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

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 *Sfi*I, 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 insertion sequences (IS) in pNRC100 were determined by Southern hybridization and limited nucleotide sequence analysis across the IS element-target site junctions: IS*H2*, a 0.5-kb element, was found in four copies; IS*H3*, a 1.4-kb heterogeneous family of elements, was present in seven copies; IS*H8*, a 1.4-kb element, was found in five copies; and IS*H50*, a 1.0-kb element, was present in a single copy. The large IRs terminated at an IS*H2* element at one end and an IS*H3* 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 archaebacterium *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$^{\delta-}$) and extremely unstable (class I), those that are Vac$^{\delta-}$ and 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* gene 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 *Hin*dIII 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 (contour-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
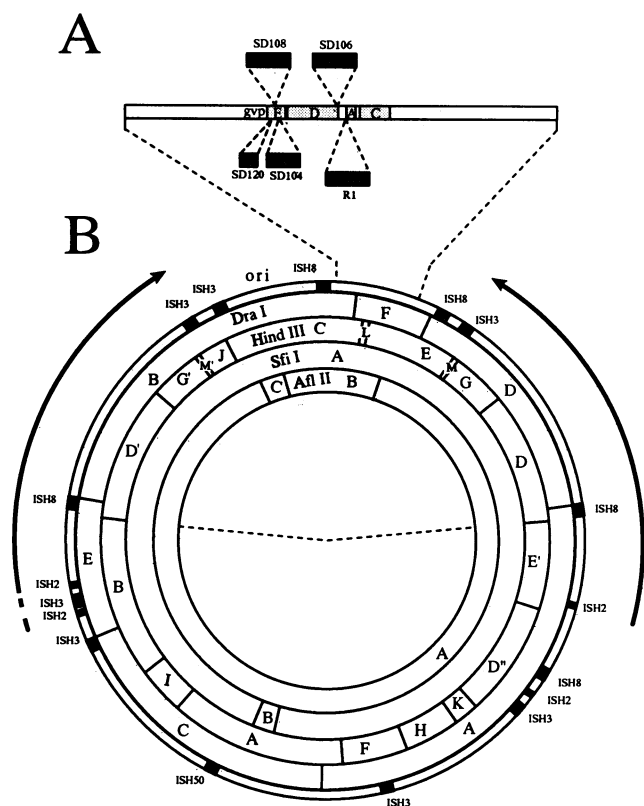
---

* Corresponding author.

A

B

FIG. 1. Physical and genetic map of the αδ 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 *Dra*I, *Hind*III, *Sfi*I, and *Afl*II 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*).

TABLE 1. Cloned pNRC100 fragments

Plasmid[a]	Insert size and identity	Insert origin	IS element(s) (no. of copies)
pNGHA1	25 kb, *Hind*III-A	pNRC100	IS*H50*
pNGHB1	23 kb, *Hind*III-B	pNRC100	IS*H2*(2) IS*H3*(2)
pNGHB*1	21 kb, *Hind*III-B*	pNRC100	IS*H2*, IS*H3*
pNGHC1	19 kb, *Hind*III-C	pNRC100	IS*H3*, IS*H8*
pNGHD1	17 kb, *Hind*III-D	pNRC100	IS*H8*
pNGHD"1	16 kb, *Hind*III-D"	pNRC100	IS*H2*, IS*H3*, IS*H8*
pNGHE1	11 kb, *Hind*III-E	pNRC100	IS*H3*, IS*H8*
pNGHE'1	11 kb, *Hind*III-E'	pNRC100	IS*H2*
pNGHF1	9.2 kb, *Hind*III-F	pNRC100	IS*H3*
pNGHG1	8.0 kb, *Hind*III-G	pNRC100	
pNGHH1	7.2 kb, *Hind*III-H	pNRC100	
pNGHI1	6.4 kb, *Hind*III-I	pNRC100	
pNGHJ1	4.2 kb, *Hind*III-J	pNRC100	IS*H3*
pNGHK1	3.5 kb, *Hind*III-K	pNRC100	
pDY1	3.9 kb, *Sst*I-*Eco*RI	pNGHE1	
pNGHJ13	1.6 kb, *Eco*RI	pNGHJ1	
pNGHBb3	5.6 kb, *Eco*RI	pNGHB1	IS*H2*(2), IS*H3*
pNGHBb32	3.3 kb, *Sal*I-*Eco*RI	pNGHBb3	IS*H2*, IS*H3*
pNG9B61	3.7 kb, *Eco*RI	pNGHB*1	IS*H2*
pNGH4A	2.5 kb, *Eco*RI-*Hind*III	pNGHE1	IS*H3*
pSK1	3.4 kb, *Eco*RI	pNGHD"1	IS*H3*
NGHBb3M6	0.96 kb, *Eco*RV	pNGHBb3	IS*H2*
NGHBb3M7	0.96 kb, *Eco*RV	pNGHBb3	IS*H2*
NGHBb3M12	2.9 kb, *Bam*HI-*Sal*I	pNGHBb3	IS*H2*, IS*H3*

[a] 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 NGHBb3M12 insert, which was in M13mp18. The 0.96-kb *Eco*RV fragment was cloned into the M13mp18 *Sma*I site in both orientations (clones designated NGHBb3M6 and -7).

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 *Hind*III library of pNRC100 was constructed by isolation of the 13 *Hind*III 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 *Hind*III-cleaved and dephosphorylated pTZ19R (18). Recombinant plasmids obtained from this library include a series of plasmids named pNGHn1 (Table 1). For most of the pNGHn1 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 *Hind*III-E, which contains two different pNRC100 fragments (E and E') and *Hind*III-B, for which there exist two polymorphic forms, with and without an IS*H2* and an IS*H3* element (B and B*, respectively). Two different clones were obtained in each of these cases. Plasmid pDY1 contains a 3.9-kb *Sst*I-*Eco*RI fragment mapping near the middle of the pNRC100 *Hind*III E fragment, while pNGHJ13 contains a 1.6-kb *Eco*RI fragment near the middle of the *Hind*III 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 IS*H50* 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 α and γ fragments, pNGHH1 for the α and β fragments, pNGHI1 for the γ and δ fragments, and pNGHJ13 for the β and δ fragments. The following IS element probes were used. For IS*H1*, a 1-kb *Tth*111I fragment internal to IS*H1* was gel purified from a bacterio-opsin gene clone, pSD17 (33). The IS*H2* probe was a 170-bp internal *Bgl*II fragment cloned in pBR322 (8). The IS*H3* probe was a 1,726-bp *Hind*III-*Stu*I fragment cloned in pKS(+) (Stratagene) which contains an entire IS*H3* element and most of the *gvpA* gene from *H. halobium* R1 (6). The IS*H50* probe was a 15-nucleotide-long synthetic oligodeoxy-

ribonucleotide (5′-ATGTCTTGATTGGCA-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 ^{32}P labeled by nick translation (26), while the ISH50 oligodeoxyribonucleotide was ^{32}P 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 oligodeoxyribonucleotide primers on double-stranded (pNGHBb32, pNGHC1, pNGHD1, pNGHD″1, pSK1, pNGHE1, pNGHE′1, pNGHF1, 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′-AAAGAGATCTCAGCC-3′ and 5′-ACTGTTCAGCCCT AC-3′ for ISH2, 5′-CGCGAACCGGCCACC-3′ and 5′-AGT TAAGGAGCTGGT-3′ for ISH3, and 5′-GAGGGAAACA GTGTA-3′ and 5′-AAAGATCCACAAGCA-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 (20), 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 pNGHE′1, 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
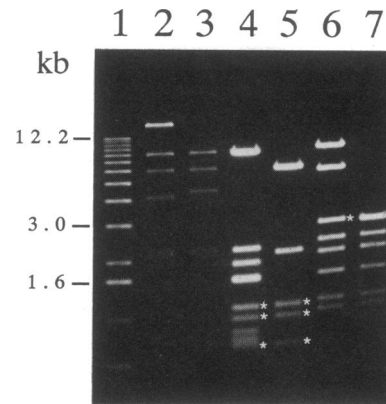

FIG. 2. Mapping of large IRs in pNRC100. An agarose gel of gel-purified fragments DraI-B (lane 2) and DraI-D (lane 3) digested with HindIII and PstI, pNGHD1 (lane 4) and pNGHG1 (lane 5) digested with DraI, HindIII, and PstI, and pNGHB1 (lane 6) and pNGHE′1 (lane 7) digested with EcoRI and HindIII is shown. Vector fragments are indicated by asterisks to the right of the bands. Lane 1 contains molecular weight markers.

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 AvaI-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 AflII, which cleaves three times within the SSC region. On the basis of restriction mapping data, we predicted that double digestion of pNRC100 with AflII and SfiI would yield fragments of 92 kb (α) and 72 kb (δ) in the αδ isomer and 89 kb (β) and 75 kb (γ) in the βγ isomer (Fig. 3). Digestion of pNRC100 with AflII plus SfiI produced four bands having approximately the same sizes as the predicted isomer fragments, indicating the presence of both isomers (Fig. 4A, lane 1). pNRC100 isomerization was confirmed by Southern hybridization analysis using four labeled clones (pDY1, pNGHH1, pNGHI1, and pNGHJ13) specific for pairs of AflII-SfiI fragments (α/γ, α/β, γ/δ, and β/δ, respectively) (see Table 1 and Materials and Methods). As predicted in Fig. 3, pDY1 hybridized with α and γ, pNGHH1 hybridized with α and β, pNGHI1 hybridized with γ and δ, and pNGHJ13 hybridized with β and δ (Fig. 4B).

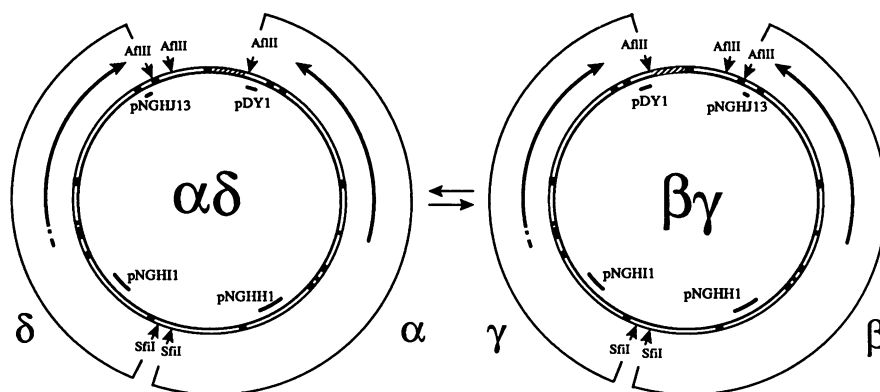**Inversion isomerization of plasmid pNRC100.** Approxi-

FIG. 3. Inversion isomerization of pNRC100. The αδ and βγ inversion isomers of pNRC100 contain four predictable *Afl*II-*Sfi*I 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 (pNGHI1), and β and δ (pNGHJ13) (short inner arcs) were used to detect the *Afl*II-*Sfi*I 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.

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. *Afl*II and *Sfi*I double digestion of these plasmid preparations resulted in a very similar restriction pattern in all cases, with both pairs of inversion isomer fragments, α/δ and β/γ (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
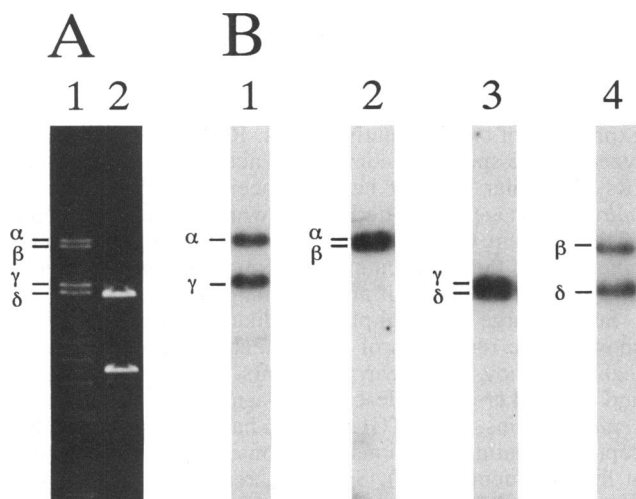

FIG. 4. Analysis of pNRC100 inversion isomers. (A) CHEF agarose gel electrophoresis of *H. halobium* NRC-1 (lane 1) and SD109 (lane 2) plasmid DNA digested with *Afl*II and *Sfi*I. α, β, γ, and δ are pNRC100 *Afl*II-*Sfi*I fragments indicated in Fig. 3. (B) Southern hybridization of multiple lanes of strain NRC-1 plasmid DNA digested with *Afl*II and *Sfi*I, using ³²P-labeled probes pDY1 (lane 1), pNGHH1 (lane 2), pNGHI1 (lane 3), and pNGHJ13 (lane 4).

resulted from transposition of the IS*H8* element in *Hin*dIII-C to *Hin*dIII-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 *Afl*II and *Sfi*I 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., IS*H2* (8), IS*H3* (5), IS*H8* (5), and IS*H50* (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 *Hin*dIII (Fig. 5). Further mapping was carried out by Southern hybridization to cloned pNRC100 fragments digested with *Hin*dIII plus *Eco*RI, *Sal*I, or *Sst*I (data not shown). A total of 17 element copies were identified: 4 copies of IS*H2*, 7 copies of IS*H3*, 5 copies of IS*H8*, and a single copy of IS*H50* (previously called IS*H4*). Interestingly, a pair of IS*H8* elements was found flanking the gas vesicle gene cluster, a structure that resembles a composite transposon. Another IS*H8* element maps within both copies of the IRs. A polymorphism was found in the IRs, with one IR copy containing an IS*H2* and an IS*H3* element not present in the other (Fig. 1B).

Eleven of the IS element copies, including the IS*H8* element found in the large IRs, were further characterized by nucleotide sequence analysis across the IS element-target site junctions. For IS*H2*, the terminal sequences contained perfect 19-bp IRs in all four copies analyzed (8). Two elements, in *Hin*dIII-E′ and -B (Fig. 6A), were flanked by 10-bp direct repeats, a target site duplication size characteristic of several other IS*H2* insertion sites. For three IS*H3*
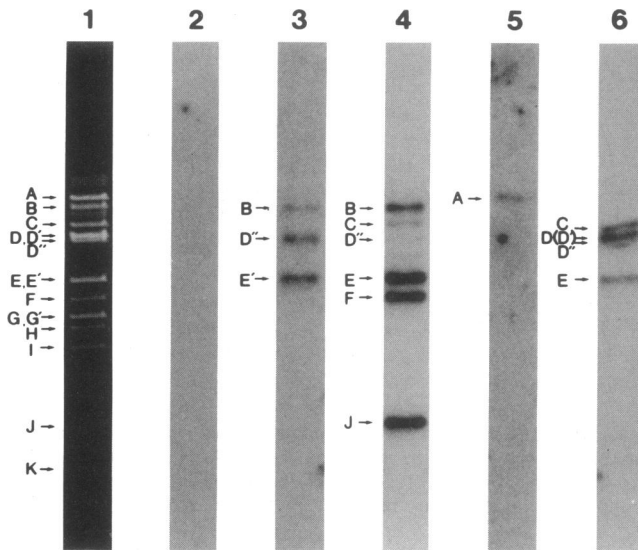
FIG. 5. Southern hybridization analysis using IS element probes. pNRC100 was digested with HindIII, fractionated on a CHEF gel (ethidium-stained lane; lane 1), and hybridized with probes specific for ISH1 (lane 2), ISH2 (lane 3), ISH3 (lane 4), ISH50 (lane 5), and ISH8 (lane 6). The identity of each hybridizing band is indicated.

copies, a member of the heterogeneous ISH51 family of elements first characterized in Haloferax 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 most 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 archaebacteria, 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 ΦH 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, Lessie 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, pHH1, 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, pHH1 is clearly distinct from pNRC100, apparently lacking the large IRs and inversion isomers.

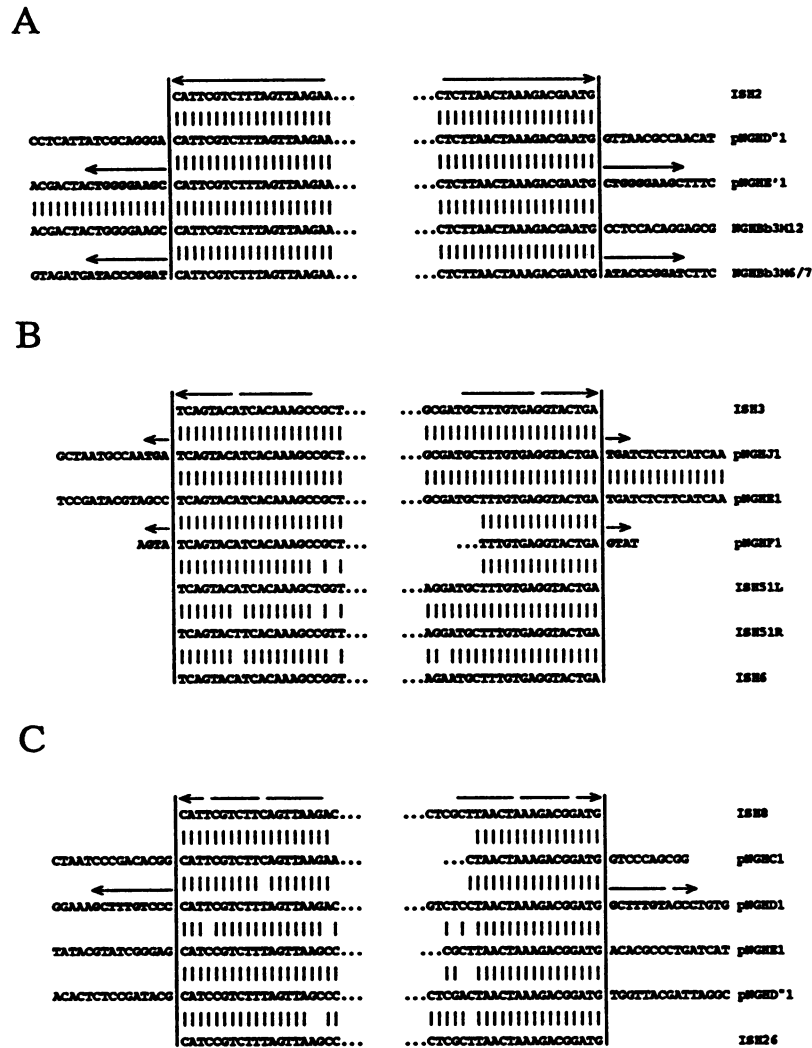Knowledge about the positions and sequences of the IS

A

B

C

FIG. 6. Nucleotide sequence of the IS*H2*, IS*H3*, IS*H8*, and related element termini and flanking sequences. (A) IS*H2* termini sequences (8) are between the vertical lines, and flanking sequences are on the outside. (B) IS*H3* termini (5) and the related elements IS*H6* (10a) and IS*H51* (36) sequences and flanking sequences. (C) IS*H8* termini and the related element IS*H26* (10) sequences and flanking sequences. The recombinant plasmids used to generate the nucleotide sequences are shown to the right. Homologous sequences are indicated by short vertical lines. IS element terminal IRs and flanking direct repeats are indicated by arrows.

elements in pNRC100 has shed light on the evolution of this unstable plasmid. For example, the large IRs terminate at IS elements, IS*H3* at one end and IS*H2* 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 IS*H2* copy in *Hind*III-E' and one of two in *Hind*III-B, and the IS*H3* copies in *Hind*III-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.

REFERENCES

1. **Altekar, W., and R. Rajagopalan.** 1990. Ribulose bisphosphate carboxylase activity in halophilic *Archaebacteria*. Arch. Microbiol. **153**:169–174.
2. **Charlebois, R. L., and W. F. Doolittle.** 1989. Transposable elements and genome structure in halobacteria, p. 297–307. *In* D. E. Berg and M. M. Howe (ed.), Mobile DNA. American Society for Microbiology, Washington, D.C.
3. **Chu, G., D. Vollrath, and R. W. Davis.** 1986. Separation of large DNA molecules by contour-clamped homogeneous electric fields. Science **234**:1582–1585.
4. **Cox, M. M.** 1988. FLP site-specific recombination system of *Saccharomyces cerevisiae*, p. 429–443. *In* R. Kucherlapati and G. R. Smith (ed.), Genetic recombination. American Society for Microbiology, Washington, D.C.
5. **DasSarma, S.** 1989. Mechanisms of genetic variability in *Halobacterium halobium*: the purple membrane and gas vesicle mutations. Can. J. Microbiol. **35**:65–72.
6. **DasSarma, S., T. Damerval, J. G. Jones, and N. Tandeau de Marsac.** 1987. A plasmid-encoded gas vesicle protein gene in a halophilic archaebacterium. Mol. Microbiol. **1**:365–370.
7. **DasSarma, S., J. T. Halladay, J. G. Jones, J. W. Donovan, P. J. Giannasca, and N. Tandeau de Marsac.** 1988. High-frequency mutations in a plasmid-encoded gas vesicle gene in *Halobacterium halobium*. Proc. Natl. Acad. Sci. USA **85**:6861–6865.
8. **DasSarma, S., U. L. RajBhandary, and H. G. Khorana.** 1983. High-frequency spontaneous mutation in the bacterio-opsin gene in *Halobacterium halobium* is mediated by transposable elements. Proc. Natl. Acad. Sci. USA **80**:2201–2205.
9. **Ebert, K., C. Hanke, H. Delius, W. Goebel, and F. Pfeifer.** 1987. A new insertion element, IS*H26*, from *Halobacterium halobium*. Mol. Gen. Genet. **206**:81–87.
10. **Ebert, K., and W. Goebel.** 1985. Conserved and variable regions in the chromosomal and extrachromosomal DNA of halobacteria. Mol. Gen. Genet. **200**:96–102.
10a. **Halladay, J. T., and S. DasSarma.** Unpublished data.
11. **Hofman, J. D., L. C. Schalkwyk, and W. F. Doolittle.** 1986. IS*H51*: a large, degenerate family of insertion sequence-like elements in the genome of the archaebacterium, *Halobacterium volcanii*. Nucleic Acids Res. **14**:6983–7000.
12. **Horne, M., and F. Pfeifer.** 1989. Expression of two gas vacuole protein genes in *Halobacterium halobium* and other related species. Mol. Gen. Genet. **218**:437–444.
13. **Hudspeth, M. E. S., D. S. Shumard, C. J. R. Bradford, and L. I. Grossman.** 1983. Organization of *Achlya* mtDNA: a population with two orientations and a large inverted repeat containing the rRNA genes. Proc. Natl. Acad. Sci. USA **80**:142–146.
14. **Jones, J. G., N. R. Hackett, J. T. Halladay, D. J. Scothorn, C.-F. Yang, W.-L. Ng, and S. DasSarma.** 1989. Analysis of insertion mutants reveals two new genes in the pNRC100 gas vesicle gene cluster of *Halobacterium halobium*. Nucleic Acids Res. **17**:7785–7793.
15. **Joshi, J. G., W. R. Guild, and P. Handler.** 1963. The presence of two species of DNA in some halobacteria. J. Mol. Biol. **6**:34–38.
16. **Lessie, T. G., M. S. Wood, A. Byrne, and A. Ferrante.** 1990. Transposable gene-activating elements in *Pseudomonas cepacia*, p. 279–291. *In* S. Silver, A. M. Chakrabarty, B. Iglewski, and S. Kaplan (ed.), *Pseudomonas*: biotransformations, pathogenesis, and evolving biotechnology. American Society for Microbiology, Washington, D.C.
17. **Lewin, B.** 1977. Gene expression, vol. 3. Plasmids and phages. John Wiley & Sons, New York.
18. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
19. **Messing, J.** 1982. New M13 vectors for cloning. Methods Enzymol. **101**:20–78.
20. **Ng, W.-L., and S. DasSarma.** *In* General and applied aspects of halophilic microorganisms, in press. Plenum Publishing Corp., New York.
21. **Palmer, J. D.** 1983. Chloroplast DNA exists in two orientations. Nature (London) **301**:92–93.
22. **Pfeifer, F., and M. Betlach.** 1985. Genome organization in *Halobacterium halobium*: a 70 kb island of more (AT) rich DNA in the chromosome. Mol. Gen. Genet. **198**:449–455.
23. **Pfeifer, F., and U. Blaseio.** 1989. Insertion elements and deletion formation in a halophilic archaebacterium. J. Bacteriol. **171**:5135–5140.
24. **Pfeifer, F., G. Weidinger, and W. Goebel.** 1981. Genetic variability in *Halobacterium halobium*. J. Bacteriol. **145**:375–381.
25. **Plasterk, R. H. A., and P. van de Putte.** 1984. Genetic switches by DNA inversions in prokaryotes. Biochim. Biophys. Acta **782**:111–119.
26. **Rigby, P. W. J., M. Dieckmann, C. Rhodes, and P. Berg.** 1977. Labeling deoxyribonucleic acid to high specific activity in vitro by nick translation with DNA polymerase I. J. Mol. Biol. **113**:237–251.
27. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5467.
28. **Sapienza, C., and W. F. Doolittle.** 1982. Unusual physical organization of the *Halobacterium* genome. Nature (London) **295**:384–389.
29. **Sapienza, C., M. R. Rose, and W. F. Doolittle.** 1982. High-frequency genomic rearrangements involving archaebacterial repeat sequence elements. Nature (London) **299**:182–185.
30. **Schnabel, H.** 1984. An immune strain of *Halobacterium halobium* carries the invertible L segment of phage ΦH as a plasmid. Proc. Natl. Acad. Sci. USA **81**:1017–1020.
31. **Shaw, J. A., W. B. Troutman, B. A. Lasker, M. M. Mason, and W. S. Riggsby.** 1989. Characterization of the inverted duplication in the mitochondrial DNA of *Candida albicans*. J. Bacteriol. **171**:6353–6356.
32. **Simon, R. D.** 1978. *Halobacterium halobium* strain 5 contains a plasmid which is correlated with the presence of gas vacuoles. Nature (London) **273**:314–317.
33. **Simsek, M., S. DasSarma, U. L. RajBhandary, and H. G. Khorana.** 1982. A transposable element from *Halobacterium halobium* which inactivates the bacteriorhodopsin gene. Proc. Natl. Acad. Sci. USA **79**:7268–7272.
34. **Southern, E. M.** 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. **98**:503–517.
35. **Weidinger, G., G. Klotz, and W. Goebel.** 1979. A large plasmid from *Halobacterium halobium* carrying genetic information for gas vacuole formation. Plasmid **2**:377–386.
36. **Xu, W.-L., and W. F. Doolittle.** 1983. Structure of the archaebacterial transposable element IS*H50*. Nucleic Acids Res. **11**:4195–4199.
37. **Zvyaga, T. A., S. A. Zozulya, and S. O. Guriev.** 1987. The nucleotide sequence of the archaebacterial transposable genetic element IS*HS1*. Bioorg. Khim. **13**:1351–1357.