Diversification and spectral tuning in marine proteorhodopsins

Dikla Man, Weiwu Wang¹, Gazalah Sabehi, 
L Aravind², Anton F. Post³, 
Ramon Massana⁴, Elena N. Spudich¹, 
John L. Spudich¹,⁵ and Oded Beja³

Department of Biology, Technion-Israel Institute of Technology, Haifa 32000, ²H. Steinitz Marine Biology Laboratory, Interuniversity Institute for Marine Sciences, Eilat 88103, Department of Microbial and Molecular Ecology, Hebrew University of Jerusalem, Jerusalem, Israel, ¹Center for Membrane Biology and Department of Biochemistry and Molecular Biology, The University of Texas Medical School, Houston, TX 77030, ³National Center for Biotechnology Information, National Library of Medicine, National Institutes of Health, Bethesda, MD 20894, USA and ⁴Departament de Biologia Marina i Oceanografia, Institut de Ciencies del Mar, CSIC, E-08003 Barcelona, Spain

Corresponding authors 
e-mail: John.L.Spudich@uth.tmc.edu or beja@tx.technion.ac.il

D. Man and W. Wang contributed equally to this work

Proteorhodopsins, ubiquitous retinylidene photoactive proton pumps, were recently discovered in the cosmopolitan uncultured SAR86 bacterial group in oceanic marine environments. Two related proteorhodopsin families were found that absorb light with different absorption maxima, 525 nm (green) and 490 nm (blue), and their distribution was shown to be stratified with depth. Using structural modeling comparisons and mutagenesis, we report here on a single amino acid residue at position 105 that functions as a spectral tuning switch and accounts for most of the spectral differences between the two pigment families. Furthermore, looking at natural environments, we found novel proteorhodopsin gene clusters spanning the range of 540–505 nm and containing changes in the same identified key switch residue leading to changes in their absorption maxima. The results suggest a simultaneous diversification of green proteorhodopsin and the new key switch variant pigments. Our observations demonstrate that this single-residue switch mechanism is the major determinant of proteorhodopsin wavelength regulation in natural marine environments.

Keywords: diversity/rhodopsin/SAR86/spectral tuning/structure modeling

Introduction

Rhodopsins are currently known to belong to two distinct retinylidene protein families: the visual rhodopsins, found in eyes throughout the animal kingdom; and microbial rhodopsins, found in extreme haloarchaea, bacteria and eukaryotic microorganisms (Spudich et al., 2000). The two families share identical topologies, characterized by seven transmembrane helices, and are covalently linked to retinal. However, they do not share any significant similarities in their sequence and possibly have a different evolutionary origin (Spudich et al., 2000).

One of the most notable distinguishing properties of retinal among the various chromophores used in photosensory receptors is the large variation in its absorption spectrum depending on interaction with the apoprotein (‘spectral tuning’: Ottolenghi and Sheves, 1989; Birge, 1990). In rhodopsins, retinal is covalently attached to the ε-amino group of a lysine residue, forming a protonated retinylidene Schiff base. In methanol, a protonated retinylidene Schiff base exhibits a λmax of 440 nm. The protein microenvironment shifts λmax (the ‘opsin shift’; Yan et al., 1995) to longer wavelengths, e.g. to 487 nm in sensory rhodopsin II and to 568 nm in bacteriorhodopsin from Halobacterium salinarum. Spectral tuning in retinylidene proteins has been studied for decades (see Kochendoerfer et al., 1999; Lin and Sakmar, 1999; Nathans, 1999; Spudich et al., 2000, and references cited therein) but is still only partially understood.

The proteorhodopsins (PRs) are light-driven proton pumps (Béja et al., 2000a; Dioumaev et al., 2002; Friedrich et al., 2002) belonging to the microbial rhodopsin superfamily, type 1 rhodopsins (Spudich et al., 2000), and associated (Béja et al., 2000a) with the ubiquitous uncultured marine bacteria SAR86 group (Mullins et al., 1995; Béja et al., 2000b; Eilers et al., 2000; Gonzalez et al., 2000; Rappé et al., 2000; Kelly and Chistoserdov, 2001; Suzuki et al., 2001; Zubkov et al., 2001; Pernthaler et al., 2002). Two related PR subgroups were found that absorb light with different absorption maxima, λmax 525 nm (green) and λmax 490 nm (blue). Their distribution was shown to be stratified with depth, with ‘green-absorbing’ pigments at the surface and ‘blue-absorbing’ pigments at deeper waters in accordance with light available at these depths (Béja et al., 2001). The PR family is an ideal model system for investigating spectral tuning in microbial rhodopsins since both ‘green-absorbing’ and ‘blue-absorbing’ PR types share >78% of their amino acid residues (200 out of 247 are identical), while their visible absorption peaks differ by almost 40 nm. In contrast, the four types of H. salinarum rhodopsins show a comparable variation in color, but share <25–50% of their residues (Spudich et al., 2000).

As part of our effort to understand the global ecological implications of PR light harvesting and energy transduction to the ocean, we present here results on the identification and variation of a single amino acid which functions as a spectral tuning switch in these widespread pigments in natural environments.

Results and discussion

By comparing theoretical structural models, we identified candidates for the spectral difference between blue-
absorbing and green-absorbing PRs (B-PRs and G-PRs, respectively). The models (Figure 1) were constructed by threading PR sequences on the 1.55 Å resolution structure of bacteriorhodopsin (Luecke et al., 1999; for details, see Béja et al., 2000a). The eBAC31A08 (Béja et al., 2000a; G-PR) and PalE6 (Béja et al., 2001; B-PR) structural models were compared in order to find residue candidates accountable for the spectral difference between the two.

As can be seen in Figure 1, almost all the residue changes between G-PRs and B-PRs are facing outward, in the opposite direction to the retinal chromophore, leaving only four polymorphic variable residues within the retinal binding pocket (residues in positions 40, 68, 105 and 106). Residues 40 and 106 are >6 Å from the Schiff base, whereas residues 68 and 105 are predicted to be positioned ~5 Å from it; therefore positions 68 and 105 were chosen for mutation analysis.

In general, the spectra of rhodopsins exhibit two forms in a proton-dependent equilibrium: an alkaline species in which the primary retinylidene Schiff base counterion (Asp97 in PRs) is unprotonated, and a red-shifted acidic species in which this residue is protonated. The PR alkaline species are active in light-driven proton ejection from the cell. To investigate the potential importance of residues 68 and 105 in PR wavelength absorption regulation, we mutated the PR pigments at these positions and compared the spectra of their acidic and alkaline forms. The residue at position 68 differs in B-PR and G-PR clades, being isoleucine in the former and valine in the latter. These residues where changed to valine in B-PR and isoleucine in G-PR. We changed these residues to valine in B-PR and isoleucine in G-PR. The results showed that the residue at position 68 has a negligible effect on the absorption spectra of both the acidic and alkaline forms (Figure 2).

The difference between the two families at position 105 is a non-polar leucine residue (G-PR) versus a polar glutamine residue (B-PR). Therefore two PR protein mutants were constructed, in which the two residues at position 105 were interchanged (G-PR L105Q and B-PR Q105L), and their absorption spectra were compared. To assess the effect of the mutations, the mutant pigments were titrated to determine their alkaline and acidic spectra (W.Wang and J.L.Spadich, in preparation). The mutations at position 105 almost completely interconverted the absorption spectra of B-PR and G-PR (Figure 3). In the alkaline forms (pH 9.0), G-PR L105Q shifted to the blue and acquired vibrational fine structure like wild-type B-PR, and B-PR Q105L shifted to the red and lost the fine structure, exhibiting spectra similar to those of G-PR. In the acid forms (pH 5.0) the interconversion was essentially complete. The mutations at this site also shift the spectral transition pH of G-PR and B-PR towards that of the other pigment (W.Wang, E.N.Spadich, O.A.Sineshchekov, D.Man, O.Béja and J.L.Spadich, in preparation). Therefore the Q–L difference between G-PR and B-PR pigments functions as a single-residue spectral tuning switch. The residue at position 105 is close to the protonated Schiff base nitrogen. A mutation at the equivalent position in bacteriorhodopsin (L93) was also shown to alter its absorption spectrum (Subramaniam et al., 1991). It is also notable that the corresponding microenvironment of the protonated retinylidene Schiff base has been found to be a site of wavelength regulation in human rhodopsin (Kochendoerfer et al., 1999).

The physical basis of the spectral tuning by variation at position 105 may derive from altered hydrogen bonding and subtle repositioning within the quaternary charge complex that appears largely to determine the energies of the ground and excited state in microbial rhodopsins.
Fig. 2. Acidic and alkaline absorption spectra of retinal-reconstituted position 68 mutant PRs in *E. coli* membranes. Spectra were generated by all-trans retinal addition to *E. coli* membranes containing apoproteins of wild-type and mutant PRs as follows: G-PR, wild-type PR from Monterey Bay (eBAC31A08 in Figure 4; Béja et al., 2000a); B-PR, wild-type deep-water PR from the HOT station (palE6 in Figure 4; Béja et al., 2001); BPR_V681 and GPR_V681, V68I and I68V mutants of G-PR and B-PR, respectively. Wild-type PR absorption spectra (in red) and mutant PR spectra (in blue) have been normalized to unity at their absorption maxima for comparison of the spectra.

(Kochendoerfer et al., 1999; Ren et al., 2001). The predicted location of the 105 residue sidechain near the protonated Schiff base supports a possible mechanism in which the polar glutamyl sidechain in clade II pigments (defined in Figure 4A) stabilizes the protonated Schiff base, thereby decreasing π-electron delocalization in the retinal and producing a higher energy gap between the ground and excited states, i.e. a blue shift. A similar mechanism has been suggested for the blue shift attributable to the introduction of Ser308 in the mouse green cone pigment (Sun et al., 1997). Moreover, a combined mutagenesis and resonance Raman spectroscopy analysis of human rod rhodopsin shows that the dielectric character and architecture of the chromophore-binding pocket are specifically altered by mutations in residues that compose the retinal-binding pocket. An increase in the number of dipolar sidechains near the protonated Schiff base of retinal was found to increase the ground–excited state energy gap via long-range dipole–dipole Coulomb interaction (Lin et al., 1998). Such interactions may explain the blue shift in PRs on substitution of the non-polar Leu with the polar Gln residue. A more detailed explanation of the Q-L switch mechanism in terms of quantum–mechanical/molecular–mechanical factors awaits atomic resolution coordinates from crystal structures of representatives from different clades.

To study the possible ecological *in vivo* significance of the *in vitro* analysis of residue 105 spectral tuning properties, we performed a search for more PR variants directly in their marine environment. Therefore we applied PR polymerase chain reaction (PCR) primers (Béja et al., 2000a, 2001) to new environments.

The new PR sequences isolated from the Mediterranean Sea (Alborán Sea; MED clones in Figure 4A) and the Red Sea (RED clones) increased the previously known PR diversity, which was based on sequences isolated from Monterey Bay, California, the Hawaii Ocean Time (HOT) station and the Palmer station in Antarctica (Béja et al., 2000a, 2001). The new sequences from the Mediterranean formed a new PR cluster (clade III in Figure 4A), based on DNA comparisons, and shared at least 88% DNA sequence identity with the Monterey Bay and shallow HOT cluster (clade I), and 69% with the Antarctic and deep HOT cluster (clade II). Nearly all DNA differences between clades I and III are synonymous substitutions; there is little difference between their predicted protein sequences. This finding might suggest that PR clade III is also encoded by SARB-like bacteria, as a similar observation was recently reported for the GSAT
(glutamate semialdehyde aminotransferase) gene from natural populations of marine *Archaea* sharing almost identical 16S rRNA (Béjà *et al.*, 2002).

New sequences from the Red Sea station fell within clade I or were associated with the new clade III (Figure 4A). Some PRs from the Red Sea station that

---

**Fig. 4.** Analysis of cloned PR genes from the Red Sea and the Mediterranean Sea. (A) DNA phylogenetic tree based on distance analysis of 740 nt positions by neighbor-joining using the Clustal X (1.81) program (Thompson *et al.*, 1997). Bootstrap values >50% are indicated above the branches. Names in bold type indicate PR variants identified in this study. The scale bar represents number of substitutions per site. All previously reported PR variants have been incorporated in the tree: eBAC31A08–AF279106 (Béjà *et al.*, 2000a) and AF349976–AF350003 (Béjà *et al.*, 2001). (B) Multiple protein alignment of PRs from the different DNA-based groups shown in (A). Residue differences between the protein variants are marked in white on red; position 105 is marked above the alignment and is indicated in all sequences for clarity. Predicted transmembrane helices A, B and C are marked in gray.
fall in clade I had isoleucine instead of valine at position 68, a change found in members from clade II (Figure 4B). As noted above, this modification in the pigments of clades I and II has no influence on the spectra, and therefore it remains to be determined whether it contributes to different properties in the Red Sea PR family. Other Red Sea variants, also integrated in clade I, had a residue change at position 105 (plus more residue changes compared with members of clade I) (Figure 4B), the same position implicated (this study) as the spectral tuning switch. Therefore, we expressed environmental Red Sea 'position 105' variants in Escherichia coli cells and examined their absorption spectra. Confirming that position 105 is a key residue for wavelength regulation, the Red Sea variants absorbed at different wavelengths ranging from 505 nm (REDr6a5a2, Glu at position 105) to 540 nm (REDr6a5a6, Val at position 105; Figure 5). These observations in natural samples strengthen our conclusion that 105 is the key residue position responsible for spectral tuning in PRs.

Low bootstrap confidence in the branching order among G-PRs and ‘position 105’ PRs (upper part of Figure 4B) suggests that G-PR and ‘position 105’ PR pigments are undergoing a rapid contemporaneous diversification of both PR mini-lineages similar to what has been suggested for marine cyanobacteria photosynthetic antenna pigments (Urbach et al., 1998; Ting et al., 2002). These natural PR variants may derive from different SAR86 strains or other planktonic microorganisms that adapt to different light conditions, or alternatively to the same bacterium armed with different adapted pigments that enable it to survive and harvest light energy in different niches. Since the identity of the bacteria harboring the PRs from clades II and III is unknown and it is not clear whether all SAR86 use this mechanism to harvest light (Fuhrman, 2002), further experiments are needed to distinguish these possibilities.

The ‘position 105’ Red Sea variants reported here were collected at a depth of 150 m, which in the Red Sea is still in the photic zone. A well-tuned pigment is expected to be more important at depths at which light is at a premium and hence to be more adapted to these conditions. Indeed, preliminary comparison with shallower samples from the same station indicates a greater abundance of ‘position 105’ variants in deeper samples (G.Sabehi, R.Massana, E.F.DeLong and O.Béja, in preparation). It has recently been shown that multiplication of antenna genes is a major adaptation strategy to low-light conditions in a low-light-adapted cyanobacterium Prochlorococcus strain (Moore et al., 1998; Garczarek et al., 2000). Therefore the finding of the different ‘position 105’ PR variants at the bottom of the photic zone could be interpreted as the adaptation of SAR86 to minimal low-light conditions.

The variation in absorption maxima in highly similar (two to three residue changes) PR pigments resembles the variation in pigments in vertebrate color vision in which different isoforms of the same chromoprotein are adapted to different wavelengths (Nathans, 1999). The PR variants with blue to yellow–green absorption properties, which firmly belong to the ‘green’ clade of the PR tree (Figure 5B), confirm that the single spectral tuning switch residue identified in this study is broadly used in PRs in natural environments. This case illustrates the value of applying a combination of environmental genomics, theoretical structure prediction and site-directed mutagenesis to reveal the mode of action of natural selection in the environment.

Materials and methods

Bacterial strains
The E.coli outer-membrane-deficient strain UT5600 [ompT'] (Elish et al., 1988) was obtained from the E.coli Genetic Stock Center at Yale University (strain 7092).

Environmental DNA collection
Red Sea samples were collected from a depth of 150 m at a station in the northern Red Sea near the entrance to the Gulf of Suez (27.17°N, 34.22°E) in February 1999. Mediterranean samples were from the Alborán Sea (36.0°N, 4.25°W) collected in May 1998.
Site-directed mutagenesis
Mutants G-PR, V68I, B-PR, R68V, G-PR, L105Q and B-PR, Q105L were constructed by introducing mutations to G-PR (eBAC31A08 variant) or B-PR (pGalE variant) by using oligonucleotide-directed [V68I66F-5’-cgtatactgctatcatgattagttc-3’, V68I66F-R-5’-aaagatcagctagttgc-3’, 168NV6-5’-cgtatactgctatcatgattagttgc-3’, I68NV6-5’-gctatactgctatcatgattagttgc-3’, Q105L sens2S-5’-gggctatcatgattagttgcc-3’, Q105L sens2-5’-gctatactgctatcatgattagttgc-3’, Q105L sens2-5’-gctatactgctatcatgattagttgc-3’, Q105L sens2-5’-gctatactgctatcatgattagttgc-3’, I68NV6-5’-gctatactgctatcatgattagttgc-3’, B-PR, P69S and P69L variants by using the in vitro transposition of the Ncol site in prev. mut. digests. Point mutation identities were verified by sequencing the length of the PCR-generated inserts in the final constructed expression vector.

Environmentsnal PCR PR amplification and expression
PRs were amplified by PCR directly from environmental samples using a modified PCR primer forward primer R68S-5’-CCATGGTTGAAAT-ATTACGTATGTTAC-3’ (Ncol site is underlined), designed based on sequence retrieved from an Antarctic fosmid containing blue PR (J.R.de la Torre and E.F.DeLong, personal communication) and the original reverse PCR primer (sBejia et al., 2006), using independently two different high-fidelity proof-reading polymerase mixes (BIO-X-ACT™ from Bioline and Takara Ex Taq™ from Takara Shuzo Co.; G.Saheli, R.Massana, E.F.DeLong and O.sBejia, in preparation). PCR products were cloned using the QIAGEN® PCR cloning kit and unique RFPL groups were sequenced. Sequences were viewed using the Sequencer 4.1.2 program (Gene Codes Corporation) and the sequences submitted to the EMBL database were expressed using the pBAD TOPO® expression kit (Invitrogen). Correctly oriented inserts were verified by Ncol digestion and religation to remove the signal sequence integrated in the commercial vector and to maintain the native signal sequence of the PR. To express PRs, O.N. cultures were diluted 1:100 and grown for 2 h at 37°C. Cells were induced with 0.2% i-arabinose for 3 h. All PCR products identical with samples handled or amplified previously in the laboratory were omitted from the analyses to avoid influence from possible contamination. The sequences reported in this study are deposited under DDBJ/EMBL/GenBank accession Nos AY210898–AY210919.

Absorption spectroscopy
Differences in the absorption spectra were measured using the retinal-generated absorption difference in E.coli membranes. Ethanolysis solutions of all-trans retinal were added to membrane suspensions and the course of pigment development monitored until completion (~60 min) in an SLM Amico DW2000 spectrophotometer as described previously (Bieszke et al., 1999). To establish the absorption spectra of retinal and acidic forms of the pigments, the retinal addition was performed with membrane suspensions adjusted to a range of pH values between 4.0 and 9.0, and the regenerated absorption spectra were recorded. The transition between acid and alkaline forms in different proprhopodopsins pigments was found to exhibit pH values in the range 6.5–8.0 (W.Wang, E.N.Spadich, O.A.Sineschekov, D.Man, O.Béja and J.L.Spadich, in preparation).

Acknowledgements
We thank E.DeLong for encouragement and for sharing unpublished data, M.Sheves for help during initial stages of the research and G.Rhodes and N.Kar for help with molecular graphics programs. We thank B.Horvitz and B.Podbielwicz for helpful discussions and critical reading of this manuscript. This research was supported by the Human Frontiers Science Program P382002 (J.L.S. and O.B.), J.S.Frankord, E. and J.Bishop, J. and A.Taub and the New York Metropolitan research funds (O.B.), National Science Foundation grant 0091287 (J.L.S.), National Institutes of Health contrasting fund (J.L.S), the Robert A.Welch Foundation (J.L.S.), EU project PICOVIT EVK3-CT1999-00021 (R.M.) and the Israel Science Foundation 525/98 (A.P.).

References
Spectral tuning in marine proteorhodopsins


Received November 26, 2002; revised February 6, 2003; accepted February 25, 2003