

Precise Determination, Cross-Recognition, and Functional Analysis of the Double-Strand Origins of the Rolling-Circle Replication Plasmids in Haloarchaea[∇]

Ligang Zhou,^{1,2} Meixian Zhou,^{1†} Chaomin Sun,^{1‡} Jing Han,^{1,2} Qiuhe Lu,^{1,2}
Jian Zhou,¹ and Hua Xiang^{1*}

State Key Laboratory of Microbial Resources, Institute of Microbiology, Chinese Academy of Sciences,¹ and
Graduate University of Chinese Academy of Sciences,² Beijing 100101, People's Republic of China

Received 30 April 2008/Accepted 9 June 2008

The precise nick site in the double-strand origin (DSO) of pZMX201, a 1,668-bp rolling-circle replication (RCR) plasmid from the haloarchaeon *Natrinema* sp. CX2021, was determined by electron microscopy and DSO mapping. In this plasmid, DSO nicking occurred between residues C404 and G405 within a heptanucleotide sequence (TCTC/GGC) located in the stem region of an imperfect hairpin structure. This nick site sequence was conserved among the haloarchaeal RCR plasmids, including pNB101, suggesting that the DSO nick site might be the same for all members of this plasmid family. Interestingly, the DSOs of pZMX201 and pNB101 were found to be cross-recognized in RCR initiation and termination in a hybrid plasmid system. Mutation analysis of the DSO from pZMX201 (DSO_Z) in this hybrid plasmid system revealed that: (i) the nucleotides in the middle of the conserved TCTCGGC sequence play more-important roles in the initiation and termination process; (ii) the left half of the hairpin structure is required for initiation but not for termination; and (iii) a 36-bp sequence containing TCTCGGC and the downstream sequence is essential and sufficient for termination. In conclusion, these haloarchaeal plasmids, with novel features that are different from the characteristics of both single-stranded DNA phages and bacterial RCR plasmids, might serve as a good model for studying the evolution of RCR replicons.

Rolling-circle replication (RCR) in bacterial plasmids has been extensively studied (5, 12, 20, 21). In the initiation step, the initiator protein (Rep) introduces a specific nick in the double-strand origin (DSO) and generates a 3'-OH end to serve as the primer for synthesis of the leading strand. After one round of replication, the Rep protein terminates the replication at the DSO by a series of concerted cleavage/joining reactions and thus generates a double-stranded plasmid for the next round of replication and a circular single-stranded DNA (ssDNA) intermediate that is converted to a double-stranded plasmid from a single-strand origin (SSO) by host enzymes. Therefore, the DSO serves as an important signal in both the initiation and termination of RCR (11, 18). Previous studies have revealed that the DSO sequences used for initiation and termination overlap, although initiation requires a longer sequence than that for termination (11, 43).

Within the DSO sequence, two functional sites have been defined. They are a binding site where the Rep protein binds to the origin and a nick site where the Rep protein cleaves the DNA to initiate or terminate the RCR (6). Among plasmids

from the same family, the nick site is highly conserved but the binding site is not, suggesting that the origin's specificity is provided by the sequence differences in the binding site (5). The DNA sequence of the binding site can be either an inverted repeat or a set of direct repeats, and the nick site is always located in the loop region of a hairpin structure (7, 31, 33), which is regarded as necessary for the Rep protein to introduce the nick, especially when the plasmid is in a tightly packed super-coiled form (4, 30, 31).

In the domain *Archaea*, two distinct RCR plasmid families have been identified according to the homologies of Rep proteins and DSO sequences. One family includes plasmids pGT5 (8), pRT1 (40), and pTN1 (36), isolated from hyperthermophilic archaea. The DSO region of pGT5 is homologous to the DSO region of the pC194 plasmid family, especially in the conserved nick sequence TCTTG/ATA (the slash denotes the nick site) (8). Like its bacterial counterparts, the Rep protein of pGT5 is revealed to have nick-closing and topoisomerase activities but, distinctively, has an additional nucleotidyl transferase activity (26–28). Another family includes more than six plasmids (pGRB1, pHSB1, pHSB2, pHGN1, pHK2, and pNB101) isolated from halophilic archaea (13–16, 19, 44). These haloarchaeal plasmids are classified into superfamily I with bacteriophage ΦX174 because the conserved motif III of their Rep proteins contains two tyrosines (17, 44). The putative DSO nick sites of pGRB1, pHGN1, and pNB101 are only predicted, by sequence comparison with other members of superfamily I (25, 44), but remain to be determined by experimental approaches.

It has been reported that in a hybrid plasmid containing two heterologous DSOs (i.e., the DSOs of pT181 and pC221), the

* Corresponding author. Mailing address: State Key Laboratory of Microbial Resources, Institute of Microbiology, Chinese Academy of Sciences, Datun Road, Chaoyang District, Beijing 100101, People's Republic of China. Phone and fax: (86) 10-6480-7472. E-mail: xiangh@sun.im.ac.cn.

† Present address: Departments of Pediatric Dentistry and Microbiology, University of Alabama at Birmingham Schools of Medicine and Dentistry, Birmingham, AL 35244.

‡ Present address: Department of Molecular and Cell Biology, University of California, Berkeley, CA 94720.

[∇] Published ahead of print on 20 June 2008.

TABLE 1. Primers used in this study

Primer ^a	Sequence (5'-3') ^b
NB4F	GGA <u>ACCATGGCGCTGGAGTTAGAGCCTGAG</u>
NB4R	GGA <u>AGGATCCGACGGAAACGACGAAACC</u>
DM1	CGT <u>TACTCCGTCCTCGTCTGTGTC</u>
DM2	CACAGCAGACCACCGCAAGC
ND92F	GGA <u>AGGTACCGCAGGCGGCGATAGGC</u>
ND92R	GGA <u>AGGATCCCGAAACCCGGCGGTGAC</u>
ZD89F	TCTTGGT <u>ACCGTGGTTCGTGAGTAGCG</u>
ZD89R	TCTTGGATCCATCCTTCGTTTTCGGC
ZD400F	GGA <u>AGGTACCTGCTACAAGAGACGTTTC</u>
ZD400R	GGA <u>AGGATCCTGGAGGTCGAGGAGATC</u>
ZD400IVF	GGA <u>AGGATCCTGCTACAAGAGACGTTTC</u>
ZD400IVR	GGA <u>AGGTACCTGGAGGTCGAGGAGATC</u>
ZDR1F	GGA <u>AGGTACCAAGCCGAGGACTCCGCC</u>
ZDR2F	GGA <u>AGGATCCTCTCGGCTATCCAAGAC</u>
ZDLR	GGA <u>AGGATCCTAGCCGAGAGGCGGAGTC</u>

^a ZDR1F or ZDR2F were paired with ZD400R and ZDLR paired with ZD400F for PCR.

^b The restriction sites are underlined.

termination sequences of both DSOs are recognized and small, derivative plasmids are generated when the Rep protein is provided in *trans* (18). Similarly, the Rep protein of pE194 is able to initiate RCR at its own DSO and terminate at the DSO of pMV158 (37). In this study, we report the determination of the precise nick site of the DSO of a novel RCR plasmid, pZMX201 (DSO_Z), from the haloarchaeon *Natrinema* sp. CX2021 by electron microscopy and DSO mapping. In addition, by establishing a hybrid plasmid system derived from two haloarchaeal RCR plasmids, pNB101 (44) and pZMX201, the cross-recognition of the haloarchaeal DSOs and the DSO sequence motifs utilized in RCR were comprehensively investigated.

MATERIALS AND METHODS

Strains, media, and transformation. *Escherichia coli* strain JM109 was used as the host for all cloning and sequencing experiments and was cultured in Luria-Bertani medium (34). When needed, ampicillin was added to a final concentration of 100 µg/ml. The halophilic archaeon *Natrinema* sp. CX2021 is the natural host of pZMX201. *Haloarcula hispanica* strain ATCC 33960 was used as the transformation host. These haloarchaeal strains were cultivated at 37°C in AS-168 medium as described previously (24). Transformation of *H. hispanica* was performed as described by Cline et al. (3). When necessary, mevinolin was added to a final concentration of 3 to 5 µg/ml. The plasmid DNA prepared from mevinolin-resistant colonies was subjected to electrophoresis and Southern blotting and introduced back into *E. coli* for further analysis (45).

DNA manipulation. Small-scale preparation of plasmid DNA from a halophilic archaeon was performed as described previously (44). Restriction analysis and recombinant DNA manipulation were performed according to standard methods (34). The plasmid pZMX201 was isolated from *Natrinema* sp. CX2021, digested with NcoI and ScaI, and inserted into pUCmT (38) to generate the recombinant plasmids pUCZ2 and pUCZ3, respectively. Both plasmids were sequenced, and the complete sequence of pZMX201 was determined. To clone the replicon of pNB101, the *repN* gene and flanking sequence were amplified by primers NB4F and NB4R (Table 1) and then inserted into pUCmT to generate recombinant plasmid pUCN4 for further construction. For DSO mapping and electron microscopy, total DNA containing the plasmid replication intermediates of pZMX201 from *Natrinema* sp. CX2021 and those of pUCZ3 from *E. coli* were isolated as described previously (29, 35).

Electron microscopy. Total DNA of pZMX201 was separated by electrophoresis in a low-melting-temperature agarose gel, and then the plasmid DNA between the supercoiled plasmid and high-molecular-weight chromosomal bands was isolated from the gel. DNA samples with or without enzyme digestion were further prepared as described previously (38). Micrographs were recorded by

using a Hitachi H-700A electron microscope. Plasmid pBR322 was used as a size marker.

DSO mapping. DSO mapping was performed as previously described (10, 42), with slight modifications. The primer was labeled with T4 polynucleotide kinase with [γ -³²P]ATP (5,000 Ci/mmol) and used for DNA sequencing and primer extension. The total DNA and plasmid DNA from *Natrinema* sp. CX2021 and *E. coli* were used as templates for primer extension with the following program: 3 min at 94°C and then 40 cycles of 1 min at 94°C, 1 min at 60°C, and 1 min at 72°C. At the same time, DNA sequencing reactions were performed with the same primers by using an fmol DNA cycle sequencing system (Promega) according to the protocols of the manufacturer. The products of primer extension and the DNA-sequencing reaction were respectively resuspended in the same sequencing stop solution, heated for 2 min at 70°C, and then analyzed on 6% acrylamide sequencing gel. Radioactive DNA fragments on the gels were visualized on X-ray film.

Construction of hybrid plasmids. The tested DSO fragment of pNB101 (ND92) was amplified by primer pairs ND92F and ND92R (Table 1) and inserted into the KpnI and BamHI sites of pWL102 (22). The resultant plasmid was digested with SphI and HindIII to excise the *E. coli* replicon and then inserted into the same sites of pUCZ2 to generate pZN92. Similarly, other tested DSO fragments of pZMX201 (ZD89, ZD400, ZD400-IV, ZDR1, ZDR2, and ZDL) were amplified by using the corresponding primer pairs (Table 1) and inserted into the KpnI and BamHI sites of pWL102. The resultant plasmids were digested with SphI and HindIII and then inserted into the same sites of pUCN4 to generate pNZ89, pNZ400, pNZ400-IV, pNZR1, pNZR2, and pNZL, respectively. Site-directed mutations, conducted as precisely described (41), were introduced into the *repZ* gene of pZMX201 in pZN92 to generate mutants pZNY297F and pZNY301F or into the TCTCGGC sequence of DSO_Z of pZMX201 (ZD89) in pNZ89 to generate the corresponding mutants. To construct the hybrid pNZ36 and its three derivative mutants pNZ36S1, pNZ36S2, and pNZ36S3, oligonucleotides containing both strands of the tested sequences (ZD36 and ZD36S1-S3) with additional restriction sites were synthesized, annealed, digested and then inserted into the KpnI and BamHI sites of pWL102, and the resultant plasmids were digested with SphI and HindIII and then inserted into the same sites of pUCN4 to generate pNZ36, pNZ36S1, pNZ36S2, and pNZ36S3, respectively.

Southern blotting. To identify the multiple DNA bands, Southern blot analysis was conducted as previously described (44). Briefly, 2% amounts of DNA samples extracted from 1 ml of cultures (optical density at 600 nm, ≈0.6) were resolved by agarose gel electrophoresis and then transferred onto a nylon membrane with prior denaturation. After prehybridization, the blots were hybridized with the DNA probes, which were labeled through random priming of the mevinolin resistance (*Mev*^r) gene or ampicillin resistance (*Amp*^r) gene with digoxigenin-dUTP (Roche) according to the manufacturer's instructions.

Accession number. The nucleotide sequences of pZMX201 and the 16S rRNA gene of *Natrinema* sp. CX2021 were deposited in GenBank under the accession numbers AY588480 and AY588481, respectively.

RESULTS

pZMX201 is a new member belonging to the haloarchaeal RCR plasmid family. The small, cryptic plasmid named pZMX201 was isolated from a haloarchaeon, *Natrinema* sp. CX2021. Sequence analysis revealed that pZMX201 is a circular DNA molecule of 1,668 bp with a GC content of 62.5% and contains only one open reading frame (*repZ*), encoding a putative replication initiator protein (RepZ) (Fig. 1A). By the results of searching the NCBI database, RepZ was revealed to be homologous (42 to 56%) to the Rep proteins of other haloarchaeal RCR plasmids. Alignment of the Rep proteins of these plasmids detected three conserved motifs (Fig. 1B), especially the two-tyrosine-containing motif III that was conserved in RCR superfamily I (17, 44). These results indicated that pZMX201 is a new member of the haloarchaeal RCR plasmid family. Moreover, several inverted repeats were detected in the noncoding region (Fig. 1A), which may contain the SSO for pZMX201 due to its potential to form long stem-loop structures, like bacterial-type SSOs (21). The presence of

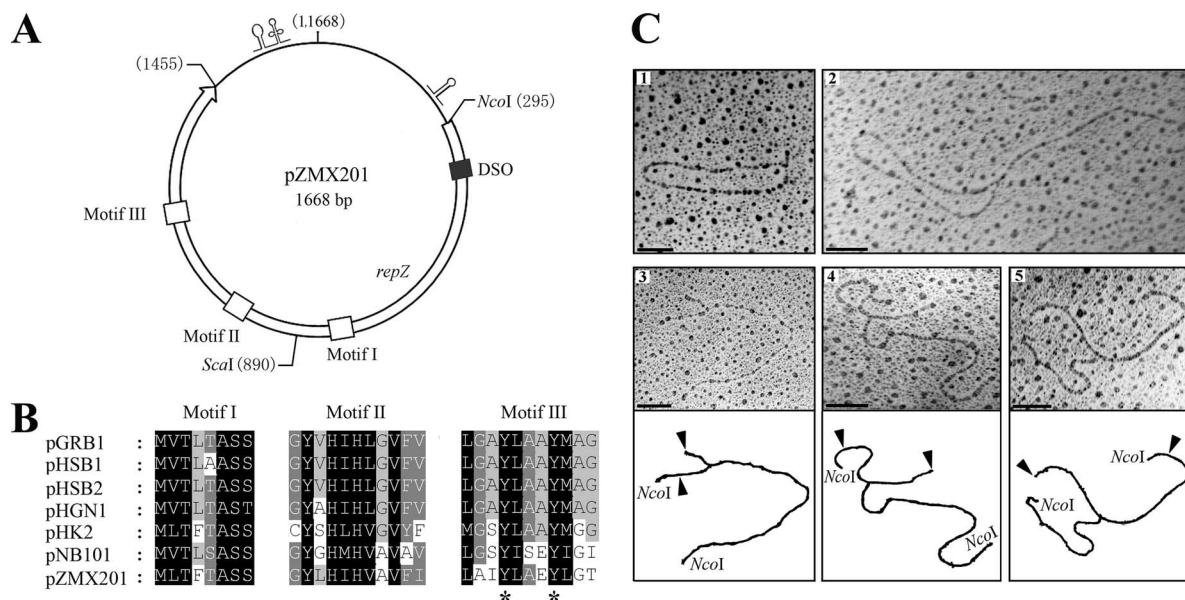


FIG. 1. (A) Physical map of pZMX201. The conserved motifs (I, II, and III) in the protein encoded by *repZ* and the DSO determined by DSO mapping are indicated. In the noncoding region (nt 1455 to 295), several inverted repeats that can potentially form long stem-loop structures are also indicated. (B) Sequence alignment of the three conserved motifs (I, II, and III) of the Rep proteins of haloarchaeal RCR plasmids. Shading in black, dark gray, or light gray indicates 100%, >80%, or 50% conserved residues, respectively. The two conserved tyrosines in motif III are labeled with asterisks. (C) Electron microscopy analysis of pZMX201. Panels 1 and 2, sigma-like replication intermediates; panels 3 to 5, Y-like molecules of *NcoI*-digested plasmids, with a schematic representation under each panel. The deduced DSO nick sites are indicated by triangles. The bars correspond to 0.2 kb.

single-stranded (ssDNA) intermediates of pZMX201 was also detected by Southern blotting (data not shown), further confirming that it is a novel haloarchaeal RCR plasmid.

Determination of the DSO nick site in pZMX201 by electron microscopy analysis and DSO mapping. Plasmid DNA recovered from low melting-temperature agarose was prepared for electron microscopy analysis. Thirty-seven sigma-like replication intermediates were detected among 675 DNA molecules (Fig. 1C, panels 1 and 2). These sigma-like intermediates have similarly sized circles with extended tails, indicating the process of replication of the leading strand. To locate the replication origin, the plasmid DNA was digested with *NcoI* before being subjected to electron microscopy analysis. Y-like molecules with various branch ends representing the *NcoI* and DSO nick sites were observed (Fig. 1C, panels 3 to 5), and the lengths of the branches from 10 representative molecules were carefully measured. The results indicate that the DSO nick site is close to the *NcoI* site.

To precisely determine the DSO nick site in pZMX201, a primer extension experiment was performed (Fig. 2A) with primer DM1 (Table 1), downstream of the *NcoI* site and complementary to the putative leading strand, the sense strand of the *repZ* gene. A strong primer extension band corresponding to G404 was generated with total DNA or plasmid extracted from *Natrinema* sp. CX2021 as the template (Fig. 2A). This band was never generated when the recombinant plasmid (pUCZ3) DNA from *E. coli* was used as template in the same experiments. These results demonstrated that the precise DSO nick site of pZMX201 should be between C403 and G404 (Fig. 2A). Moreover, when a primer (DM2; Table 1) complementary to the other strand upstream of the *NcoI* site was used, no

extension products could be detected (data not shown). Therefore, the extension signal was not ascribed to either the DNA secondary structure or the double-strand break but was indeed generated by the nicking of the Rep protein in the DSO.

After the sequence containing the nick site of pZMX201 was aligned with the sequences of other haloarchaeal RCR plasmids, a heptameric sequence (TCTC/GGC; the slash denotes the detected nick site) was found to be conserved among all these plasmids. In the adjacent nucleotides, however, only partially conserved sequence motifs, such as a CCAAG and an AACCG, were found (Fig. 2B). By DNA-folding analysis, imperfect hairpin structures were discovered in these putative DSO regions. Interestingly, although the heptameric sequence (TCTC/GGC) is usually located in the stem region of a hairpin structure, the nick site is in an unpaired position (pHGN1 and pHSB1) or near an unpaired nucleotide (pGRB1, pHK2, pHSB2, pNB101, and pZMX201) (Fig. 2C).

Cross-recognition of the DSO by Rep proteins from plasmids pZMX201 and pNB101. To verify these novel DSOs in haloarchaea, two hybrid plasmids (pNZ89 and pZN92) carrying different haloarchaeal RCR replicons and tested DSO fragments were constructed with the same additional elements, including an *Mev^r* gene, a haloarchaeal theta-type replicon (pHV2) (2), and an *E. coli* replicon (pUCmT) with an *Amp^r* gene (Fig. 3). pNZ89 contains a replicable pNB101 fragment and an 89-bp DSO_Z (ZD89) from pZMX201 (Fig. 3A), while pZN92 contains an *NcoI*-pZMX201 fragment and a 92-bp DSO (ND92) from pNB101 (DSO_N) (Fig. 3B). Based on pNZ89, various hybrid plasmids containing different DSO_Z fragments or mutants were also established (Fig. 3C).

The pHV2 replicon in these plasmids likely replicates by a

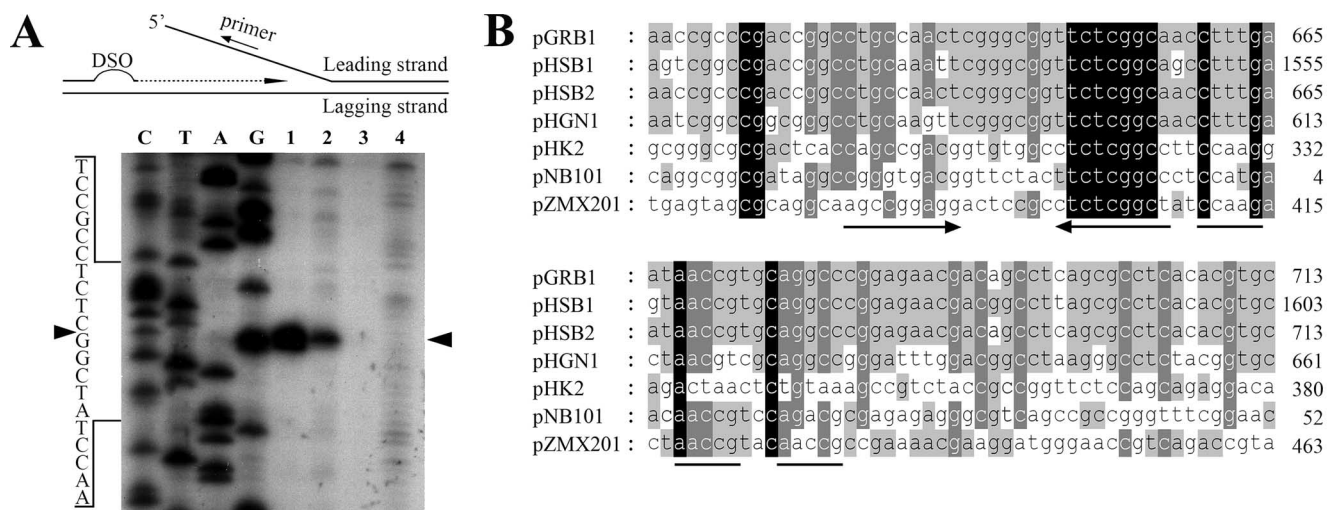


FIG. 2. (A) Mapping of the DSO nick site of pZMX201. Top panel: outline of the DSO mapping. Dotted arrow indicates the newly synthesized leading strand. The extension of the primer (solid arrow) complementary to the leading strand will stop at the 5' end nicked by the Rep protein and generate a specific product corresponding to the nick site in the sequencing gel. Bottom panel: total DNA (lane 1) and plasmid DNA (lane 2) from *Natrinema* sp. CX2021 cells containing pZMX201, and total DNA (lane 3) and plasmid DNA (lane 4) from *E. coli* cells containing the recombinant plasmid pUCZ3 were extended by the primer DM1 (Table 1). The corresponding sequence is shown on the left. The nick site is denoted by the pair of triangles. (B) Alignment of the putative DSO regions of the haloarchaeal RCR plasmids. The inverted repeat of pZMX201 is indicated by the pair of arrows. The partially conserved CCAAG and AACCG sequences are underlined. The conserved nucleotides are shaded in black (100%), gray (>80%), or light gray (>50%). (C) Results of folding analysis of the putative DSO regions of the haloarchaeal RCR plasmids using the mfold web service (46). Numbering corresponds to the published sequence of each plasmid. pHSB2, whose sequence in this DSO region is identical to that of pGRB1, is listed in parentheses. The position of the conserved TCTCGGC sequence of pZMX201 is denoted. The nick site in DSO_Z and the putative ones in other DSOs are indicated by triangles. dG, change in Gibbs free energy.

theta-type model and is compatible with RCR replicons (16), therefore supplying a platform for investigating the interaction between Rep proteins and DSOs in these hybrid plasmids. If the Rep protein, RepN from pNB101 or RepZ from pZMX201, is able to recognize the heterologous DSO as a termination signal after initiating from its own DSO, an ssDNA intermediate of the derivative plasmid containing an *Mev*^r gene and an RCR replicon with the chimeric DSO (M) will be generated. On the other hand, if the replication can initiate from the heterologous DSO and terminate at its own DSO, an ssDNA intermediate of the derivative plasmid containing the pUCmT replicon and the pHV2 replicon (E) will be generated (Fig. 3A and B). These ssDNA intermediates might be converted into the double-stranded plasmids by host enzymes.

After pNZ89 was introduced into *H. hispanica* and the plasmid DNA was extracted from the transformants and subjected

to electrophoresis, only the M derivative was detected by agarose gel electrophoresis (Fig. 4A), and a faint E-like band could be detected by Southern blot analysis (Fig. 4C). To confirm the presence of the E derivative or the full-length hybrid plasmid (H), the plasmid DNA extracted from the *H. hispanica* transformants was introduced back into *E. coli* (Fig. 4A and Table 2). Twenty *E. coli* transformants were analyzed and found to harbor the same plasmid (derivative E), which contains the pHV2 replicon and a junction site between DSO_N and DSO_Z (Fig. 4D). As the E derivative can hardly be detected by electrophoresis or Southern blotting, the number of *E. coli* transformants transformed by the plasmids from *H. hispanica* was used to evaluate the amount of the E derivative, which represented the efficiency of initiation from DSO_Z (Table 2). As the H plasmid was not detected either by Southern blotting or by screening the transformants of *E. coli* (Fig. 4 and Table 2), it is likely that efficient termination had dis-

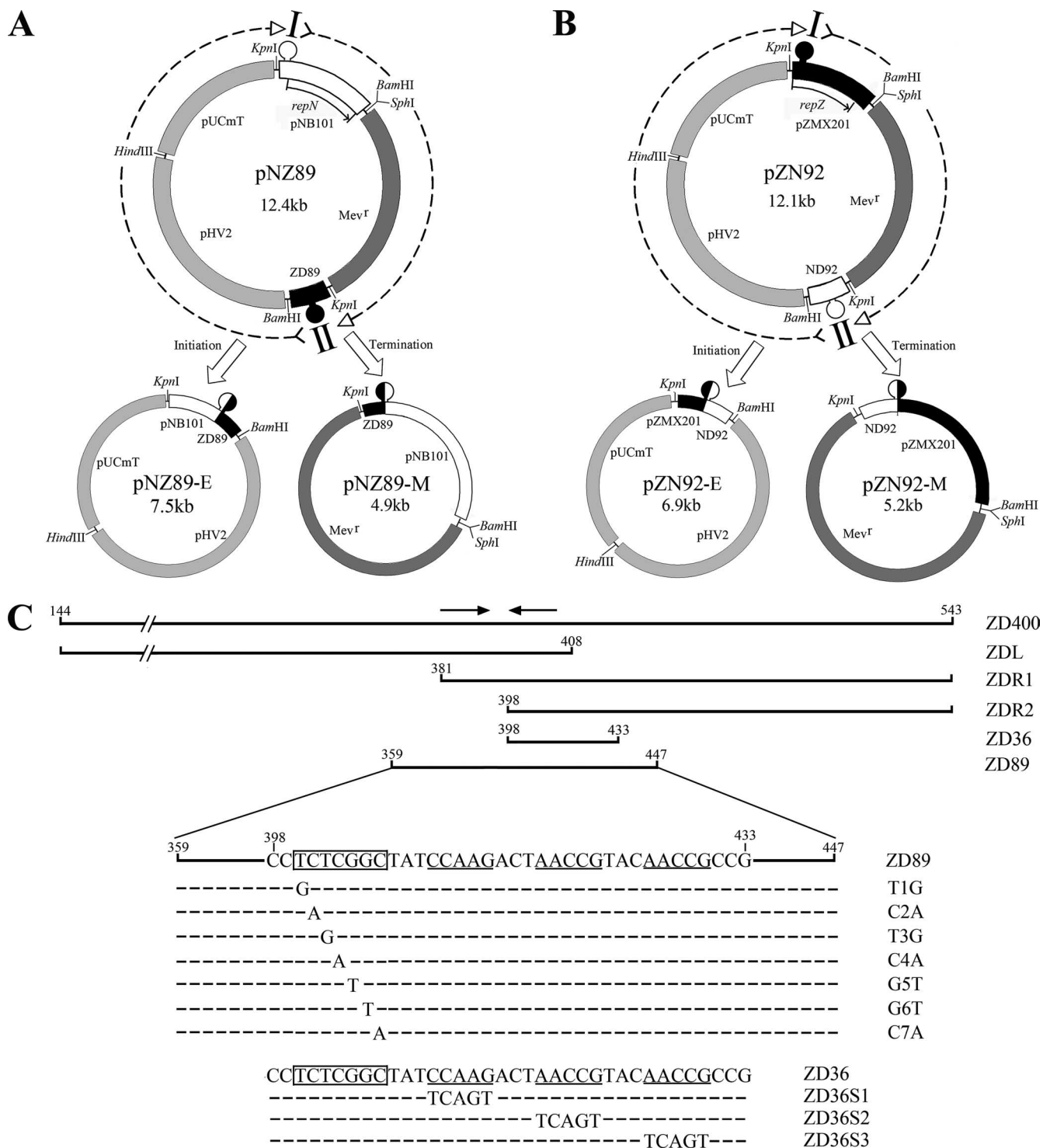


FIG. 3. (A and B) Physical map of hybrid plasmids pNZ89 and pZN92 and their derivatives. If the RCR initiates from the original DSO (I) and terminates at the tested DSO (II), a derivative (designated M) containing an *Mev^r* gene and an RCR replicon with the chimeric DSO will be generated. If the replication initiates from the tested DSO (II) and terminates at the original DSO (I), a derivative (designated E) containing the pUCmT replicon and the pHV2 replicon will be generated. The open Ω -shaped symbol represents DSO_N from pNB101, the filled Ω -shaped symbol represents DSO_Z from pZMX201, and the half-filled Ω -shaped symbol indicates the junction site of the DSO sequences from pZMX201 and pNB101. (C) Schematic representation of the various DSO_Z fragments and mutants used in hybrid plasmids. Numbering corresponds to the sequence of pZMX201 in GenBank. The arrows indicate the position of the inverted repeat in DSO_Z. The nucleotide sequence of ZD36 is shown. The conserved TCTCGGC sequence is boxed, and the sequence motifs subjected to mutation are underlined. The identical nucleotides are indicated by dashes.

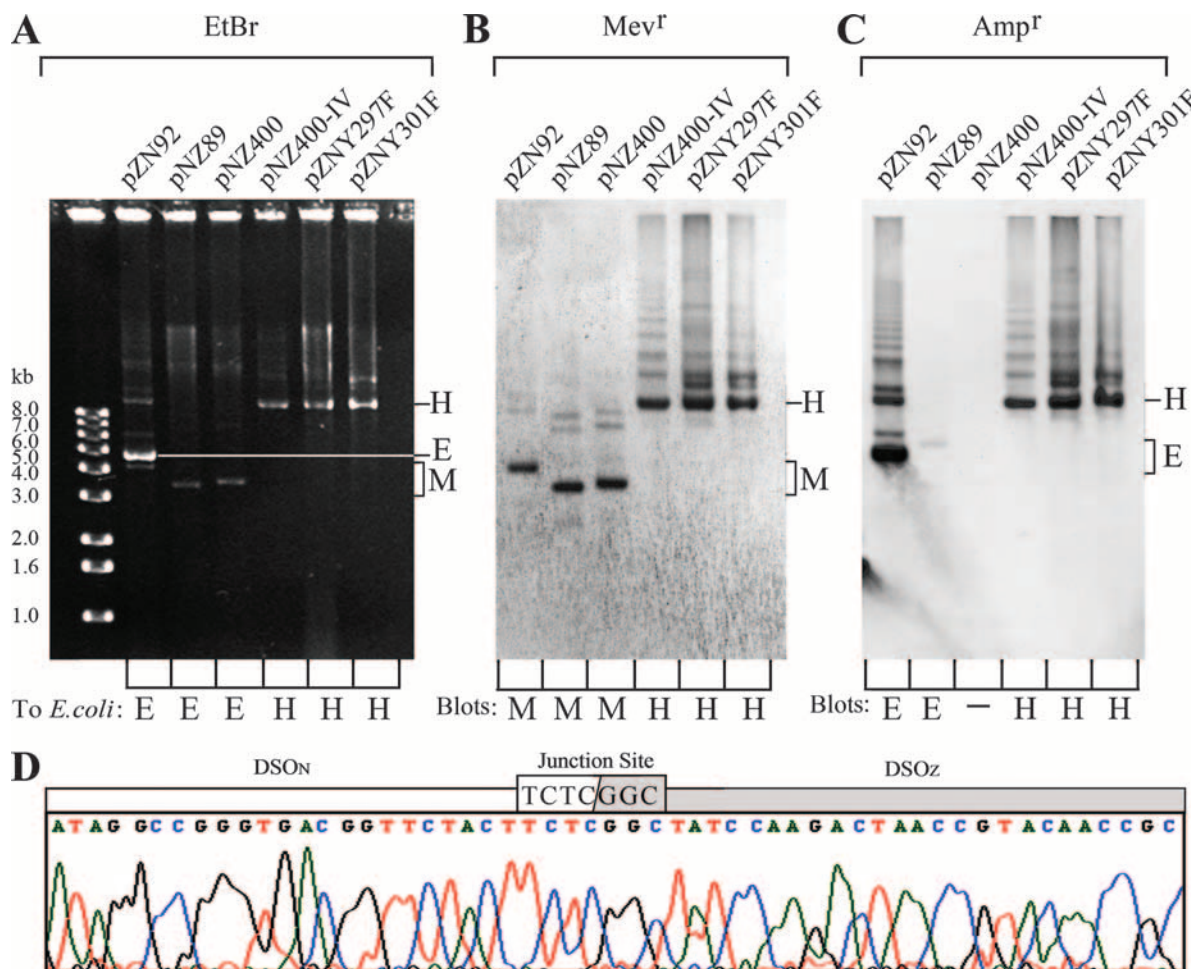


FIG. 4. (A) Agarose gel electrophoresis of hybrid plasmids extracted from corresponding *H. hispanica* transformants. The DNA was stained with ethidium bromide (EtBr). The predominant super-coiled DNA molecules are labeled with H (full-length hybrid plasmid), E (derivative containing *E. coli* replicon), and M (derivative containing *MevI* gene). DNA samples from equal amounts of cultures were subjected to electrophoresis and transformation of *E. coli*. The resultant plasmids detected in the *E. coli* transformants are indicated under the gel. The first lane contains size markers. (B and C) The blots were hybridized with probes *MevI* (B) and *AmpI* (C), respectively. Corresponding bands detected for each lane are denoted under the panel. The hybrids used for each experiment are indicated on top of each lane. -, not detectable. (D) Sequence of the DSO junction site in the E derivative of pNZ89 (pNZ89-E). The junction site (TCTC/GGC) and flanking sequences from DSO_N and DSO_Z are shown. The slash indicates the nick site.

sociated most pNZ89 plasmids, if not all of them, into small derivatives.

To exclude the possibility that the ZD89 fragment is too short to be used as an efficient initiation signal, a 400-bp DSO_Z fragment (ZD400) (Fig. 3C) of pZMX201 was used to replace the ZD89 fragment of pNZ89, resulting in pNZ400. However, similar results were obtained after introducing pNZ400 into *H. hispanica*, and the E derivative was only detectable by introducing it back into *E. coli* (Fig. 4 and Table 2), confirming that RepN could terminate efficiently at DSO_Z but initiated from it weakly.

In the case of pZN92, both the M and E derivatives were detected by electrophoresis (Fig. 4A). The M and E derivative plasmids appeared to be compatible, since their relative copy numbers had not changed after being incubated for 60 generations with or without selective pressure. Both the M and E derivatives were sequenced, and the junction sites were found to be indeed in the conserved sequence TCTCGGC (data not

shown). These results indicated that RepZ could terminate and initiate at ND92 (DSO_N). As the H plasmid was not maintained for both hybrids (Fig. 4 and Table 2), it seemed that the termination rate of RepZ at DSO_N was similar to that of RepN at DSO_Z but that, likely, the RepZ initiated more efficiently from DSO_N.

Small derivative plasmids were generated by Rep processing, but not by homologous recombination. To exclude the involvement of homologous recombination in the generation of derivative plasmids (M and E), the hybrid plasmid pNZ400-IV was constructed. The fragments in pNZ400-IV are same as those in pNZ400, except that the ZD400 fragment was inserted inversely. If homologous recombination occurred between DSO_N and DSO_Z, it would also generate two plasmids with sizes similar to the sizes of M and E. However, such plasmids were not detectable after pNZ400-IV was introduced into *H. hispanica* (Fig. 4 and Table 2). This result indicated that the small derivative plasmids (M and E) in the above investi-

TABLE 2. Summary of DSO recognition in the hybrid plasmids

Plasmid or point mutation	Plasmid(s) detected with Southern blotting	Transformation back into <i>E. coli</i> ^a		Activity of tested DSO ^c	
		Introduced plasmid ^b	No. of transformants	Termination	Initiation
pZN92	M, E	E	1.5 × 10 ⁴	+++	+++
pNZ89	M, E	E	3.8 × 10 ²	+++	++
pNZ400	M	E	3 × 10 ²	+++	++
pNZ400-IV	H	H	1.4 × 10 ⁴	-	-
pNZY297F	H	H	1.4 × 10 ⁴	-	-
pNZY301F	H	H	1.3 × 10 ⁴	-	-
T1G	M	E	<10 ²	+++	+
C2A	H	H	8 × 10 ³	-	-
T3G	H	H	7.4 × 10 ³	-	-
C4A	H	H	2.3 × 10 ³	-	-
G5T	H	H	1.9 × 10 ³	-	-
G6T	M	H	<10	+++	-
C7A	M	E	<10	+++	+
pNZL	H	H	8.9 × 10 ³	-	-
pNZR1	M	E	1.4 × 10 ²	+++	++
pNZR2	M	×	0	+++	-
pNZ36	M	×	0	+++	-
pNZ36S1	H, M	H	7 × 10 ³	++	-
pNZ36S2	H	H	8.1 × 10 ³	-	-
pNZ36S3	H	H	1.1 × 10 ⁴	-	-

^a DNA samples from equal amounts of cultures were introduced into *E. coli* to calculate the transformation rate of the H plasmid or E derivative, which represents the initiation efficiency of the tested DSO.

^b ×, not detectable.

^c -, no activity; +, ++, and +++, low-, intermediate-, and high-level activity, respectively.

gations were not generated by a homologous recombination and demonstrated that the orientation of the DSO should be consistent with the direction of replication.

To further confirm that the small derivative plasmids (M and E) were generated by the nick/closing activities of the Rep protein, the two tyrosines (Y297 and Y301) in motif III of RepZ, which were deemed to be involved in these activities of the Rep protein (16), were replaced by phenylalanine (F) in pZN92. The resulting plasmids, pZNY297F and pZNY301F,

were introduced into *H. hispanica*. Unlike pZN92, the small derivative plasmids (M and E) were not detectable in either of these two kinds of transformants (Fig. 4). Thus, the generation of M and E was indeed the result of the replication process conducted by the Rep protein, and both tyrosines in motif III of RepZ were necessary for its function.

Nucleotides in the conserved TCTCGGC sequence play variant roles in DSO recognition. Based on the construction of pNZ89, site-directed mutations were introduced into the conserved TCTCGGC sequence of ZD89 to construct mutants T1G, C2A, T3G, C4A, G5T, G6T, and C7A, respectively (Fig. 3C). For mutants C2A, T3G, C4A, and G5T, only the H plasmids were detected in all the experiments (electrophoresis, Southern blotting, and transformation), indicating that nucleotides C2, T3, C4, and G5 were essential for both initiation and termination (Fig. 5 and Table 2). In the case of mutants T1G and C7A, the M derivative was detected by electrophoresis and Southern blotting, and the E derivative was also detected, at lower levels, by introducing plasmid preparations into *E. coli* (Fig. 5 and Table 2). This indicated that mutation of T1 and C7 did not affect termination but largely decreased the efficiency of initiation. For mutant G6T, the M derivative was detectable, as in T1G and C7A, while the H plasmid, but not the E derivative, was detected at a low rate by transformation of *E. coli* (Fig. 5 and Table 2). This indicated that the substitution of G6 completely blocked initiation and partially depressed termination.

The left half of the stem-loop structure of DSO_Z is necessary for initiation, but not for termination. Three mutants, pNZR1, pNZR2, and pNZL, were constructed to identify the boundaries of the DSO_Z region. Fragment ZDL (in pNZL) consists of the stem-loop region plus the left part of the sequence of ZD400, ZDR1 (in pNZR1) consists of the stem-loop region plus the right part of the sequence of ZD400, and the ZDR2 fragment (in pNZR2) differed from ZDR1 by the omission of

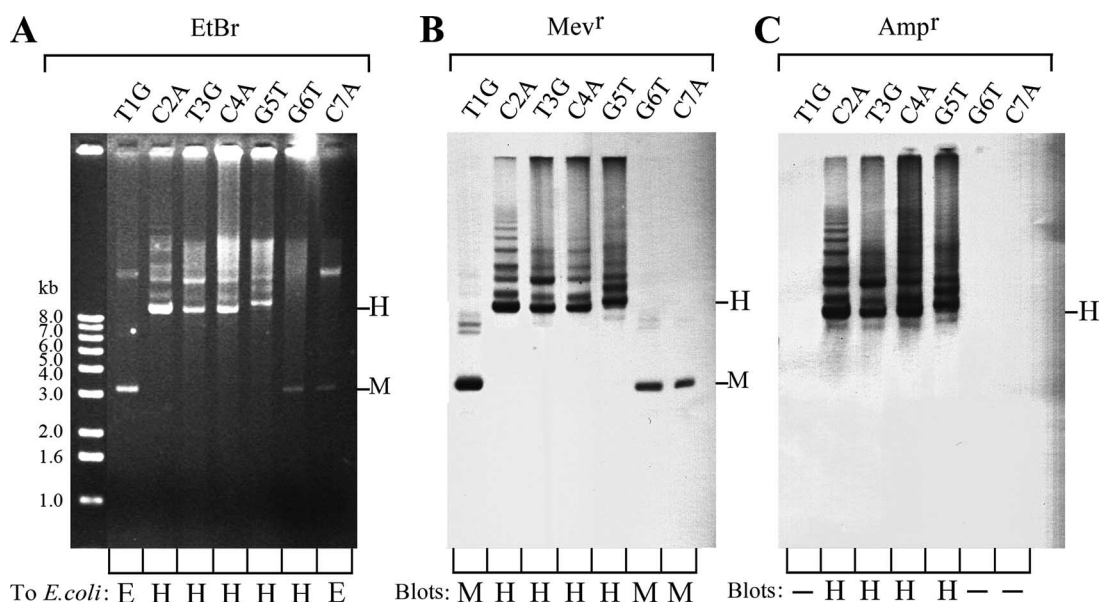


FIG. 5. Site-directed mutation of the conserved TCTCGGC sequence of the ZD89 fragment in pNZ89. Agarose gel stained with EtBr (A) and blots hybridized with probes Mev^I (B) and Amp^I (C) are shown as described in the Fig. 4 legend.

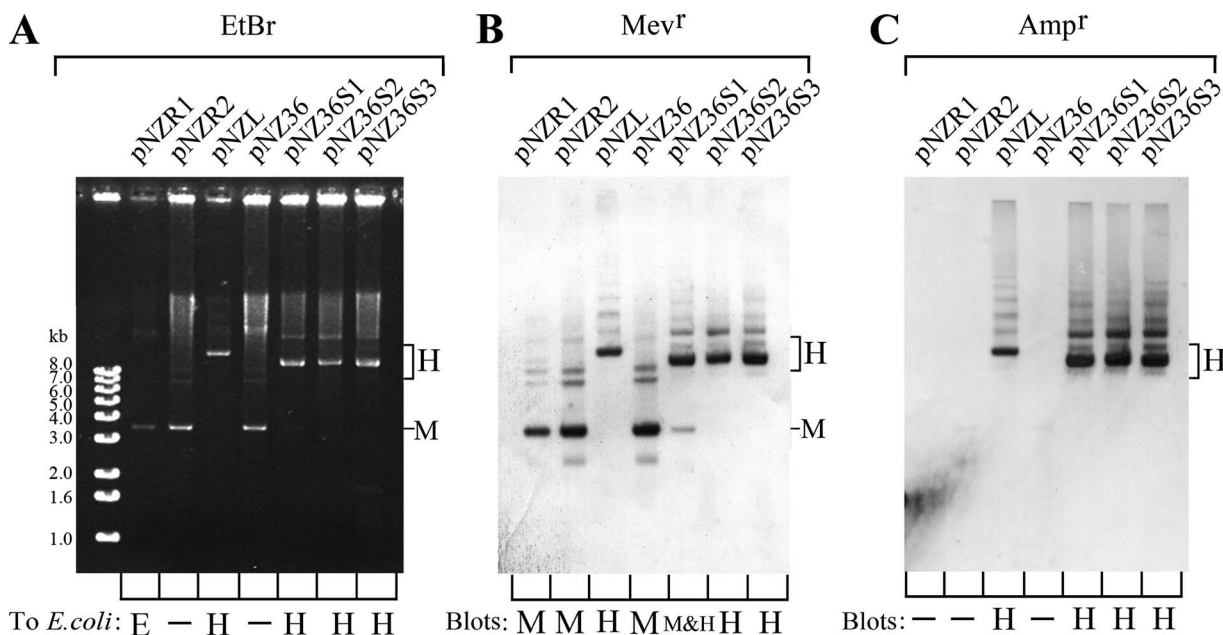


FIG. 6. Determination and mutational analysis of the necessary DSO region of DSO_Z. Agarose gel stained with EtBr (A) and blots hybridized with probes Mev^I (B) and Amp^r (C) are shown as described in the Fig. 4 legend.

the left half of the stem-loop structure (Fig. 3C). The M derivatives were detected by electrophoresis and Southern blotting for both pNZR1 and pNZR2; however, in the *E. coli* transformation experiments, the E derivative was detected for pNZR1, but no transformant was detected for pNZR2 (Fig. 6 and Table 2). These results indicated that ZDR1 was active in both initiation and termination, while ZDR2 lost the activity for initiation. For pNZL, only the H plasmid was detected in all these experiments (Fig. 6 and Table 2), indicating that the ZDL fragment was inactive for both initiation and termination. Taking these results together, the stem-loop region plus the right part of the sequence (ZDR1) were necessary and sufficient for initiation, the left part of the sequence upstream of the stem-loop region might not participate in the DSO function, and the DNA sequence in the left half of the stem-loop structure of DSO_Z was not necessary for termination but was indispensable for initiation.

A 36-bp DSO_Z fragment is sufficient for termination. Based on the results for pNZR2, hybrid plasmid pNZ36 was constructed, and the results proved that the 36-bp sequence located in the left part of fragment ZDR2, consisting of TCTCGGC plus the downstream sequence (Fig. 3C), is sufficient for termination but, again, inactive for initiation (Fig. 6 and Table 2). By comparing the 36-bp sequence of pZMX201 with the sequences of other haloarchaeal plasmids, a group of contiguous sequence motifs, one being CCAAG and two AACCG, were chosen to conduct mutation analysis. The CCAAG contains conserved C and G in all of these plasmids, while the first AACCG is conserved among five of these seven plasmids (Fig. 2B). Three mutants (ZD36S1 to ZD36S3) were constructed by replacing these motifs with a random sequence, TCAGT (Fig. 3C). Further examination indicated that the two AACCG sequences were indispensable for the termination function of pNZ36, as only the H plasmids were detected, while the re-

placement of the CCAAG sequence decreased the termination efficiency, as both H and M plasmids were detected, indicating that only part of the replication was terminated at ZD36S1 (Fig. 6 and Table 2). Thus, the 36-bp sequence may be the minimal DSO_Z for termination function.

DISCUSSION

The results of previous studies have shown that the haloarchaeal RCR plasmids belong to a novel family distantly related to the bacterial RCR plasmids and ssDNA phages (44). Although some of them have been used as shuttle vectors, the origins of replication have not been experimentally elucidated. In this work, the precise nick site of DSO_Z in pZMX201 in a conserved sequence (TCTC/GGC) was determined for the first time by both in vitro and in vivo experiments (Fig. 1 and 2). The DSO nick site sequence is conserved among the haloarchaeal family of RCR plasmids but is not homologous to those of any other RCR replicons (5). Employing a hybrid plasmid system, the DSOs of pZMX201 and pNB101 were found to be cross-recognized (Fig. 3 and 4). As the similarity between the DNA sequences of pZMX201 and pNB101 is only 34%, the chance of homologous recombination is reduced, so that the hybrid is stable when the tested DSO is inactive. In addition, the junction site of two heterologous DSOs was found, by sequencing the derivative plasmids, to be in the conserved TCTCGGC sequence (Fig. 4D), which confirmed that the nick site of DSO in vivo was consistent with the result of DSO mapping.

Mutation analysis of the conserved TCTCGGC sequence of DSO_Z revealed that nucleotides T1, G6, and C7 could be replaced without destroying the termination function of the sequence. To investigate the possibility that other regions harboring sequence motifs similar to TCTCGGC in plasmid

pZMX201 or pNB101 might also be recognized as DSO, three DSO candidates were identified and cloned, i.e., DSO_{Z2} (nt 770 to 968, containing a TCTCGAC sequence) and DSO_{Z3} (nt 35 to 263, containing a TCTCGGA sequence) from pZMX201 and DSO_{N2} (nt 1476 to 1655, containing a TCTCGGA sequence) from pNB101. After testing their initiation and termination functions in the pZN and pNZ plasmids, these DSO candidates were found to be inactive (data not shown). Thus, although the flanking sequences are not as conserved as the TCTCGGC motif, they may contain sequences or structures that are important for interaction with the Rep protein.

The structural features of the DSO regions from all the haloarchaeal RCR plasmids were analyzed (Fig. 2C). Among the hairpin structures detected, only two are structurally similar to other RCR DSOs, with the nick sites located in the unpaired region. In five other plasmids, including pZMX201, the nick site is in a stem region but near an unpaired position. As in bacterial plasmids, the nick site is always located in the loop region of a hairpin structure to facilitate the nicking by the Rep protein (30). It is possible that the hairpin structure in pZMX201 might serve as a target for recognition by the Rep protein and might melt from the unpaired position to expose the nick site upon the Rep/DSO interaction. By using deletion analysis, we found that the left half of the hairpin structure, present in pNZR1 but absent in pNZR2, was necessary for initiation but not for termination. Therefore, the left half of the hairpin structure (in pNZR1) may represent the 5' terminal boundary of DSO_Z for initiation. In addition, as ZD36 is likely the minimal DSO for termination, it is clear that the functional DSO_Z should contain the hairpin structure plus the downstream sequence. This functional DSO_Z is similar to those of bacterial plasmids (pLS1, pC194, and pT181) in structure, as the latter also contain a hairpin structure plus downstream sequence for the binding of Rep protein (7, 11, 43). However, different from the DSOs of bacterial plasmids where the hairpin structures are necessary for termination, the left half of the hairpin structure of DSO_Z is not necessary for its RCR termination. This characteristic is like the origin of ΦX174, where a 30-bp sequence without obvious secondary structures is sufficient for both initiation and termination (9, 23).

Besides these novel features of the haloarchaeal DSO sequence, the Rep proteins of the haloarchaeal RCR plasmid family contain a conserved two-tyrosine motif III in which both tyrosines were necessary for replication (Fig. 4 and Table 2). The two-tyrosine motif is characteristic of the Rep protein of some ssDNA phages and is also found in a few bacterial plasmids which have not been comprehensively studied (17, 44). The two tyrosines in the Rep protein of ΦX174 are known to function in the initiation and termination steps, respectively, and their cooperation facilitates the reinitiation after one round of replication that produces a large number of phage progeny efficiently (39). The Rep proteins of well-studied bacterial RCR plasmids, including pLS1, pC194, and pT181, contain a one-tyrosine motif III (21). Intriguingly, RepA of pC194 contains a glutamine, which is similar in both function (in the termination step) and position (separated by 3 amino acid residues from the central tyrosine) to the second tyrosine in the Rep protein of ΦX174 (32). Furthermore, it has also been demonstrated for the RepC protein of pT181 that some amino

acids other than the fixed central tyrosine may be involved in the termination step (1). Thus, the additional tyrosine function in termination may be displaced by a negatively charged amino acid (i.e., glutamine), while the central tyrosine function in initiation was conserved in evolution. These results suggest that all these RCR replicons, which have diverse DNA and protein sequences, have a similar replication mechanism and a common evolutionary ancestor.

ACKNOWLEDGMENTS

This work was supported by grants from the Ministry of Science and Technology of China (2004CB719603 and 2006AA09Z401), the National Natural Science Foundation of China (30621005 and 30671141), and the Chinese Academy of Sciences (KSCX2-YW-G-023).

REFERENCES

- Chang, T. L., M. G. Kramer, R. A. Ansari, and S. A. Khan. 2000. Role of individual monomers of a dimeric initiator protein in the initiation and termination of plasmid rolling circle replication. *J. Biol. Chem.* **275**:13529–13534.
- Charlebois, R. L., W. L. Lam, S. W. Cline, and W. F. Doolittle. 1987. Characterization of pHV2 from *Halobacterium volcanii* and its use in demonstrating transformation of an archaeobacterium. *Proc. Natl. Acad. Sci. USA* **84**:8530–8534.
- Cline, S. W., W. L. Lam, R. L. Charlebois, L. C. Schalkwyk, and W. F. Doolittle. 1989. Transformation methods for halophilic archaeobacteria. *Can. J. Microbiol.* **35**:148–152.
- del Solar, G., R. Díaz, and M. Espinosa. 1987. Replication of the streptococcal plasmid pMV158 and derivatives in cell-free extracts of *Escherichia coli*. *Mol. Gen. Genet.* **206**:428–435.
- del Solar, G., R. Giraldo, M. J. Ruiz-Echevarría, M. Espinosa, and R. Díaz-Orejás. 1998. Replication and control of circular bacterial plasmids. *Microbiol. Mol. Biol. Rev.* **62**:434–464.
- del Solar, G., M. Moscoso, and M. Espinosa. 1993. In vivo definition of the functional origin of replication (ori(+)) of the promiscuous plasmid pLS1. *Mol. Gen. Genet.* **237**:65–72.
- del Solar, G. H., A. Puyet, and M. Espinosa. 1987. Initiation signals for the conversion of single stranded to double stranded DNA forms in the streptococcal plasmid pLS1. *Nucleic Acids Res.* **15**:5561–5580.
- 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.
- Fluit, A. C., P. D. Baas, and H. S. Jansz. 1985. The complete 30-base-pair origin region of bacteriophage ΦX174 in a plasmid is both required and sufficient for in vivo rolling-circle DNA replication and packaging. *Eur. J. Biochem.* **149**:579–584.
- Grohmann, E., E. L. Zechner, and M. Espinosa. 1997. Determination of specific DNA strand discontinuities with nucleotide resolution in exponentially growing bacteria harboring rolling circle-replicating plasmids. *FEMS Microbiol. Lett.* **152**:363–369.
- Gros, M. F., H. te Riele, and S. D. Ehrlich. 1987. Rolling circle replication of single-stranded DNA plasmid pC194. *EMBO J.* **6**:3863–3869.
- Gruss, A., and S. D. Ehrlich. 1989. The family of highly interrelated single-stranded deoxyribonucleic acid plasmids. *Microbiol. Rev.* **53**:231–241.
- Hackett, N. R., and S. DasSarma. 1989. Characterization of the small endogenous plasmid of *Halobacterium* strain SB3 and its use in transformation of *H. halobium*. *Can. J. Microbiol.* **35**:86–91.
- Hackett, N. R., M. P. Krebs, S. DasSarma, W. Goebel, U. L. RajBhandary, and H. G. Khorana. 1990. Nucleotide sequence of a high copy number plasmid from *Halobacterium* strain GRB. *Nucleic Acids Res.* **18**:3408.
- Hall, M. J., and N. R. Hackett. 1989. DNA sequence of a small plasmid from *Halobacterium* strain GN101. *Nucleic Acids Res.* **17**:10501.
- Holmes, M. L., F. Pfeifer, and M. L. Dyall-Smith. 1995. Analysis of the halobacterial plasmid pHK2 minimal replicon. *Gene* **153**:117–121.
- 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.
- Iordanescu, S., and S. J. Projan. 1988. Replication termination for staphylococcal plasmids: plasmids pT181 and pC221 cross-react in the termination process. *J. Bacteriol.* **170**:3427–3434.
- Kagramanova, V. K., N. I. Derkacheva, and A. S. Mankin. 1988. The complete nucleotide sequence of the archaeobacterial plasmid pHSB from *Halobacterium*, strain SB3. *Nucleic Acids Res.* **16**:4158.
- Khan, S. A. 2005. Plasmid rolling-circle replication: highlights of two decades of research. *Plasmid* **53**:126–136.

21. Khan, S. A. 1997. Rolling-circle replication of bacterial plasmids. *Microbiol. Mol. Biol. Rev.* **61**:442–455.
22. Lam, W. L., and W. F. Doolittle. 1989. Shuttle vectors for the archaeobacterium *Halobacterium volcanii*. *Proc. Natl. Acad. Sci. USA* **86**:5478–5482.
23. Langeveld, S. A., A. D. van Mansfeld, P. D. Baas, H. S. Jansz, G. A. van Arkel, and P. J. Weisbeek. 1978. Nucleotide sequence of the origin of replication in bacteriophage ΦX174 RF DNA. *Nature* **271**:417–420.
24. Li, Y., H. Xiang, J. Liu, M. Zhou, and H. Tan. 2003. Purification and biological characterization of halocin C8, a novel peptide antibiotic from *Halobacterium* strain AS7092. *Extremophiles* **7**:401–407.
25. Liu, Y., and E. Häggard-Ljungquist. 1994. Studies of bacteriophage P2 DNA replication: localization of the cleavage site of the A protein. *Nucleic Acids Res.* **22**:5204–5210.
26. Marsin, S., and P. Forterre. 1999. The active site of the rolling circle replication protein Rep75 is involved in site-specific nuclease, ligase and nucleotidyl transferase activities. *Mol. Microbiol.* **33**:537–545.
27. Marsin, S., and P. Forterre. 1998. A rolling circle replication initiator protein with a nucleotidyl-transferase activity encoded by the plasmid pGT5 from the hyperthermophilic archaeon *Pyrococcus abyssi*. *Mol. Microbiol.* **27**:1183–1192.
28. Marsin, S., E. Marguet, and P. Forterre. 2000. Topoisomerase activity of the hyperthermophilic replication initiator protein Rep75. *Nucleic Acids Res.* **28**:2251–2255.
29. Mojica, F. J., F. Charbonnier, G. Juez, F. Rodríguez-Valera, and P. Forterre. 1994. Effects of salt and temperature on plasmid topology in the halophilic archaeon *Haloferax volcanii*. *J. Bacteriol.* **176**:4966–4973.
30. Moscoso, M., G. del Solar, and M. Espinosa. 1995. In vitro recognition of the replication origin of pLS1 and of plasmids of the pLS1 family by the RepB initiator protein. *J. Bacteriol.* **177**:7041–7049.
31. Noiro, 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.
32. Noiro-Gros, M. F., V. Bidnenko, and S. D. Ehrlich. 1994. Active site of the replication protein of the rolling circle plasmid pC194. *EMBO J.* **13**:4412–4420.
33. Puyet, A., G. H. del Solar, and M. Espinosa. 1988. Identification of the origin and direction of replication of the broad-host-range plasmid pLS1. *Nucleic Acids Res.* **16**:115–133.
34. Sambrook, J., E. F. Fritsch, and T. Maniatis. 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
35. Santamaría, D., E. Viguera, M. L. Martínez-Robles, O. Hyrien, P. Hernández, D. B. Krimer, and J. B. Schvartzman. 2000. Bi-directional replication and random termination. *Nucleic Acids Res.* **28**:2099–2107.
36. Soler, N., A. Justome, S. Quevillon-Cheruel, F. Lorieux, E. Le Cam, E. Marguet, and P. Forterre. 2007. The rolling-circle plasmid pTN1 from the hyperthermophilic archaeon *Thermococcus nautilus*. *Mol. Microbiol.* **66**:357–370.
37. Sozhamannan, S., P. Dabert, V. Moretto, S. D. Ehrlich, and A. Gruss. 1990. Plus-origin mapping of single-stranded DNA plasmid pE194 and nick site homologies with other plasmids. *J. Bacteriol.* **172**:4543–4548.
38. Sun, C., M. Zhou, Y. Li, and H. Xiang. 2006. Molecular characterization of the minimal replicon and the unidirectional theta replication of pSCM201 in extremely halophilic archaea. *J. Bacteriol.* **188**:8136–8144.
39. van Mansfeld, A. D., H. A. van Teeffelen, P. D. Baas, and H. S. Jansz. 1986. Two juxtaposed tyrosyl-OH groups participate in ΦX174 gene A protein catalysed cleavage and ligation of DNA. *Nucleic Acids Res.* **14**:4229–4238.
40. Ward, D. E., I. M. Revet, R. Nandakumar, J. H. Tuttle, W. M. de Vos, J. van der Oost, and J. DiRuggiero. 2002. Characterization of plasmid pRT1 from *Pyrococcus* sp. strain JT1. *J. Bacteriol.* **184**:2561–2566.
41. Weiner, M. P., G. L. Costa, W. Schoettlin, J. Cline, E. Mathur, and J. C. Bauer. 1994. Site-directed mutagenesis of double-stranded DNA by the polymerase chain reaction. *Gene* **151**:119–123.
42. Zechner, E. L., H. Prüger, E. Grohmann, M. Espinosa, and G. Högenauer. 1997. Specific cleavage of chromosomal and plasmid DNA strands in gram-positive and gram-negative bacteria can be detected with nucleotide resolution. *Proc. Natl. Acad. Sci. USA* **94**:7435–7440.
43. Zhao, A. C., and S. A. Khan. 1996. An 18-base-pair sequence is sufficient for termination of rolling-circle replication of plasmid pT181. *J. Bacteriol.* **178**:5222–5228.
44. Zhou, M., H. Xiang, C. Sun, Y. Li, J. Liu, and H. Tan. 2004. Complete sequence and molecular characterization of pNB101, a rolling-circle replicating plasmid from the haloalkaliphilic archaeon *Natronobacterium* sp. strain AS7091. *Extremophiles* **8**:91–98.
45. Zhou, M., H. Xiang, C. Sun, and H. Tan. 2004. Construction of a novel shuttle vector based on an RCR-plasmid from a haloalkaliphilic archaeon and transformation into other haloarchaea. *Biotechnol. Lett.* **26**:1107–1113.
46. Zuker, M. 2003. Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res.* **31**:3406–3415.