Structure of the Gas Vesicle Plasmid in *Halobacterium halobium*: Inversion Isomers, Inverted Repeats, and Insertion Sequences

WAI-LAP NG, SRINIVAS KOTHAKOTA, AND SHILADITYA DasSARMA*

Department of Microbiology, University of Massachusetts, Amherst, Massachusetts 01003

Received 21 December 1990/Accepted 15 January 1991

*Corresponding author.

**Halobacterium halobium** NRC-1 harbors a 200-kb plasmid, pNRC100, which contains a cluster of genes for synthesis of buoyant gas-filled vesicles. Physical mapping of pNRC100 by using pulsed-field gel electrophoresis showed the presence of a large (35 to 38-kb) inverted repeat (IR) sequence. Inversion isomers of pNRC100 were demonstrated by Southern hybridization analysis using two restriction enzymes, *Afl*II and *Sac*II, that cut asymmetrically within the intervening small single-copy region and the large single-copy region, respectively, but not within the large IRs. No inversion isomers were observed for a deletion derivative of pNRC100 lacking one IR, which suggests that both copies are required for inversion to occur. Additionally, the identities and approximate positions of 17 inversion sequences (ISs) in pNRC100 were determined by Southern hybridization and limited nucleotide sequence analysis across the IS element-target site junctions: ISH2, a 0.5-kb element, was found in four copies; ISH3, a 1.4-kb heterogeneous family of elements, was present in seven copies; ISH8, a 1.4-kb element, was found in five copies; and ISH50, a 1.0-kb element, was present in a single copy. The large IRs terminated at an ISH2 element at one end and an ISH3 element at the other end. pNRC100 is similar in structure to chloroplast and mitochondrial genomes, which contain large IRs and other large halobacterial and prokaryotic plasmids that are reservoirs of IS elements but lack the large IRs.

The genome of the halophilic archaeabacterium *Halobacterium halobium* is notable for its extreme instability (2). The genome consists of two physically separable components, a major G+C-rich fraction and a minor A+T-rich satellite fraction (15). The A+T-rich satellite DNA contains a large multicopy plasmid, named pNRC100 in strain NRC-1, and several minor circular DNAs related to pNRC100 (20, 35). In addition, several other A+T-rich regions have been described which originate from other genomic locations (10, 22). Plasmid pNRC100 and A+T-rich regions in general contain a disproportionate share of insertion sequences (ISs) (20, 22, 28), several of which have been shown to be relatively A+T rich by nucleotide sequence analysis (8, 9, 11, 33, 36, 37). It has been estimated that the *H. halobium* NRC-1 genome harbors about 50 different families of IS elements (28, 29) which mediate frequent DNA rearrangements (5).

Rearrangement of the *H. halobium* genome is responsible for several high-frequency phenotypic variabilities (5). Our recent efforts have been directed at understanding the mechanism of the gas vesicle (or Vac) mutations which are observed at rates of about 1% (6, 7, 14, 20). Initially, phenotypic analysis of these mutants revealed three distinguishable types or classes: those that are partially Vac− (Vac−) and extremely unstable (class I), those that are Vac− but relatively stable (class II), and those that are stable and completely Vac− (class III). Southern hybridization analysis using as a probe the major gas vesicle protein gene, *gvpA*, which is found on pNRC100, showed that class II mutants contained ISs. Nucleotide sequence analysis identified IS elements either within the *gvpA* gene promoter or further upstream, within two open reading frames, designated *gvpD* and *gvpE*, which have a transcriptional orientation opposite that of *gvpA* (Fig. 1A). Class I mutants, in contrast, showed a decrease in the copy number for the *gvpA* gene, and class III mutants contained no *gvpA* copies. We proposed that class I mutants contained two plasmid populations, pNRC100 and a deletion derivative of pNRC100 lacking the *gvpA* gene, while class III mutants were segregants of class I mutants containing only the deleted plasmid derivative (7).

To test this deletion-segregation hypothesis, we first attempted to construct a restriction map of pNRC100 by using pulsed-field gel electrophoresis (20). We found that *Hind*III and *Dra*I restriction fragments fell into either of two separate linkage groups. Interestingly, the two linkage groups could be linked to one another in either relative orientation to form two alternative circular maps (Fig. 1B). In this report, we show that this characteristic of the restriction mapping data reflects 35 to 38-kb inverted repeats in pNRC100 which allow frequent inversion of the intervening single-copy regions. We also document the identity, copy number, and terminal and flanking sequences for several IS elements present on pNRC100.

**MATERIALS AND METHODS**

**Bacterial strains and plasmid DNA.** Wild-type *H. halobium* NRC-1 and its Vac− class III mutant strain SD109 were cultured under previously described conditions (7). Plasmid DNA was prepared from *H. halobium* by a previously described alkaline-sodium dodecyl sulfate (SDS) procedure (20). *Escherichia coli* DH5α and JM103 and plasmid vectors pTZ18R and pTZ19R (Pharmacia) and M13mp18 were used in cloning and subcloning of pNRC100 restriction fragments (18, 19).

**Pulsed-field agarose gel electrophoresis.** The CHEF (counter-clamped homogeneous electric field) agarose gel electrophoresis system (3) was used to resolve large DNA fragments. Electrophoresis was carried out in 0.7% agarose (Bethesda Research Laboratories) gels in 0.5× TBE (0.044 M Tris base, 0.044 M boric acid, 1 mM EDTA) running buffer with circulation at 16°C, 120 Volts, and switching...
FIG. 1. Physical and genetic map of the \( \delta \) isomer of pNRC100. (A) Enlargement of the gas vesicle gene region, with the gvp gene positions and sites of insertion sequences in class II mutants marked. The insertion mutant strain designations (R1, SD106, SD104, SD108, and SD120) are indicated. (B) Circular plasmid map, with locations of large IRS marked by heavy arrows on the outside, the locations of IS elements indicated on the outer circle, and the positions of DraI, HindIII, SfiI, and AflII fragments indicated on the inner circles. The dashed lines radiating from the center of the circle indicate the ambiguity in the original mapping experiments. The approximate position of an origin of replication is marked (ori).

Increment for 30 steps for 23 h from 0.2 to 6 s in Fig. 4 and 10.5 h from 0.05 to 1.5 s in Fig. 5, using an MJ Research power inverter and a Davis-type apparatus (3).

Gel purification, cloning, and subcloning of pNRC100 fragments. A HindIII library of pNRC100 was constructed by isolation of the 13 HindIII fragments larger than 1 kb that are resolved on low-melting-point agarose CHEF gels, purification by phenol extraction and ethanol precipitation, and ligation into HindIII-cleaved and dephosphorylated pTZ19R (18). Recombinant plasmids obtained from this library include a series of plasmids named pNGH1 (Table 1). For most of the pNGH1 constructs, only a single cloned fragment was obtained from each gel-purified pNRC100 fragment out of approximately a dozen recombinants screened. The only exceptions were HindIII-E, which contains two different pNRC100 fragments (E and E') and HindIII-B, for which there exist two polymorphic forms, with and without an ISH2 and an ISH3 element (B and B*, respectively). Two different clones were obtained in each of these cases. Plasmid pDY1 contains a 3.9-kb SstI-EcoRI fragment mapping near the middle of the pNRC100 HindIII E fragment, while pNGHJ13 contains a 1.6-kb EcoRI fragment near the middle of the HindIII J fragment. Both pDY1 and pNGHJ13 lack IS elements. Table 1 summarizes all of the pNRC100 clones and subclones used in this study.

Southern hybridization analysis. DNA was transferred to nitrocellulose (Schleicher & Schuell, Inc.) by the Southern procedure (34) with a 20-min 0.25 N HCl treatment prior to the denaturation step. After transfer, membranes were prehybridized for at least 4 h in 3 × SSC (0.45 M NaCl, 0.045 M sodium citrate, pH 7)–0.5% SDS–5 × Denhardt solution (0.1% [wt/vol] each Ficoll, polyvinylpyrrolidone, and bovine serum albumin)–100 μg of denatured and sheared salmon sperm DNA per ml. Hybridization was carried out overnight in the same solution containing probe. Filters were washed three times with 3 × SSC at 20-min intervals and then once with 0.1 × SSC–0.5% SDS for 30 min. The prehybridization, hybridization, and washes were carried out at 68°C except when the ISH50 oligodeoxyribonucleotide was used as a probe, in which case the temperature was 25°C. For all IS element hybridizations, the high-stringency wash was omitted.

For analysis of inversion isomers, the following plasmids were used as probes after \( ^{32} \)P labeling by nick translation (26): pDY1 for the \( \alpha \) and \( \gamma \) fragments, pNGHH1 for the \( \alpha \) and \( \beta \) fragments, pNGHII for the \( \gamma \) and \( \delta \) fragments, and pNGHH13 for the \( \beta \) and \( \delta \) fragments. The following IS element probes were used. For ISH1, a 1-kb Tth1111 fragment internal to ISH1 was gel purified from a bacterio-opsin gene clone, pSD17 (33). The ISH2 probe was a 170-bp internal BglII fragment cloned in pBR322 (8). The ISH3 probe was a 1,726-bp HindIII-Stul fragment cloned in pKSI (+) (Stratagene) which contains an entire ISH3 element and most of the gvpA gene from H. halobium R1 (6). The ISH50 probe was a 15-nucleotide-long synthetic oligodeoxy-

### Table 1. Cloned pNRC100 fragments

<table>
<thead>
<tr>
<th>Plasmid</th>
<th>Insert size and identity</th>
<th>Insert origin (no. of copies)</th>
</tr>
</thead>
<tbody>
<tr>
<td>pNGHA1</td>
<td>25 kb, HindIII-A</td>
<td>pNRC100 ISH50</td>
</tr>
<tr>
<td>pNGHB1</td>
<td>23 kb, HindIII-B</td>
<td>pNRC100 ISH2(2), ISH3(2)</td>
</tr>
<tr>
<td>pNGHB*1</td>
<td>21 kb, HindIII-B*</td>
<td>pNRC100 ISH2, ISH3</td>
</tr>
<tr>
<td>pNGHC1</td>
<td>19 kb, HindIII-C</td>
<td>pNRC100 ISH3, ISH8</td>
</tr>
<tr>
<td>pNGHD1</td>
<td>17 kb, HindIII-D</td>
<td>pNRC100 ISH8, ISH9</td>
</tr>
<tr>
<td>pNGHD1*</td>
<td>16 kb, HindIII-D*</td>
<td>pNRC100 ISH2, ISH3, ISH8</td>
</tr>
<tr>
<td>pNGHE1</td>
<td>11 kb, HindIII-E</td>
<td>pNRC100 ISH3, ISH8</td>
</tr>
<tr>
<td>pNGHE1*</td>
<td>11 kb, HindIII-E*</td>
<td>pNRC100 ISH2</td>
</tr>
<tr>
<td>pNGHF1</td>
<td>9.2 kb, HindIII-F</td>
<td>pNRC100 ISH3</td>
</tr>
<tr>
<td>pNGHG1</td>
<td>8.0 kb, HindIII-G</td>
<td>pNRC100 ISH3</td>
</tr>
<tr>
<td>pNGHH1</td>
<td>7.2 kb, HindIII-H</td>
<td>pNRC100 ISH3</td>
</tr>
<tr>
<td>pNGHI1</td>
<td>6.4 kb, HindIII-I</td>
<td>pNRC100 ISH3</td>
</tr>
<tr>
<td>pNGHJ1</td>
<td>4.2 kb, HindIII-J</td>
<td>pNRC100 ISH3</td>
</tr>
<tr>
<td>pNGHK1</td>
<td>3.5 kb, HindIII-K</td>
<td>pNRC100 ISH3</td>
</tr>
<tr>
<td>pDY1</td>
<td>3.9 kb, SstI-EcoRI</td>
<td>pNGHE1</td>
</tr>
<tr>
<td>pNGHI13</td>
<td>1.6 kb, EcoRI</td>
<td>pNGHJ1</td>
</tr>
<tr>
<td>pNGHB3</td>
<td>5.6 kb, EcoRI</td>
<td>pNGHB1 ISH2(2), ISH3</td>
</tr>
<tr>
<td>pNGHB3*</td>
<td>5.3 kb, SstI-EcoRI</td>
<td>pNGHB3 ISH2, ISH3</td>
</tr>
<tr>
<td>pNGHB6</td>
<td>3.7 kb, EcoRI</td>
<td>pNGHB1 ISH2* ISH2</td>
</tr>
<tr>
<td>pNGH4A</td>
<td>2.5 kb, EcoRI-HindIII</td>
<td>pNGHE1 ISH3</td>
</tr>
<tr>
<td>pSK1</td>
<td>3.4 kb, EcoRI</td>
<td>pNGHD1* ISH3</td>
</tr>
<tr>
<td>pNGHB3M6</td>
<td>0.96 kb, EcoRV</td>
<td>pNGHB3 ISH2</td>
</tr>
<tr>
<td>pNGHB3M7</td>
<td>0.96 kb, EcoRV</td>
<td>pNGHB3 ISH2</td>
</tr>
<tr>
<td>pNGHB3M12</td>
<td>2.9 kb, BamHI-SalI</td>
<td>pNGHB3 ISH2, ISH3</td>
</tr>
</tbody>
</table>

*All fragments were cloned into the homologous sites usually in the pTZ19R multiple cloning region except for the pDY1 insert, which was cloned into pTZ18R, and the NGHBb312 insert, which was in M13mp18. The 0.96-kb EcoRV fragment was cloned into the M13mp18 SmaI site in both orientations (clones designated NGHBb316 and -7).
ribonucleotide (5'-ATGTCCTGATGCG-3') homologous to a region near the left end of the element (36). The ISH8 probe was an ~1,470-bp BglII fragment containing 128 bp of flanking gvpE sequence cloned in M13mp19 (14). The ISH1, ISH2, ISH3, and ISH8 probes were 32P labeled by nick translation (26), while the ISH50 oligodeoxiribonucleotide was 32P labeled with T4 polynucleotide kinase (18).

**DNA sequence analysis.** DNA sequence analysis was carried out by the chain termination procedure of Sanger et al. (27), using synthetic oligodeoxiribonucleotide primers on double-stranded (pNGHBb32, pNGHC1, pNGHD1, pNGHD1, pNGHE1, pNGHE1, pNGHE1, and pNGHJ1) and single-stranded (NGHBb3M6, M7, and M12) templates. The primers were designed to hybridize near the termini of the IS elements for sequencing across the IS element-target site junctions. The primer sequences were 5'-AAAGAGATCTAGCCA-3' and 5'-ACTGTTGACCCCTAC-3' for ISH2, 5'-GCGGAAACGGCCACC-3' and 5'-AGTTAAAGGAGCTCTGTT-G3' for ISH3, and 5'-GAGGGAAACA GTGTA-3' and 5'-AAAGATCCCAAGCA-3' for ISH8.

**RESULTS**

Plasmid pNRC100 contains a large repeated sequence. The presence of a large repeated sequence in pNRC100 was demonstrated by restriction mapping and Southern hybridization analysis (data not shown). Restriction digests of pNRC100 with HindIII showed the presence of several bands, e.g., HindIII-D and HindIII-G (30), staining with double the intensity of other bands, which suggested that they contained multiple fragments. We performed secondary digests on gel-purified HindIII-D and -G with PstI, which produced fragments totaling 17 kb for D and 8 kb for G. Since these are consistent with the sizes of the two HindIII bands, we concluded that each one contains multiple copies of a single species. Furthermore, only a single fragment was repeatedly recovered from each of the gel-purified HindIII bands in cloning experiments. Next, we showed that the cloned HindIII D and G fragments each probe two different DraI fragments, B (47 kb) and D (28 kb), by Southern hybridization. Since we had previously established a DraI map for pNRC100 (20) with DraI-B and -D mapping to distinct regions, we concluded that pNRC100 contains a large repeat. To illustrate the similarity between DraI-B and -D and HindIII-D and -G, the gel-purified DraI fragments digested with HindIII and PstI were electrophoresed with the cloned HindIII fragments digested with DraI, HindIII, and PstI (Fig. 2, lanes 2 to 5). Almost all restriction fragments in the digests of the cloned HindIII fragments are present in digests of both gel-purified DraI-B and -D except for the fragments mapping to the small HindIII-DraI region at one end of HindIII-D and vector fragments. These results taken together show that there is a large repeated sequence, at least 23 kb, in pNRC100 (Fig. 1B).

The large repeats exist in inverted orientation. To determine the relative orientation of the large repeat, we mapped the ends with respect to the flanking small single-copy (SSC) region and large single-copy (LSC) region. One end of the repeat was mapped by using plasmids containing cloned pNRC100 HindIII fragments B and E, pNGHB1, and pNGHE1, by digestion with HindIII plus EcoRI (Fig. 2, lanes 6 and 7). On the basis of the similar restriction patterns, we concluded that at least 8-kb regions of these HindIII fragments are homologous. Each of these fragments contains one end of the large repeat mapping next to the LSC region, a result consistent with an inverted orientation of the repeats. Similarly, the other end of each of the repeats in HindIII fragments E and J maps near the SSC region containing the gas vesicle gene cluster (7, 14). Homology at this end was shown by comparing the EcoRI-plus-HindIII or Aval-plus-HindIII double digestion patterns of the cloned pNRC100 HindIII E fragment in pNGHE1 and a 2.5-kb EcoRI-HindIII subclone of pNGHE1, pNGH4A, with that of the HindIII J fragment in pNGHJ1 (data not shown). These comparisons together with Southern hybridization analysis suggested that the inverted repeats (IRs) most likely terminate at an ISH3 element at one end and an ISH2 element at the other, a result that was confirmed by DNA sequence analysis (see below). Between the two ends of the IRs, we mapped the three HindIII fragment pairs G, D, and M (and their homologs G', D', and M') in the inverted orientation shown in Fig. 1B (20). The estimated size of the IRs is between 35 and 38 kb.

Plasmid pNRC100 exists in two inversion isomeric forms in vivo. We sought to determine whether both possible inversion isomers of pNRC100, formed by homologous or site-specific recombination between the IRs, occur. First, we screened for restriction enzymes that cleave pNRC100 in the single-copy regions but not within the IRs. Of the 12 restriction enzymes screened, only two such enzymes were found: SfiI, which cleaves twice in the LSC region, and AfII, which cleaves three times within the SSC region. On the basis of restriction mapping data, we predicted that double digestion of pNRC100 with AfII and SfiI would yield fragments of 92 kb (α) and 72 kb (β) in the αβ isomer and 89 kb (β) and 75 kb (γ) in the βγ isomer (Fig. 3). Digestion of pNRC100 with AfII plus SfiI produced four bands having approximately the same sizes as the predicted isomeric fragments, indicating the presence of both isomers (Fig. 4A, lane 1). pNRC100 isomerization was confirmed by Southern hybridization analysis using four labeled clones (pDY1, pNGHJ1, pNGH11, and pNGH13) specific for pairs of AfII-SfiI fragments (α/γ, w/β, ν/β, and β/δ, respectively) (see Table 1 and Materials and Methods). As predicted in Fig. 3, pDY1 hybridized with α and γ, pNGH11 hybridized with α and β, pNGH11 hybridized with γ and δ, and pNGHJ13 hybridized with β and δ (Fig. 4B).

**Inversion isomerization of plasmid pNRC100.** Approx-
mately equal amounts of both pNRC100 isomers were observed in the preparation of pNRC100 DNA used for mapping analysis. This could have resulted from the presence of either two populations of cells, each harboring a different pNRC100 isomer, or a single cell population containing both isomers. To distinguish between these possibilities, plasmid DNA was analyzed from 12 independent cultures of *H. halobium* NRC-1. *AfII* and *SfiI* double digestion of these plasmid preparations resulted in a very similar restriction pattern in all cases, with both pairs of inversion isomer fragments, αδ and βγ (and therefore the two inversion isomers) present in every culture (data not shown).

To determine the importance of the IRs for inversion isomerization, we examined a deletion derivative of plasmid pNRC100 from *H. halobium* SD109, a class III gas vesicle-deficient mutant of wild-type strain NRC-1. The deletion resulted from transposition of the ISH8 element in *HindIII*-C to *HindIII*-D', leaving only one copy of the IR (data not shown). In contrast to the results with pNRC100 (see above), double digestion of the deleted plasmid with *AfII* and *SfiI* resulted in only a single pair of fragments, the δ fragment plus a deleted α fragment derivative (Fig. 4A, lane 2). This result indicates that deletion of one copy of the IR sequence prevents inversion isomerization of the plasmid. The occurrence of both inversion isomers in all NRC-1 cultures but a single isomer in the deletion strain SD109 suggests that maintenance of both isomers in the wild-type occurs by recombination between the IRs.

**ISs in pNRC100.** In addition to the large IRs, pNRC100 harbors many smaller repeated elements, several of which are well-characterized IS elements, e.g., ISH2 (8), ISH4 (5), ISH8 (5), and ISH50 (36). We have mapped the approximate location of 17 IS element copies in pNRC100 by Southern hybridization analysis using IS element-specific probes and further characterized 11 of these IS elements by nucleotide sequence analysis across the termini of the elements and into the flanking sequences by using IS element-specific primers (Fig. 1, 5, and 6).

For initial mapping of IS elements, we hybridized probes specific to each element to Southern blots of pNRC100 digested with *HindIII* (Fig. 5). Further mapping was carried out by Southern hybridization to cloned pNRC100 fragments digested with *HindIII* plus *EcoRI*, *SalI*, or *SstI* (data not shown). A total of 17 element copies were identified: 4 copies of ISH2, 7 copies of ISH3, 5 copies of ISH8, and a single copy of ISH50 (previously called ISH4). Interestingly, a pair of ISH8 elements was found flanking the gas vesicle gene cluster, a structure that resembles a composite transposon. Another ISH8 element maps within both copies of the IRs. A polymorphism was found in the IRs, with one IR copy containing an ISH2 and an ISH3 element not present in the other (Fig. 1B).

Eleven of the IS element copies, including the ISH8 element found in the large IRs, were further characterized by nucleotide sequence analysis across the IS element-target site junctions. For ISH2, the terminal sequences contained perfect 19-bp IRs in all four copies analyzed (8). Two elements, in *HindIII*-E and -B (Fig. 6A), were flanked by 10-bp direct repeats, a target site duplication size characteristic of several other ISH2 insertion sites. For three ISH3

---

**FIG.** 3. Inversion isomerization of pNRC100. The αδ and βγ inversion isomers of pNRC100 contain four predictable *AfII*-SfiI fragments, α, β, γ, and δ, which are indicated by the outer arcs. Labeled clones that are predicted to hybridize with α and γ (pDY1), α and β (pNGHH1), γ and δ (pNGHJ1), and β and δ (pNGHJ13) (short inner arcs) were used to detect the *AfII*-SfiI fragments in the inversion isomers (see Fig. 4). The positions of IS elements are shaded, gvp gene regions are hatched, and the large IRs are indicated by heavy arrows.

**FIG.** 4. Analysis of pNRC100 inversion isomers. (A) CHEF agarose gel electrophoresis of *H. halobium* NRC-1 (lane 1) and SD109 (lane 2) plasmid DNA digested with *AfII* and *SfiI*. α, β, γ, and δ are pNRC100 *AfII*-SfiI fragments indicated in Fig. 3. (B) Southern hybridization of multiple lanes of strain NRC-1 plasmid DNA digested with *AfII* and *SfiI*, using 32P-labeled probes pDY1 (lane 1), pNGHH1 (lane 2), pNGHJ1 (lane 3), and pNGHJ13 (lane 4).
pairs, a member of the heterogeneous ISH5 family of elements first characterized in *Haloflexx volcanii* (11), the terminal sequences contain imperfect 17-bp IRs (Fig. 6B). Two of these elements, in *HindIII*-F and -J, are flanked by 3-bp direct repeats, suggestive of characteristic target site duplications. ISH8 is about 80% similar to the previously described element ISH26-1 (9) but lacks a 321-bp duplication at one end present in the latter (10a). The ends of both ISH8 and ISH26-1 share imperfect homology (16 of 19 bp) with the terminal IRs of ISH2 (Fig. 6A and C). Of the four ISH8 copies characterized, only one is flanked by direct repeats, but the repeat is imperfect, carrying an extra nucleotide in one copy. Sequence heterogeneity is apparent at the left end of the element, which makes these pNRC100 elements are more similar to ISH8 than to ISH26-1.

The identities of the IS elements at the termini of the large IRs established by restriction mapping (see above) were confirmed by the finding of extensive sequence homologies on the IR side and sequence divergence on the single-copy side of the elements. The sequences flanking one end of the ISH3 elements in *HindIII*-J and *HindIII*-E were homologous, confirming the IR-SSC boundary to the ISH3 elements (Fig. 6B). Similarly, sequences flanking one end of the ISH2 copies in *HindIII*-E' and -B were homologous, confirming the IR-LSC boundary to the ISH2 elements (Fig. 6A).

**DISCUSSION**

We have reported on the structure of the large gas vesicle plasmid, pNRC100, in *H. halobium* NRC-1. We have determined (i) the existence and locations of 35- to 38-kb IRs by Southern hybridization and restriction analysis of cloned and gel-purified pNRC100 restriction fragments, (ii) the existence of pNRC100 inversion isomers resulting, most likely, from recombination between the IRs by restriction analysis using *AflII* and *SfiI*, which do not cleave in the large IRs and thus discriminate between the isomers, and (iii) the locations and identities of 17 IS elements in pNRC100 by hybridization and sequence analysis. In addition, we have identified a cluster of genes involved in the synthesis of gas vesicles by the analysis of gas vesicle-deficient mutants (5, 14), and our recent transformation experiments have localized an origin of replication in the *HindIII* C fragment of pNRC100 (Fig. 1B) (unpublished data).

Our results suggest that the large IRs of pNRC100 mediate inversion of the intervening single-copy regions. The gas vesicle mutant SD109 contains a plasmid lacking one IR copy and harbors only a single isomer form. Additionally, cell lines derived from single cells of the wild-type strain NRC-1 always contain both inversion isomers. These two results together make it likely that the IRs mediate high-frequency inversion of the intervening single-copy region unless some other specific mechanism exists for maintenance of both isomers in individual cells.

The presence of large IRs in pNRC100 makes its structure unique among bacterial plasmids but similar to some chloroplast genomes and some mitochondrial genomes. For *Phaseolus vulgaris*, Palmer showed that both inversion isomers are present in the chloroplasts (21). Also, mitochondrial genomes of *Candida albicans* and *Achlya ambisexualis* were shown to contain large IRs and occur in two inversion isomeric forms (13, 31). However, rRNA genes mapping within the IRs of the organellar genomes are absent in pNRC100 (our unpublished results), suggesting the similarity in the replicon structure does not reflect homology. Interestingly, another similarity has recently been reported between chloroplasts and some halophilic archaeabacteria, the presence of ribulose bisphosphate carboxylase activity (1).

Two possible functions that have been hypothesized for large IRs are in genome stability by minimizing sequence variation in the IRs via copy correction (21) and in plasmid amplification (4). For amplification, a replicating θ structure could undergo inversion, converting the two converging replication forks into two nonconverging forks (that would move in the same direction). Multimeric forms of the plasmid would result from a single initiation event. For *H. halobium*, amplification of pNRC100 would increase gas vesicle gene copy number and expression level, resulting in increased cell buoyancy.

Many inverting elements have been described from other bacteria and phage (25). In general, these elements are only 1 to 2 kb in size and their inversion is known to regulate the expression of two alternative genes. Recombination is catalyzed by site-specific recombinases acting at sites within the IRs. A similar element has also been described in an *H. halobium* phage, a FH derivative, which carries two copies of the element ISH1.8 in inverted orientation and mediates inversion of the intervening L region (30). The function of this inverting element, if any, is unknown.

Several large bacterial plasmids like pNRC100 have been shown to be reservoirs of IS elements, with characteristic high-frequency DNA rearrangements. For example, Leslie and co-workers have described plasmids in *Pseudomonas cepacia* strains, e.g., pTGL6, which harbor many IS element copies mediating insertions, deletions, and replicon fusions at high frequencies (16). Similarly, resistance factors in *E. coli* such as R100 have multiple copies of IS elements and transposons (17). Pfeifer and others have described a plasmid, pH11, in an *H. halobium* strain related to NRC-1 which also contains gas vesicle genes and large numbers of IS elements (12, 23, 24, 32). However, pH11 is clearly distinct from pNRC100, apparently lacking the large IRs and inversion isomers.

Knowledge about the positions and sequences of the IS
elements in pNRC100 has shed light on the evolution of this unstable plasmid. For example, the large IRs terminate at IS elements, ISH3 at one end and ISH2 at the other. Presumably, these IS elements were involved in the duplication resulting in the IRs, perhaps by a mechanism involving multiple recombination events within IS elements. We can also recognize IS elements and transposons that have not been involved in rearrangements after insertion into pNRC100 by the presence of characteristic-size flanking direct repeats. For example, the ISH2 copy in HindIII-E' and one of two in HindIII-B, and the ISH3 copies in HindIII-J and -F, fall into this category. However, the putative composite transposon containing the gvp gene cluster lacks flanking target site duplications. Other composite transposons also could not be identified on the basis of target site duplications. They may have been deleted, however, by further transposition or recombination.

Our study of the structure of pNRC100 has resulted in the finding of several interesting features, including a large number of repeated sequences. These repeats apparently mediate DNA rearrangements at high frequencies and cause gas vesicle mutations and inversion isomerization. Other DNA rearrangements are also apparent in pNRC100, and knowledge about the precise restriction map of pNRC100 will now allow a thorough molecular analysis of these alterations. Additionally, it should now be possible to investigate the function of the rearrangements, including the role of the pNRC100 IRs in plasmid stability and gene amplification.

ACKNOWLEDGMENTS

This work was supported by Public Health Service grant GM41980 from the National Institutes of Health and by grant DMB-8703486 from the National Science Foundation.

We thank T. G. Lessie, N. R. Hackett, and the reviewers for critically reading the manuscript.
REFERENCES