CHAPTER ONE

Engineering *Escherichia coli* for Functional Expression of Membrane Proteins

Franz Y. Ho, Bert Poolman

Department of Biochemistry, Groningen Biomolecular Sciences and Biotechnology Institute & Zernike Institute for Advanced Materials, University of Groningen, Groningen, The Netherlands

1Corresponding author: e-mail address: b.poolman@rug.nl

Contents

1. Introduction 4
   1.1 Green fluorescent protein as a folding reporter 6
   1.2 Antibiotic resistance marker protein for selection 7
2. Preparation of Erythromycin-Sensitive E. coli Strain 8
3. Preparation of Expression Plasmid 10
   3.1 Construction of expression vector for target protein–GFP–ErmC fusion 10
   3.2 Expression test 11
4. Selecting Cells for Better Expression 14
5. Characterizing Evolved Strains 15
   5.1 Basic characterizations 15
   5.2 Plasmid curing 16
   5.3 Functional assays 16
6. Summary 17
Acknowledgments 18
References 18

Abstract

A major bottleneck in the characterization of membrane proteins is low yield of functional protein in recombinant expression. Microorganisms are widely used for recombinant protein production, because of ease of cultivation and high protein yield. However, the target proteins do not always obtain their native conformation and may end up in a nonfunctional state, in insoluble aggregates. For screening of functional protein, it is thus important to readily discriminate aggregated, mistargeted protein from globally well-folded, membrane-inserted protein. We developed a robust strategy for expression screening of functional proteins in bacteria, which is based on directed evolution. In this strategy, the C-terminus of the target membrane protein is tagged with two additional protein domains in tandem. The first one is green fluorescent protein (GFP), which functions as a reporter of the global folding state of the fusion protein. The other one is the erythromycin resistance protein (23S ribosomal RNA adenine N-6

Methods in Enzymology, Volume 556
ISSN 0076-6879
http://dx.doi.org/10.1016/bs.mie.2015.01.003

© 2015 Elsevier Inc.
All rights reserved.
methyltransferase, ErmC), which confers a means to select for enhanced expression. By gradually increasing the antibiotic concentration in the medium, we force the cells to evolve in a way that allows more functional target-GFP–ErmC to be expressed. The acquired genomic mutations can be generic or membrane protein specific. This strategy is readily adopted for the expression of any protein and ultimately yields a wealth of genomic data that may provide insight into the factors that limit the production of given classes or types of proteins.

1. INTRODUCTION

Integral membrane proteins contribute 15–30% of the open-reading frames found in the genomes of organisms from all domains of life (Bendtsen, Binnewies, Hallin, & Ussery, 2005; Wallin & von Heijne, 1998). However, our understanding of function–structure relationships of membrane proteins is low compared to water-soluble proteins (Granseth, Seppälä, Rapp, Daley, & Von Heijne, 2007). There are many difficulties in studying membrane proteins, and the challenge already starts with the expression of the genes and production of the proteins in a functional state. Recombinant expression in heterologous hosts is the most versatile strategy, but difficult to produce proteins often cause toxicity to the cells and end up themselves in insoluble aggregates (Bill et al., 2011). Expression conditions can be optimized but this is a laborious process (Francis & Page, 2010).

Large-scale production of proteins is commonly achieved using microorganisms, because of the ease of growth and many genetic and biochemical tools are available. Several factors determine the production level and functionality of the proteins, such as (i) the toxicity the proteins evoke to the host cells (Wagner et al., 2007), (ii) the posttranslational modifications required for structure and/or functionality (Grisshammer & Tate, 1995), (iii) the difference in codon usage between the recombinant gene and the expression host (Angov, Hillier, Kincaid, & Lyon, 2008), (iv) limitations of protein synthesis precursors such as tRNA and amino acids (Marreddy et al., 2010; Puri et al., 2014), (v) overloading of foldase and chaperone activities (Tate, Whiteley, & Betenbaugh, 1999), (vi) saturation of the membrane protein insertion machinery (Loll, 2003; Wagner, Bader, Drew, & de Gier, 2006), and (vii) uncoordinated protein biosynthesis kinetics (Bill et al., 2011; De Marco, 2013). Membrane proteins have a complex biogenesis pathway, requiring chaperones and machineries for membrane insertion, protein folding, and oligomerization, but often the limiting factor(s) leading to low or nonfunctional expression is not understood. Rather than
screening numbers of expression conditions, it is easier to improve the production by evolving the host.

In this method, we describe a simple procedure for engineering *Escherichia coli* by forcing the cells to produce more desired protein by fusing a folding indicator and antibiotic resistance marker to its C-terminus (Fig. 1A). The same strategy has been proven successful for the over-expression of membrane proteins in both the Gram-positive bacterium *Lactococcus lactis* and the Gram-negative bacterium *E. coli* (Gul, Linares, Ho, & Poolman, 2014; Linares, Geertsma, & Poolman, 2010).

**Figure 1** (A) Cartoon of fusion protein used for strain selection. The folding reporter GFP and the selection marker encoding the 23S ribosomal RNA adenine N-6-methyltransferase (ErmC) are fused in tandem at the C-terminus of the target membrane protein. N and C indicate the N- and C-terminus of the fusion protein. (B) Transport activity of LacSΔIIA fused with GFP in *E. coli* MC1061 cells. The expression of LacSΔIIA was controlled by the concentration of l-arabinose. (C) Expression of LacSΔIIA–GFP fusion was analyzed by SDS-PAGE. Upper panel: in-gel fluorescence; lower panel: anti-His immunoblot. White and black arrows indicate the folded and misfolded fusion protein, respectively. (D) Confocal microscopy images of *E. coli* MC1061 cells expressing LacSΔIIA–GFP to different levels; the l-arabinose concentration is indicated above the panels. Upper panel: close-up of cells to indicate the distribution of the fluorescence (scale bar, 2 μm). Lower panel: overview of the culture (scale bar, 10 μm). **Panels (B–D) are modified with permission from Geertsma, Groeneveld, Slotboom, and Poolman (2008); copyright (2008) by the National Academy of Sciences USA.**
1.1 Green fluorescent protein as a folding reporter

Protein folding can start right after the nascent chain is formed, but in case of membrane proteins, it is delayed as translation and membrane insertion are coupled and folding is completed after the last transmembrane segments have left the translocon (Luirink, von Heijne, Houben, & de Gier, 2005). In multidomain proteins, domains are tethered by either structured or disordered linkers and the rate of folding of individual domains can differ (Arviv & Levy, 2012; George & Heringa, 2002; Jappelli, Luzzago, Tataseo, Pernice, & Cesareni, 1992; Zhang & Ignatova, 2009). In the present method, green fluorescent protein (GFP) is used as a global folding reporter, which requires fusion of the fluorescent protein to the C-terminus of the target membrane protein. Thus, the synthesis of the membrane protein precedes that of the fluorescent protein. The folding of a membrane protein is typically faster than the folding and maturation of GFP, which takes about 30–90 min (Evdokimov et al., 2006; Reid & Flynn, 1997; Waldo, Standish, Berendzen, & Terwilliger, 1999). Therefore, when the target membrane protein is misfolded, it is likely to drag the not yet fully synthesized, folded, and/or matured GFP into an aggregated state. In contrast, when the target membrane protein is properly folded, the GFP β-barrel is formed and the chromophore will mature. Proper folding of the protein protects the chromophore from quenching by water dipoles, paramagnetic oxygen, or cis–trans isomerization (Tsien, 1998). The maturation process involves a series of covalent rearrangements of the amino acids that form the tripeptide chromophore (Ser/Thr65, Tyr66, and Gly67) within an α-helix that is buried inside the hollow β-barrel. When the chromophore matures, GFP becomes SDS resistant, and the protein migrates faster in SDS-PAA than the fully denatured polypeptide. The apparent difference in migration is about 10 kDa, and this difference is also observed when GFP is fused to another protein (Geertsma, Groeneveld, et al., 2008). Thus, when a membrane protein is well folded, the C-terminal GFP will fully mature, and the whole fusion protein migrates faster on SDS-PAA than the misfolded membrane protein–GFP fusion. By analyzing protein expression on immunoblots, one can easily discriminate the well-folded and misfolded proteins (Drew, Lerch, Kunji, Slotboom, & de Gier, 2006; Geertsma, Groeneveld, et al., 2008). Moreover, the fluorescence reports the absolute amount of folded fusion protein, which can be observed not only in gel but also in vivo (Drew et al., 2006; Linares et al., 2010).

We and others have observed that the in vivo activity of transport proteins correlates with the corresponding in-cell and in-gel GFP fluorescence of the
fusion proteins (Geertsma, Groeneveld, et al., 2008; Hibi et al., 2008; Schlegel et al., 2012). In brief, the glutamate transporter GltP from *E. coli* and lactose transporter LacSΔIIA from *Streptococcus thermophilus* have been fused to the N-terminus of GFP and expressed in *E. coli*. The measured transport rate of glutamate or lactose matches the GFP fluorescence intensity both in vivo and in SDS-PAA gels (Fig. 1), which is observed over a wide range of expression levels (Geertsma, Groeneveld, et al., 2008). This indicates that the fluorescence intensity of GFP can be used as a read-out of the in vivo activity of the target membrane protein. However, since the chromophore maturation requires aerial oxygen, the culture has to be well aerated when the protein is being expressed (see figure 1C in Linares et al., 2010). We emphasize that the fluorescence of GFP reports the global folding of the target molecule, but the membrane-inserted protein is not necessarily functional in every aspect. Factors like final folding steps, oligomerization, membrane lipid composition, and others can impact the observed activity (Grisshammer & Tate, 1995).

1.2 Antibiotic resistance marker protein for selection

One of the key features of our strain engineering system is the tagging of the target membrane protein–GFP fusion with an antibiotic resistance marker at the C-terminus. Thereby, cells expressing folded, membrane-inserted, fluorescent fusion protein have a growth advantage when cultivated in the presence of the antibiotic. We have tested several antibiotic resistance markers, such as chloramphenicol acetyltransferase, tetracycline efflux protein, aminoglycoside phosphotransferase, and 23S ribosomal RNA adenine N-6-methyltransferase (ErmC). Of these antibiotic resistance markers, the water-soluble ErmC has the most favorable properties and did not strongly affect the expression or activity of the membrane proteins to which it was fused. Contrary to ErmC, all the other antibiotic resistance markers function as oligomers, which is likely to impact the protein to which they are fused (Gul et al., 2014).

Proteome-wide topology studies in *E. coli* and *S. cerevisiae* indicate that ~80% of the integral membrane proteins have their C-termini at the cytosolic side (Cin topology) (Daley et al., 2005; Kim, Melén, Osterberg, & Von Heijne, 2006; Von Heijne, 2006). Our strategy of using GFP as folding reporter and ErmC to select for increased expression is not readily applicable to proteins with the C-terminus on the outside, since GFP is not fluorescent when localized in the periplasm of, e.g., *E. coli*. The localization of the
C-terminus of the target protein is readily predicted by various online protein topology prediction programs (Bernsel et al., 2008; Hennerdal & Elofsson, 2011). Thus, our selection system is based on expressing membrane protein fused with C-terminal inside GFP and ErmC.

2. PREPARATION OF ERYTHROMYCIN-SENSITIVE E. COLI STRAIN

Compared to Gram-positive bacteria such as L. lactis, enterobacteriaceae including E. coli are intrinsically resistant to erythromycin (Andremont, Gerbaud, & Courvalin, 1986). Therefore, we engineered the erythromycin sensitivity of the MC1061 strain by knocking out the acriflavine resistance B (acrB) gene. AcrB is a member of the resistance nodulation cell division (RND) superfamily and forms the AcrAB-TolC tripartite drug efflux system. It recognizes a wide variety of toxic compounds, including antibiotics, and transports these compounds out of the cell (Pos, 2009; Tal & Schuldiner, 2009). In our initial experiments, we used drug efflux mutants prepared in the BW25113 strain (Gul et al., 2014). However, E. coli MC1061 (araD139, Δ(araA-leu)7697, ΔlacX74, galU−, galK−, hsr−, hsm+, strA; Casadaban & Cohen, 1980) is generally a better host for membrane protein expression, and thus we inactivated acrB gene in this strain by targeted gene knock-out.

Gene inactivation can be achieved by various methods. We used the recombination-mediated genetic engineering method (recombineering) to create E. coli MC1061ΔacrB, which is simple and efficient for introducing site-specific mutations. It utilizes the bacteriophage λ homologous recombination proteins, called RED, to modify the genome of E. coli by using linear double-stranded (ds) DNA fragments of around 50 nucleotides that are homologous to the target locus (Datsenko & Wanner, 2000). Here, we describe the background and procedure of “recombineering” briefly; for the detailed protocol, we refer to Sawitzke et al. (2007) and Thomason, Sawitzke, Li, Costantino, and Court (2014).

Primer sequences (Table 1) for amplification of the DNA fragment, used for knocking out the acrB coding sequence, are designed according to Baba et al. (2006). Capital letters in the primer sequences are complementary to the sequences adjacent to the chromosomal acrB, while the small letters are sequences priming to pKD13. This plasmid carries a kanamycin resistance gene cassette flanked by Flp recognition targets (FRT) at both ends, which is recognized by yeast Flp recombinase (Datsenko & Wanner, 2000). Using
the primers listed in Table 1, the PCR-amplified DNA fragment for recombineering contains the kanamycin resistance marker gene flanked by FRT sites, with 50-nucleotide-long extensions corresponding to upstream and downstream of the chromosomal \textit{acrB} coding region. After replacing the \textit{acrB} gene by the FRT-flanked kanamycin resistance gene cassette, one can remove the marker by temporally expressing Flp to excise it from the genome, leaving only a short sequence of FRT as a scar. In principle, any antibiotic resistance gene without FRT sequences can be used for replacing the genomic sequence, if removal of the antibiotic marker is unnecessary (Sawitzke et al., 2007).

1. Transform pKD46 in \textit{E. coli} MC1061; pKD46 is a temperature-sensitive plasmid with coding sequences for \textit{\lambda} RED proteins.
2. Grow the transformed cells at 30 °C and shake the culture at 200 rpm. Induce the expression of \textit{\lambda} RED proteins by the addition of 0.2% (w/v) \textit{l}-arabinose.
3. Collect cells during mid-log phase and prepare electrocompetent cells.
4. Amplify the DNA fragment by PCR that corresponds to the \textit{acrB} locus.
5. Transform the cells with the amplified DNA fragment by electroporation.
6. Select colonies on LB agar supplemented with 25 \(\mu\)g/ml kanamycin and confirm the genome modification by colony PCR.
7. Remove pKD46 plasmid by growing the cells on nonselective medium (LB agar without antibiotics) at 37 °C.
8. Test the erythromycin sensitivity of the parent and \(\Delta\textit{acrB}\) strain; the recommended concentration range is 1–500 \(\mu\)g/ml. \textit{E. coli} MC1061 grows normal up to 100 \(\mu\)g/ml of erythromycin, whereas the \(\Delta\textit{acrB}\) strain is already sensitive at 5 \(\mu\)g/ml; other strains may have different erythromycin sensitivity. If the growth of \(\Delta\textit{acrB}\) cells in the presence of 5 \(\mu\)g/ml erythromycin is minimal after overnight incubation, the strain can be used for host evolution (Gul et al., 2014). Erythromycin sensitivity of MC1061 before and after \textit{acrB} inactivation is shown in Fig. 2.

<table>
<thead>
<tr>
<th>Table 1 Primers for inactivation of \textit{acrB} gene</th>
<th>Primer</th>
<th>Sequence (5’ to 3’)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Forward</td>
<td>TTACGCGGCTTAGTACACGTTATCAATGATGATCGACAGTATGctgtgctggagctgcc</td>
<td></td>
</tr>
<tr>
<td>Reverse</td>
<td>TCAGCCTGACAGTGCAAGTCTTAACCTAAACAGGACGTCAAGACATGctgaacatgagaattaa</td>
<td></td>
</tr>
</tbody>
</table>
3. PREPARATION OF EXPRESSION PLASMID

3.1 Construction of expression vector for target protein–GFP–ErmC fusion

There are wide varieties of plasmids, which can be used for protein expression in *E. coli*. In the current method, we prefer the araBAD expression system for production of membrane proteins. It uses the $P_{BAD}$ promoter, a tightly regulated bacterial promoter system, which gives minimal expression in the absence of inducer and toxicity by misfolded protein is readily controlled. Furthermore, the dynamic range of expression upon induction by L-arabinose can be more than three orders of magnitude (Guzman, Belin, Carson, & Beckwith, 1995).

Genes encoding EGFP and ErmC are first cloned at the 3′ end of the multiple-cloning site of the pBAD vector, after which the coding sequence of the target protein is inserted at the 5′ end in frame with the gene for the GFP–ErmC tandem. Another concern is the linker tethering the target membrane protein and the GFP. We tested the effect of the linker length and the folding reporter efficiency of the C-terminal fused GFP. We used a flexible (GGGS)$_n$ peptide sequence and observed no difference in whole-cell fluorescence from no linker to (GGGS)$_5$ in case the fusion protein itself has already up to 32 amino acids between the last transmembrane segment and the GFP (Linares et al., 2010). Furthermore, for purification and immunodetection of the proteins, we routinely add a 10× His tag at
the C-terminus of the fusion protein; and for functional and structural stud-
ies, we engineer a TEV protease recognition peptide sequence in the linker
region between the target protein and GFP (Geertsma, Groeneveld, et al.,
2008; Gul et al., 2014).

The expression plasmid can be converted into the ligation-independent
cloning (LIC) compatible system, which facilitates the cloning procedure.
For the details of designing and conversion of plasmids into LIC compatible
vectors, we refer to Geertsma & Poolman, 2007.

1. Design primers for amplifying codon optimized \textit{gfp} (GenBank:
\texttt{KF410615}; we use \texttt{gfp}+ that encodes EGFP; Scholz, Thiel, Hillen, &
Niederweis, 2000) and \textit{ermC} from \textit{Staphylococcus aureus} (GenBank:
\texttt{JF968525}; Weisblum, 1995). The 10× His tag coding sequence can
be prepared by annealing complementing oligonucleotides. They have
to be in-frame after construction.

2. Clone the coding sequences of \textit{gfp}, \textit{ermC}, and the 10× His in the
multiple-cloning site.

3. Clone the coding sequence of the target gene into the multiple-cloning
site. Make sure that all the sequences are in frame.

\section*{3.2 Expression test}

Before selecting cells with increased expression, it is important to first find
the optimal induction condition for the fusion protein. At least two param-
eters should be tested: the inducer concentration and the temperature of
expression. In general, the fraction of properly folded protein is higher at
low induction level and at low growth temperature, since excessive tran-
scription and translation usually overload the chaperone and translocation
systems. The optimal induction condition has to be determined experimen-
tally, but we routinely perform the experiments at 20 °C and \(\text{l-arabinose}
\) concentrations of 0.01–0.001% (w/v). The expression test is also useful after
the selection process, i.e., to find the best performing clones and to bench-
mark their expression performance.

\subsection*{3.2.1 Bacterial culture for protein expression and determination of
whole-cell fluorescence}

1. Transform \textit{E. coli} cells with pBAD bearing the gene for the target mem-
brane protein fused with GFP (with or without ErmC).

2. Pick a single colony and inoculate into LB medium supplemented with
100 μg/ml ampicillin (to maintain the pBAD vector), and grow the cul-
ture overnight at 37 °C with shaking at 200 rpm for adequate aeration.
3. Inoculate 1% (v/v) of overnight culture into fresh LB medium, supplemented with 100 μg/ml ampicillin, and grow the cells to $A_{600}$ of 0.5–0.6 (1 cm light path) at 37 °C.

4. When the cells reach $A_{600}$ of 0.5–0.6, induce protein expression by adding L-arabinose to the culture. Keep a fraction of cells for detection of un-induced expression. Initially, a broad range of induction conditions is tested, which can include varying the L-arabinose concentration, induction time, and growth temperature (Table 2).

5. At the end of the cultivation, measure the optical density of the cultures at 600 nm.

6. Collect cells by centrifugation at 5000 $\times g$ for 15 min, and wash the cells by resuspension in 0.5 $\times$ of culture volume of 50 mM potassium phosphate (KPi), pH 7 (wash buffer). Repeat the washing step and resuspend the cells in 0.2 $\times$ of the original culture volume of wash buffer.

7. After the second wash, remove the supernatant and resuspend cells in wash buffer to equal protein content (e.g., 3–5 mg/ml total protein, given to the fact that $A_{600}$ of 1 corresponds to about 0.3 mg/ml total protein).

8. Aliquot 100 μl of cell suspension (in triplicate) into wells of a 96-well plate, which is suitable for fluorescence measurements (black wells with transparent bottom). Use a fluorescence plate reader with suitable excitation and emission filters for GFP (488 nm excitation, 510 nm emission).

9. Measure the optical density of the cell suspension in the wells. These values are used for normalizing the fluorescence counts and subsequent gel-based analysis.

### Table 2: Parameters of expression optimization

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Recommendation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Expression</td>
<td>16–37 °C are commonly used e.g., 16, 20, 25, 30, and 37 °C</td>
</tr>
<tr>
<td>temperature</td>
<td></td>
</tr>
<tr>
<td>Time</td>
<td>1 h to overnight e.g., 1, 2, 4, 6 h, or overnight</td>
</tr>
<tr>
<td>Inducer</td>
<td>$10^{-4}$–$10^{-1}$% (w/v) L-arabinose for the araBAD system; strains that catabolize L-arabinose require higher concentrations of the inducer (Guzman et al., 1995; Horazdovsky &amp; Hogg, 1989) 20 μM to 0.4 mM IPTG for the lac promoter; strains that lack the lacY gene generally require a higher concentration of IPTG</td>
</tr>
</tbody>
</table>
3.2.2 *Gel-based analysis of the expressed fusion constructs*

Although whole-cell fluorescence indicates the amount of globally folded membrane protein–GFP–ErmC fusion, the fluorescence signal can be misleading when part of the fusion protein is degraded and soluble GFP is measured. Therefore, it is necessary to confirm the full length of the fusion proteins by SDS-PAGE. Moreover, folded and misfolded fusion proteins can be discriminated by SDS-PAGE followed by Western blotting (Geertsma, Groeneveld, et al., 2008). The choice of lysis buffer depends on the target proteins and subsequent purification steps; 50 mM potassium phosphate, pH 7, supplemented with 1 mM MgCl₂, 10% glycerol, 25 μg/ml DNaseI, and 1 mM PMSF is usually a good starting point.

1. Resuspend cells corresponding to 3–5 mg/ml in lysis buffer.

2. Break the cells by either heating with 0.1 μm glass beads (at 4 °C) or by probe sonication in an ice-water bath (control sonication power to prevent foaming and heating of the sample).

3. Mix 40 μl of cell lysate with 10 μl 5 × Laemmli buffer (Laemmli, 1970), incubate the mixture at 37–50 °C for 5 min (temperature should be lower than 65 °C, as otherwise the GFP will denature).

4. Separate 5–15 μl of each samples by SDS-PAGE; samples should contain 20–100 μg of total cell protein.

5. After electrophoresis, rinse the gel with water and capture its fluorescence on a proper imaging system, such as ImageAnalyzer LAS from FujiFilm or Typhoon from GE Healthcare. In the fluorescence image, only properly matured GFP and thus overall folded fusion proteins are detected.

6. Use the same gel, transfer the separated proteins onto PVDF membrane by semidry electrotransfer (Bjerrum & Schäfer-Nielsen, 1986; Gershoni, 1988).

7. After the transfer, visualize the protein expression by Western blotting, using an anti-His antibody and an appropriate secondary antibody (Gershoni, 1988).

8. Both folded and misfolded proteins are visible on the Western blot. Different culture and induction conditions can be compared at this stage, and optimal starting conditions for expression screening are chosen. An example of an expression test is shown in Fig. 3. In the example, the cAMP receptor 1 (cAR1) from *Dictyostelium* was expressed in *E. coli* BW25113B and NG3 at different inducer (l-arabinose) concentrations. The NG3 strain was obtained by selecting for enhanced production of the glutamate transporter GltP in the BW25113B strain as described in this method. It carries...
genomic mutations in the \textit{hns}, \textit{ung}, and \textit{cpxA} genes when compared to the parental BW25113B strain (Gul et al., 2014).

### 4. SELECTING CELLS FOR BETTER EXPRESSION

Improved expression strains are selected by directed evolution. The target protein is fused to GFP and ErmC. If the target protein is misfolded, it will drag GFP and ErmC in a nonfunctional state and growth will be

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{example_expression_test.png}
\caption{Example of expression test. (A) \textit{In vivo} cell fluorescence of \textit{E. coli} BW25113B cells (open square) and its evolved derivative NG3 (closed circle) expressing the \textit{Dictyostelium discoideum} cAMP receptor 1–GFP–ErmC fusion protein at different L-arabinose concentrations. The fluorescence intensity is normalized by cell density and error bars show the standard deviation (SD) of three independent measurements. (B) In-gel GFP fluorescence and immunoblots detected by anti-His tag antibody of the same gels are shown. Black and white arrows indicate the positions of the non-fluorescent (misfolded) and fluorescent (folded) protein species, respectively.}
\end{figure}
limited by the presence of erythromycin in the medium. Evolved cells that produce more folded protein will be more resistant to erythromycin and outcompete the parental strain. By culturing cells in the presence of inducer to trigger protein expression and gradually increasing the erythromycin concentration, one selects for better performing cells in terms of protein expression. We use GFP to identify among the erythromycin-resistant strains those that produce the highest fluorescence, and thus fusion protein that is inserted into the membrane rather than misfolded protein with some ErmC activity.

1. Pick a single colony and inoculate into LB medium supplemented with 100 μg/ml ampicillin (or appropriate antibiotics to maintain the plasmid). Grow overnight at 37 °C with vigorous agitation (200 rpm).

2. Dilute the overnight culture in 1–100 ratio. Grow at 37 °C to $A_{600} \sim 0.4$, and then lower the cultivation temperature to 25 °C; we occasionally perform the selection at 17 °C.

3. Induce protein expression at $A_{600}$ 0.5–0.6, by adding α-arabinose to a final concentration of 0.01% (w/v).

4. After 2 h, inoculate fresh LB-ampicillin medium with the induced culture at 2% (v/v); keep α-arabinose at 0.01% for protein expression, and add erythromycin at 5 μg/ml for selection. Continue the cultivation for 48 h.

5. Transfer 2% of the culture into fresh medium with the same supplements but an elevated concentration of erythromycin, e.g., 10 μg/ml, and continue growth for 48 h.

6. Repeat step 5 and increase the erythromycin concentration consecutively from 10 μg/ml to 20, 50, 100, and 200 μg/ml.

7. Optionally, one can continue the subculturing at 200 μg/ml erythromycin by repeating step 5 for one to five times.

8. At the end of the selection, the cells are plated onto LB agar supplemented with 100 μg/ml ampicillin, 0.01% (w/v) α-arabinose, and the highest concentration of erythromycin reached.

9. Pick the most fluorescent colonies from the plate for further analysis.

5. CHARACTERIZING EVOLVED STRAINS

5.1 Basic characterizations

5.1.1 Plasmid copy number, DNA sequencing, and transcript levels
Recombinant protein yield can be increased by altered copy number or mutations within the plasmid. We estimate changes in plasmid copy number from gels or more precisely by quantitative PCR (Skulj et al., 2008). So far,
we have never found changes in plasmid copy number or mutations in the plasmid. Whole-genome sequencing using next generation sequencing technologies (ultradeep sequencing) revealed multiple mutations in the genomes of *E. coli* and *L. lactis* (Gul et al., 2014; Linares et al., 2010). The transcription level of the target protein can be determined by qRT-PCR (Báez-Viveros et al., 2007).

### 5.2 Plasmid curing

The evolved *E. coli* cells are subsequently cured from the expression plasmid, which is done by subculturing cells at 5–7 °C above the optimal growth temperature (42–44 °C) (Trevors, 1986). Depending on the properties of the plasmid (copy number, stability), the cells may have to be subcultured up to 50 times. Cells which are sensitive to the marker antibiotic are isolated.

1. Inoculate evolved strains in LB medium without the plasmid marker antibiotic; i.e., ampicillin in case of pBAD vectors.
2. Grow the cells at 37 °C with vigorous agitation until late log phase.
3. Transfer the cells to fresh LB medium at a 5–10% (v/v) inoculum, continue cultivation nonselectively at 42–44 °C with vigorous agitation.
4. Repeat step 3 and screen colonies until the cells are sensitive to ampicillin.
5. Purify isolated colonies and characterize the strains, following transformation with fresh plasmid. The evolved expression hosts can now be tested for the expression of different target proteins.

### 5.3 Functional assays

The function of the expressed protein can sometimes already be tested *in vivo*. For example, it is possible to test the function of transport proteins by measuring the import or export of radiolabeled substrate (Geertsma, Groeneveld, et al., 2008). Similarly, it is possible to probe binding of ligands to a receptor *in vivo* when high-affinity ligands are available (Brodersen, Honoré, Pedersen, & Klotz, 1988; Detmers et al., 2000; Silhavy, Szmelcman, Boos, & Schwartz, 1975). Ultimately, one will have to purify the protein and determine if the protein isolated from the evolved host has genuine activity. For membrane transport proteins, we have previously described various simple methods to test their activity either in membrane vesicles or proteoliposomes (Geertsma, Nik Mahmood, et al., 2008; Mulligan et al., 2009; Poolman et al., 2005).
6. SUMMARY

In this chapter, we describe a straightforward procedure for increasing the level of well-folded membrane protein by directed evolution of bacteria. The method involves selection of E. coli strains that tolerate a higher concentration of erythromycin when the target membrane protein, fused to...
GFP and ErmC, is produced at elevated level. The evolved strains can have multiple mutations in their genome, which differ depending on the target protein. We emphasize that the method is not restricted to overexpressing membrane proteins but is applicable to any difficult to produce protein. We also present protocols for the basic characterization of the evolved strains. The complete procedure is summarized schematically in Fig. 4.

ACKNOWLEDGMENTS

We thank Prof. Shimon Schuldiner for the gift of E. coli strain BW25113 and the isogenic drug efflux-deficient derivatives. For developing the system, we thank Dr. Eric Geertsma, Dr. Daniel Linares, and Dr. Nadia Gul for their pioneering work. We also thank the Netherlands Science Foundation (NWO; TOP-GO program grant, project number 700.10.53) and the Michael J. Fox foundation (Rapid Response Innovation Awards, 2014) for financial support.

REFERENCES


