genomes are actively and accurately distributed between daughter cells at cell division. In general these systems consist of three elements: two trans-acting genes organized in an operon and an adjacent cis-required centromere-like sequence (the determinant of incompatibility). Stability is ensured only by the presence of all three components of the system. pTAV400 carries only a fragment of the first gene of a potential operon, encoding a protein that is homologous to Walker-type partitioning ATPases, which are involved in autoregulation of the par operons (Gerdes et al., 2000). The potential protein ParA of pTAV400 shows the greatest similarity to the RepA proteins, coded by replicons of the repABC type. In the repABC replicons, the genes encoding active partitioning proteins (repA and repB) and the structural genes for a replicator protein (repC) are clustered in an operon, thus ensuring a unique regulatory relation between genes involved in both these processes (Ramirez-Romero et al., 2001; Bartosik et al., 2001).

In plasmid pTAV3, the replicator region and potential active partitioning system also are in close proximity and a key role in their regulation seems to be played by an intergenic sequence located between the genes rep and paraA. This sequence, besides the origin of replication, also contains potential promoters for both genes as well as strong determinant(s) of incompati- bility, whose introduction in trans results in the loss of the residing megaplasmid pTAV3. Taking into account that both the partitioning ATPases as well as many other Rep proteins bind to operator sequences in the promoter regions of their structural genes and consequently negatively regulate the level of their expression, it seems likely that similar interactions may also occur in the intergenic region of pTAV400. The promoter regions of the paraA and rep genes contain direct (DR1, DR2) and inverted (IR5, IR5′) repeated sequences (Fig. 4), which may constitute potential sites of interaction between DNA and plasmid-encoded proteins. We have initiated studies to identify the remaining elements of the partitioning system of pTAV3. This will be a starting point for further analysis aimed at determining the interactions of the ParA and Rep proteins within the intergenic sequence and identifying their effect on the expression of genes involved in plasmid replication or partitioning.

In the terminal part of pTAV400, an incomplete ORF has been identified. Its protein product shows significant
similarity to the family of conserved transcriptional activators of $\alpha^{41}$-dependent promoters. The greatest similarity was to the AcoR proteins, which in the case of Ralstonia eutropha and Bacillus subtilis regulate the expression of genes involved in the catabolism of acetoin (Krüger & Steinbüchel, 1992; Huang et al., 1999). However, it will be necessary to obtain the complete sequence of ORF3 (acoR) of pTAV3, as well as of the DNA region located upstream of this gene, to allow the study of the putative role of AcoR in the regulation of expression of pTAV3-encoded genes.

The megaplasmid-free strains obtained by incompatibility allowed us to start a search for functions encoded by pTAV3. We found that the strains deprived of pTAV3 are missing some undefined genes required for growth in minimal media. The definition of a plasmid implies being dispensable for the host cell, so it should not carry any ‘housekeeping’ genes. Our observations prompt the question as to whether the genetic element designated as pTAV3 should be regarded as a plasmid or as a so-called mini-chromosome. An additional element that impedes the definition of the functions encoded by individual megaplasmids is the possible co-operation of genes situated on different replicons, as exemplified by genes encoding rhizobial symbiotic functions. We have not tested pTAV3 for the presence of genes encoding rRNA. However, when one takes into account the report of Kunnimalaiyaan et al. (2001) on the identification of rRNA operons within non-essential plasmid DNA in Bacillus megaterium, the criterion of the presence or absence of genes encoding rRNA as differentiating between chromosome and plasmid seems invalid.

Winterstein & Ludwig (1998) showed that the genomes of two of the species being phylogenetically closest to P. versutus (Paracoccus demiridrius and P. pantotrophus) consist of three to four ‘large DNA species’ designated CI–CIV, in the size range 0.5–2.2 Mb, some of which could be linear. The presence of pTAV3 and pTAV4 in P. versutus suggests that the multireplicon genomic organization may be characteristic for the entire genus Paracoccus. This view is consistent with suggestions of unconventional genomic organization in $\alpha$-Proteobacteria (Jumas-Bilak et al., 1998).

ACKNOWLEDGEMENTS

This work was done within the MECBAD co-operative project (BIO4-CT-0099). Financial support was obtained from Warsaw University, Poland (BW 420/17/98 and BW 1455/6/99).

REFERENCES


Improved broad-host-range plasmids for DNA cloning in Gram-negative bacteria. Gene 70, 191–197.


Received 6 July 2001; revised 21 November 2001; accepted 22 November 2001.