An Empirical Test of Convergent Evolution in Rhodopsins

Kristine A. Mackin,¹ Richard A. Roy,¹ and Douglas L. Theobald*¹

¹Department of Biochemistry, Brandeis University
*Corresponding author: E-mail: dtheobald@brandeis.edu.
Associate editor: Jeffrey Thorne

Abstract

Rhodopsins are photochemically reactive membrane proteins that covalently bind retinal chromophores. Type I rhodopsins are found in both prokaryotes and eukaryotic microbes, whereas type II rhodopsins function as phototransduced G-protein coupled receptors (GPCRs) in animal vision. Both rhodopsin families share the seven transmembrane α-helix GPCR fold and a Schiff base linkage from a conserved lysine to retinal in helix G. Nevertheless, rhodopsins are widely cited as a striking example of evolutionary convergence, largely because the two families lack detectable sequence similarity and differ in many structural and mechanistic details. Convergence entails that the shared rhodopsin fold is so especially suited to photosensitive function that proteins from separate origins were selected for this architecture twice. Here we show, however, that the rhodopsin fold is not required for photosensitive activity. We engineered functional bacteriorhodopsin variants with novel folds, including radical noncircular permutations of the α-helices, circular permutations of an eight-helix construct, and retinal linkages relocated to other helices. These results contradict a key prediction of convergence and thereby provide an experimental attack on one of the most intractable problems in molecular evolution: how to establish structural homology for proteins devoid of discernible sequence similarity.

Key words: homology, convergence, bacteriorhodopsin, fold, permutation, lysine.

Introduction

Type I and type II rhodopsin families share several salient structural and functional features (Lanyi 1999; Luecke 2000; Spudich et al. 2000; Smith 2010). All rhodopsins of known structure adopt a seven transmembrane (7TM) α-helical fold, the G-protein coupled receptor (GPCR) fold, which is defined by a particular spatial arrangement and connectivity of the α-helices (see fig. 1). This seven-helix architecture forms an internal pocket for the retinal moiety, which is attached via a Schiff base linkage to the ε-amino group of a conserved lysine residue in the middle of the seventh α-helix (helix G). In both rhodopsin families, the retinal chromophore undergoes a light-induced cis–trans isomerization across a double bond in the polyene chain. In all rhodopsins, including both the sensory and ion-transporting proteins, the photoinduced retinal isomerization is associated with a mechanistically critical deprotonation of the positively charged Schiff base and subsequent transfer of the proton to another residue in the retinal-binding pocket.

Despite the striking similarities between the type I and type II rhodopsins, the two families are distinguishable by sequence, structure, and mechanism (Spudich et al. 2000). Sequence analysis has consistently failed to detect significant similarities between these rhodopsin families. Furthermore, type I rhodopsins purportedly contain a weak internal sequence repeat between helices A–C and E–G, suggestive of an ancient gene duplication event; type II rhodopsins apparently lack such an internal repeat (Taylor and Agarwal 1993; Larusso et al. 2008). The two families are also taxonomically distinct. Type I rhodopsins are found throughout prokaryotes (including Actinobacteria, Bacteroidetes, Chloroflexi, Cyanobacteria, Deinococcus-Thermus, Firmicutes, Planctomycetes, Proteobacteria, and archaeal Halobacteria) and in a few single-celled eukaryotes (Alveolata, Fungi, and various algae including Chlorophyta, Cryptophyta, Glaucophyta, Haptophyceae, and Streptophyta) (Sharma et al. 2009). Until recently, type II rhodopsins were thought to be exclusively eumetazoan, though now they have also been found in fungal genomes (Heintzen 2012). The spatial arrangement of the seven transmembrane α-helices differs between the families (fig. 1), particularly in the packing of helix C against helix E and in the angle of the helices relative to the plane of the membrane. In type II rhodopsins, helices B and E–G are distorted by severe mid-helix kinks, whereas type I helices are relatively linear. In type I rhodopsins, the retinal is bound in a pocket solely composed of helical residues; in type II rhodopsins, the extracellular side of the retinal binding pocket is formed by a small β-hairpin connecting helices D and E. There are also mechanistic differences between the two families. Most notably, in type I rhodopsins, the retinal photoisomerizes from all-trans to 13-cis, whereas in the type II family the retinal converts from 11-cis to all-trans.

Did these two protein families diverge from an ancient common ancestral protein or have they converged on the same protein fold from independent origins? This question has been a subject of controversy for over 40 years (Oesterhelt and Stoeckenius 1971; Hargrave et al. 1983; Rao et al. 1983; Findlay and Pappin 1986; Dohlman et al. 1987; Oesterhelt and Tittor 1989; Henderson and Schertler 1990; Hibert et al. 1991; Pardó et al. 1992; Taylor and Agarwal 1993; Soppa 1994;
In both type I and type II rhodopsins, the seven transmembrane helices adopt a particular spatial arrangement that is necessary for chromophore binding and spectral tuning (Yan et al. 1995; Kochendoerfer et al. 1999; Yokoyama 2008). Free in solution, the retinal chromophore has an absorbance maximum (λ\text{max}) of 380 nm. The transmembrane helices form a specific pocket for binding the retinal cofactor, typically resulting in a large redshift of the retinal λ\text{max}. In Halobacterium salinarum type I bacteriorhodopsin (Hs br), for example, the absorption maximum is shifted nearly 200 nm to λ\text{max} = 568 nm. Similarly, the type II bovine rhodopsin has a λ\text{max} = 550 nm. A key component of the retinal binding site is the negatively charged “counterion” residue, typically an aspartic or glutamic acid, that promotes protonation of the lysine–retinal Schiff base. When the Schiff base deprotonates, the protein-bound chromophore has a λ\text{max} of ~410 nm. On the other hand, if the protein is denatured while maintaining a protonated Schiff linkage (e.g., during an “acid-trap” experiment), the water-exposed retinal has a λ\text{max} of ~440 nm (Fasick et al. 1999). Even conservative mutations in the retinal binding pocket generally result in large changes in activity and spectral characteristics (usually a blueshift of the λ\text{max}) (Yokoyama 2008). Therefore, the rhodopsin λ\text{max} is a very sensitive gauge for detecting structural perturbations of the retinal binding site and helical packing.

Rhodopsin’s seven transmembrane helices are connected by loops in a specific order that is characteristic of the GPCR fold (Murzin et al. 1995). These connecting loops are relatively far removed from the retinal binding pocket, and various lines of biochemical and phylogenetic evidence indicate that the loops are largely dispensable for function. Because type I and type II rhodopsins share the GPCR fold, they also share the same loop connectivity among the helices. However, there are 144 possible different connectivities (and corresponding protein folds) for these seven helices that nevertheless could maintain their observed spatial arrangement and preserve the retinal pocket. Evolutionary processes are unlikely to converge on the same connectivity (or fold) by sheer chance, due to the large number of possible protein folds (Grishin 2001; Sadowski and Taylor 2010). Why then is only one of these possible connectivities seen in nature?

According to the convergent hypothesis, both rhodopsin families share the specific GPCR connectivity and retinal linkage because these particular structural features are essential for function. That is, the photosensitive, retinal-dependent function is highly physically constrained and requires the rhodopsin fold. The tight coupling between rhodopsin structure and function has allowed selection for photosensitive function to lead proteins from different origins to the rhodopsin fold independently (Spudich et al. 2000; Larusso et al. 2008). To test the convergent hypothesis, we therefore experimentally assessed the question: Is the observed rhodopsin fold in fact required for rhodopsin activity? Remarkably, the answer is no—neither the observed GPCR connectivity nor the conserved retinal linkage in helix G is necessary for bacteriorhodopsin photosensitive function.
Results

Permutations of the Helices Are Functional

Our experiments are based on derivatives of a His-tagged bacteriorhodopsin (bR) from *Haloterrigena turkmenica* (light-adapted $\lambda_{\text{max}} = 548$ nm), which can be recombinantly overexpressed with high yield as a functional proton pump in the *Escherichia coli* membrane (Kamo et al. 2006; Hara et al. 2011). Using the *Hal. turkmenica* bR (Ht bR) as a template, we designed two classes of artificial bacteriorhodopsins with novel protein folds: 1) noncircular permutations of the native seven $\alpha$-helices and 2) circular permutations of an eight-helix construct.

Circularly permuted proteins occur naturally, though rarely, and permutations have been used to manipulate the tertiary architectures of soluble proteins (Heinemann and Hahn 1995; Lindqvist and Schneider 1997; Grishin 2001; Yu and Lutz 2011; Bliven and Prlic 2012). However, circular permutations of the 7TMGPCR fold cannot be constructed directly, because an odd number of helices dictates that the N- and C-termini of the fold reside on opposite sides of the membrane. It is therefore only possible to construct noncircular permutations of the seven-helix architecture (see fig. 1). To create circularly permuted mutants, we inserted an artificial eighth helix between the C- and N-termini of the wild-type (WT) bR based on a WALP21 transmembrane peptide (Holt and Killian 2010). Circular permutations can then be generated by choosing new termini in the loops between helices, and any of the helices can be placed at the N-terminus (fig. 2). In total, 11 different permuted constructs were evaluated; seven of these overexpress successfully in the *E. coli* membrane. The remaining constructs either do not express or express as inclusion bodies, and they were not pursued further.

Overexpression of the WT Ht bR in the presence of retinal imparts a notable bright pinkish purple hue to the bacterial cell pellets, due to the functional reconstitution of the retinalidene protein in the bacterial cell membrane (supplementary fig. S1, Supplementary Material online) (Kamo et al. 2006). Overexpression of several of our mutant proteins also produces cell pellets with a similar pink coloring, suggesting that the mutant proteins insert into the membrane in a properly folded, retinal-conjugated form. Nascent Hs bR binds to the signal recognition particle (SRP) and is targeted to the translocon for membrane insertion (Dale and Krebs 1999; Dale et al. 2000; Curnow et al. 2011). In vitro studies have shown that recombinantly expressed Hs bR also uses an SRP-dependent mechanism for insertion in the *E. coli* membrane (Raine et al. 2003). Therefore, a pink pellet suggests that the helices do not significantly affect the ability of a given permutation construct to interact with the SRP, target to the membrane, or fold in a native-like conformation. Constructs in which helices A, B, C, D, F, and G are located at the N-terminus are all able to target the protein to the membrane. Thus, multiple “signal sequences” appear capable of initiating SRP-dependent membrane insertion in our recombinant system.

The absorption spectrum of the purified permuted constructs closely recapitulates that of the WT protein (fig. 3, $546 \text{ nm} < \lambda_{\text{max}} < 551$ nm), despite the fact that rhodopsin absorption spectra are exquisitely sensitive to changes in the local retinal environment in the protein interior. Acid-trap experiments indicate that the permuted constructs covalently bind the retinal via a Schiff base linkage (Oesterhelt and Stoeckenius 1971; Fasick et al. 1999) (supplementary fig. S3, Supplementary Material online). The native-like spectra imply that the binding pocket residues form native-like contacts with the retinal and further suggest that the helices pack with minimal structural perturbation.

We assessed proton-pumping activity of our bR constructs using a proteoliposome assay in which bR is reconstituted in unilamellar liposome vesicles (Oesterhelt and Stoeckenius 1971; Hackett et al. 1987). Upon illumination, WT Ht bR transports protons from the exterior to the interior of the liposome, consistent with the “inside—out” orientation of bR previously observed in reconstituted systems from *H. salinarum* (Huang et al. 1980). Trypsin digestion experiments with the Ht bR are also consistent with an inside—out protein orientation in our vesicle assays. All the purified permutation constructs are functionally competent in proteoliposome proton-pumping assays, though with varying levels of activity (fig. 3). Surprisingly, two constructs apparently show even higher activity than the WT Ht bR.

bR Folding Information Is Encoded Locally

The permutation mutants fold correctly in the membrane and actively pump protons, demonstrating that no particular primary helical order is essential for the folding and function of bR. Bacteriorhodopsin helix assembly and activity is therefore largely governed by helical packing interactions alone, being remarkably insensitive to both structural perturbations of the inter-helical loops and to helical connectivity, consistent with previous observations (Kahn and Engelman 1992; Kataoka et al. 1992; Marti 1998; Kim et al. 2001). Currently, the principles of membrane protein folding are poorly understood relative to soluble proteins. Because radically permuted helices can nevertheless fold competently into native-like arrangements, local sequence elements encode bR tertiary information independently of global connectivity, similar to many soluble proteins (Viguer et al. 1995).

According to the preferred model for bR folding, the N-terminal five helices (A–E) are independently stable elements within the membrane that insert first and associate via a two-stage mechanism (Popot and Engelman 1990; Booth 2000). Only after helices A–E have formed a stable transmembrane core can helices F and G then insert and pack to form the apo bacterioopsin. Finally, the retinal spontaneously enters the protein core and reacts to make the protonated Schiff base linkage. However, two of our constructs, GBCDEFA and CDEFGAB, are unable to fold by this mechanism. In these mutants, helices A and F are adjacent in sequence, and therefore, helix A cannot assemble into the stable folding core without the prior or simultaneous insertion of helix F.
Therefore, the hypothesized bR folding mechanism may be specific to the native helix connectivity or the mechanism may require revision. Our GBCDEFA and CDEFABG permutation mutants necessarily fold by a different mechanism, and hence the folding pathway does not appear to present a significant barrier to the evolution of alternative helix connectivities and folds.

Fig. 2. Bacteriorhodopsin transmembrane helix permutation constructs. For each permutation construct, a secondary structure schematic is shown above emphasizing differing connectivities. Below is the primary sequence structure, with transmembrane helices colored as in figure 1. Panels B–E represent the noncircular permutations. Panels F and G represent the circular permutations with the eighth additional "WALP21" helix shown in gray. (A) Wild-type bR, (B) GBCDEFA, (C) CDEFGBA, (D) GFABCD, (E) CDEFGAB, (F) FGWABC, (G) DEFGBWC, (H) BCDEFGW. See also supplementary table S2 and figure S1, Supplementary Material online.
Schiff Base Linkage Functions in Alternate Helices

The functionally critical lysine (K216, *H. salinarum* residue numbering) found in helix G, which forms a Schiff base with the retinal chromophore, is an extraordinarily conserved feature of all known bR homologs (excepting a small number of fungal homologs of unknown function). Although the lysine–retinal Schiff base itself may be essential for activity, its specific location in primary sequence may not be; the lysine could potentially reside in structural elements other than helix G and still form the Schiff base without altering the retinal conformation and function. If type I and type II rhodopsins are convergent, the shared lysine position in helix G is likely a functionally necessary feature of the rhodopsin fold. Alternatively, if rhodopsins are homologous, the strict conservation may be historical detritus inherited from an ancient common ancestral rhodopsin protein. To distinguish between these hypotheses, we constructed mutant bRs in which the lysine Schiff base linkage was relocated to helices B and C and assessed their folding and function.

**Discussion**

The rising number of membrane protein crystal structures includes many surprising examples of structural similarity between proteins initially thought to be unrelated due to a lack of sequence similarity (Theobald and Miller 2010). Like watersoluble proteins, a membrane protein’s tertiary fold and function apparently can be encoded by a large number of vastly different sequences—a remarkable biophysical feature of protein polymers (Gherardini et al. 2007; Omelchenko et al. 2010). Due to the implausibility of independently converging on similar sequences, either by chance or via selection, significant sequence similarity between two proteins provides strong support for homology (i.e., divergent evolution from a common molecular ancestor) (Theobald 2011). On the other hand, sequence dissimilarity is an unpersuasive evidence for convergence. Billions of years of accumulated amino acid substitutions can erase any residual sequence similarity between homologous proteins, even while preserving overall tertiary structure and function. Compared with protein sequences, the total number of protein folds is relatively small (<10,000). Evolutionary convergence to similar folds is therefore much more plausible than converging on similar sequences.

Given two transmembrane proteins with identical folds, yet no sequence similarity, how then could we distinguish convergence from homology? This fundamental question is one of the longest standing problems in molecular evolution (Zuckerkandl and Pauling 1965; Doolittle 1994; Murzin 1998; Grishin 2001; Cheng et al. 2008). Discriminating between
these hypotheses is notoriously difficult to address empirically, and to date, theoretical arguments have failed to satisfactorily resolve the controversy. Rhodopsins have been at the center of the homology–convergence controversy for nearly half a century (Oesterhelt and Stoeckenius 1971; Hargrave et al. 1983; Rao et al. 1983; Findlay and Pappin 1986; Dohlman et al. 1987; Oesterhelt and Tittor 1988; Henderson and Schertler 1990; Hibert et al. 1991; Pardo et al. 1992; Taylor and Agarwal 1993; Soppa 1994; Metzger et al. 1996; Spudich et al. 2000; Larusso et al. 2008). Fortunately, convergence hypotheses predict the existence of specific, functionally essential structural constraints, and in principle, these constraints can be experimentally investigated.

These results show that bacteriorhodopsin can withstand radical remodeling of its fold and structural relocation of the critical, highly conserved Schiff base linkage. In other work on the type II bovine rhodopsin, we have shown that the conserved active-site lysine can also be relocated to three other secondary structure elements (other than helix G) while maintaining WT-like function (Devine et al. 2013). These variant architectures have never been observed in nature; each of our seven Ht bR permutation constructs represents a novel transmembrane fold. The apparent lack of strong structural constraints thus challenges the widely held view that type I and II rhodopsins are convergent evolutionary inventions, having arrived at their signature structural features separately. Because the naturally occurring rhodopsin fold is unnecessary for functional competence, it is unlikely that selection would lead unrelated proteins to this particular architecture independently.

Other proteins display a lack of structural constraint similar to bR, though relevant studies have been less extensive than those reported here. To our knowledge, our constructs are the first engineered permutations of a helical membrane protein of known structure, and they represent the only noncircular permutations of any membrane protein. There are two other examples of functional permutations of helical membrane proteins, although both cases lack high resolution structural information (Gutknecht et al. 1998; Beutler et al. 2000). For type II rhodopsins, there are examples of split bovine rhodopsins that function in trans (Yu et al. 1995). We speculate, therefore, that certain helical permutations would likely result in functional type II rhodopsins, but several of the type II loops are important for protein–protein interactions and could not be perturbed so readily without detrimental functional consequences. As mentioned above, in bovine rhodopsin, the active site lysine can reside in four different structural elements with retention of function (Devine et al. 2013), indicating that type II rhodopsins may be even less constrained than type I rhodopsins with respect to the location of the Schiff base linkage. Furthermore, many soluble proteins can be circularly permuted with retention of function (Yu and Lutz 2011; Bliven and Prlic 2012). However, noncircular permutations have been successfully engineered in only one other protein, the water-soluble green fluorescent protein (Reeder et al. 2010). Taken together, the large and diverse number of known functional protein permutations indicates that it is in general unlikely for protein evolution to converge to the same fold as a result of structural and functional constraints.

Several of our mutant bR constructs exhibit proton translocation rates of only 10–50% of the WT Ht bR rate. One might argue that these are in fact significant decreases in activity indicating strong structural constraints on bR function. However, even our slowest active mutant (As3K/K216A, 11% of WT rate) should be considered fully functional for the following reasons.

Any measurable photoactivated proton transport is remarkable and biologically significant, considering that in over 3 billion years evolution has found only two different ways to harvest light and convert it to chemical energy (Bryant and Frigaard 2006). Phototrophy is an extraordinary biomolecular feat. Bacteriorhodopsin actively builds up a transmembrane electrochemical potential against a gradient—an extremely thermodynamically unfavorable process—by mechanistically coupling proton transport to the favorable entropy changes of a star 93 million miles distant (Brittin and Gamow 1961; Albarrán-Zavala and Angulo-Brown 2007). There is no “background” proton-pumping activity; a transmembrane electrochemical gradient does not spontaneously accumulate but dissipates. For bR to pump a proton, the chromophore must absorb a photon, enter into a high-energy state, and then relax back to the ground state with the dissipation of energy. During this process, the protein must somehow couple the chromophore’s photocycle to the translocation of a single proton through the center of the protein and across the lipid bilayer. The precise mechanism for how this efficient coupling is accomplished is still largely unknown (Hirai and Subramaniam 2009; Hirai et al. 2009), but there are many more pathways for the coupled reaction cycle to fail than for it to work (Hill 2005). The chromophore could absorb the wrong wavelength of light; the proper wavelength could be absorbed by destroying the chromophore; the chromophore could enter and leave the high-energy state by simply dissipating the energy as heat or by forming a nonproductive adduct with the protein; the proton could be picked up from the intracellular side and deposited back on the same side, rather than translocating; or the protons could flow back down the gradient through the very protein channel they just traversed. Therefore, a protein construct that exhibits any detectable light-activated proton-pumping activity is necessarily a complex and unlikely molecular device.

---

Table 1. Proton Translocation Activity of Lysine Swap Mutants.

<table>
<thead>
<tr>
<th>Mutant</th>
<th>$\lambda_{\text{max}}$</th>
<th>Lysine Linkage Helix</th>
<th>Relative Initial Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>WT</td>
<td>548</td>
<td>G</td>
<td>1.0</td>
</tr>
<tr>
<td>A53K/K216A</td>
<td>520</td>
<td>B</td>
<td>0.11</td>
</tr>
<tr>
<td>A53K/K216V</td>
<td>510</td>
<td>B</td>
<td>ND</td>
</tr>
<tr>
<td>A53K/K216I</td>
<td>510</td>
<td>B</td>
<td>ND</td>
</tr>
<tr>
<td>T89K/K216T</td>
<td>600</td>
<td>C</td>
<td>ND</td>
</tr>
</tbody>
</table>

Note.—Rates are given relative to WT $H.\ turb.\ bR$. ND indicates no detectable activity in our assays.
Most importantly, the measured proton-pumping rates of our mutant constructs are well within the observed range of naturally occurring type I rhodopsins. The proton-pumping rate of WT Ht bR is ~4-fold faster than the classic Hs bR, based on their photocycles (3 vs. 10 ms time constants, respectively) (Kamo et al. 2006). Because one proton is pumped per cycle, the proton-pumping rate should scale proportionally with the photocycle. We therefore estimate that our slowest mutant (the active site lysine swap construct A53K/K216A, 11% of WT pumping rate) has a ~30 ms photocycle. This value is in par with the pumping rates of the widespread aquatic and marine green proteorhodopsins, which contribute significantly to the energy needs of their microbes and are considered to be highly efficient proton pumps (Beja et al. 2000, 2001; Fuhrman et al. 2008, Gomez-Consarnau et al. 2010). Furthermore, deep water blue proteorhodopsins are about 10- to 50-fold slower than Hs bR (i.e., a time constant of 150–600 ms). An even slower photocycle (hundreds of milliseconds) is found in sensory type I rhodopsins, which function in phototaxis to avoid UV damage, in photoregulation of metabolism, and as possible daytime sensors or water depth gauges (Fuhrman et al. 2008).

Unlike the permutation constructs, the \( \lambda_{\text{max}} \) of the A53K/K216A lysine swap mutant is blue-shifted by about 30 nm, indicating a significant perturbation of the retinal binding site. What is the reason for this shift in the absorbance maximum? Although a small fraction of the A53K/K216A mutant may be unfolded, several lines of evidence indicate that the protein is mostly folded and that the observed blueshift is due to perturbations in the folded protein fraction. Free retinal has a maximum absorption of 380 nm. It takes a very specific electrostatic molecular environment to redshift the retinal absorbance to 550 nm (Vasileiou et al. 2007; Yokoyama 2008; Wang et al. 2012). When the WT protein is unfolded (for instance, by SDS denaturation), the retinal is exposed to water, hydrolyzes, and has a \( \lambda_{\text{max}} \) of ~380 nm (supplementary fig. S3, Supplementary Material online). It is possible to acid-trap the Schiff base linkage at low pH by protonating it during denaturation. This prevents hydrolysis and keeps the water-exposed retinal covalently bound to the denatured protein. When acid-trapped, the protonated retinal has a \( \lambda_{\text{max}} \) of ~440 nm (supplementary fig. S3, Supplementary Material online). The A53K/K216A lysine swap mutant behaves similarly to WT in denaturation and acid-trapping experiments. With the A53K/K216A construct, the 520 nm state unfolds to the 380 nm state at neutral pH, and at pH 2 it unfolds to the acid-trapped 440 nm. All three states have distinct, easily distinguishable spectra and characteristic \( \lambda_{\text{max}} \) values. The large redshift from 380 to 520 nm in the non-denatured A53K/K216A construct indicates that the retinal must reside in a specific, intact binding pocket. Taking all the data into account, we conclude that the A53K/K216A construct is a wellfolded protein with a blue-shifted absorbance maximum. In this construct, the Schiff base linkage has been relocated to an adjacent helix. Hence, the blueshift is likely due to minor electrostatic perturbations resulting from alteration of the conformation of the protonated Schiff base relative to the counterion.

All of our constructs have apparently altered pumping rates, relative to WT, suggesting that their respective photocycles are also perturbed. We currently have no experimental data directly addressing the photocycles of our mutant constructs, so we are reluctant to speculate extensively on the matter. However, all the permutation mutants have a \( \lambda_{\text{max}} \) identical to WT (within experimental error), and they all have apparent pumping rates greater than 30% of WT. Hence, we predict that the photocycle has not changed dramatically in these constructs; most likely, the photocycle time constant is also within 30% of WT. The A53K/K216A construct is perhaps more interesting, with its blueshifted \( \lambda_{\text{max}} \) and slower rate. We suspect that A53K/K216A may have larger changes in the photocycle than the rest, primarily because in this mutant the critical Schiff base linkage has been perturbed, which may affect the photocycle M intermediates. We are currently in the process of characterizing the mutant photocycles using laser flash photolysis. In any case, from an evolutionary perspective, it does not matter whether different rhodopsins have different photocycles, as long as each protein pumps protons across the membrane at a rate beneficial to the organism.

Our biochemical characterization of the restructured bRs does not directly address other selective pressures that likely exist on bR in nature, such as thermodynamic stability, folding kinetics, and protein degradation. We emphasize, however, that our redesigned bRs were engineered very crudely and no effort was made at optimization. It is highly likely that compensatory mutations could be found, which restore the activity of our mutant constructs to that of the WT. In contrast, natural rhodopsins have been crafted by selection for millions of years in individual species (a point that applies to both convergent and divergent hypotheses). Any potential compensatory mutations would be accessible to convergent evolution if the rhodopsins were not homologous. Selective factors vary among species and environments and likely can be accommodated by only a few mutations in each variant bR architecture. For example, our lysine swap mutant A53K/K216A has a significantly blueshifted \( \lambda_{\text{max}} \) (520 nm). This absorbance maximum is similar to the vast majority of bR homologs in the biosphere, which are aquatic proteorhodopsins having a \( \lambda_{\text{max}} \) in the blue/green region (490–530 nm) (Fuhrman et al. 2008). Although a \( \lambda_{\text{max}} \) of 520 nm may be suboptimal in certain environments and species, minor residue substitutions have large consequences for tuning the bR absorbance spectrum. It is relatively easy to modify, say, the chromophore \( \lambda_{\text{max}} \) or the stability of a protein; it is considerably more difficult to find a protein architecture that can transform the energy from a green photon into a transmembrane electrochemical gradient. For testing the convergent evolution of the rhodopsin fold, the pertinent concern is whether our unoptimized alternative architectures have any experimentally detectable photosensitive function. Our naive attempts at re-engineering the type I rhodopsin architecture easily found functional variants while apparently nature has not—a fact that highlights the implausibility of convergent evolution to the same rhodopsin fold.
We have radically perturbed extraordinarily conserved elements of bR with minimal functional consequence. When inspecting a sequence alignment of bRs, one of the most striking features is the conserved lysine in helix G. Extreme sequence conservation is generally interpreted as indicating functional necessity. However, the structure–function relationship is highly degenerate; many sequences/structures are capable of performing the same biochemical function. In terms of the protein fitness landscape metaphor, our results suggest that the natural bR is effectively confined to a local fitness maximum. There are apparently other local maxima with the Schiff base lysine in helix B, but crossing fitness valleys is problematic for gradual divergent evolution—though in principle this is not a barrier to convergent evolution. Moving from one of these maxima to another would involve two improbable and precise mutations at specific positions. In any particular species, the novel mutant protein would likely have a significant selective disadvantage relative to the native protein. The rhodopsin architecture thus appears to be phylogenetically constrained, rather than physically constrained (McGhee 2008).

One of the most cogent pieces of evidence for the convergent evolution of rhodopsins involves their distinct taxonomic distribution. Until recently, it was thought that type I rhodopsins were exclusively prokaryotic, whereas type II rhodopsins (and other class A GPCRs) evolved from opisthokonts. However, type II rhodopsins have now been found in microbial eukaryotes (the basal fungi Blastocladiomycota and Chytridiomycota), and phylogenetic analysis strongly indicates that rhodopsins evolved from a non-retinylidene cAMP receptor GPCR near the origin of opisthokonts (Krishnan et al. 2012). Furthermore, type I rhodopsins have likewise been identified in numerous eukaryotic microbes, including plants, animals, fungi, and alveolates—an observation that suggests the GPCR architecture evolved first and later gave rise to a retinylidene type II rhodopsin in the opisthokont ancestor, independently of prokaryotic rhodopsins. However, type II rhodopsins have now been found in microbial eukaryotes (the basal fungi Blastocladiomycota and Chytridiomycota), and phylogenetic analysis strongly indicates that rhodopsins evolved from a non-retinylidene cAMP receptor GPCR near the origin of opisthokonts (Krishnan et al. 2012).

Materials and Methods

Design of Permuted bR Constructs

Homology models, multiple sequence alignment, and structures of the canonical Hs bR were used to identify residues constituting each α-helix. Helices were shuffled within the primary sequence, maintaining native loops where appropriate. Constructs requiring loops with greater distances than WT were spanned with repeating SSG motifs (see supplementary information, Supplementary Material online, for protein sequences). In this article, we use H. salinarum residue numbering for the Hal. turkmenica residues. Hs bR residues A53, D85, T89, and K216 correspond to Ht bR residues A61, D93, T97, and K225, respectively.

Expression and Purification

Genes for bR permutations were synthesized by GenScript (Piscataway, NJ) in pET-21c vectors (except the WT and GBCDEFA, which were subcloned into pET-21b). BL21(DE3) pLys cells were transformed with these vectors, and cells were
grown to an OD$_{600}$ between 0.3 and 0.6 and induced by adding 1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) and 10 μM all-trans retinal. Cells were harvested after 1–3 h by centrifugation.

The cell pellet was resuspended in buffer (50 mM Tris-Cl, pH 8.0, 5 mM MgCl$_2$) after centrifugation. Benzonase (EMD Millipore) was added and the suspension was sonicated, and the membrane fraction was isolated by centrifugation. The membrane pellet was resuspended in buffer (300 mM NaCl, 50 mM MES, pH 6.5). Dodecyl maltoside (DDM) was added to 1.0% w/v. The sample was incubated with gentle shaking for an hour or more at room temperature. It was then purified using a cobalt affinity column and eluted with 300 mM NaCl, 300 mM imidazole, 50 mM MES, pH 6.5 containing 0.1% DDM. The detergent solubilized protein was further purified by sizing column gel filtration into 150 mM NaCl, 20 mM Tris-HCl, pH 7.5 with 0.1% DDM. Samples were concentrated to ~10 mg/ml, flash-frozen in liquid nitrogen, and stored at −80 °C. The absorbance spectra of these concentrated samples were recorded from 2 μl samples using a Nanodrop 1000-C.

Liposome Reconstitution and Proton-Pumping Measurements

Purified protein was reconstituted into soybean liposomes. For proton-pumping liposome assays, samples were illuminated to initiate proton pumping and monitored by a digitally connected pH meter.

Soybean lipids were reconstituted at 2% w/v into 50 mM KCl, 100 mM KPi, pH 7.0 with 14 mM octyl glucoside. Protein was added at a range between 10 and 40 μg protein/mg lipid. Liposomes were formed by dialysis; the first two incubations contained 50 mM KCl, 100 mM KPi, pH 7.0, whereas the third contained 1.9 M KCl, 100 mM KPi, pH 7.0. All dialysis changes proceeded 8 h to overnight. Protein samples were removed and the volume change recorded. Samples were flash-frozen in liquid nitrogen and stored at −80 °C.

For proton-pumping experiments, the reconstituted samples were removed and subjected to three freeze–thaw cycles in liquid nitrogen. They were then passed through a 0.4 m filter 21 times, using the Liposofast device from Avestin (Ottawa, ON, Canada). Liposomes were then passed over a column containing Sephadex G50 beads suspended in 2 M KCl. Liposome samples were added to make ~2.0 μM protein in 2 ml of 2 M KCl. Valinomycin dissolved in dimethyl sulfoxide was added to a final concentration of 2 μg/ml. The sample was then illuminated under saturating conditions with a 300 W halogen lamp, and the pH change was recorded as relative millivolts using an Ionalyzer analog pH meter with signal digitized with a DataQ (Akron, OH) digitizer. At the end of each experiment, 2 μg/ml of carbonyl cyanide-4-(trifluoromethoxy)phenylhydrazone (FCCP) was added to diffuse the liposome pH gradient. The system was finally calibrated with 1, 2, and 5 μl of 10 mM HCl. Rates were determined from the initial slope as pH increases with time. The high variability in the rates we observe (approximately ±50% relative standard deviation) is typical of these types of vesicle reconstitution pumping experiments (Hackett et al. 1987).

Supplementary Material

Supplementary figures S1–S3 and tables S1 and S2 are available at Molecular Biology and Evolution online (http://www.mbe.oxfordjournals.org/).

Acknowledgments

The authors thank Chris Miller and Daniel Oprian for critical commentary. This work was supported by the National Institutes of Health (grant numbers 1R01GM094468 and 1R01GM096053).

References

Divergent Evolution of Rhodopsins · doi:10.1093/molbev/mst171


