

Characterization of Plasmid pRT1 from *Pyrococcus* sp. Strain JT1

Donald E. Ward,^{1*} Ingrid M. Revet,¹ Renu Nandakumar,² Jon H. Tuttle,³ Willem M. de Vos,¹ John van der Oost,¹ and Jocelyne DiRuggiero²

Laboratory for Microbiology, Wageningen University, 6703CT Wageningen, The Netherlands¹; Department of Cell Biology and Molecular Genetics, University of Maryland, College Park, Maryland 20742²; and Chesapeake Biological Laboratory, Center for Environmental and Estuarine Studies, University of Maryland System, Solomons, Maryland 20688³

Received 30 August 2001/Accepted 7 February 2002

We discovered a 3,373-bp plasmid (pRT1) in the hyperthermophilic archaeon *Pyrococcus* sp. strain JT1. Two major open reading frames were identified, and analysis of the sequence revealed some resemblance to motifs typically found in plasmids that replicate via a rolling-circle mechanism. The presence of single-stranded DNA replication intermediates of pRT1 was detected, confirming this mode of replication.

In recent years, genetic elements (plasmids, viruses, and insertion sequence elements) have been found throughout the hyperthermophilic bacteria and archaea, some of which have been characterized in detail (for recent reviews, see references 16, 20, and 22). While these elements can serve as excellent models for the study of various aspects of genetics, such as DNA replication and gene expression, their greatest potential may be to serve as the templates for the building of genetic systems for the hyperthermophiles. Many hyperthermophilic bacteria and archaea have been found to harbor an endogenous plasmid(s). They can range in size from the 0.8-kb plasmid pRQ7 from *Thermotoga maritima* (8) to the over-40-kb conjugative plasmid pNOB8 from *Sulfolobus* (17, 18). Despite significant effort to link the plasmids to phenotypic traits, the majority of plasmids characterized to date appear to be cryptic. A study to assess the plasmid diversity in the *Thermococcales* revealed the presence of a number of plasmids in isolates of *Pyrococcus* and *Thermococcus*, ranging in size from 3 to 26.8 kb (2). A total of 57 strains were examined for the presence of extrachromosomal elements; 11 of the examined strains harbored plasmids, suggesting that plasmids are more diverse and widely distributed among the *Thermococcales* than previously thought (2). In the *Thermococcales*, only the small cryptic plasmid pGT5 from *Pyrococcus abyssi* has been characterized in detail (6, 7). This plasmid is 3.4 kb in size and contains two open reading frames (ORFs), the larger of which encodes the Rep protein necessary for plasmid replication. Based on the similarity of the Rep protein to the Rep proteins of plasmids from that pC194/pUB110 family of rolling-circle plasmids and the identification of single-stranded replication intermediates, it was determined that pGT5 replicated via the rolling-circle mechanism. We here describe the isolation and characterization of a second pyrococcal plasmid, from *Pyrococcus* sp. strain JT1. The 3.4-kb plasmid (pRT1) was sequenced and found to be distinct from pGT5. The plasmid encodes two proteins, both which are expressed. Evidence is presented that, similarly

to pGT5, this plasmid replicates via the rolling-circle mechanism.

Pyrococcus sp. strain JT1 was isolated from samples from the Juan de Fuca Ridge in the North Pacific and was grown anaerobically at 85°C in a sea salt medium which contained 5 g of tryptone/liter and 1 g of yeast extract/liter as the carbon and energy source and 1% elemental sulfur (21). Analysis of the 16S rRNA showed 98% identity to those of other pyrococci (*P. glycovorans* and *P. abyssi*). The organism grows optimally at 95°C and requires S⁰ for growth. During the purification of total DNA, a faster-migrating band was identified that was believed to be a plasmid. The plasmid pRT1 was isolated from 250 ml of *Pyrococcus* sp. strain JT1 cells that had been grown until stationary phase. The cells were harvested by centrifugation, washed with 5 ml of 300 mM NaCl, and resuspended in 0.5 ml of ice-cold Tris-EDTA buffer to which 3.75 ml of guanidine thiocyanate solution had been added. After a 5-min incubation at room temperature, 375 μ l of 2 M sodium acetate (pH 4.5) and an equal volume of phenol-chloroform (5:1 [vol/vol]; pH 4.5) were added. The mixture was vortexed and incubated on ice for 5 min, and phase separation was obtained by centrifugation (10,000 \times g for 20 min at 4°C). The aqueous phase was extracted twice with equal volumes of phenol-chloroform-isoamyl alcohol (25:24:1 [vol/vol]; pH 8) and once with chloroform-isoamyl alcohol (24:1 [vol/vol]). The aqueous phase was removed, and the plasmid was precipitated by the addition of 1/10 volume of 3 M sodium acetate (pH 5.5) and 2.5 volume of 96% ethanol followed by a 2-h incubation at –80°C. The pelleted plasmid DNA was washed three times with 70% ethanol, dried, and resuspended in 50 μ l of 10 mM Tris (pH 8.5). The sample was then treated with DNase-free RNase for 1 h at 37°C to remove contaminating RNA.

The plasmid was digested with either *Pst*I or *Sau*3A1, and the fragments were cloned into the similarly digested plasmid pUC19. Sequencing reactions were carried out on a LiCor 4000L sequencer with a Thermo Sequenase fluorescence-labeled primer cycle sequencing kit with 7-deaza-dGTP (Amersham) and infrared-labeled oligonucleotides M13forward and M13reverse (MWG-Biotech). Gaps in the sequence were joined by means of oligonucleotides designed to span the ambiguous regions; the oligonucleotides were either cloned into

* Corresponding author. Present address: Department of Chemical Engineering, North Carolina State University, Campus Box 7270, Raleigh, NC 27695-7270. Phone: (919) 515-4452. Fax: (919) 515-3465. E-mail: deward@unity.ncsu.edu.

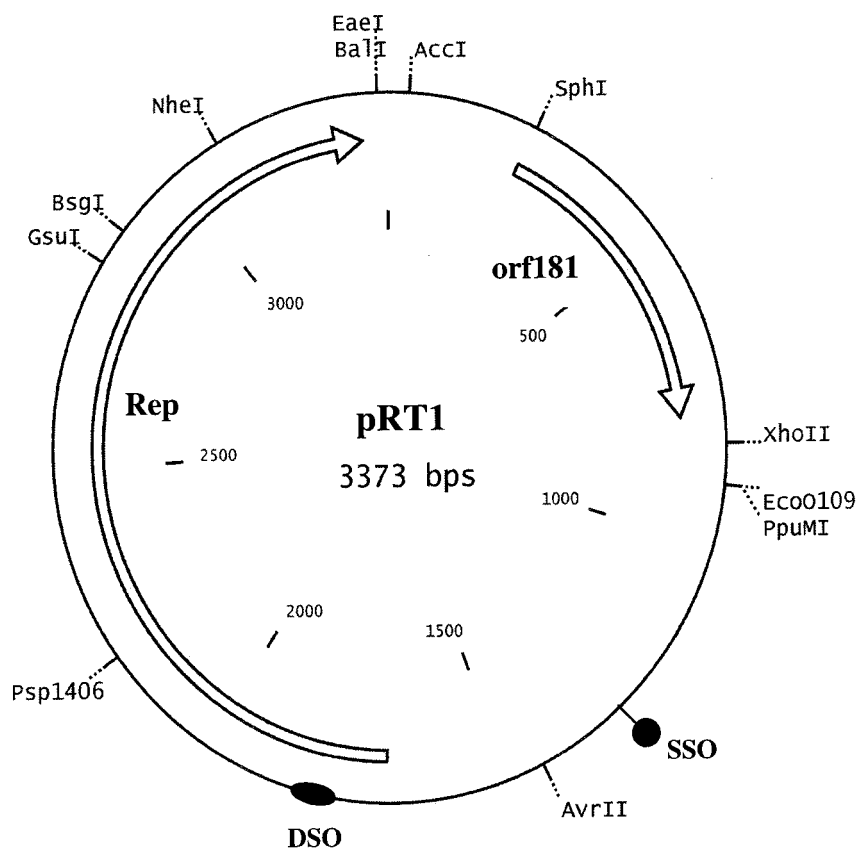


FIG. 1. Map of pRT1. The genes encoding the putative Rep, ORF181, are indicated, as are the putative DSO and SSO.

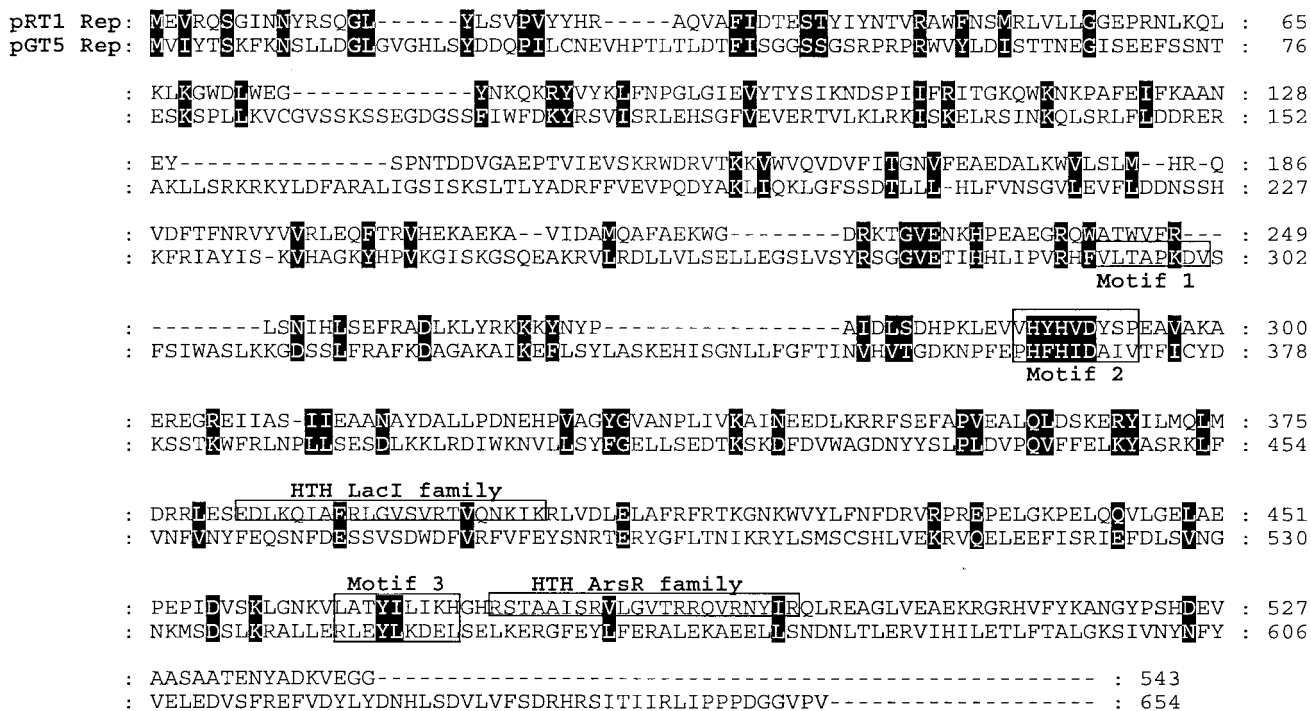
pUC19 or sequenced directly by using specific oligonucleotides.

Pyrococcus sp. strain JT1 harbors a single plasmid of approximately 3.4 kb; a structural map is shown in Fig. 1. The G+C content is 43%, which is in good agreement with those of the genomes of various pyrococcal species (38 to 45%). Comparison to the *P. abyssi* plasmid pGT5 revealed that while the plasmids were similar in comparisons based on their genetic organizations, the plasmids were only 41% identical at the DNA level, which clearly distinguishes the two plasmids. Sequence analysis identified eight ORFs of at least 50 amino acids in size, two of which were larger than 150 amino acids in size and were found on the same strand. The largest ORF, designated Rep, encodes a 62.9-kDa protein, and TFASTA and BLAST analysis of the translated sequence did not reveal any homologs in the database. However, a low level of identity to transcriptional regulators of the ArsR family was observed in the C-terminal region. A search for motifs by using ScanProsite (<http://ca.expasy.org/tools/scnpsite.html>) revealed the presence of two helix-turn-helix (HTH) motifs of the LacI and ArsR families in the C-terminal region (Fig. 2A). By use of the HTH motif prediction (http://npsa-pbil.ibcp.fr/cgi-bin/npsa_automat.pl?page=/NPSA/npsa_hth.html) devised by Dodd and Egan (5), both HTH domains were identified, with scores of 5.77 for the LacI domain and 5.90 for the ArsR domain. Both scores signify a 100% probability that HTH domains are found in these areas of the protein. Closer visual inspection of the Rep protein sequence also revealed sequences that resem-

ble those of motif II and motif III, which are found in Rep proteins that replicate via the rolling-circle mechanism (4, 12). Motif II (consensus, xxHYHUUUxx) is believed to be involved in binding of the metal ions Mg^{2+} or Mn^{2+} , as these are usually required for the activity of the Rep. Motif III [consensus, uxxYux(K,H)xx] contains the active site (tyrosine) that initiates the nucleophilic attack at the double (leading)-strand origin (DSO), which results in a nick in the DNA, thereby generating a free 3' hydroxyl for the host's DNA polymerase (4, 9). Aside from the Rep protein, rolling-circle plasmids require a DSO and a single (lagging)-strand origin (SSO) (4, 12). The SSO is usually found in a noncoding region of the plasmid, while the DSO is located upstream or even within the gene encoding the Rep protein. Numerous direct repeats and a single inverted repeat were detected. Two sets of 8-bp direct repeats were located between *orf181* and the *rep* gene. An inverted repeat which could serve as the SSO was identified upstream of *rep* (at nucleotides 1280 to 1330). Furthermore, a putative DSO [TTTATT(G/A)TA] was also identified within the gene encoding Rep (nucleotides 1786 to 1795) and is also found on the same coding strand as the Rep protein, as is common among the DSOs. The DSO of pRT1 exhibited significant identity to the DSOs of both pGT5 and pC194, further suggesting that the mode of replication of this plasmid is via the rolling-circle mechanism (Fig. 2).

The second ORF, designated ORF181, encodes a putative protein of 29.1 kDa. Similarly to the Rep protein, homologs to ORF181 could not be found after TFASTA and BLAST anal-

A.



B.

pRT1	DSO	GGGTGCTTTTATTG	ATA
pGT5	DSO	GTTGGGTTTATCTTG	ATA
pC194	DSO	TTCTTTCTTATCTTG	ATA

FIG. 2. (A) CLUSTAL alignment of the Rep proteins from the pyrococcal plasmids pRT1 and pGT5. The motifs involved in rolling-circle replication are boxed. Black boxes indicate conserved residues. (B) Comparison of the putative DSOs of pRT1, pGT5, and pC194.

yses and ORF181 was only 16.5% identical to the ORF2 of pGT5. A search for known structural motifs by using ScanProsite did not identify any known motifs. However, a putative motif was identified during the BLAST searches. It was found that ORF181 has a low level of identity to ORF80 of the pRN family of plasmids. This family includes pRN1 (ORF80) and pRN2 (ORF81) from *Sulfolobus* (10, 11) and pDL10 (ORF71) from *Acidianus ambivalens* (13). This level of identity was due to the presence of a 35-amino-acid region in the C-terminal region of these ORFs that is approximately 50% identical across the four proteins (Fig. 3). Interestingly, this motif is absent in ORF2 of pGT5. Based on the basic character of the predicted protein and the conservation across the three-pRN family of plasmids, it has been suggested that it might be a DNA-binding protein (13). However, typical DNA-binding motifs such as HTH domains are not present in this region or in any other part of the protein. Hence, an anticipated role of this protein in copy-number control or plasmid segregation remains to be elucidated.

Total DNA and plasmid preparations from *Pyrococcus* sp.

strain JT1 were analyzed for the presence of single-stranded DNA (ssDNA) replication intermediates of pRT1. ssDNA was detected in total and plasmid DNA samples as described by Noirot-Gros and Ehrlich (15). A 3.0-kb probe was generated by PCR using the primers BG910 (5'-CTTGAGGCTCTCTT TAAACGC-3') and BG908 (5'-TCATCCCCCTTCTACCTT ATC-3'). The PCR product was purified by using Qiaquick (Qiagen) and labeled with [α -³²P]dATP by nick translation. Three forms of the plasmid, namely, open circle, supercoiled, and single stranded, were observed in the samples that were not treated with S1 nuclease under denaturing conditions (Fig. 4). The addition of S1 nuclease resulted in a loss of the third band, confirming that it is the single-stranded form of the plasmid. Under non-denaturing conditions, only the single-stranded form of the plasmid was observed and addition of S1 nuclease again resulted in disappearance of the band. The fact that the single-stranded form appears as a faint smear instead of a distinct band may be due to the fact that the cells were growing rapidly and the plasmid was replicating readily; therefore, the smearing might have been caused by different forms

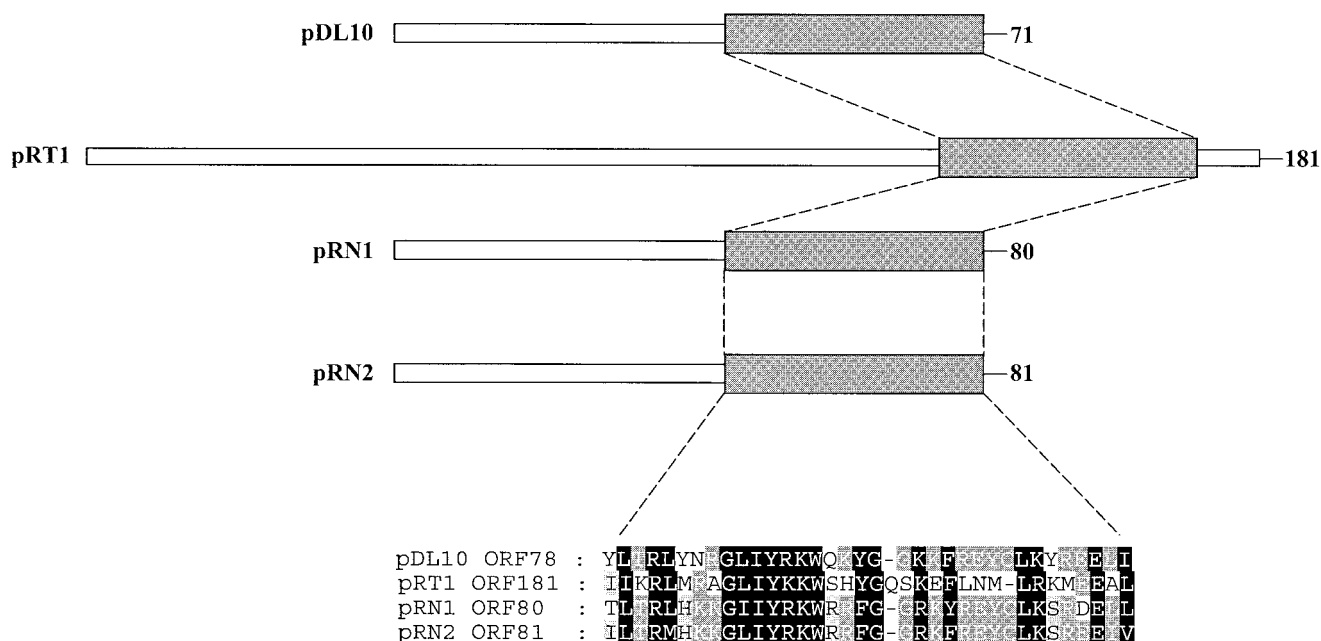


FIG. 3. Comparison of ORF181 from pRT1 to ORF80 of the pRN family of plasmids.

of the single-stranded intermediate undergoing lagging strand synthesis.

In an effort to identify the transcripts encoded by pRT1, total RNA was isolated from 250 ml of exponentially growing *Pyrococcus* sp. strain JT1 as described previously (21). Initial attempts to detect transcripts of either *rep* or *orf181* by Northern analysis were unsuccessful. This is not surprising, as the levels of these transcripts are generally very low. Consequently, reverse transcription-PCR (RT-PCR) was used to determine if both *rep* and *orf181* were transcribed and, if so, whether they were transcribed as distinct monocistronic or as polycistronic transcripts. RT-PCR was carried out by using the Access RT-PCR kit (Promega, Madison, Wis.) according to the manufacturer's instructions. A typical reaction contained 100 ng of total RNA as a template, and the following oligonucleotide pairs were used for *rep*: BG863 (5'-GCGCGGATCCCTCAGATAAGTGAATATTG-3') and BG909 (5'-GCAAACAGTGGAAAAATAGCC-3') and *orf181*: BG885 (5'-CGCGGAATTCTTGAATGATAACGCCTTTGCC-3') and BG910 (5'-CTTGAGGCTCTCTTTAAACGC-3'). The reverse transcription reaction was carried out at 42°C for 30 min, after which the PCR was started (1 min at 95°C, 1.5 min at 48°C, and 2 min at 72°C for 40 cycles). RT-PCR gave rise to PCR products of the correct size for both *rep* and *orf181* (Fig. 5A). However, RT-PCR products were not obtained when using oligonucleotides that would demonstrate the presence of an *orf181-rep* transcript (data not shown). We conclude that both *orf181* and *rep* are transcribed during replication of the plasmid as distinct monocistronic mRNAs.

The 5' end of *rep* was mapped by using the oligonucleotide 5'-GTATCAATAAAAGCAACCTGTGCC-3' as described previously (21). This oligonucleotide was labeled with a fluorescent label, IRD800, by the manufacturer, Biolegio (Malden, The Netherlands). A single major apparent mRNA 5' end was identified upstream of *rep* by primer extension (Fig. 5B). The

transcriptional start was a pyrimidine base, rather than the more common purine, and was located 6 bp upstream of the ATG start codon. Analysis of the upstream sequence indicated that it was relatively G+C rich and did not reveal the presence of features common to archaeal promoters, which include the TATA box found 25 bp upstream of the transcription initiation site and the transcription factor B-responsive element or the transcription factor IIB-responsive element (1). Moreover, no obvious ribosome-binding site was present on the transcript.

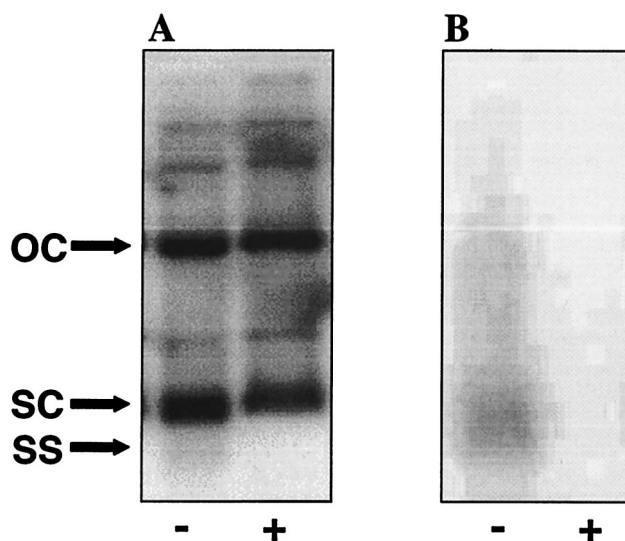


FIG. 4. Detection of the pRT1 ssDNA intermediate. Purified plasmid DNA from *Pyrococcus* sp. strain JT1 was electrophoresed on a 1.5% agarose gel with (+) or without (-) prior S1 nuclease treatment. The DNA was (A) or was not (B) denatured prior to transfer to a nylon membrane. Bands representing open circle (OC), supercoiled (SC), and single-stranded (SS) forms of pRT1 are indicated.

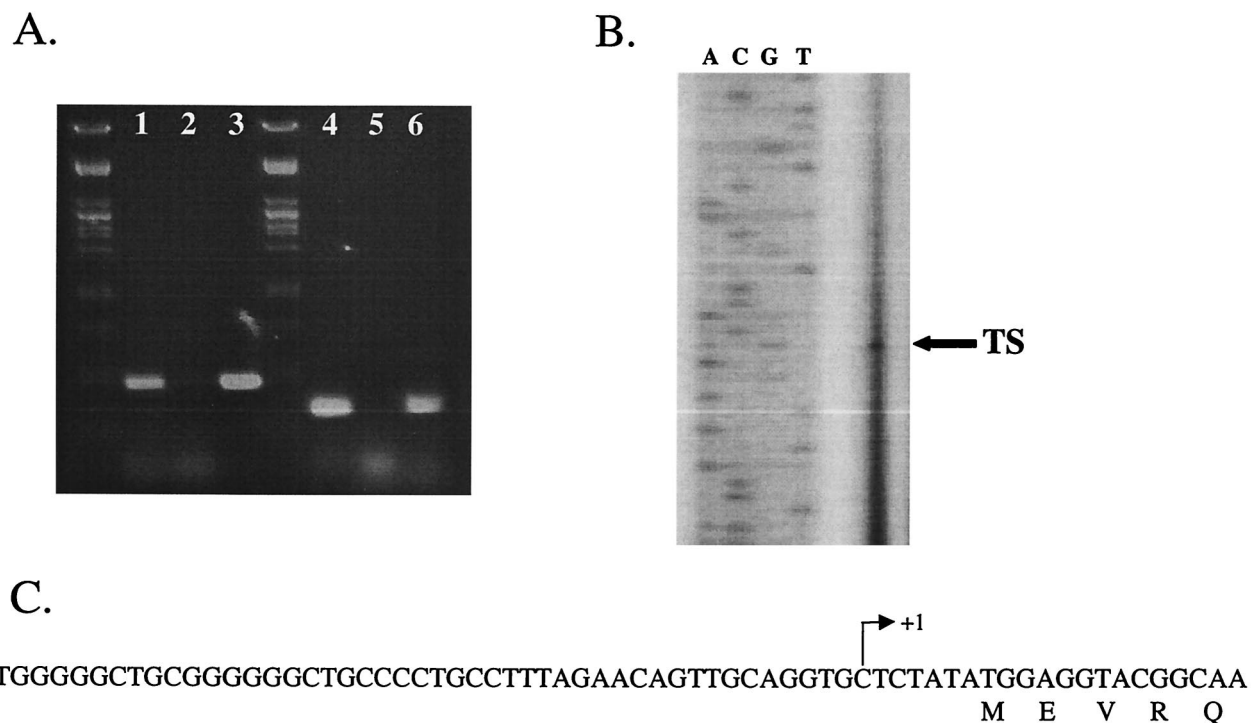


FIG. 5. (A) RT-PCR analysis of the *rep* and *orf181* genes from pRT1. Lane 1, *rep* RT-PCR; lane 2, *rep* negative control; lane 3, *rep* positive control; lane 4, *orf181* RT-PCR; lane 5, *orf181* negative control; lane 6, *orf181* positive control. (B) Identification by primer extension analysis of the transcriptional start site for the *rep* gene; the arrow indicates the primer extension product. (C) Sequence of the *rep* promoter region. The transcriptional start site, as determined by primer extension analysis, is designated by an arrow.

The apparent lack of a ribosome-binding site in the leader might suggest that it was instead located downstream of the translation start, possibly at nucleotide positions +3 to +7 (GGAGG), as has previously been reported for haloarchaea (14). Alternatively, translation of the *rep* transcript might occur via a leaderless mechanism. Leaderless mechanisms have been observed in both *Sulfolobus solfataricus* (3) and *Pyrobaculum aerophilum* (19). Attempts to identify the promoter of *orf181* were unsuccessful.

The ultimate goal of these studies is the development of a genetic system for the *Thermococcales*. pRT1 could serve as an excellent template for the generation of a *Pyrococcus-Escherichia coli* shuttle vector. The presence of a unique *AccI* site between *rep* and *orf181* is ideal for the insertion of an *E. coli* plasmid, as it does not interrupt any of the ORFs and is distant from the predicted DSO and SSO. Furthermore, the insertion should not interrupt transcription of the ORFs, since RT-PCR analysis could not detect a transcript in this region and an insertion in this site is also sufficiently upstream (210 bp) of *orf181* that the promoter of *orf181* should not be affected. In an effort to avoid instability of the construct in *E. coli*, a low-copy (7 to 12 copies) cloning plasmid will be used. Efforts to make this construct, as well as develop a selective marker, are presently under way.

Nucleotide sequence accession numbers. The pRT1 sequence reported here and the 16S rRNA sequence of *Pyrococcus* sp. strain JT1 have been submitted to GenBank with the accession numbers AF393813 and AF411292, respectively.

The research was partly supported by the EU project Extremophiles as Cell Factory (contract BIO-CT96-0488) and by grants from NSF (MCB-0196309) and NASA (NCC8-175) to J.D.R.

REFERENCES

- Bell, S. D., and S. P. Jackson. 2001. Mechanism and regulation of transcription in archaea. *Curr. Opin. Microbiol.* **4**:208–213.
- Benbouzid Rollet, N., P. Lopez Garcia, L. Watrin, G. Erauso, D. Prieur, and P. Forterre. 1997. Isolation of new plasmids from hyperthermophilic archaea of the order Thermococcales. *Res. Microbiol.* **148**:767–775.
- Condo, I., A. Ciannaruconi, D. Benelli, D. Ruggero, and P. Londei. 1999. *cis*-acting signals controlling translational initiation in the thermophilic archaeon *Sulfolobus solfataricus*. *Mol. Microbiol.* **34**:377–384.
- del Solar, G., R. Giraldo, M. J. Ruiz-Echevarria, M. Espinosa, and R. Diaz-Orejas. 1998. Replication and control of circular bacterial plasmids. *Microbiol. Mol. Biol. Rev.* **62**:434–464.
- Dodd, I. B., and J. B. Egan. 1990. Improved detection of helix-turn-helix DNA-binding motifs in protein sequences. *Nucleic Acids Res.* **18**:5019–5026.
- Erauso, G., F. Charbonnier, T. Barbeyron, P. Forterre, and D. Prieur. 1992. Preliminary characterization of a hyperthermophilic archaeobacterium with a plasmid, isolated from a north Fiji basin hydrothermal vent. *C. R. Acad. Sci. Ser. III Sci. Vie* **314**:387–393.
- Erauso, G., S. Marsin, N. Benbouzid-Rollet, M.-F. Baucher, T. Barbeyron, Y. Zivanovic, D. Prieur, and P. Forterre. 1996. Sequence of plasmid pGT5 from the archaeon *Pyrococcus abyssi*: evidence for rolling-circle replication in a hyperthermophile. *J. Bacteriol.* **178**:3232–3237.
- Harriott, O. T., R. Huber, K. O. Stetter, P. W. Betts, and K. M. Noll. 1994. A cryptic miniplasmid from the hyperthermophilic bacterium *Thermotoga* sp. strain RQ7. *J. Bacteriol.* **176**:2759–2762.
- Ilyina, T. V., and E. V. Koonin. 1992. Conserved sequence motifs in the initiator proteins for rolling circle DNA replication encoded by diverse replicons from eubacteria, eucaryotes and archaeobacteria. *Nucleic Acids Res.* **20**:3279–3285.
- Keeling, P. J., H. P. Klenk, R. K. Singh, O. Feeley, C. Schleper, W. Zillig, W. F. Doolittle, and C. W. Sensen. 1996. Complete nucleotide sequence of the *Sulfolobus islandicus* multicopy plasmid pRN1. *Plasmid* **35**:141–144.
- Keeling, P. J., H. P. Klenk, R. K. Singh, M. E. Schenk, C. W. Sensen, W. Zillig, and W. F. Doolittle. 1998. *Sulfolobus islandicus* plasmids pRN1 and

- pRN2 share distant but common evolutionary ancestry. *Extremophiles* **2**:391–393.
12. Khan, S. A. 1997. Rolling-circle replication of bacterial plasmids. *Microbiol. Mol. Biol. Rev.* **61**:442–455.
 13. Kletzin, A., A. Lieke, T. Ulrich, R. L. Charlebois, and C. W. Sensen. 1999. Molecular analysis of pDL10 from *Acidianus ambivalens* reveals a family of related plasmids from extremely thermophilic and acidophilic archaea. *Genetics* **152**:1307–1314.
 14. Ng, W. V., S. P. Kennedy, G. G. Mahairas, B. Berquist, M. Pan, H. D. Shukla, S. R. Lasky, N. S. Baliga, V. Thorsson, J. Sbrogna, S. Swartzell, D. Weir, J. Hall, T. A. Dahl, R. Welti, Y. A. Goo, B. Leithauser, K. Keller, R. Cruz, M. J. Danson, D. W. Hough, D. G. Maddocks, P. E. Jablonski, M. P. Krebs, C. M. Angevine, H. Dale, T. A. Isenbarger, R. F. Peck, M. Pohlschroder, J. L. Spudich, K. W. Jung, M. Alam, T. Freitas, S. Hou, C. J. Daniels, P. P. Dennis, A. D. Omer, H. Ebhardt, T. M. Lowe, P. Liang, M. Riley, L. Hood, and S. DasSarma. 2000. Genome sequence of *Halobacterium* species NRC-1. *Proc. Natl. Acad. Sci. USA* **97**:12176–12181.
 15. Noiro-Gros, M., and S. D. Ehrlich. 1994. Detection of single-stranded plasmid DNA. *Methods Mol. Genet.* **3**:370–379.
 16. Noll, K. M., and M. Vargas. 1997. Recent advances in genetic analyses of hyperthermophilic archaea and bacteria. *Arch. Microbiol.* **168**:73–80.
 17. Schleper, C., I. Holz, D. Janekovic, J. Murphy, and W. Zillig. 1995. A multicopy plasmid of the extremely thermophilic archaeon *Sulfolobus* effects its transfer to recipients by mating. *J. Bacteriol.* **177**:4417–4426.
 18. She, Q., H. Phan, R. A. Garrett, S. V. Albers, K. M. Stedman, and W. Zillig. 1998. Genetic profile of pNOB8 from *Sulfolobus*: the first conjugative plasmid from an archaeon. *Extremophiles* **2**:417–425.
 19. Slupska, M. M., A. G. King, S. Fitz-Gibbon, J. Besemer, M. Borodovsky, and J. H. Miller. 2001. Leaderless transcripts of the crenarchaeal hyperthermophile *Pyrobaculum aerophilum*. *J. Mol. Biol.* **309**:347–360.
 20. Sowers, K. R., and H. J. Schreier. 1999. Gene transfer systems for the archaea. *Trends Microbiol.* **7**:212–219.
 21. Ward, D. E., C. J. Donnelly, M. E. Mullendore, J. van der Oost, W. M. de Vos, and E. J. Crane III. 2001. The NADH oxidase from *Pyrococcus furiosus*. Implications for the protection of anaerobic hyperthermophiles against oxidative stress. *Eur. J. Biochem.* **268**:5816–5823.
 22. Zillig, W., H. P. Arnold, I. Holz, D. Prangishvili, A. Schweier, K. Stedman, Q. She, H. Phan, R. Garrett, and J. K. Kristjansson. 1998. Genetic elements in the extremely thermophilic archaeon *Sulfolobus*. *Extremophiles* **2**:131–140.