

RESEARCH ARTICLE

# An Approach to Heterologous Expression of Membrane Proteins. The Case of Bacteriorhodopsin

Dmitry Bratanov<sup>1,2</sup>✉, Taras Balandin<sup>1</sup>✉, Ekaterina Round<sup>1,3,4,5</sup>, Vitaly Shevchenko<sup>1,6</sup>, Ivan Gushchin<sup>1,3,4,5,6</sup>, Vitaly Polovinkin<sup>1,3,4,5,6</sup>, Valentin Borschhevskiy<sup>1,6</sup>, Valentin Gordeliy<sup>1,3,4,5,6\*</sup>

**1** Institute of Complex Systems (ICS), ICS-6: Structural Biochemistry, Research Centre Jülich, Jülich, Germany, **2** Institute of Crystallography, University of Aachen (RWTH), Jägerstrasse 17–19, Aachen, Germany, **3** Univ. Grenoble Alpes, IBS, Grenoble, France, **4** CNRS, IBS, Grenoble, France, **5** CEA, IBS, Grenoble, France, **6** Research-Educational Centre “Bionanophysics”, Moscow Institute of Physics and Technology, Dolgoprudny, Russia

✉ These authors contributed equally to this work.

\* [valentin.gordeliy@ibs.fr](mailto:valentin.gordeliy@ibs.fr)



**OPEN ACCESS**

**Citation:** Bratanov D, Balandin T, Round E, Shevchenko V, Gushchin I, Polovinkin V, et al. (2015) An Approach to Heterologous Expression of Membrane Proteins. The Case of Bacteriorhodopsin. PLoS ONE 10(6): e0128390. doi:10.1371/journal.pone.0128390

**Academic Editor:** Dimitrios Fotiadis, University of Bern, SWITZERLAND

**Received:** January 9, 2015

**Accepted:** April 24, 2015

**Published:** June 5, 2015

**Copyright:** © 2015 Bratanov et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** All relevant data are within the paper.

**Funding:** The work was supported by the program "Chaires d'excellence" edition 2008 of ANR France, CEA(IFS) – HGF(FZJ) STC 5.1 specific agreement; the MC grant (Marie Curie, FP7-PEOPLE-2007-1-1-ITN, project SBMPs); and an EC FP7 grant for the EDICT consortium (HEALTH-201924). Part of this work was supported by BMBF (PhoNa – Photonic Nanomaterials). The work was partially supported by the 5 top100-program of the Ministry for science and education of Russia and by RFBR, research project

## Abstract

Heterologous overexpression of functional membrane proteins is a major bottleneck of structural biology. Bacteriorhodopsin from *Halobium salinarum* (bR) is a striking example of the difficulties in membrane protein overexpression. We suggest a general approach with a finite number of steps which allows one to localize the underlying problem of poor expression of a membrane protein using bR as an example. Our approach is based on constructing chimeric proteins comprising parts of a protein of interest and complementary parts of a homologous protein demonstrating advantageous expression. This complementary protein approach allowed us to increase bR expression by two orders of magnitude through the introduction of two silent mutations into bR coding DNA. For the first time the high quality crystals of bR expressed in *E. Coli* were obtained using the produced protein. The crystals obtained with *in meso* nanovolume crystallization diffracted to 1.67 Å.

## Introduction

The membrane protein (MP) genes comprise about one third of the human genes encoding proteins. In spite of the recent success with GPCRs [1, 2], membrane proteins (MPs) still only comprise about 1% of total structures in the PDB [3, 4]. One of the bottlenecks in structural biology for MPs is the difficulty of production of pure and functional MPs, especially of human origin. However even bacterial and archaea MPs often do not express heterologously. One of the striking examples is the expression of bacteriorhodopsin from *Halobium salinarum* (bR) in *E. coli*.

Bacteriorhodopsin is a light-driven pump that provides proton vectorial transport across the cell membrane of the archaea [5, 6]. It consists of 7 transmembrane  $\alpha$ -helices (7TM) with the retinal molecule bound covalently to the Lys 216 residue [7] and belongs to the 7TM

No. 13-04-91320. The authors greatly acknowledge support of this work by ONEXIM, Russia. The work used the platforms of the Grenoble Instruct centre (ISBG; UMS 3518 CNRS-CEA-UJF-EMBL) with support from FRISBI (ANR-10-INSB-05-02) and GRAL (ANR-10-LABX-49-01) within the Grenoble Partnership for Structural Biology (PSB). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing Interests:** The funding from a commercial source (ONEXIM) does not alter the authors' adherence to PLOS ONE policies on sharing data and materials.

protein superfamily. As part of one of the simplest energy production mechanisms in the cell, bR is of great interest for bioenergetics. Due to the high level of homologous expression, combined with the ease of purification from the natural source [8] and high thermal and chemical stability [9] this protein has become the most studied MP. BR is a widely used model for expression, folding, crystallization of MPs and X-ray crystallography. Highly ordered 3D crystals of bR [10, 11] significantly improved understanding of the molecular mechanism of MP function and of vectorial proton transport in particular [12]. This protein and its mutants are in high demand in studies regarding their applications in bioelectronics, optics, and optogenetics [13–15]. Unfortunately, homologous production of bR mutants and many other light-driven proteins is laborious, time- and resource-consuming or is simply not possible at all.

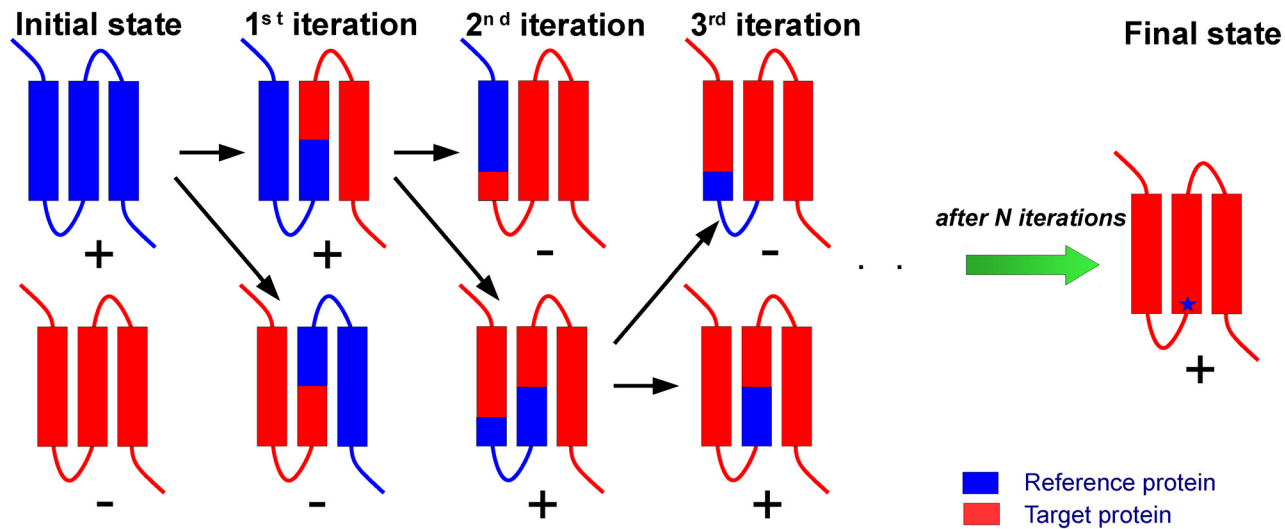
The bR functional expression in *E. coli* would be ideal as it is the most simple, robust, and inexpensive system [16]. However, for almost 30 years continued efforts towards this goal have been unsuccessful. The first studies showed a low level of wild type bR expression in *E. coli* due to severe degradation of the newly synthesized protein [17, 18]. Application of exogenous *N*-terminal tags has been used to stabilize the protein and prevent its degradation thereby increasing the bR yield to 17 mg of protein per litre of culture [18, 19]. Further improvements in yield were from the use of fusion proteins giving expression of 100–200 mg/l [20–22]. However, despite the high yield the above mentioned systems did not result in functional bR expression requiring refolding of the purified protein. Such behavior of bR is surprising, as a number of closely related retinal proteins were expressed in a functional form in *E. coli*, namely halorhodopsin and sensory rhodopsin II from *Natronomonas pharaonis* (hR and SRII, respectively) [23, 24], rhodopsin from *Exiguobacterium sibiricum* (ESR) [25], deltarhodopsin from *Haloterrigena turkmenica* [26], and bacteriorhodopsin from *Haloarcula marismortui* [27].

This raises the questions why are the other 7TM retinal proteins readily overexpressed in *E. coli* and why is bR not? What is the difference between bR and SRII that allows functional expression of SRII in *E. coli*? We addressed these questions by applying the complementary protein approach (CPA, shown in Fig 1). Chimeric proteins were constructed which comprised a part of bR and a complementary part of a reference protein SRII. SRII was chosen as this retinal protein was expressed functionally in *E. coli* [23, 24] and has high sequence identity to bR (27.49%) plus the high resolution 3D crystal structure is available [28, 29]. This approach allowed us to localize quickly the reason for the lack of bR expression in *E. coli* and we suggest that it may have a general application.

## Results

### Identification of the part of bacteriorhodopsin amino acid sequence responsible for failure of its expression in *E. coli* and strong influence of positively charged Arg7 at N-terminus

We constructed chimeric proteins by combining complementary parts of bR and SRII (Fig 2). The names of chimeric proteins indicate the complementary residues which have been replaced in subscript. For example, we replaced the initial 43 amino acids of bR by the corresponding 36 amino acids of SRII to make chimeric protein SR<sub>1-37</sub>bR<sub>Δ(0-43)</sub>. To elucidate the influence of extracellular, transmembrane and cytoplasmic regions independently we constructed SR<sub>1-8</sub>bR<sub>Δ(0-9)</sub>, SR<sub>26-37</sub>bR<sub>Δ(28-43)</sub>, SR<sub>1-8,26-37</sub>bR<sub>Δ(0-9,28-43)</sub>, respectively (Fig 2). The proteins were expressed in *E. coli*, solubilized and purified from membranes isolated by ultracentrifugation using metal affinity chromatography under denaturing conditions. Identity of the purified proteins was confirmed by Western blotting with anti-*His*-tag antibodies. Replacement of the first 9 amino acids on the *N*-terminus of bR by the counterparts from SRII significantly improves the expression of the native bR gene (Fig 3). The yields of SR<sub>1-8</sub>bR<sub>Δ(0-9)</sub> and the native bR were quantified



**Fig 1. Representation of the complementary protein approach.** Two homologous MPs reference protein (blue) and target protein (red) which have high and low (or zero) expression in a selected organism respectively. One can construct chimeric proteins (1<sup>st</sup> iteration) comprising parts of the target protein and the reference protein. Analyzing expression yields of the chimeric proteins one can determine which part of the target protein is responsible for a heterologous expression failure. In the next step, one can divide this part of the protein and construct the next chimeric proteins (2<sup>nd</sup> iteration). Acting in the same manner one can finally localize the problem of the lack of high yield expression of the target protein (showed as blue star) in a limited number of steps. One can estimate the required number of genetic constructs as  $2 \cdot \log_2 N$  instead of  $2^N$  point mutations in the case of a random search for the problematic parts of a protein responsible for the failure of the expression, where N is the number of amino acids in the target protein.

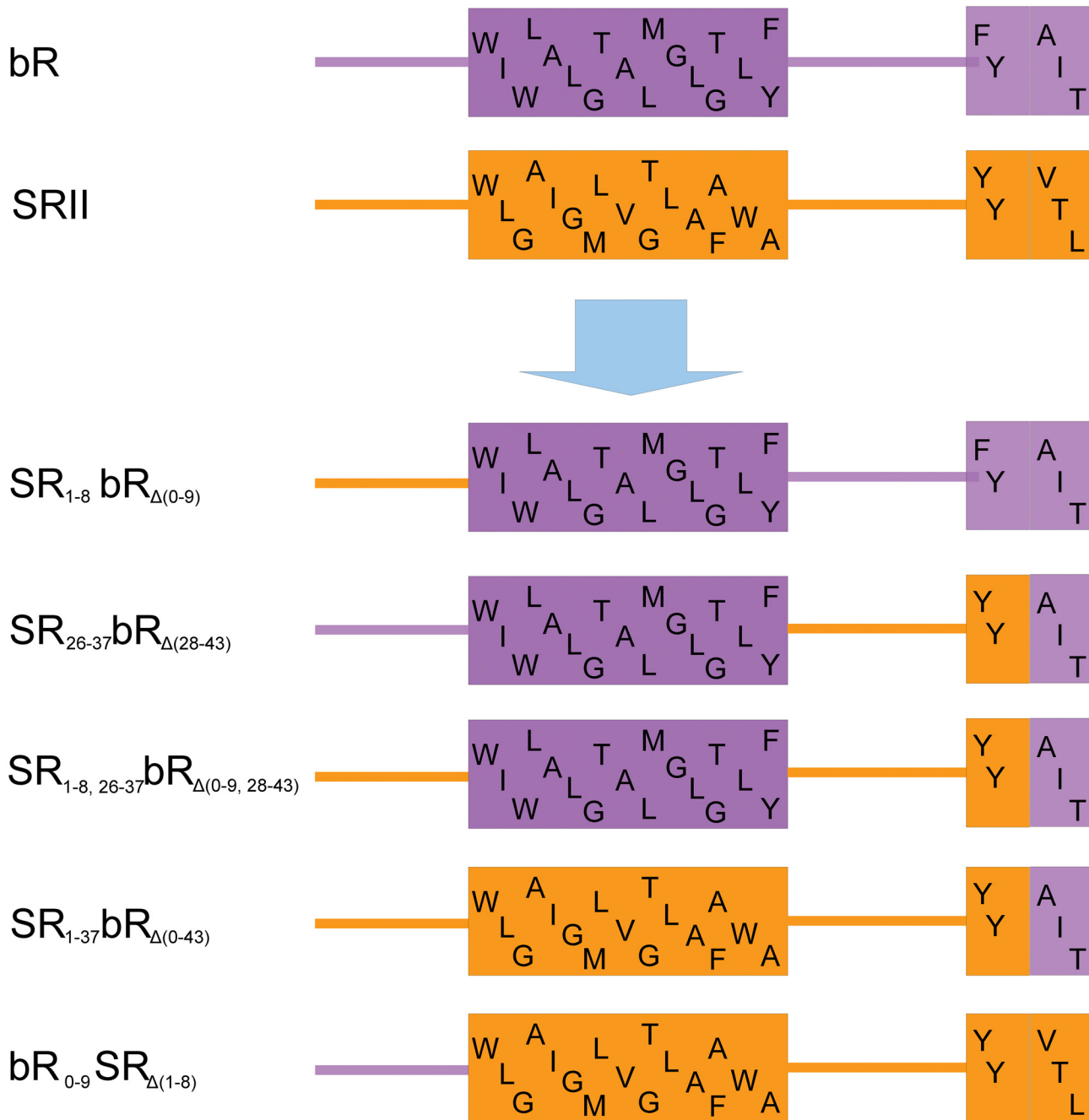
doi:10.1371/journal.pone.0128390.g001

by a BCA protein assay as  $7.0 \pm 1.0$  and  $0.14 \pm 0.04$  mg/l, respectively. The yield of chimeric protein is comparable with that of the native bR in *Halobium salinarum* (30 mg/litre of culture, [8]) and sufficient for structural studies and other purposes. Both chimeric proteins SR<sub>1-8</sub>bR $_{\Delta(0-9)}$  and SR<sub>1-37</sub>bR $_{\Delta(0-43)}$  purified under non-denaturing conditions were not functional, but were renatured using an established procedure of bR retinalization in DMPC/CHAPS micelles [30]. The reverse construct bR<sub>0-9</sub>SR $_{\Delta(1-8)}$  containing the first 10 amino acids of bR instead of their counterparts on the N-terminus of SR<sub>II</sub> (Fig 2) yielded  $6.0 \pm 1.8$  mg/l comparing to  $18.1 \pm 2.8$  mg/l for SR<sub>II</sub>.

The point mutants R7Q and R7E eliminating the only positive charge on extracellular N-terminus that do not satisfy “positive inside” rule were expressed and purified following the denaturing protocol with the yields of  $0.96 \pm 0.20$  mg and  $0.37 \pm 0.10$  mg of protein per litre of culture, respectively (Fig 3), which are correspondingly 7 and 3 times higher than that for the native bR but still 7 times lower than for a chimeric gene SR<sub>1-8</sub>bR $_{\Delta(0-9)}$ .

### Optimization of mRNA remarkably increases the yield of wild type bacteriorhodopsin

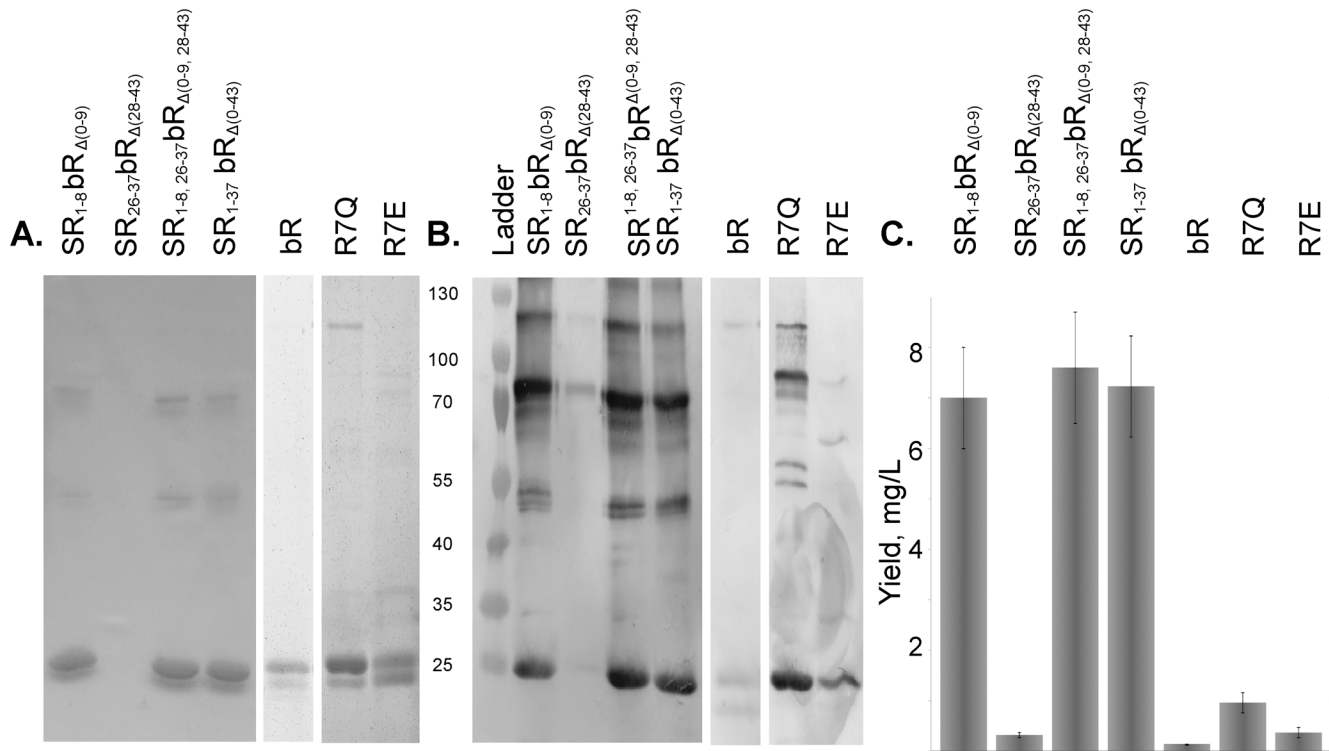
Using the mRNA structure prediction software *mRNAshapes* [31] we found a putative stem structure at the start of the bR gene (Fig 4). As the stem stretches beyond nucleotide 37, we had to expand the region under consideration to (-4, +47) comparing to the previous study [32]. This region corresponds to amino acids from Met0 to Leu15. The free folding energy of mRNA for the (-4, +47) region of bR is predicted to be  $-22.4$  kcal/mol (Fig 4), while the corresponding value for SR<sub>II</sub>, ESR, and hR genes are  $-12.3$ ,  $-5.2$ , and  $-13.1$  kcal/mol respectively, indicating the decreased stability of the mRNA structures at the 5'-terminus near the ribosome binding site. The change of the first 9 codons of bR to the counterparts of SR<sub>II</sub> considerably weakened the interactions between the 6–15 and 33–41 regions of this stem and increased the free energy by 13.0



**Fig 2. Design of chimeric proteins between bR and SR11.** N-terminus, first  $\alpha$ -helix and the beginning of the second  $\alpha$ -helix of bR (purple) and SR11 (orange) with the corresponding amino acid sequence are shown. Subscript in the name of the construct indicates amino acids of bR that were replaced by the counterparts from SR11 and vice versa.

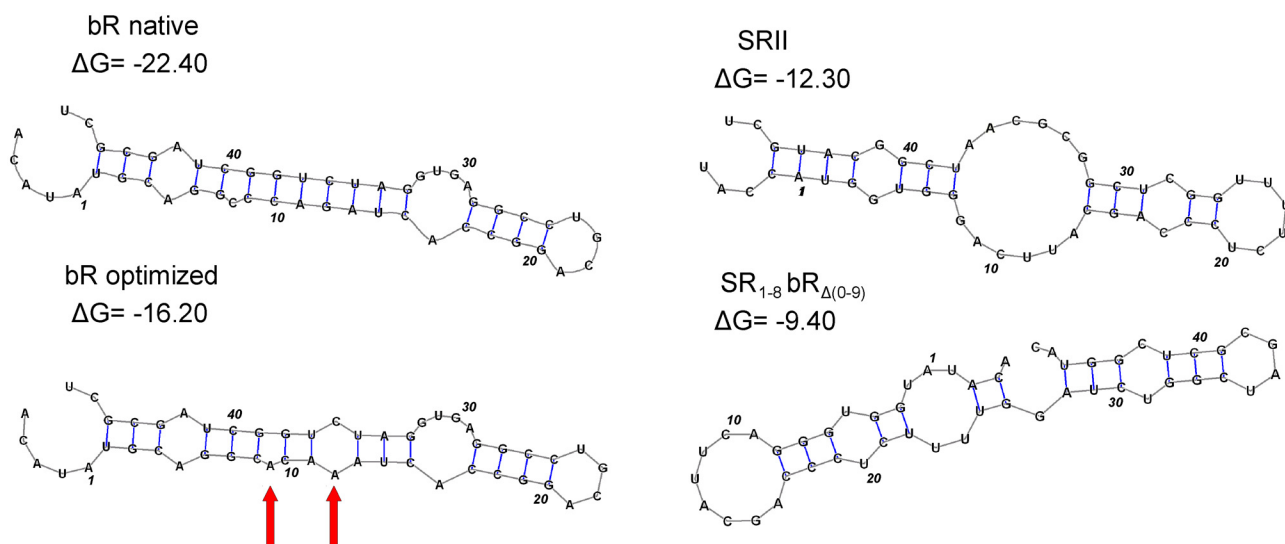
doi:10.1371/journal.pone.0128390.g002

*kcal/mol*. Conversely, the folding energy of bR<sub>0-9</sub>SR<sub>11Δ(1-8)</sub> is -15.5 *kcal/mol* which accounts for the lower yield of this chimera. Since the expression levels of the native bR, SR11, SR<sub>1-8</sub>bR<sub>Δ(0-9)</sub>, and bR<sub>0-9</sub>SR<sub>11Δ(1-8)</sub> correlated with the stability of the (-4, +47) region of mRNA structure, we introduced into the wild type bR gene two silent mutations (C9A, G12A corresponding to amino acids Ala2 and Gln3) that increase the free energy by 6.2 *kcal/mol* and significantly reduce the



**Fig 3. SDS-PAGE analysis of chimeric proteins, point mutants purified by metal affinity chromatography under denaturing conditions.** A. coomassie-stained gels. B. corresponding anti-*His*-tag antibody stained immunoblots. The amount of the protein loaded was normalized to the total amount of cell culture. C. The yields of the chimeric proteins and point mutants as quantified by the BCA protein assay, band density on SDS-PAGE and Western blotting. Chimeric proteins containing extracellular *N*-terminus of SR<sub>II</sub> have shown a much better yield.

doi:10.1371/journal.pone.0128390.g003



**Fig 4. The mRNA shapes of the native bR, SR<sub>II</sub>, chimeric protein SR<sub>1-8</sub>bR $\Delta$ (0-9), and optimized bR gene.** The region (-4, +47) with numbering starting from ATG start codon is shown. This region corresponds to amino acids from Met0 to Leu15. Free energies presented are expressed in *kcal/mol*. The values of mRNA free energy for SR<sub>II</sub> and SR<sub>1-8</sub>bR $\Delta$ (0-9) are considerably smaller than those of the native bR gene. The mRNA structure optimization of bR by silent mutations increased the free energy of the region (-4, +47) by 6.2 *kcal/mol*. The red arrows indicate the mutated nucleotides.

doi:10.1371/journal.pone.0128390.g004

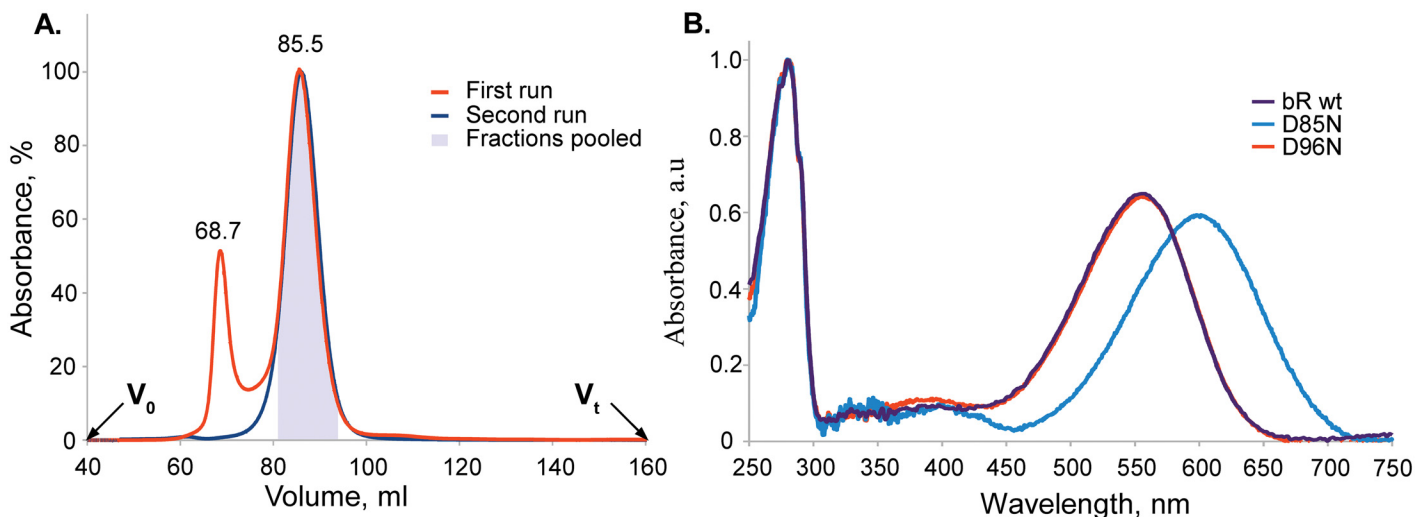
stem stability (Fig 4). The optimized bR gene was expressed in *E. coli* and purified following the denaturing protocol. The yield of the protein in case of the optimized gene was quantified as  $9.2 \pm 3.5$  mg per litre of culture, when measured by BCA protein assay, being essentially the same as that of SR<sub>1-8</sub>bR $\Delta_{(0-9)}$  ( $p = 0.74$ ) in contrast to the low yield of native gene (Fig 3).

### Purification and characterization of functional wild-type bacteriorhodopsin under non-denaturing conditions

When we utilized a mild detergent for solubilization of *E. coli* membranes and the non-denaturing protocol for protein purification, bR expressed from the optimized gene retained its purple color. The first step of purification by metal affinity chromatography led to mixed functional and nonfunctional bR preparations. The total yield of the protein after affinity purification was  $7.6 \pm 2.9$  mg per litre of culture, as measured by BCA protein assay. Sample homogeneity was confirmed by coomassie-stained SDS-PAGE.

As solubilized bR is unstable at alkaline pH and imidazole is harmful for the protein, we removed imidazole and adjusted pH to 6.0 by dialysis. During the pH exchange the protein heavily precipitated. The pellet was not colored and constituted of bR according to SDS-PAGE analysis suggesting aggregation of the misfolded protein. The pellet was discarded leaving the functional bR in the supernatant, however UV-Vis spectroscopy showed that the sample still contained aggregates and protein contaminants.

Further purification of bR utilised size-exclusion chromatography (SEC) yielding two distinct peaks at 69 ml and 86 ml corresponding to bR aggregates and functional bR respectively (Fig 5). The colored fractions from the latter peak were pooled, mixed and concentrated. The UV-Vis absorbance spectrum exhibited the retinal absorption peak at  $555.5 \pm 1.0$  nm. A peak ratio at  $A_{280}/A_{\lambda_{ret}}$  of 1.5 was achieved (Fig 5), demonstrating that the purity of the protein is consistent with that of bR solubilized from native purple membranes of *H. salinarum*. Storage stability of the protein was analysed 5 days after purification by repeat SEC on the same column.



**Fig 5. Characterization of bR expressed in *E. coli*.** A. Elution profile of the wild type bR on Sephacryl S200HR. The initial run is shown in red and the repeat run (after storage) in blue. The peak at 68.68 ml in the first run corresponds to the aggregated colorless protein. The second peak at 85.54 ml is the functional target protein. The fractions corresponding to the second peak were colored and those in the grey box were pooled. The repeat run after 5 days does not show any significant amount of aggregates. B. UV-Vis absorption spectra of the samples of the wild type bR and D85N and D96N mutants normalized by absorbance at 280 nm. Proteins were solubilized in DDM and purified using non-denaturing protocol. The spectrum of the wild type bR exhibited the retinal absorption peak at  $555.5 \pm 1.0$  nm with a peak ratio  $A_{280}/A_{\lambda_{ret}}$  of 1.5 which corresponds to the highest purity of bR.

doi:10.1371/journal.pone.0128390.g005

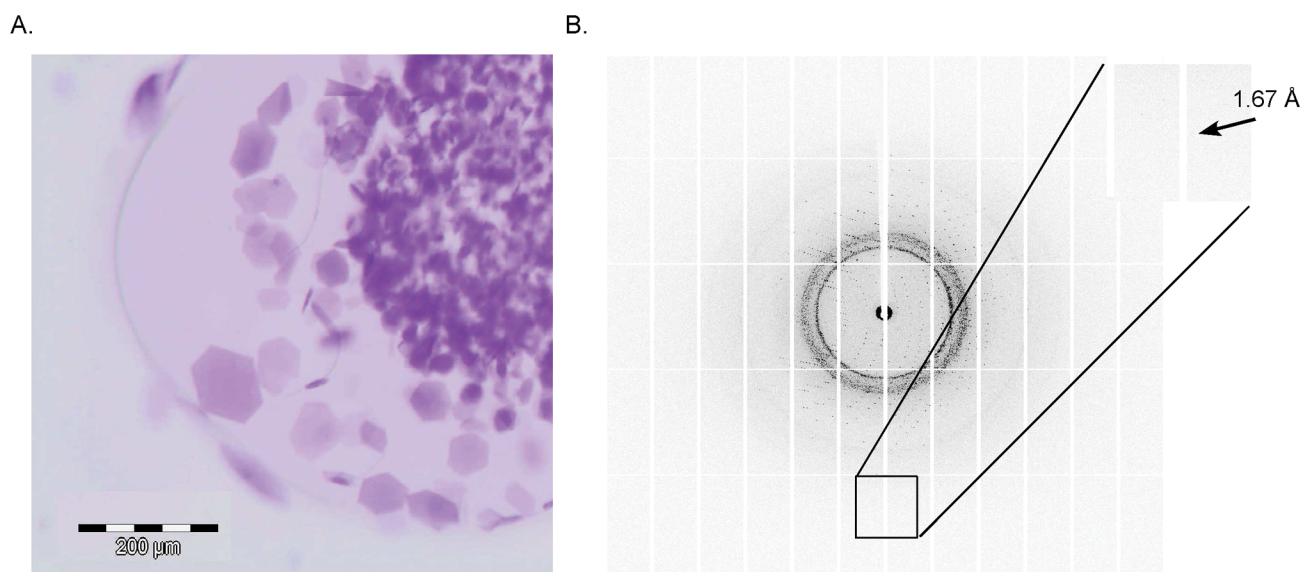
Bacteriorhodopsin eluted as a single and symmetric (asymmetry index 1.05) peak indicating size homogeneity of the final product. This confirmed that the aggregates were removed completely from the samples and the purified protein had no tendency to denature or form aggregates during storage. Prolonged storage indicated that the protein half-life at 4°C exceeds 120 days. The final yield of the purified functional wild type bR expressed in *E. coli* was  $2.4 \pm 1.3$  mg of the protein per litre of culture corresponding to 15–35% of the total synthesized bR.

### Expression, functional purification, and characterization of D85N and D96N mutants of bacteriorhodopsin

Using the mRNA-optimized bR gene we have introduced mutations D85N and D96N and utilizing the non-denaturing purification protocol we readily obtained functional mutant proteins with the final yields of 3.8 and 8.8 mg per litre of culture, respectively. The mutant proteins had essentially the same gel-filtration elution profiles as the wild-type bR. D96N UV-Vis spectrum was similar to the wild-type bR, while D85N mutant exhibited the characteristic maximum retinal absorbance for this protein at 598.0 nm (Fig 5).

### Crystallization of the purified proteins

With the wild type bR, D85N and D96N mutant proteins we set up nanovolume *in meso* crystallization trials [10, 33] which yielded crystals within 5–10 days for all proteins. The crystals of wild type protein and D96N mutant had a shape of thin hexagonal plates of 120  $\mu\text{m}$  (Fig 6). The single crystals obtained without optimization were tested at a synchrotron beam line and gave diffraction up to 1.67 Å resolution (Fig 6). The structure was solved with a resolution of 1.9 Å (see Table 1 for Crystallographic data statistics) and deposited in the PDB with ID code of 4XXJ. The crystals belonged to the spacegroup C2 with three monomers in the asymmetric unit that form a trimer. Typically for *in meso* crystallization, the protein molecules form membrane-like layers, with purple membrane-like packing within the layer (Fig 7). Despite the absence of native *Halobacterium salinarum* lipids, the intertrimer distance is essentially identical to that of



**Fig 6. Crystals of bR D96N mutant.** A. Crystals of bR D96N mutant expressed in *E. coli* were obtained using the *in meso* nanovolume crystallization approach. The largest crystals had the size of 120  $\mu\text{m}$ . B. A crystal (without further optimization of crystallization conditions) showed diffraction to 1.67 Å (inset).

doi:10.1371/journal.pone.0128390.g006

**Table 1. Crystallographic data collection and refinement statistics.**

| <b>Data collection</b>                              |                       |
|---|-----------------------|
| Space group   | C2                    |
| Cell dimensions                                     |                       |
| <i>a</i> , <i>b</i> , <i>c</i> (Å)                  | 105.84, 60.76, 113.36 |
| $\alpha$ , $\beta$ , $\gamma$ (°)                   | 90, 99.78, 90         |
| Resolution (Å)                                      | 49.2–1.9 (1.95–1.9)*  |
| <i>R</i> <sub>merge</sub> (%)                       | 7.1 (110.2)*          |
| <i>I</i> / $\sigma$ <i>I</i>                        | 13.7 (1.41)*          |
| Completeness (%)                                    | 98.8 (98.3)*          |
| Multiplicity  | 4.5 (4.5)*            |
| <b>Refinement</b>                                   |                       |
| Resolution  | 49.2–1.9 Å            |
| No. reflections                                     | 56113 (2796**)        |
| <i>R</i> <sub>work</sub> / <i>R</i> <sub>free</sub> | 18.35% / 21.39%       |
| Number of atoms                                     |                       |
| Protein   | 5306                  |
| Retinal   | 60                    |
| Lipid   | 248                   |
| Water   | 90                    |
| Average B-factor (Å <sup>2</sup> )                  |                       |
| Protein   | 34.4                  |
| Retinal   | 25.6                  |
| Lipid   | 51.2                  |
| Water   | 39.0                  |
| R.m.s. deviations                                   |                       |
| Bond lengths  | 0.005 Å               |
| Bond angles   | 0.9°                  |

\* Values in parentheses are for the highest-resolution shell.

\*\* Number of reflections that are not used for refinement (free reflections)

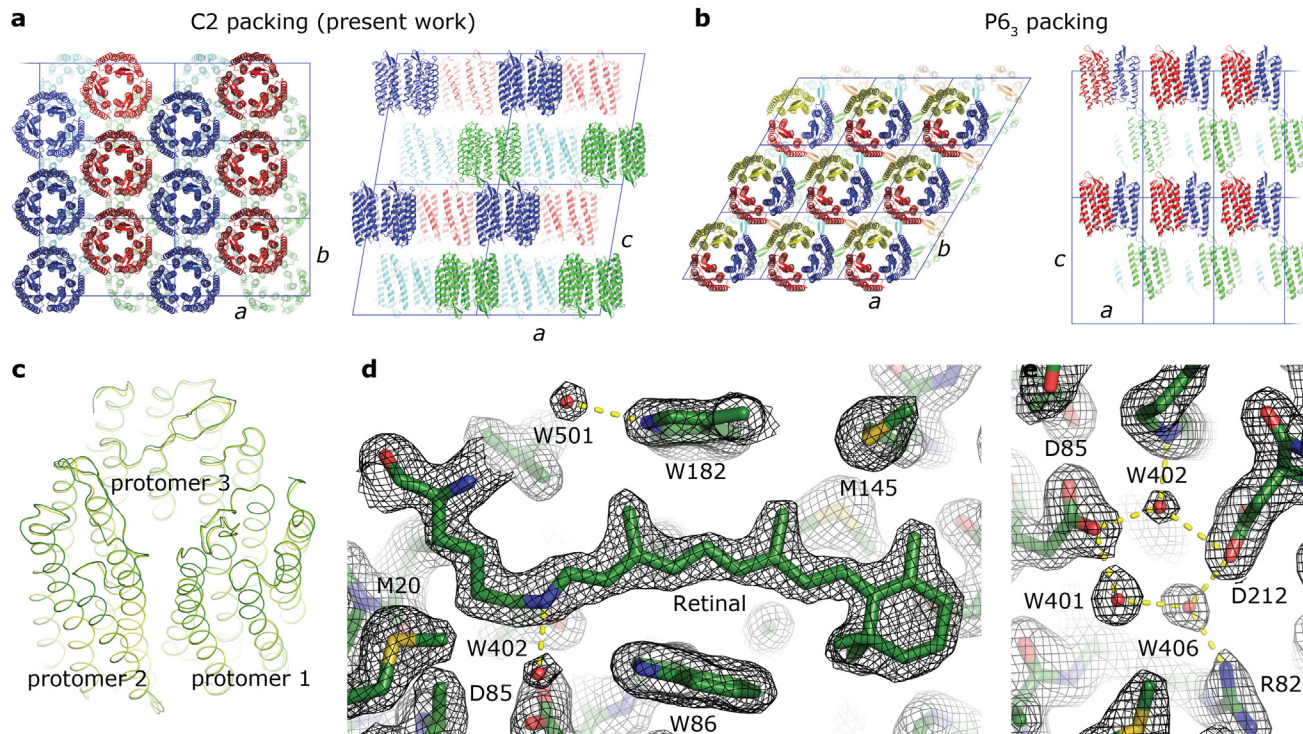
doi:10.1371/journal.pone.0128390.t001

the crystals obtained with purple membrane-derived bacteriorhodopsin. The structure of the protein and its trimer is also identical to that of purple membrane-derived bacteriorhodopsin (Fig 7). Data reveal the essential features of the structure, including the retinal cofactor and the hydrogen bond-linked cluster of water molecules close to the Schiff base (Fig 7).

## Discussion

In spite of the high level of sequence identity bR did not show functional expression in *E. coli* whereas to SRII, hR, and ESR did. To determine why this is the case we looked more coldly at the differences in the amino acid sequences (Fig 8). The “positive inside” rule states that MP topogenesis is controlled by positively charged amino acids [34] that keep the cytoplasmic parts of MP in the cytoplasm. This is due to the interactions with negatively charged head-groups of anionic lipids [35, 36] and these amino acids can therefore withstand both Sec-dependent and Sec-independent translocation against membrane potential. As observed *in vitro*, via the use of lysates and inner membrane vesicles (prepared from *E. coli*) bR is inserted cotranslationally into the *E. coli* membrane in a Sec-dependent manner [37] proving the amino acid sequence of the first transmembrane helix and adjacent regions to be particularly important for the insertion of integral MPs into membrane.





**Fig 7. Crystallographic structure of *E. coli*-expressed bacteriorhodopsin.** A. Packing of *E. coli*-expressed bacteriorhodopsin in the C2 crystals (three molecules per asymmetric unit). The inter-trimer distance is 60.76 Å. B. Packing of bacteriorhodopsin from purple membrane in the P6<sub>3</sub> crystals (one molecule per asymmetric unit). The inter-trimer distance is 60.63 Å [11]. C Comparison of the trimer structure in the C2 (green) and P6<sub>3</sub> (yellow) crystals. The structures are essentially identical with r.m.s.d. of the atom positions of less than 0.3 Å. D and E 2F<sub>o</sub>-F<sub>c</sub> electron densities around the retinal D and its Schiff base E contoured at the level of 1.5 σ. Hydrogen bonds are shown in yellow.

doi:10.1371/journal.pone.0128390.g007

The alignment of bR, SRII, ESR, and hR revealed two distinct features of bacteriorhodopsin in the region of the first transmembrane helix. First, the positively charged Arg7 on the extracellular *N*-terminus of bacteriorhodopsin deviates from the "positive inside" rule. Second, bacteriorhodopsin has a different positively charged amino acid pattern in the cytoplasmic loop A-B, where lysines substitute for the arginines in the vicinity of the hydrophobic-hydrophilic interface. Such a charge distribution on the membrane flanking regions of the first helix may compromise the folding and expression rate of bR. Thus, this preliminary considerations helped us to select the bR region that may contain parts of the protein problematic for *E. coli* expression. To investigate whether the charge distribution in the helix A region influences bR expression level in *E. coli* and the insertion of the protein into bacterial membrane we

|             |  |
|-------------|--|
| <i>hR</i>   | MTETLPPVTESAVALQAEVTQRELFEFVLNDPLLASSLYINIALAGLSI--LLFVFMTRGL---D-DPRAKL |
| <i>ESR</i>  | M-----EEVNLLVLATQYMFVWVGVGMAAGTLYFLVE-RNSLA--PEYRSTA                     |
| <i>SRII</i> | M-----VGLTTLFWLGAIGMLVGTLAFWAGRDAG--SGE-RRYY                             |
| <i>bR</i>   | M-----QAQITGRPEWIWLALGTALMGLGTYFLV--KGMGVSDDPAKKFY                       |

|-----|
|-----|
|-----|

**Extracellular *N*-terminus**
**Transmembrane**
**Loop A-B**

**Fig 8. Sequence alignment of HR, ESR, SRII and bR amino acids of the helix A.** Positively charged, negatively charged and hydrophobic amino acids are marked with blue, red and green colors respectively. Transmembrane domains of helix A as determined from the structure are marked with light grey. These specific features of bR, which may affect expression of the protein, were revealed in bR and are underlined.

doi:10.1371/journal.pone.0128390.g008

constructed several chimeric proteins by combining complementary parts of bR and SRII (Fig 2), as the latter is expressed in *E. coli* functionally with a high yield [23] and satisfies the "positive inside" rule (Fig 8).

There are already literature data on the construction of chimeric proteins between homologous proteins to find, for example, amino acid residues determining functional differences between these homologs (e.g. [38]). Moreover, on occasions when the chimeric proteins have a high expression yield or increased stability, they are expressed and even crystallized instead of their ancestors [39]. However, the systematic studies of the reason for the *E. coli* MP expression failure that involve the use of the chimeric proteins were lacking. Here, using bR as an example we show how the described approach can be employed for efficient MP production. A gain of over 50 times in the yield of bR expressed in *E. coli* was reached by the replacement of extracellular N-terminus of bR with the SRII amino acid sequence. It should be stressed that the part of the protein "responsible" for the failure of bR expression in *E. coli* was localized in 2 steps using 4 chimeric proteins. In general, CPA allows one to reduce the required number of genetic constructs from  $2^N$  point mutations to  $2 \cdot \log_2 N$  chimeric proteins, where N is the number of amino acids in the target protein.

Interestingly, the construct bR<sub>0-9</sub>SR $\Delta$ (1-8) containing the first 10 amino acids of bR instead of their counterparts on the N-terminus of SRII (Fig 2) yielded  $6.0 \pm 1.8$  mg/l comparing to  $18.1 \pm 2.8$  mg/l for SRII. Spectral tuning of SRII by constructing chimeric proteins between SRII and bR [40] produced in small amount the functional chimera comprising A-C helices of bR and D-G helices of SRII expressed in *E. coli*. In the case of the bR<sub>0-9</sub>SR $\Delta$ (1-8) construct under non-denaturing conditions a functional protein was not obtained. We hypothesize that the introduction of charged residues into the hydrophobic interface of helix A brakes its hydrophobic interactions with the other helices leading to the destabilization of the SRII structure and thus misfolding of the bR<sub>0-9</sub>SR $\Delta$ (1-8) chimera. Therefore, the suppressing influence of the extracellular N-terminus of bR has been supported by deterioration of functional SRII expression in *E. coli* when the N-terminus of SRII was replaced with its counterpart from bR.

Since the low expression level of wild type bR in *E. coli* comparing to SR<sub>1-8</sub>bR $\Delta$ (0-9) could be caused by the unfavorable mRNA structure of the bR gene near the ribosome binding site or the presence of positively charged Arg7 at the N-terminus side of the helix A of bR, we studied the influence of these two factors separately. First, we replaced Arg7 of the native bR gene by either neutral or negatively charged residue and constructed mutants R7Q and R7E respectively. The observed relationship between expression levels of native bR, R7Q and R7E mutants, and the SR<sub>1-8</sub>bR $\Delta$ (0-9) chimera in *E. coli* (Fig 3) could be explained in part by the "positive inside" rule as the removal of the positive charge on the N-terminus of bR improves the yield of the protein. However, these experiments showed that it is not the only reason for the low expression level of bR in *E. coli*.

The stability of mRNA folding near the ribosome binding site is known to have a strong impact on the protein expression level [32]. It has been shown that the variation of the nucleotide sequence near the ribosome binding site significantly influences the expression level of bR in *E. coli*, but functional incorporation of the protein into membrane has not been achieved [18]. Using the mRNA structure prediction software *mRNASHapes* [31] we found a putative stem structure at the start of the bR gene. The optimized bR gene yielded essentially the same amount of protein as the chimeric proteins in contrast to the native gene as demonstrated by Fig 3. Thereby, we showed that destabilization of the putative stem structure in the 5'-terminus of mRNA leads to a significant increase in the expression level of the protein.

The optimization of mRNA has also significantly increased the bR expression rate in the previous study with the use of pJP plasmid [18]. Additionally, in the mRNA 5' untranslated region the hairpin structure was found. It encloses the *lac* operator and hence impairs mRNA

translation initiation [18], whereas in case of the pSCodon plasmid we have not observed any adverse influence of the *lac* operator on the protein yield. This discrepancy could be explained by an insufficient spacing between the *lac* operator and the ribosome binding site in the pJP plasmid. The total yield of an unfolded bR was approximately 2 times lower than in the present study, because of a rapid degradation of the newly synthesized protein under the expression conditions.

Karnik et al. have shown that only 1–2% of bR synthesized in *E. coli* bound retinal, when cells were incubated at 37°C after induction with IPTG [18]. We suppose that the main factor that influences the yield of the functional bR in the case of *E. coli* expression may be the rate of protein synthesis. Indeed, it is recommended to express MPs at lower temperatures (20–30°C) to reduce the rate of the protein synthesis, facilitate its membrane insertion and proper folding [41]. As shown *in vitro* bR insertion into membrane occurs co-translationally in a Sec-dependent manner [37]. Its overexpression can overload the cell translocation system resulting in a misfolded protein and an increased rate of protein degradation. Induction of protein synthesis by lactose and cultivation of the cells at lowered temperature reduce the protein synthesis rate and thus favour an accumulation of the properly folded protein. Using similar mutations to optimize mRNA structure (-16.2 vs. -12.3 kcal/mol) we obtained the doubled translational yield of bR comparing to the previous study of Karnik et al. [18]. Since Karnik et al. employed a denaturing protein purification protocol, they had to renature the protein using DMPC/CHAPS vesicles, whereas our protocol takes advantage of non-denaturing conditions, thus allowing a straightforward production of the functional protein. The high rate of bR misfolding can also be attributed to an unfavorable lipid composition of the *E. coli* membrane. The PE lipids, a major component of the *E. coli* membrane, were shown to decrease the bR regeneration yield *in vitro* [42]. Also the folding of bR is described by a two-stage model with at least one transition state [43]. Therefore, *E. coli* lipids appear to have a significant effect on the transition state and favor bR misfolding.

Nevertheless, the protocol for high-yield bR production in *E. coli* introduced here allows one to obtain functional protein under non-denaturing conditions in quantities sufficient for structural biology and biochemical studies. Whereas other protocols require protein extraction with organic solvents and solubilization of the protein in denaturing detergents followed by protein renaturation as well as the necessity of the expression drivers [18–22], bR obtained using the presented protocol is stable, homogenous and resembles a native bR from purple membranes, and thus satisfies the requirements for the use of this protein in different applications in science and industry.

One of the main advantages of the *E. coli* expression system over *H. salinarum* is the considerably reduced time required to produce the mutants of interest facilitating intense studies of the target protein. We proved that the suggested approach is also efficient for fast production of bR mutants. Asp85 is primary proton acceptor and Asp96 is primary proton donor, therefore the D85N and D96N are the key mutants that are intensively used to trap the intermediate states of bR. The higher expression level of D85N and D96N mutants compared to wild type bR deserves additional attention. It has been previously reported that mutation D94N in the bacteriorhodopsin from *Haloarcula marismortui* (analogous to D96N in bR) raised 10-fold the yield of the functional protein [44]. In addition, it was shown that insertion of the bR helix C into membrane *in vitro* is impeded by two aspartic acid residues within its transmembrane region [45]. In the present work D85N and D96N mutations led to the 1.5 to 4-fold increase in the yield of the functional protein relative to the wild type bR, the fraction of a properly folded protein increased as well from 25±10% to 35% and 60% of the total transcriptional yield of D85N and D96N mutants, respectively. This increase may be explained by improved incorporation of the newly synthesized protein into the *E. coli* membrane that reduces protein degradation and facilitates the correct folding of the bR mutants.

The growth of the first highly ordered 3D crystals of bR expressed in *E. coli* and the similarity of its structure to the structure of bR obtained from *H. salinarum* show the aptness of the presented approach for the expression of bR and its mutants suitable for all scientific and industrial applications. Moreover, the obtained high-quality crystals of bR produced in *E. coli* lacking native lipids of *H. salinarum* addresses a long-standing question in crystallization of membrane proteins. Despite the role of native lipids for MP crystallization being considered important, there is little information published on this matter [46]. Particularly, archae lipids that stabilize bR molecules in the trimer inside the 2D and 3D crystals were assumed to be highly specific [47, 48]. These findings were supporting the idea that native lipids of *H. salinarum* are required to form highly ordered bR crystals. Our work demonstrates that bR expressed in *E. coli* is readily crystallized by *in meso* approach and, therefore, native lipids surrounding the protein are not an absolute requirement for growing well diffracting bR crystals. Although we have demonstrated high efficiency of the CPA in the case of microbial rhodopsin family we believe that the CPA may be applied to other membrane proteins.

## Materials and Methods

### Materials

All the salts and media components were purchased either from AppliChem (Darmstadt, Germany) or Sigma-Aldrich (Taufkirchen, Germany) of analytical quality or higher. All enzymes were from Fermentas (part of Thermo Fisher Scientific, St. Leon-Rot, Germany). DDM was from Affymetrix (Santa Clara, USA), Sarkosyl from AppliChem, retinal, DMPC, and CHAPS from Sigma-Aldrich.

### DNA manipulations

The coding region of bR gene excluding the leader signal peptide sequence was amplified from pEF191 plasmid kindly provided by D. Oesterhelt [49]. The factor Xa cleavage site and eight histidine purification tag coding sequences were appended to the bR 3'-terminus using synthetic oligonucleotides resulting in GSGIEGRSGAPHHHHHHHH extension. The obtained bR gene was cloned into *Nde*I and *Xho*I sites of pSCodon vector (Delphi Genetics S.A., Charleroi, Belgium). The pSOPII plasmid bearing the SRII gene in pET27bmod vector was provided kindly by M. Engelhard [23]. All bR-SRII chimeric genes (Fig 2) and bR mutant genes were produced by PCR.

The mRNA folding energies were calculated using *mRNashapes* software [31]. Bacteriorhodopsin coding sequence was optimized by introducing two silent mutations C9A and G12A corresponding to amino acids Ala2 and Gln3.

### Protein expression and membrane isolation

*E. coli* SE1 [50] cells were transformed with pSCodon-derived plasmids, and *E. coli* BL21(DE3) cells were transformed with pET27-derived plasmids. Cells were plated over LB-agar and then grown in 450 ml rich ZYP-5052 autoinduction media [51] with an appropriate antibiotic in baffled shaking 2-L flasks at 120 rpm. The cells were incubated at 37°C until OD<sub>600</sub> reached 1.0–1.2 AU, when 400 µl of 50 mM *all-trans* retinal solution in ethanol was added, and cells were further cultivated at 20°C. Cells from 1 litre of overnight culture were harvested by centrifugation at 5000 rpm, resuspended in 50 ml of 20 mM Tris-HCl pH 8.0, 5% glycerol, supplemented with 10 mg of lysozyme and 1 mg of DNaseI, and incubated for 2 h at 4°C with stirring. Lysate was obtained by passing the suspension 3 times through micro-fluidizer M-110P from Microfluidics (Westwood, USA). Then 5M NaCl was added to a final concentration

of 200 mM and suspension was layered over a glycerol cushion (1 ml—90%, 1 ml—80%, 1 ml—60%) in two 32 ml tubes. The total membranes were isolated by ultracentrifugation in SW-32Ti rotor (Beckman Coulter, Krefeld, Germany) at 28 000 rpm for 1h. The supernatant was discarded and the glycerol cushion containing membranes was resuspended in 50 ml of 20 mM Tris-HCl pH 8.0, 100 mM NaCl.

### Denaturing protein purification

Sarkosyl was added to the final concentration of 2% and membranes were solubilized overnight with stirring at 4°C. Insoluble material was removed by ultracentrifugation in Ti-70 (Beckman Coulter) rotor at 35 000 rpm for 1h. Supernatant was 5 times diluted with 20 mM Tris-HCl pH 8.0, 100 mM NaCl buffer and 10 mM imidazole was added. Suspension was loaded on the 5 ml of Ni-NTA resin (Qiagen, Hilden, Germany) equilibrated with the same buffer. The column was first washed with 3 CV of 50 mM Tris-HCl pH 8.0, 100 mM NaCl, 0.25% Sarkosyl, 20 mM imidazole buffer, then 5 CV of 100 mM Na<sub>2</sub>HPO<sub>4</sub> pH 8.0, 100 mM NaCl, 0.2% SDS, and 3 CV of 100 mM Na<sub>2</sub>HPO<sub>4</sub> pH 8.0, 0.2% SDS buffer. The protein was eluted with 3 CV of 100 mM Na<sub>2</sub>HPO<sub>4</sub> pH 8.0, 0.2% SDS, 300 mM imidazole. To remove imidazole the samples were dialysed against 0.8 L of 50 mM NaH<sub>2</sub>PO<sub>4</sub> pH 6.0, 0.2% SDS. When necessary purified proteins were refolded in mixed DMPC/CHAPS micelles [30]. Reliability of the protein yields comparison was estimated by Student's test and presented as *p* value.

### Non-denaturing protein purification

DDM was added to the final concentration of 1% and membranes were solubilized overnight with stirring at 4°C. Insoluble material was removed by ultracentrifugation in Ti-70 rotor at 35 000 rpm for 1h. Supernatant was 5 times diluted with 20 mM Tris-HCl pH 8.0, 100 mM NaCl buffer and 10 mM imidazole was added. Suspension was loaded on the 5 ml of Ni-NTA resin equilibrated with the same buffer. The column was washed with 10 CV of 50 mM NaH<sub>2</sub>PO<sub>4</sub> pH 6.0, 100 mM NaCl, 0.2% DDM, 30 mM imidazole. The protein was eluted with 3 CV of 50 mM NaH<sub>2</sub>PO<sub>4</sub> pH 7.4, 100 mM NaCl, 0.2% DDM, 300 mM imidazole. Only coloured fractions were pooled. To remove imidazole the samples were immediately dialysed against 0.6 L of 50 mM NaH<sub>2</sub>PO<sub>4</sub> pH 6.0, 100 mM NaCl for 2 hours and removed from dialysis buffer, then after 10 hours dialysis was continued for additional 2 hours against fresh buffer.

After the dialysis the protein heavily precipitated. The white pellet was separated by centrifugation at 5000 rpm from colored solution and discarded. Protein was concentrated to the volume of 2 ml by ultrafiltration and applied to 165 ml Sephacryl S200HR (GE Healthcare, Germany) column equilibrated with 50 mM NaH<sub>2</sub>PO<sub>4</sub> pH 6.0, 100 mM NaCl, 0.1% DDM. Peak of colored functional protein could be easily separated from the peak of the aggregated protein.

The samples were protected from the light. Total protein concentration was measured by BCA Protein Assay Kit (Thermo Fisher Scientific, Schwerte, Germany) following supplier protocol. To access the protein purity the samples were analyzed on 8–16% gradient SDS-PAGE. UV-Vis absorbance spectrum was measured on UV-2450 spectrophotometer (Shimadzu, Duisburg, Germany). The fraction of the functional protein was accessed as absorbance ratio  $A_{280}/A_{\lambda_{\text{ret}}}$ , where  $A_{\lambda_{\text{ret}}}$  is maximum absorbance of retinal in the protein [52].

### Crystallization, data collection, and structure determination

The crystals were grown using *in meso* approach using nanovolume robotic system Formulatrix NT8 (Waltham, USA) [10, 33], similarly to our previous work [28, 53, 54]. The purified protein in crystallization buffer was added to the monooleyl-based lipid mesophase. The best crystals were obtained using the protein concentration of 20 mg/ml and 1.5M Na/K-Pi pH 5.6

precipitation solution. The crystals were grown at 22°C and reached 120 µm in size in approximately one week.

X-ray diffraction data (wavelength 0.976 Å) were collected at the beamline ID23-1 of the European Synchrotron Radiation Facility (ESRF, Grenoble, France) using a PILATUS 6M detector (Dectris Ltd., Baden, Switzerland). Diffraction patterns integrated using XDS and the reflexes' intensities were scaled using and XSCALE [55], and the space group was determined to be C2. The data statistics are presented in the [Table 1](#). Initial phases were successfully obtained in the space group C2 by molecular replacement method using MOLREP [56], with PDB ID 1C3W [11] as a search model. There are three monomers in the asymmetric unit in this space group. The initial molecular replacement model was then iteratively refined using REFMAC5 [57] and Coot [58].

## Accession numbers

The atomic coordinate and structure factors have been deposited in the PDB with ID code 4XXJ.

## Acknowledgments

The diffraction experiments were performed on the ID14-1 and ID23-1 beamlines at the European Synchrotron Radiation Facility (ESRF), Grenoble, France. We are grateful to A. Popov and other local contacts at ESRF for providing assistance in using these beamlines.

## Author Contributions

Conceived and designed the experiments: DB TB VG. Performed the experiments: DB TB ER VS IG VP VB VG. Analyzed the data: DB TB VS IG VG. Wrote the paper: DB TB IG ER VG.

## References

1. Cherezov V, Rosenbaum DM, Hanson MA, Rasmussen SG, Thian FS, Kobilka TS, et al. High-resolution crystal structure of an engineered human beta2-adrenergic G protein-coupled receptor. *Science*. 2007; 318(5854): 1258–65. PMID: [17962520](#)
2. Rasmussen SG, DeVree BT, Zou Y, Kruse AC, Chung KY, Kobilka TS, et al. Crystal structure of the β2 adrenergic receptor-Gs protein complex. *Nature*. 2011; 477: 549–55. doi: [10.1038/nature10361](#) PMID: [21772288](#)
3. Tusnády GE, Dosztányi Z, Simon I. Transmembrane proteins in the Protein Data Bank: identification and classification. *Bioinformatics*. 2004; 20: 2964–72. PMID: [15180935](#)
4. White SH. Biophysical dissection of membrane proteins. *Nature*. 2009; 459: 344–6. doi: [10.1038/nature08142](#) PMID: [19458709](#)
5. Kayushin LP, Skulachev VP. Bacteriorhodopsin as an electrogenic proton pump: reconstitution of bacteriorhodopsin proteoliposomes generating delta psi and delta pH. *FEBS Lett*. 1974; 39: 39–42. PMID: [4850730](#)
6. Oesterhelt D, Stoeckenius W. Functions of a new photoreceptor membrane. *Proc Natl Acad Sci U S A*. 1973; 70: 2853–7. PMID: [4517939](#)
7. Huang KS, Liao MJ, Gupta CM, Royal N, Biemann K, Khorana HG. The site of attachment of retinal in bacteriorhodopsin. The epsilon-amino group in Lys-41 is not required for proton translocation. *J Biol Chem*. 1982; 257: 8596–9. PMID: [6284736](#)
8. Lorber B, DeLucas LJ. Large scale preparation of homogeneous bacteriorhodopsin. *FEBS Lett*. 1990; 261: 14–8. PMID: [2407551](#)
9. Brouillette CG, McMichens RB, Stern LJ, Khorana HG. Structure and thermal stability of monomeric bacteriorhodopsin in mixed phospholipid/detergent micelles. *Proteins*. 1989; 5: 38–46. PMID: [2748571](#)
10. Landau EM, Rosenbusch JP. Lipidic cubic phases: a novel concept for the crystallization of membrane proteins. *Proc Natl Acad Sci U S A*. 1996; 93(25): 14532–5. PMID: [8962086](#)
11. Luecke H, Schobert B, Richter HT, Cartailler JP, Lanyi JK. Structure of bacteriorhodopsin at 1.55 Å resolution. *J Mol Biol*. 1999; 291(4): 899–911. PMID: [10452895](#)

12. Lanyi JK. Bacteriorhodopsin. *Annu Rev Physiol.* 2004; 66: 665–88. PMID: [14977418](#)
13. Birge RR, Gillespie NB, Izaguirre EW, Kusnetzow A, Lawrence AF, Singh D, et al. Biomolecular Electronics: Protein-Based Associative Processors and Volumetric Memories. *J Phys Chem B.* 1999; 103 (49): 10746–66.
14. Boyden ES, Zhang F, Bamberg E, Nagel G, Deisseroth K. Millisecond-timescale, genetically targeted optical control of neural activity. *Nat Neurosci.* 2005; 8(9): 1263–8. PMID: [16116447](#)
15. Hampp N. Bacteriorhodopsin as a Photochromic Retinal Protein for Optical Memories. *Chem Rev.* 2000; 100: 1755–76. PMID: [11777419](#)
16. Francis DM, Page R. Strategies to optimize protein expression in *E. coli*. *Curr Protoc Protein Sci.* 2010; Chapter 5: Unit 5.24.1–29.
17. Karnik S, Doi T, Molday R, Khorana HG. Expression of the archaeobacterial bacterio-opsin gene with and without signal sequences in *Escherichia coli*: the expressed proteins are located in the membrane but bind retinal poorly. *Proc Natl Acad Sci U S A.* 1990; 87: 8955–9. PMID: [2247471](#)
18. Karnik SS, Nassal M, Doi T, Jay E, Sgaramella V, Khorana HG. Structure-function studies on bacteriorhodopsin. II. Improved expression of the bacterio-opsin gene in *Escherichia coli*. *J Biol Chem.* 1987; 262: 9255–63. PMID: [3298253](#)
19. Shand RF, Miercke LJ, Mitra AK, Fong SK, Stroud RM, Betlach MC. Wild-type and mutant bacterioopsins D85N, D96N, and R82Q: high-level expression in *Escherichia coli*. *Biochemistry.* 1991; 30: 3082–8. PMID: [2007142](#)
20. Chen GQ, Gouaux JE. Overexpression of bacterio-opsin in *Escherichia coli* as a water-soluble fusion to maltose binding protein: efficient regeneration of the fusion protein and selective cleavage with trypsin. *Protein Sci.* 1996; 5: 456–67. PMID: [8868482](#)
21. Nekrasova OV, Wulfson AN, Tikhonov RV, Yakimov SA, Simonova TN, Tagvey AI, et al. A new hybrid protein for production of recombinant bacteriorhodopsin in *Escherichia coli*. *J Biotechnol.* 2010; 147: 145–50. doi: [10.1016/j.jbiotec.2010.03.019](#) PMID: [20363267](#)
22. Pompejus M, Friedrich K, Teufel M, Fritz HJ. High-yield production of bacteriorhodopsin via expression of a synthetic gene in *Escherichia coli*. *Eur J Biochem.* 1993; 211: 27–35. PMID: [7916683](#)
23. Hohenfeld IP, Wegener AA, Engelhard M. Purification of histidine tagged bacteriorhodopsin, pharaonis halorhodopsin and pharaonis sensory rhodopsin II functionally expressed in *Escherichia coli*. *FEBS Lett.* 1999; 442: 198–202. PMID: [9929001](#)
24. Shimono K, Iwamoto M, Sumi M, Kamo N. Functional expression of pharaonis phoborhodopsin in *Escherichia coli*. *FEBS Lett.* 1997; 420: 54–6. PMID: [9450549](#)
25. Petrovskaya LE, Lukashev EP, Chupin VV, Sychev SV, Lyukmanova EN, Kryukova EA, et al. Predicted bacteriorhodopsin from *Exiguobacterium sibiricum* is a functional proton pump. *FEBS Lett.* 2010; 584: 4193–6. doi: [10.1016/j.febslet.2010.09.005](#) PMID: [20831870](#)
26. Kamo N, Hashiba T, Kikukawa T, Araiso T, Ihara K, Nara T. A light-driven proton pump from *Haloterrigena turkmenica*: functional expression in *Escherichia coli* membrane and coupling with a H<sup>+</sup> co-transporter. *Biochem Biophys Res Commun.* 2006; 341: 285–90. PMID: [16413498](#)
27. Fu HY, Lin YC, Chang YN, Tseng H, Huang CC, Liu KC, et al. A novel six-rhodopsin system in a single archaeon. *J Bacteriol.* 2010; 192: 5866–73. doi: [10.1128/JB.00642-10](#) PMID: [20802037](#)
28. Gordeliy VI, Labahn J, Moukhametzanov R, Efremov R, Granzin J, Schlesinger R, et al. Molecular basis of transmembrane signalling by sensory rhodopsin II-transducer complex. *Nature.* 2002; 419 (6906): 484–7. PMID: [12368857](#)
29. Royant A, Nollert P, Edman K, Neutze R, Landau EM, Pebay-Peyroula E, et al. X-ray structure of sensory rhodopsin II at 2.1-Å resolution. *Proc Natl Acad Sci U S A.* 2001; 98: 10131–6. PMID: [11504917](#)
30. Booth PJ, Riley ML, Flitsch SL, Templer RH, Farooq A, Curran AR, et al. Evidence that bilayer bending rigidity affects membrane protein folding. *Biochemistry.* 1997; 36: 197–203. PMID: [8993334](#)
31. Steffen P, Voss B, Rehmsmeier M, Reeder J, Giegerich R. RNAshapes: an integrated RNA analysis package based on abstract shapes. *Bioinformatics.* 2006; 22: 500–3. PMID: [16357029](#)
32. Kudla G, Murray AW, Tollervey D, Plotkin JB. Coding-sequence determinants of gene expression in *Escherichia coli*. *Science.* 2009; 324: 255–8. doi: [10.1126/science.1170160](#) PMID: [19359587](#)
33. Caffrey M, Cherezov V. Crystallizing membrane proteins using lipidic mesophases. *Nat Protoc.* 2009; 4: 706–31. doi: [10.1038/nprot.2009.31](#) PMID: [19390528](#)
34. von Heijne G. Control of topology and mode of assembly of a polytopic membrane protein by positively charged residues. *Nature.* 1989; 341: 456–8. PMID: [2677744](#)
35. Bogdanov M, Heacock P, Guan Z, Dowhan W. Plasticity of lipid-protein interactions in the function and topogenesis of the membrane protein lactose permease from *Escherichia coli*. *Proc Natl Acad Sci U S A.* 2010; 107: 15057–62. doi: [10.1073/pnas.1006286107](#) PMID: [20696931](#)

36. van Klompenburg W, Nilsson I, von Heijne G, de Kruijff B. Anionic phospholipids are determinants of membrane protein topology. *EMBO J.* 1997; 16: 4261–6. PMID: [9250669](#)
37. Raine A, Ullers R, Pavlov M, Luirink J, Wikberg JE, Ehrenberg M. Targeting and insertion of heterologous membrane proteins in *E. coli*. *Biochimie.* 2003; 85: 659–68. PMID: [14505821](#)
38. Wang H, Sugiyama Y, Hikima T, Sugano E, Tomita H, Takahashi T, et al. Molecular determinants differentiating photocurrent properties of two channelrhodopsins from *Chlamydomonas*. *J Biol Chem.* 2009; 284: 5685–96. doi: [10.1074/jbc.M807632200](#) PMID: [19103605](#)
39. Kato HE, Zhang F, Yizhar O, Ramakrishnan C, Nishizawa T, Hirata K, et al. Crystal structure of the channelrhodopsin light-gated cation channel. *Nature.* 2012; 482: 369–74. doi: [10.1038/nature10870](#) PMID: [22266941](#)
40. Shimono K, Hayashi T, Ikeura Y, Sudo Y, Iwamoto M, Kamo N. Importance of the broad regional interaction for spectral tuning in *Natronobacterium pharaonis* phoborhodopsin (sensory rhodopsin II). *J Biol Chem.* 2003; 278: 23882–9. PMID: [12690098](#)
41. Samuelson J. *Production of Membrane Proteins: Strategies for Expression and Isolation.* Wiley-VCH Verlag GmbH & Co.; 2011.
42. Curran AR, Templer RH, Booth PJ. Modulation of folding and assembly of the membrane protein bacteriorhodopsin by intermolecular forces within the lipid bilayer. *Biochemistry.* 1999; 38: 9328–36. PMID: [10413507](#)
43. Booth PJ, Clarke J. Membrane protein folding makes the transition. *Proc Natl Acad Sci U S A.* 2010; 107: 3947–8. doi: [10.1073/pnas.0914478107](#) PMID: [20173094](#)
44. Hsu MF, Yu TF, Chou CC, Fu HY, Yang CS, Wang AH. Using Haloarcula marismortui Bacteriorhodopsin as a Fusion Tag for Enhancing and Visible Expression of Integral Membrane Proteins in *Escherichia coli*. *PLoS One.* 2013; 8: e56363. doi: [10.1371/journal.pone.0056363](#) PMID: [23457558](#)
45. Hunt JF, Rath P, Rothschild KJ, Engelman DM. Spontaneous, pH-dependent membrane insertion of a transbilayer alpha-helix. *Biochemistry.* 1997; 36: 15177–92. PMID: [9398245](#)
46. Cartailleur JP, Luecke H. X-ray crystallographic analysis of lipid-protein interactions in the bacteriorhodopsin purple membrane. *Annu Rev Biophys Biomol Struct.* 2003; 32: 285–310. PMID: [12598369](#)
47. Belrhali H, Nollert P, Royant A, Menzel C, Rosenbusch JP, Landau EM, et al. Protein, lipid and water organization in bacteriorhodopsin crystals: a molecular view of the purple membrane at 1.9 Å resolution. *Structure.* 1999; 7(8): 909–17. PMID: [10467143](#)
48. Essen L, Siegert R, Lehmann WD, Oesterhelt D. Lipid patches in membrane protein oligomers: crystal structure of the bacteriorhodopsin-lipid complex. *Proc Natl Acad Sci U S A.* 1998; 95: 11673–8. PMID: [9751724](#)
49. Ferrando E, Schweiger U, Oesterhelt D. Homologous bacterio-opsin-encoding gene expression via site-specific vector integration. *Gene.* 1993; 125: 41–7. PMID: [8383625](#)
50. Szpirer CY, Milinkovitch MC. Separate-component-stabilization system for protein and DNA production without the use of antibiotics. *Biotechniques.* 2005; 38: 775–81. PMID: [15945374](#)
51. Studier WF. Protein production by auto-induction in high density shaking cultures. *Protein Expr Purif.* 2005; 41: 207–34. PMID: [15915565](#)
52. Miercke LJ, Ross PE, Stroud RM, Dratz EA. Purification of bacteriorhodopsin and characterization of mature and partially processed forms. *J Biol Chem.* 1989; 264: 7531–5. PMID: [2708376](#)
53. Gushchin I, Chervakov P, Kuzmichev P, Popov AN, Round E, Borshchevskiy V, et al. Structural insights into the proton pumping by unusual proteorhodopsin from nonmarine bacteria. *Proc Natl Acad Sci U S A.* 2013; 110(31): 12631–6. doi: [10.1073/pnas.1221629110](#) PMID: [23872846](#)
54. Shevchenko V, Gushchin I, Polovinkin V, Round E, Borshchevskiy V, Utrobin P, et al. Crystal Structure of *Escherichia coli*-Expressed *Haloarcula marismortui* Bacteriorhodopsin I in the Trimeric Form. *PLoS One.* 2014; 9(12): e112873. doi: [10.1371/journal.pone.0112873](#) PMID: [25479443](#)
55. Kabsch W. XDS. *Acta Crystallogr D Biol Crystallogr.* 2010; 66(Pt 2): 125–32. doi: [10.1107/S0907444909047337](#) PMID: [20124692](#)
56. Vagin A, Teplyakov A. Molecular replacement with MOLREP. *Acta Crystallogr D Biol Crystallogr.* 2010; 66(Pt 1): 22–5. doi: [10.1107/S0907444909042589](#) PMID: [20057045](#)
57. Murshudov GN, Skubák P, Lebedev AA, Pannu NS, Steiner RA, Nicholls RA, et al. REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr D Biol Crystallogr.* 2011; 67(Pt 4): 355–67. doi: [10.1107/S0907444911001314](#) PMID: [21460454](#)
58. Emsley P, Cowtan K. Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr.* 2004; 60: 2126–32. PMID: [15572765](#)