Characterization of a Second Gene Involved in Bacterio-Opsin Gene Expression in a Halophilic Archaeobacterium

DIANE LEONG,^{1†} FELICITAS PFEIFER,² HERBERT BOYER,¹ AND MARY BETLACH^{1*}

Department of Biochemistry and Biophysics, University of California, San Francisco, San Francisco, California 94143,¹ and Max-Planck-Institut für Biochemie, Munich, Federal Republic of Germany²

Received 29 April 1988/Accepted 22 July 1988

Southern blot analysis and nucleotide sequencing of DNA from three bacterio-opsin-deficient mutants of the archaeobacterium *Halobacterium halobium* (M86, W105, and W109) revealed that they each contain an alteration in a region 2,000 to 3,800 base pairs (bp) upstream of the bacterio-opsin gene (*bop*). Nucleotide sequence analysis of this region, which is also located downstream of the previously characterized *brp* gene, revealed that it contains an open reading frame (ORF) of 2,022 bp. This 2,022-bp ORF has a start codon which overlaps the stop codon of the *brp* gene and is read in the same direction. The ORF could encode an acidic protein of 73,334 daltons (674 amino acids) with a predicted secondary structure typical of a soluble protein. Bop mutant M86 contains a 1,883-bp deletion extending from bp 351 of the ORF, to 197 bp beyond the stop codon. Mutant W105 has an ISH2 element integrated at bp 1239 of the ORF, and mutant W109 has an ISH26 element integrated at bp 1889. Our results suggest that the ORF is a gene (designated *bat* for bacterio-opsin activator gene) involved in *bop* gene expression.

The purple membrane of the archaeobacterium *Halobacterium halobium* functions as a light-driven proton pump and is composed of lipids and bacteriorhodopsin (32). Bacteriorhodopsin consists of the bacterio-opsin protein complexed stoichiometrically with the retinal chromophore. Physiological conditions of low oxygen tension in the presence of light induce the formation of purple membrane, a process in which the syntheses of bacterio-opsin and retinal appear to be coordinated (22, 33, 34).

Bacterio-opsin is encoded by a 786-base-pair (bp) gene (*bop*) located in the more guanine-plus-cytosine-rich (G + C-rich) fraction of the *H. halobium* genome (4, 10). A focus of our investigations has been the regulation of this gene. Until a DNA transformation system has been developed for *H. halobium*, our approach to analyzing *bop* gene regulation has been to characterize Bop mutants. Spontaneous Bop mutants occur at a high frequency (10^{-4} [29]), and the most common mutational event resulting in a Bop⁻ phenotype is the integration of insertion elements in or near the *bop* gene (3, 25, 27). ISH2 (520 bp) and ISH1 (1,118 bp) are the most frequently found insertion elements (23) in the Bop mutants analyzed thus far. Mapping of the integration sites of various ISH elements in Bop mutants led to the discovery of another gene, designated *brp*, which is located 526 bp upstream of the *bop* gene and is transcribed in the opposite direction (3). It is clear that *brp* gene expression affects *bop* gene expression or purple membrane formation or both, but the mechanism remains unclear.

The only Bop revertant described thus far has been isolated from Bop mutant IV-4, which contains an ISH24 insertion element of 3 kbp near the 3' terminus of the *brp* gene (1, 27). In revertant reIV-41, a 588-bp segment of DNA ("ISH25") is found adjacent to ISH24, presumably due to a homologous recombination event between the ISH24 copy within the *brp* gene and an ISH24 copy located in plasmid pHH1 DNA (27). Spontaneous Bop mutants derived from

the revertant arise at the same frequency as do those derived from the wild type. A series of eight such mutants of the revertant with insertions in the *bop* gene, in the *brp* gene or between these two genes has been described (26).

In an effort to identify other sites required for *bop* gene expression or purple membrane assembly, we have characterized three Bop mutants derived either from the wild type or from the Bop revertant. These three mutants contain alterations in a region of DNA from 2,000 to 3,800 bp upstream of the *bop* gene and immediately downstream of the *brp* gene. Our data imply that this region is also involved in *bop* gene expression. The properties of a putative gene of 2,022 bp within which the alterations lie are described.

MATERIALS AND METHODS

Materials. [γ -³²P]ATP (>7,000 Ci/mM), [α -³²P]CTP (>400 Ci/mM) and α -³²P-labeled deoxynucleotide triphosphates (400 Ci/mM) were obtained from Amersham Corp., Arlington Heights, Ill. Deoxynucleotide triphosphates were obtained from P-L Biochemicals, Inc., Milwaukee, Wis. SP6 RNA polymerase and the in vitro transcription kit were from Promega Biotec, Madison, Wis. The Sequenase DNA sequencing kit was obtained from United States Biochemical Corp., Cleveland, Ohio. Restriction endonucleases and all other DNA modifying enzymes were obtained from New England BioLabs, Beverly, Mass.

Halobacterial strains and mutants. Bop mutants, such as M86, W105, and W109, were isolated as follows. Purple membrane-deficient (Pum⁻) mutants were isolated on solid media as white colonies in a background of purple colonies. In principle, the Pum⁻ phenotype can arise from mutations in either retinal synthesis (a Ret⁻ phenotype) or in *bop* gene expression or purple membrane assembly (a Bop⁻ phenotype). These two phenotypes were distinguished by growing the mutants in the presence of exogenous retinal. Pum⁻ mutants determined to be Ret⁺ were assumed to be Bop⁻. Bop mutants were derivatives of either strain II-7, which is a bacterioruberin-deficient derivative of wild-type *H. halobium* NRC817, or strain R1, which is a vacuole-deficient derivative of wild-type *H. halobium* NRC34020. The Bop

* Corresponding author.

† Present address: Advanced Genetic Sciences, Oakland, CA 94608.

TABLE 1. Mutants used in this study

Strain	Phenotype	Description	Insertion or deletion (location in or near <i>bop</i> gene)	Reference
II-7	Bop ⁺ Ret ⁺	Parent of IV-4	None	26, 28
IV-3	Bop ⁻ Ret ⁺	Bop mutant of II-7	ISH2 (3' end of <i>brp</i> gene)	3
IV-4	Bop ⁻ Ret ⁺	Bop mutant of II-7	ISH24 (3' end of <i>brp</i> gene)	26
reIV-41	Bop ⁺ Ret ⁺	Revertant of IV-4	ISH24, "ISH25" (3' end of <i>brp</i> gene)	26
M86 ^a	Bop ⁻ Ret ⁺	Bop mutant of R1	Deletion (upstream of <i>bop</i> gene)	This study
M133	Bop ⁻ Ret ⁺	Bop mutant of R1	Unidentified insertion (in <i>Pst</i> I 1.1-kb fragment)	This study
M140	Bop ⁻ Ret ⁺	Bop mutant of R1	ISH26 (in <i>bop</i> gene)	11
W105 ^b	Bop ⁻ Ret ⁺	Bop mutant of reIV-41	ISH2 (in <i>Pst</i> I 1.1-kbp fragment)	This study
W109	Bop ⁻ Ret ⁺	Bop mutant of reIV-41	ISH26 (in <i>Pst</i> I 1.1-kbp fragment)	This study
JW-5 ^c	Bop ⁺ Ret ⁻	Ret mutant	None	35
S9Flx3R	Bop ⁺ Ret ⁻	Ret mutant	None	31
M144	Bop ⁺ Ret ⁻	Ret mutant	None	This study

^a All M mutants were the gift of D. Oesterhelt and were isolated as described previously (21).

^b All W mutants also retain insertions of reIV-41.

^c JW-5 was isolated as a spontaneous vacuole-deficient derivative of mutant W5002-1.

and Ret mutants which were used in this study are listed and described in Table 1. Halobacterial growth conditions were as described previously (25). Bop⁺ natural isolate YC81819-9 and halobacterial species GN101, *Halobacterium salinarum* K, and *Halobacterium cutirubrum* have been described previously (24, 25).

DNA isolation and cloning. *Escherichia coli* plasmid DNA was isolated by the alkaline lysis method (5), whereas total halobacterial DNA was isolated as described previously (4). Derivatives of *E. coli* plasmids pKG2 and pKGS containing various fragments of halobacterial DNA were isolated from genomic libraries constructed as described by Kuhn et al. (14) and Pfeifer et al. (25). Transformants containing the restriction fragments desired were identified by colony hybridization with the appropriate nick-translated DNA fragment.

Southern hybridization. Southern blots of total halobacterial DNA were prepared as described previously (25). Nick-translated probes used for mapping the mutations in M86, W105, and W109 and for analysis of the open reading frame (ORF) region in various Pum⁺ halobacterial strains were made and hybridized to the blots under the conditions used by Pfeifer et al. (25). The ORF region RNA probes were made from DNA templates which were derivatives of plasmids pSP64 or pSP65 containing the restriction fragment indicated. The probes were synthesized in vitro from the SP6 promoter by using SP6 RNA polymerase and hybridized to Southern blots under conditions recommended by the manufacturer (Promega Biotec) (20).

DNA sequencing. DNA sequencing of the ORF and flanking regions in the wild type and M86 was done by the method of Maxam and Gilbert as modified by Maniatis (18, 19) or by the Sequenase method with conditions recommended by the manufacturer (United States Biochemical Corp). The sources of restriction fragments used for the former method were pSP64 or pSP65 containing the 955-bp *Pst*I-*Bam*HI fragment, pEMBL8 containing the 1,134-bp *Pst*I fragment, pKGS containing the approximately 550-bp *Pst*I-*Bgl*II fragment, and pEMBL9 containing the 2,160-bp *Bam*HI fragment. The templates used in the Sequenase reactions were single-stranded M13 derivatives containing the 1,134-bp *Pst*I fragment or the approximately 550-bp *Pst*I-*Bgl*II fragment. The Sanger sequencing method as modified by Chen and Seeburg (7) was used for the mapping of the insertions in Bop mutants W105 and W109. The templates were double-stranded DNA from pKGS or pKG2 containing the *Pst*I fragments from W105 and W109 corresponding to the 1,134-bp *Pst*I fragment from the wild type.

DNA sequence analysis. Sequence analysis was performed by using computer programs designed by H. Martinez (University of California, San Francisco) as well as the DFASTN program (17) and a protein secondary structure prediction program developed by J. Finer-Moore and R. Stroud (University of California, San Francisco) (12).

RESULTS

Restriction mapping of mutations. The Bop mutants M86, W105, and W109 were analyzed in this study. In order to characterize the mutations in these mutants, the *bop* gene region of each mutant was mapped by Southern blot analysis. None of the three mutants, M86, W105, and W109, showed any detectable alterations in the *bop* gene, the *brp* gene or the 526-bp area of DNA between the two genes. Mutant M86 was derived from strain R1, a Bop⁺ gas vacuole-deficient derivative of the wild type. Genomic Southern blot hybridizations with 40 kilobases (kb) of DNA containing the *bop* gene and flanking regions suggested that the only alteration present in M86 was a deletion located upstream of the *bop* gene. Restriction mapping and Southern hybridizations of the region from approximately 1,500 to 3,800 bp upstream of the *bop* gene supported this interpretation.

Mutants W105 and W109 are members of a recently isolated series of 11 white spontaneous Bop mutants (W101 to W111) derived from a Bop revertant, reIV-41 (unpublished data). Genomic Southern hybridizations with a 1.1-kbp *Pst*I fragment located downstream of the *brp* gene (see Fig. 1) revealed that mutants W105 and W109 contained insertions of approximately 500 and 1,500 bp, respectively, within this region. These insertions were present in addition to the two normally found in the *brp* gene of the parental revertant, which are ISH24 and "ISH25".

The *Pst*I fragments of W105 and W109 which corresponded to the 1.1-kbp *Pst*I fragment of the wild type were cloned in the positive selection vector pKGS (14). In order to identify the insertions, a series of insertion sequence-specific DNA probes was made from previously characterized Bop mutants and used to probe the cloned *Pst*I fragments of the mutants. A 270-bp *Mbo*II fragment derived from the ISH2 copy in the *brp* gene of mutant IV-3 (3) hybridized with the insertion in the W105 *Pst*I fragment. A 2,029-bp *Ava*I fragment containing the ISH26 copy in the *bop* gene of mutant M140 (11) hybridized with the insertion in the W109 *Pst*I fragment. These results indicated that the insertions in the region downstream of the *brp* gene in W105 and W109

FIG. 1. Restriction map of wild-type DNA illustrating the locations of alterations in mutants M86, W105, and W109 and the nucleotide-sequencing strategy for mutant and wild-type DNA. The coding regions of the *brp* gene and part of the *bop* gene are indicated by bars, with wavy arrows below the bars denoting the 5' termini of the *brp* and *bop* mRNAs. The locations of insertions in mutants W105 and W109 are marked with vertical lines, whereas the deletion in M86 is marked by a bracket above the map. Restriction sites are indicated as follows: □, *Bam*HI; ●, *Pst*I; △, *Nru*I; ▽, *Mlu*I; ○, *Ava*I; ■, *Bgl*II. Not all sites are shown. ORFs downstream of the *brp* gene are denoted by straight arrows directly below the map. The 3' terminus of the 381-bp ORF is dotted, indicating that the start of this ORF has not yet been determined. The nucleotide-sequencing strategy used for wild-type DNA is summarized by the cluster of arrows beneath the map, which indicate the length and strand sequenced. Most of the sequence was determined by using the Maxam-Gilbert method, with the exception of the sequence downstream from the most distal *Pst*I site, which was determined by the Sequenase method (United States Biochemical Corp.). The sequencing strategy used for M86 DNA is illustrated above the wild-type map. The restriction fragments used for Maxam and Gilbert method sequencing which span the ends of the deletion in M86 are listed at the left, while the arrows indicate the extent and direction of sequence obtained. The wild-type nucleotide sequence is shown in Fig. 2.

are the halobacterial insertion elements ISH2 and ISH26, respectively. We also corroborated the type of halobacterial insertion element present in W105 and W109 by DNA sequence analysis of the ends of the insertions as described below and by comparison with the previously determined nucleotide sequence of ISH2 and ISH26 (9, 11).

DNA sequencing of wild-type and mutant DNA. The locations of the alterations in the Bop mutants described above raised the possibility that the region downstream of the *brp* gene was somehow involved in *bop* gene regulation. To characterize this region, DNA sequence analysis was performed in the wild type and in mutants M86, W105, and W109. A 2,276-bp region from the wild type was sequenced (see Fig. 1 and 2). Analysis of the sequence revealed that the longest ORF is 2,022 bp and reads in the same direction as the *brp* gene (Fig. 1). The ATG start codon of the 2,022-bp ORF (designated ORF 2022) overlaps the TGA stop codon of the *brp* gene (ATGA). Northern (RNA) blot analysis with single-stranded RNA probes (described in the accompanying paper [16]) revealed a transcript of an appropriate size and orientation corresponding to the 2,022-bp ORF. Other ORFs in the same direction are no longer than 351 bp.

In the opposite direction, the longest ORF is 1,059 bp long and overlaps the 2,022-bp ORF (Fig. 1 and 2). A shorter ORF of 381 bp for which the start codon has not yet been determined was observed at the end of the sequenced region distal to the *bop* and *brp* genes and the 2,022-bp ORF (Fig. 1). Preliminary analysis indicated that only the latter of these two ORFs appears to be transcribed (data not shown).

Parts of the region containing ORF 2022 in mutants M86, W105, and W109 were sequenced in order to further characterize the defects in these mutants (Fig. 1). Restriction mapping of the regions flanking the deletion in M86 indicated that a 254-bp *Hinf*I-*Mlu*I fragment and a 503-bp *Mlu*I-*Bam*HI fragment span the sites where the deletion occurred. Sequencing of these fragments by the method of Maxam and Gilbert (19) in the directions indicated on the map (Fig. 1) allowed us to determine the nucleotide sequence at the

deletion boundaries. Comparison with the wild-type sequence indicated that the deletion in M86 encompasses a total of 1,883 bp starting at bp 351 of ORF 2022 and ending 197 bp beyond the stop codon of ORF 2022 (Fig. 2). In addition, there is an 8-bp direct repeat of wild-type sequences at the boundaries of the deletion (Fig. 2); in M86, a single copy of the repeated sequence is present. No inverted repeats at the ends of the deleted region were observed.

In order to determine the exact location of the insertions in the region downstream of the *brp* gene in mutants W105 and W109, oligonucleotides specific to one end of ISH2 and ISH26, respectively, were used as primers in a Sanger DNA sequencing reaction. The *Pst*I fragments of W105 and W109 cloned in pKGS were used as templates (7). This DNA sequence analysis revealed that the ISH2 insertion element in mutant W105 occurs at bp 1239 of ORF 2022 and the ISH26 element in mutant W109 occurs at bp 1889 (Fig. 2). In the portions of ORF 2022 sequenced in M86, W105, and W109, there were no other nucleotide changes observed relative to the wild-type sequence. The above analysis established that all of the alterations in the three Bop mutants disrupt ORF 2022.

Characterization of the 2,022-bp ORF. The regions flanking ORF 2022 were examined for potential signals for transcription and translation. A consensus promoter sequence for halobacterial structural genes has yet to be defined due to the paucity of gene sequences for comparison and the lack of an assay for promoter function. However, comparison of the upstream region of ORF 2022 (Fig. 2) to that upstream of the five halobacterial structural genes sequenced so far (3, 6, 8, 10, 15) revealed that the region upstream of ORF 2022 has homology only to *bop* and *brp* upstream sequences. This homology is considerably less than that observed between the upstream regions of the *bop* and *brp* genes (data not shown). The significance of the homology of sequences upstream of ORF 2022 to those upstream of the *bop* and *brp* genes remains uncertain until a transcription initiation site is determined for ORF 2022. No homology was observed to the

FIG. 2. Nucleotide sequence upstream of the *bop* gene in *H. halobium* wild-type NRC817 and translation of the 2,022-bp ORF. The figure shows 2,798 bp of sequence. Numbering starts with the first base of the start codon of the ORF. The 357 bp of sequence upstream of ORF 2022 (which constitutes *brp* gene sequences) is numbered negatively backward from the ORF 2022 start codon at position 1. The sequence of ORF 2022 is of the same strand as the *brp* gene, and the *brp* gene stop codon which overlaps the start codon of the ORF is underlined. The ORF terminates at nucleotide 2022, corresponding to a possible 674 amino acid residues. Sequence up to the *Mlu*I site at position 2276 (shown underlined) was determined on both strands. Sequence beyond the *Mlu*I site was determined for one strand only. The 8-bp direct repeats at the boundaries of the deletion in M86 are boxed. The locations of the ISH2 insertion in mutant W105 and the ISH26 insertion in W109 are indicated with vertical arrows. A 14-bp imperfect inverted repeat at the 3' terminus of the ORF coding sequence which constitutes a possible transcription termination signal is indicated by horizontal arrows beneath the sequence. The start codon of the untranscribed 1,059-bp ORF which is on the opposite strand as ORF 2022 is indicated with an arrow (position 1977), and the stop codon is boxed (position 921).

	<i>bop</i>	<i>brp</i>	<i>hop</i>	ORF
UUU phe	0	1	2	0
UUC	13	18	8	20
UUA leu	2	0	0	1
UUG	6	6	4	12
CUU	2	4	2	2
CUC leu	12	17	12	15
CUA	2	0	1	0
CUG	15	19	19	19
AUU	0	1	0	1
AUC ile	15	12	14	27
AUA	0	0	1	0
AUG met	10	9	10	8
GUU	3	2	0	2
GUC val	8	26	9	33
GUA	3	0	1	1
GUG	9	16	24	31

	<i>bop</i>	<i>brp</i>	<i>hop</i>	ORF
UCU	0	0	0	0
UCC ser	3	2	4	12
UCA	0	1	2	1
UCG	5	12	10	13
CCU	0	0	0	0
CCC pro	3	5	3	5
CCA	3	3	0	1
CCG	6	8	4	18
ACU	0	2	0	2
ACC thr	7	4	8	30
ACA	2	0	0	3
ACG	10	8	10	24
GCU	1	1	0	3
GCC ala	8	27	7	29
GCA	5	5	0	3
GCG	16	24	32	39

	<i>bop</i>	<i>brp</i>	<i>hop</i>	ORF
UAU tyr	3	1	2	1
UAC	8	12	5	14
UAA OC	0	0	0	0
UAG AM	0	0	0	0
CAU	0	2	0	2
CAC his	0	3	1	11
CAA gln	0	0	1	2
CAG	4	8	3	27
AAU asn	0	1	0	0
AAC	3	2	4	11
AAA lys	3	1	0	1
AAG	4	1	1	6
GAU asp	2	0	2	11
GAC	8	15	7	64
GAA glu	4	0	0	9
GAG	7	5	6	49

	<i>bop</i>	<i>brp</i>	<i>hop</i>	ORF
UGU cys	0	0	0	4
UGC	0	0	2	2
UGA OP	1	1	1	1
UGG trp	8	9	10	4
CGU	3	0	0	2
CGC arg	3	6	4	23
CGA	0	2	1	2
CGG	1	9	5	18
AGU ser	1	0	5	1
AGC	5	6	3	10
AGA arg	0	0	1	0
AGG	0	1	0	0
GGU	4	7	1	0
GGC gly	10	22	7	27
GGA	4	3	4	3
GGG	8	10	12	15

FIG. 3. Codon usage of *bop*, *brp*, and *hop* genes and ORF 2022 (3, 6, 10).

consensus sequences proposed thus far for archaeobacterial promoters or halobacterial rRNA promoters (13).

The upstream region and the 5' terminus of ORF 2022 were analyzed for stem loop structures in which the loop contains homology to the 3' terminus of *H. halobium* 16S rRNA. Such structures have been identified at the 5' termini of the *bop*, *brp*, and halo-opsin (*hop*) genes and have been proposed to function in translation initiation (3, 6, 10). No such stem loop structures were observed in the region at or immediately upstream of the 5' terminus of ORF 2022 (Fig. 2).

A 14-bp imperfect inverted repeat which may function as a potential transcriptional terminator was identified immediately downstream of the 3' terminus of ORF 2022 (Fig. 2) and is followed by a 28-bp region of 46% G + C content. Stem loops followed by a relatively A + T-rich region have been identified downstream of the *bop* and *hop* genes (6, 10).

ORF 2022 has a G + C content of 68%. In comparison, the *bop*, *brp*, and *hop* genes have G + C contents of 62, 68, and 67%, respectively (3, 6, 10). This is in contrast to the relatively A + T-rich regions located between the *bop* and *brp* genes (49% G + C) (3). The ORF 2022 sequence possesses no significant homology to any gene found in GenBank, although it has a 656-bp region of 46.8% homology to the *brp* gene.

The codon usage of ORF 2022 and the *bop*, *brp*, and *hop* genes is presented in Fig. 3. In ORF 2022, 91% of the codons end in G or C, which is similar to the *bop*, *brp*, and *hop* genes (82, 90, and 91%, respectively). The stop codon of ORF 2022 and of the other three genes is UGA. ORF 2022 has an RNY triplet (R, purine; N, any nucleotide; Y, pyrimidine) count of 37.2%. The other two reading frames of the same orientation in this 2,022-bp region have RNY counts of 17 and 26.4%. Thus, like the *brp* gene, ORF 2022 follows the highest RNY

rule of Shepherd (30), which is characteristic of most protein-encoding sequences.

The 2,022-bp ORF could encode 674 amino acids, giving a putative protein with a molecular weight of 73,334. About 20% of the amino acids in the putative ORF 2022 product are acidic, contributing to a predicted isoelectric point of 3.9. Analysis of the deduced amino acid sequence was performed by using a secondary structure prediction program developed by Finer-Moore and Stroud (12). The analysis indicated that there is a central region of the putative protein which contains alternating alpha helices and hydrophobic beta sheets.

Southern blot analysis of different halobacterial species and isolates. Conservation of the region containing ORF 2022 in other Pum⁺ halobacterial strains was examined by Southern blot analysis. Three probes of 550 to 1,134 bp, corresponding to the 5' terminus, the middle, and the 3' terminus of ORF 2022 as well as to a 1,600-bp region downstream of the 3' terminus of the ORF, were synthesized in vitro from ribo-probe vectors. The probes were hybridized to Southern blots of total genomic DNA from strains *H. salinarium* K, *H. cutirubrum*, YC81819-9, and GN101. The former two strains are very closely related to *H. halobium* and are currently thought to be different isolates of the same species. The latter strains are more distantly related to *H. halobium*. Our hybridization results showed that *H. salinarium* K and *H. cutirubrum* were identical to the wild type in the region containing ORF 2022. The only alteration detectable in Southern blots of strains GN101 and YC81819-9 was a *Pst*I site polymorphism in the region containing ORF 2022.

A number of strains previously characterized as Ret⁻ were examined by Southern blot analysis to determine if they contained alterations in ORF 2022. The region containing ORF 2022 in strains JW-5, Flx3R, and M144 was analyzed by using the probes described above. No alterations in the ORF 2022 of these mutants were detected.

DISCUSSION

Southern blot analysis and DNA sequencing of Bop mutants M86, W105, and W109 established that each of these mutants contain alterations in the region downstream of the *brp* gene. It is possible that undetected concurrent point mutations in the *bop* or *brp* genes are responsible for the Bop⁻ phenotype in these mutants. However, spontaneous Bop mutants derived from both the wild type (i.e., M86) and the revertant (i.e., W105 and W109) arise at the same frequency (10⁻⁴) (26, 29), and the majority of these mutants contain insertion elements in the *bop* gene or flanking regions (3, 25–27). Since such alterations are readily detectable by Southern blot analysis, it is more likely that the alterations characterized here are the only ones present in M86, W105, and W109 and thus are responsible for the Bop⁻ phenotype.

M86 is the only Bop mutant characterized so far in which the mutational defect is a deletion instead of an insertion. The presence of direct repeats in the wild-type sequence at the boundaries of the deletion suggests that the deletion in M86 was generated by homologous recombination analogous to that observed in *E. coli* (1). Moreover, the absence of inverted repeats within the deleted region makes it unlikely that the deletion in M86 represents an insertion element.

The two Bop mutants, W105 and W109, which contain insertions in ORF 2022 are derived from the revertant reIV-41. Bop mutants with insertions in ORF 2022 can also be derived from the wild type. A recently characterized

mutant, M133, is derived from strain R1, a gas vacuole-deficient derivative of the wild type. Preliminary analysis indicated that M133 contains an insertion in the 1.1-kbp *Pst*I fragment within ORF 2022, proximal to the site of the insertion in W105. Hence, insertions within ORF 2022 appear to have a similar phenotypic effect in both the wild type and the revertant.

Mutants M86, W105, W109, and M133 represent the only Bop mutants known to date which contain alterations in the 2,022-bp ORF. An earlier study of 33 Bop mutants revealed that 22 of these Bop mutants contain insertions in either the *bop* gene or its putative promoter region (3). Of the 22 mutants, 10 have the same insertion element in the same site, indicating that the *bop* gene contains an insertional hot spot. Eight mutants in this same study have insertions in the *brp* gene (3), while two mutants, M86 and M133, contain alterations in ORF 2022, and one mutant, IV-5, thus far has no detectable alterations. Thus, the *bop* and *brp* genes are the primary targets for insertion mutations, whereas ORF 2022 appears to be a secondary target. It is unclear why more Bop mutants have insertions in the *bop* and *brp* genes than in ORF 2022.

Southern blot analysis of various *Pum*⁺ species and isolates other than *H. halobium* indicated that the region containing ORF 2022 is relatively conserved. These results are in agreement with previous Southern analyses of the *bop* and *brp* gene regions using a larger (40 kbp) DNA probe (24, 25). The earlier studies also indicated that the entire 40-kbp region containing the *bop* and *brp* genes and ORF 2022 is missing in *Halobacterium volcanii* and *Halobacterium vallismortis*, strains which do not make purple membrane.

Southern blot analysis of several retinal-deficient (*Ret*⁻) mutants indicated that the 2,022-bp ORF of these mutants is not detectably altered. Similar analysis of the *brp* gene yielded the same results (unpublished observations). These results suggest that mutations occurring at sites other than the *bop* or *brp* genes or ORF 2022 confer the *Ret*⁻ phenotype. Hence, the coordination which appears to exist between bacterio-opsin and retinal synthesis (33, 34) is most likely not mediated by the *brp* gene or ORF 2022 or their putative products.

A number of observations lead us to propose that ORF 2022 represents a gene which is expressed in halobacteria. The codon usage of the ORF 2022 is similar to that of other halobacterial structural genes. The high percentage of acidic residues in the putative product of the ORF is characteristic of halobacterial proteins (2). Also, the secondary structure predicted for the ORF 2022 product is indicative of a typical alpha-beta-type soluble protein. Significantly, ORF 2022 is transcribed in halobacterial cells, as described in the accompanying paper (16).

Since mutants with alterations in ORF 2022 have a Bop⁻ phenotype, ORF 2022 or its product must somehow be required for *bop* gene expression or purple membrane assembly. Determination of the pattern of transcription of the *bop* gene, the *brp* gene, and ORF 2022 in a number of Bop mutants described in the accompanying paper (16) suggests that ORF 2022 is involved in activating *bop* and *brp* gene expression. Thus, we designate this putative gene *bat* for bacterio-opsin activator gene.

ACKNOWLEDGMENTS

We thank Kathleen McGinnis for her help in DNA sequencing, Dieter Oesterhelt for his gift of the M mutants, and Richard Shand for critical reading of the manuscript. We would also like to acknowledge Colin Clarke for his observation at the 1985 Archae-

bacterial meeting in Munich that 168 bp of available sequence distal to the 3' terminus of the *brp* gene contained an ORF following the RNY rule.

This work was supported by Public Health Service grant GM31785 from the National Institute of General Medical Sciences. D.L. was supported by Public Health Service National Research Service Award F 32 GM10333 from the National Institute of General Medical Sciences.

LITERATURE CITED

- Albertini, A., M. Hofer, M. Carlos, and J. Miller. 1982. On the formation of spontaneous deletions: the importance of short sequence homologies in the generation of large deletions. *Cell* 29:319-328.
- Bayley, S., and R. Morton. 1978. Recent developments in the molecular biology of extremely halophilic bacteria. *Crit. Rev. Microbiol.* 6:151-205.
- Betlach, M., J. Friedman, H. Boyer, and F. Pfeifer. 1984. Characterization of a halobacterial gene affecting bacterio-opsin gene expression. *Nucleic Acids Res.* 12:7949-7959.
- Betlach, M., F. Pfeifer, J. Friedman, and H. W. Boyer. 1983. Bacterio-opsin mutants of *Halobacterium halobium*. *Proc. Natl. Acad. Sci. USA* 80:1416-1420.
- Birnboim, H., and J. Doly. 1979. A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucleic Acids Res.* 7:1513-1523.
- Blanck, A., and D. Oesterhelt. 1987. The halo-opsin gene. II. Sequence, primary structure of halorhodopsin and comparison with bacteriorhodopsin. *EMBO J.* 6:265-273.
- Chen, E., and P. Seeburg. 1985. Supercoil sequencing: a fast and simple method for sequencing plasmid DNA. *4:165-170.*
- DasSarma, S., T. Darmerval, J. Jones, and N. Tandeau de Marsac. 1987. A plasmid-encoded gas vesicle protein gene in a halophilic archaeobacterium. *Mol. Microbiol.* 1:365-370.
- 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.
- Dunn, R., J. McCoy, M. Simsek, A. Majumdar, S. Chang, U. RajBhandary, and H. G. Khorana. 1981. The bacteriorhodopsin gene. *Proc. Natl. Acad. Sci. USA* 78:6744-6748.
- Ebert, K., C. Hanke, H. Delius, W. Goebel, and F. Pfeifer. 1987. A new insertion element, ISH26, from *Halobacterium halobium*. *Mol. Genet.* 206:81-87.
- Finer-Moore, J., and R. Stroud. 1984. Amphipathic analysis and possible formation of the ion channel in an acetylcholine receptor. *Proc. Natl. Acad. Sci. USA* 81:151-159.
- Hamilton, P., and J. Reeve. 1986. Molecular biology of archaeobacteria, p. 338-348. In L. Leive (ed.), *Microbiology*—1986. American Society for Microbiology, Washington, D.C.
- Kuhn, I., F. Stephenson, H. Boyer, and P. Greene. 1986. Positive-selection vectors utilizing lethality of the *Eco*RI endonuclease. *Gene* 44:253-263.
- Lechner, J., and M. Sumper. 1987. The primary structure of a procaryotic glycoprotein. *J. Biol. Chem.* 262:9724-9729.
- Leong, D., H. Boyer, and M. Betlach. 1988. Transcription of genes involved in bacterio-opsin gene expression in mutants of a halophilic archaeobacterium. *J. Bacteriol.* 170:4910-4915.
- Lipman, D., and W. Pearson. 1985. Rapid and sensitive protein similarity searches. *Science* 22:1435-1441.
- Maniatis, T., E. F. Fritsch, and J. Sambrook. 1982. *Molecular cloning: a laboratory manual*. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
- Maxam, A. M., and W. Gilbert. 1980. Sequencing end-labeled DNA with base-specific chemical cleavages. *Methods Enzymol.* 65:499-560.
- Melton, D., P. Krieg, M. Rebagliati, T. Maniatis, K. Zinn, and M. Green. 1984. Efficient *in vitro* synthesis of biologically active RNA and RNA hybridization probes from plasmids containing a bacteriophage SP6 promoter. *Nucleic Acids Res.* 12:7035-7056.
- Oesterhelt, D., and G. Krippahl. 1983. Phototrophic growth of halobacteria and its use for isolation of photosynthetically-deficient mutants. *Ann. Inst. Pasteur Microbiol.* 134B:137-150.

22. Oesterhelt, D., and W. Stoerkenius. 1973. Functions of a new photoreceptor membrane. *Proc. Natl. Acad. Sci. USA* **70**:2853–2857.
23. Pfeifer, F. 1988. Genetics of halobacteria, p. 105–133. In F. Rodriguez-Valera (ed.), *Halophilic bacteria*. CRC Press, Inc., Boca Raton, Fla.
24. Pfeifer, F., and M. Betlach. 1985. Genome organization of *Halobacterium halobium*: a 70 kb island of more AT rich DNA in the chromosome. *Mol. Gen. Genet.* **198**:449–455.
25. Pfeifer, F., M. Betlach, R. Martienssen, J. Friedman, and H. Boyer. 1983. Transposable elements of *Halobacterium halobium*. *Mol. Gen. Genet.* **191**:182–188.
26. Pfeifer, F., H. Boyer, and M. Betlach. 1985. Restoration of bacterio-opsin gene expression in a revertant of *Halobacterium halobium*. *J. Bacteriol.* **164**:414–420.
27. Pfeifer, F., J. Friedman, H. Boyer, and M. Betlach. 1984. Characterization of insertions affecting the expression of the bacterio-opsin gene in *Halobacterium halobium*. *Nucleic Acids Res.* **12**:2489–2497.
28. Pfeifer, F., G. Weidinger, and W. Goebel. 1981. Plasmids in halobacteria. *J. Bacteriol.* **145**:369–374.
29. Pfeifer, F., G. Weidinger, and W. Goebel. 1981. Genetic variability of *Halobacterium halobium*. *J. Bacteriol.* **145**:375–381.
30. Shepherd, J. 1981. Method to determine the reading frame of a protein from the purine/pyrimidine genome sequence and its possible evolutionary justification. *Proc. Natl. Acad. Sci. USA* **78**:1596–1600.
31. Spudich, E., and J. Spudich. 1982. Control of transmembrane ion fluxes to select halorhodopsin-deficient and other energy-transduction mutants of *Halobacterium halobium*. *Proc. Natl. Acad. Sci. USA* **79**:4308–4312.
32. Stoerkenius, W., and R. Bogomolni. 1982. Bacteriorhodopsin and related pigments of halobacteria. *Annu. Rev. Biochem.* **52**:587–616.
33. Sumper, M., and G. Herrmann. 1976. Biogenesis of purple membrane: regulation of bacterio-opsin synthesis. *FEBS Lett.* **69**:149–152.
34. Sumper, M., and G. Herrmann. 1976. Biosynthesis of purple membrane: control of retinal synthesis by bacterio-opsin. *FEBS Lett.* **71**:333–336.
35. Weber, J., and R. Bogomolni. 1981. P588 a second retinal containing pigment in *Halobacterium halobium*. *Photochem. Photobiol.* **33**:601–608.