The primary structures of helices A to G of three new bacteriorhodopsin-like retinal proteins

JUN OTOMO,* YASUKO URABE, HIROAKI TOMIOKA and HIROYUKI SASABE

Frontier Research Program, RIKEN Institute, Wako, Saitama 351-01, Japan

(Received 12 May 1992; revised 29 June 1992; accepted 4 August 1992)

The primary structures of helices A to G of all bacteriorhodopsin (BR)-like retinal proteins identified in newly isolated halobacteria have been determined from the nucleotide sequence of the BR-like protein genes. Using PCR methods, gene fragments encoding the A- to G-helix region of BR-like proteins were directly amplified from the total genomic DNA of the seven new halobacterial strains. Oligonucleotide primers corresponding to highly conserved regions in the helices A to G were designed from the nucleotide sequences of bacterioopsin (*bop*) and archaeopsin-I (*aop*-I), and some primers were effective for the amplification of the gene encoding C- to G-helix region of all new BR-like proteins. The primer corresponding to A-helix region was designed either from the nucleotide sequence of *bop* and *aop*-I or from the N-terminus amino acid sequence of a BR-like protein. Three new BR-like proteins were identified from the amino acid sequence, which was deduced from the nucleotide sequence of the genes encoding A- to G-helix region of the BR-like proteins. It was found that not only the amino acid sequence, but also the nucleotide sequence of the gene encoding the C- and G-helix region, in which a number of important residues for proton translocation are located, is highly conserved in three new BR-like proteins. Analysis of the primary structures of the A- to G-helix region of new BR-like proteins revealed that one has about 85% homology with aR-I and aR-II, and the rest have about 55% homology with halobium BR, aR-I and aR-II. From the results of the sequence analysis, we suggest that BR and BR-like proteins (functioning as light-driven proton pumps) can be classified into three types (BR type, aR type and a new type), and each of these types has 50–55% homology to each other in amino acid sequence.

Introduction

Bacteriorhodopsin (BR), a retinal protein found in the extremely halophilic archaeobacterium *Halobacterium halobium*, functions as a light-driven proton pump (Oesterhelt & Stoekenius, 1971; Stoekenius *et al.*, 1979). It consists of a retinal and a single polypeptide chain of amino acids whose sequence is known by both protein and gene sequencings (Khorana *et al.*, 1979; Ovchinnikov *et al.*, 1979; Dunn *et al.*, 1981). The mechanism of the light-driven proton pump has been extensively studied by site-directed mutagenesis (Khor-

ana, 1988), and Asp 85, Asp 96, Asp 212, and probably Arg 82 are thought responsible for proton translocation during the photocycle (Mogi *et al.*, 1988; Otto *et al.*, 1990). Point mutational analysis using halobacterial strain GRB also identified Asp 85 and Asp 96 as the essential residues for proton pumping (Soppa *et al.*, 1989). In addition to BR, other retinal proteins, archaerhodopsin-I and -II (aR-I and aR-II) having the function of a light-driven proton pump, were found in other halobacterial species isolated from Western Australia (Mukohata *et al.*, 1988; Uegaki *et al.*, 1991). The amino acid sequences of these proteins, which were deduced from the structural genes coding the polypeptides of aR-I and aR-II, also indicated that all essential residues for proton pumping are conserved (Sugiyama *et al.*, 1989; Uegaki *et al.*, 1991).

BR is a membrane protein that aggregates *in vivo* to form a highly ordered two-dimensional crystalline patch called purple membrane in the cell membrane. This formation is not only advantageous for retaining the high stability of BR, but also for fine structural analysis using

* Author for correspondence. Tel. 81 484 62 1111 ext. 6333; fax 81 484 65 8048.

Abbreviations: BR, bacteriorhodopsin; aR, archaerhodopsin; *bop*, bacterioopsin; *aop*, archaeopsin.

The nucleotide sequence data reported in this paper have been submitted to DDBJ/GenBank/EMBL and have been assigned the accession numbers D11056, D11057 and D11058.

electron cryo-microscopy to near-atomic resolution (Henderson *et al.*, 1990). It is known that in the purple membrane, BR trimers aggregate to form a hexagonal lattice and this trimer structure is likely to be most important for retaining purple membrane structure. Therefore, the elucidation of this trimer structure is probably important in understanding the two-dimensional crystallization mechanism of BR. Several amino acid residues are probably involved in this crystallization mechanism; therefore a mutational approach to this problem seems impractical, as numerous amino acid alterations would be required. An alternative approach would be to compare the amino acid sequences of BR in purple membrane and in other forms, and this is expected to reveal which residues are essential for the formation of purple membrane areas in the cell membrane.

We are conducting a study on the structure and function of bacterial rhodopsins isolated from new halobacterial strains (Otomo *et al.*, 1992). In order to study the mechanism of two-dimensional crystallization of BR, several BR-like proteins that do not make purple membrane in the cell membrane have been isolated. Determination of their primary structure is the most important characterization for these new BR-like proteins, and such studies require rapid and efficient DNA sequence analysis of their BR-like protein genes. In this study, we designed several oligonucleotide primers at the A-, C- and G-helix parts of BR, and succeeded in amplifying the gene encoding the A- to G-helix region of BR-like protein by the PCR method. This amplified DNA was then directly analysed using an automated DNA sequencer.

Methods

Oligonucleotides. The oligonucleotide primers used for PCR amplification and sequencing were obtained from Intertech (Japan).

Strains. Seven new halobacterial strains (damp, mex, mac, port, shark-I, shark-II and mex-II) were isolated from crude solar salts commercially produced in Mexico and Australia (Otomo *et al.*, 1992). *Natrobacterium pharaonis* (DSM 2160) was obtained from the Deutsche Sammlung von Mikroorganismen.

Isolation of DNA. Chromosomal DNA from *Halobacterium halobium* R1, *N. pharaonis* and halobacterial strains damp, mex, mac, port, shark-I, shark-II and mex-II was isolated and purified according to Vogelsang *et al.* (1983).

PCR amplification and agarose gel electrophoresis. PCR amplification was performed in a 100 µl reaction mixture containing 20 µM Tris/HCl (pH 8.3), 1.5 mM-MgCl₂, 25 mM-KCl, 1 µg gelatin, 50 µM each dNTP, 1 unit of Taq DNA polymerase (Takara) and 20 pmol of each primer (Saiki *et al.*, 1988). Chromosomal DNA (1 µg) was used as a template. All PCRs were performed for 25 cycles of denaturation at 94 °C for 1 min, primer annealing at 60 °C for 2 min and extension at 72 °C for 3 min. The thermal cycling was performed in a programmable heat

block (Thermal cyclic reactor, Model TC-100, Houei). The PCR products were fractionated by electrophoresis on a 4.0 % NuSieve 3:1 Agarose gel (Takara).

DNA sequencing. PCR DNA product was purified from the agarose gel using GENE CLEAN II (Funakoshi), and approx. 40 ng purified DNA was used for the sequence template. Sequencing reactions were performed using Taq DyeDeoxy Terminator Cycle Sequencing Kit (Applied Biosystems) and the nucleotide sequence of the PCR DNA product was determined by an automated DNA sequencing system (Applied Biosystems, Model 373A).

Isolation of BR-like protein from halobacterial strain port. Purification of port BR was performed using almost the same method for the halorhodopsin purification as described (Duschl *et al.*, 1988) with the following modifications. PIPES/NaOH buffer was used instead of Tris/HCl buffer. MEGA-9 was used for the elution of port BR, because port BR is more stable in MEGA-9 than in octylglucoside.

Protein sequence analysis. The amino acid sequence of the N-terminus of port BR was determined by automated Edman degradation. Purified port BR was separated by SDS-PAGE. Proteins were blotted on a polyvinylidene difluoride membrane (Millipore). Electroblothing was carried out at a constant current for 2 h at 4 °C in a blotting buffer composed of 30 mM-Tris, 17 mM-boric acid, 0.01 % SDS and 20 % (v/v) methanol using an electroblotter (TEFCO, Japan). Transferred proteins were stained with 0.1 % Amido black 10b (Bio-Rad) in 50 % (v/v) methanol and 10 % (v/v) acetic acid for 1 min and destained with distilled water. The stained band corresponding to port BR was cut out and subjected to protein sequencing (Applied Biosystems, Model 477A).

Results

A comparison of the amino acid sequence of halobium BR and that of aR-I, which has the same transport function, reveals that amino acids in the C-helix and G-helix are highly conserved (Dunn *et al.*, 1981; Sugiyama *et al.*, 1989). In the C-helix and G-helix regions, primers shown in Table 1 (haloBR-C-f and haloBR-G-r) were designed. Because Lys 216 (AAG) is the conserved amino acid as a retinal binding site in BR and aRs (Dunn *et al.*, 1981; Sugiyama *et al.*, 1989; Uegaki *et al.*, 1991), haloBR-G-r primer was designed as Lys 216 (AAG) to be the 3' end position of the oligonucleotide primer (Table 1). The 3' end of this primer should have high specificity for BR-like genes. Using these two oligonucleotide primers, a DNA fragment of about 400 bp could be amplified directly from the genomic DNA of strains R1, damp, mex and port (Fig. 1a). After recovery and purification of this fragment from the agarose gel, the DNA sequences of these amplified products were analysed (Fig. 4). Deduced amino acid sequences in this region are found to be highly homologous with that of halobium BR, and all the essential residues for proton pumping are conserved in these strains.

On the basis of amino acid sequence and nucleotide sequence of halobium BR and aR-I, the oligonucleotide

Table 1. Oligonucleotide primers for PCR

Name*	Sequence
haloBR-G-r	5'-AGG'ATG'A(G,A)(C,G)'CCG'AA(C,G)'CCG'ACC'TT-3'
haloBR-C-f	5'-CCG'CTG'(C,T)TG'(C,T)TG'(C,T)T(A,C)'GAC'CTC'G-3'
haloBR-A-f	5'-GG(A,T)'CG(C,T)'CC(C,G)'GAG'(A,T)(C,G)G'(C,A)TC'TGG'(C,T)T-3'
portBR-Nt-f-I	5'-ATG'GA(G,A)'CC(A,C,G)'GG(C,G)'TC(A,C,G)'GA(G,A)'GC-3'
portBR-Nt-f-II	5'-ATG'GA(G,A)'CC(A,C,G)'GG(C,G)'AG(C,T)'GA(G,A)'GC-3'
portBR-Nt-f-III	5'-ATG'GA(G,A)'CC(A,C,G)'GG(A,T)'(A,T)(C,G)(A,C,G,T)'GA(G,A)'GC-3'
portBR-A-f	5'-CC(A,C,G)'GG(C,G)'AG(C,T)'GA(G,A)'GCG'ATA'TGG-3'
portBR-D-r	5'-GAT'CAT'CAG'CAC'GTC'GAG'GCT'G-3'

* The designations f and r at the end of the name indicate sense and antisense strands, respectively.

Fig. 1. (a) Agarose gel electrophoresis of PCR products with the primers haloBR-G-r and haloBR-C-f. (b) Agarose gel electrophoresis of PCR products with the primers haloBR-G-r and haloBR-A-f. The PCR products were separated on a 4.0 % (w/v) NuSieve 3:1 Agarose gel. Lanes: 2, strain R1; 3, strain damp; 4, strain mex; 5, strain port, 1 and 6, a *Hae*III digest of YX174 DNA size marker, sizes (bp) are on the left. The arrow indicates the corresponding partial *bop* gene fragment.

primer in the A-helix region was designed from the highly conserved region (Table 1). Using these haloBR-A-f and haloBR-G-r primers, a DNA fragment of about 650 bp in length was amplified from the genomic DNA of strains R1, damp and mex, but not port (Fig. 1b). The corresponding DNA fragment from the genomic DNA of strain port could not be amplified under any PCR conditions with these primers, indicating low homology of A-helix region of port BR with halobium BR and aR-I.

Fig. 2. Agarose gel electrophoresis of PCR products using the chromosomal DNA from strain port as a template. The PCR products were separated on a 4.0 % (w/v) NuSieve 3:1 Agarose gel. PCR was performed with primer portBR-Nt-f-I (lane 2), portBR-Nt-f-II (lane 3) and portBR-Nt-f-III (lane 4). As a reverse primer, haloBR-G-r was used. An *Hae*III digest of YX174 DNA size markers is shown in lanes 1 and 5, sizes (bp) are on the left. The arrow indicates the corresponding partial *bop* gene fragment.

To design a primer of A-helix region of port BR, the protein was purified, and its amino acid sequence determined by automated Edman degradation. The N-terminus sequence of the non-cleaved polypeptide was determined to be MEPGSEAIWLWLGTAGMFLGMLYFIARGWGETDS. This amino acid sequence was similar to that of the A-helix region of both halobium BR and aR-I. From the amino acid sequence, we designed and synthesized three different mixtures of oligonucleotide primers (portBR-Nt-f-I, portBR-Nt-f-II, portBR-Nt-f-III) coding the first 7 amino acid residues of port BR (Table 1). Using these primers and the haloBR-G-r primer, DNA amplification by PCR was performed. Only portBR-Nt-f-II primer was successful in amplifying a DNA fragment of about 650 bp in length (Fig. 2). DNA sequencing of this fragment showed the sequence of C- to G-helix region to be 100% identical to that of the C-G fragment amplified with haloBR-C-f and haloBR-G-r primers.

Fig. 3. Agarose gel electrophoresis of PCR products with the primers haloBR-G-r and haloBR-A-f. The PCR products were separated on a 4.0 % (w/v) NuSieve 3:1 Agarose gel. Lanes: 2, strain R1; 3, strain mac; 4, strain shark-I; 5, strain shark-II; 6, strain mex-II; 7, *N. pharaonis*; 1 and 8, *Hae*III digest of YX174 DNA size marker, sizes (bp) are on the left. The arrow indicates the corresponding partial *bop* gene fragment.

Using haloBR-C-f and haloBR-G-r primers, the gene fragment encoding C- to G-helix region of BR-like proteins from other halobacterial strains (mac, shark-I, shark-II and mex-II) isolated from crude solar salts (Otomo *et al.*, 1992) was also amplified by PCR (Fig. 3). However, because *N. pharaonis* does not have any BR-like protein (Bivin & Stoeckenius, 1986), no PCR-amplified product was obtained for it. From the DNA sequence analysis of the gene fragment encoding C- to G-helix region of these BR-like proteins, we have found that the BR-like protein gene of strain mac is identical to that of strain mex, and that those of strains shark-I, shark-II and mex-II are about 95% homologous to that of strain port. Therefore, a common primer of portBR-A-f for strains shark-I, shark-II and mex-II was designed from the nucleotide sequence of BR-like protein gene from strain port (Table 1). Using this primer and a portBR-D-r primer whose nucleotide sequence is conserved in BR-like protein genes from strains port, shark-I, shark-II and mex-II, the gene fragment encoding A- to D-helix region of BR-like proteins from these strains could be amplified (data not shown). The gene fragment encoding A- to G-helix region of BR-like protein from strain mac was also amplified with haloBR-A-f and haloBR-G-r primers (data not shown).

The nucleotide sequence of A- to G-helix region of BR-like protein gene from all newly isolated halobacterial strains was determined using an automated DNA sequencing system and the amino acid sequence of all BR-like proteins was deduced from their nucleotide sequences of A- to G-helix region of BR-like protein genes. Although the nucleotide sequences of BR-like protein genes from the new halobacteria are not identical to each other, the amino acid sequences are identical in the A- to G-helix region between the two strains mex and

(a) mex bacteriorhodopsin

GG(A,T)CG(C,T)CC(C,G)GAG(A,T)(C,G)G(C,A)TCTGG(C,T)T										1	GGT
										G	G
ATC	GGC	ACG	CTA	CTA	ATG	CTC	ATC	GGG	ACC	TTC	TAC
I	G	T	L	L	M	L	I	G	T	F	Y
										31	TTC
										G	F
GGT	TGG	GGC	GTG	ACG	GAC	AAG	AAG	GCC	CGT	GAG	TAC
G	W	G	V	T	D	K	K	A	R	E	Y
										61	TTC
										G	F
ATC	CTC	GTG	CCG	GGG	ATC	121	TCG	GCA	GCG	TAT	CTG
I	L	V	P	G	I	GCG	A	A	A	L	S
										151	TTC
										G	F
GGC	ATC	GGG	CTG	ACG	ACG	GTC	GAG	GTG	GCC	181	TTC
G	I	G	L	T	T	V	E	V	A	G	M
										211	TTC
										G	F
GAA	ATC	TAC	TAC	GCC	CGG	TAT	GCG	GAC	TGG	CTG	TTC
E	I	Y	Y	A	R	Y	A	D	W	L	F
										241	TTC
										G	F
CTG	CTG	CTT	GAC	CTC	GCG	CTG	CTC	GCG	AAC	GCT	GAC
L	L	L	D	L	A	L	L	A	N	A	D
										301	TTC
										G	F
GGG	ACG	CTC	ATC	GGC	GTC	GAC	GCG	CTG	ATG	ATC	GTC
G	T	L	I	G	V	D	A	L	M	I	V
										331	TTC
										G	F
GGT	GCG	CTC	TCG	CAC	ACG	CCG	CTC	GCG	CGG	TAC	ACC
G	A	L	S	H	T	P	L	A	R	T	W
										361	TTC
										G	F
391	ACG	ATC	GCG	TTC	CTG	TTC	GTG	CTG	TAC	TAC	CTC
S	T	I	A	F	L	F	V	L	Y	Y	L
										421	TTC
										G	F
CGC	AGC	GCG	GCC	GCG	GAG	CTC	TCG	GAG	GAC	GTG	CAG
R	S	A	A	A	E	L	S	E	D	V	Q
										451	TTC
										G	F
ACG	CTA	ACG	GCG	CTC	GTC	GCC	GTC	CTC	TGG	ACG	GCG
T	L	L	L	L	V	A	V	L	W	T	A
										511	TTC
										G	F
TGG	ATC	ATC	GGG	ACC	GAA	GGL	GCC	GGC	GTC	GTC	GGC
W	I	I	G	T	E	G	A	G	V	V	S
										541	TTC
										G	F
ACG	CTC	GCG	TTC	ATG	GTG	CTC	GAC	GTG	ACG	613	TTC
T	L	A	F	M	V	D	D	V	T	G	G

mac, and among the three strains shark-I, shark-II and mex-II. Fig. 4 shows the nucleotide sequences and the deduced amino acid sequences of the A- to G-helix region of three new BR-like proteins. The nucleotide and the amino acid sequence of damp BR were identical to that of halobium BR.

Discussion

We have amplified the gene encoding C- to G-helix region of all BR-like proteins from seven new halobacterial strains without *in vivo* gene amplification. Therefore, these two primers may be useful for the gene amplification of all unidentified BR-like proteins. The mutational studies on BR have clearly indicated that the regions of two oligonucleotide primers (haloBR-C-f and haloBR-G-r) shown in Table 1 are essential for their proton-pumping function (Khorana, 1988). In addition, using a degenerate primer, we have succeeded in amplifying the gene encoding A- to G-helix region of port BR. This result shows that a mixture of primers

(b) port bacteriorhodopsin

ATGGA (G, A) CC (A, C, G) GG (C, G) AG (C, T) GA (G, A) GC 1
 A I W L
 TGG TTA GGT ACA GCG GGC 31
 W L G T A G M F L G G M T C T C G G C A T G C T C T A C T T C A T C G C G
 61
 C G C G G C T G G G G T G A G A C C G A C A G T A G G C G T C A A A A G T T C T A C A T C G C G
 R G G W G E T D S R R Q K F Y I A
 121
 A C A A T A C T T A C A G C A T C G C G T T C G T A A T T A C T C G C 151
 T I L I T A I A F V N Y L A M A
 181
 C T T G G C T T C G G G T T G A C G A T C G T T G A G T T C G C C G A G A G C A C C C C
 L G F G L T I V E F A G E E H P
 211
 A T C T A C T G G G C C G A T A C A G T G A C T G G T T G T T C A C G A C A C C G T T G C T G
 I Y W A R Y S D W L F T P L L
 271
 T T G T A C G A C C T C G G G T T G C T T G C T G G G G C A G A C C G T A A C A C T A T C A C C
 L Y D L G L L A G A D R N T I T
 331
 T C A C T C G T A G C C T C G A C G T G C T G A T G A T C G G G A C C G G C T G V C G C
 S L V S L D V L M I G T G L V A
 361
 A C G C T L A G C G A G G A A G C G G C G T G C T G C G C G G T A G C G A C G G C T G
 T S S A G G S G V L S G A G E R L
 421
 G T C T G G T G G G G C A T C A G T A C C G C C T T C T G C T G T C T G C T G T A C T T C
 V W W G I S T A F L L V L L Y F
 481
 C T G T T C A G C T C G T C C G G T C G G G T T G C A G A C T T G C C C A G T G A C A C C
 L F S L S L S R V A D L P C S D T
 511
 C G T A G C A C C T T C A A G A C G C T A C R A A C C T C G T G A C C G T G T G G T T G
 R S T F K T L R N L V A C T V V W L
 571
 541
 G T G T A C C G G T G T G T G G T G C T A T C G G G A C C G A G G G A T C G G C C T C G T C
 V Y F V W W L I G T E G I S L V
 601
 G G G A T C G G T A T C G A G A C G C C G G C T T C A T G G T C A T C G A C C T G A C C G C C
 G I G I E T A G F M V I D L T A
 631 634
A A G G T C G G (C, G) T T C G G (C, G) (C, T) T C A T C C T
 K V G F G

(c) shark bacteriorhodopsin

CC (A, C, G) GG (C, G) AG (C, T) GA (G, A) GCGATATGG 1
 A I W L TGG TTA
 GGT ACA GCG GGC ATG TTC CTC 31
 G T A G M F L G G M L Y F I A C G C G C G
 61
 TGG GGC GAA ACT GAC GGC CGG CGT CAG AAG TTC 91
 W G G E T D G R R Q K F Y I A T A T A
 121
 C T G A T C A C G C A A T T G C G T T C G T G A A T T A C T C G C G A T G C G C C T T G G C
 L I T A I A F V N Y L A M A L G
 181
 T T C G G C C T G A C G T T C A T C G A G T T C G G C G G C G A A C A C C A C C A T C T A C
 F G L T T F I E F G G E Q H P I Y
 241
 T G G G C G C G A T A C C G A C T G G C T G T T C A C A C G C C A T T G C T G C T G T A T
 W A R Y T D W L F T T P L L L Y
 271
 G A C C T C G G A C T G C T T G C A G G T G C A G A C C G C A A C A C C A T T T A C T C G C T C
 D L G L L A G A D R N T I Y S L
 331
 301
 G T C A G C C T C G A C G T G C T G A T G A T C G G T A C T G G T G T G G T C G A C G C T A
 V S L D V L M I G T G L V A T G V V A T L
 361
 A G T G C G G G C A G T G G C G T G C T G T C T G C C A G G T G C G A A C G G C T G T G C T G G
 S A G S G V L S A G E R L
 421
 T G G G G T A T C A G C A C T G C C T T C T G C T A G T C T G T A C T T C T G T T C
 W G I S T A F L L V L Y F L L
 481
 A G T T C G C T G T C C G G C G T T G C A A A C C T G C C C A G T G A C A C T R A G C
 S S L S G R V A N L P S D A C T R A G C
 511
 A C G T T C A A G A C G T T G C G C A A T C T C G T C A C T G T C G T G T G C T G T C T A C
 T F K T L R N L V A T V V W L V Y
 571
 541
 C C A G T C T G G T G C T G T C G T C G G C A G T G A G G A C T T G G C C T C G T C G G T A T
 P V W W L V G S E G L G L V G I
 601
 G G G A T C G A G A C G C C G G C T T C A T G G T C A T T G A C C T C G T C G C G C
 G I E T A G F M V I D L V A
 627
A A G G T C G G (C, G) T T C G G (C, G) (C, T) T C A T C C T
 K V G F G

Fig. 4. Nucleotide and the deduced amino acid sequences of bacterioopsin genes of strains mex (a), port (b) and shark (c). The single-letter amino acid code is used. The underlined positions of oligonucleotides indicate the part of primers for PCR amplification.

derived from an amino acid sequence is sufficient for PCR amplification of the target DNA fragment. Therefore, if the N-terminal amino acid sequence of unidentified BR-like proteins is determined, it is likely that PCR amplification of the gene encoding the A- to G-helix region will be possible, and that sequence analysis can be performed immediately.

We have compared the amino acid sequences from helices A to G of halobium BR and five BR-like proteins (including aR-I and aR-II) to estimate the amino acid residue identity. Fig. 5 compares the amino acid sequences of the new BR-like proteins identified in this study with those of halobium BR, aR-I and aR-II. It is possible to classify BR and BR-like proteins into three types: halobium BR, mex BR (or aR) or port BR. About 50–60% of amino acid residues are identical among each group, and more than 80% residues are identical in the same group. These three types are also characterized from the sequence of the segment between helices D and

E, where the kind and number of amino acids in this interhelical segment are completely different. However, they are identical in the same group, except for K128 of aR-II. The difference in the proton-pumping function among these three types is not known, but it is unlikely that this interhelical segment is important for this function, since it is relatively far from the Schiff base of the retinal (Henderson *et al.*, 1990). On the other hand, it is well known that halobium BR forms purple membrane in the cell membrane. BR from halobacterial strain damp and strain GRB (Soppa *et al.*, 1989), whose amino acid sequence is identical to that of halobium BR, also makes purple membrane. The BRs of strains mex and mac make a claret-coloured patch rather than purple membranes (Otomo *et al.*, 1992). This patch is probably the same as the 'claret membrane' in which aR-I or aR-II exists (Mukohata *et al.*, 1988; Mukohata *et al.*, 1991). We have also found that the density and the X-ray diffraction pattern of the claret membrane are different from those

PM	claret membrane				others		BR (%)	PM
	halobium GRB damp	aus-1	aus-2	mex mac	port	shark-1,2 mex-2		
	59	56	56	58	57	halobium GRB damp		
		86	83	54	51	aus-1		claret membrane
			89	54	52	aus-2		
				55	52	mex mac		
					93	port		others
						shark-1,2 mex-2		

Fig. 5. Comparisons of the amino acid sequence of all known BRs. Percentages of homology at amino acid level are shown. The black names are new halobacterial strains whose sequences were determined in this study. The percentage identity above 80% is indicated by shadow. BRs are classified into three types: purple membrane (PM), claret membrane and others. The percentages are calculated from positions 10 to 220 numbered as in Khorana *et al.*, (1979).

of purple membrane. Therefore, the two-dimensional aggregation of the BR pigment in the cell membrane is probably slightly different in form between a purple membrane and a claret membrane. Although port BR contains about the same amount of BR-like protein in the total membrane fraction as that of strain mex, neither purple membrane nor claret membrane was isolated by a wash in water (Otomo *et al.*, 1992). These results imply that port BR probably does not form two-dimensional crystalline patches in its cell membrane. Therefore, the sequence differences among the three protein types may reflect the character of their two-dimensional aggregation in the cell membrane.

On the basis of the result of Henderson *et al.* (1990), a helical wheel projection map for BR is shown in Fig. 6, indicating several identical amino acids for all BRs and all known bacterial rhodopsins (Dunn *et al.*, 1981; Sugiyama *et al.*, 1989; Uegaki *et al.*, 1991; Blanck & Oesterhelt, 1987; Blanck *et al.*, 1989; Lanyi *et al.*, 1990). Except for Pro 91, all conserved amino acid residues face towards the retinal in all bacterial rhodopsins. Moreover, most of them are located around the β -ionone ring of the

Fig. 6. A helical wheel projection model for halobium bacteriorhodopsin. The amino acids in the putative helical segments (A-G) suggested by Henderson *et al.* (1990) and the all-trans retinal are shown. Black labels indicate conserved residues in halobium BR, five BR-like proteins (including aR-I and aR-II), two HRs (halobium and pharaonis) and halobium SR. Grey labels indicate conserved residues in halobium BR and five BR-like proteins (including aR-I and aR-II). The positions from 221 to 228 are compared among halobium BR, aR-I and aR-II.

retinal. This indicates that these amino acid residues are important for binding the retinal to the apoprotein and most of them are suggested as residues of the retinal pocket by Henderson *et al.* (1990). They have also suggested that Val 49 and Met 145 could be retinal pocket residues. However, these are not conserved in all BRs and bacterial rhodopsins, and are probably not as important in retinal pocket formation. Although Pro 91 is conserved in all bacterial rhodopsins, it does not face the retinal. Site-directed mutagenesis substituting each proline residue has indicated that Pro 91 is important for chromophore regeneration but not indispensable for the proton-pumping function (Mogi *et al.*, 1989). They suggested that it might have a role in the three-dimensional interaction between the helical segment and retinal. The kink of helix C is at the position Pro 91 (Henderson *et al.*, 1990) and therefore may be important in the interaction between apoprotein and retinal.

A large number of identical amino acid residues common to all known BRs are located around the Schiff base of the retinal, and none of them except the conserved residues in all bacterial rhodopsins (black residues in Fig. 6) are conserved in halorhodopsin functioning as the light-driven chloride pump (Blanck & Oesterheld, 1987; Lanyi *et al.*, 1990). This strongly indicates that these residues (except the residues conserved in all bacterial rhodopsins) are important only for proton translocation. Therefore, these residues around the Schiff base of the retinal could play a role in ion selection, and are most likely to be important for proton translocation. On the other hand, almost all the amino acid residues at the external face in the helical wheel projection map and the interhelical segments are not conserved between all BRs. These residues are probably not important for proton pump function and retinal binding.

We have described a rapid and efficient method of DNA sequencing using PCR. Although the DNA sequence of whole structural genes of new BR-like proteins was not determined, almost all of the important amino acid sequence for the structure and function (i.e., A- to G-helix region of BR) was obtained. For comparative studies, we have isolated a number of new halobacterial strains (Otomo *et al.* 1992). As a first step in characterization, we identified the presence of BR-like pigment in these strains using flash spectrophotometry. However, determination of the amino acid sequence of their BR-like protein using the conventional gene cloning and the nucleotide sequencing is laborious. It takes several weeks to sequence one gene encoding a BR-like protein. We can not be sure whether the BR-like protein is new or not until the gene has been cloned and sequenced. Using the primers haloBR-C-f and haloBR-G-r, it was possible to amplify and sequence the gene

fragment encoding the C- to G-helix region of BR-like pigments from all new halobacterial strains in a few days. We have also amplified the gene fragment from a small amount of genomic DNA, which was obtained from the cells of a single colony by an easy DNA purification procedure (unpublished data). Therefore, PCR using these primers can be used for rapid screening of a large number of halobacterial strains for new BR-like proteins on the basis of the sequence of the C-to G-helix region.

We would like to thank Dr Y. Narahashi for peptide sequencing and Mr J. Honda for critical reading of the manuscript. This work was supported in part by a grant for 'Special Researchers' Basic Science Program' from the Science and Technology Agency of the Japanese Government.

References

- BIVIN, D. B. & STOECKENIUS, W. (1986). Photoactive retinal pigments in haloalkaliphilic bacteria. *Journal of General Microbiology* **132**, 2167–2177.
- BLANCK, A. & OESTERHELT, D. (1987). The halo-opsin gene. II. Sequence, primary structure of halorhodopsin and comparison with bacteriorhodopsin. *EMBO Journal* **6**, 265–273.
- BLANCK, A., OESTERHELT, D., FERRANDO, E., SCHEGK, E. S. & LOTTSPREICH, F. (1989). Primary structure of sensory rhodopsin I, a prokaryotic photoreceptor. *EMBO Journal* **8**, 3963–3971.
- DUNN, R., MCCOY, J., SIMESK, J. M., MAJUMDAR, A., CHANG, S. H., RAJBHANDARY, U. L. & KHORANA, H. G. (1981). The bacterioopsin gene. *Proceedings of the National Academy of Sciences of the United States of America* **78**, 6744–6748.
- DUSCHL, A., MCCLOSKEY, M. A. & LANYI, J. K. (1988). Functional reconstitution of halorhodopsin. *Journal of Biological Chemistry* **263**, 17016–17022.
- HENDERSON, R., BALDWIN, J. M., CESKA, T. A., ZEMLIN, F., BECKMANN, E. & DOWNING, K. H. (1990). Model for the structure of bacteriorhodopsin based on high-resolution electron cryo-microscopy. *Journal of Molecular Biology* **213**, 899–929.
- KHORANA, H. G. (1988). Bacteriorhodopsin, a membrane protein that uses light to translocate protons. *Journal of Biological Chemistry* **263**, 7439–7442.
- KHORANA, H. G., GERBER, G. E., HERLIHY, W. C., GRAY, C. P., ANDEREGG, R. J., NIHEI, K. & BIEMANN, K. (1979). Amino acid sequence of bacteriorhodopsin. *Proceedings of the National Academy of Sciences of the United States of America* **76**, 5046–5050.
- LANYI, J. K., DUSCHL, A., HATFIELD, G. W., MAY, K. & OESTERHELT, D. (1990). The primary structure of a halorhodopsin from *Natronobacterium pharaonis*. *Journal of Biological Chemistry* **265**, 1253–1260.
- MOGI, T., STERN, L. J., MARTI, T., CHAO, B. H. & KHORANA, H. G. (1988). Aspartic acid substitutions affect proton translocation by bacteriorhodopsin. *Proceedings of the National Academy of Sciences of the United States of America* **85**, 4148–4152.
- MOGI, T., STERN, L. J., MARTI, T., CHAO, B. H. & KHORANA, H. G. (1989). Structure-function studies on Bacteriorhodopsin. VIII. Substitutions of the membrane-embedded prolines 50, 91, and 186: the effects are determined by the substituting amino acids. *Journal of Biological Chemistry* **264**, 14192–14196.
- MUKOHATA, Y., SUGIYAMA, Y., IHARA, K. & YOSHIDA, M. (1988). An Australian halobacterium contains a novel proton pump retinal protein: archaerhodopsin. *Biochemical and Biophysical Research Communications* **151**, 1339–1345.
- MUKOHATA, Y., IHARA, K., UEGAKI, K., MIYASAKA, Y. & SUGIYAMA, Y. (1991). Australian halobacteria and their retinal-protein ion pumps. *Photochemistry and Photobiology* **54**, 1039–1045.

- OESTERHELT, D. & STOECKENIUS, W. (1971). Rhodopsin-like protein from the purple membrane of *Halobacterium halobium*. *Nature New Biology* **233**, 149–152.
- OTOMO, J., TOMIOKA, H. & SASABE, H. (1992). Bacterial rhodopsins of newly isolated halobacteria. *Journal of General Microbiology* **138**, 1027–1037.
- OTTO, H., MARTI, T., HOLZ, M., MOGI, T., STERN, L. J., ENGEL, F., KHORANA, H. G. & HEYN, M. P. (1990). Substitution of amino acids Asp-85, Asp-96, and Arg-82 in bacteriorhodopsin affects the proton release phase of the pump and pK of the Schiff base. *Proceedings of the National Academy of Sciences of the United States of America* **87**, 1018–1022.
- OVCHINNIKOV, YU. A., ABDULAEV, N. G., FEIGINA, M. Y., KISELEV, A. V. & LOBANOV, N. A. (1979). The structural basis of the functioning of bacteriorhodopsin: an overview. *FEBS Letters* **100**, 219–224.
- SAIKI, G. K., GELFAND, D. H., STOFFEL, S., SCHARF, S., HIGUCHI, R., HORN, G. T., MULLIS, K. B. & ERLICH, H. A. (1988). Primer-directed enzymatic amplification of DNA with thermostable DNA polymerase. *Science* **239**, 487–491.
- SOPPA, J., OTOMO, J., STRAUB, J., TITTOR, J., MEESSEN, S. & OESTERHELT, D. (1989). Bacteriorhodopsin mutants of *Halobacterium* sp. GRB. II. Characterization of mutants. *Journal of Biological Chemistry* **264**, 13049–13056.
- STOECKENIUS, W., LOZIER, R. H. & BOGOMOLNI, R. A. (1979). Bacteriorhodopsin and the purple membrane of Halobacteria. *Biochimica et Biophysica Acta* **505**, 215–278.
- SUGIYAMA, Y., MAEDA, M., FUTAI, M. & MUKOHATA, Y. (1989). Isolation of a gene that encodes a new retinal protein, archaerhodopsin, from *Halobacterium* sp. aus-1. *Journal of Biological Chemistry* **264**, 20859–20862.
- UEGAKI, K., SUGIYAMA, Y. & MUKOHATA, Y. (1991). Archaerhodopsin-2 from *Halobacterium* sp. aus-2 further reveals essential amino acid residues for light-driven proton pumps. *Archives of Biochemistry and Biophysics* **286**, 107–110.
- VOGELSANG, H., OERTEL, W. & OESTERHELT, D. (1983). The isolation of the bacterioopsin gene. *Methods in Enzymology* **97**, 226–241.