JOURNAL OF BACTERIOLOGY, May 1994, p. 2759–2762
0021-9193/94/$04.00+0

Vol. 176, No. 9

A Cryptic Miniplasmid from the Hyperthermophilic Bacterium *Thermotoga* sp. Strain RQ7

OLIVIA T. HARRIOTT,[1] ROBERT HUBER,[2] KARL O. STETTER,[2] PAUL W. BETTS,[3] AND KENNETH M. NOLL[1]*

Department of Molecular and Cell Biology[1] and Biotechnology Center,[3] University of Connecticut, Storrs, Connecticut 06269, and Institute of Microbiology, University of Regensburg, Regensburg, Germany[2]

Received 10 December 1993/Accepted 27 February 1994

An 846-bp cryptic plasmid has been discovered in the hyperthermophilic bacterium *Thermotoga* sp. strain RQ7. This is the first plasmid described for an organism from this ancient bacterial lineage and the smallest plasmid described to date for any organism. Nucleotide sequencing revealed a single open reading frame possibly encoding a 25,460-Da basic protein (212 amino acids). Upstream of the putative promoter lie five 11-bp direct repeats, each separated by 1 to 4 bp, while between the promoter and the open reading frame lies an 11-bp palindromic sequence. Its mode of replication is unknown, but its sequence bears similarities to those of plasmids which replicate by a rolling-circle mechanism.

Species of the genus *Thermotoga* are among the most slowly evolving organisms belonging to the domain *Bacteria* (1). Analyses of 16S rRNA gene sequences demonstrate that members of the order *Thermotogales* diverged prior to the emergence of all bacterial lineages except the *Aquifex* lineage (4). During our search for a cloning vector for *Thermotoga* species, we discovered a miniplasmid only 846 bp in size in *Thermotoga* sp. strain RQ7. This is the smallest replicon so far described for any organism. Basic replicons (miniplasmids) encoding only the minimal information necessary for their replication are typically 1.3 to 5 kp (8, 12). This plasmid is much smaller than those previously described yet is stably maintained in its host. The minimal genetic information necessary for stable plasmid replication includes at least one origin of replication and frequently one or more Rep (replication) proteins (12). On the basis of the sequence of this plasmid, we postulate the existence of a Rep protein and speculate as to the origin of replication upon which it acts.

Members of the genus *Thermotoga* are strict anaerobes that ferment carbohydrates and grow at temperatures of up to 90°C. They are found in geothermal environments around the world. The phylogenetic position of *Thermotoga* species makes them an attractive subject for comparative analyses. Although they are bacteria, their physiology is similar to those of thermophilic archaea such as *Pyrococcus* and *Thermococcus* species which appeared early in archaeal evolution (2). As such, *Thermotoga* species can serve as an important benchmark for tracing the progress of bacterial evolution and in understanding the events that gave rise to the division between bacteria and archaea.

Plasmid isolation. During the preparation of chromosomal DNA from a number of *Thermotoga* isolates by cesium chloride gradient centrifugation, extrachromosomal DNA was observed. One of these preparations, from *Thermotoga* sp. strain RQ7, yielded a diffuse extrachromosomal band following agarose gel electrophoresis. This band migrated to a position just above that where RNA typically migrates. When DNA was isolated from this strain by an alkaline lysis miniprep method (9) and analyzed by gel electrophoresis, two bands were observed (Fig. 1). The upper (more slowly migrating) band was often barely visible in ethidium bromide-stained gels (Fig. 1). The yield of total plasmid DNA was typically about 1 μg of DNA from a minipreparation of a 20-ml overnight culture (approximately 5×10^9 cells). This suggests that each cell contains on average a minimum of 200 copies of the plasmid.

We tested a number of restriction endonucleases with 4- and 6-bp recognition sites for the ability to digest the plasmid. *Eco*RV and *Hin*dIII were among the few enzymes capable of digesting the plasmid, but *Hin*dIII failed to do so reproducibly. The digested plasmid migrated as a single band to a position between the two bands of the minipreparations (Fig. 1). The digest band migrated to a position indicating that it contained DNA of approximately 850 bp.

Cloning and sequencing pRQ7. To clone the plasmid for sequencing, plasmid DNA was isolated by an alkaline lysis miniprep method (9). This DNA was digested with *Eco*RV, and the linearized plasmid was ligated to *Eco*RV-digested pBluescript KS+ (Stratagene Cloning Systems, La Jolla, Calif.). Plasmid preparations from the resulting white colonies were examined electrophoretically for the presence of a vector containing an insert (Fig. 1). One of these, pOH19, was used for further analyses.

To examine whether the DNA cloned into pOH19 was pRQ7 or was derived from the chromosome of strain RQ7 and to assess the distribution of pRQ7 among other *Thermotoga* strains, we prepared a hybridization probe from the cloned insert of pOH19 and probed a nylon membrane containing chromosomal DNA and DNA obtained from the minipreparation procedure from four *Thermotoga* isolates (Fig. 2). Chromosomal DNAs were digested with *Eco*RI, which does not cut pRQ7. The probe hybridized only with DNA from strain RQ7, which migrated to the same position as pRQ7 DNA. Therefore, the cloned DNA is not a fragment of the chromosome or any other replicon in strain RQ7. DNA sharing significant sequence identity with the plasmid does not appear to be present in the chromosomes of any of the strains tested, nor do any of the other strains contain pRQ7. Both bands of DNA isolated from strain RQ7 by the minipreparation procedure hybridized to the probe. It appears that during the preparation of chromosomal DNA, the plasmid is found

* Corresponding author. Phone: (203) 486-4688. Fax: (203) 486-4331. Electronic mail address: noll@uconnvm.uconn.edu.

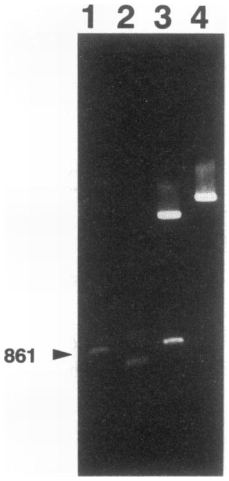

FIG. 1. Electrophoretic analyses of pRQ7 and pOH19. DNA samples were applied to a 1% agarose gel and subjected to electrophoresis in 1× Tris-borate-EDTA buffer. DNA was visualized by ethidium bromide staining. The size marker was a 123-bp ladder (Gibco BRL, Gaithersburg, Md.). The location of the 861-bp marker is indicated. The samples are undigested pRQ7 (lane 1), pRQ7 digested with *Eco*RV (lane 2), pOH19 digested with *Eco*RV (lane 3), and pBluescript (3.4 kbp) digested with *Eco*RV (lane 4).

primarily in the form corresponding to the upper band in the uncut minipreparations.

The plasmid was sequenced by the dideoxy chain termination method using both fluorescent primers (with a Pharmacia ALF sequencer at the University of Connecticut Biotechnology Center) and radiolabelled primers. Both strands were sequenced. Although the plasmid was cloned into the vector at its *Eco*RV site, the sequence of the plasmid shown in Fig. 3 begins with the first direct repeat for clarity. The plasmid consists of 846 bp with a G+C content of 40%, which is slightly below that of chromosomal DNA from other *Thermotoga* strains (41 to 45%). Consistent with our experimental observations, the sequence revealed sites for *Eco*RV and *Hin*dIII and no sites for *Eco*RI, *Sau*3AI, or *Bam*HI. Interestingly, a single site for *Spe*I was found. This enzyme cuts the chromosomes of *Thermotoga neapolitana* and *Thermotoga maritima* very infrequently (unpublished observations).

DNA sequences were analyzed by using Genetics Computer Group sequence analysis software and compared with sequences in the GenBank (release 72.0), EMBL (release 31.0), VecBase (release 3.0), PIR-Nucleic (release 32.0), and PIR-Protein (release 32.0) data bases (5). BLAST analyses were performed at the National Center for Biotechnology Information by using the BLAST network service (3).

Examination of the sequence revealed five direct repeats of 11 bp (GATAAGGNTWW) separated by 1 bp, except the third and fourth repeats, which are separated by 4 bp. These repeats are followed by a potential promoter consisting of a −35 region (GTGAC) differing from the *T. maritima* consensus (TTGAC) by one nucleotide and a −10 region (TATAAT) identical to the consensus sequence (10). Thirty base pairs downstream from the −10 region are two 11-bp inverted repeats separated by 6 bp which together could form a cruciform structure or encode an mRNA with a hairpin loop. Twelve base pairs further downstream, a sequence encoding a ribosome binding site contains 11 of 12 nucleotides complementary to the 3′ end of the 16S rRNA of *T. maritima* (1, 6, 10, 14).

All six reading frames were examined, and only one large open reading frame was found. A second region following the putative promoter could encode a polypeptide of 20 amino

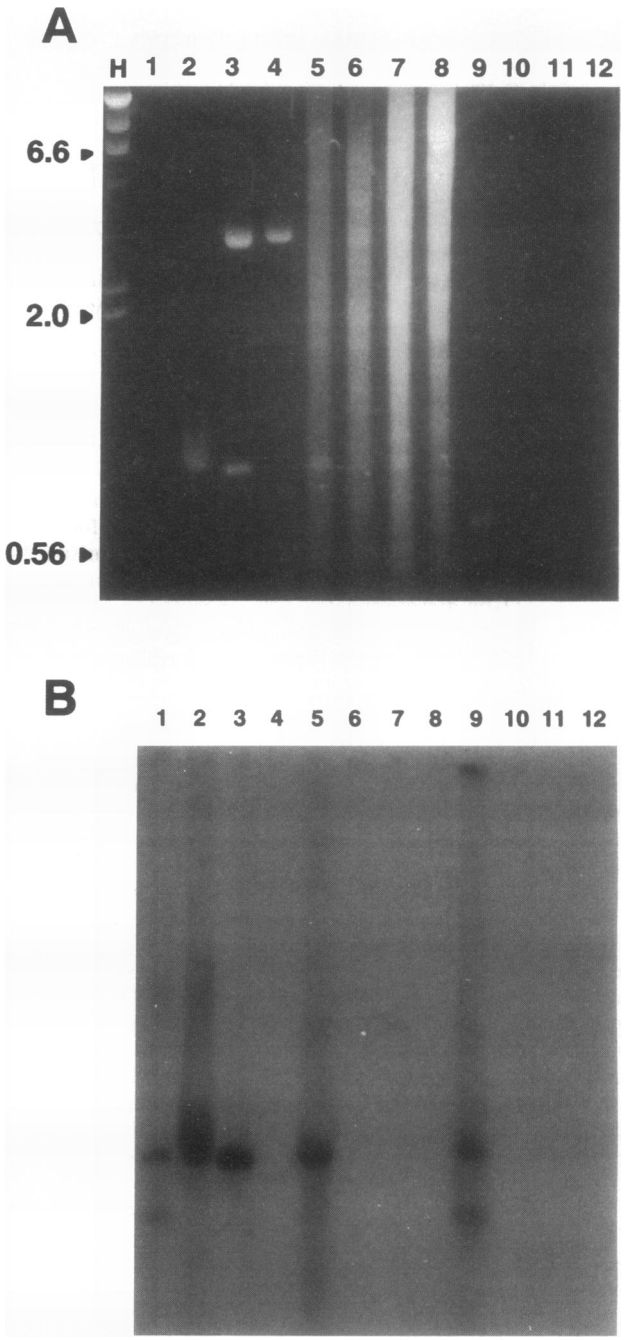

FIG. 2. (A) Electrophoretic analyses of plasmids and chromosomal DNAs. DNA samples were applied to a 0.8% agarose gel and subjected to electrophoresis in 1× Tris-borate-EDTA buffer. Fragments of bacteriophage lambda DNA generated by *Hin*dIII digestion were used as size markers (lane H). Sizes (at left) are in kilobase pairs. Lanes 1 to 4 are identical to those described for Fig. 1. Chromosomal DNA digested with *Eco*RI (lanes 5 to 8) and plasmid minipreparations (lanes 9 to 12) were prepared with DNA from four *Thermotoga* strains. *Thermotoga* sp. strains RQ7 and SG7 were isolated from samples collected from thermal sites in the Azores and Indonesia, respectively. *T. maritima* MSB8 and *T. neapolitana* NSE have been described previously (7a, 7b). The lanes contain *Thermotoga* sp. strain RQ7 (lanes 5 and 9), *T. maritima* MSB8 (lanes 6 and 10), *T. neapolitana* NSE (lanes 7 and 11), and *Thermotoga* sp. strain SG7 (lanes 8 and 12). (B) Southern blot analyses of pRQ7. DNA resolved in the gel shown in panel A was transferred to a nylon membrane (12) and probed with pRQ7 DNA. Lane numbers correspond to those in panel A. Note that under the electrophoretic conditions employed for this gel, linear and one form of uncut pRQ7 migrate to similar positions. The two fragments are distinct in Fig. 1.

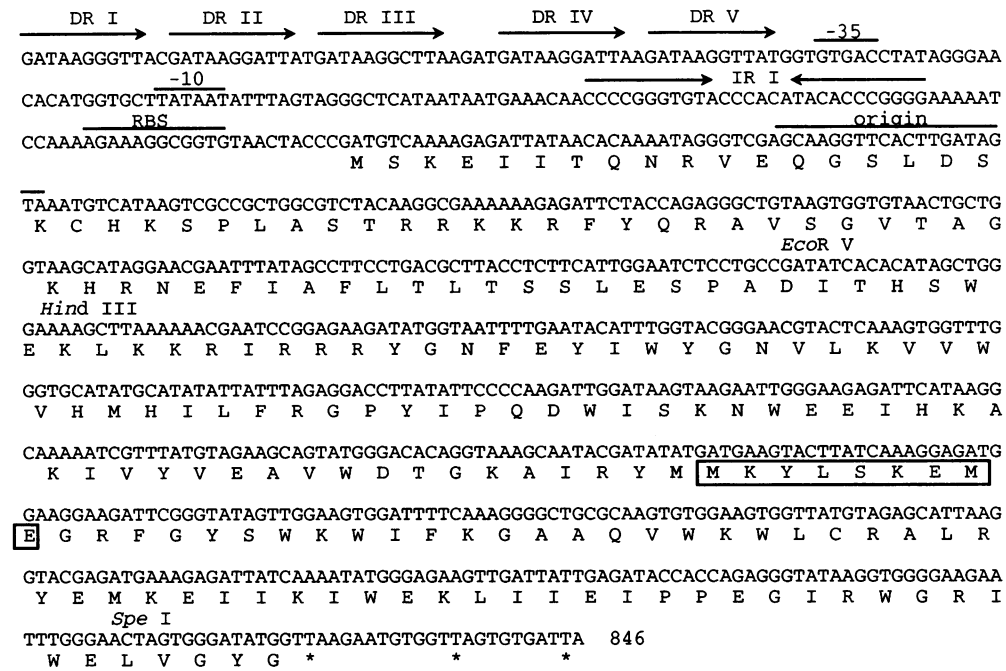

FIG. 3. Nucleotide sequence of pRQ7 and its putative protein. The five direct repeats (DR), promoter (−35 and −10), inverted repeat (IR), ribosome binding site (RBS), and three restriction endonuclease recognition sites are indicated. The sequence of the putative protein is shown beneath the coding region and is followed by three stop codons (*). Also indicated are the regions similar to the plus-strand origin (origin) and Rep protein sequence (boxed sequence) depicted in Table 1.

acids, but it has no identifiable ribosome binding site upstream of the coding region. The single open reading frame encodes a putative 25,460-Da protein with a calculated pI of 9.9. No significant sequence similarity covering large regions was found when the DNA sequence and translated protein sequence were compared with those in a number of data bases (except see below) (5). The coding region is followed by three stop codons. The final codon overlaps the first direct repeat by one nucleotide. No transcriptional terminator was evident in the sequence. Sequences resembling rho-dependent and rho-independent terminators have been found in *Thermotoga* species (6, 10, 14).

Possible modes of replication. Since there is only one protein encoded on pRQ7, and it is basic, we hypothesize that this is a Rep protein which binds to the plasmid and directs its replication. On the basis of sequence similarities between pRQ7 and members of a family of replicons that replicate by a rolling-circle (RC) mechanism, we hypothesize that pRQ7 may replicate by an RC mechanism. RC replicons include the *Escherichia coli* single-stranded DNA bacteriophage ΦX174, many plasmids found in gram-positive bacteria, two plasmids from gram-negative hosts, and a plasmid from the archaeon *Halobacterium* sp. strain GRB (7, 16). Members of the pC194 family of RC replicons share sequence similarities in their origins of replication and in their Rep proteins (7). Similar sequences are found in pRQ7 (Table 1). The putative plus-strand origin shown in Table 1 lies within the open reading frame (Fig. 3). These origins frequently are found within or just upstream of the *rep* genes in RC plasmids. Members of the pT181 family of RC plasmids use an inverted repeat structure as the plus origin (11). Perhaps the inverted repeat in pRQ7 (IR I in Fig. 3) may serve this purpose. In addition, RC plasmids have 130- to 300-bp regions containing imperfect palindromic sequences (7, 17). These palindromes function as minus-strand replication origins. pRQ7 has several imperfect palindromes extending from the ribosome binding site to the *Eco*RV site.

A characteristic feature of RC plasmids is the presence of

TABLE 1. Sequence similarities between pRQ7 and single-stranded DNA replicons[a]

Replicon	Plus-strand replication origin sequence	Rep protein sequence	Organism
pRQ7	gC-aaggTTcaCTTGATA-gTA	MKYLSKEME	*Thermotoga* sp. strain RQ7
ΦX174	gC-TccCccAaCTTGATA-tTA	AKYVNKKSD	*Escherichia coli*
pC194	TC-TTTCTTATCTTGATA-ATA	AKYSGKDSD	*Staphylococcus aureus*
pUB110	TC-TTTCTTATCTTGATACATA	AKYPVKDTD	*Staphylococcus aureus*
pBAA1	TCtTTTCTTATCTTGATAC-TA	SKYPVKDTD	*Bacillus subtilis*
pCB101	TC-TTTCTTATCTTGATA-ATA	FKYMTKVTG	*Clostridium butyricum*
pLP1	TC--TTCTTATCTTGATAC-TA	AKYGVKSAD	*Lactobacillus plantarum*

[a] The nucleotide sequence and putative amino acid sequence from pRQ7 were aligned with those published in Gruss and Ehrlich, 1989. Highly conserved nucletoides are shown in upper case. Amino acids conserved in all proteins are shown in bold face type.

single-stranded replication intermediates in the cells (7). We examined strain RQ7 for a single-stranded form of pRQ7 by performing two Southern hybridizations of pRQ7 DNA which had been resolved by gel electrophoresis. One sample was transferred to the membrane following alkaline denaturation, and the other was transferred under nondenaturing conditions. Both membranes were probed with labelled cloned pRQ7 DNA. The results showed that neither membrane contained a band of pRQ7 DNA which migrated faster than the covalently closed circular form, which is where one would expect to find the single-stranded form (data not shown) (13). Although single-stranded forms of RC plasmids are often revealed by this technique, some, such as that from *Halobacterium* sp. strain GRB (15), are present in very few copies and so may escape detection.

If pRQ7 is a member of this class of replicons, it is the simplest member of the family. Unlike RC plasmids, it encodes no obvious selectable traits. The organization of its putative leading strand replication and control region (16) is unlike that of RC plasmids, to which it seems most related. Indeed, it seems to share features found in different families of RC plasmids, a situation one might expect of a plasmid more like an ancestral RC plasmid. This possibility is consistent with the phylogenetic position of *Thermotoga* species.

Nucleotide sequence accession number. The sequence data in Fig. 3 appear in the EMBL, GenBank, and DDBJ nucleotide sequence libraries under accession number L19928.

We acknowledge the assistance and helpful discussions provided by David Benson and Philip Youderian.

This research was supported by grants from the Office of Naval Research (N00014-91-J-1358), the National Science Foundation (DCB 91-05403), and the University of Connecticut Research Foundation.

REFERENCES

1. **Achenbach-Richter, L., R. Gupta, K. O. Stetter, and C. R. Woese.** 1987. Were the original eubacteria thermophiles? Syst. Appl. Microbiol. **9:**34–39.
2. **Achenbach-Richter, L., R. Gupta, W. Zillig, and C. R. Woese.** 1988. Rooting the archaebacterial tree: the pivotal role of *Thermococcus celer* in archaebacterial evolution. Syst. Appl. Microbiol. **10:**231–240.
3. **Altschul, S. F., W. Gish, W. Miller, E. W. Myers, and D. J. Lipman.** 1990. Basic local alignment search tool. J. Mol. Biol. **215:**403–410.
4. **Burggraf, S., G. J. Olsen, K. O. Stetter, and C. R. Woese.** 1992. A phylogenetic analysis of *Aquifex pyrophilus*. Syst. Appl. Microbiol. **15:**352–356.
5. **Devereux, J., P. Haeberli, and O. Smithies.** 1984. A comprehensive set of sequence analysis programs for the VAX. Nucleic Acids Res. **12:**387–395.
6. **Engel, A. M., Z. Cejka, A. Lupas, F. Lottspeich, and W. Baumeister.** 1992. Isolation and cloning of Ompα, a coiled-coil protein spanning the periplasmic space of the ancestral eubacterium *Thermotoga maritima*. EMBO J. **11:**4369–4378.
7. **Gruss, A., and S. D. Ehrlich.** 1989. The family of highly interrelated single-stranded deoxyribonucleic acid plasmids. Microbiol. Rev. **53:**231–241.

7a. **Huber, R., T. A. Langworthy, H. König, M. Thomm, C. R. Woese, U. B. Sleytr, and K. O. Stetter.** 1986. *Thermotoga maritima* sp. nov. represents a new genus of unique extremely thermophilic eubacteria growing up to 90°C. Arch. Microbiol. **144:**324–333.

7b. **Jannasch, H. W., R. Huber, S. Belkin, and K. O. Stetter.** 1988. *Thermotoga neapolitana* sp. nov. of the extremely thermophilic eubacterial genus *Thermotoga*. Arch. Microbiol. **150:**103–104.

8. **Khan, S. A., and R. P. Novick.** 1982. Structural analysis of plasmid pSN2 in *Staphylococcus aureus*: no involvement in enterotoxin B production. J. Bacteriol. **149:**642–649.
9. **Lee, S.-Y., and S. Rasheed.** 1990. A simple procedure for maximum yield of high-quality plasmid DNA. BioTechniques **9:**676–679.
10. **Liao, D. Q., and P. P. Dennis.** 1992. The organization and expression of essential transcription translation component genes in the extremely thermophilic eubacterium *Thermotoga maritima*. J. Biol. Chem. **267:**22787–22797.
11. **Noirot, P., J. Bargonetti, and R. P. Novick.** 1990. Initiation of rolling-circle replication in pT181 plasmid: initiator protein enhances cruciform extrusion at the origin. Proc. Natl. Acad. Sci. USA **87:**8560–8564.
12. **Nordström, K.** 1985. Control of plasmid replication: theoretical considerations and practical solutions, p. 189–214. *In* D. R. Helinski, S. N. Cohen, D. B. Colwell, D. A. Jackson, and A. Hollaender (ed.), Plasmids in bacteria. Plenum Press, New York.
13. **Riele, H. T., B. Michel, and S. D. Ehrlich.** 1986. Single-stranded plasmid DNA in *Bacillus subtilis* and *Staphylococcus aureus*. Proc. Natl. Acad. Sci. USA **83:**2541–2545.
14. **Sanangelantoni, A. M., G. Forlani, F. Ambroselli, P. Cammarano, and O. Tiboni.** 1992. The *glnA* gene of the extremely thermophilic eubacterium *Thermotoga maritima*—cloning, primary structure, and expression in *Escherichia coli*. J. Gen. Microbiol. **138:**383–393.
15. **Sioud, M., G. Baldacci, P. Forterre, and A.-M. de Recondo.** 1988. Novobiocin induces accumulation of a single strand of plasmid pGRB-1 in the archaebacterium *Halobium* GRB. Nucleic Acids Res. **16:**7833–7842.
16. **Solar, G. D., M. Moscoso, and M. Espinosa.** 1993. Rolling circle-replicating plasmids from Gram-positive and Gram-negative bacteria: a wall falls. Mol. Microbiol. **8:**789–796.
17. **Vujcic, M., and L. Topisirovic.** 1993. Molecular analysis of the rolling-circle replicating plasmid pA1 of *Lactobacillus plantarum* A112. Appl. Environ. Microbiol. **59:**274–280.