Sequence Analysis of the Cryptic Plasmid pMG101 from *Rhodopseudomonas palustris* and Construction of Stable Cloning Vectors

MASAYUKI INUI, JUNG HYEON ROH, KENNETH ZAHN, AND HIDEAKI YUKAWA*

Research Institute of Innovative Technology for the Earth, Soraku, Kyoto 619-0292, Japan

Received 30 August 1999/Accepted 22 October 1999

A 15-kb cryptic plasmid was obtained from a natural isolate of *Rhodopseudomonas palustris*. The plasmid, designated pMG101, was able to replicate in *R. palustris* and in closely related strains of *Bradyrhizobium japonicum* and phototrophic *Bradyrhizobium* species. However, it was unable to replicate in the purple nonsulfur bacterium *Rhodobacter sphaeroides* and in *Rhizobium* species. The replication region of pMG101 was localized to a 3.0-kb *SalI*-*XhoI* fragment, and this fragment was stably maintained in *R. palustris* for over 100 generations in the absence of selection. The complete nucleotide sequence of this fragment revealed two open reading frames (ORFs), ORF1 and ORF2. The deduced amino acid sequence of ORF1 is similar to sequences of Par proteins, which mediate plasmid stability from certain plasmids, while ORF2 was identified as a putative *rep* gene, coding for an initiator of plasmid replication, based on homology with the Rep proteins of several other plasmids. The function of these sequences was studied by deletion mapping and gene disruptions of ORF1 and ORF2. pMG101-based *Escherichia coli*-*R. palustris* shuttle cloning vectors pMG103 and pMG105 were constructed and were stably maintained in *R. palustris* growing under nonselective conditions. The ability of plasmid pMG101 to replicate in *R. palustris* and its close phylogenetic relatives should enable broad application of these vectors within this group of α -proteobacteria.

Purple nonsulfur bacteria (PNSB) are an assemblage of phenotypically diverse species. Under anaerobic conditions in the light, all species grow photoheterotrophically when supplied with various organic substrates or photoautotrophically with CO₂ as a sole carbon source. Under microaerobic to aerobic conditions in the dark, many representatives can grow chemoheterotrophically, and some grow chemoautotrophically (40).

To develop a new CO₂-fixing bioprocess, we have been performing biochemical and genetic analyses of intermediary metabolism, including CO₂ fixation, underlying the complex modes of growth in the PNSB, using *Rhodopseudomonas palustris* as a model microorganism (4, 23, 24). For this purpose, development of a versatile host-vector system would be helpful.

In *R. palustris* and other PNSB, broad-host-range vectors have been used to provide the tools for gene transfer. The most widely used vectors are derivatives of RK2 such as pRK415 (25) and pLAFR1 (14). Cloning vector pRK415 has been utilized for genetic analyses of several *R. palustris* genes (8, 17), and cosmid vector pLAFR1 has been used to make a library of *R. palustris* DNA (8). However, these plasmids were unstable in *R. palustris* under nonselective conditions (M. Inui, unpublished data), as also observed in *R. sphaeroides* (9) and in *Rhodospirillum rubrum* (42).

Other vectors derived from the broad-host-range plasmid RSF1010 such as pDSK519 (25) can also replicate in PNSB, including *R. palustris* (Inui, unpublished data). But this vector was also unstable under nonselective conditions in *R. palustris* and *R. sphaeroides* (Inui, unpublished data). Thus, these broad-host-range plasmids show segregational instability in *R. palustris* as well as in other PNSB.

R. palustris species in general contain no endogenous plasmids, but it has been reported that several other PNSB may contain one or more such plasmids. *R. sphaeroides* may carry up to six cryptic plasmids ranging in size from 42 to 140 kb (13). *Rhodobacter capsulatus* strains also have either one or two plasmids, with a size range similar to that found in *R. sphaeroides* (21, 57). All strains of *R. rubrum* tested appear to include a single 55-kb plasmid (30). So far there has been no report on use of these large endogenous plasmids to construct vectors for PNSB.

On the other hand, shuttle vectors for the marine photosynthetic bacterium *Rhodobacter marinus* have been constructed by connecting an endogenous plasmid to an *Escherichia coli* cloning vector, because broad-host-range vectors like those mentioned above are ineffective in this organism (5, 32, 33). These plasmids have not been demonstrated to function in other PNSB.

To establish a versatile vector system to facilitate genetic analysis in *R. palustris*, we have screened PNSB for endogenous plasmids of relatively small size. Here we report on the isolation, replication range, and stability of plasmid pMG101 and on sequence analysis of the replication region. Using this information, we have been able to construct stable *R. palustris*-*E. coli* shuttle vectors.

MATERIALS AND METHODS

Bacterial strains and plasmids. Bacterial strains and plasmids used in this study are listed in Table 1.

Culture conditions. *E. coli* strains were grown aerobically at 37°C in Luria-Bertani medium (44). PNSB were cultivated aerobically at 30°C in van Niel's complex medium (53). *Rhizobium* and *Bradyrhizobium* species were grown aerobically at 30°C in 805 medium (1 g of yeast extract, 5 g of mannitol, 0.7 g of K₂HPO₄, 0.1 g of KH₂PO₄, and 1 g of MgSO₄ · 7H₂O per liter, adjusted pH 7.0 to 7.2). When appropriate, media were supplemented with antibiotics. For *E. coli*, ampicillin and kanamycin were each used at a final concentration of 50 µg ml⁻¹; for *R. palustris*, the final concentration of kanamycin was 200 µg ml⁻¹; for other PNSB and *Rhizobium* and *Bradyrhizobium* species, the final concentration of kanamycin was 50 µg ml⁻¹.

* Corresponding author. Mailing address: Research Institute of Innovative Technology for the Earth, 9-2 Kizugawadai, Kizu, Soraku, Kyoto 619-0292 Japan. Phone: 81-774-75-2308. Fax: 81-774-75-2321. E-mail: yukawa@rite.or.jp.

TABLE 1. Bacterial strains and plasmids used in this study

Strain or plasmid	Relevant characteristic(s)	Source or reference ^a
Strains		
<i>E. coli</i> JM109	<i>recA1 endA1 gyrA96 thi hsdR17 supE44 relA1 Δ(lac-proAB)/F'[traD36 proAB⁺ lacI^q lacZΔM15]</i>	Takara
<i>R. palustris</i> No. 7	Alcohol-assimilating PNSB	15
<i>R. palustris</i> S55	Natural isolate of strain bearing plasmid pMG101	This work
<i>R. palustris</i> ATCC 17000	Alcohol-assimilating PNSB	ATCC
<i>R. palustris</i> ATCC 17001	Type strain	ATCC
<i>R. sphaeroides</i> ATCC17023	Type strain	ATCC
<i>R. etli</i> IFO15573	Type strain (=ATCC 51251)	IFO
<i>R. leguminosarum</i> IFO14778	Type strain (=ATCC 10004)	IFO
<i>B. japonicum</i> ATCC 10324	Type strain	ATCC
USDA 4362	Phototrophic <i>Bradyrhizobium</i> species	USDA
USDA 4377	Phototrophic <i>Bradyrhizobium</i> species	USDA
Plasmids		
pUC118	Ap ^r ; α -lac/multicloning site, M13 <i>ori</i>	Takara
pHSG298	Km ^r ; α -lac/multicloning site, M13 <i>ori</i>	Takara
pHSG299	Km ^r ; α -lac/multicloning site (reverse direction of pHSG298), M13 <i>ori</i>	Takara
pHSG298X	Km ^r ; α -lac/multicloning site, M13 <i>ori</i> , addition of <i>XhoI</i> site in pHSG298 by PCR	This work
pMG101	Natural isolate of 15-kb plasmid from <i>R. palustris</i> S55	This work
pMG101Km	Km ^r ; pHSG298 with a 15-kb <i>SalI</i> fragment of pMG101	This work
pMG101A	Km ^r ; pHSG298 with a 8.5-kb <i>EcoRI</i> fragment of pMG101	This work
pMG101B	Km ^r ; pHSG298 with a 6.5-kb <i>EcoRI</i> fragment of pMG101	This work
pMG101C	Km ^r ; pHSG298 with a 4.5-kb <i>SalI-EcoRI</i> fragment of pMG101	This work
pMG101D	Km ^r ; pHSG298 with a 4.0-kb <i>EcoRI-SalI</i> fragment of pMG101	This work
pMG101E	Km ^r ; pHSG298 with a 4.0-kb <i>PstI-EcoRI</i> fragment of pMG101	This work
pMG101F	Km ^r ; pHSG298 with a 1.5-kb <i>XhoI-EcoRI</i> fragment of pMG101	This work
pMG101G	Km ^r ; pHSG298 with a 2.3-kb <i>PstI-XhoI</i> fragment of pMG101	This work
pMG102	Km ^r ; pHSG298 with a 3.0-kb <i>SalI-XhoI</i> fragment of pMG101	This work
pMG102-1 to pMG102-9	Km ^r ; series of pMG102 deletion mutant plasmids	This work
pMG102-10	Km ^r ; frameshift mutant of the <i>parA</i> gene in pMG102	This work
pMG102-11	Km ^r ; frameshift mutant of the <i>repA</i> gene in pMG102	This work
pMG103	Km ^r ; α -lac/multicloning site, pHSG298 with a 3.0-kb <i>SalI-XhoI</i> fragment of pMG101	This work
pMG105	Km ^r ; α -lac/multicloning site, pHSG299 with a 3.0-kb <i>SalI-XhoI</i> fragment of pMG101	This work
pMG103- <i>pckA</i>	Km ^r ; pMG103 with a 2.3-kb <i>BamHI-PstI</i> fragment containing <i>R. palustris</i> No. 7 <i>pckA</i> gene	This work

^a ATCC, American Type Culture Collection; IFO, Institute for Fermentation, Osaka.

DNA manipulations. Plasmid DNA was isolated by the alkaline lysis procedure (44). Restriction endonucleases and T4 ligase were obtained from Takara and used according to the manufacturer's instructions. *E. coli* strains were transformed by the CaCl₂ method (44). PNSB, *Rhizobium*, and *Bradyrhizobium* species were transformed by electroporation (9), with the following modification to obtain optimum transformation efficiency. A Bio-Rad Gene Pulser apparatus was used, with a pulse controller and 0.1-cm-gap electroporation cuvette. A resistance of 200 ohms on the pulse controller and settings of 1.75 kV and 25 μ F were used. These settings generate a field strength of 17.5 kV/cm. Optimal electroporation frequencies are obtained with one pulse per sample. With this method, we routinely achieve approximately 5×10^5 to 10×10^5 drug-resistant cells of *R. palustris* per μ g of plasmid DNA. Restriction fragments were isolated, when required, from agarose gels (1%, wt/vol) with a Prep-a-Gene matrix (Bio-Rad, Richmond, Calif.) according to the manufacturer's instructions.

Isolation of an endogenous plasmid from PNSB. As elsewhere reported (15), PNSB were isolated from enrichment cultures under anaerobic light conditions, using 1-propanol as a carbon source. Isolates were aerobically grown, and plasmid DNA was isolated as described above.

DNA sequencing. To determine the complete nucleotide sequence of the replicator region of plasmid pMG101, overlapping fragments were subcloned in pUC118, and DNA was sequenced on both strands in an automated 373A DNA sequencer (Applied Biosystems/Perkin-Elmer, Foster City, Calif.). DNA sequence data were analyzed with the INHERIT (Applied Biosystems/Perkin-Elmer) and Genetyx (Software Development, Tokyo, Japan) programs.

Isolation and analysis of 16S rDNA. Two oligonucleotide primers, fD1 (5'-A GAGTTTGATCCTGGCTCAG-3') and rD1 (5'-AAGGAGGTGATCCAGCC-3'), were used to PCR amplify nearly full-length 16S ribosomal DNA (rDNA) (55). DNA sequences of PCR-amplified 16S rDNA from strains S55, USDA 4362 (BTai1), and USDA 4377 were determined as described above. The rDNA sequences of strain S55 and the reference strains covering a total of 1,229 nucleotide positions were aligned with the *E. coli* sequence by using the CLUSTAL V program (20). A phylogeny was reconstructed by the neighbor-joining method (43) from K_{nuc} values (27). Variable domains and unidentified base positions were not taken into consideration in the analysis.

Segregational plasmid stability. *R. palustris* cells containing pMG101 derivatives were first grown in selective liquid medium (van Niel's medium containing kanamycin). Cells in the late exponential growth phase were diluted 1,000-fold in van Niel's medium without the antibiotic and grown at 35°C. At 2-day intervals,

cultures were diluted 1,000-fold. At each dilution step, cell density was measured to estimate the number of generations, and the cells were plated on nonselective plates. From these plates, 200 colonies each were transferred to selective and nonselective plates to determine the frequency of plasmid loss based on the percentage of kanamycin-sensitive colonies. Similarly, plasmid stability in *Bradyrhizobium japonicum* and phototrophic *Bradyrhizobium* species was analyzed after growth in 805 medium with or without kanamycin.

RNA isolation and primer extension. Template RNA was extracted from an *R. palustris* No. 7 culture in late exponential phase as previously described (23). Primers were synthesized with 5'-end positions 300, 320, and 354 for the *parA* gene and 1390, 1410, and 1526 for the *repA* gene, complementary to the mRNA sequence. Primer extension was performed as described elsewhere (23).

Construction of *E. coli-Rhodospseudomonas* shuttle cloning vectors. Plasmid pHSG298X was made by PCR amplification of plasmid pHSG298, using primers containing an *XhoI* site. The primers P1 (5'-CGCTCGAGGGAGCCACGGTT GA-3') and P2 (5'-CGCTCGAGCAACACCTTCTTCACGA-3') were used, and the PCR product was ligated to itself and transformed into *E. coli* JM109. The resulting plasmid has a unique *XhoI* site between the pHSG298 origin and kanamycin resistance gene. Plasmid pMG103 was constructed by cloning a 3.0-kb *SalI-XhoI* fragment from plasmid pMG101 into the *XhoI* site of plasmid pHSG298X. For the construction of plasmid pMG105, a 3.8-kb *ApaLI-Clal* fragment containing the pMG101 origin from plasmid pMG103 was ligated with the 1.8-kb *ApaLI-Clal* fragment containing the β -galactosidase (*lacZ*) gene, which has the opposite orientation of the polylinker site versus that of plasmid pMG103.

Nucleotide sequence accession numbers. The DDBJ/EMBL/GenBank accession numbers for the sequences reported in this paper are AB031076 (the *SalI-XhoI* 3.0-kb fragment of pMG101), D84187 (rDNA for strain S55), D86354 (rDNA for USDA 4362), D86355 (rDNA for USDA 4377), AB031077 (plasmid pMG103), and AB031078 (plasmid pMG105).

RESULTS

Isolation of plasmid pMG101 and characteristics of the host strain. To develop a vector system for PNSB, we screened various strains for small (<20-kb) plasmids. Among 400 strains

FIG. 1. Unrooted distance matrix tree showing phylogenetic relationships based on 16S rDNA sequences of selected strains of proteobacteria. The topology of the phylogenetic tree was evaluated by bootstrap analysis with 1,000 replicates. Numbers indicate percentage bootstrap values of 1,000 replicates. USDA 4362 and USDA 4377 are phototrophic *Bradyrhizobium* species. Abbreviations where relevant and, in parentheses, accession numbers: strain S55 (D84187), *R. palustris*, *Rhodospseudomonas palustris* (D25312); *B. japonicum*, *Bradyrhizobium japonicum* (D11345); *R. roseus*, *Rhodoplanes roseus* (D25313); *R. acidophila*, *Rhodospseudomonas acidophila* (M34128); *R. viridis*, *Rhodospseudomonas viridis* (D25314); *R. rubrum*, *Rhodospirillum rubrum* (D30778); *R. sphaeroides*, *Rhodobacter sphaeroides* (D16425); *R. capsulatus*, *Rhodobacter capsulatus* (D16428); *R. salicigenes*, *Rhodospirillum salicigenes* (M59070); *R. tennisi*, *Rhodocyclus tennisi* (D16208); *R. purpureus*, *Rhodocyclus purpureus* (M34132); *R. gelatinosus*, *Rhodocyclus gelatinosus* (D16214); *E. coli*, *Escherichia coli* (V00348); *R. etli*, *Rhizobium etli* (U28916); *R. leguminosarum*, *Rhizobium leguminosarum* (U73208).

isolated from sewage, rice fields, and other places, one strain (S55) contained an endogenous plasmid about 15 kb in size. This plasmid was designated pMG101. The loss of pMG101 from strain S55 did not induce any detectable phenotypic difference (data not shown).

Strain S55 was gram negative and nonsporulating, and it formed rods with rounded ends. Cell multiplication occurred by budding. Cultures were red to brown-red under anaerobic light conditions and pale pink to white under aerobic dark conditions. The *in vivo* spectrum of cell suspensions grown anaerobically in the light exhibited absorption maxima at 375, 590, 805, and 860 nm, which is characteristic for bacteriochlorophyll *a*, and at 465, 498, and 530 nm, indicating the presence of carotenoids of the normal spirilloxanthin series (40). These results indicate that strain S55 is a species of the genus *Rhodospseudomonas*.

To confirm the phylogenetic position of strain S55, we determined the rDNA sequence of 1,481 nucleotides of strain S55. The phylogenetic position of strain S55, as well as that of two species of phototrophic *Bradyrhizobium*, was analyzed with data for sequences longer than those used for previous analyses of these strains (60, 62) (Fig. 1). Strain S55 formed a tight cluster with *R. palustris*, *B. japonicum*, and phototrophic *Bradyrhizobium* species and was less similar to other PNSB and *Rhizobium* species. A bootstrap analysis confirmed the monophyly of the strain S55 cluster in 100% of trees generated (11). These results confirmed that strain S55 is a strain of *R. palustris* and is phylogenetically close to *B. japonicum* and phototrophic *Bradyrhizobium* species. Our data are in agreement with previous reports that *R. palustris*, phototrophic *Bradyrhizobium* species which produce bacteriochlorophyll (12), and *Bradyrhizobium japonicum* show high sequence similarity and form a branch on 16S rRNA phylogenetic trees (60, 62).

Replication range and stability of plasmid pMG101. Plasmid pMG101 is the first plasmid to be identified in *R. palustris* and is smaller than known endogenous plasmids in other PNSB. Restriction mapping of pMG101 was carried out (Fig. 2), and the shuttle vector pMG101Km was constructed. pMG101Km can replicate in *R. palustris* (strain No. 7 [15], ATCC 17001, and ATCC 17000), *B. japonicum* ATCC 10324, and phototrophic *Bradyrhizobium* species (USDA 4362 and USDA 4377). It cannot replicate in *R. sphaeroides* ATCC 17023 or in nonpho-

trophic *Rhizobium* species, *R. etli* IFO 15573 and *R. leguminosarum* IFO 14778. Species-specific replication of pMG101Km clearly reflects the phylogenetic relationships of the *Rhodospseudomonas* group (Fig. 1).

To determine the stability of plasmid pMG101, *R. palustris*, *B. japonicum*, and phototrophic *Bradyrhizobium* species cultures containing pMG101Km were grown for 30 generations under nonselective conditions. The plasmid was stably maintained in *R. palustris* (100% maintenance), less stable in phototrophic *Bradyrhizobium* species (50% maintenance), and least stable in *B. japonicum* (0% maintenance). However, pMG101Km was stably maintained in all of these hosts when kanamycin was added. Plasmid replication is thought to be determined by the interaction of host- and plasmid-encoded factors with the replication region of the plasmid (54). The differences in stability might reflect differences in host factors in each strain.

Localization of the replication and stability region. To identify the replication and stability region of pMG101, we constructed a series of pMG101 derivatives by subcloning restriction fragments into plasmid pHSG298 (Fig. 2), which contains a kanamycin resistance gene, and these plasmids were transformed into *R. palustris*. Plasmid pHSG298, which contains a ColE1 origin, was not able to replicate in *R. palustris* (data not shown). The plasmid pMG101 derivatives pMG101A, pMG101C, pMG101E, pMG101G, and pMG102 could replicate in *R. palustris*. However, when pMG101B, pMG101D, and pMG101F were transformed in *R. palustris*, no kanamycin-resistant colonies were obtained.

To determine the location of the stability region of pMG101, *R. palustris* cells harboring the pMG101 derivatives were grown for 50 generations without selection, and retention of kanamycin resistance was evaluated by transferring at least 200 isolates to selective and nonselective media (Fig. 2). Plasmids pMG101A, pMG101C, and pMG102 were stably maintained in *R. palustris*; in contrast, plasmids pMG101E and pMG101G were gradually lost, and the percentage of plasmid-containing cells decreased to 10 to 11% after 50 generations. Moreover, no plasmid loss was observed in *R. palustris* containing pMG102 for over 100 generations without selective pressure (data not shown). The 3.0-kb *SalI*-*XhoI* fragment therefore appears to include both replication and stability functions of pMG101, and digestion at

FIG. 2. Localization of the replication and stability regions. Construction of pMG101 derivatives is described in Table 1. The heavy horizontal lines indicate regions of plasmid pMG101 that were inserted into pHSG298. Replication was considered positive (+) when transformants were present and plasmid DNA was detected after alkaline lysis and gel electrophoresis and negative (-) when no transformants were obtained in three experiments. Plasmid stability in *R. palustris* was analyzed as described in Materials and Methods.

a *PstI* restriction site within this fragment destabilizes the plasmid.

DNA sequence of the 3.0-kb *SalI-XhoI* fragment of pMG101. To clarify the plasmid partitioning and replication functions of pMG101, the complete nucleotide sequence of the 3.0-kb *SalI-XhoI* fragment was determined. The G+C content of this fragment was 57.9%, which is below the range (from 64.8 to 66.3%) previously reported for the genome of *R. palustris* (22). Computer analysis revealed the presence of two potential open reading frames (ORFs), ORF1 and ORF2.

ORF1 (positions 236 to 889) consists of 654 nucleotides,

corresponding to 217 amino acids and a predicted molecular weight of 22,450. The deduced amino acid sequence shows weak but significant sequence similarities to sequences of the *parA* gene products of plasmid pTAR from *Agrobacterium tumefaciens* (28.5%) (16 [accession no. X05121]), pFAJ2600 from *Rhodococcus erythropolis* (27.7%) (7 [accession no. AF015088]), and a putative Par-like ORF from the genome of *Helicobacter pylori* (25.8%) (50 [accession no. AE000608]). ORF1 was therefore designated the *parA* gene. The result of the best alignment with ParA proteins from pMG101 and pTAR is shown in Fig. 3A. The motif close to the N-terminal region of A-type Sop/

A

pMG101-ParA	MKTIALVTQKGGAGKTTTAAATAVAATEAGESVIALDLDPQGSAAWGDQRAQAQREKPAVDALPPDKVPQLPQILKALA	80
pTAR-ParA	MPVVVVASKGGAGKSTTAVVLGTELAHKGVPVMTLDCDPNRSLTIWANA-GEV-FEN--ITALSVDVTESSIVKTIKQHD	76
	* * * * *	
pMG101-ParA	GKGYTALAILDCPGVASTATSLAMQAADLSLVPARPTRLD--IQATR-PTI-QALLKLQRPFF--LFVLNQCPFITRSGRAS	154
pTAR-ParA	VDG-AVVIVDLEGVASRMVSRRAISQADLVLIIPMRPKALDATTGAQSLQLIAEEEEAIDRKIAHAVVFTMVSPAIRSHEYT	155
	* * * * *	
pMG101-ParA	EAAVGLSALG-SLAEPMLVSRADFQDAIA-AG--QGVTEYAPS-GKAAEEARGLWRSIQHKLKTKAVA	217
pTAR-ParA	GIKASLIENGVEIIEPPLVERTAYSALFQFGGNLSHMSKSKQGNMAAIIENAEAFAMAIFKKL-TEALR	222
	* * * * *	

B

pMG101-RepA	MPRERRRRSII---REVIENSPPSPADVRLHLSVLAVCGMPYDRQPLEVREFERKQGNMALDLSAGFLRDGEGKKILQPL	77
pRM21-RepA	MPRTRALDALLKAAREIAEQDAREADAIAVMARLLTQATIP-HREFRGQTTYRREAGDLVLEI-QG--R--E--K--Y--GL	70
	* * * * *	
pMG101-RepA	PFGPKARLVMMHLCSEAIRQKTPTTIEIAESFTAFVREMGSFSDSGGRKGPLTAFKEQINALAACTMRVSVWDGVRRAKT-KT	156
pRM21-RepA	PYGSYPRLLLAWMCTEAVKTRSRKLYLGASLSEFMDKGLVPSGGRWGTIPRLRDQAKRLFFSRWFVDVSRMEAKGWRA	150
	* * * * *	
pMG101-RepA	-VN-PFD-EV---E-VWLPPHLDQRLWPS-T-VTFNPFVFFESLQQRALPLNANAVRAFGASARKLDLYFWLGYRLHNIR	227
pRM21-RepA	WWDLGFQIALSRSTLLWDVKDPDQLALFEEHSYVELTQEFFEAVLLRPVPLDARALRAKQSPALDLAYWLSYRASYAQ	230
	* * * * *	
pMG101-RepA	HPVTMSWKALSEQFGQGFTRPRAFAQLIEELTHIKEVLPKPLKLTQYGISIEPADPAVLALPAIRV-----RKK	298
pRM21-RepA	KDLAVTWQQLHAQFGADYENVDEFARKVKQAIKIHLLWPDRLRVYPK--GRVLVKATTQVHVFPRAIGSNDKGSARKK	307
	* * * * *	

FIG. 3. Sequence alignments of the putative ParA and RepA proteins. (A) Comparison of amino acid sequences of ParA proteins from plasmids pMG101 and pTAR. A modified NTP-binding motif is indicated by a plus sign above the sequence. (B) Comparison of amino acid sequences of RepA proteins from plasmids pMG101 and pRM21. Identical residues are marked with asterisks below the sequence alignments.

A		B		C	
1	AGCGCTGTAGCGCT	1	CCGCAGACTTTGTTGCG	1	TAGAACA
2	TAATAATGCTATAGCATTTTA	2	CCGCAGACTTGGTGGCA	2	CAGAACA
3	GCTATGCC	3	CCGCAGACTTGGTGGCG	3	CAGAACA
4	GCTATGCC	4	CCGCAGACTTGGTGGCG	4	CAGAACA
5	GCTATGCC	5	CCGCAGACTTGGTAGCG	5	CAGAACA
6	GCTATGGC	6	CCGCAGACTTGGTGGCG	6	GAACA
7	GCTTTGAC	7	CCGCAG		
8	GCTTTAGT			Consensus	CAGAACA
9	GCTATAGT	Consensus	CCGCAGACTTGGTGGCG		T
	Consensus GCTATGCC		T T A		
	G AGT		A		
	T A				

FIG. 5. Sequence alignments of the nine 8-bp direct repeats (DR I) (A), the six 17-bp imperfect direct repeats (DR II) and a partial homologous sequence (B), and the five 7-bp imperfect direct repeats (DR III) and a partial homologous sequence (C). The consensus sequences are shown below the sequence alignments. Two pairs of IRs, which are overlapped with DRs, are noted by rightward and leftward arrows above DR sequences in panel A.

control and also to genetically engineer the plasmid for use as a vector. To localize the *parA* and *repA* promoters, the transcriptional start sites of the *parA* and *repA* genes were determined by using primer extension with three 20-mer oligonucleotides having different 5'-end positions (see Materials and Methods).

Primer extension analysis of the *parA* gene revealed one major extension product with all three primers corresponding to the A of the AUG start codon of the *parA* gene. The result with the primer at 320 is shown in Fig. 6A. Putative -10 and -35 promoter sequences, similar to the *E. coli* σ^{70} promoter consensus (19) separated by a typical 17-nucleotide interval, were identified upstream of the transcriptional start point, and the putative promoter sequence is overlapped with DR I sequences (Fig. 4A).

Reverse transcription of the *repA* mRNA yielded two major bands with all three primers, corresponding to 5' ends at 172 and 173 bases upstream of the translation initiation codon. The result with the primer at 1390 is shown in Fig. 6B. Upstream of the transcriptional start site, putative -10 and -35 promoter sequences, separated by a typical 17-nucleotide interval, were observed, and a partial copy of DR II at positions 1275 to 1280 was found between the -10 and the -35 promoter sequences (Fig. 4B and 5B). In addition, a partial copy of DR III at positions 1429 to 1435 also occurs between the promoter sequence and the translation initiation codon (Fig. 4B and 5C).

Deletion mapping and gene disruption analysis. To define more precisely the replication and stability functions of pMG101, various fragments were amplified by PCR, cloned into the *E. coli* vector pHSG298, and then tested for the ability to promote plasmid replication and stability in *R. palustris* (Fig. 7). Plasmid pMG102 was used as a positive control in these experiments.

Plasmid pMG102-1, containing the 2,954-bp fragment (positions 100 to 3053) which lacked a possible upstream ORF within the 3.0-kb *SalI-XhoI* fragment, can replicate and be maintained in *R. palustris* as well as a control plasmid, pMG102. Plasmid pMG102-2, containing the 2,844-bp fragment (positions 210 to 3053) which lacked eight of nine DR I sequences and a putative -35 promoter sequence of the *parA* gene, was able to replicate but was not maintained (25% plasmid maintenance) in *R. palustris* growing under nonselective conditions (Fig. 4 and 7). It is likely that DR I sequences and a putative *parA* promoter region are required for pMG102 stability. In addition, the *parA* product is likely to be essential for pMG102 stability, because a frameshift mutation created by insertion of 4 bp into the *MunI* site within the *parA* sequence in plasmid pMG102-10 resulted in instability in *R. palustris* under nonselective growth conditions (Fig. 7).

Plasmid pMG102-3, containing the 2,054-bp fragment (positions 1000 to 3053), can replicate in *R. palustris*, but plasmid pMG102-4, which contains the 1,954-bp fragment (positions 1100 to 3053), was not able to replicate, suggesting that the 100-bp sequence at positions 1000 to 1100 is essential for pMG102 replication. However, DNA sequence analysis did not detect any known sequence motifs, specific DNA structures, or homologous nucleotide or amino acid coding sequences in this region.

Analysis of the deletion mutant plasmids pMG102-5, pMG102-6, and pMG102-7 indicated that a region containing six DR II, four DR III, and an AT-rich sequence was sufficient for pMG102 replication (Fig. 7). Plasmid pMG102-9, which included all DR II, DR III, AT-rich, and GC-rich sequences, replicated and was maintained stably in *R. palustris* as well as a control plasmid, pMG102. Plasmid pMG102-8, containing the 2,700-bp fragment (positions 1 to 2700), which lacked a part of the GC-rich sequence and downstream sequences from the

FIG. 6. Mapping of the transcriptional start sites of the *parA* (A) and *repA* (B) genes by primer extension analysis. High-resolution electrophoresis of 5'-labeled cDNA products are shown aligned with a sequencing ladder generated from DNA from the same region and with the same primer. Dideoxy sequencing marker lanes are indicated with CTAG, and cDNA synthesis from the RNA samples is shown in the next lane (PE). Migration positions of the primer extension products are indicated with arrows. Portions of the DNA sequence are indicated at the left; asterisks indicate the putative transcriptional start site.

FIG. 7. Deletion mapping and gene disruption analysis of the 3.0-kb *SalI*-*XhoI* fragment containing the pMG101 replication and stability region. The restriction cleavage map of this fragment is shown at the top. Arrows depict the *parA* and *repA* genes and the C-terminal region of a possible ORF. DR I, DR II, and DR III are indicated by open leftward, closed leftward, and open rightward arrowheads, respectively. The partial DR II and DR III are represented by DR II* (shaded leftward arrowhead) and DR III* (dotted rightward arrowhead). Putative -10 and -35 promoter sequences of the *parA* and *repA* genes are shown by small closed boxes. AT- and GC-rich sequences are indicated by shaded and bold underlines, respectively. The scale is graduated in base pairs from the left (*SalI* site) to the right (*XhoI* site), corresponding to the nucleotide sequence of the 3.0-kb *SalI*-*XhoI* fragment. pMG102 deletion and frameshift mutant derivatives are listed at the left; each fragment is represented by a horizontal bold line. Deletion positions in base pairs are shown at the end of each horizontal bold line. Frameshift positions are indicated by downward arrowheads. Construction of these plasmids is described in Table 1. Assay of replication in *R. palustris* was as described in the legend to Fig. 2. ± indicates that transformation efficiency is low and that it takes a longer incubation time to obtain transformants than is the case for the control plasmid pMG102. Plasmid stability in *R. palustris* was evaluated as described in Materials and Methods.

FIG. 8. Physical and genetic maps of pMG103 and pMG105. Notation: *lacZ*, β -galactosidase α -peptide gene from pHSG298 or pHSG299; *Km^r*, kanamycin resistance gene; pMG101 ori, the 3.0-kb *SalI*-*XhoI* fragment for pMG101 origin of replication, which includes *parA* and *repA* genes; pHSG298 ori and pHSG299 ori, origins of replication from pHSG298 and pHSG299, respectively. Arrows indicate direction of transcription. Unique restriction sites are shown.

3.0-kb *SalI*-*XhoI* fragment, can replicate in *R. palustris*, but transformation efficiency is 1,000 times lower and an incubation time of 5 days is required to obtain transformants, compared with 2 days for the pMG102 control. Moreover, plasmid pMG102-8 showed segregational instability in *R. palustris* under nonselective growth conditions. These data suggest that the GC-rich sequence plays an important role in the replication or stability of plasmid pMG102. Furthermore, a frameshift mutation created by insertion of 2 bp into the *NspV* site in the *repA* sequence in plasmid pMG102-11 blocked replication in *R. palustris*, suggesting that the *repA* gene product is essential for pMG102 replication (Fig. 7).

Construction of *E. coli*-*R. palustris* shuttle cloning vectors. To further improve the versatility of the *E. coli*-*R. palustris* shuttle cloning vector, plasmids pMG103 and pMG105, which are 5,680 bp in size and contain a kanamycin resistance marker, were constructed as described in Materials and Methods. The shuttle vectors have a polylinker containing unique *EcoRI*, *SacI*, *KpnI*, *BamHI*, *SalI*, *XbaI*, *Sse8387I*, and *SphI* sites, and insertional inactivation of the *lacZ* gene can be used to identify cloned inserts by blue/white colony screening in the *E. coli* host strain JM109 when plated on isopropyl- β -D-thiogalactopyranoside-5-bromo-4-chloro-3-indolyl- β -D-galactopyranoside (IPTG-X-Gal) agar plates (Fig. 8). Both of these plasmids were stably maintained in *R. palustris* for over 100 generations without selective pressure.

To estimate the copy number of pMG103 per chromosome of *R. palustris*, we constructed plasmid pMG103-*pckA*, containing the *R. palustris* No. 7 *pckA* gene. This gene is present in only a single copy on the chromosome (24). Southern blotting of *R. palustris* total DNA harboring plasmid pMG103-*pckA*, digested with *SmaI*, *BamHI*, and *PstI*, was performed. Hybridization with the *pckA* gene probe detected a 2.5-kb *SmaI*-*PstI* fragment representing the *pckA* gene encoded on the *R. palustris* No. 7 chromosome and a 2.2-kb *BamHI*-*PstI* fragment representing the *pckA* gene from plasmid pMG103-*pckA*. The ratio of two bands was evaluated from densitometry of the relative intensities of the hybridization signals (data not shown), and pMG103 copy number was estimated to be between three to five per *R. palustris* chromosome.

In previous reports, *R. palustris* cells containing the vector pMG102 bearing a cloned *R. palustris pckA* or *ppc* gene exhibited enzyme activities four times greater than that of the wild-type strain (23, 24). These data agree with the proposed copy number of plasmid pMG101 based on hybridization.

DISCUSSION

We have isolated a 15-kb cryptic plasmid, pMG101, from a natural isolate of PNSB, identified the partition and replication region in a 3.0-kb *SalI*-*XhoI* subfragment, determined its DNA sequence and transcriptional start sites, and used this information to construct stable *R. palustris*-*E. coli* shuttle vectors. These shuttle vectors have already proved useful in studies of gene function in *R. palustris* (23, 24).

Sequence studies presented here suggest that the closest relationship is between the Par protein of pMG101 and that of pTAR, a plasmid isolated from *Agrobacterium* which, like *R. palustris*, is an α -proteobacterium. Portions of the adjacent sequence exhibit a repetitive sequence organization like that of the *cis*-acting *par* locus of pTAR.

The pTAR plasmid-partitioning system is a member of the *sop/par* family, which includes *sop* of plasmid F and *par* of plasmid P1 (56). Upstream of the pTAR *parA* gene, the *cis*-acting partition site contains a set of twelve 7-bp repeat sequences and the promoter region at which ParA binds to regulate its own transcription (16). Upstream of the *parA* gene of plasmid pMG101, there is an array of nine 8-bp repeats (DR I) that overlap the putative *parA* promoter (Fig. 4). The data (Fig. 4) suggest that the region upstream of the pMG101 *parA* gene, which includes the nine repeats, is the *cis*-acting recognition site at which ParA interacts to bring about partitioning and regulation of its own transcription, as proposed for the partition system of plasmid pTAR.

Structural and expression analyses of the *parA* gene revealed that the transcriptional start site corresponds precisely to the position of the deduced translation initiation codon for this gene (Fig. 4 and 6). Upstream of the site, we find an *E. coli* σ^{70} -like promoter but no distinct Shine-Dalgarno (SD) sequence. In addition, a 4-bp insertion into the *MunI* site, which is located just downstream of the deduced initiation codon, destabilizes the plasmid in *R. palustris* under nonselective conditions. This finding suggests that the insertion created the expected frameshift mutation in the *parA* gene, which is predicted to be initiated only three codons upstream of this insertion (Fig. 7). Moreover, comparison of the deduced amino acid sequences with sequences of the pMG101 *parA* and pTAR *parA* gene products showed similar lengths and significant similarity along the length of the protein as well as a putative NTP-binding motif close to the N-terminal region (Fig. 3).

The pMG101 *parA* gene is unlikely to be translated by the usual system of SD interaction of the mRNA with 16S rRNA during initiation, because in contrast to most *R. palustris* genes

(23, 24) and those of other PNSB, we find that the *parA* mRNA does not possess a SD sequence. In several prokaryotes, genes in which transcription and translation are initiated from the same nucleotides have been found (1, 2, 6, 10, 28, 31, 41, 49, 58). It was proposed that in *E. coli*, these genes lack SD sequences but possess a downstream box (DB) located just downstream of initiation codon which is complementary to a region (anti-downstream box [ADB]) in 16S rRNA, and the DB can function as a translation initiation signal (46).

Although no homology was observed between the *E. coli* ADB in 16S rRNA and the corresponding region of *R. palustris* 16S rRNA, a region just downstream of the *parA* initiation codon shows significant sequence complementarity to another region at the 5' end of *R. palustris* 16S rRNA (data not shown). However, the significance of the DB-ADB complementarity in translation of *E. coli* genes which lack SD sequences has recently been questioned (37). Further studies are necessary to determine whether the leaderless *parA* transcript of plasmid pMG101 is translated by this or another mechanism.

According to molecular and transcriptional analyses of the replication region of plasmid pMG101, the *repA* gene was transcribed from 172 and 173 bp upstream of the translation initiation codon, and a putative promoter sequence is found 5' of the transcriptional start site. 3' of the *repA* gene, two series of direct repeats (DR II and DR III), followed by AT-rich and GC-rich sequences, are found.

Clusters of direct repeats, termed iterons (34), exist in the replication origin regions of several plasmids, including mini-F (36), λ dv (18, 34), R6K (47), and RK2 (48). Moreover, plasmid pSa, whose RepA protein is similar to that of pMG101 as described above, also contains six 17-bp direct repeats downstream of the *repA* gene (39). These repeat sequences or iterons are known to be the binding sites for Rep-like proteins (35, 51, 63), and the 17-bp direct repeats (DR II) of pMG101 are likely to represent the binding sites for RepA protein in this plasmid.

AT-rich regions have low thermal stability and are known to be the sites of strand separation at a variety of replication origins (3, 45, 52). Specialized sequences involved in strand separation at origins have been proposed to exist here (3, 45). In the case of the *E. coli* origin, four AT-rich 13-mers are proposed to be the sites of strand separation, and similar sets of repeats are found in plasmid origins, where they appear to serve the same purpose. The DR III 7-mer repeats are not homologous to these sequences but might function similarly.

GC-rich sequences have high thermal stability and could be important in limiting the extent of strand separation at the adjacent AT-rich regions. This might concentrate the strand separation locally and promote loading of the DNA helicase at this site. A similar sequence organization, consisting of a cluster of direct repeats followed by AT-rich and GC-rich regions, has been observed in the origin regions of RK2 (48) and pSa (39). In each case, deletion of these GC-rich regions strongly affects replication.

Additionally, there are partial copies of DR II and of DR III in the pMG101 *repA* promoter sequence and between the promoter sequence and the *repA* initiation codon, respectively (Fig. 4 and 5). It is possible that these partial direct repeats are involved in autoregulation of *repA* gene expression, as proposed for plasmid R6K (26).

Deletion analysis of the pMG101 replication region indicated that a 100-bp sequence which is located upstream of the *repA* promoter (at positions 1000 to 1100) was also essential for replication (Fig. 8). The absence of any specific structures or sequences in the 100-bp region gives no clues to its function. However, the existence of auxiliary *cis*-acting sequences, or

enhancers, affecting replication has been demonstrated in several systems (38, 61, 64). Replication enhancers bind a variety of proteins, and the pMG101 sequence may function in the same way (38, 61, 64). Further studies are needed to understand the functions of the genes and the sites and mechanisms controlling pMG101 replication.

The stable *E. coli*-*R. palustris* shuttle cloning vectors pMG103 and pMG105 have provided us with a versatile cloning system for PNSB. We have found that pMG101 derivatives can replicate not only in *R. palustris* but also in the *Bradyrhizobium* species, and this host range observation clearly reflects the phylogenetic relationships of these species in the α -proteobacteria group. This is the first report that an endogenous plasmid from a photosynthetic bacterium replicates in other genera, especially the phenotypically distinct *Bradyrhizobium* species. These observations support the proposal that photosynthetic bacteria may have been ancestors of the *Rhizobiaceae* (59). Though the natural transfer of this plasmid to *B. japonicum* and phototrophic *Bradyrhizobium* species has not been demonstrated, the possibility that it is involved in horizontal gene transfer between these genera and the possible transmissibility of plasmid pMG101 in this process should be considered.

ACKNOWLEDGMENTS

We are especially grateful to Peter van Berkum (Agriculture Research Service, USDA) for the gift of strains USDA 4362 and USDA 4377. We thank Yasuyoshi Nakagawa (Institute for Fermentation, Osaka, Japan) for his help in phylogenetic analysis and László Puskás for the construction of plasmid pHSG298X and for valuable discussions.

This work was supported by a grant from the New Energy and Industrial Technology Development Organization.

REFERENCES

1. Betlach, M., J. Friedman, H. W. Boyer, and F. Pfeifer. 1984. Characterization of a halobacterial gene affecting bacterio-opsin gene expression. *Nucleic Acids Res.* **12**:7949–7959.
2. Blanck, A., and D. Oesterhelt. 1987. The halo-opsin gene. II. Sequence, primary structure of halorhodopsin and comparison with bacteriorhodopsin. *EMBO J.* **6**:265–273.
3. Bramhill, D., and A. Kornberg. 1988. Duplex opening by *dnaA* protein at novel sequences in initiation of replication at the origin of the *E. coli* chromosome. *Cell* **52**:743–755.
4. Brenner, V., M. Inui, N. Nunoura, K. Momma, and H. Yukawa. 1998. Studies on CO₂ fixation in PNSB: utilization of waste as the additional source of carbon for CO₂ fixation by PNSB, p. 593–596. In T. Inui, M. Anpo, K. Izui, S. Yanagida, and T. Yamaguchi (ed.), *Advances in chemical conversions for mitigating carbon dioxide. Studies in surface science and catalysis*, vol. 114. Elsevier Science B.V., Amsterdam, The Netherlands.
5. Burgess, J. G., H. Sudo, K. Sode, and T. Matsunaga. 1993. Efficient gene transfer of high copy number mobilizable plasmids containing a marine *Rhodobacter* origin of replication. *Curr. Microbiol.* **26**:105–108.
6. Christie, G. E., and R. Calendar. 1985. Bacteriophage P2 late promoters. II. Comparison of the four late promoter sequences. *J. Mol. Biol.* **181**:373–382.
7. De Mot, R., I. Nagy, A. De Schrijver, P. Pattanapitpaisal, G. Schoofs, and J. Vanderleyden. 1997. Structural analysis of the 6 kb cryptic plasmid pFAJ2600 from *Rhodococcus erythropolis* NI86/21 and construction of *Escherichia coli*-*Rhodococcus* shuttle vectors. *Microbiology* **143**:3137–3147.
8. Dispensa, M., C. T. Thomas, M. K. Kim, J. A. Perrotta, J. Gibson, and C. S. Harwood. 1992. Anaerobic growth of *Rhodospseudomonas palustris* on 4-hydroxybenzoate is dependent on AadR, a member of the cyclic AMP receptor protein family of transcriptional regulators. *J. Bacteriol.* **174**:5803–5813.
9. Donohue, T. J., and S. Kaplan. 1991. Genetic techniques in Rhodospirillaceae. *Methods Enzymol.* **174**:459–485.
10. Dunn, R., J. McCoy, M. Simsek, A. Majumdar, S. H. Chang, U. L. Raj-Bhandry, and H. G. Khorana. 1981. The bacteriorhodopsin gene. *Proc. Natl. Acad. Sci. USA* **78**:6744–6748.
11. Felsenstein, J. 1985. Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* **39**:783–791.
12. Fleischman, D. E., W. R. Evans, and I. M. Miller. 1995. Bacteriochlorophyll-containing *Rhizobium* species. *Adv. Photosyn.* **2**:123–136.
13. Fornari, C. S., M. Watkins, and S. Kaplan. 1984. Plasmid distribution and analyses in *Rhodospseudomonas sphaeroides*. *Plasmid* **11**:39–47.
14. Friedman, A. M., S. R. Long, S. E. Brown, W. J. Buikema, and F. M. Ausubel.

1982. Construction of a broad host range cosmid cloning vector and its use in the genetic analysis of *Rhizobium* mutants. *Gene* **18**:289–296.
15. Fujii, T., A. Nakazawa, N. Sumi, H. Tani, A. Ando, and M. Yabuki. 1983. Utilization of alcohols by *Rhodopseudomonas* sp. No. 7 isolated from n-propanol-enrichment cultures. *Agric. Biol. Chem.* **47**:2747–2753.
16. Gallie, D. R., and C. I. Kado. 1987. *Agrobacterium tumefaciens* pTAR *parA* promoter region involved in autoregulation, incompatibility and plasmid partitioning. *J. Mol. Biol.* **193**:465–478.
17. Gibson, J., M. Dispensa, G. C. Fogg, D. T. Evans, and C. S. Harwood. 1994. 4-Hydroxybenzoate-coenzyme A ligase from *Rhodopseudomonas palustris*: purification, gene sequence, and role in anaerobic degradation. *J. Bacteriol.* **176**:634–641.
18. Grosschedl, R., and G. Hobom. 1979. DNA sequences and structural homologies of the replication origins of lambdoid bacteriophages. *Nature* **277**:621–627.
19. Hawley, D. K., and W. R. McClure. 1983. Compilation and analysis of *Escherichia coli* promoter DNA sequences. *Nucleic Acids Res.* **11**:2237–2255.
20. Higgins, D. G., A. J. Bleasby, and R. Fuchs. 1992. CLUSTAL V: improved software for multiple sequence alignment. *Comput. Appl. Biosci.* **8**:189–191.
21. Hu, N. T., and B. L. Marrs. 1979. Characterization of the plasmid DNAs of *Rhodopseudomonas capsulata*. *Arch. Microbiol.* **121**:61–69.
22. Imhoff, J. F. 1995. Taxonomy and physiology of phototrophic purple bacteria and green sulfur bacteria. *Adv. Photosyn.* **2**:1–15.
23. Inui, M., V. Dumay, K. Zahn, H. Yamagata, and H. Yukawa. 1997. Structural and functional analysis of the phosphoenolpyruvate carboxylase gene from the purple nonsulfur bacterium *Rhodopseudomonas palustris* No. 7. *J. Bacteriol.* **179**:4942–4945.
24. Inui, M., K. Nakata, J. H. Roh, K. Zahn, and H. Yukawa. 1999. Molecular and functional characterization of the *Rhodopseudomonas palustris* No. 7 phosphoenolpyruvate carboxylase gene. *J. Bacteriol.* **181**:2689–2696.
25. Keen, N. T., S. Tamaki, D. Kobayashi, and D. Trollinger. 1988. Improved broad-host-range plasmids for DNA cloning in Gram-negative bacteria. *Gene* **70**:191–197.
26. Kelley, W., and D. Bastia. 1985. Replication initiator protein of plasmid R6K autoregulates its own synthesis at the transcriptional step. *Proc. Natl. Acad. Sci. USA* **82**:2574–2578.
27. Kimura, M. 1980. A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *J. Mol. Evol.* **16**:111–120.
28. Klock, G., and W. Hillen. 1986. Expression, purification and operator binding of the transposon TnT721-encoded Tet repressor. *J. Mol. Biol.* **189**:633–641.
29. 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.
30. Kuhl, S. A., D. W. Nix, and D. C. Yoch. 1983. Characterization of a *Rhodospirillum rubrum* plasmid: loss of photosynthetic growth in plasmidless strains. *J. Bacteriol.* **156**:737–742.
31. Lopez, P., S. Martinez, A. Diaz, M. Espinosa, and S. A. Lacks. 1989. Characterization of the *polA* gene of *Streptococcus pneumoniae* and comparison of the DNA polymerase I it encodes to homologous enzymes from *Escherichia coli* and phage T7. *J. Biol. Chem.* **264**:4255–4263.
32. Matsunaga, T., N. Matsunaga, K. Tsubaki, and T. Tanaka. 1986. Development of a gene cloning system for the hydrogen-producing marine photosynthetic bacterium *Rhodopseudomonas* sp. *J. Bacteriol.* **168**:460–463.
33. Matsunaga, T., K. Tsubaki, H. Miyashita, and J. G. Burgess. 1990. Chloramphenicol acetyltransferase expression in marine *Rhodobacter* sp. NKPB 0021 by use of shuttle vectors containing the minimal replicon of an endogenous plasmid. *Plasmid* **24**:90–99.
34. Moore, D. D., K. Denniston-Thompson, K. E. Kruger, M. E. Furth, B. G. Williams, D. L. Daniels, and F. R. Blattner. 1979. Dissection and comparative anatomy of the origins of replication of the lambdoid phages. Cold Spring Harbor Symp. Quant. Biol. **43**:155–163.
35. Mukherjee, S., I. Patel, and D. Bastia. 1985. Conformational changes in a replication origin induced by an initiator protein. *Cell* **43**:189–197.
36. Murotsu, T., K. Matsubara, H. Sugisaki, and M. Takanami. 1981. Nine unique repeating sequences in a region essential for replication and incompatibility of the mini-F plasmid. *Gene* **15**:257–271.
37. O'Connor, M., T. Asai, C. L. Squires, and A. E. Dahlberg. 1999. Enhancement of translation by the downstream box does not involve base pairing of mRNA with the penultimate stem sequence of 16S rRNA. *Proc. Natl. Acad. Sci. USA* **96**:8973–8978.
38. Ohkubo, S., and K. Yamaguchi. 1995. Two enhancer elements for DNA replication of pSC101, *par* and a palindromic binding sequence of the Rep protein. *J. Bacteriol.* **177**:558–565.
39. Okumura, M. S., and C. I. Kado. 1992. The region essential for efficient autonomous replication of pSa in *Escherichia coli*. *Mol. Gen. Genet.* **235**:55–63.
40. Pfennig, N., and H. G. Trüper. 1989. Anoxygenic phototrophic bacteria, p. 1635–1709. In J. T. Staley, M. P. Bryant, N. Pfennig, and J. G. Holt (ed.), *Bergey's manual of systematic bacteriology*, vol. 3. Williams & Wilkins, Baltimore, Md.
41. Ptashne, M., K. Backman, M. Z. Humayun, A. Jeffrey, R. Maurer, B. Meyer, and R. T. Sauer. 1976. Autoregulation and function of a repressor in bacteriophage lambda. *Science* **194**:156–161.
42. Saegesser, R., R. Ghosh, and R. Bachofen. 1992. Stability of broad host range cloning vector in the phototrophic bacterium *Rhodospirillum rubrum*. *FEMS Microbiol. Lett.* **95**:7–12.
43. Saitou, N., and M. Nei. 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* **4**:406–425.
44. Sambrook, J., E. F. Fritsch, and T. Maniatis. 1989. *Molecular cloning: a laboratory manual*, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
45. Schnös, M., K. Zahn, R. B. Inman, and F. R. Blattner. 1988. Initiation protein induced helix destabilization at the λ origin: a prepriming step in DNA replication. *Cell* **52**:385–395.
46. Sprengart, M. L., E. Fuchs, and A. G. Porter. 1996. The downstream box: an efficient and independent translation initiation signal in *Escherichia coli*. *EMBO J.* **15**:665–674.
47. Stalker, D. M., R. Kolter, and D. R. Helinski. 1982. Plasmid R6K DNA replication. I. Complete nucleotide sequence of an autonomously replicating segment. *J. Mol. Biol.* **161**:33–43.
48. Stalker, D. M., C. M. Thomas, and D. R. Helinski. 1981. Nucleotide sequence of the region of the origin of replication of the broad host range plasmid RK2. *Mol. Gen. Genet.* **181**:8–12.
49. Strohl, W. R. 1992. Compilation and analysis of DNA sequences associated with apparent streptomycete promoters. *Nucleic Acids Res.* **20**:961–974.
50. Tomb, J. F., and 40 other authors. 1997. The complete genome sequence of the gastric pathogen *Helicobacter pylori*. *Nature* **388**:539–547.
51. Tsurimoto, T., and K. Matsubara. 1981. Purified bacteriophage λ O protein binds to four repeating sequences at the λ replication origin. *Nucleic Acids Res.* **9**:1789–1799.
52. Umek, R. M., and D. Kowalski. 1988. The ease of DNA unwinding as a determinant of initiation at yeast replication origins. *Cell* **52**:559–567.
53. van Niel, C. B. 1944. The culture, general physiology, and classification of the non-sulfur purple and brown bacteria. *Bacteriol. Rev.* **8**:1–118.
54. Wada, C., and T. Yura. 1979. *Escherichia coli* mutants incapable of supporting replication of F-like plasmids at high temperature: isolation and characterization of *mafA* and *mafB* mutants. *J. Bacteriol.* **140**:864–873.
55. Weisburg, W. G., S. M. Barns, D. A. Pelletier, and D. J. Lane. 1991. 16S ribosomal DNA amplification for phylogenetic study. *J. Bacteriol.* **173**:697–703.
56. Williams, D. R., and C. M. Thomas. 1992. Active partitioning of bacterial plasmids. *J. Gen. Microbiol.* **138**:1–16.
57. Willison, J. C., J. P. Magnin, and P. M. Vignais. 1987. Isolation and characterization of *Rhodobacter capsulatus* strains lacking endogenous plasmids. *Arch. Microbiol.* **147**:134–142.
58. Winzler, E., and L. Shapiro. 1997. Translation of the leaderless *Caulobacter dnaX* mRNA. *J. Bacteriol.* **179**:3981–3988.
59. Woese, C. R. 1987. Bacterial evolution. *Microbiol. Rev.* **51**:221–271.
60. Wong, F. Y. K., E. Stackebrandt, J. K. Ladha, D. E. Fleischman, R. A. Date, and J. A. Fuerst. 1994. Phylogenetic analysis of *Bradyrhizobium japonicum* and photosynthetic stem-nodulating bacteria from *Aeschynomene* species grown in separated geographical regions. *Appl. Environ. Microbiol.* **60**:940–946.
61. Wu, F., I. Levchenko, and M. Filutowicz. 1994. Binding of DnaA protein to a replication enhancer counteracts the inhibition of plasmid R6K γ origin replication mediated by elevated levels of R6K π protein. *J. Bacteriol.* **176**:6795–6801.
62. Young, J. P., H. L. Downer, and B. D. Eardly. 1991. Phylogeny of the phototrophic rhizobium strain BTail by polymerase chain reaction-based sequencing of a 16S rRNA segment. *J. Bacteriol.* **173**:2271–2277.
63. Zahn, K., and F. R. Blattner. 1985. Binding and bending of the λ replication origin by the phage O protein. *EMBO J.* **4**:3605–3616.
64. Zhang, Q., S. S. De Oliveira, R. Colangeli, and M. L. Gennaro. 1997. Binding of a novel host factor to the pT181 plasmid replication enhancer. *J. Bacteriol.* **179**:684–688.