Available online at www.sciencedirect.com
ScienceDirect
Protein Expression and Purification 58 (2008) 103–113
Protein Expression & Purification
www.elsevier.com/locate/yprep

# Optimized *in vitro* and *in vivo* expression of proteorhodopsin: A seven-transmembrane proton pump

Pontus Gourdon [a], Anna Alfredsson [b], Anders Pedersen [b,c], Erik Malmerberg [b], Maria Nyblom [a], Mikael Widell [a], Ronnie Berntsson [a,d], Jarone Pinhassi [e], Marc Braiman [f], Örjan Hansson [b], Nicklas Bonander [b], Göran Karlsson [c], Richard Neutze [b,*]

[a] *Department of Chemical and Biological Engineering, Molecular Biotechnology, Chalmers University of Technology, P.O. Box 462, SE-40530 Göteborg, Sweden*
[b] *Department of Chemistry, Biochemistry and Biophysics, Göteborg University, P.O. Box 462, SE-40530 Göteborg, Sweden*
[c] *Swedish NMR Centre, Göteborg University, P.O. Box 465, SE-40530 Göteborg, Sweden*
[d] *Department of Biochemistry, University of Groningen, 9747 AG Groningen, The Netherlands*
[e] *Marine Microbiology, Department of Biology and Environmental Sciences, University of Kalmar, SE-39182 Kalmar, Sweden*
[f] *Department of Chemistry, Syracuse University, Syracuse, New York 13244-4100, USA*

Received 9 August 2007, and in revised form 15 October 2007
Available online 20 November 2007

**Abstract**

Proteorhodopsin is an integral membrane light-harvesting proton pump that is found in bacteria distributed throughout global surface waters. Here, we present a protocol for functional *in vitro* production of pR using a commercial cell-free synthesis system yielding 1.0 mg purified protein per milliliter of cell lysate. We also present an optimized protocol for *in vivo* over-expression of pR in *Escherichia coli*, and a two-step purification yielding 5 mg of essentially pure functional protein per liter of culture. Both approaches are straightforward, rapid, and easily scalable. Thus either may facilitate the exploitation of pR for commercial biotechnological applications. Finally, the implications of some observations of the *in vitro* synthesis behavior, as well as preliminary results towards a structural determination of pR are discussed.


*Keywords:* Proteorhodopsin; Membrane protein; Cell-free; Functional expression

Proteorhodopsin[1] (pR) is a heptahelical, light-harvesting and growth-stimulating membrane bound proton pump that is found throughout global surface waters [1–3]. It is present in groups of bacterial organisms that are among the most prevalent on Earth (such as marine α-, β- and γ-proteobacteria and bacteroidetes) [1–6]. Considering the vast biomass of bacterioplankton, it has been speculated that pR-mediated phototrophism may have a significant impact on carbon and energy flux in the ocean [1,7]. Global analyses have also revealed a large sequence diversity of pR-like genes, reflecting their spectral tuning in accordance with the light availability at different depths in their natural environment [7–9] as well as implied variety in function [10].

The best characterized of all pR related proteins is bacteriorhodopsin (bR) from the archaeon *Halobacterium salinarium* [11]. These two proteins share many properties, including retinal based light-sensitivity, topology, and reaction mechanism. The photosensitivity, cyclicity, and high stability of bR [12], have made it a popular subject for studies for potential commercialization. bR has been used for developing optoelectronic and biotechnological applications, including: biomimetic ATP production for powering of molecular motors or direct solar-to-hydrogen production [13]; photovoltaics to directly convert solar

* Corresponding author. Fax: +46 31 773 3910.
*E-mail address:* richard.neutze@chem.gu.se (R. Neutze).
[1] *Abbreviations used:* pR, proteorhodopsin; bR, bacteriorhodopsin; CV, column volumes.

1046-5928/$ - see front matter © 2007 Elsevier Inc. All rights reserved.
doi:10.1016/j.pep.2007.10.017

radiation into electrical energy [13]; optical data storage [14]; and numerous additional areas [12,15]. However, the use of bR in commercial applications has been hampered by, for example, the limited availability of genetic systems for optimization of strain and protein properties in its native source [14,16]. Attempts to circumvent this difficulty by producing bR heterologously in *Escherichia coli* or cell-free systems have only given very low yields [17,18] unless fusion-constructs were used [19]. In contrast, pR holds great promise in addressing a number of these concerns and hence it is an attractive candidate for replacing bR in these efforts. pR is also very stable and unlike bR, which progressively loses its chromophore following solubilization [20], pR remains intact in detergent [21]. In addition, pR is more easily expressed functionally in *E. coli* than its archaeal homologue, albeit still at modest levels [22]. Therefore, exploiting the existing know-how surrounding technological applications of bR, along with the natural diversity of pRs, the use of pR for commercial exploitation could be expedited.

Here, we present a protocol for functional production of pR *in vitro* using a commercial cell-free over-production system. The yield, 1.0 mg soluble protein per mL cell-lysate, is superior to what has been previously obtained for any membrane protein using batch-based systems [23,24]. It is also 40 times higher than reported for bR from wheat germ extract [25] and recently from an *E. coli* based commercially available continuous exchange system [18]. An optimized procedure for functional *in vivo* expression in *E. coli* is also presented representing a significant improvement in yield of purified protein hitherto described [22]. This protocol can reproducibly provide ~5 mg/L cell culture of active protein purified to homogeneity. Both selected approaches, an easy-to-setup batch-based cell-free system and an *in vivo* expression with subsequent his-tag based purification are straightforward, rapid, and easily scalable, and may facilitate the development and use of pR. Finally, the implications of some observations of the *in vitro* synthesis behavior, as well as preliminary results towards a structural determination of pR are discussed.

## Materials and methods

### *Gene constructs, plasmid and strains*

The initial SAR86 pR gene originating from an uncultured marine γ-proteobacterium, had been cloned into an arabinose-inducible pBAD-TOPO vector (Invitrogen) including a N-terminal leader sequence and a C-terminal V5-epitope followed by a hexa-histidine tag (construct #1) [1]. The V5-epitope was removed and constructs with and without a factor Xa protease cleavage site (IEGR) between the histidine-tag and the gene were obtained by vector PCR. The MED134 pR gene from a Mediterranean Sea bacterium was cloned into the pBAD-TOPO vector as described previously [2]. These constructs were further subcloned between the NcoI and SmaI sites of the pIVEX2.3d (Roche) plasmid, a T7-promotor based system containing a C-terminal histidine-tag. In this process, two additional SAR86 constructs with shortened N-termini were created. In one, the N-terminal signal sequence (the first 17 residues) predicted by the program SignalP [26] was removed, and the other was adjusted to be similar to the length of pR from MED134. All constructs are summarized in Table 1. The *E. coli* strain UT5600 (Invitrogen) was used for *in vivo* expression.

### *Spectroscopic analysis of pR*

Correctly folded pR from both SAR86 and MED134 absorbs light around 535 nm, giving the protein a pink color and allowing concentration determination and purity estimation ($Abs_{280\ nm}/Abs_{535\ nm}$).

Table 1
Different constructs utilized in the work

| Construct | Vector | LS | SP | Gene | T | Tag |
|---|---|---|---|---|---|---|
| 1 | pBAD | X | X | cf-SAR | No | KGQLE + V5 + RTG + 6His |
| 2 | pBAD | X | X | cf-SAR | No | GP + 6His |
| 3 | pBAD | X | X | cf-SAR | No | GP + FXa + 6His |
| 4 | pBAD | X | X | cf-SAR | 11C | GP + 6His |
| 5 | pBAD | X | X | wt-MED | No | GP + 6His |
| 6 | pBAD | No | X | wt-MED | No | GP + 6His |
| 7 | pIVEX | — | X | cf-SAR | No | 6His |
| 8 | pIVEX | — | No | cf-SAR | 12N = M | 6His |
| 9 | pIVEX | — | No | cf-SAR | 17N = MG | 6His |
| 10 | pIVEX | — | X | wt-MED | No | 6His |

Vector indicates the used vector. pBAD = pBAD-TOPO and pIVEX = pIVEX2.3d. SP indicates whether the constructs have the predicted signal peptide (only accounts for SAR86 since MED134 lack such signal). LS indicates additional leader sequence to improve expression originating from the pBAD-TOPO vector. In construct 1–4 LS and SP are cleaved off upon membrane insertion, verified by N-terminal aminoacid sequencing. Gene defines whether the cysteine lacking SAR86 or the wild-type MED134 gene was employed. T specifies whether the gene was N- or C-terminally truncated and how many residues that were removed or replaced, e.g., 17N = MG, seven residues at the N-terminus are replaced by Met-Gly residues. Tag shows which C-terminal tags that were present.