Part One
Expression Systems

1
Bacterial Systems

James Samuelson

1.1
Introduction

The study of membrane protein structure and function is limited by various challenges. In native cells, membrane protein copy number is often very low, so the study of individual proteins is often not feasible. Alternatively, overexpression of these hydrophobic molecules in heterologous hosts is not a routine endeavor as it is for many water-soluble proteins. Most modern bacterial expression systems have been engineered for maximal output of recombinant protein. This characteristic is ideal for well-behaved soluble proteins, but less desirable when the target protein normally resides within a lipid environment. A compounding problem in the study of membrane proteins is that the isolated target protein may exhibit polydispersity, meaning that diverse oligomeric complexes can spontaneously accumulate. This latter concern may be influenced by the expression method, but primarily depends on the detergent/lipid and buffer used for solubilization. This chapter highlights preferred strategies for membrane protein expression in bacteria that will increase the likelihood of isolating adequate amounts of homogenous target protein. Many sections will also detail the features of expression strains that are relevant to the yield and quality of expressed protein.

In this chapter, the term membrane protein will generally be used to represent α-helical membrane proteins that reside within a phospholipid bilayer environment of either eukaryotic or prokaryotic cells. Such integral membrane proteins are the most difficult to manipulate since each contains hydrophobic transmembrane (TM) regions as well as hydrophilic extramembrane regions or domains. In the case of single-spanning membrane proteins, often the catalytic domain is a water-soluble entity that may be studied by expression of a ΔTM variant. However, multispanning membrane proteins such as ion channels must be expressed without gross deletions of hydrophobic residues.

Membrane proteins with β-barrel structure such as those found in the Gram-negative bacterial outer membrane or the mitochondrial outer membrane are typically expressed at high levels as inclusion bodies within the *Escherichia coli* cytoplasm. Isolation and washing of these inclusion bodies often leads to a

Production of Membrane Proteins: Strategies for Expression and Isolation, First Edition.
Edited by Anne Skaja Robinson.

relatively pure sample of recombinant protein and the literature contains many examples of refolding of β-barrel proteins, such as Omp proteins from *E. coli* [1]. In contrast, refolding of α-helical integral membrane protein is quite a difficult challenge, although some successes have been reported [2–4]. The default method of expressing α-helical membrane proteins should be to direct them to the membrane fraction of the host cell and to perform purification procedures beginning with isolation of the cellular membrane fraction.

1.2 Understanding the Problem

Each recombinant membrane protein clone should be assumed to be "toxic" to the host cell. This is particularly true when bacterial hosts are employed. It is well established that uncontrolled expression of most membrane proteins in *E. coli* will lead to induction of cellular stress responses and occasionally cell death. In some cases, the plasmid transformation step may fail because the transformed cell cannot recover due to the uncontrolled expression of membrane protein. Therefore, the first bit of advice in designing expression clones is to use a vector that propagates at 40 copies or less per cell (pMB1+*rop*, oriV, p15A, pSC101 replication origins). Accordingly, a vector with a pUC-derived origin should be avoided. Secondly, the promoter driving protein expression should be controllable (inducible). Much of this chapter is allocated to describing appropriate host/vector/promoter combinations (see Table 1.1 for a summary).

In bacteria, passage through the inner membrane Sec translocase [5] is recognized as the primary bottleneck during the overexpression of recombinant membrane protein. Yet, many other factors may contribute to a limited expression yield. There are reports of Sec-independent membrane translocation, but true host protein-independent membrane assembly by a heterologous protein has not been clearly substantiated in the literature. For example, membrane assembly of Mistic fusion proteins [6] may be initiated by the affinity of the Mistic protein for the cytoplasmic face of the *E. coli* inner membrane; however, proper membrane assembly of the fused protein of interest must still require assistance from the Sec translocase when large extracellular hydrophilic domains need to be translocated across the inner membrane.

Our lab has investigated several possible modes of Sec-independent membrane assembly without arriving at any evidence that a heterologous integral membrane protein can bypass the Sec translocase (unpublished data). Furthermore, we have attempted to increase the efficiency of membrane integration by overexpressing the endogenous YidC protein that is thought to aid the Sec translocase or act independently as a membrane insertase [7]. We specifically chose to study the effect of YidC on the membrane integration of phage M13 p8 fusion proteins, as p8 protein by itself requires YidC for inner membrane assembly [8]. To our surprise, a 10-fold increased level of YidC had no effect on the membrane translocation of p8-derived fusion proteins containing a C-terminal PhoA domain as a

reporter. One conclusion of this experiment is that the activity of SecA ATPase may be the limiting factor for the translocation of the large hydrophilic PhoA domain. Recently we determined that the p8 fusion partner (p8CBDek described in Luo *et al*. 2009 [9]) utilizes the cotranslational signal recognition particle (SRP) pathway [10–12], the route traveled by most endogenous membrane proteins. During cotranslational membrane protein assembly, there is less opportunity for hydrophobic amino acid segments to aggregate after emerging from the ribosome tunnel. Perhaps the limiting factor in p8 fusion protein expression and the overexpression of most membrane proteins is simply the rate of protein translation (or efficiency of translation initiation) at the ribosome. With this thought in mind, we tested various ribosomal binding sites (RBSs) and found a distinct difference in the efficiency of p8CBDek-mediated polytopic membrane protein assembly. Strikingly, the clone containing the much weaker RBS (AGGACGGCCGGatg) produced a greater level of protein per cell after a 20-h expression period at 20 °C. In contrast, the stronger RBS provided more protein per cell in the first stage of expression, but also resulted in jamming of the translocation pathway and cessation of culture growth. Thus, the take-home message from our recent work is to express recombinant membrane proteins "in moderation." This advice may seem obvious, but many expression systems do not allow for careful control of expression. The solution of genetically engineering the appropriate RBS for the protein of interest may not be a preferred method of optimization. Instead, a much simpler solution for expression optimization is to employ a promoter that allows fine control of the level of mRNA encoding the membrane protein of interest.

1.3 Vector/Promoter Types

The most-studied bacterial promoters are those controlling operons for sugar metabolism (*lacZYA*, *araBAD*, *rhaBAD*). Many variants of the *lac* promoter have been isolated but all suffer to some degree from the inability to completely shut off expression with the LacI repressor protein. The wild-type *lac* promoter is a good choice for membrane protein expression due to its moderate strength. However, very few expression vectors encode the unmodified *lac* promoter. Vectors pUC18/pUC19 carry a simple *lac* promoter, but again pUC derivatives are not good choices due to high copy number and overproduction of β-lactamase (AmpR) that enables the growth of cells lacking plasmid. Vectors utilizing modified *lac* promoters are highlighted in Table 1.2. The *lacUV5* promoter has two mutations within the –10 region of the *lac* promoter. In addition, a mutation is present at –66 within the catabolite gene activator protein (CAP) binding site. These mutations increase the promoter strength relative to the wild-type *lac* promoter and expression from *lacUV5* is less subject to catabolite repression [13]. The *tac* promoter was first described by deBoer *et al*. [14–15]. This strong promoter is a hybrid of the –10 region of the *lacUV5* promoter and the –35 region of the *trp* promoter. Amann *et al*. reported that the *tac* promoter is at least 5 times more efficient than the

Table 1.1 Recommended *E. coli* strains for membrane protein expression.

Strain	Source	Distinguishing features	Growth/expression guidelines	Compatible expression vectors
BL21(DE3)	NEB; Novagen; Invitrogen; Stratagene; Genlantis; Lucigen	chromosomal DE3 prophage expresses T7 RNA polymerase under control of the *lacUV5* promoter; BL21 derivatives lack Lon and OmpT proteases, which may stabilize expression of some recombinant protein	exhibits significant basal T7 expression, thus addition of 1% glucose to the growth medium is recommended for toxic clones	pET or T7 vectors with *lacI* gene and *T7–lac* promoter
BL21(DE3) pLysS	Novagen; Invitrogen; Stratagene; Genlantis; Lucigen	same as above, plus the pLysS plasmid produces wild-type T7 lysozyme to reduce basal T7 expression of the gene of interest; pLysS is compatible with plasmids containing the ColE1 or pMB1 origin (most pET vectors)	Cam 34μg/ml to maintain pLysS	pET or T7 vectors with *lacI* gene and *T7–lac* promoter
Lemo21(DE3)	NEB; Xbrane Bioscience	pLemo plasmid produces amidase-negative T7 lysozyme (*lysY*) from a tunable promoter (*Prha*); pLemo is compatible with plasmids containing the ColE1 or pMB1 origin (most pET vectors)	Cam 34μg/ml to maintain pLemo; expression trials typically benefit from rhamnose addition	pET or T7 vectors with *lacI* gene and *T7–lac* promoter
T7 Express LysY/I^q	NEB	T7 RNA polymerase gene encoded within the *lac* operon; combination of *lysY* and $lacI^q$ control offers strict control of basal T7 expression	incompatible with CamR expression vectors	pET or other T7 vector, presence of *lacI* gene is less important
BL21-AI	Invitrogen	T7 RNA polymerase gene is controlled by the *araBAD* promoter; if using a pET vector, IPTG is also required for induction to titrate LacI repressor away from the *T7–lac* promoter on the vector	T7 gene 1 is induced by L-arabinose; 0.2% glucose will repress	recommended for use with Invitrogen pDEST vectors
BL21	Various companies	protein expression from non-T7 promoters metabolizes arabinose, but still suitable for membrane expression from *ParaBAD*	none	nearly all non-T7 vectors
NEB Express I^q	NEB	BL21 derivative, same features as above except additional control of IPTG-inducible promoters suitable for membrane expression from *ParaBAD*	incompatible with CamR expression vectors	non-T7 vectors

TOP10; LMG194	Invitrogen	K-12 strains that do not metabolize L-arabinose may provide slight improvement when expressing membrane protein directly from *ParaBAD*	0.2% glucose will repress expression of toxic proteins	pBAD vectors
EXP strains	UCLA, J. Bowie	these are TOP10 mutants selected for improved expression of recombinant membrane protein; DE3 prophage has been added to original isolates	none	nearly all vectors with DE3 strains
Single Step KRX	Promega	T7 RNA polymerase gene is controlled by *PrhaBAD*; induce with 0.1% L-rhamnose; if using pET, pF1A or pF1K vectors, IPTG is also required for induction to titrate LacI repressor away from the *T7–lac* promoter on the vector	0.4% glucose in LB plates or starter culture to stabilize toxic clones	pET vectors T7 Flexi vectors (Promega)
C41(DE3); C43(DE3)	Lucigen	BL21(DE3) derivatives with much lower levels of T7 RNA polymerase under noninducing and inducing conditions	glucose addition is not necessary as with BL21(DE3)	pET or other T7 vector, presence of *lacI* gene is less important
Tuner; Tuner(DE3)	Novagen	*lacZY* derivatives of BL21; the *lac* permease mutation (*lacY1*) allows uniform entry of IPTG into all cells in the population, which produces a concentration-dependent level of induction		pET or T7 vectors with *lacI* gene and *T7–lac* promoter
Tuner (DE3) pLysS	Novagen	expresses T7 lysozyme to control T7 expression in addition to the *lac* permease mutation		pET or T7 vectors with *lacI* gene and *T7–lac* promoter
Rosetta; Rosetta2	Novagen	Rosetta host strains are BL21 *lacZY* (Tuner) derivatives designed to enhance the expression of proteins that contain codons rarely used in *E. coli*; these strains express tRNAs for rare codons on a compatible CamR plasmid	Cam 34µg/ml	pET or T7 vectors with *lacI* gene and *T7–lac* promoter
Rosetta pLysS; Rosetta2 pLysS	Novagen	in the pLysS strains, the rare tRNA genes and T7 lysozyme gene are carried by the same plasmid	Cam 34µg/ml	pET or T7 vectors with *lacI* gene and *T7–lac* promoter
BL21 CodonPlus; (DE3)-RIPL	Stratagene	RIPL indicates that the cells carry extra copies of the *argU*, *ileW*, *leuY*, and *proL* tRNA genes for increased recognition of the AGA/AGG (Arg), AUA (Ile), CUA (Leu), and CCC (Pro) codons	Cam 50µg/ml	pET or other T7 vectors

Table 1.2 Common vectors/promoters/types of regulation (for more options, a comprehensive vector database is maintained by the EMBL Protein Expression and Purification Core Facility: http://www.pepcore.embl.de/strains_vectors/vectors/bacterial_expression.html).

Promoter	RNA polymerase	Inducer/relative strength	Distinguishing features	Vectors	Source
lac	*E. coli*	IPTG/+	repressed by glucose; controlled by LacI	pUC; NR	NEB
lacUV5	*E. coli*	IPTG/++	repressed by glucose; controlled by LacI		
tac	*E. coli*	IPTG/+++	CAP site absent, not affected by glucose well controlled by LacI	pMAL; p8CBDek; pKK223-3	NEB
trc	*E. coli*	IPTG/+++	CAP site absent, not affected by glucose well controlled by LacI	pTrc99a; pPROEX; pKK233-2	
T5	*E. coli*	IPTG/+++	with pQE, need to supply expression of LacI repressor (multicopy pREP4 or a strain carrying $lacI^q$ gene)	pQE	Qiagen
araBAD	*E. coli*	L-arabinose/++	repressed by 0.2% glucose; tunable from 0.001–0.2% L-arabinose	pBAD	Invitrogen
rhaBAD	*E. coli*	L-rhamnose/++	repressed by 0.2% glucose; very tunable when vector carries *rhaRS* genes; inducer range = 10–2000 μM L-rhamnose	pRHA-67	Xbrane Bioscience
tetA	*E. coli*	aTc/++	pASK75 carries *tetR*; DH5αZ1 strain or other *tetR* strain is recommended	pASK75; pZ series	A. Skerra; H. Bujard/Expressys
T7 plain	T7	IPTG/+++	very high basal expression; controlled by T7 lysozyme	pET3, 9, 14, 17, 20, 23	Novagen
T7–lac	T7	IPTG/+++	high basal expression; controlled by T7 lysozyme	pET21, etc.; Various	Novagen

+, ++, +++ relative strength of induction.
NR, not recommended for membrane protein expression.

lacUV5 promoter [16]. The *trc* promoter is equivalent to the *tac* promoter since the 1-bp difference in spacing between the −35 and −10 consensus sequences does not affect promoter strength [17]. Note that the *tac* and *trc* promoters are not subject to catabolite repression as the CAP binding site is missing. *Ptac* and *Ptrc* systems are generally well controlled by LacI repression. When employing any type of modified *lac* promoter, LacI should be overexpressed from a *lacI* or *lacI*^{q} gene carried by the expression vector. Also, isopropyl-β-D-thiogalactopyranoside (IPTG) induction should be tested in the low range (e.g., 0, 10, 100 versus 400 μM). The *lacI*^{q} mutant was reported by Calos in 1978 and this mutation is simply an "up" promoter mutation resulting in a 10-fold enhancement of LacI repressor expression [18].

The pQE vectors from Qiagen utilize the phage T5 promoter that is controlled by two *lac* operator sequences. The T5 promoter is recognized by the *E. coli* RNA polymerase and induction is accomplished by IPTG addition to release the Lac repressor from the dual operator sequence. Since pQE vectors do not carry the *lacI* gene, the host strain must supply an excess of Lac repressor. Two options exist for LacI supplementation: copropagation of multicopy pREP4 (QIAexpress manual) or use of a strain that carries the *lacI*^{q} gene. Many K-12 strains (e.g., JM109) carry the *lacI*^{q} gene, but few B strains offer LacI overexpression. One recommendation is NEB Express I^q, which is a BL21 derivative that carries a miniF-*lacI*^{q} which does not require antibiotic selection (Table 1.1).

Guzman *et al.* characterized the *araBAD* promoter in exquisite detail in 1995 [19], and the resulting the pBAD vector series offers many options for gene cloning and expression using L-arabinose induction. Note that some pBAD vectors do not encode RBS sites, so the gene insert must contain an appropriate translation initiation sequence. When glucose is added to the outgrowth media, expression from *araBAD* is essentially shut off (Table 1 in Guzman *et al.* [19]). For many years, the *araBAD* system was a first choice for tightly regulated expression, as protein output appears to correlate very well with the amount of inducer (Figure 4 in Guzman *et al.* [19]) However, careful studies of the *araBAD* promoter by Siegele *et al.* [20] and Giacalone *et al.* [21] both agreed that at subsaturating levels of L-arabinose, protein expression cultures contain a mixed population with only some of the cells expressing protein. In addition, the potential for protein overexpression is generally lower when a pBAD vector is compared to T7-mediated expression from a pET construct.

A more recently characterized sugar promoter is derived from the rhamnose operon. The *rhaBAD* promoter is induced by L-rhamnose. When protein is expressed directly from *PrhaBAD*, the expression level within each cell falls within a range that correlates very well with the amount of inducer added to the culture [21]. In fact, Giacalone *et al.* presents convincing data that the pRHA-67 vector is more tunable and is capable of higher output than a high-copy vector containing the *araBAD* promoter. The pRHA-67 vector is commercially available from Xbrane Bioscience. Data presented by Haldimann *et al.* [22] indicates that expression from the *rhaBAD* promoter is very tightly regulated, yet this system also offers the potential for 5800-fold induction when glycerol is used as the primary carbon source.

The tetracycline inducible system is also very tightly regulated. Although we do not have experience with this system, Skerra *et al.* [23] reports that the pASK75 vector utilizing the *tetA* promoter/operator and encoding the cognate repressor gene (*tetR*) displays tightly regulated and high-level expression of heterologous protein in several *E. coli* K-12 and B strains. Induction is accomplished with low concentrations of anhydrotetracycline (aTc) and the induction potential is comparable to that of the *lacUV5* promoter. Lutz and Bujard [24] described additional aTc-inducible vectors that make use of the engineered $P_{LtetO-1}$ promoter, which is also controlled by TetR repression. The pZ vectors offer low, medium or high level expression from $P_{LtetO-1}$ corresponding to the copy number dictated by the pSC101, p15A, or ColE1 origins of replication, respectively. The aTc-inducible pZ vectors require expression in strains overexpressing TetR (e.g., DH5αZ1). The pSC101 version offers the most strictly regulated expression with an induction/repression ratio of 5000.

1.4 T7 Expression System

Over the last 20 years, the most common vector series for bacterial protein expression is the pET series (*p*lasmid for *e*xpression by *T*7 RNA polymerase). The T7 expression system was developed primarily by F. William Studier and colleagues at Brookhaven National Laboratory [25]. The T7 system is best recognized for the capacity to generate a high level of recombinant protein as the phage T7 RNA polymerase is very active and also very selective for phage T7 promoters (e.g., ϕ10). Therefore, T7 transcription within a bacterial cell can be specifically directed at a single promoter within the pET vector carrying the gene of interest. In most T7 expression strains, the chromosomal DE3 prophage carries the T7 RNA polymerase gene (T7 gene 1), which is expressed from the *lacUV5* promoter. Since this promoter is not completely shut off by LacI, some molecules of T7 RNA polymerase are continuously expressed and are able to make considerable amounts of target mRNA in the absence of IPTG. With respect to membrane protein expression, this is an unacceptable situation. An early partial solution to this problem was to include the *lacI* repressor gene on the multicopy pET vector. Thus, LacI repressor protein is produced in large excess relative to its operator binding site present in the *lacUV5* promoter driving T7 gene 1. Another partial solution to leaky T7 expression was the introduction of the *T7–lac* hybrid promoter to the pET vector series. In vectors beginning with pET-10, the *lac* operator sequence overlaps the T7 promoter so that excess LacI is able to inhibit T7-mediated transcription of the target gene. However, even with this improvement uninduced expression is observed in many experiments employing BL21(DE3). Uninduced expression of even mildly toxic gene products may be lethal to BL21(DE3) at the transformation step.

A very effective means to control T7 expression is to coexpress T7 lysozyme, the natural inhibitor of T7 RNA polymerase. Until recently, three types of lysozyme

strains were available and all were designed to produce lysozyme at a relatively constant level: pLysS and pLysE express wild-type T7 lysozyme from a low-copy plasmid, and in NEB *lysY* strains, an amidase-negative variant of T7 lysozyme (K128Y) is expressed from a single-copy miniF plasmid. The K128Y variant does not degrade the peptidoglycan layer of the *E. coli* cell wall [26] and, accordingly, *lysY* results in greater overall culture stability when membrane proteins are targeted to the cell envelope. In constitutive lysozyme systems, the level of lysozyme is sufficient to sequester the basal level of T7 RNA polymerase by a 1 : 1 protein interaction. When IPTG is added, the level of T7 RNA polymerase is present in large excess and target protein expression proceeds. If a membrane protein expression plasmid does not yield transformants when using BL21(DE3) or other basic T7 expression strains, the first response should be to test transformation into a lysozyme strain. LysY or pLysS strains may yield normal colonies and express the protein of interest at moderate to high levels. Finally, it should be noted that the choice of lysY or pLysS should take into account downstream processing of cells. Strains expressing active lysozyme often lyse spontaneously upon one freeze–thaw cycle and the resulting cell pellets may be difficult to process.

1.5 Tunable T7 Expression Systems

A recent development in T7 expression is the ability to tune the level of expression. Tunable expression provides a means for optimizing the traffic flow into the membrane translocation pathway. Four commercial strains promote this feature: Tuner™ from Novagen, BL21-AI from Invitrogen, the KRX strain from Promega, and the Lemo21(DE3) strain from New England Biolabs.

- The Tuner strain does not express lac permease (*lacY*) and this allows more uniform uptake of IPTG. However, T7 expression in Tuner strains may still be too robust for membrane protein expression unless the plasmid has a *T7–lac* promoter and lysozyme is coexpressed.
- BL21-AI offers greater potential for expressing toxic gene products as the *araBAD* promoter controls the expression of the T7 RNA polymerase. The associated pDEST expression vectors contain a plain T7 promoter (no *lac* operator site).
- In the Single Step KRX strain, T7 gene 1 expression is controlled by the *rhaBAD* promoter, so greater potential for toxic protein expression is expected. This K-12 strain has been designed for cloning and protein expression.
- The Lemo21(DE3) strain [27] is a tunable T7 expression strain derived from BL21(DE3). Lemo means “less is more” as often less expression results in more protein produced in the desired form. The Lemo strain is distinct from other T7 host strains since the *fraction of functionally active* T7 RNA polymerase is regulated by varying the level of T7 lysozyme (lysY). Fine-tuning is possible

since the LysY inhibitor protein is expressed from the L-rhamnose inducible promoter. The wide-ranging expression potential of Lemo21(DE3) is sampled to find the appropriate level for each target membrane protein. When using Lemo21(DE3), expression media should lack glucose since this carbon source affects lysozyme expression from *PrhaBAD*.

1.6 Other Useful Membrane Protein Expression Strains

C41(DE3) and C43(DE3) have been employed as membrane protein expression strains since their isolation from parent strain BL21(DE3) in 1996 [28]. Recently, Wagner *et al*. [26] reported that these two strains carry mutations within the promoter driving expression of the T7 RNA polymerase. Therefore, the characteristic robust T7 expression of DE3 strains is attenuated in C41(DE3) and C43(DE3), and this accounts for the advantage observed in the expression of some toxic proteins.

More recently, the TOP10 strain was subjected to a genetic selection procedure that produced several mutant strains exhibiting improved expression of heterologous membrane protein. This work was completed by Elizabeth Massey-Gendel *et al*. under the direction of James Bowie at the University of California at Los Angeles (UCLA). Target membrane proteins were expressed with a C-terminal cytoplasmic fusion to mouse dihydrofolate reductase (DHFR) (providing resistance to trimethoprim) or to a kanamycin resistance protein. A positive hit in the selection was obtained when a mutant strain was capable of expressing both fusion proteins at a level sufficient to provide resistance to both drugs. Five of the selected strains have been characterized in some detail [29] and the genomes of two such strains have been sequenced. At the January 2010 Peptalk meeting in San Diego, Professor Bowie reported that his lab is currently investigating the relevance of the mutations identified in the TOP10 derivatives designated as EXP-Rv1337-1 and EXP-Rv1337-5. The results of this investigation are widely anticipated. The DE3 prophage has been added to the EXP strains so that T7 expression is possible.

The Single Protein Production System (SPP System™) was developed by Masayori Inouye [30] and is marketed by Takara Bio. This is a two-vector system suitable for use in most *E. coli* strains. The target protein is expressed from a vector with the cold-inducible *E. coli cspA* promoter, which is of course consistent with membrane protein expression. The unique, enabling feature of the SPP System is the inducible expression of a site-specific mRNA interferase (MazF) from a second plasmid, which degrades endogenous mRNA by acting at ACA sites. Accordingly, the gene of interest must be synthesized to lack ACA sequences. The net result is that the target mRNA persists and becomes a preferential substrate for the translation machinery. The elimination of most host-derived mRNA is reported to create a quasidormant cell where expression of the target membrane protein is sustained. If Sec translocase function is also sustained, then this system may offer an advantage, as the target protein should encounter less competition from endogenous proteins on the membrane translocation pathway.

1.7
Clone Stability

When expressing membrane proteins, clone stability should always be a concern. The first indication of clone toxicity is often realized during the initial cloning/transformation step. Poor transformation results may indicate that mutant genes are being selected during the cloning step, so sequence verification is always advised and is absolutely critical if the gene has been amplified by polymerase chain reaction. If a clone is suspected to be toxic, certain precautions should be followed. First, lower growth temperatures are often stabilizing. Also, it is beneficial to include 0.1% glucose in selection plates in many situations. Glucose will repress basal expression from *Plac*, *PlacUV5*, *ParaBAD*, and *Prha*. (Note: *Ptac* and *Ptrc* are not subject to glucose repression as the CAP binding site is absent from these promoters). Glucose containing plates are also advantageous when transforming clones into T7 Express and DE3 expression strains, as the T7 RNA polymerase gene is controlled by *Plac* and *PlacUV5* in these strains, respectively. One exception is transformation into Lemo21(DE3) where glucose repression is not stabilizing. When transforming extremely toxic clones into Lemo21(DE3), 500 µM rhamnose addition to selection plates and starter cultures will reduce the basal expression to an undetectable level (Figure 1.1).

During the outgrowth stage for protein expression, plasmid maintenance should be examined. This is especially critical when propagating AmpR vectors, as the resistance protein (β-lactamase) is secreted and ampicillin may be completely degraded. Plasmid maintenance is easily checked by plating cells at the point of induction onto drug containing plates versus nondrug plates. If a significantly lower number of colonies are counted on the drug plates (below 80% the number counted on nondrug plates), then modifications to the protocol or the clone may be necessary. If plasmid maintenance is an issue with AmpR constructs, increasing the level of ampicillin to 200 µg/ml is recommended. Alternatively, initiate growth with 100 µg/ml ampicillin and then spike in another dose (100 µg/ml) at

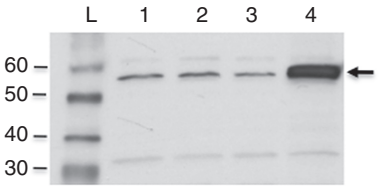

Figure 1.1 T7 expression is tightly regulated in Lemo21(DE3) cells. Whole-cell lysates were subjected to SDS–PAGE, and target protein was detected using anti-YidC serum that recognizes both endogenous wild-type YidC and recombinant 6His-YidC membrane protein expressed from pET28c. (1) No vector control indicating endogenous YidC level; (2) cells containing pET28-6hisyidC, no IPTG, no rhamnose; (3) cells containing pET28-6hisyidC, 500 µM rhamnose, no IPTG; (4) cells containing pET28-6hisyidC, 500 µM rhamnose, 400 µM IPTG. Arrow indicates YidC target.

mid-log stage. Carbenicillin (at 50–200 μg/ml) may be used in place of ampicillin. According to the Novagen pET system manual, pET (AmpR) clones may be stabilized by using high concentrations of carbenicillin and by changing the medium twice prior to induction. Carbenicillin is more stable than ampicillin in low pH conditions, which may be encountered after extended fermentation periods.

Vectors expressing KanR or CamR are preferred for creating membrane protein clones. One versatile KanR vector is pET28, which allows for simple construction of genes tagged at either end with the polyhistidine coding sequence. pBAD33 (CamR) is also a good choice as expression is tightly regulated. (Note: when cloning into the pBAD33 polylinker, a translation initiation signal (RBS site) must be included with the gene insert). In extreme cases plasmid maintenance systems can be incorporated. For example, the *hok/sok* system has been utilized by groups expressing G-protein-coupled receptors (GPCRs) in *E. coli* [31–32].

1.8 Media Types

The type of media is also an important consideration. Although the use of LB is commonly cited, we generally observe a greater level of membrane protein expression in Terrific Broth (TB). This conclusion was made after multiple expression trials using a *tac* promoter, which is insensitive to glucose repression. TB is a rich broth buffered by potassium phosphate and containing glycerol as a carbon source [33]. When the target protein is expressed from *Plac*, *PlacUV5*, *ParaBAD*, or *Prha*, a rich media containing a low-level of glucose may be more appropriate. With respect to controlling expression in BL21(DE3), Pan and Malcolm [34] found that 1% glucose addition to either TB or M9 starter cultures minimized basal expression to a level equal to pLysS-containing strains. These researchers further demonstrated that glucose addition is less important in a strain expressing lysozyme to control basal T7 expression. When target protein expression is driven directly from a sugar promoter, then glucose repression is advised. For example, pBAD constructs may be stabilized by growth in media containing 0.1% glucose, which should be metabolized by the point of induction with arabinose. Such a protocol leads to a discussion of "autoinduction" media [35] marketed as the Overnight Express™ autoinduction system for simplified T7 expression. The advantages of this system are: (i) manual IPTG induction is not required and (ii) expression trials are more reproducible as growth is carried out in a defined media containing a mix of carbon sources (generally glucose, lactose, and glycerol). When glucose is depleted, lactose serves to induce expression of the T7 RNA polymerase from the *lacUV5* promoter in DE3 strains. The actual inducer molecule is allolactose that is produced by β-galactosidase (the *lacZ* gene product). Thus, Studier points out that autoinduction should be performed in strains encoding an intact *lac* operon. Note: The T7 Express line of strains (NEB) are not suitable for autoinduction protocols as the T7 RNA polymerase gene disrupts the *lacZ* open reading frame (ORF).

1.9
Fusion Partners/Membrane Targeting Peptides

A first step in cloning or characterizing heterologous membrane protein ORFs is the analysis of membrane topology using more than one algorithm. Four common predictors are: SPOCTOPUS [36] (octopus.cbr.su.se), TopPred (http://mobyle.pasteur.fr/cgi-bin/portal.py?form=toppred), Phobius [37] (phobius.sbc.su.se), and TargetP 1.1 [38] (http://www.cbs.dtu.dk/services/TargetP/). Nielsen *et al.* [39] showed more than 10 years ago that eukaryotic secretory proteins (e.g., membrane receptors) are expressed with N-terminal signals that resemble cleavable signal peptides found in bacteria. *E. coli* proteins exported to the periplasm or outer membrane are expressed with a signal peptide that is cleaved by signal peptidase. In contrast, most endogenous *E. coli* membrane proteins are expressed with a more hydrophobic N-terminal signal that remains uncleaved (signal anchor). A more hydrophobic N-terminal signal increases the probability of *E. coli* SRP recognition and targeting of a protein to the cotranslational membrane insertion pathway [40]. Thus, when designing constructs for *E. coli* expression, the hydrophobicity of the N-terminal residues should be evaluated. If necessary, the N-terminal signal of the protein of interest may be replaced by a relatively hydrophobic signal from a different protein. For example, Chang *et al.* [41] tested eight different membrane-targeting peptides to maximize the *E. coli* expression of a plant derived P450 enzyme 8-cadinene hydroxylase (CAH). The results varied widely and surprisingly the signal from a bovine CAH performed the best. One note of caution regarding heterologous signal peptides is that rare codons may significantly impact expression results. Although using strains that correct for rare codons is an option (Table 1.1) the selection of an appropriate signal sequence is an empirical process because fully optimized translation at the 5′-end of the message is not necessarily advantageous. Other options are to replace the native signal sequence with an N-terminal fusion partner that travels the SRP pathway. For example, the *E. coli* GlpF protein has been used as a fusion partner [42] or the rationally designed P8CBDek fusion partner will also facilitate membrane targeting/expression of foreign membrane proteins in *E. coli* [9].

Maltose binding protein (MBP) from *E. coli* is a tried-and-true N-terminal fusion partner for enhancing the expression/solubility of heterologous protein. Native MBP is a periplasmic protein that is exported via the Sec pathway. Therefore, proteins fused to MBP (containing its native signal peptide) are targeted to the Sec translocase and, as such, have the opportunity to be integrated into the inner membrane upon completion of MBP export. This method has been applied to facilitate the expression of several eukaryotic membrane proteins. For example, the human cannabinoid receptor CB2 [32], human serotonin 5-HT_{1A} [43], rat neurotensin receptor [31], and the prokaryotic Glvi proton-gated ion channel [44] have been expressed in *E. coli* in functional form as fusions to MBP. Several studies have established that a large fraction of MBP molecules are delivered to the membrane and SecA in a post-translational manner after recognition by the cytoplasmic chaperone SecB. Accordingly, a reasonable assumption is that MBP–membrane

protein fusion expression might be improved by the overexpression of chaperones such as SecB, DnaK/DnaJ, and GroEL/GroES in order to protect hydrophobic segments from aggregation within the cytoplasm before they engage the Sec translocase.

1.10 Chaperone Overexpression

Many studies have demonstrated that cytoplasmic chaperone overexpression improves the expression of heterologous soluble proteins in *E. coli*. This approach has also been tested to aid the overexpression of CorA, an atypical membrane protein from *E. coli*. CorA lacks an N-terminal hydrophobic signal and is believed to integrate into the inner membrane by a post-translational process. Chen *et al.* [45] carried out a comprehensive study to determine the optimal conditions for CorA overexpression and to determine the effect of overexpressing several different *E. coli* factors relevant to protein biogenesis. A conclusion of this study was that increasing DnaK 8-fold resulted in a 4-fold increase of membrane integrated CorA when expression was carried out at 37 °C. However, the same net result (13–15 mg CorA) was obtained by simply lowering the expression temperature to the range of 15–30 °C. Thus, the underlying conclusion of this work was that the yield of membrane protein per cell may be increased by reducing the synthesis rate of the target protein. In fact, most studies indicate that membrane protein expression is optimal at 20–30 °C, with 20 °C being more favorable for more difficult polytopic membrane proteins. Chen *et al.* calculated the translation rate of CorA at the extremes: in wild-type *E. coli*, the average rate of CorA synthesis was estimated to be 600 molecules/cell/min at 15 °C and 5500 molecules/cell/min at 37 °C, whereas export of proOmpA through the Sec translocase was estimated to be 450–900 molecules/cell/min at 37 °C [46]. Since protein synthesis rate can easily exceed the Sec translocase capacity, this clearly points out the fact that the Sec translocase is a bottleneck for membrane protein expression in *E. coli*. Engineering a strain with greater translocase activity has not been achievable so far, so the most practical option is to express recombinant membrane proteins "in moderation." Moderate expression at lower temperatures helps to ensure that the chaperone pool is not exceeded and also that the translation rate does not far exceed the capacity of the Sec translocase.

Link *et al.* [47] recently examined the overexpression of the type 1 cannabinoid receptor CB1 and found a positive effect from several different helper proteins. In this study, DnaK/DnaJ coexpression from a plasmid again showed promise in increasing membrane protein yield. Furthermore, overexpression of Ffh (SRP protein component) and Trigger Factor (cotranslational chaperone) provided some benefit (2- to 3-fold enhancement). Most remarkably, the overexpression of FtsH (an inner membrane protease) resulted in up to 8-fold enhancement of GPCR expression and nearly a 2-fold higher cell density after 36 h at 12 °C. Link *et al.* also state in their Discussion that "FtsH overexpression also increases the bacterial

production of native *E. coli* membrane proteins such as YidC." The rationale behind the FtsH effect will be discussed in the next section.

1.11
Cautionary Notes Related to Chaperone Overexpression

DnaK overexpression may result in a situation where a fraction of the target protein is isolated as a complex with DnaK. This is a common occurrence during overexpression and Ni-NTA isolation of soluble protein and this outcome was also reported in the CorA–DnaK study. DnaK contamination after one or more chromatography steps most certainly indicates that a fraction of the target protein is not folded properly. Furthermore, attempts to remove DnaK contamination may be futile (personal experience with overexpression of soluble protein). DnaJ chaperone expression in wild-type strains apparently acts to inhibit the expression some membrane proteins according to Skretas and Georgiou [48]. This conclusion was made after a DnaJ null strain displayed a large increase in CB1–Green Fluorescent Protein (GFP) fluorescence and in the production of membrane-integrated CB1. When designing constructs for helper protein (e.g., DnaK) expression, do not employ high-copy vectors with inducible promoters. A simple and effective method is to clone the helper protein gene with its native promoter onto a low-copy vector such as pACYC184. Then the helper protein will be moderately overexpressed and its expression may be naturally regulated upon the induction of cellular stress responses.

1.12
Emerging Role of Quality Control Proteases

Although bacterial cells are extremely efficient factories for protein production, some recombinant proteins fall off or become stalled on the pathway to their final folded destination. This is especially reasonable to imagine for heterologous, hydrophobic proteins. So what happens to such proteins? One outcome is that such proteins serve as aggregation targets and the continuous supply of induced recombinant protein accumulates as inclusion bodies within the cytoplasm. During this process, the cell responds by upregulating chaperones *and* proteases to take care of the state of disarray [49]. Thus, if endogenous membrane protein assembly fails, protease degradation is a natural response. One bit of direct evidence for this comes from the work of van Bloois *et al.* [50] where FtsH protease was able to be crosslinked to the YidC, a protein with a loosely defined membrane protein chaperone activity [51–52]. FtsH is capable of processive, ATP-dependent degradation of *E. coli* membrane substrates such as YccA and SecY [53]. In cases of recombinant membrane protein overexpression, this activity apparently clears away those molecules that become stalled during the post- or cotranslational integration process. Other proteases must certainly play a role in clearing the membrane translocation pathway. In retrospect, at least two studies suggested that the

cytoplasmic Lon protease acts as a quality control factor during membrane protein biogenesis as well as soluble protein biogenesis. The first clue came from work in Tom Silhavy's lab published in 1992, where William Snyder was studying the *prlF1* host mutation, which suppresses Sec pathway jamming by the LamB–LacZ fusion protein [54]. Nearly 20 years ago, it was difficult to explain why the PrlF1 phenotype was only observed in *prlF1/lon*$^{+}$ strains. However, now it is certain that Lon functions to clear away fully translated proteins that are misfolded or partially translated proteins that are stalled at the ribosome [55]. The relevance of the *prlF* gene product is still a bit ambiguous, but recently Schmidt *et al.* [56] showed that PrlF is an antitoxin that counteracts the bacteriostatic effect of its toxin partner YhaV. The *prlF1* 7-bp insertion has two effects: (i) the downstream *yhaV* toxin gene is expressed to a much lesser extent and (ii) the PrlF1 antitoxin is stabilized by amino acids changes at its C-terminus. The overall net result appears to be that deactivation of the toxin system allows cells to better recover from jamming of the Sec pathway. We have tested the hypothesis that Lon is a key factor in rescuing ribosomes that stall during the cotranslational membrane insertion process. Our preliminary data indicates that Lon complementation improves the growth of cells overexpressing a polytopic membrane protein. However, so far an improved yield of membrane protein per cell has not been demonstrated. A role for Lon in membrane protein biogenesis is also suggested by Harris Bernstein's characterization of the SRP pathway. In a study published in 2001, Bernstein and Hyndman [57] reported that proteases Lon and ClpQ become essential when the SRP level is reduced, suggesting that SRP-deficient cells require an increased capacity to degrade mislocalized inner membrane proteins.

Other cell envelope proteases may be essential for membrane protein biogenesis. For example, Wang *et al.* [58] studied YidC-depleted cells and found that several cell envelope proteases were elevated. One such enzyme is HtpX, an inner membrane protease regulated by the Cpx envelope stress response [59].

Contrary to the immediate discussion, deletion of some host proteases may be beneficial. For example, OmpT acts at dibasic sequences and it may be detrimental during the protein isolation stage. Thus, preferred protein expression strains (e.g., BL21 derivatives) are OmpT minus. Note that most K-12 strains express OmpT. However, the K-12 KRX strain has been engineered to lack OmpT protease. Also, uncharacterized differences in protease expression between B and K-12 strains may affect the quality of expressed protein. One specific example is demonstrated by Luo *et al.* [9] where the same protein was resistant to proteolysis when expressed in NEB Express (a BL21 derivative), but was less stable during expression in MC1061, a robust K-12 strain.

1.13 Tag Selection

The selection of affinity tags/detection epitopes is an important consideration when constructing a recombinant membrane protein clone. As expression levels

are characteristically low, Western analysis is a routine procedure. However, extremely hydrophobic proteins may run anomalously on standard gel systems or may transfer poorly to Western detection membranes. Before performing sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS–PAGE) analysis, it may be necessary to heat samples at 37 °C rather than 95 °C to avoid aggregation of expressed membrane protein. One alternative to Western analysis is to monitor the expression/purification by expressing the protein of interest as a C-terminal fusion to GFP [60].

Metal-affinity chromatography is undoubtedly the preferred method for isolating recombinant membrane proteins from *E. coli*. Many researchers make use of eight histidines to improve the yield from low expressing clones. This may also reduce the background of *E. coli* metal binding proteins, as the protein of interest should be eluted at a higher imidazole concentration. If the isolation procedure includes an ultracentrifugation step to pellet the membrane fraction, then soluble *E. coli* metal binding proteins are less of a concern. The position of a polyhistidine tag may affect expression levels, behavior in solution and the propensity for membrane protein crystal formation [61]. Furthermore, if the protein N-terminus may be subject to signal peptidase cleavage, then of course the affinity tag needs to be located elsewhere. A polyhistidine tag may even be placed within a cytoplasmic loop of the target protein [62]. Histidine-rich sequences may affect membrane translocation, so a cytoplasmic location is recommended. Other effective purification tags include the Strep tag, the FLAG tag, or other immunoaffinity sequences [63]. For a more complete discussion of suitable fusion tags, consult the recent reviews by Xie *et al* [64]. The FLAG tag is rich with charged residues DYKDDDDK yet this tag is efficiently translocated across the inner membrane in the context of the P8CBD fusion partner [9]. The FLAG tag conveniently encodes the enterokinase protease site DDDDK and this protease works well in detergent-containing buffer. Thrombin is also recommended for removal of tags from membrane proteins as this protease shows reliable activity in many detergents. In contrast, tobacco etch virus (TEV) protease suffers from poor activity in several common detergents such as octyl-glucopyranoside [65].

1.14
Potential Expression Yield

Most importantly, have modest expectations. Always remember that quality is more important than quantity when attempting to overexpress membrane proteins in bacteria. Regarding prokaryotic proteins: any yield of membrane-integrated protein above 3 mg/l of culture is a good result. According to the NEB catalog 2007–2008, this level corresponds to approximately 2% of total cellular protein. Levels in the 10–20 mg/l range might be obtained for native *E. coli* membrane proteins (e.g., CorA study). In contrast, proteins from higher organisms will be expressed at much lower levels in most cases. An expression yield (in the membrane fraction) approaching 1 mg/l is an outstanding achievement for eukaryotic

polytopic membrane proteins. This discrepancy even after gene/codon optimization is not completely understood, but the following factors may be responsible:

i) The eukaryotic protein translation rate is generally lower than the rate in bacteria even when grown at low temperatures. Thus, eukaryotic membrane proteins may have evolved to require different chaperone requirements or folding timescales.

ii) The molecular composition of eukaryotic membranes varies considerably from the inner membrane composition of *E. coli*. For example, the bacterial plasma (inner) membrane is less rigid than the plasma membrane of mammalian cells due to the lack of cholesterol. Accordingly, membrane protein stability may be influenced.

iii) Wild-type *E. coli* cells do not offer the possibility of post-translational modifications (e.g., glycosylation) that may be necessary to stabilize some eukaryotic proteins.

iv) Bacteria express a different repertoire of membrane proteases that may act on some sequences/structures presented in heterologous proteins.

1.15 Strategies to Overcome Protein Instability

Protein instability may be the result of many factors: unproductive membrane insertion may lead to degradation by cellular proteases. In contrast, proteins failing to assemble properly in the membrane may aggregate *in vivo* or, postexpression, the protein of interest may aggregate as a result of nonoptimal buffer conditions during membrane solubilization or affinity chromatography. Expression and isolation of "stable" membrane protein is still a very empirical process. If a detergent/buffer screen fails to give a protein sample amenable to characterization, then it may be wise to screen homologs, truncation mutants, or point mutants. Screening proteins from multiple organisms is a common practice in order to find one member of a family that behaves well enough for characterization. Proteins from thermophilic organisms may be better behaved during expression and/or crystallization but this is not a general rule. Screening truncation mutants is a common practice with soluble proteins, but this type of systematic approach is not well tested with membrane proteins. Finally, point mutations may provide stabilization but the study of a random variant protein is less desirable. The radical approach of engineering a stabilization domain (T4 lysozyme) into a cytoplasmic loop of the human β_2-adrenergic receptor aided in expression within Sf9 insect cells and influenced the formation of quality protein crystals [66]. Another proven method for stabilizing the same GPCR is cocrystallization of a monoclonal antibody that binds to an inherently flexible region of the receptor [67]. Perhaps similar approaches will be fruitful for bacterial expression and/or crystallization of unstable polytopic membrane proteins.

Acknowledgments

Lab members Jianying Luo, Carine Robichon, and Julie Choulet are recognized for many contributions including: sharing of unpublished data, critical reading of the manuscript, and assistance with preparation of figures/tables. Additionally, Bill Jack and Elisabeth Raleigh of New England Biolabs are thanked for critical reading of the manuscript.

Abbreviations

aTc	anhydrotetracycline
CAH	8-cadinene hydroxylase
CAP	catabolite gene activator protein
DHFR	dihydrofolate reductase
GFP	Green Fluorescent Protein
GPCR	G-protein-coupled receptor
IPTG	isopropyl-β-D-thiogalactopyranoside
MBP	Maltose binding protein
RBS	ribosomal binding site
SDS	sodium dodecyl sulfate
SRP	signal recognition particle
PAGE	polyacrylamide gel electrophoresis
TB	Terrific Broth
TEV	tobacco etch virus
TM	transmembrane
UCLA	University of California at Los Angeles

References

1 Buchanan, S.K. (1999) Beta-barrel proteins from bacterial outer membranes: structure, function and refolding. *Curr. Opin. Struct. Biol.*, **9**, 455–461.

2 Kiefer, H. (2003) *In vitro* folding of alpha-helical membrane proteins. *Biochim. Biophys. Acta*, **1610**, 57–62.

3 Rogl, H., Kosemund, K., Kuhlbrandt, W., and Collinson, I. (1998) Refolding of *Escherichia coli* produced membrane protein inclusion bodies immobilized by nickel chelating chromatography. *FEBS Lett.*, **432**, 21–26.

4 Sarramegna, V., Muller, I., Milon, A., and Talmont, F. (2006) Recombinant G protein-coupled receptors from expression to renaturation: a challenge towards structure. *Cell. Mol. Life Sci.*, **60**, 1–16.

5 Xie, K. and Dalbey, R.E. (2008) Inserting proteins into the bacterial cytoplasmic membrane using the Sec and YidC translocases. *Nat. Rev. Microbiol.*, **6**, 234–244.

6 Roosild, T.P., Greenwald, J., Vega, M., Castronovo, S., Riek, R., and Choe, S. (2005) NMR structure of Mistic, a membrane-integrating protein for membrane protein expression. *Science*, **307**, 1317–1321.

7 Serek, J., Bauer-Manz, G., Struhalla, G., van den Berg, L., Kiefer, D., Dalbey, R., and Kuhn, A. (2004) *Escherichia coli* YidC

is a membrane insertase for Sec-independent proteins. *EMBO J.*, **23**, 294–301.
8 Samuelson, J.C., Chen, M., Jiang, F., Moller, I., Wiedmann, M., Kuhn, A., Phillips, G.J., and Dalbey, R.E. (2000) YidC mediates membrane protein insertion in bacteria. *Nature*, **406**, 637–641.
9 Luo, J., Choulet, J., and Samuelson, J.C. (2009) Rational design of a fusion partner for membrane protein expression in *E. coli*. *Protein Sci.*, **18**, 1735–1744.
10 Luirink, J., von Heijne, G., Houben, E., and de Gier, J.-W. (2005) Biogenesis of inner membrane proteins in *Escherichia coli*. *Annu. Rev. Microbiol.*, **59**, 329–355.
11 Bahari, L., Parlitz, R., Eitan, A., Stjepanovic, G., Bochkareva, E.S., Sinning, I., and Bibi, E. (2007) Membrane targeting of ribosomes and their release require distinct and separable functions of FtsY. *J. Biol. Chem.*, **282**, 32168–32175.
12 Weiche, B., Bürk, J., Angelini, S., Schiltz, E., Thumfart, J.O., and Koch, H.G. (2008) A cleavable N-terminal membrane anchor is involved in membrane binding of the *Escherichia coli* SRP receptor. *J. Mol. Biol.*, **377**, 761–773.
13 Hirschel, B.J., Shen, V., and Schlessinger, D. (1980) Lactose operon transcription from wild-type and L8-UV5 *lac* promoters in *Escherichia coli* treated with chloramphenicol. *J. Bacteriol.*, **143**, 1534–1537.
14 deBoer, H.A., Comstock, L.J., Yansura, D.G., and Heynecker, H.L. (1982) Construction of a tandem *trp–lac* promoter and a hybrid *trp–lac* promoter for efficient and controlled expression of the human growth hormone gene in *Escherichia coli*, in *Promoters: Structure and Function* (eds R.L. Rodriguez and M.J. Chamberlin), Praeger, New York, pp. 462–481.
15 deBoer, H.A., Comstock, L.J., and Vasser, M. (1983) The *tac* promoter: a functional hybrid derived from *trp* and *lac* promoters. *Proc. Natl. Acad. Sci. USA*, **80**, 21–25.
16 Amann, E., Brosius, J., and Ptashne, M. (1983) Vectors bearing a hybrid *trp–lac* promoter useful for regulated expression of cloned genes in *Escherichia coli*. *Gene*, **25**, 167–178.
17 Mulligan, M.E., Brosius, J., and Clure, W.R. (1985) Characterization *in vitro* of the effect of spacer length on the activity of *Escherichia coli* RNA polymerase at the *tac* promoter. *J. Biol. Chem.*, **260**, 3529–3538.
18 Calos, M.P. (1978) DNA Sequence for a low-level promoter of the *lac* repressor gene and an "up" promoter mutation. *Nature*, **274**, 762–765.
19 Guzman, L.M., Belin, D., Carson, M.J., and Beckwith, J. (1995) Tight regulation, modulation, and high-level expression by vectors containing the arabinose PBAD promoter. *J. Bacteriol.*, **177**, 4121–4130.
20 Giacalone, M.J., Gentile, A.M., Lovitt, B.T., Berkley, N.L., Gunderson, C.W., and Surber, M.W. (2006) Toxic protein expression in *Escherichia coli* using a rhamnose-based tightly regulated and tunable promoter system. *Biotechniques*, **40**, 355–364.
21 Siegele, D.A. and Hu, J.C. (1997) Gene expression from plasmids containing the araBAD promoter at subsaturating inducer concentrations represents mixed populations. *Proc. Natl. Acad. Sci. USA*, **94**, 8168–8172.
22 Haldimann, A., Daniels, L.L., and Wanner, B.L. (1998) Use of new methods for construction of tightly regulated arabinose and rhamnose promoter fusions in studies of the *Escherichia coli* phosphate regulon. *J. Bacteriol.*, **180**, 1277–1286.
23 Skerra, A. (1994) Use of the tetracycline promoter for the tightly regulated production of a murine antibody fragment in *Escherichia coli*. *Gene*, **151**, 131–135.
24 Lutz, R. and Bujard, H. (1997) Independent and tight regulation of transcriptional units in *Escherichia coli* via the *lac*R/O, the TetR/O and AraC/I1-I2 regulatory elements. *Nucleic Acids Res.*, **25**, 1203–1210.
25 Studier, F.W., Rosenberg, A.H., Dunn, J.J., and Dubendorff, J.W. (1990) Use of T7 RNA polymerase to direct expression of cloned genes. *Methods Enzymol.*, **185**, 60–89.
26 Cheng, X., Zhang, X., Pflugrath, J.W., and Studier, F.W. (1994) The structure of

bacteriophage T7 lysozyme, a zinc amidase and an inhibitor of T7 RNA polymerase. *Proc. Natl. Acad. Sci. USA*, **91**, 4034–4038.

27 Wagner, S., *et al.* (2008) Tuning *Escherichia coli* for membrane protein overexpression. *Proc. Natl. Acad. Sci. USA*, **105**, 14371–14376.

28 Miroux, B. and Walker, J.E. (1996) Over-production of proteins in *Escherichia coli*: mutant hosts that allow synthesis of some membrane proteins and globular proteins at high levels. *J. Mol. Biol.*, **260**, 289–298.

29 Massey-Gendel, E., Zhao, A., Boulting, G., Kim, H.-Y., Balamotis, M.A., Seligman, L.M., Nakamoto, R.K., and Bowie, J.U. (2009) Genetic selection system for improving recombinant membrane protein expression in *E. coli*. *Protein Sci.*, **18**, 372–383.

30 Suzuki, M., Mao, L., and Inouye, M. (2007) Single protein production (SPP) system in *Escherichia coli*. *Nat. Protoc.*, **2**, 1802–1810.

31 Tucker, J. and Grisshammer, R. (1996) Purification of a rat neurotensin receptor expressed in *Escherichia coli*. *Biochem. J.*, **317**, 891–899.

32 Yeliseev, A.A., Wong, K.K., Soubias, O., and Gawrisch, K. (2005) Expression of human peripheral cannabinoid receptor for structural studies. *Protein Sci.*, **14**, 2638–2653.

33 Tartof, K.D. and Hobbs, C.A. (1988) New cloning vectors and techniques for easy and rapid restriction mapping. *Gene*, **30**, 169–182.

34 Pan, S.H. and Malcolm, B.A. (2000) Reduced background expression and improved plasmid stability with pET vectors in BL21(DE3). *Biotechniques*, **29**, 1234–1237.

35 Studier, F.W. (2005) Protein production by auto-induction in high-density shaking cultures. *Protein Expr. Purif.*, **41**, 207–234.

36 Viklund, H., Bernsel, A., Skwark, M., and Elofsson, A. (2008) SPOCTOPUS: a combined predictor of signal peptides and membrane protein topology. *Bioinformatics*, **24**, 2928–2929.

37 Käll, L., Krogh, A., and Sonnhammer, E.L. (2007) Advantages of combined transmembrane topology and signal peptide prediction – the Phobius web server. *Nucleic Acids Res.*, **35**, W429–W432.

38 Emanuelsson, O., Brunak, S., von Heijne, G., and Nielsen, H. (2006) Locating proteins in the cell using TargetP, SignalP and related tools. *Nat. Protoc.*, **2**, 95371.

39 Nielsen, H., Engelbrecht, J., Brunak, S., and von Heijne, G. (1997) Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. *Protein Eng.*, **10**, 1–6.

40 Lee, H.C. and Bernstein, H.D. (2001) The targeting pathway of *Escherichia coli* presecretory and integral membrane proteins is specified by the hydrophobicity of the targeting signal. *Proc. Natl. Acad. Sci. USA*, **98**, 3471–3476.

41 Chang, M.C.Y., Eachus, R.A., Trieu, W., Ro, D.-K., and Keasling, J.D. (2007) Engineering *Escherichia coli* for production of functionalized terpenoids using plant P450s. *Nat. Chem. Biol.*, **3**, 274–277.

42 Neophytou, I., Harvey, R., Lawrence, J., Marsh, P., Panaretou, B., and Barlow, D. (2007) Eukaryotic integral membrane protein expression utilizing the *Escherichia coli* glycerol-conducting channel protein (GlpF). *Appl. Microbiol. Biotechnol.*, **77**, 375–381.

43 Bertin, B., Freissmuth, M., Breyer, R.M., Schutz, W., Strosberg, A.D., and Marullo, S. (1992) Functional expression of the human serotonin 5-HT1A receptor in *Escherichia coli*. *J. Biol. Chem.*, **267**, 8200–8206.

44 Bocquet, N., Prado de Carvalho, L., Cartaud, J., Neyton, J., Le Poupon, C., Taly, A., Grutter, T., Changeux, J.P., and Corringer, P.J. (2007) A prokaryotic proton-gated ion channel from the nicotinic acetylcholine receptor family. *Nature*, **445**, 116–119.

45 Chen, Y., Song, J., Sui, S.-f., and Wang, D.-N. (2003) DnaK and DnaJ facilitated the folding process and reduced inclusion body formation of magnesium transporter CorA overexpressed in *Escherichia coli*. *Protein Exp. Purif.*, **32**, 221–231.

46 De Keyzer, J., van der Does, C., and Driessen, A.J. (2002) Kinetic analysis of the translocation of fluorescent precursor

proteins into *Escherichia coli* membrane vesicles. *J. Biol. Chem.*, **277**, 46059–46065.
47 Link, A.J., Skretas, G., Strauch, E.-M., Chari, N.S., and Georgiou, G. (2008) Efficient production of membrane-integrated and detergent-soluble G protein-coupled receptors in *Escherichia coli*. *Protein Sci.*, **17**, 1857–1863.
48 Skretas, G. and Georgiou, G. (2009) Genetic analysis of G protein-coupled receptor expression in *Escherichia coli*: inhibitory role of DnaJ on the membrane integration of the human central cannabinoid receptor. *Biotechnol. Bioeng.*, **102**, 357–367.
49 Wagner, S., Baars, L., Ytterberg, A.J., Klussmeier, A., Wagner, C.S., Nord, O., Nygren, P.A., van Wijk, K.J., and de Gier, J.W. (2007) Consequences of membrane protein overexpression in *Escherichia coli*. *Mol. Cell Proteomics*, **6**, 1527–1550.
50 van Bloois, E., Dekker, H.L., Fröderberg, L., Houben, E.N., Urbanus, M.L., de Koster, C.G., de Gier, J.W., and Luirink, J. (2008) Detection of cross-links between FtsH, YidC, HflK/C suggests a linked role for these proteins in quality control upon insertion of bacterial inner membrane proteins. *FEBS Lett.*, **582**, 1419–1424.
51 Nagamori, S., Smirnova, I.N., and Kaback, H.R. (2004) Role of YidC in folding of polytopic membrane proteins. *J. Cell Biol.*, **165**, 53–62.
52 Kol, S., Nouwen, N., and Driessen, A.J. (2008) Mechanisms of YidC-mediated insertion and assembly of multimeric membrane protein complexes. *J. Biol. Chem.*, **283**, 31269–31273.
53 Kihara, A., Akiyama, Y., and Ito, K. (1999) Dislocation of membrane proteins in FtsH-mediated proteolysis. *EMBO J.*, **18**, 2970–2981.
54 Snyder, W.B. and Silhavy, T.J. (1992) Enhanced export of beta-galactosidase fusion proteins in prlF mutants is Lon-dependent. *J. Bacteriol.*, **174**, 5661–5668.
55 Choy, J.S., Aung, L.L., and Karzai, A.W. (2007) Lon protease degrades transfer-messenger RNA-tagged proteins. *J. Bacteriol.*, **189**, 6564–6571.
56 Schmidt, O., Schuenemann, V.J., Hand, N.J., Silhavy, T.J., Martin, J., Lupas, A.N., and Djuranovic, S. (2007) *prlF* and *yhaV* encode a new toxin–antitoxin system in *Escherichia coli*. *J. Mol. Biol.*, **372**, 894–905.
57 Bernstein, H.D. and Hyndman, J.B. (2001) Physiological basis for conservation of the signal recognition particle targeting pathway in *Escherichia coli*. *J. Bacteriol.*, **183**, 2187–2197.
58 Wang, P., Kuhn, A., and Dalbey, R.E. (2010) Global change of gene expression and cell physiology in YidC-depleted *E. coli*. *J. Bacteriol.*, **192**, 2193–2209.
59 Shimohata, N., Chiba, S., Saikawa, N., Ito, K., and Akiyama, Y. (2002) The Cpx stress response system of *Escherichia coli* senses plasma membrane proteins and controls HtpX, a membrane protease with a cytosolic active site. *Genes Cells*, **7**, 653–662.
60 Drew, D., Lerch, M., Kunji, E., Slotboom, D.-J., and de Gier, J.-W. (2006) Optimization of membrane protein overexpression and purification using GFP fusions. *Nat. Methods*, **3**, 303–313.
61 Mohanty, A.K. and Wiener, M.C. (2004) Membrane protein expression and production: effects of polyhistidine tag length and position. *Protein Expr. Purif.*, **33**, 311–325.
62 Paetzel, M., Strynadka, N.C., Tschantz, W.R., Casareno, R., Bullinger, P.R., and Dalbey, R.E. (1997) Use of site-directed chemical modification to study an essential lysine in *Escherichia coli* leader peptidase. *J. Biol. Chem.*, **272**, 9994–10003.
63 Takayama, H., Chelikani, P., Reeves, P.J., Zhang, S., and Khorana, H.G. (2008) High-level expression, single-step immunoaffinity purification and characterization of human tetraspanin membrane protein CD81. *PLoS ONE*, **3**, e2314.
64 Xie, H., Guo, X.-M., and Chen, H. (2009) Making the most of fusion tags technology in structural characterization of membrane proteins. *Mol. Biotechnol.*, **42**, 135–145.
65 Mohanty, A.K., Simmons, C., and Wiener, M.C. (2003) Inhibition of tobacco

etch virus protease activity by detergents. *Protein Expr. Purif.*, **27**, 109–114.

66 Cherezov, V., *et al.* (2007) High-resolution crystal structure of an engineered human beta2-adrenergic G protein-coupled receptor. *Science*, **318**, 1258–1265.

67 Rasmussen, S.G.F., *et al.* (2007) Crystal structure of the human beta2-adrenergic G protein-coupled receptor. *Nature*, **450**, 383–387.