Proc. Natl. Acad. Sci. USA
Vol. 79, pp. 7268–7272, December 1982
Biochemistry

A transposable element from *Halobacterium halobium* which inactivates the bacteriorhodopsin gene

(archaebacteria/duplicative transposition/site-specific insertion/open reading frame/transposable element specific RNA)

MEHMET SIMSEK*†‡, SHILADITYA DASSARMA*, UTTAM L. RAJBHANDARY*, AND H. GOBIND KHORANA*†

Departments of *Biology and †Chemistry, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

Contributed by H. Gobind Khorana, September 24, 1982

ABSTRACT **We describe the characterization of a transposable element from an archaebacterium. The bacteriorhodopsin genes from the wild-type and two mutant *Halobacterium halobium* strains have been cloned as *Bam*HI fragments in pBR322. The cloned DNA fragments from the two mutants both contain a 1.1-kilobase-pair insertion sequence (ISH*1*) near the NH_2 terminus of the bacteriorhodopsin coding sequence. ISH*1* is present in the two mutants in an identical palindromic site but in opposite orientations. The complete sequence of ISH*1* has been determined; it is 1,118 nucleotides long, it has 8-base-pair interrupted inverted repeats at the ends, and it duplicates an 8-base-pair (A-G-T-T-A-T-T-G) target sequence upon insertion. As for most eukaryotic and some prokaryotic transposable elements, the sequence of the ISH*1* begins with T-G and ends in C-A. ISH*1* contains an open reading frame 810 nucleotides long and codes for an RNA approximately 900 nucleotides long. The copy number of ISH*1* ranges from one to five or more in different *H. halobium* strains. In at least one of the strains, one copy of ISH*1* is present also on a plasmid DNA.**

We have previously reported on the identification and characterization of the gene for the purple membrane (Pum) protein, bacteriorhodopsin (BR), from *Halobacterium halobium*, an archaebacterium (1). During the cloning of this gene, a variant clone (pMSb1) was also obtained; this most likely represented a minor spontaneous mutant (Pum^-) in the culture used for isolation of the DNA. We have now isolated another spontaneous Pum^- mutant (SD17) and cloned the BR gene from this strain. In this paper, we report on the cloning and characterization of these two mutant BR genes. We show that the two mutants contain a single transposable element (ISH*1*), which has inserted into the same site within the BR gene but in opposite orientations. The nucleotide sequence of ISH*1* and several of its features are described and these are compared with those of insertion elements from other systems. The presence of "mobile elements" in halobacteria has been suggested previously by Sapienza and Doolittle (2) and Pfeifer *et al.* (3). The present work gives a detailed characterization of a transposable element from halobacteria and demonstrates its involvement in inactivation of a specific gene.

While a large number of transposable elements have been characterized from eukaryotes and prokaryotes, ISH*1* represents the only such element to be characterized from archaebacteria.

MATERIALS AND METHODS

Strains. *H. halobium* strain S9 was obtained from W. Stoeckenius and strains NRC-1 and R1 were gifts of F. Doolittle. These strains are all Pum^+. The Pum^- strains R1mR and L33 were from Y. Mukohata and J. Lanyi, respectively. The strain SD17, also Pum^-, is a spontaneous orange mutant of the purple strain S9 isolated in this laboratory.

Materials. Restriction enzymes and phage T4 DNA polymerase were purchased from Bethesda Research Laboratories. T4 DNA ligase and polynucleotide kinase were either from Bethesda Research Laboratories or purified as described (4). Deoxyribonucleoside [α-^{32}P]triphosphates were from Amersham and the inorganic [^{32}P]phosphate used was from New England Nuclear.

DNA Probes for Hybridization. The preparation of a *Bam*HI fragment [≈80 base pairs (bp)] of cloned cDNA corresponding to the amino-terminal region of the BR gene has been described (1, 5). The probe for ISH*1* was a 1-kbp *Tth* I fragment from within the pMSb1 element (Fig. 1*C*). The labeling of double-stranded DNA probes was carried out by O'Farrell's procedure (6). Single-stranded DNA probe specific for ISH*1* was prepared by using M13 mp7 recombinant virion DNA. A 269-bp *Hpa* II fragment (nucleotides 338–607 in Fig. 4 below) derived from ISH*1* was cloned in the *Acc* I site of M13 mp7 replicative form (RF) DNA (7). Due to the extensive symmetry at the universal cloning site, the single-stranded viral DNA, complementary to the open reading frame (ORF), could be cut with *Bam*HI. ^{32}P labeling of the probe was accomplished by filling in the cohesive end with T4 DNA polymerase and [α-^{32}P]dGTP.

Cloning. *Bam*HI fragments of *H. halobium* DNA were inserted into the *Bam*HI site of pBR322 DNA as described (1). Colonies were screened by the procedure of Hanahan and Meselson (8).

Southern Blot Hybridization. Genomic DNA (2 μg per lane) or plasmid DNA (0.5 μg per lane) was digested with appropriate restriction enzymes and the digests were fractionated on 1% agarose slab gels and blotted to nitrocellulose (9). Filters were baked at 80°C for 2 hr under reduced pressure and processed as described by Jeffreys and Flavel (10). The filters were pretreated overnight at 68°C with 3× NaCl/Cit (1× NaCl/Cit is 0.15 M NaCl/0.015 M sodium citrate) containing 0.2% bovine serum albumin, 0.2% polyvinylpyrrolidone, and 0.2% Ficoll (solution A) and denatured sonicated calf thymus DNA at 100 μg/ml. Hybridization with the ^{32}P-labeled probe (1.5–6 × 10^6 dpm) was carried out overnight at 68°C in the same solution. The filters were washed (4 times, 15 min each) at 68°C with solution A and then for 30 min at 68°C with 0.1× NaCl/Cit containing 0.1% sodium dodecyl sulfate and autoradiographed at −80°C with an intensifying screen.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U. S. C. §1734 solely to indicate this fact.

Abbreviations: BR, bacteriorhodopsin; bp, base pair(s); kb, kilobase(s); ORF, open reading frame; IS, insertion sequence; ISH, insertion sequence from *Halobacterium halobium*; Pum, purple membrane phenotype; Tn, transposable element; RF, replicative form; NaCl/Cit, 0.15 M NaCl/0.015 M sodium citrate.

‡ Present address: Dept. of Biological Sciences, Middle East Technical Univ., Ankara, Turkey.

RNA Electrophoresis, Blotting, and Hybridization. *H. halobium* cells were grown to midlogarithmic phase and total RNA was prepared and centrifuged into a CsCl cushion as described (5). Electrophoresis of the RNA was carried out on a 1.25% agarose glyoxal gel as described (11). RNA was blotted onto nitrocellulose filter in 20× NaCl/Cit (11) and hybridized with DNA probes as described by Alwine *et al.* (12).

In vivo ^{32}P-labeled RNA used for Southern hybridization was isolated from *H. halobium* cells that had been grown in a medium containing [^{32}P]phosphate. Total RNA prepared as described (5) was depleted of BR mRNA by passage through a BR cDNA column (this preparation was provided by Alokes Majumdar).

Heteroduplexes for Electron Microscopy. Heteroduplexed DNAs were prepared for electron microscopy as described by Davis *et al.* (13). pMSb1 and pSD17 DNAs (1 μg each) were cleaved with *Bam*HI. The fragments produced from the two plasmids were mixed, denatured with 0.15 M NaOH, and annealed in a solution containing 300 mM Tris·HCl (pH 7.5) and 50% (vol/vol) formamide. For electron microscopy, the DNA was spread in presence of cytochrome *c*, stained with uranyl acetate, and shadowed with platinum.

DNA Sequence Analysis. Sequences were determined by using the method of Maxam and Gilbert (14) and the chain termination method of Sanger *et al.* (15).

RESULTS

The Two Mutant BR Genes that Contain a 1.1-kbp Insert (ISH*1*) Within the Gene. We have previously reported the cloning of the BR gene from *H. halobium* strain S9 as a 5.3-kbp *Pst* I fragment (1). In further attempts to generate a *Bam*HI clone bank from the same *Halobacterium* strain, a clone, designated pMSb1, was obtained. This contained a 7.5-kbp *Bam*HI fragment. In Fig. 1*A* are shown the results of hybridization of BR gene-specific probe to *Bam*HI digests of S9 genomic DNA (lane 1) and pMSb1 DNA (lane 2). The genomic DNA yields two bands, a major one 6.4 kbp long and a minor one 7.5 kbp long. Thus, pMSb1 represents the cloning of the BR gene from a minor spontaneous Pum$^-$ mutant of S9. The BR gene from the second spontaneous Pum$^-$ mutant, SD17 (*Materials and Methods*), was also cloned as a 7.5-kbp *Bam*HI fragment. Comparison of pMSb1 and pSD17 by restriction analysis showed that both contain, within the BR gene, a 1.1-kbp insert but in opposite orientations (Fig. 1 *B* and *C*). Southern hybridization with a BR gene-specific probe to *Xma* III digests of *H. halobium* genomic and cloned DNAs (Fig. 1*B*) showed that the 1.1-kbp fragment present in digests of Pum$^+$ DNA (lanes 4 and 6) has shifted to 2.2 kbp in digests of pMSb1 (lane 3), pSD17 (lane 5), and SD17 genomic DNA (lane 7). pMSb1 and pSD17 were digested with *Bam*HI, *Pst* I, and *Kpn* I. In every case the BR gene-containing fragments obtained were of equal size (*Bam*HI, 7.5 kbp; *Pst* I, 5.2 kbp, and *Kpn* I, 2.2 kbp). Further restriction analysis showed that the inserts within the BR gene in pMSb1 and in pSD17 both contain single *Sma* I and *Hinc*II sites and produce fragments of identical size with *Tth* I (1 kbp), *Dde* I (850 bp), and *Hin*fI (430, 250, 200, and 170 bp). Finally, mapping using *Kpn* I (data not shown), which cleaves near the middle of the BR gene (1), localizes the insertions in both pMSb1 and pSD17 to the 500-bp region between the *Xma* III and *Kpn* I sites (Fig. 1*C*).

ISH*1* Is Inserted into the BR Gene in the Two Pum$^-$ Mutants at the Same Site but in Opposite Orientations. Fig. 2 shows an electron microscopic analysis of heteroduplexes formed between the 7.5-kbp *Bam*HI fragments derived from pMSb1 and pSD17. Two major classes of heteroduplexes were seen. In the first class (I), a 1.1-kilobase (kb) single-stranded substitution loop is flanked by 400 bp and 6 kbp of duplex DNA. In the second class (II), a 1.1-kbp duplex DNA is flanked by a pair of short (400 bases) and long (6 kb) single-stranded tails. The simplest interpretation of these results, consistent with the restriction mapping and hybridization data, is that a homologous DNA sequence has inserted in opposite orientations at a site(s) close to the NH_2-terminal coding region of the BR gene.

Sequence Analysis of ISH*1*. (a) *Junction sequences.* The most direct evidence that insertion of ISH*1* has occurred within the BR gene, at the same site but in opposite orientations, was provided by analysis of the junction sequences between the BR gene and the ISH*1* element by using the method of Maxam and Gilbert (14) (Fig. 3). For pMSb1, the *Sma* I site was used for determining the sequence of the left junction and the *Kpn* I site for the right junction (Fig. 1*C*). For pSD17, the *Sma* I site was used for determining the sequence of the right junction and the *Tth* I site (near the left end of ISH*1* in pSD17) for the left junction.

The terminal sequences of the inserts in the two mutants are identical but are in opposite orientations (Fig. 3 *B* and *C*). The element is inserted within the region that codes for the first few amino acids at the NH_2 terminus of BR (1).

Several features of ISH*1* and the region within the BR gene where it is inserted are noteworthy. (*i*) ISH*1* contains an interrupted 8-bp inverted repeat T-G-C-C-T- -G-T-T at its termini (Fig. 3*B*). (*ii*) Insertion of ISH*1* leads to the duplication of an 8-bp target sequence A-G-T-T-A-T-T-G. The duplicated sequence is found as a direct repeat that flanks ISH*1*. These two features, presence of terminal inverted repeats and flanking

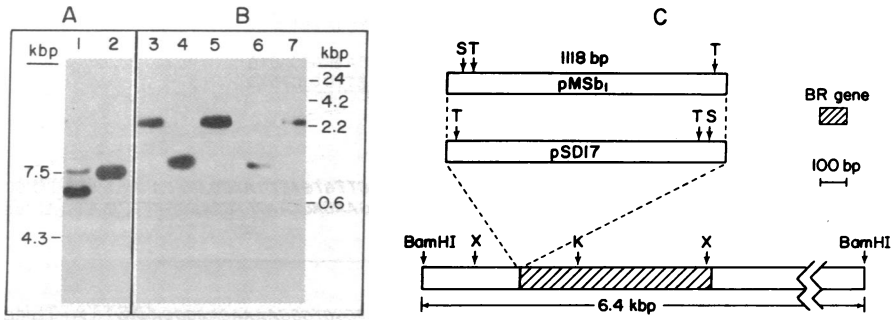

FIG. 1. (*A* and *B*) Identification of BR gene in restriction digests of genomic and cloned DNAs from Pum$^+$ and Pum$^-$ strains of *H. halobium*. (*A*) Southern hybridization of BR cDNA probe to *Bam*HI digests of genomic DNA from S9 (lane 1) and pMSb1 (lane 2). (*B*) Southern hybridization of a BR cDNA probe to *Xma* III digests of DNAs from pMSb1 (lane 3), pR1 (lane 4), pSD17 (lane 5), R1 (lane 6), and SD17 (lane 7). Numbers on the sides indicate DNA size markers. (*C*) Partial restriction map of *Bam*HI fragments containing the BR gene in different clones. The BR structural gene is shown as hatched region. The wild-type BR gene is in a 6.4-kbp *Bam*HI fragment in pR1. pMSb1 and pSD17 contain an insert at the same site within the BR gene. T, *Tth* I; S, *Sma* I; X, *Xma* III; and K, *Kpn* I.

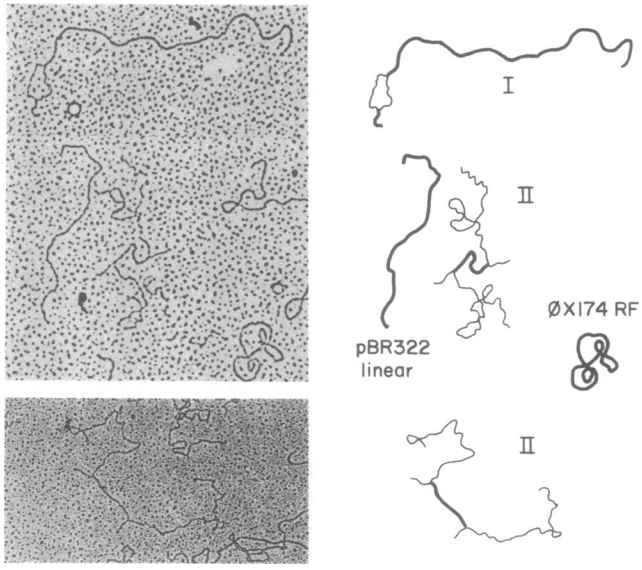

FIG. 2. Electron micrographs of heteroduplexes formed between pMSb1 and pSD17 DNAs. Plasmid DNAs were cleaved with *Bam*HI and the DNA fragments were denatured and annealed. For details of I and II see text. RF DNA from phage ϕX174 was present as a size standard.

duplication of the insertion site, are typical of transposable elements (16, 17).

Interestingly, the duplicated target sequence in the BR gene is itself flanked by an interrupted 9-bp inverted repeat (Fig. 3A), which is partially homologous to the inverted repeats of ISH*1* (indicated by dots in Fig. 3 *B* and *C*). The extent of homology between the target site and the ends of the element and the symmetrical arrangement of the target site are unique among all transposable elements whose sequences have so far been determined.

(b) *Complete sequence.* The complete sequence of ISH*1* in pMSb1 was determined (Fig. 4) by using the chain termination method (15). A 1.8-kbp *Bam*HI/*Kpn* I fragment (Fig. 1*C*) was isolated and cleaved separately with *Sau*3A, *Taq* I, and *Hpa* II. The resulting fragments were inserted into either the *Bam*HI or the *Acc* I site of M13 mp7 RF DNA (7). The sequences of both strands were determined for more than 85% of ISH*1*. Sequences in some regions where band compressions occurred were confirmed by using 10% polyacrylamide/90% formamide gels (18).

An additional interesting feature of this sequence is the presence in it of a large number of sequences with inverted repeats. Those that contain eight or more contiguous nucleotides are indicated in Fig. 4 by half arrows.

```
                                                                      60
TGCCTTGTTT TGCCACCGAT TGAGGGAAGT TTCAGACTCT CTCCCGGGAA GATTCCGTCA
                                                                     120
AGCTAACCAG GAATTGGACG CCGTCTGGCG ATATGGCATC GCTCAGACGG CTTGCTTGGA
                                                                     180
TGTGTCGAAA CCTTGCCAAA CAGCACGTTG ACGATCCGGA CGTACCCGCC GCGCCGTCCG
                                                                     240
GCGCGGGCGG GTACGCCGAG TGGGTGCAGA TCGCGTTGAT TCTGTACCGT GTCGAACTGG
                                                                     300
AAAAGAGCCT CCGTGAATCC GAGGACTACC TCAACGAGAT GCCCGGTGTT CTTGCCGTGT
                                                                     360
TTGGACTTGA CGAAGCACCA CACTACAGCT CGTTCTGCCG GTGGGAAAAC GAGTATCGAA
                                                                     420
TGCGTGAGCT CCGCCGCCTG CTCCGCGCTT CGGCGGAGCA GGCGGGCTGG AGTGGCGAAG
                                                                     480
CCGCGATTGA CGCGAGCGGC TTCCAGCGCG ATCAAACCAG CTACCACTAC CGCGACCGCG
                                                                     540
CGAATTACTC GTTCCAGTCG ATGAAGACGA CGATCTTGAT CGACGTGAAC TCGCTAGCGA
                                                                     600
TCAAGGACGT TCACTACACG ACGCAGAAGC CTGGGACGGC CACATTGGGA TGCAGGTCTT
                                                                     660
CCGCCGGAAA CGCGGAAGAC CTGCGGGTGC TGTCTGCTGA CGCGAACTAC TCGTGGAGCG
                                                                     720
ACCTCCGTGA GGAGTGTCGC TCCGAATCAA CGCGACCGTT GATCAAGCAC AGGGAGCAAA
                                                                     780
CACCGTTGCA GAAGGCTCAC CACGCCCGGA TGAACGAGGA CTACAACCAA CGCTGGATGA
                                                                     840
GTGAAACCGG CTTCTCGCAG TTGAAGGAAG ACGACGGCGA GAAGCTGCGC TCCGGAGCTG
                                                                     900
GCAGGCCAGT TCCGGAGCTG ACTCGGAAGT GCATCATCCA TAACCTGACG CAGGCGGCGA
                                                                     960
GTTAAGGGCT CGCCGCCTGC TCGCTTTCTC CGTACGTATC CGGAGAGGCA TCGCCGTCGT
                                                                    1020
CATCGGAACA ACGAAGCAAG ATACCATAGT TGTGACCCTT CAGCAACCGC CGTGAGTGAC
                                                                    1080
AGCTACTGCA TCTTCTGAGG TCAAGAACCC GTCTCTGACG CTGTGAAACT GCGAATAGTC
TTCCCTACCC CGACGCTGTC TTGTGATTCA ACGAGGCA
```

FIG. 4. Complete sequence of ISH*1* in pMSb1 DNA. The sequence shown is of the strand containing the long ORF (nucleotides 93–902). Sequences corresponding to the initiator and terminator codons are indicated by * and **, respectively. ATG and GTG sequences that are in phase with the long ORF are underlined. Inverted repeat sequences eight nucleotides or longer are indicated by half arrows.

ISH*1* Has Long ORFs. The sequence shown in Fig. 4 contains ORFs of various lengths in both strands, most of which are short except for one which is very long (810 nucleotides; 93–902 of Fig. 4), representing 72% of the total ISH*1* sequence. This ORF can code for a 30-kilodalton protein that is slightly less basic than the putative proteins encoded by *Escherichia coli* IS2 and IS5 (19–21).

If GUG is used as an initiator codon in *H. halobium* (22), the sequence complementary to that shown in Fig. 4 also has a fairly long ORF. This ORF (nucleotides 467–66) is contained almost totally within the longer ORF in the other strand and has the

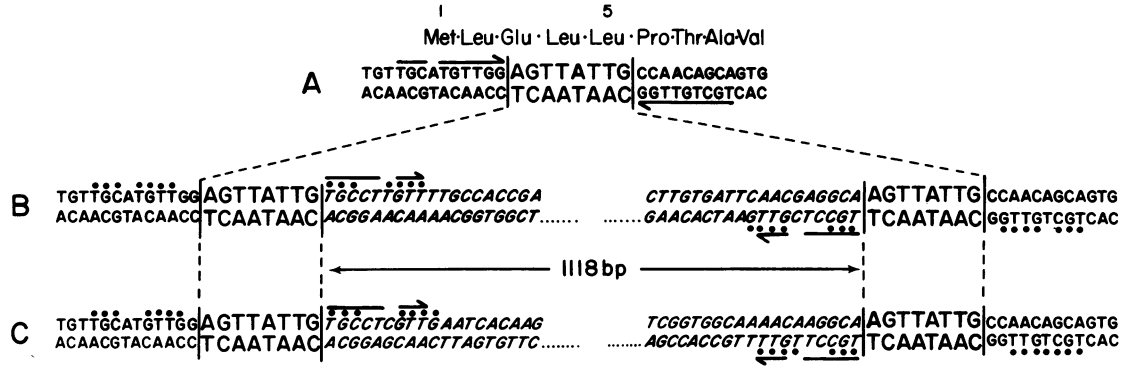

FIG. 3. Junction sequences at the ISH*1* insertion site within the BR gene. (*A*) DNA sequence of the NH_2 terminus of the BR gene. (*B* and *C*) BR gene–ISH*1* junction sequences in pMSb1 and in pSD17 DNAs, respectively. Shown in larger letters is the 8-bp sequence at the insertion site (*A*) that flanks insertions of ISH*1* as direct repeats (*B* and *C*). Inverted repeat sequences in the BR gene and at the ends of ISH*1* are indicated by half arrows. Dots indicate homology in DNA sequence between the BR gene and ends of ISH*1*.

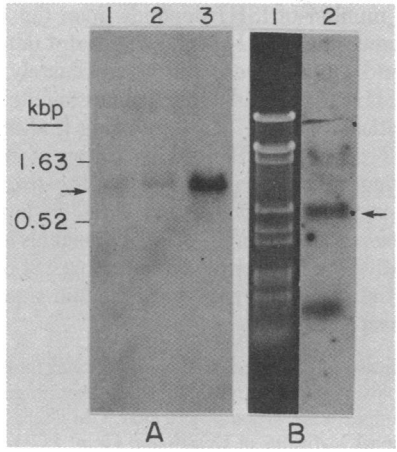

FIG. 5. Detection of transcripts corresponding to ISH*1* in *H. halobium*. (*A*) RNA gel transfer hybridization of RNA isolated from *H. halobium* strains S9 (lane 1), R1 (lane 2), and R1mR (lane 3) with a single-stranded ISH*1*-specific DNA as probe. Numbers on the left specify DNA size markers. Arrow on the left indicates the location of BR mRNA on the same gel (as detected by using BR cDNA as probe). (*B*) Southern blot hybridization of ^{32}P-labeled *H. halobium* RNA to a combined *Tth* I, *Kpn* I, and *Bam*HI digest of pMSb1 DNA. Lane 1, ethidium bromide staining pattern; lane 2, hybridization pattern. Arrow on the right indicates hybridization of RNA to a 1-kb *Tth* I fragment derived from ISH*1* (Fig. 1*C*).

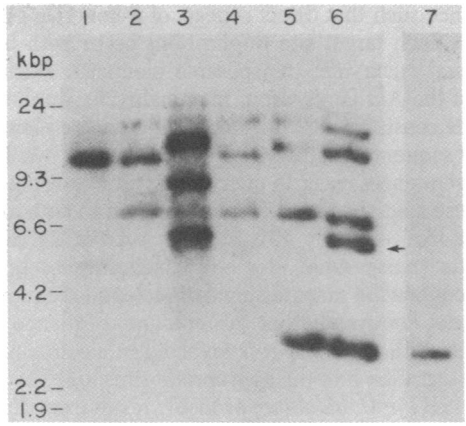

FIG. 6. Southern blot hybridization of ISH*1* specific probe to *Pst* I digests of DNA from different strains of *H. halobium*. Total DNA from NRC-1, R1, R1mR, L33, S9, and SD17 (lanes 1–6, respectively) and plasmid DNA from S9 (lane 7) were used for digestion. Arrow indicates the position of the new copy of ISH*1* formed by duplicative transposition of ISH*1* within the BR gene.

same reading frame as the longer ORF. This situation is analogous to that observed in several *E. coli* insertion elements in which both strands of an element have ORFs in which codons are read in the same frame. In the case of *E. coli* IS*5* it has been demonstrated recently (21) that the ORFs in both strands are expressed *in vivo* to produce 12.3- and 37-kilodalton proteins.

A Stable Transcript Homologous to ISH*1* Is Present in *H. halobium*. Fig. 5*A* shows the results of RNA gel transfer hybridization using a single-stranded DNA complementary to the long ORF in ISH*1* probe against RNA from three *H. halobium* strains. Each strain contains one RNA band that hybridizes to the probe. The intensity of hybridization for strain R1mR (lane 3) is, however, significantly greater than that of the other two strains tested. This finding is consistent with the suggestion (see below) that R1mR has the largest copy number of ISH*1* sequences. The size of the ISH*1* specific transcript is slightly larger than that of the mRNA for BR (arrow indicates the position of BR mRNA on the same gel) and is estimated to be around 900 nucleotides. This finding is consistent with the possibility that the ISH*1* specific RNA may code for a protein roughly 270 amino acids long.

Fig. 5*B* shows the results of Southern hybridization using uniformly ^{32}P-labeled RNA from *H. halobium* strain S9 as probe against a combined *Tth* I, *Kpn* I, *Bam*HI digest of pMSb1 (the ^{32}P-labeled RNA used had been depleted of BR mRNA by passage through a BR cDNA column). Among the many *H. halobium* DNA fragments present in the digest, only two hybridized to the RNA, and one of these (indicated by arrow in Fig. 5*B*) corresponds to the 1-kbp *Tth* I fragment that covers almost all of ISH*1* sequence (Fig. 1*C*).

ISH*1* Is Present in Multiple Copies in the *H. halobium* DNA. To examine the copy number and mode of transposition of ISH*1*, *Pst* I digests of total DNA from six *H. halobium* strains (NRC-1, R1, R1mR, L33, S9, and SD17) were analyzed by Southern hybridization using ISH*1* specific DNA (1-kbp *Tth* I fragment; Fig. 1*C*) as probe. Fig. 6 shows that ISH*1* is present from one to five or more copies per genome, depending upon the strain. Wild-type strain NRC-1 contains one strong band, whereas strains R1 and L33 contain two additional bands. R1mR, S9, and SD17 contain four to six bands that hybridize to the probe. At least some of these bands are derived from plasmid DNAs of *H. halobium*. For example, the ≈3.2-kbp fragment in digests of S9 and SD17 DNA is also present in *Pst* I digests of purified plasmid DNA from strain S9 (lanes 5–7). In addition, the relatively strong intensity of three of the hybridization bands in digests of R1mR suggests that these may also be derived from plasmid DNA(s). Because plasmids in halobacteria are present in about four copies (23), this would mean that strain R1mR has the largest copy number of ISH*1* sequences in its DNA.

Comparison of hybridization patterns for S9 (Pum$^+$) and SD17 (Pum$^-$) indicates that SD17 has an additional copy of ISH*1* compared to its parent S9 (Fig. 6, lanes 5 and 6). This band, indicated by an arrow, corresponds to the ISH*1* that has inserted within the BR gene. Thus, the mutation leading to Pum$^-$ phenotype of SD17 is due to a duplicative transposition of ISH*1* into the BR gene.

DISCUSSION

Pfeifer *et al.* (3) have reported on the high frequency (10^{-4} to 10^{-2}) of spontaneous Pum$^-$, gas vacuole, and bacterioruberin mutations in halobacteria and the association of these mutations with changes in restriction patterns of plasmid DNA. This genetic instability has been ascribed to the presence of a large number of mobile elements (2, 3). However, a correlation between a single phenotype and a specific change at the DNA level has not been demonstrated. We have now characterized a transposable element from *H. halobium* and the inactivation of the BR gene by its insertion at a specific site. It is not known whether this mechanism accounts for the formation of a majority of spontaneous Pum$^-$ mutants; analysis of several Pum$^-$ mutants should provide this knowledge. Such a study should also provide some information on the generality of insertional inactivation of genes in halobacteria.

ISH*1* has many features similar to those found in prokaryotic and eukaryotic transposable elements: (*i*) It contains interrupted 8-bp inverted repeats at its ends. These are the shortest known for prokaryotic transposable elements except for IS*50*-R and the bacteriophage Mu (16). In eukaryotic transposable elements, the size of the inverted repeats again varies and may be as short as 2 bp (24). (*ii*) Insertion of ISH*1* leads to duplication of an 8-

bp sequence such that direct repeats of it now flank the ISH*1* sequence. Such target site duplications occur with both prokaryotic and eukaryotic transposable elements. With the exception of the *Alu* family from mammals (25), duplications in most cases consist of 5–9 bp, depending on the element. (*iii*) The ISH*1* sequence begins with T-G and ends in C-A. The same terminal sequences occur in every eukaryotic transposable element whose sequence has been determined to date except for the "copia-like" element, 297, from *Drosophila* (24, 26). Many prokaryotic transposable elements, including bacteriophage Mu, also contain the same terminal dinucleotide sequences (16, 17). It is not known whether this common sequence has evolutionary significance or if it is involved in a common mechanism of integration into the appropriate sites. (*iv*) The multiple copies of ISH*1* in *H. halobium* probably result from duplicative transposition of an existing element as has been shown for *E. coli* elements (27). The copy number of ISH*1* depends upon the *H. halobium* strain examined. NRC-1, the wild-type strain, contains a single copy, whereas the mutant strains contain more (Fig. 6). (*v*) Some transposable elements in *E. coli* occur both on the chromosome and on the F plasmid (28). The present work shows that, in strain S9 and, possibly, R1mR, ISH*1* occurs on both the chromosomal and the plasmid DNA. It has been suggested that integration of *H. halobium* plasmids into the chromosome occurs fairly frequently (3). If so, it is likely that this process is facilitated by the presence of ISH*1* in the plasmid and the chromosome. This is analogous to the proposed mechanism for integration into and excision of F factor from the *E. coli* chromosome (28).

ISH*1* inserts into the BR gene at the same site but in opposite orientations in the two Pum^- mutants studied. The possibility must therefore be considered that ISH*1* insertions into the BR gene are site specific. In the transposable elements studied there is a wide range of specificity in sites of insertion (16). Thus, bacteriophage Mu inserts essentially randomly whereas elements such as IS*4*, Tn*554*, Tn*7*, Tn*10*, and the *Drosophila* copia-like element 297 (26) are quite sequence specific. The factors that determine site specificity of insertion are not known. In a few cases—e.g., Tn*10* and copia-like *Drosophila* elements—insertions occur within a symmetrical sequence (26, 29). In the case of ISH*1*, the duplicated target sequence, A-G-T-T-A-T-T-G, is not symmetrical. However, it is interesting that the target sequence in the BR gene is flanked by an almost perfect 9-bp inverted repeat (Fig. 3*A*). Furthermore, nucleotides within this inverted repeat show significant homology to the ends of ISH*1* (Fig. 3 *B* and *C*). A similar situation exists also in the case of Tn*7*, which inserts at a unique site on the *E. coli* chromosome and on ColE1 DNA (30). Whether ISH*1* inserts into *H. halobium* DNA in a site-specific manner and whether these features in the target DNA contribute to the specificity of insertion must await cloning and junction sequence analysis of other DNA fragments with ISH*1* inserts.

ISH*1* has, similar to *E. coli* insertion elements, a long ORF with the potentiality of coding for a 30-kilodalton protein. The finding of an ISH*1* specific transcript indicates that the long ORF in ISH*1* is also expressed *in vivo*. The size of the transcript corresponding to ISH*1* (≈900 nucleotides) suggests that it could code for a 30-kilodalton protein. Another interesting feature of the ISH*1* sequence is the presence within it of a large number of inverted repeats (Fig. 4). In particular, the region around the start codon (nucleotides 93–95) of the ORF can form a stable stem–loop (13-bp stem, 5-base loop) structure. A similar stem–loop structure can occur in BR mRNA (1). However, the AUG codon of ISH*1* RNA would be in the loop rather than in the stem as for the BR mRNA (1).

The copy number of ISH*1* depends upon the *H. halobium* strain. The amount of ISH*1* specific transcript detected on gel transfer hybridization corresponds approximately to the copy number of ISH*1*. This would imply that most of the ISH*1* copies are transcriptionally active. An interesting finding is that, although ISH*1* copies are flanked by different sequences in strains carrying multiple copies of ISH*1*, the size of the transcript is the same in all strains. This would suggest that ISH*1* contains its own promoter and terminator signals for transcription. Alternatively, if ISH*1* RNA is made as part of read-through transcripts, the ISH*1* sequence must contain signals for RNA processing near its termini.

We thank Christopher Bazinet and Erika Hartwieg for their help with the heteroduplexing and electron microscopy experiment. This investigation was supported by Grants GM 17151, GM 28289, and AI 11479 from the National Institutes of Health and Grant PCM-8110992 from the National Science Foundation.

1. Dunn, R., McCoy, J., Simsek, M., Majumdar, A., Chang, S. H., RajBhandary, U. L. & Khorana, H. G. (1981) *Proc. Natl. Acad. Sci. USA* **78**, 6744–6748.
2. Sapienza, C. & Doolittle, W. F. (1982) *Nature (London)* **295**, 384–389.
3. Pfeifer, F., Weidinger, G. & Goebel, W. (1981) *J. Bacteriol.* **145**, 375–381.
4. Panet, A., van de Sande, J. H., Loewen, P. C., Khorana, H. G., Lillehaug, J. R., Maniatis, T., Fritsch, E. F. & Kleppe, K. (1973) *Biochemistry* **12**, 5045–5050.
5. Chang, S., Majumdar, A., Dunn, R., Makabe, O., RajBhandary, U. L., Khorana, H. G., Ohtsuka, E., Tanaka, T., Taniyama, Y. O. & Ikehara, M. (1981) *Proc. Natl. Acad. Sci. USA* **78**, 3398–3402.
6. Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) *Molecular Cloning* (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY), pp. 117–121.
7. Messing, J., Crea, R. & Seeburg, P. H. (1981) *Nucleic Acids Res.* **9**, 309–323.
8. Hanahan, D. & Meselson, M. (1980) *Gene* **10**, 63–67.
9. Southern, E. M. (1975) *J. Mol. Biol.* **98**, 503–517.
10. Jeffreys, A. J. & Flavel, R. A. (1977) *Cell* **12**, 429–439.
11. Thomas, P. S. (1980) *Proc. Natl. Acad. Sci. USA* **77**, 5201–5205.
12. Alwine, J. C., Kemp, D. J., Parker, B. A., Reiser, J., Renart, J., Stark, G. R. & Wahl, G. M. (1979) *Methods Enzymol.* **68**, 220–242.
13. Davis, R. W., Simon, M. & Davidson, N. (1971) *Methods Enzymol.* **21**, 413–428.
14. Maxam, A. M. & Gilbert, W. (1980) *Methods Enzymol.* **65**, 499–560.
15. Sanger, F., Nicklen, S. & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 5463–5467.
16. Kleckner, N. (1981) *Annu. Rev. Genet.* **15**, 341–404.
17. Calos, M. P. & Miller, J. H. (1980) *Cell* **20**, 579–595.
18. Lockard, R. E. & Kumar, A. (1981) *Nucleic Acids Res.* **9**, 5125–5140.
19. Ghosal, D., Sommer, H. & Saedler, H. (1979) *Nucleic Acids Res.* **6**, 1111–1121.
20. Kroger, M. & Hobom, G. (1982) *Nature (London)* **297**, 159–162.
21. Rak, B., Lusky, M. & Hable, M. (1982) *Nature (London)* **297**, 124–128.
22. Bayley, S. T. (1976) in *Extreme Environments: Mechanisms of Microbial Adaptation*, ed. Heinrich, M. R. (Academic, New York), p. 119.
23. Weidinger, G., Klotz, G. & Goebel, W. (1977) *Plasmid* **2**, 377–386.
24. Temin, H. M. (1980) *Cell* **21**, 599–600.
25. Schmid, C. W. & Jelinek, W. R. (1982) *Science* **216**, 1065–1070.
26. Ikenaga, H. & Saigo, K. (1982) *Proc. Natl. Acad. Sci. USA* **79**, 4143–4147.
27. Read, H. A., DasSarma, S. & Jaskunas, S. R. (1980) *Proc. Natl. Acad. Sci. USA* **77**, 2514–2518.
28. Deonjer, R. C. & Hadley, R. C. (1980) *Plasmid* **3**, 48–64.
29. Halling, S. M. & Kleckner, N. (1982) *Cell* **28**, 155–163.
30. Lichtenstein, C. & Brenner, S. (1982) *Nature (London)* **297**, 601–603.