Bacterioopsin, Haloopsin, and Sensory Opsin I of the Halobacterial Isolate *Halobacterium* sp. Strain SG1: Three New Members of a Growing Family

JÖRG SOPPA,* JUTTA DUSCHL, AND DIETER OESTERHELT
Max-Planck-Institut für Biochemie, Am Klopferspitz, W8033 Martinsried, Germany
Received 24 August 1992/Accepted 18 January 1993

The genes coding for bacterioopsin, haloopsin, and sensory opsin I of a halobacterial isolate from the Red Sea called *Halobacterium* sp. strain SG1 have been cloned and sequenced. The deduced protein sequences were aligned to the previously known halobacterial retinal proteins. The addition of these new sequences lowered the number of conserved residues to only 23 amino acids, or 8% of the alignment. Data base searches with two highly conserved peptides as well as with an alignment profile yielded no significant similarity to any other protein, so the halobacterial retinal proteins should be regarded as a distinct protein family. The protein alignment was used to make predictions about the structure of the retinal proteins as well as about the amino acids in contact with retinal proteins. These results were in excellent agreement with the structural model of bacteriorhodopsin of *Halobacterium halobium* as well as with mutant studies, indicating that (i) structure predictions based on the sequences of a membrane protein family can be quite accurate; (ii) halorhodopsin and sensory rhodopsin I have tertiary structures similar to that of bacteriorhodopsin; (iii) conserved amino acids do not take part in reactions specific for one group of proteins, e.g., proton translocation for bacterio- rhodopsins, but have a crucial role in determining the conformation and reactions of the chromophore; and (iv) the general mode of action (light-induced chromophore and protein movements) is the same for all halobacterial retinal proteins, ion pumps as well as sensors.

Three different types of retinal proteins have been found in halobacteria: bacteriorhodopsin (BR), a light-driven proton pump (35); halorhodopsin (HR), a light-driven chloride pump (43); and sensory rhodopsins (SR) I and II, which are involved in phototaxis (5, 50). For BR of *Halobacterium halobium*, a model for the tertiary structure based on electron diffraction methods has been proposed (23). *H. halobium* has recently been renamed *Halobacterium salinarium*, but in this report the old name will be used because all work on retinal proteins has been done with this designation. In addition to the structural study, much information on the reaction mechanism and functionally important amino acids has been gathered through the analysis of wild-type and mutant proteins, using different biophysical methods (31, 32, 51). In contrast, much less is known about the other three proteins. Nevertheless, it has become evident that the proton pump BR and the chloride pump HR share a number of characteristic features such as the photocycle. Therefore, it has been proposed that the principal mode of action is the same for the anion and cation transporters (36). The primary structures of the apoproteins of BR, HR, and SR I from *H. halobium*, HR of *Natronobacter pharaonis*, and BR of two isolates from Australia have been determined (3, 4, 11, 29, 48, 53). The sequences showed a high degree of similarity, making it likely that the halobacterial retinal proteins belong to a protein family which has evolved from a common ancestor. Knowledge of additional members of this family should lower the number of conserved amino acids between all sequences and thereby reveal the minimal number of amino acids involved in retinal protein functions, e.g., retinal binding or color determination, as well as group-specific amino acids which are only conserved among the ion pumps or the sensors.

We report the cloning and sequencing of the genes coding for the apoproteins bacterioopsin (Bop), haloopsin (Hop), and sensory opsin I (Sop I) of a new halobacterial isolate from the Red Sea. A comparison of all known halobacterial retinal proteins is given, and the implications for the understanding of retinal protein structure and function are discussed. In addition, the alignment was used to search the available data bases for other proteins showing similarity to this protein family.

MATERIALS AND METHODS

*Halobacterium* sp. strain SG1. *Halobacterium* sp. strain SG1 was isolated from Sabkha el gavish in the Sinai and will be described in detail elsewhere (33). It is a rod-shaped bacterium and grows optimally at pH 8. This strain possesses BR, HR, and SR I.

Cloning and sequencing. For identification of the genes, standard techniques for cloning and sequencing were used (42, 45). Genomic DNA of isolate SG1 was cut with different restriction enzymes (Boehringer Mannheim, Gibco BRL, AGS). After electrophoretic separation, it was transferred to nylon filters (Biodyne A; Pall) by capillary blotting by the method of Southern (46). Cloned fragments of the Bop, Hop, and Sop I genes of *H. halobium* were radioactively labeled with [*32*P]dCTP (Amersham) by random priming (Boehringer Mannheim) and used as probes in hybridization experiments under different conditions of stringency. The best results were obtained by hybridizing for 4 days at 32°C in 40% formamide-2.5× SSC (1× SSC is 0.15 M NaCl plus 0.015 M sodium citrate)-0.1% sodium dodecyl sulfate-5× Denhardt's solution.

After preparative digestion of genomic DNA, DNA frag-
ments of the desired sizes were isolated by electrophoresis from agarose gels. They were ligated into the vector pGem3 (Promega Biotech) which had been dephosphorylated with calf intestinal phosphatase (Boehringer Mannheim). Positive clones were detected by colony hybridization. The following fragments were ligated into pGem3: (i) a 9.5-kbp BamHI and a 1.1-kbp PvuII fragment hybridizing with the Bop-specific probe, (ii) an 8.5-kbp BamHI and a 1.1-kbp PvuII fragment hybridizing with the Hop-specific probe, and (iii) a 4.0-kbp BamHI and a 1.3-kbp SacI fragment hybridizing with the Sop I-specific probe. The ends of the fragments were sequenced with the T7 and SP6 primers, and nucleotide sequences corresponding to part of the Hop and Sop I genes were identified on the smaller fragments. To locate the Bop gene on the 9.5-kbp BamHI fragment, a restriction map was constructed and fragments which hybridized with the Hop probe were identified. A 0.9-kbp HincII fragment was subcloned, and part of the nucleotide sequence was determined with vector-specific primers. The remaining parts of the three genes were sequenced by gene walking, using synthetic oligonucleotide primers. Both strands of the genes were sequenced.

**Computer programs.** For analysis of the sequences, various programs of software packages of the Protein Identification Resource (PIR; National Institutes of Health, Bethesda, Md.) and the Genetics Computer Group (GCC; Milwaukee, Wis.) were used on a Microvax computer. The sequence alignment was done by a combination of computer comparison and manual editing as described below. Other programs include Argos (2) for sequence comparison and PAUP (phylogenetic analysis using parsimony [49]) and the software package of Felsenstein (15) for calculating evolutionary trees and distances.

**Alignment of deduced protein sequences.** The alignment of the deduced protein sequences was done by a combination of computer alignment and manual editing. First, the 21 pairwise alignments of the seven nonidentical retinal proteins were computed with the program Piralign (PIR). The actual alignment was constructed on the basis of these alignments by manual editing, using the multiple sequence editor Lineup (GCC). The rationale of the alignment followed the concept of progressive sequence alignment (16); i.e., first the sequences within a group were aligned and then the sequence groups were aligned. During the course of this study another BOP sequence from a second isolate from Australia (*Halobacterium* sp. strain aus-2) became available (53) and was introduced into the alignment solely by manual editing.

**Nomenclature.** The growing number of halobacterial retinal protein sequences can make the discussion of the role of single amino acids confusing. The nomenclature we use here and propose for use during discussions on the retinal protein family is the following. On the basis of protein family alignment, the residue number follows the amino acid name, and in parentheses the corresponding amino acid in a specific sequence is identified by the abbreviations of species and sequence name and the amino acid residue and number. For example, P75 (Hhbp P50) would describe a proline, which in the conventional numbering of Bop of *H. halobium* is residue 50.

**Nucleotide sequence accession numbers.** The sequences were deposited in the EMBL nucleotide sequence data base and were given the following accession numbers: Bop, X70291; Hop, X70292; Sop I, X70290. They are also available from us in printed form or on a floppy disk, upon request.

### RESULTS

**Cloning and sequence determination, alignment of the deduced protein sequences, and similarities of genes and proteins.** The genes coding for Bop, Hop, and Sop I of the halobacterial isolate strain SG1 could be cloned by using the respective genes of *H. halobium* as heterologous probes, as described above. We found it crucial to adjust the hybridization conditions carefully since the signals were more than 10 times weaker than with genomic DNA of *H. halobium* as a homologous control. The assignments of the sequences to Bop, Hop, and Sop I were based on two criteria: (i) by using the appropriate hybridization conditions, a single unique band hybridized with each of the probes; and (ii) each of the deduced protein sequences had a higher degree of similarity to one of the known retinal protein groups than to the others. The degree of sequence similarity between Sop I and Sop II sequences was also determined with the Sop II sequence obtained recently (26a). First, we tried to set up an alignment of the deduced protein sequences and the previously known sequences of halobacterial retinal proteins with the help of various computer programs, but the results in less conserved parts of the proteins were very sensitive to the comparison parameters used. We therefore constructed the final alignment by a combination of computer alignment and manual editing. We found that the use of short sequence motifs conserved between some or all of the sequences could improve the results considerably. The final alignment is shown in Fig. 1.

The similarities of all sequence pairs were calculated by dividing the number of identical residues by the comparable length of the sequences. The results are summarized in Fig. 2. The degree of amino acid sequence identity within one sequence group varied between 54 and 84%; pairs of proteins from different groups of retinal proteins shared between 19 and 32% identical amino acid residues. The similarity between Bops and Hops was slightly higher than when either was compared with Sop I. In addition, many amino acids were replaced by similar residues. The conserved amino acids were not evenly distributed in regions thought to form membrane-spanning helices for BR of *H. halobium* (23). The same is true for more than 80% of the amino acids conserved between two of the three protein groups (Fig. 1).

The nucleotide sequences of the retinal protein genes were aligned in accordance with the protein sequence alignment. This alignment was used to calculate genetic distances and to build an evolutionary tree, using the software of Felsenstein and the PAUP program (data not shown). It showed, regardless of the algorithms used, that there was a greater distance between the Sop I genes and the Bop or Hop genes than between the last two groups. However, because the halobacterial retinal proteins were not related to any other proteins (see below), the tree could not be rooted.

**Searching for similar proteins.** We used the information contained in the protein alignment in two ways to search for similar proteins in nucleotide sequence data bases. First, we chose two peptides which corresponded to the most highly conserved parts of the proteins, namely, RY...W...TP (no. 123 to 132) and F...L... . . . . . . W...YP...W...G...G (no. 213 to 237). With these sequences the combined protein sequence database MIPSX was searched repeatedly, allowing an increasing number of mismatches. No protein of meaningful similarity could be found. The second approach took advantage of the program Profilemake (GCC), which generates a "profile" of a sequence group. The profile has the
length of the alignment, and each position is weighted by the degree of conservation of an amino acid residue at that position. The protein sequence data bank was searched by using the profile obtained from the halobacterial retinal proteins; again, no other profile showed significant similarity to this group. Thus, the halobacterial retinal proteins must be regarded as a separate protein family.

Predicting membrane-spanning helices. Two approaches were used to predict the location of membrane-spanning helices. The first approach is based on the average hydrophobicity of the retinal proteins. The hydrophobicity profiles of all nonidentical halobacterial retinal proteins were calculated with the Peptide program (21) in conjunction with the hydrophobicity scale of Engelman et al. (14). The mean hydrophobicities of the helices were calculated, and regions with higher hydrophobicity were interpreted to be membrane-spanning regions. For each of the seven helices, the boundaries determined for the individual sequences were averaged. The results are listed in Table 1.

The second approach is based on the assumption that for a family of membrane proteins with functional sites within the plane of the membrane, e.g., transport proteins, the location of conserved amino acids can be used to predict the location of transmembrane helices. We used the following assumptions for our predictions: (i) each membrane-spanning helix is 25 amino acids long; (ii) the conserved amino acids are mostly located in the middle of the membrane; (iii) glycines, serines, asparagines, and aspartates are often located at the N-caps and glycines are located at the C-caps of

<table>
<thead>
<tr>
<th>TABLE 1. Prediction of helix boundariesa</th>
</tr>
</thead>
<tbody>
<tr>
<td>Avg hydrophobicity</td>
</tr>
<tr>
<td>-------------------</td>
</tr>
<tr>
<td>34-57</td>
</tr>
<tr>
<td>68-95</td>
</tr>
<tr>
<td>127-143</td>
</tr>
<tr>
<td>148-171</td>
</tr>
<tr>
<td>177-198</td>
</tr>
<tr>
<td>216-239</td>
</tr>
<tr>
<td>245-269</td>
</tr>
</tbody>
</table>

* The prediction methods are explained in the text; the numbering is the same as in Fig. 1.
TABLE 2. Amino acid composition

<table>
<thead>
<tr>
<th>Amino acid</th>
<th>%a</th>
<th>Ret</th>
<th>Prot</th>
<th>Eco</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gly</td>
<td>92.0</td>
<td>84.4</td>
<td>89.9</td>
<td></td>
</tr>
<tr>
<td>Ala</td>
<td>131.0</td>
<td>101.3</td>
<td>88.3</td>
<td></td>
</tr>
<tr>
<td>Val</td>
<td>98.6</td>
<td>85.5</td>
<td>82.2</td>
<td></td>
</tr>
<tr>
<td>Leu</td>
<td>144.3</td>
<td>73.0</td>
<td>83.5</td>
<td></td>
</tr>
<tr>
<td>Ile</td>
<td>61.2</td>
<td>47.3</td>
<td>60.4</td>
<td></td>
</tr>
<tr>
<td>Phe</td>
<td>46.1</td>
<td>24.6</td>
<td>28.5</td>
<td></td>
</tr>
<tr>
<td>Tyr</td>
<td>39.3</td>
<td>24.4</td>
<td>30.2</td>
<td></td>
</tr>
<tr>
<td>Trp</td>
<td>28.3</td>
<td>6.5</td>
<td>8.0</td>
<td></td>
</tr>
<tr>
<td>Pro</td>
<td>37.0</td>
<td>40.7</td>
<td>35.7</td>
<td></td>
</tr>
<tr>
<td>His</td>
<td>4.5</td>
<td>19.5</td>
<td>15.0</td>
<td></td>
</tr>
<tr>
<td>Ser</td>
<td>55.9</td>
<td>60.7</td>
<td>51.1</td>
<td></td>
</tr>
<tr>
<td>Thr</td>
<td>74.0</td>
<td>68.0</td>
<td>54.4</td>
<td></td>
</tr>
<tr>
<td>Met</td>
<td>29.8</td>
<td>20.2</td>
<td>24.7</td>
<td></td>
</tr>
<tr>
<td>Cys</td>
<td>3.4</td>
<td>4.6</td>
<td>6.5</td>
<td></td>
</tr>
<tr>
<td>Glu</td>
<td>36.3</td>
<td>95.1</td>
<td>72.9</td>
<td></td>
</tr>
<tr>
<td>Gln</td>
<td>11.0</td>
<td>31.8</td>
<td>39.2</td>
<td></td>
</tr>
<tr>
<td>Asp</td>
<td>40.8</td>
<td>92.7</td>
<td>64.9</td>
<td></td>
</tr>
<tr>
<td>Asn</td>
<td>12.1</td>
<td>36.4</td>
<td>42.5</td>
<td></td>
</tr>
<tr>
<td>Lys</td>
<td>17.0</td>
<td>27.3</td>
<td>62.6</td>
<td></td>
</tr>
<tr>
<td>Arg</td>
<td>31.7</td>
<td>53.0</td>
<td>56.6</td>
<td></td>
</tr>
<tr>
<td>Asp + Glu</td>
<td>77.1</td>
<td>187.8</td>
<td>137.8</td>
<td></td>
</tr>
</tbody>
</table>

Acidic preference

- a Ret, halobacterial retinal proteins; Prot, halobacterial proteins without the retinal proteins; Eco, highly expressed proteins of E. coli.
- b Sum of acidic residues.
- c Difference of acidic and basic residues.

helices (40); and (iv) each halobacterial retinal protein contains seven membrane-spanning helices, as known from BR of H. halobium. The proposed helix boundaries based on these assumptions and the sequence alignment (Fig. 1) are shown in Table 1.

**Amino acid composition.** The amino acid compositions of the retinal proteins and all other halobacterial proteins available in the data bases were calculated. The results are shown in Table 2 together with the amino acid composition of highly expressed proteins of Escherichia coli. The membrane-bound retinal proteins are enriched in alanine, leucine, phenylalanine, and tryptophan, whereas the charged amino acids and histidine are underrepresented. In contrast, the soluble halobacterial proteins are specifically enriched in acidic residues (about 20%). The protein groups deviate most in the preference for acidic residues, which is about 11% for soluble halobacterial proteins, 3% for the retinal proteins, and 2% for the highly expressed proteins of E. coli. Interestingly, the frequency of proline is nearly the same for all three protein groups.

**DISCUSSION**

Similarity of the proteins, sequence alignment, the search for similar proteins, and evolutionary considerations. Including the three deduced protein sequences of this study, 10 retinal protein sequences are now known: Bops of H. halobium (11), Halobacterium sp. strain GRB (45), Halobacterium sp. strain aus-1 (48), Halobacterium sp. strain aus-2 (53), and Halobacterium sp. strain SG1; Hops of H. halobium (3), N. pharaonis (29), and Halobacterium sp. strain SG1; and Sop I of H. halobium (4) and Halobacterium sp. strain SG1. The Bop sequences of Halobacterium sp. strain SG1 (isolated in Israel) and Halobacterium sp. strain aus-1 (isolated in Australia) are identical. The same is true for the Bop sequences of Halobacterium sp. strain GRB (isolated in France [12]) and H. halobium. The history of H. halobium is somewhat obscure. Several strains were isolated early in this century from salted fish in different parts of the world (20). The identity of these two pairs of sequences is not due to constraints on the primary sequence by the protein function, because other Bops have only about 50% identical amino acids (Fig. 2). It might reflect the history of invasion of different halobacterial habitats. Investigating the similarity of the four species with molecules suited for phylogenetic studies such as rRNA would give further insight.

Apart from the two pairs of identical sequences, the nonidentical retinal proteins also show a high degree of similarity. In addition, many of the nonidentical positions are occupied by similar amino acids in different sequences. The conserved amino acids are not evenly distributed, but are concentrated in regions which had been identified as membrane-spanning helices in the Bop of H. halobium.

The identification of other halobacterial retinal proteins (especially the second Sop I and the third Hop) contributed to the following results: (i) it underscored that the halobacterial retinal proteins form a protein family, which had been proposed previously (4); (ii) it allowed a refinement of the sequence alignment; (iii) it reduced the number of amino acids conserved between all sequences to 8%; and (iv) it reduced the number of conserved amino acids within sequence groups. The impact of these results for the prediction of structural features and amino acids of functional importance is discussed below.

Sequence comparisons between H. halobium BR and other proteins, specifically, the visual pigments and seven helix G-protein coupled receptors, have been made, and despite some analogies, e.g., the number of helices and the chromophore attachment, no similarities in the primary structures were found (24). In this study the emergent information content of the sequence alignment of the retinal protein family was used for search for related proteins in the data base. Neither the short “signature” peptides nor the alignment profile, however, yielded any other protein with significant similarity. Thus, the halobacterial retinal proteins must be regarded as a distinct protein family not related to any known protein. The lack of similarity to other proteins led also to the inability to root an evolutionary tree on the basis of the gene sequences. Assuming equal mutational rates, the greater distance between the Sop I genes and the other two groups could be explained by an early division of a primordial ion (proton?) pump and a sensory protein before the diversion of the pump into a cation and anion derivative. But the possibility that the differences in protein function led to faster accumulation of mutations in one branch of the retinal protein family cannot be ruled out. Therefore, elucidation of the evolution of the retinal proteins must wait until these proteins can be linked to newly discovered proteins, e.g., rhodopsins of other species.

Using a family of membrane proteins to predict structural features and important amino acids. The determination of the three-dimensional structure of membrane proteins is difficult, and only a very few structures have been solved. Therefore, indirect methods such as hydrophobicity profiles of the primary structure, studies of proteolytic cleavage, and antibody-binding sites or spectroscopic methods are commonly used as bases for structural models of membrane proteins. This had also been done by various groups for BR of H. halobium (1, 13, 14, 26, 38). However, these models differed dramatically in some parts of the protein, placing the
FIG. 3. Helical wheel representation of a halobacterial retinal protein. The helices are oriented so that the conserved amino acids point towards the inner space formed by the helical bundle. The angular orientation corresponds to the one determined experimentally for BR of *H. halobium* (23).

same amino acids either in the interior of the membrane or in solvent accessible loops. We found it interesting to use the information for a protein family rather than for a single sequence to predict the structure of the proteins and important amino acids. The results can be compared with a refined model for BR of *H. halobium* based on electron microscopy (23), which can give an accuracy estimate of the prediction methods based on membrane protein families.

We used two approaches to predict the location of membrane-spanning helices: (i) the average of the helix boundaries predicted for each of the retinal proteins on the basis of hydrophobicity calculations and (ii) the location of conserved amino acids in combination with "helix end rules." The results of both methods are in good agreement with results based on the structural model of BR of *H. halobium*: the average differences are less than three and less than two amino acids, respectively. In addition, it has to be taken into account that the resolution of the structural model is still limited, and Henderson et al. (23) state that "the residue number at which each helix terminates is uncertain at the top and the bottom of each helix by at least one residue." It will be interesting to determine the helix boundaries in a high-resolution structure and see whether the helix end rules (40) derived from helices of soluble proteins of known structure also hold true for membrane-spanning helices.

The conserved amino acids should play a role in a function common to all halobacterial retinal proteins, especially for binding retinal and influencing its conformation and electronic state. Thereby, important properties of the proteins such as Schiff's base pK, color, and restriction of the light-induced isomerization of the retinal to just the 13–14 bond as well as a lowered energy barrier for the thermal back-isomerization around this bond are determined. If one assumes that prolines have another functional role and that not all amino acids which are conserved over the whole length of helix F are in contact with retinal, one can predict that about 16 of the 23 conserved amino acids are located near the retinal in all halobacterial retinal proteins (Fig. 1). Of these, 13 were described by Henderson et al. (23) to be part of the retinal-binding pocket of *H. halobium* BR. The locations of the conserved amino acids make it possible to predict the orientation of the helices. Fig. 3 shows a helical wheel representation of a prototypic retinal protein in which the conserved amino acids thought to interact with retinal are oriented towards the interhelical space. This orientation is again consistent with the structural model of BR of *H. halobium*.

**Discussion of several groups of conserved amino acids.** We will now discuss some of the 23 conserved amino acids with respect to their possible functions. This includes results from studies on BR of *H. halobium*, e.g., the structural model, mutant BRs, and Fourier transform infrared spectroscopic data. Three tryptophan residues are conserved: W127 (Hhbop W86), W224 (Hhbop W182), and W231 (Hhbop W189). They were all found to be in the retinal-binding pocket of BR of *H. halobium* (23). Earlier, it had been shown that three to four tryptophans are excitonically coupled to retinal, i.e., must be in close contact with retinal (39). In BR of *H. halobium*, a fourth tryptophan at position 180 (Hhbop W138) is found in the retinal-binding pocket (23), and its replacement by arginine leads to a large color shift of the chromoprotein (44, 45). At this position an aromatic amino acid is found in all retinal proteins, albeit a different residue for every group: tryptophan for BRs, tyrosine for HRs, and phenylalanine for SR I. It is, therefore, most probable that, as in *H. halobium* BR, the retinal is sandwiched between aromatic residues in all halobacterial retinal proteins and that its conformation is thereby stabilized.

Aspartic acid residues can be used to demonstrate the difference between general functions of conserved amino acids in contrast to specific functions of nonconserved essential amino acids. Two aspartic acids, D156 (Hhbop D115) and D255 (Hhbop D212), are conserved in all retinal proteins. Replacement of these residues in the BR of *H. halobium* showed that they are not directly involved in proton translocation. Nevertheless, they do have important functions for the reaction of the chromophore; e.g., when D255 (Hhbop D212) was exchanged, no M-like intermediate could be detected, and mutations of D156 (Hhbop 115) show a large color shift and a slowed relaxation from the O intermediate to the initial state (47).

Aspartic acids, D126 (Hhbop D85) and D137 (Hhbop D96), have been shown to be directly involved in proton transfer; i.e., they undergo deprotonation and reprotonation during the catalytic cycle (7, 19, 25, 37, 52). These residues are not conserved in all retinal proteins. D137 (Hhbop D96), which is involved in the reprotonation of the Schiff's base from the cytoplasmic side, is found only in BRs. D126 (Hhbop D85), which is involved in the deprotonation of the Schiff's base to the external side, is not found in HRs, and, correspondingly, no deprotonated intermediate is found in the photocycle of HR of *H. halobium* (22, 34). This residue is found in all BRs and SR I. Although deprotonation of the Schiff's base has not been demonstrated in SR I, a blue-shifted intermediate absorbing at 380 nm could be taken as an indication that it is involved in SR I function. The absence of an internal proton donor for the reprotonation of the Schiff's base could explain the slow photocycle of SR I; this has also been observed in BR if only D126 (Hhbop D85) is present and D137 (Hhbop D96) is missing (30, 44, 45).

Similar to studies on the conserved aspartic acid residues, studies on mutants of conserved tyrosines had shown previously that these residues in *H. halobium* BR are not de- or reprotonated themselves. Instead, they play important roles in reactions of the chromophore, e.g., light or dark adaptation (not discussed here).

Three prolines are conserved among the retinal proteins. They are located in helices B, C, and F (cf. Fig. 1 and 3). It
had been noticed that prolines are excluded from intramembranous parts of nontransport membrane proteins, but rather are found inside the membrane in proteins with transport function (6). This seems to indicate that they are involved in the mechanism of transport. Mechanistically, it had been proposed that cis-trans isomerization is involved (6), that the partial negative charge of their carbonyl bond might act as an intramembranous ligand-binding site, that pK changes due to conformational changes are important (10), or that their role is to induce a kink in a helix which is necessary for substrate (here, retinal) binding. Structural changes of prolines could indeed be detected with BR of \textit{H. halobium} by Fourier transform infrared spectroscopy (18, 41). It is possible that these structural changes of prolines are related to movements of part of the protein during the photocycle which could be detected by time-resolved X-ray diffraction (27).

It seems to be trivial to note that the lysine to which the retinal is covalently bound is conserved in all sequences (K259 [Hhbop K2161]). The conservation of arginine R123 (Hhbop R82) in all retinal proteins, however, was unexpected. In the HRs, only two positively charged residues are conserved, namely, R123 (Hhbop 108) and R/K217 (Hhbop R200). They were both thought to be involved in chloride binding (36) since two different chloride-binding sites have been described for HR of \textit{H. halobium} (for a review, see reference 28). The conservation of a residue involved in an HR-specific function seems to contradict a general role, as would be proposed for a conserved amino acid. At the moment one can only speculate that in the course of evolution the arginine first had a general function, e.g., participation in an intramembranous ion pair, and only later started to serve as a chloride-binding site. Thus, it would fulfill a dual function in HRs. In this context it is interesting to note that a chloride-binding site has been demonstrated also for BR (17). In addition, at low pH, chloride-dependent photocurrents have been measured with BR of \textit{H. halobium}, indicating that the proton pump BR is also able to pump chloride (8, 9).

At three positions, amino acids with hydroxyl groups are conserved: S/T130 (Hhbop T89), T131 (Hhbop T90), and S183 (Hhbop S141). The replacement of all threonines and serines in BR of \textit{H. halobium} yielded active proton pumps, but mutations at the three conserved positions affected the reactions of the chromophore, e.g., light or dark adaptation (31). The exact roles of these residues as well as that of the conserved M159 (Hhbop M118) remain to be investigated.

**Conclusions.** We think that the protein alignment presented here demonstrates that (i) the halobacterial retinal proteins form a distinct protein family not related to other known proteins; (ii) the principles of structure, chromophore isomerizations, and protein movements should be the same for all members of this family; and (iii) the conserved amino acids do not take part in reactions specific for one of the protein groups but have general functions. This is of great heuristic value, because if conserved amino acids are involved, results gained with BR of \textit{H. halobium}, still the main object of investigation, can serve as a first model for HR and SR I also. In principle, it should be possible to use the sequence alignment to predict amino acids which are important for a function specific to one of the protein groups or for the sensors in contrast to the ion pumps. Unfortunately, the sequences discussed in this report are too similar for this purpose; thus, predictions of this kind must wait until more sequences with less similarity are known.

**REFERENCES**