Extremophiles (2004) 8:91–98
DOI 10.1007/s00792-003-0366-z

ORIGINAL PAPER

Meixian Zhou · Hua Xiang · Chaomin Sun · Yun Li
Jingfang Liu · Huarong Tan

Complete sequence and molecular characterization of pNB101, a rolling-circle replicating plasmid from the haloalkaliphilic archaeon *Natronobacterium* sp. strain AS7091

Received: 27 June 2003 / Accepted: 20 October 2003 / Published online: 13 December 2003

Abstract A new plasmid was isolated from the haloalkaliphilic archaeon, *Natronobacterium* sp. strain AS7091 and named pNB101. Sequence analysis revealed that pNB101 consists of 2,538 bp, in which three major open reading frames (ORF1, ORF2, and ORF3) were identified in the same strand. The ORF1 encodes a putative replication (Rep) protein with three typical motifs (I, II, and III) found in rolling-circle (RC) replicating plasmids. The putative double-stranded origin (DSO) and single-stranded origin (SSO) were detected within ORF3 and downstream of ORF1, respectively. S1 nuclease digestion and Southern blot analysis demonstrated the existence of the single-stranded DNA (ssDNA) intermediate from pNB101, which corresponds to the leading strand in RC replication and was further confirmed by strand-specific RNA probes. A single transcript for ORF1 (*rep*) was detected by Northern blotting, and the 5′ end of this transcript was determined by primer extension. Both results indicate that the three motifs (I–III) are located at the very end of the N-terminal of this Rep protein. Northern blot analysis also revealed that the ORF3 was transcribed at a very high level, which may play an important role in plasmid replication because the putative DSO is located in this gene. Together, our results indicate that pNB101, the first plasmid isolated from haloalkaliphilic *Archaea*, represents a novel RC-replicating plasmid.

Keywords *Natronobacterium* · Plasmid · Rolling-circle replication

Communicated by K. Horikoshi

M. Zhou · H. Xiang (✉) · C. Sun · Y. Li · J. Liu · H. Tan
State Key Laboratory of Microbial Resources and Center for Molecular Microbiology, Institute of Microbiology, Chinese Academy of Sciences, 100080 Beijing, People's Republic of China
E-mail: xiangh@sun.im.ac.cn
Tel.: +86-10-62656916
Fax: +86-10-62656916

Introduction

The extremely halophilic *Archaea*, *Halobacteriaceae*, a group of microorganisms with unusual genetic, biochemical, and physiological characteristics, and unique evolutionary history, are becoming an increasingly attractive subject of study for microbiologists (Cowan 1992; Litchfield 1998; Russell 1989; Wise et al. 2002). So far, at least 14 different genera have been identified in this group (Ochsenreiter et al. 2002), and several multicopy plasmids in two genera, *Haloferax* and *Halobacterium*, have been described. The plasmid pGRB1 (1.8 kb) from *Halobacterium* sp. strain GRB and its relative plasmids (pHGN1, pHSB1, and pHSB2) from other *Halobacterium* species (Hackett et al. 1990; Hall and Hackett 1989; Kagramanova et al. 1988), as well as pHV2 (6.4 kb) from *Haloferax volcanii* (Charlebois et al. 1987) have been completely sequenced. The minimal replication region of the plasmid pHK2 (10.5 kb) from *Haloferax* sp. strain Aa2.2 has also been defined and sequenced (Holmes et al. 1995). All the above plasmids except for pHV2 were believed to replicate via a RC mechanism, because they encode homologous Rep proteins containing conserved motifs typically found in rolling-circle replicating (RCR) plasmids and phages (Holmes et al. 1995). Some experimental support for this argument is provided by the fact that the ssDNA form of the plasmid pGRB1, which is the hallmark of RCR plasmids, was detected in *Halobacterium* sp. strain GRB following novobiocin induction (Sioud et al. 1988). Regardless of their modes of replication, some of these plasmids including pHV2 have been successfully used for the construction of shuttle vectors (Holmes et al. 1994; Lam and Doolittle 1989; Mankin et al. 1992), which have contributed to both genetic studies and biomolecular engineering applications in extremely halophilic *Archaea*.

In the remaining 12 genera, very few plasmids have been isolated. Moreover, no plasmid has yet been described in the subgroup of haloalkaliphilic *Archaea*.

Within this subgroup, the genus *Natronobacterium* differs from most other halophiles in several respects. They are capable of tolerating both hypersaline and alkaline environments. They grow at a high pH level (between pH 8.5 and 11, where the optimum is at 9.5) and a high NaCl concentration, but a very low (less than 10 mM) Mg^{2+} concentration (Tindall et al. 1980). They also contain unique diether lipids (Upasani et al. 1994).

In recent years, several specific genes and proteins have been isolated from *Natronobacteria* (Mayr and Pfeifer 1997; Polosina et al. 1998), and the sensory rhodopsin II from a haloalkaliphilic archaeon, *Natronobacterium pharaonis*, has been studied in some detail (Hein et al. 2003; Schmies et al. 2000). However, genetic study and molecular engineering of *Natronobacteria* remain virtually unexplored fields, which is mainly due to no suitable genetic elements (e.g. plasmids) having been isolated and no genetic tools developed in these special organisms. In order to better understand the organisms in hypersaline and alkaline environments, we have initiated the screening of the native plasmids and the establishment of genetic systems in the haloalkaliphilic *Archaea*, and a multicopy plasmid pNB101 (2.5 kb) was isolated from *Natronobacterium* sp. strain AS7091 out of 29 different haloalkaliphilic archaeal strains. In this paper, we report the complete nucleotide sequence and the replication model of this plasmid. Our results suggest that pNB101, the first plasmid isolated from the haloalkaliphilic *Archaea*, is a novel RCR plasmid.

Materials and methods

Strains and media

The haloalkaliphilic archaeal strains used for plasmid screening were from *Natronobacterium*, *Natronococcus*, and *Natrialba*, and were supplied by the Type Culture Collection of the Chinese Academy of Sciences (Beijing, P.R. China). These strains, including *Natronobacterium* sp. strain AS7091, which harbors the plasmid pNB101, were grown in rich medium containing (per liter) 200 g of NaCl, 7.5 g of casamino acids (Difco), 10 g of yeast extract (Oxoid), 3 g of trisodium citrate dehydrate, 0.1 g of $MgSO_4{\cdot}7H_2O$, 2 g of KCl, 36 mg of $FeCl_2{\cdot}4H_2O$, 0.36 mg of $MnCl_2{\cdot}4H_2O$, and 8 g of Na_2CO_3, pH 9.0–9.5. Cultures were grown in liquid media at 37°C with constant shaking (200 rpm), or on 1.5% agar plates prepared with the same media in plastic containers.

Plasmid isolation and DNA sequencing analysis

The plasmid pNB101 was isolated from *Natronobacterium* sp. strain AS7091 by the alkaline lysis method (Ng et al. 1995). For DNA sequencing, the pNB101 DNA was digested at either the *Hin*d III or the *Sac*I site to generate two kinds of linear plasmid DNA molecules, which were inserted into pBluescript KS- (Stratagene) at the corresponding sites to yield two different recombinant clones. The pNB101 sequences in both clones were determined by primer walking with the dideoxynucleotide termination cycle sequencing method (Strauss et al. 1986) in an ABI automated DNA sequencer as described previously (Liu et al. 1998). Sequence analysis was performed with a DNAStar sequence analysis software package (DNAStar, Madison, Wis., USA). Nucleotide and amino-acid sequence homologies were searched with the NCBI BLAST programs.

The nucleotide sequence of pNB101 described in this study was deposited in GenBank under the accession number AF474073.

Detection of ssDNA form of pNB101

Total DNAs from mid-log-phase cells of *Natronobacterium* sp. strain AS7091 were analyzed for the presence of ssDNA intermediate generated from pNB101. For DNA preparation, the cultures (1.5 ml) were collected and resuspended in 0.5 ml of lysis buffer (25 mM Tris-Cl, 25 mM EDTA, 0.3 M sucrose, 1 mg/ml lysozyme, pH 8.0). After a 20-min incubation at 37°C, 0.25 ml of 3% SDS was added and the mixture was incubated for additional 2 h. The mixture was then extracted with neutral phenol/chloroform and the DNA in the supernatant was precipitated with isopropanol. The DNA pellet was air-dried and dissolved in 0.5 ml of TE (10 mM Tris-Cl, 1 mM EDTA, pH 8.0) containing 50 μg/ml RNase A, and incubated at 37°C for 30 min. Afterwards, the DNA sample was treated with proteinase K (150 μg/ml in 0.5% SDS) at 56°C for 1 h, and was extracted again with phenol/chloroform and precipitated with ethanol. After these procedures, the extracted DNA was ready for further analysis.

To detect the ssDNA intermediate, Southern blot analysis was conducted. Briefly, 5 μg of the total DNAs was treated with or without 5 U of S1 nuclease. These DNA samples were resolved by agarose gel electrophoresis, and then transferred to a nylon membrane with or without prior denaturation. After prehybridization, the blot was hybridized with double-stranded DNA probes, which were labeled through random priming of the whole pNB101 plasmid with digoxigenin (DIG)-dUTP (Roche) according to the manufacturer's instructions.

Identification of the leading strand of pNB101 in RC replication

The total DNA samples from *Natronobacterium* sp. strain AS7091 were subjected to agarose gel electrophoresis and transferred to a nylon membrane. The blot was hybridized with strand-specific RNA probes, which were labeled with [α-^{32}P]rUTP in vitro by T7 RNA polymerase (Sambrook et al. 1989), using two plasmids containing the entire pNB101 sequence in different orientation in pBluescript KS- as the templates.

RNA isolation and Northern blot analysis

Total RNAs from exponentially grown cells of *Natronobacterium* sp. strain AS7091 were prepared with TRIzol reagent (Gibco BRL) according to the manufacturer's instructions. The RNA pellets were dissolved in DEPC-treated water, and 20 μg of total RNAs was used for Northern blot analysis. For probe preparation, the DNA sequence for each probe was first amplified by polymerase chain reaction (PCR). The following primers were used: 5′-GGCGAGGAACTTGTGGCAG-3′ (forward) and 5′-CGATGCTGGGACCGCACA-3′ (reverse) for probe 1 (P1, nt 1961–2232 in ORF3); 5′-GTCCAGACGCGAGAGAGG-3′ (forward) and 5′-TGTTCTCAAGCGGTGTGC-3′ (reverse) for probe 2 (P2, nt 11–328 in ORF1); 5′-CGCGAGAAAGGCGCTACAC-3′ (forward) and 5′-TCACGCCACTGGTCAAACG-3′ (reverse) for probe 3 (P3, nt 513–992 in ORF1). The PCR products were then purified and labeled with [α-^{32}P]dCTP through five additional PCR cycles. The other procedures, such as gel electrophoresis, prehybridization, hybridization, washing, and exposure, were conducted as described previously (Xiang et al. 2000).

Determination of the transcriptional start site of ORF1

Two primers, 5′-GACGGAGATCGGGCGGAAC-3′ (nt 84–102) and 5′-CTCCCAGACCTTCGCGTACT-3′ (nt 557–576), which were complementary to the sequence downstream of the ATG-1

and ATG-421 of ORF1, respectively, were labeled at the 5′-end with [γ-^{32}P]ATP, and used for DNA sequencing and primer extension. For primer extension analysis, 30 μg of total RNAs and 30 pmol of [γ-^{32}P]ATP-labeled primer were mixed in a volume of 12 μl, denatured at 70°C for 10 min, and then quickly chilled on ice. Afterwards, 4 μl of 5×first strand buffer (Gibco BRL), 2 μl of 0.1 M DTT, 1 μl of 10 mM dNTP and 1 μl of Superscript RT II (200U, Gibco BRL) were added to the RNA/primer mixture and incubated at 42°C for 1 h. The reaction was terminated by incubation at 70°C for 15 min. Then, 50 μl RNase reaction mix (100 μg/ml salmon sperm DNA, 20 μg/ml RNase A) was added, and the mixture was incubated at 37°C for 15 min. The primer extension products were extracted with phenol/chloroform and precipitated with ice-cold ethanol. At the same time, DNA-sequencing reactions were performed with the same primers by using the fmol DNA cycle sequencing system (Promega) according to the protocols of the manufacturer. The products of primer extension and DNA-sequencing reaction were respectively resuspended in the same sequencing stop solution and heated for 2 min at 70°C, then analyzed on 6% acrylamide sequencing gel. Radioactive DNA fragments on the gels were visualized on X-ray film.

Results

Isolation and sequence characterization of pNB101

After screening the native plasmids in 29 different haloalkaliphilic archaeal strains from three genera, *Natronobacterium*, *Natronococcus*, and *Natrialba*, one multicopy plasmid was isolated from the *Natronobacterium* sp. strain AS7091 and named pNB101. Subsequent restriction enzyme mapping of pNB101 revealed the presence of two unique restriction enzyme sites, *Sac*I and *Hin*d III, in this plasmid. These two sites were then utilized to clone the entire pNB101 sequence into an *E. coli* cloning vector, pBluescript KS-, to generate two different plasmids, pNB1 and pNB2. Both plasmids were then used for DNA sequencing analysis and construction of a detailed map of pNB101 (see Fig. 1).

DNA sequence analysis showed that pNB101 is a circular molecule of 2,538 bp with a G + C content of 63.7%, which is similar to the high G + C content of the genome or plasmids of other haloarchaea (Charlebois et al. 1987; Hackett et al. 1990; Hall and Hackett 1989; Holmes et al. 1995). Examination of the six possible phases of the plasmid sequence revealed the presence of three major ORFs (ORF1, ORF2, and ORF3) in the same strand, which cover about 80% of the entire sequence (Fig. 1). These ORFs were supposed to encode three polypeptides with 388, 95, and 181 amino acids. The first base in the putative translation initiation codon (ATG) of ORF1 was arbitrarily defined as the position 1 of this plasmid. While no significant homologies were found in the NCBI protein databases for the deduced proteins encoded by ORF2 and ORF3, the deduced protein encoded by ORF1 shared 31%–37% identity (45%–54% similarity) with the Rep proteins of the small plasmids (pHSB1, pGRB1, pHSB2, and pHGN1) from *Halobacterium* strains, and pHK2 from *Haloferax* sp. Aa2.2, suggesting that these plasmids isolated from different members of halophilic *Archaea* may belong to the same family. Because no significant

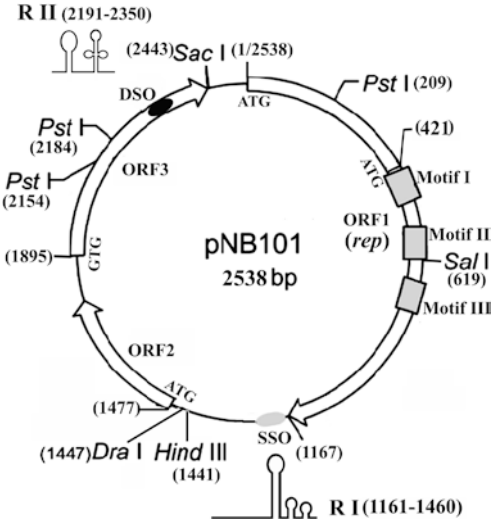

Fig. 1 Map of pNB101. The three major open reading frames, namely ORF1 (*rep*), ORF2, and ORF3, are indicated with *open arrows*. The three motifs (Motifs I–III) common to Rep proteins of RC plasmids are indicated by *rectangles*. The putative DSO and SSO are indicated by *black and gray ellipses*, respectively. The Region I (*RI*) and II (*RII*), and the restriction sites for appropriate enzymes are also indicated. The *numbers in parentheses* indicate positions in base pairs from the *1* point at the top of the map, arbitrarily defined by the first ATG codon of the largest open reading frame (ORF1)

promoter sequence can be detected upstream of ATG-1 of the ORF1, it was deduced that another possible start codon (ATG-421) could be used by the plasmid at the further downstream site (Fig. 1).

As well as the above three open reading frames, two regions with specific structural features were detected. Region I (RI, nt 1161–1460) is located between the ORF1 and ORF2 (Fig. 1). A putative hairpin structure with a GC-rich stem produced by the inverted repeats was detected immediately downstream of the ORF1. Following this hairpin structure there is a 200-bp-long AT-rich (44.8% in this region vs 35.4% for the rest) section. Three direct repeats, GTACAG, CGGCCG, and GGTTTC, were repeated three times, respectively, in this section. While the repeat GGTTTC was mainly found in this section, the other two direct repeats were dispersed in the whole plasmid with about ten occurrences. Another special region, named Region II (RII, nt 2191–2350), is mainly located within the ORF3 (Fig. 1). Region II also contained a significant amount of direct or inverted repeats.

Comparative analysis of the pNB101 structure with RCR plasmids

Three general replication models: theta, strand displacement, and RC replication, have been described for circular plasmids (del Solar et al. 1998). RCR plasmids are widespread in *Bacteria*, and have also recently been found in species of *Archaea*. Two plasmids, pGT5 and pRT1 (Erauso et al. 1996; Ward et al. 2002), isolated

from the hyperthermophilic archaeon *Pyrococcus*, were identified to replicate via the RCR model. Their ssDNA intermediates in RC replication were determined by Southern blot analysis (Erauso et al. 1996; Ward et al. 2002). The plasmid pGRB1, pHSB1, pHSB2, and pHGN1 isolated from *Halobacterium*, as well as pHK2 from *Haloferax*, were also believed to replicate via a RCR mechanism because their putative Rep proteins contain the three motifs (I, II, and III) conserved in RC-replicating plasmids or phages (Holmes et al. 1995). These conserved motifs were also detected in the putative Rep protein from the ORF1 of pNB101 (Figs. 1 and 2).

The Rep proteins have a type I-topoisomerase-like activity to cleave and join plasmid DNA, which are involved in initiation and termination of leading strand synthesis in RC replication. The motif III of the Rep protein is the catalytic motif. It contains the conserved tyrosines (Y) involved in the nucleophilic attack of the plasmid DNA, and thereby introduces a site-specific nick on the plus strand at the DSO region, and leaves a 3′-OH end as a primer for leading-strand synthesis (Khan 2000). RC replicons can be divided into two superfamilies according to the sequence of motif III (Koonin and Ilyina 1993). Superfamily I, which includes coliphage Φ-X174 and several bacterial plasmids, such as pRQ7 from *Thermotoga* sp. strain RQ7 (Harriott et al. 1994), is characterized by two conserved tyrosines in this motif. The pNB101 and the known RCR plasmids isolated from extremely halophilic *Archaea* all have two tyrosines in this motif (Fig. 2), suggesting that they belong to this superfamily. Other RCR plasmids such as bacterial plasmid pC194 (Seery et al. 1993) and archaeal plasmid pGT5 (Erauso et al. 1996), which harbor a single tyrosine in motif III (Fig. 2), can be grouped to the superfamily II. This unique feature in motif III seems to be correlated with the active forms of Rep protein (dimeric or monomeric) during the nicking–closing process of leading-strand synthesis (Khan 2000). Thus, pNB101 and pGT5 may represent different types of the archaeal RCR plasmids. Another important motif for Rep protein is the HUH motif (motif II), which usually has two conserved histidines, and is thought to be the metal-binding domain. This typical motif was also detected in pNB101 (Fig. 2). In contrast to the above two motifs, motif I is less conserved among the RCR plasmids, and its function is yet to be explored. Nevertheless, as with motifs II and III, the pNB101 motif I was also highly conserved with the corresponding motif found in other halophilic archaeal RCR plasmids (Fig. 2), further suggesting that these plasmids are closely related.

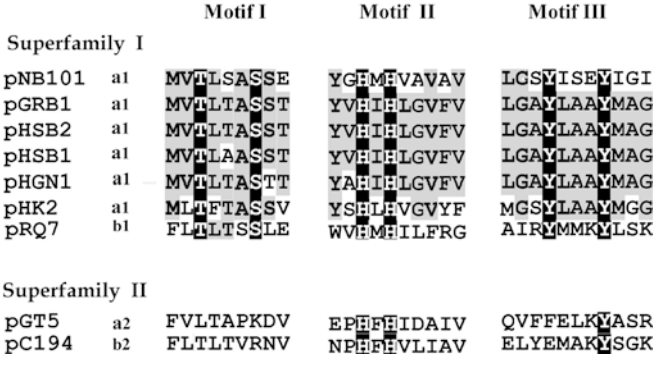

Fig. 2 Identification of the three motifs universally conserved in RC Rep proteins within the amino-acid sequence of pNB101 Rep protein. These motifs are aligned with the homologous regions of the Rep proteins from several RCR plasmids isolated from halophilic *Archaea* (*a1*), *Bacteria* (*b1 or b2*) or thermophilic *Archaea* (*a2*). Residues are *boxed* where most sequences have identical amino acids

Within the nucleotide sequence of pNB101, a putative double-stranded origin (DSO) of replication was identified in Region II at the 3′-end of ORF3 (Fig. 1). The DSO in RCR plasmids is a specific sequence in which Rep protein introduces the initial nick. It has two loci, namely the *bind* and *nick* sites. A typical feature of RCR plasmids is that the *nick* sites are highly conserved among replicons of the same family, whereas the *bind* sites are usually different (Khan 2000). Although no *nick* sites in the halophilic archaeal RCR plasmids have been experimentally identified yet, it has already been inferred in pGRB1 based on sequence comparison (Erauso et al. 1996). The *nick* site of the putative pNB101 DSO showed strong similarity to those found in pGRB1 and its relatives. The GA dinucleotide at the putative *nick* site is totally conserved, and nucleotides surrounding this dinucleotide are also highly conserved in pNB101 (Fig. 3a). In addition, as in other DSO, the putative DSO in pNB101 contained several direct repeats (DR I–IV) and inverted repeats (IR I–III) contiguous to or separated from the *nick* site (Fig. 3b). These repeats presumably serve as the Rep protein-binding site.

The single-stranded origin (SSO) of replication is usually a specific non-coding region with strong secondary structure. Two of the conserved sequences, RNA polymerase binding site (RS_B) and 6-nt sequence (CS-6), have been shown to be critical for binding of RNA polymerase and termination of RNA primer synthesis in SSO*A*-type origins (Khan 2000; Kramer et al. 1997). The CS-6 motif for the SSO*A*-type (pMV158) is TAGCGt/a (del Solar et al. 1998); while for the archaeal plasmid pGT5 it is likely TAGTGG (Erauso et al. 1996). The putative SSO of pNB101 with these specific characters was identified within the Region I immediately downstream of the ORF1 (Fig. 1). This putative SSO in pNB101 generates the most stable stem–loop structure in this plasmid (Fig. 4b). Within the loop region, it also contains a putative CS-6 sequence, TAGGGG, which shares four or five out of six nucleotides with that of pMV158 or pGT5, respectively. The sequences surrounding the CS-6 were also highly conserved in both pNB101 and pGT5 (Fig. 4a).

In summary, our comparative analysis revealed the presence of the putative Rep gene (*rep*), DSO, and SSO in pNB101, suggesting that this plasmid is likely to replicate through a RC mechanism. Since the DSO, SSO, and *rep* gene are localized on the same strand of pNB101, we defined this strand as the plus strand.

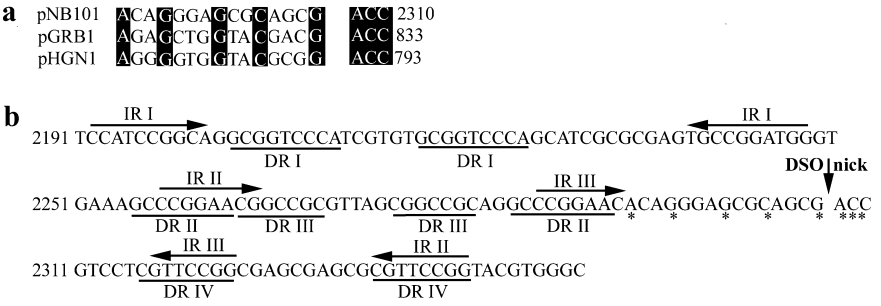

Fig. 3a, b Identification of the putative pNB101 DSO. **a** Alignment of the sequence of the putative DSO of pNB101 with those of two different halophilic archaeal RCR plasmids, pGRB1 and pHGN1. Identical nucleotides in all sequences are *boxed*. A *space* indicates the putative cleavage site. **b** Sequence of the pNB101 DSO region. The inverted repeats (*IR*) are indicated with *arrows above the sequences*, and the direct repeats (*DR*) are *underlined*. The conserved nucleotides near the putative DSO cleavage site (DSO *nick*) are indicated by *asterisks*

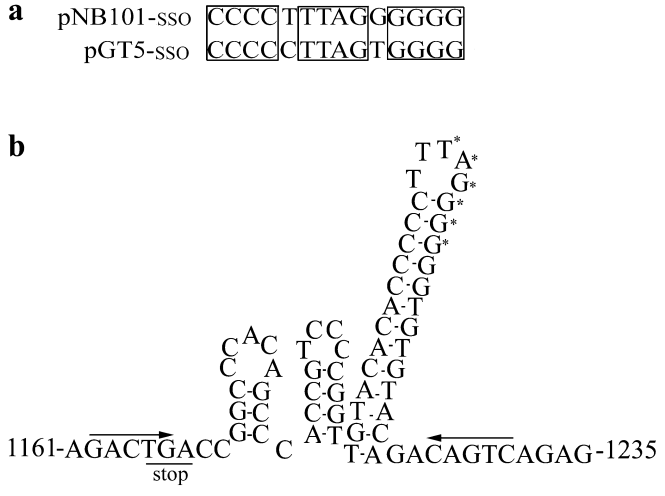

Fig. 4a, b Identification of the putative SSO of pNB101. **a** Alignment of the central sequence of the putative SSO of pNB101 with that of the archaeal RCR plasmid pGT5. The identical nucleotides in both sequences are *boxed*. **b** Secondary structure of putative SSO of pNB101. The hexanucleotide sequence TAGGGG, which is similar to the conserved six-nucleotide sequence (*CS-6*, see text) in SSO*A* (del Solar et al. 1998), is indicated with *asterisks*. The inverted sequences separated by the stem–loop structures are indicated with *arrows*. The stop codon of *rep* at the upstream of the SSO is also indicated

Identification of the pNB101 ssDNA intermediate in RC replication

To determine whether pNB101 actually replicates through the RC mechanism, the possible existence of the ssDNA intermediate of pNB101 in RC replication was examined with Southern blot analysis. Total DNA extracted from *Natronobacterium* sp. strain AS7091 with or without S1 nuclease treatment was resolved in agarose gel electrophoresis, and was transferred to a nylon membrane with or without prior denaturation. After hybridization, one band migrating faster than the supercoiled double-stranded DNA band was detected with the pNB101 probes under denaturing conditions (left lane of Fig. 5a). This band disappeared after S1 nuclease treatment (right lane of Fig. 5a). Its sensitivity to S1 nuclease digestion indicates that this band should be a single-stranded DNA. Moreover, only this single stranded form of the plasmid was efficiently hybridized to the probes under non-denaturing conditions, and addition of S1 nuclease again resulted in disappearance of this band (Fig. 5b), which further confirms that it is the ssDNA intermediate of pNB101 in replication, and thus pNB101 indeed replicates via a RC mechanism.

Determination of the leading strand of pNB101 in RC replication

The ssDNA intermediate of RCR plasmid corresponds to the leading strand of this plasmid in RC replication, which is cleaved by Rep protein and displaced by the DNA polymerase. To identify which strand is the leading strand, we hybridized the ssDNA intermediate with two RNA probes corresponding to either strand of pNB101. The RNA probe corresponding to the minus strand was able to hybridize with the ssDNA intermediate (right lane of Fig. 5c). In contrast, hybridization with the RNA probe equivalent to the plus strand could not detect the same band (left lane of Fig. 5c). These results indicate that the ssDNA intermediate (leading strand) is generated from the plus strand. As in most other RCR plasmids (Khan 1997), our results also showed that the *rep* gene of pNB101 is on the leading strand, and the transcription of the *rep* gene proceeds in the same direction as the leading-strand synthesis does.

Determination of the transcription pattern of pNB101

Although pNB101 was identified as a RCR plasmid, its organization differs from the other RCR plasmids from halophiles. For example, pGRB1 (1.8 kb) contains only one major open reading frame (*rep*) and the putative DSO is within the *rep* gene. In contrast, pNB101 contains three major open reading frames, and the DSO is located in ORF3, which is separated from the putative *rep* gene (ORF1). To understand the replication

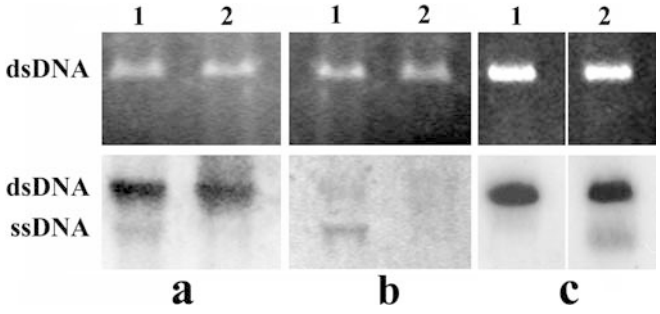

Fig. 5a–c Detection of the ssDNA intermediate in the replication of pNB101. The top panel of each figure (**a**–**c**) is ethidium bromide staining of the DNA gel to show the equal loading of two samples, while the bottom panel of each figure is the Southern blots after hybridization. **a**, **b** Detection of the pNB101 ssDNA intermediate with pNB101 dsDNA probes. Purified total DNAs from *Natronobacterium* sp. strain AS7091 were resolved on 0.8% (w/v) agarose gels with (*lane 2*) or without (*lane 1*) prior to S1 nuclease treatment. The DNA was denatured (**a**) or not denatured (**b**) prior to transfer to a nylon membrane. **c** Identification of the pNB101 leading strand in RC replication with strand-specific RNA probes corresponding to the plus strand (*lane 1*) or minus strand (*lane 2*)

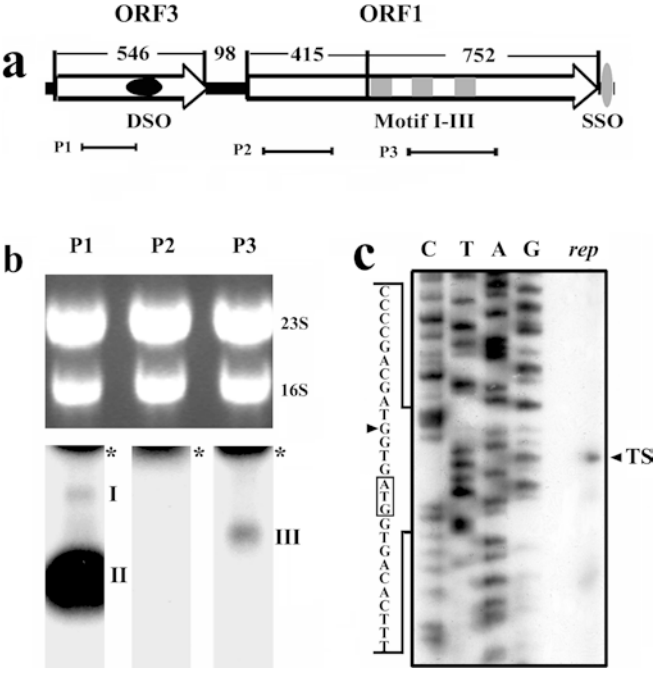

Fig. 6a–c Determination of the transcription pattern of ORF1 and ORF3. **a** Genetic organization of the replicon of pNB101 including ORF1 (*rep*), ORF3 (DSO), and SSO. The *numbers* indicate the size (in bp) of each DNA fragment. The positions of the probes (P1, P2, and P3) for Northern blot analysis (**b**) were shown. **b** Detection of the transcripts of ORF1 and ORF3 by Northern blot analysis. *Top panel* ethidium bromide staining of the RNA gel to show the equal loading of the three samples. The positions of 23S and 16S RNA were indicated. *Bottom panel* Northern blotting detection of the transcripts of ORF3 (*band II in the left lane*) and ORF1 (*band III in the right lane*) by probe P1 and P3, respectively. Note that probe P2 did not detect any transcript (*middle lane*) while P1 also detected a larger transcript (*band I in the left lane*). The contaminated DNA of the plasmid pNB101 detected by all the three probes at the same position is indicated with *asterisks*. **c** Determination of the transcriptional start sites of the *rep* gene by primer extension analysis. The relevant sequence is shown *on the left*. The transcriptional start point (*TS*) is indicated with a *triangle*, while the putative translation start codon (ATG-421) is *boxed*

mechanism of pNB101, we have analyzed the transcription of these two main open reading frames (ORF1 and ORF3) using Northern blot analysis.

Total RNAs were extracted from 10 ml liquid cultures of exponentially grown *Natronobacterium* sp. strain AS7091. A sample of 20 μg of total RNAs was used for each lane in the RNA blot, and was hybridized with the specific probes for ORF1 and ORF3 (Fig. 6a). The ORF3 that harbors the DSO could independently transcribe at a very high level, resulting in a transcript of about 0.5 kb (band II in the left lane of Fig. 6b), which was detected by the ORF3 probe P1. In addition, a larger transcript (about 1 kb) that may be generated from the co-transcription with ORF2 was also detected with the same probe (band I in the left lane of Fig. 6b). This co-transcription had been confirmed with an OFR2-specific probe (data not shown). When the DNA probe P2 which represents the coding sequence for the first ~300 bp of ORF1 (*rep* gene) was used for hybridization, however, no signal was detected (middle lane of Fig. 6b). It suggests that the transcriptional start site of *rep* may be located in the downstream sequence. Indeed, the downstream probe P3 detected a single transcript with a calculated size of about 0.75 kb (band III in the right lane of Fig. 6b). In all the three blots, the contaminated DNA of pNB101 was detected at the same position by all the three probes (Fig. 6b), which was confirmed by loading the plasmid DNA in separate lanes and recognized by all the three probes (data not shown), thus it can be used as a size reference and a positive hybridization control of each probes. Together, our results indicated that the three ORFs of this plasmid were actively expressed in replication of this plasmid. Moreover, unlike pGRB1 and its relatives, the DSO and *rep* gene of pNB101 were transcribed separately.

Determination of the transcriptional start site of the *rep* gene

To understand the results from Northern blot analysis of the *rep* gene, the transcription initiation site was determined by primer extension experiments. For this purpose, we have designed two primers (nt 84–102 and nt 557–576) located downstream of the two putative translational start codons, ATG-1 and ATG-421, respectively. In agreement with the Northern blot results, no extension product was detected with the first primer (data not shown). In contrast, the primer extension product with the second primer allowed us to assign the *rep* transcriptional start site at G-417 (Fig. 6c), one or four bases upstream of the putative start codon GTG-418 or ATG-421, respectively. This confirms that the transcription of the *rep* gene is initiated in the middle of the ORF1, and the transcript of this *rep* also lacks a 5′-untranslated region as do many other archaeal

mRNAs. Considering that the ATG-421 encode the first amino acid (M) of the Rep motif I (MVTLSASSE), our results also indicate that the Rep motifs of this protein are located at the very end of the N-terminal. It will be intriguing to study the structure and function of this novel type of Rep protein. A putative promoter sequence ATAGAAC (nt 388–394) was also detected at a distance of 23 bp upstream from G-417, which further confirmed that G-417 represents a truly transcriptional start site. This result is consistent with the 0.75 kb *rep* transcript (nt 417–1167) revealed by Northern blot analysis (Fig. 6b).

Discussion

Several RCR plasmids have been reported in halophilic *Archaea* (Hackett et al. 1990; Hall and Hackett 1989; Holmes et al. 1995; Kagramanova et al. 1988), and some of these have been used for the development of vectors (Holmes et al. 1994). However, these plasmids were mainly isolated from *Haloferax* and *Halobacterium*, and their genetic features remained to be further investigated. In this paper, we report the DNA sequence and molecular characterization of the first plasmid isolated from the haloalkaliphilic *Archaea*, the pNB101 from *Natronobacterium* sp. strain AS7901. Several lines of evidence indicate that pNB101 may replicate through the RC replication mechanism. First, a homologous RC replication protein with conserved motifs was found in this plasmid. Second, the putative DSO with conserved nick sequence and appropriate second structure was identified. Third, the putative SSO with typical stem–loop structure and CS-6-like sequence was revealed. Finally, the ssDNA intermediate of pNB101 in RC replication was demonstrated by S1 nuclease digestion, and the leading strand was verified by Southern blot analysis with strand-specific RNA probes. Comparative analysis of the DSO and Rep features in the RCR plasmids from halophiles strongly suggests that the pNB101 from *Natronobacterium*, the pGRB1 and its relatives from *Halobacterium*, and the pHK2 from *Haloferax* are closely related. However, pNB101 differs from those plasmids in gene organization and possibly in transcription pattern and may represent a new and unique model of replication control.

In most RCR plasmids, the plasmid genes are located in the same strand. Transcription of the *rep* gene and replication of the leading-strand usually proceed at the same direction (Khan 1997). These features are shared by pNB101. Two small putative genes, ORF2 and ORF3, are located on the plus strand, as is the *rep* gene. These genes are transcribed at the same direction as the leading strand replication proceeds.

The *rep* gene of pNB101 seems to encode a novel Rep protein because the three RC Rep motifs are located at the very end of the N-terminal, which had never been reported before. These motifs are highly conserved within the RCR plasmids isolated from the halophiles and presumably contribute to nicking–closing activities in their conserved DSO *nick* sites. Compared with the Rep proteins from other RCR plasmids, the pNB101 Rep protein is relatively small (about 28 kDa), and should be a good candidate for studying the replication mechanism of the RCR plasmid in halophilic *Archaea*.

Besides the *rep* gene, the functions of ORF2 and ORF3 are currently unknown. However, the ORF3 seems to be involved in the initiation of the plasmid replication too, because the DSO is located within ORF3 near the 3′-end, and Northern blot analysis revealed that this gene transcribed at a very high level in replication of the plasmid. It would be interesting to investigate this high-level transcription besides investigation of the protein function. A plausible mechanism for high-level transcription around DSO is that it may relax the supercoiled DNA, and help to generate the cruciform structure at DSO to expose the *nick* site for Rep protein. Similar situations were found in the chromosome replication, such as in lambda phage as well as in *E. coli*, where transcription of the origin, named as transcriptional activation of replication, is the crucial step for initiation of the replication of the chromosome DNA (Baker and Kornberg 1988; Taylor and Wegrzyn 1995). This mechanism may be utilized by RCR plasmids because most DSOs are located very near the transcribed *rep* gene or even within the *rep* gene. The sequences around the DSOs are certainly transcribed in those RCR plasmids. Although the DSO of pNB101 is relatively far from its *rep* gene, our results indicated that they were both efficiently transcribed. However, whether transcription of or around DSO is a common theme for RCR plasmid replication initiation remains to be investigated, and pNB101 may be a good model for this study because transcriptions of the DSO and the *rep* gene in this plasmid were respectively controlled.

The ultimate goal of our study is the development of a genetic tool for the *Natronobacteria*. The plasmid pNB101 could serve as an excellent candidate for this purpose. The unique *Hin*d III site may be ideal for the insertion of an antibiotic resistant gene and other necessary DNA fragments, as it is distant from the SSO, DSO, and the *rep* gene. However, it may affect the expression of the uncharacterized ORF2. Another possible site for insertion of the foreign DNA is the unique *Sac*I site, which does not interrupt any of the ORFs and has a necessary distance from SSO, DSO, and the *rep* gene. Construction of a vector based on this plasmid and subsequent introduction of this construct into the *Natronobacteria* will provide novel experimental results to help us understand the genes and structures in this plasmid.

Acknowledgments We thank Dr. David W.C. Li (The Hormel Institute, University of Minnesota) for helpful comments during preparation of this manuscript. This work was supported by grants from the National Natural Science Foundation of China (No. 30270026), the National "863" Program of China (No. 2002AA629020), and the Chinese Academy of Sciences (No. KSCX2-3-01-02).

References

Baker TA, Kornberg A (1988) Transcriptional activation of initiation of replication from the *E. coli* chromosomal origin: an RNA-DNA hybrid near *oriC*. Cell 55:113-123

Charlebois RL, Lam WL, Cline SW, Doolittle WF (1987) Characterization of pHV2 from *Halobacterium volcanii* and its use in demonstrating transformation of an archaebacterium. Proc Natl Acad Sci USA 84:8530-85344

Cowan DA (1992) Biotechnology of the *Archaea*. Trends Biotechnol 10:315-323

del Solar G, Giraldo R, Ruiz-Echevarria MJ, Espinosa M, Diaz-Orejas R (1998) Replication and control of circular bacterial plasmids. Microbiol Mol Biol Rev 62:434–464

Erauso G, Marsin S, Benbouzid-Rollet N, Baucher MF, Barbeyron T, Zivanovic Y, Prieur D, Forterre P (1996) Sequence of plasmid pGT5 from the archaeon *Pyrococcus abyssi*: evidence for rolling-circle replication in a hyperthermophile. J Bacteriol 178:3232–3237

Hackett NR, Krebs MP, DasSarma S, Goebel W, RajBhandary UL, Khorana HG (1990) Nucleotide sequence of a high copy number plasmid from *Halobacterium* strain GRB. Nucleic Acids Res 18:3408

Hall MJ, Hackett NR (1989) DNA sequence of a small plasmid from *Halobacterium* strain GN101. Nucleic Acids Res 17:10501

Harriott OT, Huber R, Stetter KO, Betts PW, Noll KM (1994) A cryptic miniplasmid from the hyperthermophilic bacterium *Thermotoga* sp. strain RQ7. J Bacteriol 176:2759–2762

Hein M, Wegener AA, Engelhard M, Siebert F (2003) Time-resolved FTIR studies of sensory rhodopsin II (NpSRII) from *Natronobacterium pharaonis*: implications for proton transport and receptor activation. Biophys J 84:1208–1217

Holmes M, Pfeifer F, Dyall-Smith M (1994) Improved shuttle vectors for *Haloferax volcanii* including a dual-resistance plasmid. Gene 146:117–121

Holmes ML, Pfeifer F, Dyall-Smith ML (1995) Analysis of the halobacterial plasmid pHK2 minimal replicon. Gene 153:117–121

Kagramanova VK, Derckacheva NI, Mankin AS (1988) The complete nucleotide sequence of the arcaebacterial plasmid pHSB from *Halobacterium* strain SB3. Nucleic Acids Res 16:4158

Khan SA (1997) Rolling-circle replication of bacterial plasmids. Microbiol Mol Biol Rev 61:442–455

Khan SA (2000) Plasmid rolling-circle replication: recent developments. Mol Microbiol 37:477–484

Koonin EV, Ilyina TV (1993) Computer-assisted dissection of rolling circle DNA replication. Biosystems 30:241–268

Kramer MG, Khan SA, Espinosa M (1997) Plasmid rolling circle replication: identification of the RNA polymerase-directed primer RNA and requirement for DNA polymerase I for lagging strand synthesis. EMBO J 16:5784–5795

Lam WL, Doolittle WF (1989) Shuttle vectors for the archaebacterium *Halobacterium volcanii*. Proc Natl Acad Sci USA 86:5478–5482

Litchfield CD (1998) Survival strategies for microorganisms in hypersaline environments and their relevance to life on early Mars. Meteor Planet Sci 33:813–819

Liu J, Xiang H, Zhou Y, Zhu Z, Zhang J (1998) A new technique of screening gene mutation: fluorescent fingerprinting. Genet Anal Biomol Eng 14:113–116

Mankin AS, Zyrianova IM, Kagramanova VK, Garrett RA (1992) Introducing mutations into the single-copy chromosomal 23S rRNA gene of the archaeon *Halobacterium halobium* by using an rRNA operon-based transformation system. Proc Natl Acad Sci USA 89:6535–6539

Mayr A, Pfeifer F (1997) The characterization of the nv-gvpACNOFGH gene cluster involved in gas vesicle formation in *Natronobacterium vacuolatum*. Arch Microbiol 168:24–32

Ng WL, Yang CF, Halladay JT, Arora P, DasSarma S (1995) Isolation of genomic and plasmid DNAs from *Halobacterium halobium*. In: DasSarma S, Fleischmann EM (eds) Halophiles, archaea: a laboratory manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., pp 181–184

Ochsenreiter T, Pfeifer F, Schleper C (2002) Diversity of *Archaea* in hypersaline environments characterized by molecular-phylogenetic and cultivation studies. Extremophiles 6:267–274

Polosina YY, Jarrell KF, Fedorov OV, Kostyukova AS (1998) Nucleoside diphosphate kinase from haloalkaliphilic archaeon *Natronobacterium magadii*: purification and characterization. Extremophiles 2:333–338

Russell NJ (1989) Adaptive modifications in membranes of halotolerant and halophilic microorganisms. J Bioenerg Biomembr 21:93–113

Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning: a laboratory manual, 2nd edn. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.

Schmies G, Luttenberg B, Chizhov I, Engelhard M, Becker A, Bamberg E (2000) Sensory rhodopsin II from the haloalkaliphilic *Natronobacterium pharaonis*: light-activated proton transfer reactions. Biophys J 78:967–976

Seery LT, Nolan NC, Sharp PM, Devine KM (1993) Comparative analysis of the pC194 group of rolling circle plasmids. Plasmid 30:185–196

Sioud M, Baldacci G, Forterre P, Recondo AM de (1988) Novobiocin induces accumulation of a single strand of plasmid pGRB-1 in the archaebacterium *Halobacterium* GRB. Nucleic Acids Res 16:7833–7842

Strauss EC, Kobori JA, Siu G, Hood LE (1986) Specific-primer-directed DNA sequencing. Anal Biochem 154:353–360

Taylor K, Wegrzyn G (1995) Replication of coliphage lambda DNA. FEMS Microbiol Rev 17:109–119

Tindall BJ, Mills AA, Grant WD (1980) An alkaliphilic red halophilic bacterium with a low magnesium requirement from a Kenyan soda lake. J Gen Microbiol 116:257–260

Upasani VN, Desai SG, Moldoveanu N, Kates M (1994) Lipids of extremely halophilic archaeobacteria from saline environments in India: a novel glycolipid in *Natronobacterium* strains. Microbiology 140:1959–1966

Ward DE, Revet IM, Nandakumar R, Tuttle JH, Vos WM de, Oost J van der, DiRuggiero J (2002) Characterization of plasmid pRT1 from *Pyrococcus* sp. strain JT1. J Bacteriol 184:2561–2566

Wise KJ, Gillespie NB, Stuart JA, Krebs MP, Birge RR (2002) Optimization of bacteriorhodopsin for bioelectronic devices. Trends Biotechnol 20:387–394

Xiang H, Wang J, Mao YW, Li DW (2000) hTERT can function with rabbit telomerase RNA: regulation of gene expression and attenuation of apoptosis. Biochem Biophys Res Commun 278:503–510