Minireview

Revolutionizing membrane protein overexpression in bacteria

Susan Schlegel,[1†] Mirjam Klepsch,[1†]
Dimitra Gialama,[1] David Wickström,[1]
Dirk Jan Slotboom[2] and Jan-Willem de Gier[1]*

[1]*Center for Biomembrane Research, Department of Biochemistry and Biophysics, Stockholm University, SE-106 91 Stockholm, Sweden.*
[2]*Department of Biochemistry, University of Groningen, Nyenborg 4, 9747 AG Groningen, the Netherlands.*

Summary

The bacterium *Escherichia coli* is the most widely used expression host for overexpression trials of membrane proteins. Usually, different strains, culture conditions and expression regimes are screened for to identify the optimal overexpression strategy. However, yields are often not satisfactory, especially for eukaryotic membrane proteins. This has initiated a revolution of membrane protein overexpression in bacteria. Recent studies have shown that it is feasible to (i) engineer or select for *E. coli* strains with strongly improved membrane protein overexpression characteristics, (ii) use bacteria other than *E. coli* for the expression of membrane proteins, (iii) engineer or select for membrane protein variants that retain functionality but express better than the wild-type protein, and (iv) express membrane proteins using *E. coli*-based cell-free systems.

Introduction

In both pro- and eukaryotes 20–30% of all genes encode membrane proteins, which usually form supramolecular complexes acting in many different and often essential capacities (Wallin and von Heijne, 1998). Membrane proteins play key roles in many diseases and around 70% of all drug targets are membrane proteins (Lundstrom,

2007). The natural abundance of membrane proteins is often too low to isolate sufficient material for *in vitro* functional and structural studies. Furthermore, the use of natural sources excludes the possibility of genetically modifying proteins to facilitate their detection and/or purification, and efficiently labelling them for Nuclear Magnetic Resonance (NMR) and crystallographic studies.

Two classes of membrane proteins exist: β-barrel and helical bundle membrane proteins (von Heijne, 1999). β-Barrel membrane proteins can often be expressed in inclusion bodies in *Escherichia coli* from which they are readily isolated and refolded in their native conformation (Bannwarth and Schulz, 2003). In contrast, for helical bundle membrane proteins the isolation of functional material from inclusion bodies is seldom successful. Therefore, helical bundle membrane proteins must be overexpressed in such a way that they properly insert in the membrane from which they can be purified after detergent extraction or wherein they can be studied directly. Here, we will deal only with the overexpression of helical bundle membrane proteins, hereafter referred to as membrane proteins.

Escherichia coli is the most widely used host when attempting to overexpress membrane proteins. Usually, a variety of strains, culture conditions and expression regimes are screened for to find the optimal overexpression strategy (reviewed by Wang *et al.*, 2003). However, yields – especially of eukaryotic membrane proteins – are usually not sufficient for functional and structural studies. Furthermore, eukaryotic overexpression hosts usually do not perform better for production of eukaryotic membrane proteins (Wagner *et al.*, 2006).

Recent studies have shown that overexpression yields of membrane proteins can be greatly improved by (i) engineering or selecting for *E. coli* strains with strongly improved membrane protein overexpression characteristics, (ii) using bacterial expression hosts other than *E. coli*, (iii) engineering or selecting for better expressing membrane protein variants that retain functionality, and (iv) employing *E. coli*-based cell-free systems. New analytical methods for monitoring membrane protein overexpression have been crucial for some of these developments.

Monitoring membrane protein overexpression

Monitoring the localization, quantity and quality of overexpressed membrane proteins is important to allow assessment and optimization of overexpression yields. It is unpredictable whether overexpressed membrane proteins end up in the lipid bilayer or in inclusion bodies. Therefore, the first step in monitoring membrane protein overexpression is usually the fractionation of the overexpression host into an insoluble fraction (inclusion bodies), a soluble fraction (combination of the cytoplasm and periplasm) and a membrane fraction (see e.g. Korepanova *et al.*, 2005).

Coomassie/silver-stained SDS-polyacrylamide gels are usually used to detect membrane proteins in (subfractionated) overexpression hosts. They allow assessing the integrity and the quantity of overexpressed membrane proteins. Western blotting, using antibodies against, for example, an expression or purification tag, is also widely used to detect proteins. Due to the hydrophobic nature of membrane proteins their transfer from a gel to blotting membrane can be troublesome, making Western blotting less suitable for quantitative purposes.

The above described approaches to monitor membrane protein overexpression are very laborious, time-consuming and often not very accurate. Therefore alternative methods have been developed.

Fusing green fluorescent protein (GFP) to the C-terminus of a membrane protein enables monitoring the levels of overexpressed protein in whole cells (Drew *et al.*, 2001; 2005; 2006). The GFP moiety does not fold properly (and is therefore non-fluorescent) if the membrane protein GFP fusion ends up in inclusion bodies. But it becomes fluorescent if the membrane protein is correctly inserted into the membrane. Using whole cells as starting material, the membrane protein GFP fusion can be visualized directly through in-gel fluorescence in SDS-polyacrylamide gels (Drew *et al.*, 2006). This enables to rapidly assess the integrity of the overexpressed material quantitatively, and is a welcome alternative for Western blotting.

To easily monitor overexpression of large numbers of soluble proteins, the so-called Colony Filtration (CoFi) blotting method was developed (Cornvik *et al.*, 2005). In this method colonies are transferred, induced and lysed on a filter membrane. Upon cell lysis, the soluble proteins diffuse through the filter membrane and are captured on a nitrocellulose membrane, whereas inclusion bodies cannot pass through the filter. Subsequently, the nitrocellulose membrane is incubated with antibodies or probes that specifically recognize the protein of interest. Recently, the CoFi blotting method was modified so that it can also be used to monitor the expression of detergent solubilized membrane proteins (Martinez Molina *et al.*, 2008).

The above described methods do not provide any information on the functionality of the overexpressed proteins; integration into the membrane is no guarantee for this. To monitor the functionality of a membrane protein, its function should be known and also an activity assay should be available. Depending on the protein, for example, binding assays with fluorescent or radioactive ligands or transport assays have been used (see e.g. Wang *et al.*, 2003).

Selecting and engineering *E. coli* strains with improved membrane protein overexpression characteristics

Obviously, during evolution there has not been selective pressure on *E. coli* towards overexpression of (heterologous) membrane proteins. However, this does not mean that it is impossible to create *E. coli* strains that can efficiently produce membrane proteins. More than a decade ago, in the laboratory of John Walker a simple screening approach was used to isolate *E. coli* strains with improved membrane protein overexpression characteristics (Miroux and Walker, 1996). *Escherichia coli* BL21(DE3) cells expressing toxic membrane proteins were plated on solid medium with inducer to select for survivors, i.e. cells that can cope with the toxic effects of membrane protein overexpression. The C41(DE3) and C43(DE3) strains – which have commonly become known as the Walker strains – were isolated in this way and are now widely used to overexpress membrane proteins. It should be noted that the Walker strains do not show improved yields for all membrane proteins tested (Wagner *et al.*, 2008).

BL21(DE3), from which the Walker strains were derived, was developed for the overexpression of soluble proteins (Studier *et al.*, 1990). Overexpression in BL21(DE3) is driven by the bacteriophage T7 RNA polymerase (T7 RNAP), which transcribes much faster than *E. coli* RNAP (Iost *et al.*, 1992). Expression of the gene encoding T7 RNAP is governed by the *lac*UV5 promoter, which is a mutant that is more powerful than the wild-type *lac* promoter (Arditti *et al.*, 1968). The rationale behind BL21(DE3) is very simple; the more mRNA produced, the more protein can be overexpressed. However, this simple relation does not hold for all proteins, and certainly not for all membrane proteins.

It has been shown that mutations in the *lac*UV5 promoter governing expression of T7 RNAP are key to the improved membrane protein overexpression characteristics of the Walker strains (Wagner *et al.*, 2008). These mutations result in the production of much lower amounts of T7 RNAP upon induction of expression than in BL21(DE3). Subsequent lower production rates of the mRNA for the membrane protein ensure that the capacity of the membrane protein biogenesis machinery is

sufficient to integrate the overexpressed proteins in the cytoplasmic membrane.

Based on the characterization of the Walker strains, a derivative strain of *E. coli* BL21(DE3), termed Lemo21(DE3), was engineered in which the activity of the T7 RNAP can be precisely controlled by its natural inhibitor T7 lysozyme (T7 Lys) (Wagner *et al.*, 2008). In Lemo21(DE3) the gene encoding the T7 Lys is on a plasmid under control of a rhamnose promoter, which is extremely well titratable and covers a broad range of expression intensities (Giacalone *et al.*, 2006). Lemo21(DE3) is tunable for membrane protein overexpression and conveniently allows optimizing overexpression of any given membrane protein by using only a single strain rather than a multitude of different strains. The combination of the *lac*UV5- and the rhamnose promoters governing expression of T7 RNAP from the chromosome and T7 Lys from a plasmid, respectively, guarantees the widest window of expression intensities possible. Therefore, in Lemo21(DE3) the amount of membrane protein produced can be easily harmonized with the membrane protein biogenesis capacity of the cell. The development of this strain was sped up tremendously by using GFP fusions to monitor membrane protein overexpression.

A sophisticated selection strategy to isolate mutant strains of *E. coli* with improved membrane protein overexpression characteristics was recently described (Massey-Gendel *et al.*, 2009). The gene encoding the membrane protein of interest was simultaneously expressed from two separate plasmids. Each plasmid was constructed such that upon expression the membrane protein was C-terminally fused to a different selection marker. Cells containing both plasmids were exposed to treatments that randomly introduce mutations and subsequently selected for growth on selective medium. An increased resistance towards the two selection drugs indicated an increased expression of the target protein. Importantly, the use of a dual selection strategy considerably lowers the risk of obtaining mutations that confer resistance to both drugs without increasing membrane protein expression levels. Furthermore, a rapid method for curing isolated strains of the plasmids was used during the selection process; the plasmids were removed by *in vivo* digestion with the homing endonuclease I-*Cre*I. This allowed rapid isolation of potentially interesting mutant strains for retesting of expression. It has been shown that in one of the isolated strains the copy number of the expression plasmids is considerably lower, suggesting that in this strain – like in the Walker strains – a lowered expression level improves membrane protein overexpression yields.

Coexpression of chaperones is routinely used to improve overexpression yields of soluble proteins (see e.g. Kolaj *et al.*, 2009). Chen and colleagues (2003) pioneered this approach for membrane proteins. They tried to improve expression of the magnesium transporter CorA in *E. coli* by coexpressing various components involved in membrane protein biogenesis and protein folding. CorA overexpression levels could be improved by the coexpression of the cytoplasmic DnaK/J chaperone system. The CorA transporter is a homopentamer and each protomer of CorA consists of a large N-terminal cytoplasmic domain and two transmembrane segments at the C-terminus. The architecture of CorA suggests that unlike most other membrane proteins it is not co-translationally targeted via the SRP/Sec-translocon pathway, but is targeted post-translationally (Luirink *et al.*, 2005). Thus, it is likely that the DnaK/J chaperone system is involved in the targeting and folding of CorA and could explain why coexpression of this system improves CorA yields.

Link and colleagues (2008) used a coexpression approach to improve expression of four different human class I GPCRs in *E. coli*. They co-overexpressed various components involved in the biogenesis and quality control of membrane proteins in *E. coli*. The use of GFP fusions allowed them to rapidly assess the effect of co-overexpression on GPCR expression levels using flow cytometry. Coexpression of the membrane-bound protease FtsH greatly enhanced the expression of all four GPCRs. However, it was not shown if improved yields also resulted in more functional material, and it is not yet understood how FtsH improves overexpression of GPCRs.

It is also possible that there are host proteins that hamper membrane protein overexpression. Therefore, Georgiou and co-workers monitored overexpression of the human GPCR CB1 fused to GFP in an *E. coli* transposon library (Skretas and Georgiou, 2009). Fluorescence-activated cell sorting (FACS) was used to identify and isolate cells with improved overexpression yields for further characterization. A transposon insertion in the *dnaJ* gene resulted in an increase in CB1-GFP fluorescence and an enhancement in production of membrane-integrated CB1. Thus, the chaperone/co-chaperone DnaJ seems to inhibit overexpression of CB1. However, again it was not shown if improved yields also resulted in more functional material. DnaJ did not inhibit expression of another GPCR tested. This suggests that – at least in some cases – the optimal strain background for the overexpression of a membrane protein may be protein-specific.

The examples above show that different strategies have been used to create or select for *E. coli* strains with improved membrane protein overexpression characteristics. It would be very interesting to compare these strains. This could also shed more light on the mechanism(s) that underlie their improved performance. It is very likely that what we have seen so far is just the beginning of the

engineering of *E. coli* strains with improved overexpression characteristics for membrane proteins.

Bacteria other than *E. coli* as membrane protein overexpression hosts

Why has *E. coli* become the most widely used host for membrane protein overexpression? The answer must be that this is largely for historical reasons. The organism was isolated over a century ago and turned out to be particularly easy to cultivate, which paved the way for *E. coli* to become a favourite model organism. Today a very large toolbox of genetic, molecular biological and biochemical methods is at hand. But it is good to realize that *a priori* there is no reason why *E. coli* should be more suitable for the heterologous overexpression of membrane proteins than any other bacterium. In contrast, there may be good arguments why other bacteria could be more suitable for expression of certain membrane proteins. For example, bacteria with slower translation rates than *E. coli* may be able to deal better with 'difficult folders'. Similarly, bacteria that have different repertoires of chaperones from *E. coli*, like Gram-positive bacteria which express two copies of the integral membrane chaperone YidC [e.g. *Bacillus subtilis/Lactococcus lactis* (Luirink *et al.*, 2001; Zweers *et al.*, 2008; Funes *et al.*, 2009)], may perform better at insertion and assembly of heterologous membrane proteins. Therefore, instead of optimizing *E. coli* for overexpression of 'difficult' membrane proteins, a viable alternative strategy is to look for different expression hosts that have distinct properties and may perform better for the protein of interest.

One promising bacterial host for overexpression of membrane proteins is *L. lactis*. *Lactococcus lactis* is a Gram-positive lactic acid bacterium, which is used in the dairy industry. Because of the industrial interest in the organism, its physiology has been studied in great detail. The organism is genetically accessible, and a variety of expression plasmids, both high- and low-copy-number plasmids, and inducible promoters are available (Kuipers *et al.*, 1997). The basics of the expression system have been reviewed by Kunji and coworkers (2003; 2005). There are several notable cases in which functional overexpression of eukaryotic as well as bacterial membrane proteins [e.g. the human KDEL receptor and Na⁺/tyrosine transporter (Tyt1) of *Fusobacterium nucleatum*] could be achieved in *L. lactis*, but not in *E. coli* (Kunji *et al.*, 2003; 2005; Monne *et al.*, 2005a; 2007; Quick and Javitch, 2007). Among the potential advantages of *L. lactis* are a slower growth rate than *E. coli* (~1 doubling h^{-1} versus ~2 doublings h^{-1} for *E. coli*), which could be beneficial for expression of proteins that do not fold easily; the presence of a single membrane only, which facilitates functional characterization; and the lack of excessive

proteolytic activity, which may help to prevent breakdown of expressed proteins. Just like in *E. coli*, coexpression of multiple proteins from different plasmids is possible in *L. lactis* and can be used for chaperone production alongside the membrane protein of interest (Rodionov *et al.*, 2009). Two recent studies have shown that *L. lactis* also can be used for efficient incorporation of amino acid derivatives in expressed proteins (El Khattabi *et al.*, 2008; Berntsson *et al.*, 2009). This makes the organism a more complete alternative to *E. coli* for production of proteins for X-ray crystallography, where Seleno-methionine incorporation is routinely used for phase determination.

Besides *E. coli* and *L. lactis*, other bacterial or archaeal organisms may also be good expression hosts for certain membrane proteins (e.g. *Halobacterium salinarum* and *B. subtilis*). An important question is: which organism to choose for the protein of interest? There is no general answer to this question, because it is impossible to predict how heterologously expressed proteins will behave in different hosts. However, when expressing prokaryotic proteins it is likely best to choose a host that is as closely related to the natural host of the protein to be expressed as possible, so as to mimic homologous expression as closely as possible. For example, when Surade and colleagues (2006) compared the expression hosts *E. coli* and *L. lactis* for overexpression of membrane proteins from *Salmonella* Typhimurium, *E. coli* – not surprisingly – performed better than *L. lactis*, because *E. coli* is a member of the closest known genus to *Salmonella* (McClelland *et al.*, 2001).

Engineering and isolating membrane protein variants with improved overexpression characteristics

Rather than engineering or changing the expression host, an alternative strategy to improve membrane protein overexpression yields in bacteria is to engineer the membrane protein of interest. It has been known for a long time that modifying N- and C-termini can improve expression yields significantly. Recently, also random mutagenesis approaches have been used to isolate better-expressing membrane protein variants.

Modifying N- and C-termini

Most pro- and eukaryotic membrane proteins have N-terminal tails that have to be translocated across the membrane (Wallin and von Heijne, 1995; Daley *et al.*, 2005). Translocation of an N-terminal tail depends on the ability of the N-terminus to remain in a translocation-competent conformation, the number of positively charged residues in the tail region, and the 'strength' of the first transmembrane segment, i.e. the charge differ-

ence, the length and the overall hydrophobicity of the reverse-signal anchor (Monne *et al.*, 2005b). The inability to efficiently translocate the N-terminal tail of a membrane protein may severely hamper its overexpression. Indeed, the functional overexpression of the yeast mitochondrial carrier AAC2 (ADP/ATP exchanger) can be increased in *L. lactis* if the N-terminus is shortened, or the N-terminal tail is swapped with a shorter one taken from the isoform AAC3 (Monne *et al.*, 2005a). Likewise, functional expression of, for example, the neurotension receptor (NTR) from rat is improved in *E. coli* by fusing the maltose-binding protein (MBP) with its secretory signal sequence to its N-terminus (Tucker and Grisshammer, 1996). Finally, the N-terminal addition of the *B. subtilis* protein 'Mistic' was reported to improve expression of *E. coli* histidine kinase receptors and a number of eukaryotic membrane proteins in *E. coli* (Roosild *et al.*, 2005; Kefala *et al.*, 2007). In addition, it has been shown that a 'Mistic' orthologue/chimera can improve expression of a voltage-gated potassium channel from *Aplysia californica* (aKv1.1) (Dvir & Choe, 2009).

Grisshammer and co-workers tested the ability of various additions and combinations of C-terminal tags (e.g. biotinylation, poly histidine, flag and strep tags), or single additions of either a polyhistidine/c-Myc epitope or *E. coli* thioredoxin (aa 2–109) fusion, to improve the functional overexpression of the aforementioned MBP–NTR fusion (Tucker and Grisshammer, 1996). The most significant improvement of overexpression was obtained with thioredoxin, which was attributed to the remarkable stability of the globular protein. The combined use of an N-terminal MBP and a His-tagged C-terminal thioredoxin fusion has also been successfully employed for the overexpression of other GPCRs in *E. coli* (Weiss and Grisshammer, 2002; White *et al.*, 2004; Yeliseev *et al.*, 2005). Interestingly, the addition of GFP to the C-terminus of the human KDEL receptor improved functional expression levels considerably in *L. lactis* (Drew *et al.*, 2005). Like thioredoxin, GFP is an exceptionally stable protein that very well could protect overexpressed proteins from proteases.

Random mutagenesis

Expression levels of membrane proteins in bacteria can be improved not only by modifying N- and C-termini, but also random mutagenesis approaches can be used to this end. Recently, random mutant libraries for nine membrane proteins were screened for expression in *E. coli* using the CoFi blot method (Martinez Molina *et al.*, 2008). For five out of the nine proteins tested, one cycle of random mutagenesis resulted in significant improvements of yields of detergent solubilized membrane protein. Among the five proteins was the human microsomal

glutathione *S*-transferase 2. Notably, the mutations that improved expression of this variant of the protein did not interfere with its activity. It should be noted that it is not known how the various mutations contribute to the higher overexpression yields.

A similar approach was used to isolate NTR variants from rat with improved expression characteristics in *E. coli* (Sarkar *et al.*, 2008). The gene encoding NTR was randomly mutagenized and expressed as an N-terminal MBP and C-terminal thioredoxin fusion (see above). NTR expression was monitored directly using a fluorescently labelled ligand that only properly folded protein can bind. Cells expressing the largest number of functional receptors exhibited the greatest fluorescence and were isolated using FACS. This way a NTR variant, DO3, that shows much higher expression in *E. coli* but is still fully functional was isolated. The improved overexpression yields are most likely due to improved co-translational folding and insertion (Sarkar *et al.*, 2008). The beauty of this approach is that the fluorescently labelled ligand allows screening for variants that not only express better, but are also still properly folded in a single experiment.

For DO3 it has been shown that the protein is also more stable when solubilized and purified than wild-type NTR (Sarkar *et al.*, 2008). This suggests that this variant may be more amenable for functional and structural studies, just like the turkey β1-adrenergic and human adenosine A2a receptor variants that were recently selected for improved thermostability by Ala scanning using *E. coli* as expression platform (Magnani *et al.*, 2008; Serrano-Vega *et al.*, 2008).

Producing membrane proteins using cell-free expression systems

Cell-free protein expression systems have been developed to circumvent the toxic effects of protein production on the overexpression host, and to facilitate the labelling of proteins for NMR and crystallographic studies. Systems based on both *E. coli* extracts and purified *E. coli* components are now available (Shimizu *et al.*, 2001; Schwarz *et al.*, 2008). Since cell-free systems are 'open systems', ongoing protein synthesis reactions can be easily manipulated, which not only has made it possible to improve expression yields considerably but also greatly facilitates the labelling of proteins.

Membrane proteins from both pro- and eukaryotic origin have been expressed in *E. coli*-based cell-free expression systems. In the absence of a possible scaffold (detergent or lipid bilayer), membrane proteins have been expressed as precipitates that have to be solubilized afterwards using detergents (e.g. Klammt *et al.*, 2004). They also have been expressed in a soluble form in the presence of detergents (e.g. Ishihara *et al.*, 2005; Klammt *et al.*, 2005)

Table 1. Overview of the different strategies used to improve membrane protein overexpression in bacteria.

	Selecting and engineering *Escherichia coli* strains	
Overexpression based strain selection	Screening for BL21(DE3) derivatives that survive the overexpression of toxic membrane proteins led to the isolation of the C41(DE3) and C43(DE3) strains.	[Miroux and Walker, 1996]
Mutagenesis/overexpression based strain selection	Mutagenesis of cells combined with a selection procedure based on a two plasmid expression system that allows the simultaneous expression of the target membrane protein fused to two different selection markers led to the isolation of the EXP strains.	[Massey-Gendel et al., 2009]
Transposon mutagenesis	Identification of the cytoplasmic chaperone/co-chaperone DnaJ as a factor that inhibits membrane integration of the human GPCR CB1.	[Skretas and Georgiou, 2009]
Tuning expression	In the Lemo21(DE3) strain, in which T7 RNA polymerase activity can be controlled by its natural inhibitor T7 lysozyme, the expression level of the target membrane protein can be precisely regulated.	[Wagner et al., 2008]
Coexpression of chaperones	Coexpression of the cytoplasmic DnaK/J chaperone system increased soluble overexpression yields of the magnesium transporter CorA in the cytoplasm and in the cytoplasmic membrane.	[Chen et al., 2003]
Coexpression of proteases	Coexpression of the membrane-bound AAA+ protease FtsH enhanced yields of membrane-integrated class I GPCRs.	[Link et al., 2008]
	Strains other than *E. coli*	
Lactococcus lactis	Successful overexpression of the human KDEL receptor (hKDEL), various bacterial transporters, and mitochondrial carriers.	[Kunji et al., 2003; Kunji et al., 2005; Monne et al., 2005a; Monne et al., 2007; Quick and Javitch, 2007]
	Altering the target membrane protein	
Engineering the N-terminus	Shortening the translocated N-terminus of the yeast mitochondrial carrier AAC2 led to improved overexpression levels in *L. lactis*.	[Monne et al., 2005a]
N-terminal fusions	Fusing the maltose binding protein (MBP) to the N-terminus of the rat neurotensin receptor (NTR) led to increased overexpression yields in *E. coli*.	[Tucker and Grisshammer, 1996]
	Fusing the *Bacillus subtilis* protein 'Mistic' to the N-terminus increased overexpression yields of several *E. coli* histidine kinase receptors and various eukaryotic membrane proteins in *E. coli*.	[Roosild et al., 2005; Kefala et al., 2007]
	A 'Mistic' orthologue/chimera improved expression levels of a voltage gated potassium channel from *Aplysia californica*.	[Dvir and Choe, 2009]
C-terminal fusions	Fusion of various tags/epitopes to MBP-NTR/other GPCRs led to identification of thioredoxin (Trx) as the tag that improves expression in *E. coli* most.	[Tucker and Grisshammer, 1996; Weiss and Grisshammer, 2002; White et al., 2004; Yeliseev et al., 2005]
	Fusion of green fluorescent protein to the hKDEL improved expression levels in *L. lactis*.	[Drew et al., 2005]
Random mutagenesis	Screening random mutant libraries resulted in the identification of membrane protein variants with increased expression levels for five out of the nine tested proteins.	[Martinez Molina et al., 2008]
	Random mutagenesis of the NTR gene enabled the isolation of a fully functional MBP-NTR-Trx variant with improved overexpression characteristics.	[Sarkar et al., 2008]
	E. coli based cell-free expression	
Expression in the absence of a scaffold	Expression of membrane proteins as precipitates that are solubilized afterwards.	[Klammt et al., 2004]
Expression in the presence of a scaffold	Expression of membrane proteins in the presence of detergent and/or lipids, or inverted membrane vesicles.	[Klammt et al., 2005; Ishihara et al., 2005; Wuu and Swartz, 2008]
	Expression of membrane proteins in the presence of nanolipoprotein particles.	[Cappuccio et al., 2008; Katzen et al., 2008]

and/or lipids as a scaffold. Fluorinated and hemifluorinated detergents have successfully been used as alternatives to commonly used detergents for membrane protein expression in cell-free synthesis (Park *et al.*, 2007). These detergents allow the easy and efficient reconstitution of a membrane protein into a lipid bilayer. Membrane proteins expressed in a cell-free system can also be inserted into a planar phospholipid membrane bilayer surrounded by an apolipoprotein. Such a structure is called a nanolipoprotein particle or NLP. Nanolipoprotein particles can be added to the reaction from the start or the apolipoproteins required for the formation of NLPs can be coexpressed with the membrane protein in the presence of lipids (Cappuccio *et al.*, 2008; Katzen *et al.*, 2008). Both bacterial and eukaryotic membrane proteins have been produced in *E. coli* extracts using NLPs. Surprisingly, membrane proteins can assemble properly in detergents/lipids/NLPs in the absence of many of the factors that assist their biogenesis into the membrane *in vivo*. Compared with the *in vivo* situation, membrane protein expression/biogenesis in cell-free system is a very slow process. This may make it possible for membrane proteins to fold properly in a more or less unassisted fashion. Only recently, reasonable expression yields have been reported for a cell-free based system with purified inner membrane vesicles (Wuu and Swartz, 2008).

For quite a number of membrane proteins expressed in cell-free systems it has been shown that they are functional (Schwarz *et al.*, 2008). To determine the structure of the *E. coli* membrane protein EmrE cell-free expression was used to produce a Seleno-methionine derivative that was used to solve the phases (Chen *et al.*, 2007). A recent screen, which compared the expression of more than 100 *E. coli* membrane proteins in a cell-free expression system versus expression *in vivo*, showed that more proteins could be expressed using the cell-free system (Savage *et al.*, 2007). Thus, *E. coli*-based cell-free expression systems have become a serious alternative for the production of membrane proteins.

Future perspectives

Recently, impressive progress has been made using bacteria to produce membrane proteins, including some notoriously difficult to express eukaryotic proteins (Table 1). However, there are still some major challenges ahead; for example, most eukaryotic membrane proteins are glycosylated when inserted into the endoplasmic reticulum (ER) membrane, and for a considerable number of them this is essential for proper folding, stability and also function (Helenius and Aebi, 2004). The bacteria used thus far to produce membrane proteins are not able to glycosylate them at all. Recently, based on the glycosylation machinery from the bacterium *Campylobacter jejuni*, Wacker and

colleagues (2002) constructed an *E. coli* strain that can add sugar moieties to proteins on the periplasmic face of the cytoplasmic membrane, creating a situation similar to the one in the ER. Unfortunately, the type of glycosylation does not resemble any of the kinds of glycosylation that occur in eukaryotic systems. In this respect it should be noted that glycosylation patterns are organism and cell type dependent. However, Wacker and colleagues (2002) may very well have made the first step towards engineering *E. coli* strains that can glycosylate proteins similar to eukaryotes. For the yeast *Pichia pastoris* it has been shown – using a sophisticated strategy – that it is possible to engineer strains that can glycosylate proteins similar to humans (Hamilton *et al.*, 2006). Also the lipid composition of the membrane can be important for proper folding, stability and function of overexpressed membrane proteins (see e.g. Hanson *et al.*, 2008; Powl *et al.*, 2008; Phillips *et al.*, 2009). Our rapidly growing knowledge of the biosynthesis of lipids/membranes may make it possible to also engineer strains with an optimal lipid composition for the overexpression of specific membrane proteins.

What we have seen so far is just the beginning of the revolution of the overexpression of membrane proteins in bacteria, which will greatly stimulate functional and structural studies of this important class of proteins.

References

Arditti, R.R., Scaife, J.G., and Beckwith, J.R. (1968) The nature of mutants in the *lac* promoter region. *J Mol Biol* **38:** 421–426.

Bannwarth, M., and Schulz, G.E. (2003) The expression of outer membrane proteins for crystallization. *Biochim Biophys Acta* **1610:** 37–45.

Berntsson, R.P., Oktaviana, N.A., Fusetti, F., Thunnissen, A.M., Poolman, B., and Slotboom, D.J. (2009) Selenomethionine incorporation in proteins expressed in *Lactococcus lactis*. *Protein Sci* **18:** 1121–1127.

Cappuccio, J.A., Blanchette, C.D., Sulchek, T.A., Arroyo, E.S., Kralj, J.M., Hinz, A.K., *et al.* (2008) Cell-free co-expression of functional membrane proteins and apolipoprotein, forming soluble nanolipoprotein particles. *Mol Cell Proteomics* **7:** 2246–2253.

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 Expr Purif* **32:** 221–231.

Chen, Y.J., Pornillos, O., Lieu, S., Ma, C., Chen, A.P., and Chang, G. (2007) X-ray structure of EmrE supports dual topology model. *Proc Natl Acad Sci USA* **104:** 18999–19004.

Cornvik, T., Dahlroth, S.L., Magnusdottir, A., Herman, M.D., Knaust, R., Ekberg, M., and Nordlund, P. (2005) Colony filtration blot: a new screening method for soluble protein expression in *Escherichia coli. Nat Methods* **2:** 507–509.

Daley, D.O., Rapp, M., Granseth, E., Melen, K., Drew, D., and von Heijne, G. (2005) Global topology analysis of the *Escherichia coli* inner membrane proteome. *Science* **308:** 1321–1323.

Drew, D.E., von Heijne, G., Nordlund, P., and de Gier, J.W.L. (2001) Green fluorescent protein as an indicator to monitor membrane protein overexpression in *Escherichia coli. FEBS Lett* **507:** 220–224.

Drew, D., Slotboom, D.J., Friso, G., Reda, T., Genevaux, P., Rapp, M., *et al.* (2005) A scalable, GFP-based pipeline for membrane protein overexpression screening and purification. *Protein Sci* **14:** 2011–2017.

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.

Dvir, H., and Choe, S. (2009) Bacterial expression of a eukaryotic membrane protein in fusion to various Mistic orthologs. *Protein Exp Purif* **68:** 28–33.

El Khattabi, M., van Roosmalen, M.L., Jager, D., Metselaar, H., Permentier, H., Leenhouts, K., and Broos, J. (2008) *Lactococcus lactis* as expression host for the biosynthetic incorporation of tryptophan analogues into recombinant proteins. *Biochem J* **409:** 193–198.

Funes, S., Hasona, A., Bauerschmitt, H., Grubbauer, C., Kauff, F., Collins, R., *et al.* (2009) Independent gene duplications of the YidC/Oxa/Alb3 family enabled a specialized cotranslational function. *Proc Natl Acad Sci USA* **106:** 6656–6661.

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.

Hamilton, S.R., Davidson, R.C., Sethuraman, N., Nett, J.H., Jiang, Y., Rios, S., *et al.* (2006) Humanization of yeast to produce complex terminally sialylated glycoproteins. *Science* **313:** 1441–1443.

Hanson, M.A., Cherezov, V., Griffith, M.T., Roth, C.B., Jaakola, V.P., Chien, E.Y., *et al.* (2008) A specific cholesterol binding site is established by the 2.8 Å structure of the human beta2-adrenergic receptor. *Structure* **16:** 897–905.

von Heijne, G. (1999) Recent advances in the understanding of membrane protein assembly and structure. *Q Rev Biophys* **32:** 285–307.

Helenius, A., and Aebi, M. (2004) Roles of N-linked glycans in the endoplasmic reticulum. *Annu Rev Biochem* **73:** 1019–1049.

Iost, I., Guillerez, J., and Dreyfus, M. (1992) Bacteriophage T7 RNA polymerase travels far ahead of ribosomes *in vivo. J Bacteriol* **174:** 619–622.

Ishihara, G., Goto, M., Saeki, M., Ito, K., Hori, T., Kigawa, T., *et al.* (2005) Expression of G protein coupled receptors in a cell-free translational system using detergents and thioredoxin-fusion vectors. *Protein Expr Purif* **41:** 27–37.

Katzen, F., Fletcher, J.E., Yang, J.P., Kang, D., Peterson, T.C., Cappuccio, J.A., *et al.* (2008) Insertion of membrane proteins into discoidal membranes using a cell-free protein expression approach. *J Proteome Res* **7:** 3535–3542.

Kefala, G., Kwiatkowski, W., Esquivies, L., Maslennikov, I., and Choe, S. (2007) Application of Mistic to improving the expression and membrane integration of histidine kinase receptors from *Escherichia coli. J Struct Funct Genomics* **8:** 167–172.

Klammt, C., Lohr, F., Schafer, B., Haase, W., Dotsch, V., Ruterjans, H., *et al.* (2004) High level cell-free expression and specific labeling of integral membrane proteins. *Eur J Biochem* **271:** 568–580.

Klammt, C., Schwarz, D., Fendler, K., Haase, W., Dotsch, V., and Bernhard, F. (2005) Evaluation of detergents for the soluble expression of alpha-helical and beta-barrel-type integral membrane proteins by a preparative scale individual cell-free expression system. *FEBS J* **272:** 6024–6038.

Kolaj, O., Spada, S., Robin, S., and Wall, J.G. (2009) Use of folding modulators to improve heterologous protein production in *Escherichia coli. Microb Cell Fact* **8:** 9.

Korepanova, A., Gao, F.P., Hua, Y., Qin, H., Nakamoto, R.K., and Cross, T.A. (2005) Cloning and expression of multiple integral membrane proteins from *Mycobacterium tuberculosis* in *Escherichia coli. Protein Sci* **14:** 148–158.

Kuipers, O.P., de Ruyter, P.G., Kleerebezem, M., and de Vos, W.M. (1997) Controlled overproduction of proteins by lactic acid bacteria. *Trends Biotechnol* **15:** 135–140.

Kunji, E.R., Slotboom, D.J., and Poolman, B. (2003) *Lactococcus lactis* as host for overproduction of functional membrane proteins. *Biochim Biophys Acta* **1610:** 97–108.

Kunji, E.R., Chan, K.W., Slotboom, D.J., Floyd, S., O'Connor, R., and Monne, M. (2005) Eukaryotic membrane protein overproduction in *Lactococcus lactis. Curr Opin Biotechnol* **16:** 546–551.

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.

Luirink, J., Samuelsson, T., and de Gier, J.W. (2001) YidC/Oxa1p/Alb3: evolutionarily conserved mediators of membrane protein assembly. *FEBS Lett* **501:** 1–5.

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.

Lundstrom, K. (2007) Structural genomics and drug discovery. *J Cell Mol Med* **11:** 224–238.

McClelland, M., Sanderson, K.E., Spieth, J., Clifton, S.W., Latreille, P., Courtney, L., *et al.* (2001) Complete genome sequence of *Salmonella enterica* serovar Typhimurium LT2. *Nature* **413:** 852–856.

Magnani, F., Shibata, Y., Serrano-Vega, M.J., and Tate, C.G. (2008) Co-evolving stability and conformational homogeneity of the human adenosine A2a receptor. *Proc Natl Acad Sci USA* **105:** 10744–10749.

Martinez Molina, D., Cornvik, T., Eshaghi, S., Haeggstrom, J.Z., Nordlund, P., and Sabet, M.I. (2008) Engineering membrane protein overproduction in *Escherichia coli*. *Protein Sci* **17:** 673–680.

Massey-Gendel, E., Zhao, A., Boulting, G., Kim, H.Y., Balamotis, M.A., Seligman, L.M., *et al.* (2009) Genetic selection system for improving recombinant membrane protein expression in *E. coli*. *Protein Sci* **18:** 372–383.

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.

Monne, M., Chan, K.W., Slotboom, D.J., and Kunji, E.R. (2005a) Functional expression of eukaryotic membrane proteins in *Lactococcus lactis*. *Protein Sci* **14:** 3048–3056.

Monne, M., Hessa, T., Thissen, L., and von Heijne, G. (2005b) Competition between neighboring topogenic signals during membrane protein insertion into the ER. *FEBS J* **272:** 28–36.

Monne, M., Robinson, A.J., Boes, C., Harbour, M.E., Fearnley, I.M., and Kunji, E.R. (2007) The mimivirus genome encodes a mitochondrial carrier that transports dATP and dTTP. *J Virol* **81:** 3181–3186.

Park, K.H., Berrier, C., Lebaupain, F., Pucci, B., Popot, J.L., Ghazi, A., and Zito, F. (2007) Fluorinated and hemifluorinated surfactants as alternatives to detergents for membrane protein cell-free synthesis. *Biochem J* **403:** 183–187.

Phillips, R., Ursell, T., Wiggins, P., and Sens, P. (2009) Emerging roles for lipids in shaping membrane-protein function. *Nature* **459:** 379–385.

Powl, A.M., East, J.M., and Lee, A.G. (2008) Importance of direct interactions with lipids for the function of the mechanosensitive channel MscL. *Biochemistry* **47:** 12175–12184.

Quick, M., and Javitch, J.A. (2007) Monitoring the function of membrane transport proteins in detergent-solubilized form. *Proc Natl Acad Sci USA* **104:** 3603–3608.

Rodionov, D.A., Hebbeln, P., Eudes, A., ter Beek, J., Rodionova, I.A., Erkens, G.B., *et al.* (2009) A novel class of modular transporters for vitamins in prokaryotes. *J Bacteriol* **191:** 42–51.

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.

Sarkar, C.A., Dodevski, I., Kenig, M., Dudli, S., Mohr, A., Hermans, E., and Pluckthun, A. (2008) Directed evolution of a G protein-coupled receptor for expression, stability, and binding selectivity. *Proc Natl Acad Sci USA* **105:** 14808–14813.

Savage, D.F., Anderson, C.L., Robles-Colmenares, Y., Newby, Z.E., and Stroud, R.M. (2007) Cell-free complements *in vivo* expression of the *E. coli* membrane proteome. *Protein Sci* **16:** 966–976.

Schwarz, D., Dotsch, V., and Bernhard, F. (2008) Production of membrane proteins using cell-free expression systems. *Proteomics* **8:** 3933–3946.

Serrano-Vega, M.J., Magnani, F., Shibata, Y., and Tate, C.G. (2008) Conformational thermostabilization of the beta1-adrenergic receptor in a detergent-resistant form. *Proc Natl Acad Sci USA* **105:** 877–882.

Shimizu, Y., Inoue, A., Tomari, Y., Suzuki, T., Yokogawa, T., Nishikawa, K., and Ueda, T. (2001) Cell-free translation reconstituted with purified components. *Nat Biotechnol* **19:** 751–755.

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.

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.

Surade, S., Klein, M., Stolt-Bergner, P.C., Muenke, C., Roy, A., and Michel, H. (2006) Comparative analysis and 'expression space' coverage of the production of prokaryotic membrane proteins for structural genomics. *Protein Sci* **15:** 2178–2189.

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

Wacker, M., Linton, D., Hitchen, P.G., Nita-Lazar, M., Haslam, S.M., North, S.J., *et al.* (2002) N-linked glycosylation in *Campylobacter jejuni* and its functional transfer into *E. coli*. *Science* **298:** 1790–1793.

Wagner, S., Bader, M.L., Drew, D., and de Gier, J.W. (2006) Rationalizing membrane protein overexpression. *Trends Biotechnol* **24:** 364–371.

Wagner, S., Klepsch, M.M., Schlegel, S., Appel, A., Draheim, R., Tarry, M., *et al.* (2008) Tuning *Escherichia coli* for membrane protein overexpression. *Proc Natl Acad Sci USA* **105:** 14371–14376.

Wallin, E., and von Heijne, G. (1995) Properties of N-terminal tails in G-protein coupled receptors – a statistical study. *Protein Eng* **8:** 693–698.

Wallin, E., and von Heijne, G. (1998) Genome-wide analysis of integral membrane proteins from eubacterial, archaean, and eukaryotic organisms. *Protein Sci* **7:** 1029–1038.

Wang, D.N., Safferling, M., Lemieux, M.J., Griffith, H., Chen, Y., and Li, X.D. (2003) Practical aspects of overexpressing bacterial secondary membrane transporters for structural studies. *Biochim Biophys Acta* **1610:** 23–36.

Weiss, H.M., and Grisshammer, R. (2002) Purification and characterization of the human adenosine A(2a) receptor functionally expressed in *Escherichia coli*. *Eur J Biochem* **269:** 82–92.

White, J.F., Trinh, L.B., Shiloach, J., and Grisshammer, R. (2004) Automated large-scale purification of a G protein-coupled receptor for neurotensin. *FEBS Lett* **564:** 289–293.

Wuu, J.J., and Swartz, J.R. (2008) High yield cell-free production of integral membrane proteins without refolding or detergents. *Biochim Biophys Acta* **1778:** 1237–1250.

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.

Zweers, J.C., Barák, I., Becher, D., Driessen, A.J., Hecker, M., Kontinen, V.P., *et al.* (2008) Development of *Bacillus subtilis* as a cell factory for membrane proteins and protein complexes. *Microb Cell Fact* **7:** 10.