Available online at www.sciencedirect.com
SCIENCE DIRECT®
Biochimica et Biophysica Acta 1610 (2003) 23–36
BIOCHIMICA ET BIOPHYSICA ACTA
BBA
www.bba-direct.com

Review

Practical aspects of overexpressing bacterial secondary membrane transporters for structural studies

Da-Neng Wang*, Markus Safferling [1], M. Joanne Lemieux, Heather Griffith, Yong Chen [2], Xiao-Dan Li [3]

Skirball Institute of Biomolecular Medicine and Department of Cell Biology, New York University School of Medicine, 540 First Avenue, New York, NY 10016, USA

Received 26 July 2002; accepted 4 November 2002

Abstract

Membrane transporter proteins play critical physiological roles in the cell and constitute 5–10% of prokaryotic and eukaryotic genomes. High-resolution structural information is essential for understanding the functional mechanism of these proteins. A prerequisite for structural study is to overexpress such proteins in large quantities. In the last few years, over 20 bacterial membrane transporters were overexpressed at a level of 1 mg/liter of culture or higher, most often in *Escherichia coli*. In this review we analyzed those factors that affect the quantity and quality of the protein produced, and summarized recent progress in overexpression of membrane transporters from bacterial inner membrane. Rapid progress in genome sequencing provides opportunities for expressing several homologues and orthologues of the target protein simultaneously, while the availability of various expression vectors allows flexible experimental design. Careful optimization of cell culture conditions can drastically improve the expression level and homogeneity of the target protein. New sample preparation techniques for mass spectrometry of membrane proteins have enabled one to identity the rigid protein core, which can be subsequently overexpressed. Size-exclusion chromatography on HPLC has proven to be an efficient method in screening detergent, pH an other conditions required for maintaining the stability and monodispersity of the protein. Such high-quality preparations of membrane transporter proteins will probably lead to successful crystallization and structure determination of these proteins in the next few years.

Keywords: Membrane protein overexpression; Secondary membrane transporter; Membrane protein structure

1. Introduction

To survive and function properly, cells exchange substances such as nutrients, ions, and metabolites with their environment. These tasks are performed by transporter proteins that are embedded in the cell membrane. Membrane transporters account for 5–10% of proteins in bacterial genomes [1,2] and 3% of the human genome [3,4] (Table 1). A majority of these transport proteins use either ion or solute gradients as the driving force to translocate substrates across the membrane, and they are called secondary transporters. Based on phylogenetic analysis, secondary transporter proteins have been classified into over 100 families and superfamilies [1,2,5]. The larger family, the major facilitator superfamily (MFS), currently has more than 1000 identified members, and they account for about 25% of all transporter proteins [6].

Abbreviations: $C_{12}E_8$, octyloxyethylene dodecylether; CMC, critical micellar concentration; Cymal-5, cyclohexyl-pentylmaltoside; Cymal-6, cyclohexyl-hexylmaltoside; DDM, dodecylmaltoside; DG, decylglucoside; DM, decylmaltoside; G3P, glycerol-3-phosphate; HPLC, high-performance liquid chromatography; HRP, horseradish peroxidase; IPTG, isopropyl-β-D-thiogalactopyranoside; LDAO, lauryldimethylamine oxide; MALDI-TOF, matrix-assisted laser desorption/ionization time-of-flight; Ni^{2+}–NTA, nickel–nitrilotriacetic acid; NM, nonylmaltoside; OD, optical density; PAGE, polyacrylamide gel electrophoresis; PCR, polymerase chain reaction; SDS, sodium dodecyl sulfate; UDM, undecylmaltoside

* Corresponding author. Tel.: +1-212-263-8634; fax: +1-212-263-8951.

E-mail address: wang@saturn.med.nyu.edu (D.N. Wang).

[1] Present address: Bayer CropScience, BCS-R-RL-TT, 51368 Leverkusen, Germany.

[2] Present address: Department of Biological Sciences, Tsinghua University, Beijing 100084, China.

[3] Present address: Department of Microbiology, HHSC 1216, Columbia University, 70 W 168th Street, New York, NY 10032, USA.

0005-2736/03/$ - see front matter © 2003 Elsevier Science B.V. All rights reserved.
doi:10.1016/S0005-2736(02)00709-5

Table 1
Number of transporters identified in genomes to date

	Genes	Transporters	Transporter families	MFS[a]	APC[b]
Escherichia coli [90]	4282	297	59	66	22
Bacillus subtilis [91]	4100	262	51	65	18
Thermoplasma acidophilum [92]	1509	142	n.a.[c]	n.a.[c]	>12
Saccharomyces cerevisiae [93]	6141	258	42	78	24
Human [3,4]	26,383	1006	n.a.[c]	n.a.[c]	n.a.[c]

[a] Major facilitator superfamily.
[b] Amino-proline-choline transporter family.
[c] Information not yet available.

In human cells, membrane transporters play vital roles in physiological processes and their dysfunction is implicated in a variety of diseases. For example, type II diabetes is due to the impairment of the glucose uptake process in muscle and fat cells, which is mediated by the glucose transporter Glut4 [7–9]. Glycogen storage disease type I is caused by a deficiency in the glucose-6-phosphate transporter (G6PT) [10]. In amyotrophic lateral sclerosis, Huntington's disease and Alzheimer's disease, the number of glutamate transporters per unit mass of brain tissue is decreased [11,12]. Again in the brain, cocaine exerts its addictive effects by directly blocking the dopamine transporter [13]. Equally important for human health is bacterial antibiotic resistance, which is caused largely by transporter proteins acting as drug efflux pumps [14,15].

Mutagenesis and functional studies on membrane transporters have yielded detailed information on the individual amino acid residues responsible for substrate binding and translocation [16–20]. However, because no high-resolution structure has been determined for any secondary transporter protein from either eukaryotic sources or bacterial inner membrane, substrate specificity and diversity as well as the molecular mechanisms underlying membrane transport for these proteins remain poorly understood.

An overexpression system for membrane transporters is a prerequisite for structural studies. This is because tens, and sometimes hundreds, of milligrams of highly purified samples are needed for those experiments, and few transporter proteins are abundant in natural sources. In addition, protein produced using an overexpression system is more homogeneous than that from natural sources, and therefore can yield crystals more readily. The target protein can also be manipulated at the gene level in an expression system, allowing crystallization attempts of different protein constructs. At present, almost no such overexpression system exists for any mammalian transporter [21], and for that matter, for membrane proteins from mammalian sources. However, a bacterial homologue can be used as a prototype to obtain the essential information on the structure and mechanism for its mammalian homologue. Indeed, over two-dozen secondary transporter proteins have been overexpressed and purified in milligram quantities (Table 2). Their substrates range from ions, sugars, sugar–phosphates, amino acids, to peptides. Successful expression and purification, together with technical advances in membrane protein crystallization [22,23], will probably lead to structure determination of a number of secondary transporters at high resolution in the near future.

In the current review, we discuss practical aspects that are important for producing membrane transporters for structural studies. Readers can consult other reviews on principles of protein expression [24,25] and protocols for overexpressing membrane proteins [26–29]. The expression yield and protein quality are critical to the crystallization experiment and both are affected by multiple factors at the stages of transcription, translation, and membrane insertion. Thus, we have grouped the factors into those important for the expression level, and those critical for the protein quality. Understandably, such a classification scheme is somewhat artificial, because, for example, cell culture conditions affect both the quantity and quality of the expressed protein. Still, this scheme is particularly helpful to readers with practical questions like: 'How do I increase the yield of my membrane protein expression system?', 'How can I improve the quality of my protein by optimizing cell culture conditions?' or, 'How can I stabilize my transporter protein after purification?'. We focus on secondary transporters from bacterial sources and primarily discuss the *Escherichia coli* system, because almost all studies on overexpression of secondary membrane transporters have been carried out in this system [30–32]. In addition, only those studies where a purification yield of 0.5 mg of protein per liter of bacterial culture or higher was obtained will be considered useful for structural work. Finally, we discuss future directions for overexpression of bacterial and eukaryotic membrane transporters in *E. coli*.

2. Factors important for expression level

E. coli is the system of choice for producing bacterial membrane transporters for structural studies. Most importantly, the likelihood of overexpressing bacterial membrane transporters in *E. coli* is relatively high. In addition, its low-cost, high growth rate, well-characterized genetics, and the availability of large number of cloning vectors make *E. coli* particularly attractive for large-scale protein expression [25,33].

2.1. Choice of protein target

Sequence data from genomics projects, the availability of genomic DNA and development in polymerase chain reaction (PCR) cloning techniques have made it possible to clone almost any transporter protein. As a result, multiple members of the protein family of interest can be selected for overexpression and crystallization simultaneously using a strategy commonly referred to as *searching of crystallization space by cloning* [23,34,35].

Table 2
Overexpressed bacterial secondary membrane transporters of various families

Protein and organism[a]	Expression host and vector	Cell culture	Protein purification	Expression level and purification yield	Activity assay	Comments
MFS (2.A.1)[b]						
AraE *Escherichia coli*	*E. coli* AD5827 (K12 derivative), pAD284	Growth to $OD_{600}=0.45$ at 33 °C in minimal medium, rapid shift to 42 °C, harvest 3 h post induction	No purification	7% of total membrane protein	Transport assay in cells	[80]
AraE *Klebsiella oxytoca*	*E. coli* MM23, pUC18	Growth in the presence of 10 mM L-arabinose as inducer	No purification	n.a.[c]	Transport assay in cells	[94]
Bmr *Bacillus subtilis*	*E. coli* NO2947, pTTQ18, *tac* promoter	n.a.[c]	Nickel–nitrilotriacetic acid (Ni^{2+}–NTA) affinity column	20% of inner membrane protein	n.a.[c]	No tight regulation of promoter, expression without inducer [37]
FucP *Escherichia coli*	*E. coli* AR120 and BL21 (DE3) pLys, pSPT19	Growth to $OD_{600}=0.4$ at 37 °C in 2xTY, induction with 40 μg/ml nalidixic acid, harvest 1–6 h post induction	n.a.[c]	20% of total inner membrane protein	Transport assay in cells	Poor expression under control of T7 promoter, only λ OLPL successful [95]
GalP *Escherichia coli*	*E. coli* JM1100, pBR322	Growth at 37 °C, no induction	Sucrose gradient, Ni^{2+}–NTA affinity and ion exchange columns	35–55% of total inner membrane protein	Reconstitution and transport assay	Expression under control of *galP* [28,36,96,97]
GlpT *Escherichia coli*	*E. coli* LMG194, pBAD	Growth 3 h at 37 °C, change to 25 °C 0.5 h before induction (0.1% arabinose, $OD_{600}=1$), harvest 1.5 h post induction at $OD_{600}=1.5–1.8$	Ni^{2+}–NTA affinity and size-exclusion columns	1.5–1.8 mg purified per liter culture	Fluorescence quenching, reconstitution and transport assay	Enhanced stability by C-terminal truncation, identified by limited proteolysis [42]
LacY (I) *Escherichia coli*	In vitro translation using T7 polymerase, pT75	–	Sucrose gradient	n.d.[d]	Reconstitution	Membrane-free translation system, co-expression with GroEL [98]
LacY (II) *Escherichia coli*	*E. coli* XL blue, cassette version of lacY (Acc X56095), *lac* promoter	Growth to $OD_{600}=0.8$ at 37 °C, induction with 0.3 mM isopropyl-β-D-thiogalactopyranoside (IPTG), harvest at $OD=2$	Ni^{2+}–NTA affinity column	1.2 mg per liter culture	Transport assay in cells	Expressed as cytochrome b_{562} fusion protein [63,81]
LacY (III) *Escherichia coli*	*E. coli* T184, pGM21 (derived from pBR322)	Growth to $OD_{420}=0.2$ at 37 °C, induction with 0.1 mM IPTG, harvest 2 h post induction at $OD_{420}=1.2–1.6$	avidin affinity column	~ 10% of total membrane protein, 2 mg purified from 20 g cells	Reconstitution and transport assay	Expressed as biotin acceptor domain fusion protein [86,99]
NorA *Staphylococcus aureus*	*E. coli* DH10B, pTrcHis2C	Growth to $OD_{600}=0.8$ at 37 °C. Induction with 0.5 m M IPTG, harvest 4 h post induction	Ni^{2+}–NTA affinity column	5–10% of total membrane protein, 0.5–1 mg purified per liter culture	Reconstitution and transport assay	[100]
OxlT *Oxalobacter formigenes*	*E. coli* XL3 (*E. coli* XL1 blue harboring pMS421), pBluescript II SK	Growth to $OD_{600}=0.2–0.3$ at 37 °C, induction with 1 mM IPTG, harvest 4 h post induction	Ni^{2+}–NTA affinity column	5% of total membrane protein, 0.5–1 mg purified per liter culture	Reconstitution and transport assay	His_{10}-tag, higher yield with C-terminal tag [71,101]

(continued on next page)

Table 2 (*continued*)

Protein and organism[a]	Expression host and vector	Cell culture	Protein purification	Expression level and purification yield	Activity assay	Comments
MFS (2.A.1)[b]						
ProP *Escherichia coli*	*E. coli* WG389 (K12 derivative), pBR322 (modified *galP* promoter)	Growth at 37 °C for 18 h to final OD_{600} = 6 in minimal medium	Ni^{2+}–NTA affinity column	30% of total membrane protein, 0.5 mg from 50 mg membrane protein	Reconstitution and transport assay	Significant reduced expression in rich medium [36,102]
TetA *Escherichia coli*	*E. coli* DH5α pET21b, co-transformed with pACT7	Growth to OD_{530} = 0.8 at 37 °C, induction with 0.1 mM IPTG, harvest 1.5–2 h post induction	Ni^{2+}–NTA affinity column	10–15% of total membrane protein, 1–1.5 mg purified from 1 l culture	Fluorescence binding	[44]
TetL *Bacillus subtilis*	*E. coli* BL21 (DE3) pLysS, pET15b	Growth 2.5 h at 30 °C, change to 20 °C 0.4 h before induction with 1 mM IPTG at OD_{600} = 1, harvest 1.8 h post induction at OD_{600} = 1.4–1.7	Co^{2+}-Talon affinity and size exclusion columns	1 mg purified after Co^{2+}-Talon and 0.2–0.3 mg after size exclusion column per liter culture	Reconstitution and transport assay	[45]
UhpT *Escherichia coli*	*E. coli* JM109, p261 (derived from pBR322), co-transformed with pMS421	Growth 3.5 h at 37 °C, change to 34 °C, induction with 1 mM IPTG, harvest 2–3 h post induction	Ni^{2+}–NTA affinity column	25 μg purified from 4 mg membrane protein	Reconstitution and transport assay	Purification in presence of 50 mM G6P, N-terminal His_{10}-tag [103]
XylE *Escherichia coli*	*E. coli* NO2947, pTTQ18	Growth at 37 °C, induction with 1 mM IPTG	n.a.[c]	20% of inner membrane protein	n.a.[c]	Highest expression with *tac* promoter, with PL expression 5–8% [28,36]
GPH (2.A.2)						
GusB *Escherichia coli*	*E. coli* NO2947, pTTQ18	Growth at 37 °C, induction with 0.6 mM IPTG	n.a.[c]	20% of inner membrane protein	n.a.[c]	[28,36]
LacS (I) *Streptococcus thermophilus*	*E. coli* NO2947, pSKE8	Growth in M9 minimal medium at 37 °C	No purification	8% of total membrane protein	n.d.[d]	Expression lower and solubilization less efficient than in *S. thermophilus* [30]
LacS (II) *Streptococcus thermophilus*	*S. thermophilus* ST11 and ST11 (ΔlacS), pGK13	Growth in Elliker Broth to OD_{660} = 0.8–1 at 42 °C, semi-anaerobic conditions, induction with lactose	Ni^{2+}–NTA affinity and ion-exchange columns	25–30% of total membrane protein, 4–5 mg purified from 1 l culture	Reconstitution and transport assay	[30]
MelB *Escherichia coli*	*E. coli* DW2-R, pKK223	Growth to OD_{600} = 2 at 30 °C in M9 minimal medium	Ni^{2+}–NTA affinity and ion-exchange columns	15% of total membrane protein, 20–30 mg purified from 10 g cells	Transport assay and fluorescence quenching of reconstituted proteoliposomes	N-terminally truncated form expressed [104,105]
XylP *Lactobacillus pentosus*	*L. lactis* NZ9000, pNZ8048	Growth to OD_{600} = 3 at 30 °C, induction with 4 ng/ml nisin, harvest 1 h post induction	Ni^{2+}–NTA affinity column	10% of total membrane protein	Reconstitution and transport assay	Fermentor was used for cell culture [32]

Table 2 (*continued*)

Protein and organism[a]	Expression host and vector	Cell culture	Protein purification	Expression level and purification yield	Activity assay	Comments
APC (2.A.3)						
GabP *Escherichia coli*	*E. coli* LMG194, pBAD	Growth to OD_{600} = 0.5–0.6 at 37 °C for 2 h, induction with 0.2% arabinose, change to 25 °C, harvest 2 h post induction at OD_{600} = 1.0–1.1	Co^{2+}-Talon affinity and size-exclusion columns	0.3–0.5 mg purified per liter culture	Fluorescence quenching upon substrate binding	Identification of a rigid protein core using limited proteolysis [41]
CSC (2.A.17)						
DtpT *Lactococcus lactis*	*L. lactis* MG1363 and AG300, dtpT	Growth in M17 broth at 28 °C, no induction, harvest in late exponential phase	Ni^{2+}–NTA affinity column	10% of total membrane protein, 2 mg purified from 100 mg membrane protein	Transport assay in cells and in reconstituted proteoliposomes	Overexpression of the protein in *E. coli* was lethal [31]
SSS (2.A.21)						
PutP *Escherichia coli*	*E. coli* WG170, pT7-5	Growth to mid log phase at 37 °C, induction with 0.3 mM IPTG, harvest 3 h post induction	Ni^{2+}–NTA affinity and ion-exchange columns	6 mg purified from 90 mg of total membrane proteins	Reconstitution and transport assay	[106]
DAACS (2.A.23)						
GltT *Bacillus stearothermophilus*	*E. coli* DH5, pBlueScript II KS	Growth to OD_{660} = 0.7 at 37 °C, harvest without induction	Ni^{2+}–NTA affinity column	0.7% of total membrane protein, 3 mg purified from 400 mg membrane protein	Reconstitution and transport assay	No IPTG induction, "leaky" expression [107]
CCS (2.A.24)						
CitS *Klebsiella pneumoniae*	*E. coli* C43(DE3), pET16b	Growth in terrific broth to OD_{600} = 0.6 at 30 °C, induction with 0.7 mM IPTG, change to 25 °C, harvest after 18–22 h at OD_{600} = 10	Ni^{2+}–NTA affinity column	6.5 mg purified per liter culture	Reconstitution and transport assay	N-terminal His_{10}-tag, lower yield in LB or with biotinylated protein [43]
AGCS (2.A.25)						
AlcP *Thermophilic bacillus* PS3	*E. coli* TB1, pMAL-c2	Growth in NZCY medium to late log phase at 37 °C, induction with 1 mM IPTG, harvest 3 h post induction	Amylose affinity column	20% of total membrane protein, 5 mg fusion protein from 50 mg membrane protein	Reconstitution and transport assay	Expressed as maltose binding protein fusion protein [108]
NhaA (2.A.33)						
NhaA *Escherichia coli*	*E. coli* TA16 (K12 derivative), pEP3T	Growth in minimal medium to OD_{600} = 0.5 at 30 °C, induction with 0.5 mM IPTG, harvest 2 h post induction	Hydroxylapatite and ion-exchange columns	438 μg from 10 mg cells	Reconstitution and transport assay	[83]

(continued on next page)

Table 2 (*continued*)

Protein and organism[a]	Expression host and vector	Cell culture	Protein purification	Expression level and purification yield	Activity assay	Comments
NhaB (2.A.34) NhaB *Escherichia coli*	*E. coli* TA15 (K12 derivative), pT7-6	Growth to $OD_{600}=0.6$ at 30 °C, change to 42 °C for 15 min, after addition of rifampicin back to 30 °C for 1 h	No purification	Band detected in autoradiography	Reconstitution and transport assay	Expression only for [^{35}S]-labeling [109]
MIT (2.A.45) CorA *Escherichia coli*	*E. coli* LMG 194, pBAD	Growth to $OD_{600}=0.6$ at 37 °C, induction with 0.1% arabinose, change to 15 °C, harvest at $OD_{600}=2.5$	Ni^{2+}–NTA affinity column	11% of total membrane protein, 10 mg purified per liter culture	n.d.[d]	Lower temperature reduced inclusion bodies and increased protein insertion into membrane

[a] Unless indicated otherwise, growth medium was Luria Broth (LB), reported protein purity was at minimum 95%, and His-tag contained six tandem histidines.
[b] The system of transporter families by Paulsen et al. [1,2] and Saier [5] was used for the classification of the transporter proteins.
[c] Not available.
[d] Not determined.

Among the members of the chosen protein family, transporters that have been biochemically and functionally characterized are worth considering, because such information may help crystallization experiments. Members of the transporter family with various loop lengths can be selected for they may provide different crystal contacts. However, it is not yet possible to predict the expression levels of a particular protein from its sequence. It is thought that proteins with multiple transmembrane α-helices tend to give lower expression than those with fewer transmembrane domains [25]. Recent studies, however, suggest that this may not be the case. Indeed, the Mg^{2+} transporter (CorA) and the mechanosensitive channel (MscL) from *E. coli*, both predicted to have two to three transmembrane α-helices, were expressed to a level of 10–15 mg protein per liter of *E. coli* culture, from which 5–10 mg protein was purified (Chen, Auer and Wang, unpublished results). For comparison, expression of several transporter proteins with as many as 12 predicted transmembrane α-helices produced levels of 20–50% of total membrane protein [28,36,37], although often only 1–2 mg protein was purified per liter of cell culture (Table 2). Therefore, high number of transmembrane helices does not seem to be a limiting factor for protein expression.

2.2. Expression vector

The vector is one of the most important factors for overexpression of membrane transporter proteins. Whether a protein target expresses in *E. coli* or not largely depends on the vector. Ideally, the vector should have a tightly regulated, moderately strong promoter, and should have a wide range of usable inducer concentrations. Tight regulation prevents leaky expression, which can lead to in vivo proteolysis, or even cell death when expressing toxic membrane proteins like certain amino acid transporters. Very strong promoters tend to result in inclusion body formation [38,39]. On the other hand, a wide range of applicable inducer concentrations allows better control of expression. Finally, the inducer should not be prohibitively expensive, making scaling-up impossible. A number of expression vectors meet the above criteria and have been successfully used for overexpressing membrane transporter proteins (Table 2).

The pBAD vectors from Invitrogen, containing the arabinose araBAD operon, are particularly suitable for overexpression of membrane proteins [40]. The promoter is repressed 1200-fold, which allows tight regulation of expression. It is moderately strong, and can be induced by L-arabinose at concentrations between 0.00002% and 0.2%. This expression system has been used to overexpress a number of membrane proteins to a yield of 1 mg purified protein per liter of culture or higher. These include the mechanosensitive channel MscL, the GABA transporter GabP [41], the G3P transporter GlpT [42], and the magnesium transporter CorA (Chen and Wang, unpublished results). These proteins belong to four different protein families (Table 2), demonstrating the general applicability of the pBAD system.

Similarly, overexpression of several membrane transporters has been achieved with the pET vector series [43–45]. The gene expression in this system is under the control of the T7 promoter, which is not recognized by *E. coli* RNA polymerase. As a result, almost no expression occurs until T7 RNA polymerase is provided with the vector [46]. The entire vector series includes 36 pET vectors that are tailored for different expression levels and have a variety of tags for protein purification and detection. However, because the T7

system has a relatively strong promoter, overexpression of membrane proteins with the system sometimes causes bacterial cell death [47].

The pTTQ18 vector and its variants were used by Ward et al. [28,36,37] to overexpress a number of prokaryotic membrane transporters for sugars and sugar derivatives in *E. coli* (Table 2). Almost 100% of the 16 transporters studied were expressed at levels of 20–50% of total membrane protein and purification yields of 1–2 mg per liter of culture were obtained. Protein expression in this system is under the control of the *tac* promoter. Interestingly, several of these transporter proteins that were expressed using the *tac* promoter did not express with a T7 promoter [36]. The pTTQ18 vector, however, is not a tightly controlled system, and IPTG-independent expression occurs in the stationary phase [28].

Most of the currently available expression vectors, including those mentioned above, are constructed with affinity purification tags. The 5′ end of the coding region is believed to be important for effective initiation of translation, and thus, for expression [25]. The position of the purification tags is therefore critical. The target gene can be cloned into a vector in two different constructs that have the tags at either end of the protein. If neither N- nor C-terminal tagging leads to reasonable expression, the first 10 amino acids of the target protein may be replaced with those from either a homologue or an anthologue for that high expression is already established.

2.3. Host strain

Because expression levels can vary as much as 2- to 5-fold between strains [42], several host strains should be tested during initial trials. Several *E. coli* strains are compatible with the pET and pQE vectors, including those deficient in proteases or with extra codons. By switching from *E. coli* BL21(DE3) strain to C43(DE3), a strain originally selected for it high capacity of producing mitochrondrial membrane proteins [47], Kästner et al.[43] were able to increase the protein yield for the Na^+-dependent citrate carrier of *Klebsiella pneumoniae* from 1 to 6.5 mg per liter of bacterial culture (Table 2). Another strain worth considering is the methionine auxotroph B834. It allows efficient labeling of the target protein with selenomethionine—a feature that facilitates structure determination by the multiwavelength anomalous diffraction method.

The pBAD vector requires the host cell to be deficient in arabinose metabolism. The *E. coli* LGM194 strain has the compatible phenotypes, but is highly susceptible to bacteriophage infection, particularly to the T1 phage, the most common type of bacteriophages found in a research laboratory. Alternatively, BL21–pLysS may be used as the host strain with the pBAD vector. This strain is more resistant to T1 phage infection but is not deficient in arabinose metabolism. If the time for post-induction cell growth does not exceed 2–3 h, the expression is not affected, because the arabinose levels only decrease significantly after 6 h. Another advantage of the BL21–pLysS strain is that, when used with the T7 vector system, it produces lysozyme that inhibits leaky transcription during cell culture and facilitates cell breakage at later membrane preparation stages.

When expressing proteins from thermophilic bacteria in *E. coli*, host strains that carry plasmids encoding tRNAs that recognize rare codons should be considered, because thermophilic archaebacteria use rare codons for translation [48]. Knol et al. [30] found that LacS, the lactose transporter from *Streptococcus thermophilus*, expressed at lower levels in *E. coli* than in *S. thermophilus* (Table 2). Introducing additional tRNAs for rare codons into *E. coli* can increase the expression levels of archaea proteins [49]. Codon usage between *Thermoplasma acidophilum* and *E. coli*, for example, differs in one major aspect: the former uses the two arginine-coding codons, AGG and AGA, whereas *E. coli* uses them less frequently. These same codons are also frequently used by eukaryotic organisms [50]. The Codon-Plus *E. coli* strain from Stratagene carries plasmids coding for these two tRNA codons, and may be useful for expressing *T. acidophilum* proteins. The BL21–Codon-Plus–RIL strain overexpresses tRNAs for leucine and isoleucine, and may be suitable for expressing proteins with multiple transmembrane fragments rich in these two amino acids.

2.4. Colony selection

Colonies need to be selected for large-scale expression experiments, because expression levels can vary substantially among colonies of transformed cells. Based on a colony blot protocol from Qiagen [51], we have developed a simple and systematic approach for colony selection. Colonies on a master plate containing freshly transformed cells are transferred to a nitrocellulose membrane. The colonies bound to the nitrocellulose are then placed on an LB plate containing the inducer. The cells are lysed on the nitrocellulose after a few hours of growth, followed by protein denaturation and binding to the membrane. The protein of interest is then probed with the appropriate antibody or conjugate, for example, Ni^{2+}–horseradish peroxidase (HRP) probe, on the nitrocellulose membrane (Fig. 1). High-expressing colonies are then selected from the original plate for large-scale expression. At expression levels above 0.1 mg/l, the signal from the target protein is far stronger than the background signal caused by the binding of the Ni^{2+}–HRP probe to histidine-rich endogenous *E. coli* proteins. Therefore, the selection of the high-expression colonies is not affected by the weak false-positive signal and the method is simple and reliable.

2.5. Cell culture

Besides the vector, cell culture conditions are the most important factors affecting expression of membrane proteins. Expression levels depend on the culture medium, cell

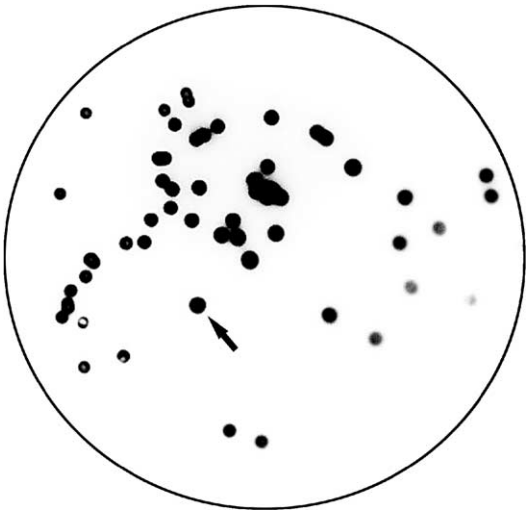

Fig. 1. Analysis of homologous expression level of the glycerol-3-phosphate (G3P) transporter (GlpT) in *E. coli*. A Ni^{2+}–HRP Western blot of a colony-blot preparation revealed various GlpT expression levels. Colonies with the highest signals gave a purification yield of 1.5 mg/l culture. Note that all colonies transferred to the nitrocellulose membrane were of similar size.

density at induction and harvesting, inducer concentration and the temperature profile, particularly, the post-induction time and temperature. Optimization of cell culture and induction conditions can increase the expression level 3- to 5-fold [41,42]. A cell growth curve can be used to guide the optimization process. Induction is often carried out at the mid log phase of cell growth, and cells are harvested before they reach the stationary phase. A rich medium typically leads to higher cell density and more protein expressed. On the other hand, prolonged post-induction cell growth can result in proteolytic degradation. For inducible promoters, it is often beneficial to screen a wide concentration range to find the optimal inducer concentration. Ultimately, the protein should be expressed at a level that allows its detection by Coomassie blue-stained sodium dodecyl sulfate (SDS)–polyacrylamide gel electrophoresis (PAGE) of membrane samples [41,42]. If the initial expression level is below 0.1 mg/l cell culture, one should consider switching to another vector, which may result in higher expression levels.

Optimizing cell culture conditions can also reduce inclusion body formation—which often occurs when the promoter is strong and the machinery for membrane protein translation and insertion becomes oversaturated [25]. Although it is possible to solubilize the inclusion bodies with high concentrations of urea or stronger chaotropic reagents and subsequently refold the membrane protein in vitro, the final yield of refolded membrane protein is often very low [52]. It is therefore best to avoid the formation of inclusion bodies in the first place, by optimizing cell culture conditions.

Inclusion body formation depends on the rate of protein synthesis and the rate of protein folding [53]. By decreasing the cell growth temperature, the formation of inclusion bodies for a number of soluble proteins expressed in bacteria was reduced [54]. This approach was recently applied to expression studies of the magnesium transporter CorA in the pBAD/LMG194 system (Chen and Wang, unpublished results). Sixty milligrams of CorA protein per liter of *E. coli* culture was produced at 37 °C. However, most of the protein was trapped in inclusion bodies, with only 3 mg found in the membrane fraction (Fig. 2). Reducing the post-induction temperature to 15 °C essentially eliminated inclusion body formation. Approximately 15 mg CorA/l cell culture was found in the membrane fraction, and of this, 10 mg was purified using Ni^{2+}–NTA affinity chromatography.

Unless one is certain there is no inclusion body formation in the expression system, a low-speed centrifugation step following cell breakage is necessary [55]. Small amounts of unfolded membrane protein can act as seeds for aggregation

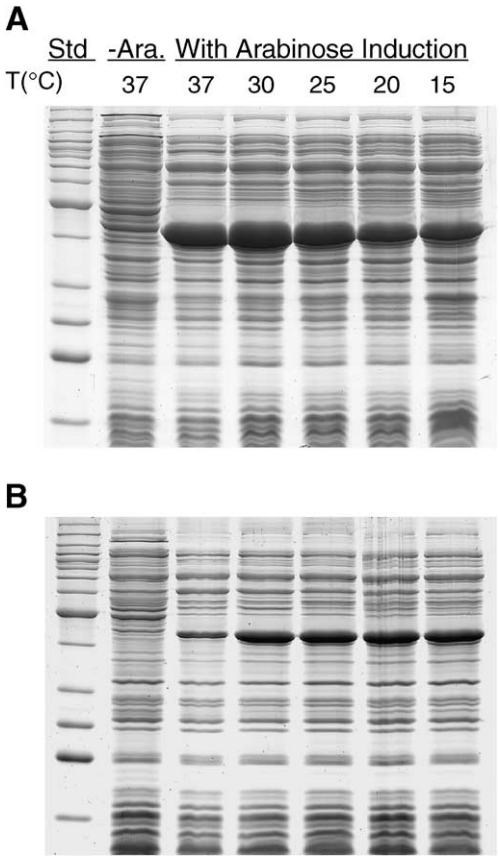

Fig. 2. Effect of post-induction cell growth temperature to homologous expression of magnesium transporter (CorA) from *E. coli*. *E. coli* LMG194 cells carrying the CorA gene in the pBAD vector were grown at 37 °C to OD_{600} ~ 0.6, followed by induction of gene expression with 0.01% arabinose. Cells cultured at different post-induction temperatures were all grown to OD_{600} ~ 2.5. Following cell breakage and fractionation, samples were analyzed by Coomassie blue-stained SDS-PAGE. (A) Total cell lysate. (B) Solubilized membrane. The total amount of CorA protein produced, about 60 mg/l *E. coli* culture, decreased only slightly at lower post-induction temperature. Lowering the post-induction temperature, however, resulted in significantly more CorA insertion into the membrane (~ 15 mg/l culture).

and, therefore, hinders crystallization. A low-speed centrifugation step prevents contamination by inclusion bodies in subsequent membrane preparations, and thus minimizes co-purification of unfolded polypeptide along with the properly folded protein [42].

3. Factors important for protein quality

Protein quality is as important as quantity for structural studies, and a system with the highest expression levels may not produce the best quality protein for crystallization [42,56]. Therefore, once an expression system is established that produces reasonable amounts of the membrane transporter protein, the protein construct, cell culture conditions, and purification proteins should all be optimized. At this point, the protein's stability and functionality should be also analyzed to prepare for crystallization experiments.

3.1. Protein construct

The most important parameter in protein crystallization is the protein itself [57]. Flexible termini and loops in a protein often hinder the formation of high-quality crystals. After the removal of such flexible termini, the protein core tends to crystallize more readily. As a result, it is now a standard procedure for improving crystal quality by identifying the protein core and redesigning the protein construct accordingly [22,57,58].

Mass spectrometry, in combination with limited proteolysis, provides a useful tool for identifying polypeptides. Previously, the technique's application in the analysis of membrane proteins was limited [59]; the presence of detergent and lipid in the membrane protein sample precluded accurate molecular mass measurements [28]. Recently, Cadene and Chait [60] introduced a wash step with trifluoroacetic acid that removes lipid and detergent from the protein. This method allowed them to determine molecular mass of 20–50 kDa membrane proteins with an accuracy of ± 2–4 Da using matrix-assisted laser desorption/ionization time-of-flight (MALDI-TOF) mass spectrometry, and thus, to identify membrane proteins unequivocally. The rigid core of both the GlpT (Fig. 3) and GabP proteins was identified using this procedure, and their plasmids redesigned and recloned accordingly [41,42]. The resulting GlpT protein core was resistant to proteases and retained full activity [42,61]. One limitation of this entire procedure, however, is that it cannot easily identify and eliminate flexible loops in the middle of the protein sequence.

At the protein level, one can increase its solvent-accessible surface areas available for making crystal contacts by conformation-specific antibody binding [62] or by fusing a soluble protein into the membrane protein sequence as a 'carrier' for crystallization. Cytochrome b_{562}, an easily crystallizable copper-binding protein, and other soluble proteins of various sizes were inserted into loops of the

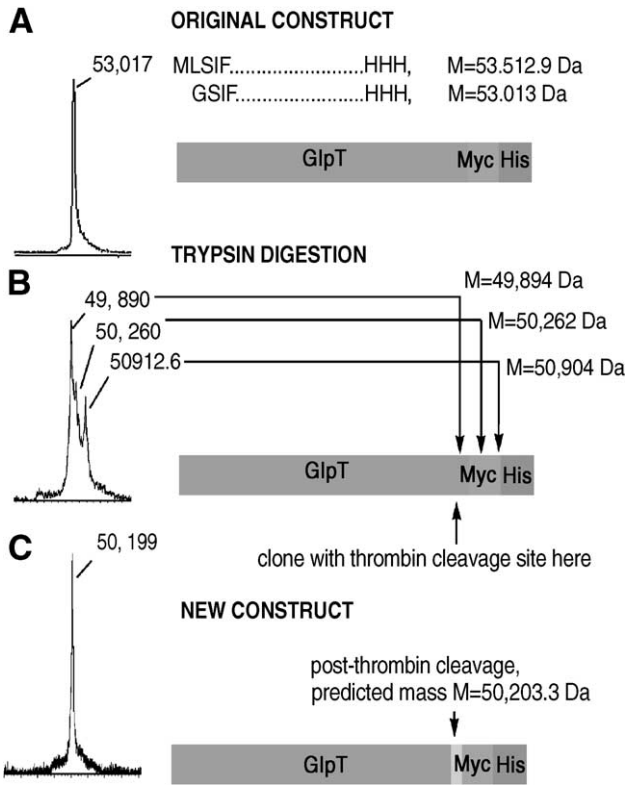

Fig. 3. Identification of flexible termini and optimization of GlpT constructs. (A) Initially, the wild-type GlpT was expressed in the pBAD vector, with a *myc*-epitope and a His_6-tag fused to its C terminus. In addition to the Leu2 → Gly2 mutation introduced to facilitate cloning, the N-terminal methionine (Met1) was found to be proteolytically removed in vivo. The molecular weight of the expressed protein, therefore, was calculated to be 53,013. This was confirmed by MALDI-TOF mass spectroscopy measurements, which yielded a value of 53,017. (B) Ni^{2+}–NTA purified wild-type GlpT protein was subjected to tryptic digestion at different enzyme-to-protein ratios for various periods of time. Three digestion products ending with residue Arg449, Lys453 and Lys459, respectively, were identified by SDS-PAGE followed by mass spectroscopy. (B) A new DNA construct was then engineered, containing residues 1–448, the protein core, followed by a thrombin-specific proteolytic and the *myc*-His-tags.

Lac permease (LacY) from *E. coli* [63,64]. Several of the fusion proteins retained their ability to transport lactose. However, it is still unknown if these fused proteins will yield crystals more readily. A potential problem is the issue of flexibility between the target and 'carrier' protein which may hinder high-quality crystal formation.

3.2. Cell culture

The importance of cell culture conditions on protein quality and the potentials for their optimization are illustrated by the studies on GlpT [42]: initially, GlpT protein expression level was 8 mg/l culture, which yielded 4 mg purified protein. However, protein microheterogeneity, presumably due to proteolytic digestion and post-translational modification, was detected by SDS-PAGE of purified samples (Fig. 4). In addition, two proteins co-purified with

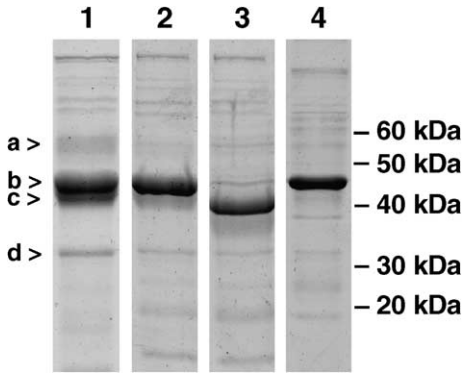

Fig. 4. Optimization of the *E. coli* cell culture for GlpT overexpression. The protein core of GlpT expressed in pBAD vector migrated on SDS-PAGE to a position corresponding to a molecular weight of 45,000, as indicated by letter 'b'. Lane 1: When the *E. coli* cells were cultured at 37 °C to OD_{600} of 1.0 and then induced with 0.2% arabinose, growth at 37 °C for an additional 2 h to OD_{600} of 2.3 produced a number of contamination protein bands that could not be isolated from GlpT by Ni^{2+}–NTA or size-exclusion column. Two unidentified proteins at 58 and 32 kDa, labeled by letters 'a' and 'd', co-purified with GlpT on the Ni^{2+}–NTA resin and could not be separated from the transporter protein by a subsequent size-exclusion or ion-exchange column. The protein bands indicated by letter 'c' consisted of several proteolyzed GlpT fragments and a broad 'shadow band,' appearing 2–3 kDa below the GlpT protein. Lane 2: Reducing the arabinose concentration for induction from 0.2% to 0.1% and lowering the temperature after induction to 25 °C minimized the expression of the 58- and 32-kDa proteins. Changing the induction OD_{600} from 1.0 to 0.5 and shortening the post-induction period from 2 to 1.5 h markedly reduced the proteolysis fragments. Lane 3: Following thrombin digestion, both GlpT and the shadow band shifted downwards by the same distance, suggesting that the shadow band was either a post-translational modification or an N-terminal truncation of GlpT. Lane 4: Lowering the temperature from 37 to 25 °C half hour before induction resulted in essentially the complete disappearance of the shadow band. The OD_{600} at cell harvest was 1.5. All protein samples were purified by Ni^{2+}–NTA affinity chromatography before SDS-PAGE analysis (reproduced from Ref. [42] with permission).

GlpT on Ni^{2+}–NTA affinity column that could not be separated from the transporter by an additional chromatography step. Purification with a third chromatography column results in protein aggregation due to delipidation, as often occurs with polytopic membrane proteins [65,66]. However, multiple cycles of cell culture condition optimization, including inducer concentration and the time and temperature for cell growth, solved most of these problems and resulted in 1.8 mg of purified GlpT per liter of *E. coli* culture [42].

3.3. Protein homogeneity

A protein is more likely to crystallize if it is in a homogeneous state [57]. Heterogeneities are either conformational or biochemical in nature. Conformational heterogeneity is related to protein flexibility and multiple conformations required for the protein's function. Besides protein impurity and variations in polypeptide sequence, biochemical heterogeneity is also related to detergent and co-purified lipid. While the protein purity and microheterogeneity can be readily detected using SDS-PAGE and mass spectrometry [42,61], the detergent and co-purified lipid should also be monitored for crystallization experiments.

Conformational heterogeneity due to intrinsic structural flexibility of membrane transporter proteins may interfere with the crystallization process. In contrast to proteins whose reaction cycle involves only the prosthetic groups bound [67,68], transporter proteins are highly flexible. This is particularly true for transporters that translocate large solute molecules across the membrane. Studies on the Lac permease, a prototype for 12 α-helices membrane transporters, illustrate the intrinsic structural flexibility of membrane transporters [69]. Upon substrate binding, the distance between helices II and VII of the protein changed by more than 4 Å. The average tilt angle of helices from the membrane normal, measured by Fourier transform infrared spectroscopy, changed from 33° to 51° with a decreased lipid-to-protein ratio [70]. The long loop connecting helices 6 and 7 in GlpT also displayed conformational changes upon substrate binding [61]. Conformational heterogeneity due to structural flexibility can be reduced by substrate and inhibitor binding. Using circular dichroism spectroscopy of purified protein, Fu and Maloney [71] observed markedly increased structural stability of the *E. coli* oxalate transporter (OxlT) in the presence of substrate in a detergent/lipid micellar environment. Similarly, the binding of GlpT to its substrate, G3P, locked the transporter into a particular conformation, protecting the long central loop of the protein from tryptic digestion [61].

Another source of biochemical heterogeneity comes from free, unbound detergent in solution, which may prevent proper packing of protein–detergent micelles in the crystal lattice. Detergent concentration, therefore, needs to be kept at a minimum that maintains protein solubility. For low critical micellar concentration (CMC) detergents like dodecylmaltoside (DDM) or octyloxyethylene dodecylether ($C_{12}E_8$), a final concentration of 0.1–0.2% appears to be appropriate for a protein concentration of 5–10 mg/ml during crystallization. This detergent concentration can be achieved by careful control of the detergent concentration used in the last purification column, followed by a concentration step using filters with an appropriate molecular weight cut-off.

Lipids typically co-purify with membrane proteins, leading to additional biochemical heterogeneity. On the other hand, complete delipidation of membrane transporter proteins often causes denaturation and aggregation [42,65,66]. It is therefore necessary to identify the lipids that are critical for the protein's stability and subsequently to control their composition in the purified sample [65]. Following extraction using organic solvent, phospholipids co-purifying with the protein can be identified by two-dimensional thin-layer chromatography, and their concentrations measured by phosphate analysis [65,66,72]. Analysis of protein samples purified with two successive

chromatography columns showed that five to eight phospholipid molecules typically co-purify with one transporter molecule of 12–14 transmembrane α-helices [65,66,72]. Lipid type and concentration were critical for two- and three-dimensional crystal formation of several membrane proteins [66,73].

3.4. Protein stability and monodispersity

Crystallization requires a protein to be stable and monodisperse at 4 or 20 °C over several days or weeks [74–76]. Due to their surface duality, being both hydrophilic and hydrophobic, membrane proteins tend to aggregate and precipitate quickly in the presence of unfavorable detergents, precluding crystallization. This is a challenge especially for transporter proteins because of their high number of multiple transmembrane spans and larger areas of hydrophobic surface [65].

While each individual protein demonstrates unique sensitivities to its environment, detergent, lipid, pH, and temperature have been shown to affect the stability and monodispersity of many transporter proteins. In addition, glycerol at concentrations of 20% is highly effective in increasing the stability of membrane transporters [27,41,42,65,77], in agreement with its stabilizing effect for soluble proteins [78]. Using analytical size-exclusion chromatography on high-performance liquid chromatography (HPLC) [79], a large number of conditions can be screened efficiently. Those conditions that keep the protein stable and monodisperse can be applied to both purification and crystallization.

Using the above procedure, the GlpT protein was found to be monomeric between pH 4.5 and 8.5, and in 9 of the 15 detergents tested (Fig. 5) [42]. In contrast, the GlpT homologue from *Haemophilus influenzae*, with a 56% sequence identity to the *E. coli* protein, was not stable in DDM (Lemieux, Song and Wang, unpublished data), a mild detergent routinely used for solubilization and purification of membrane transporters [27,28,41,42,66]. The *E. coli* GABA transporter stayed monodisperse in two detergents, DDM and Fos-Choline-12 [41]. The glutamate transporter (GltT) from *Bacillus stearothermophilus*, in contrast, proved to be an extremely stable membrane protein. It remained dimeric from pH 4 to 9 in a number of detergents, including DDM, undecylmaltoside (UDM), decylmaltoside (DM), nonylmaltoside (NM), cyclohexyl-hexylmaltoside (Cymal-6), cyclohexyl-pentylmaltoside (Cymal-5), Fos-Choline-12, Fos-Choline-10, MEGA-10, OG and lauryldimethylamine oxide (LDAO) (Lemieux, Slotboom, Lolkema and Wang, unpublished results). Furthermore, after incubation at 37 °C for 2 h, no aggregation or change in oligomeric state was detected. Such unusual stability is probably related to the fact that the protein is from a thermophilic bacterium. It remains to be seen whether other transporter proteins from thermophilic organisms are equally stable.

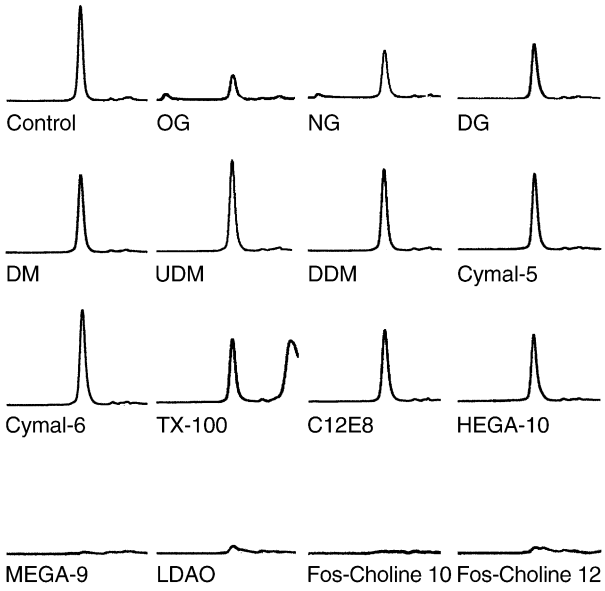

Fig. 5. Monodispersity of the GlpT protein in the presence of different detergents. The control was kept at 4 °C. All other samples were incubated with different detergents for 2 h at 25 °C, followed by analytical size-exclusion chromatography on HPLC. Eluted as single peak, the protein remained monomeric and monodisperse in a number of detergents: decylglucoside (DG), decylmaltoside (DM), undecylmaltoside (UDM), dodecylmaltoside (DDM), Cymal-5, Cymal-6, Triton X-100, $C_{12}E_8$ and HEGA-10 (reproduced from Ref. [42] with permission).

3.5. Protein activity

Activity and integrity of the expressed transporter should be monitored to ensure that the structure determined is of physiological relevance. Functionality can be studied in membrane and detergent solution, each revealing properties of the protein in different environments. Measuring transporter activity in whole bacterial cells provides a direct and relatively quick activity assay. Activity for a particular substrate is measured before and after induction of protein expression, and the difference in transport activity is correlated to the functionality of the expressed protein [31,63,80,81]. This approach, however, often does not yield reliable transport kinetics of the expressed protein. This is due most likely to: low-level leaky expression, the difficulties in estimating the number of transporter molecules expressed per cell, or possible multiple transport systems for a particular substrate. *E. coli*, for example, has four transport systems for inorganic phosphate (P_i) [17]. In addition, at high expression levels, the total substrate flux in cells can be so high, that the transport assay becomes oversaturated. The main advantage of the whole-cell measurement techniques therefore lies in rapid evaluation of the correct folding and functionality of the transporter in expression experiments.

More reliable measurement of the transport kinetics requires reconstitution of the purified protein into proteoliposomes, followed by transport activity assays [28]. These experiments have been conducted for many transporters

(Table 2). Typically, reconstitution of the transporter is achieved by mixing preformed lipid vesicles and purified protein in the presence of detergent, followed by detergent removal with Bio-Beads [82]. Substrate transport across the membrane of the proteoliposomes is measured using an enzymatic or radioactivity assay. These measurements yield the V_{max}, K_d, turnover rate, and other parameters of the transport kinetics [30,42,83].

The activity of the transporter in whole cell or proteoliposomes only reflects its properties in a lipid environment—but it is the functionality of the transporter in detergent solution that is particularly useful to protein crystallization. Direct measurement of the transport activity of membrane transporters in solution is impossible because of the vectorial nature of the transport process and the isotropy of the solubilized protein. Nonetheless, substrate binding to the solubilized protein can be studied in detergent solution, using various biophysical and spectroscopic techniques. Fluorescence quenching upon substrate binding was used to study GalP, GlpT, and GabP proteins (Table 2) [41,42,84], and the results indicate that the proteins retain their substrate-binding capability in detergent solution. Caution, however, may be necessary when drawing conclusions from the binding constants obtained from these experiments. Transporter proteins, even those with a single substrate-binding site, can adopt at least two conformations, with the binding site facing either side of the membrane [85]. In solution, the protein sample consists of a population of mixed conformations, and therefore the substrate can bind from either side, typically with different affinities. Measurements in solution cannot easily separate the two, unless the substrate binds to one conformation with a far greater affinity than it does to the other. This problem can be overcome by using whole-cell measurements, or using inside-out and right-side-out vesicles [86,87]. While a complete picture of a transporter's activity can only be derived from a combination of different functional assays, the protein functionality should be demonstrated with at least one experimental technique before crystallization.

4. Summary and outlook

The emphasis of membrane transporter overexpression for structural studies should be as much on the quality of the protein as on its quantity. Optimization, therefore, is required at several stages of expression and purification. The following steps are particularly important: identification and construction of a rigid protein core, optimization of cell culture conditions to improve protein quantity and quality, and screening for conditions that maintain protein stability and monodispersity. In addition, the issues raised in this review will probably apply to other types of prokaryotic membrane proteins.

Many of the factors discussed here may also help investigators to overcome difficulties with overexpressing transporter proteins from mammalian sources. To date, efforts to overexpress mammalian membrane proteins in bacterial systems have brought little success. One major reason is that we do not fully understand several key steps in the transcription, translation and membrane insertion processes for the *E. coli* system. For example, we know that the messenger RNA sequence and stability are critical for protein expression [88,89], but we cannot predict where secondary RNA structures are likely to form that may cause translation termination. Nor can we predict which sequences are particularly susceptible to degradation by nucleases. Although we know that the formation of inclusion bodies is a result of oversaturation of the cell's translation and membrane insertion machinery, we do not yet know which steps are responsible. It is not clear if the problem is related to solubility in the cytosol, membrane targeting, protein insertion into the membrane, or all combined. Can one coexpress the membrane protein with chaperones to increase the protein expression levels? If so, which chaperones should one choose? Finally, the total amount of membrane in a bacterial cell is limited. What is the upper limit for expression levels of fully folded membrane proteins? Once these basic cellular processes are better understood, we should be able to overexpress almost any bacterial membrane protein, perhaps at higher levels, and most importantly, we may achieve the ultimate goal of producing mammalian membrane proteins in *E. coli* for structural studies.

Acknowledgements

We are grateful to Ms. Yun Lu and Dr. Thomas Neubert for mass spectrometry measurements, and to Dr. Regina Goetz for critical reading of the manuscript. Other current and former members of our group who contributed to the work described in this review include Manfred Auer, Jonathan Boulter, Colleen Gownley, Myong Jin Kim, Jinmei Song and Anthony Villa. The research in the authors' laboratory was supported by grants from the NIH (RO1 DK53973, RO1 GM052837 and R21 DK60841), the New York State Department of Health Diabetes Research Bridging Fund Program, and the Robert Leet and Clara Guthrie Patterson Trust.

References

[1] I.T. Paulsen, M.K. Sliwinski, M.H. Saier Jr., J. Mol. Biol. 277 (1998) 573–592.
[2] I.T. Paulsen, L. Nguyen, M.K. Sliwinski, R. Rabus, M.H. Saier Jr., J. Mol. Biol. 301 (2000) 75–100.
[3] J.C. Venter, M.D. Adams, E.W. Myers, P.W. Li, et al., Science 291 (2001) 1304–1351.
[4] International Human Genome Sequencing Consortium, Nature 409 (2001) 860–921.
[5] M.H. Saier Jr., J. Cell. Biochem., Suppl. (1999) 84–94.

[6] M.H. Saier Jr., J.T. Beatty, A. Goffeau, K.T. Harley, W.H. Heijne, S.C. Huang, D.L. Jack, P.S. Jahn, K. Lew, J. Liu, S.S. Pao, I.T. Paulsen, T.T. Tseng, P.S. Virk, J. Mol. Microbiol. Biotechnol. 1 (1999) 257–279.
[7] B.B. Kahn, I.A. Simpson, S.W. Cushman, J. Clin. Invest. 82 (1988) 691–699.
[8] B.B. Kahn, J. Clin. Invest. 89 (1992) 1367–1374.
[9] P.R. Shepherd, B.B. Kahn, N. Engl. J. Med. 341 (1999) 248–257.
[10] J.Y. Chou, D. Matern, B.C. Mansfield, Y.T. Chen, Curr. Mol. Med. 2 (2002) 121–143.
[11] Y. Kanai, C.P. Smith, M.A. Hediger, FASEB J. 7 (1993) 1450–1459.
[12] D.J. Rossi, T. Oshima, D. Attwell, Nature 403 (2000) 316–321.
[13] D.N. Volkow, G.-J. Wang, M.W. Fischman, R.W. Foltin, F.W. Fowler, N.N. Abmurad, S. Vilkun, J. Logan, S.J. Gatley, N. Pappas, R. Shea, C.E. Shea, Nature 386 (1997) 827–830.
[14] S.B. Levy, Lancet 358 (2001) 1100–1111.
[15] S.B. Levy, J. Antimicrob. Chemother. 49 (2002) 25–30.
[16] P.J.F. Henderson, Curr. Opin. Cell Biol. 5 (1993) 708–712.
[17] P.C. Maloney, S.V. Ambudkar, V. Anatharam, L.A. Sonna, A. Varadhachary, Microbiol. Rev. 54 (1990) 1–17.
[18] H.R. Kaback, J. Wu, Q. Rev. Biophys. 30 (1997) 333–364.
[19] S.C. King, S.R. Fleming, C.E. Brechtel, J. Biol. Chem. 270 (1995) 19893–19897.
[20] T.H. Wilson, P.Z. Ding, Biochim. Biophys. Acta 1505 (2001) 121–130.
[21] M. Quick, E.M. Wright, Proc. Natl. Acad. Sci. U. S. A. 99 (2002) 8597–8601.
[22] D.A. Doyle, J. Morais Cabral, R.A. Pfuetzner, A. Kuo, J.M. Gulbis, S.L. Cohen, B.T. Chait, R. MacKinnon, Science 280 (1998) 69–77.
[23] G. Chang, C.B. Roth, Science 293 (2001) 1793–1800.
[24] D.V. Goeddel, Methods in Enzymology, vol. 185, Academic Press, San Diego, 1990, pp. 1–681.
[25] R. Grisshammer, C.G. Tate, Q. Rev. Biophys. 28 (1995) 315–422.
[26] G.W. Gould, Membrane Protein Expression Systems: A User's Guide, Portland Press, London, 1994, pp. 1–306.
[27] B. Poolman, J. Knol, Biochem. Soc. Trans. 27 (1999) 912–917.
[28] A. Ward, N.M. Sanderson, J. O'Reilly, N.G. Rutherford, B. Poolman, P.J.F. Henderson, in: S.A. Baldwin (Ed.), Membrane Transport, Oxford Univ. Press, Oxford, 2000, pp. 141–166.
[29] C.G. Tate, FEBS Lett. 504 (2001) 94–98.
[30] J. Knol, L. Veenhoff, W.J. Liang, P. Henderson, G. Leblanc, B. Poolman, J. Biol. Chem. 271 (1996) 15358–15366.
[31] A. Hagting, J. Knol, B. Hasemeier, M.R. Streutker, G. Fang, B. Poolman, W.N. Konings, Eur. J. Biochem. 247 (1997) 581–587.
[32] E.H. Heuberger, E. Smits, B. Poolman, J. Biol. Chem. 276 (2001) 34465–34472.
[33] F. Baneyx, Curr. Opin. Biotechnol. 10 (1999) 411–421.
[34] G. Chang, R.H. Spencer, A.T. Lee, M.T. Barclay, D.C. Rees, Science 282 (1998) 2220–2226.
[35] K.P. Locher, A.T. Lee, D.C. Rees, Science 296 (2002) 1091–1098.
[36] A. Ward, J. O'Reilly, N.G. Rutherford, S.M. Ferguson, C.K. Palmer, S.L. Palmer, J.L. Clough, H. Venter, H. Xie, G.J. Martin, G.E. Martin, J.M. Wood, P.E. Roberts, M.A. Groves, W.J. Liang, A. Steel, B.J. McKeown, P.J. Henderson, Biochem. Soc. Trans. 27 (1999) 893–899.
[37] A. Ward, C. Hoyle, S. Palmer, J. O'Reilly, J. Griffith, M. Pos, S. Morrison, B. Poolman, M. Gwenne, P. Henderson, J. Mol. Microbiol. Biotechnol. (2001) 193–200.
[38] M.J. Weickert, D.H. Doherty, E.A. Best, P.O. Olins, Curr. Opin. Biotechnol. 7 (1996) 494–499.
[39] H. Lilie, E. Schwarz, R. Rudolph, Curr. Opin. Biotechnol. 9 (1998) 497–501.
[40] L.M. Guzman, D. Belin, M.J. Carson, J. Beckwith, J. Bacteriol. 177 (1995) 4121–4130.
[41] X.D. Li, A. Villa, C. Gownley, M.J. Kim, J.M. Song, M. Auer, D.N. Wang, FEBS Lett. 494 (2001) 165–169.
[42] M. Auer, M.J. Kim, M.J. Lemieux, A. Villa, J. Song, X.D. Li, D.N. Wang, Biochemistry 40 (2001) 6628–6635.
[43] C.N. Kästner, P. Dimroth, K.M. Pos, Arch. Microbiol. 174 (2000) 67–73.
[44] M.L. Aldema, L.M. McMurry, A.R. Walmsley, S.B. Levy, Mol. Microbiol. 19 (1996) 187–195.
[45] J. Cheng, D.B. Hicks, T.A. Krulwich, Proc. Natl. Acad. Sci. U. S. A. 93 (1996) 14446–14451.
[46] F.W. Studier, B.A. Moffatt, J. Mol. Biol. 189 (1986) 113–130.
[47] B. Miroux, J.E. Walker, J. Mol. Biol. 260 (1996) 289–298.
[48] S. Kanaya, Y. Yamada, Y. Kudo, T. Ikemura, Gene 238 (1999) 143–155.
[49] T. Wakagi, T. Oshima, H. Imamura, H. Matsuzawa, Biosci. Biotechnol. Biochem. 62 (1998) 2408–2414.
[50] J.F. Kane, Curr. Opin. Biotechnol. 6 (1995) 494–500.
[51] J.F. Qiagen, QIAgen Detection and Assay Handbook, Qiagen, Hilden, Germany, 2001.
[52] H. Kiefer, K. Maier, R. Vogel, Biochem. Soc. Trans. 27 (1999) 908–912.
[53] T. Kiefhaber, R. Rudolph, H.H. Kohler, J. Buchner, Biotechnology (NY) 9 (1991) 825–829.
[54] C.H. Schein, M.H.M. Noteborn, Bio/Technology 6 (1988) 291–294.
[55] J. Sambrook, E.F. Fritsch, T. Maniatis, Molecular Cloning, 2nd ed., A Laboratory Manual, vol. 3, CSHL Press, New York, 1989.
[56] C. Luna-Chavez, T.M. Iverson, D.C. Rees, G. Cecchini, Protein Expr. Purif. 19 (2000) 188–196.
[57] P.D. Kwong, R. Wyatt, E. Desjardins, J. Robinson, J.S. Culp, B.D. Hellmig, R.W. Sweet, J. Sodroski, W.A. Hendrickson, J. Biol. Chem. 274 (1999) 4115–4123.
[58] L. Wei, S.R. Hubbard, W.A. Hendrickson, L. Ellis, J. Biol. Chem. 270 (1995) 8122–8130.
[59] J.P. Whitelegge, J. le Coutre, J.C. Lee, C.K. Engel, G.G. Prive, K.F. Faull, H.R. Kaback, Proc. Natl. Acad. Sci. U. S. A. 96 (1999) 10695–10698.
[60] M. Cadene, B. Chait, Anal. Chem. 72 (2000) 5655–5658.
[61] J.M. Lemieux, J.M. Song, D.N. Wang, 2002, submitted for publication.
[62] C. Ostermeier, S. Iwata, B. Ludwig, H. Michel, Nat. Struct. Biol. 2 (1995) 842–846.
[63] G.G. Prive, G.E. Yerner, C. Weitzman, K.H. Zen, D. Eisenberg, H.R. Kaback, Acta Crystallogr., D 50 (1994) 375–379.
[64] C.K. Engel, L. Chen, G.G. Prive, Biochim. Biophys. Acta 1564 (2002) 47–56.
[65] J.M. Boulter, D.N. Wang, Protein Expr. Purif. 22 (2001) 337–348.
[66] M.J. Lemieux, R. Reithmeier, D.N. Wang, J. Struct. Biol. 137 (2002) 322–332.
[67] J. Deisenhofer, O. Epp, K. Miki, R. Huber, H. Michel, Nature 318 (1985) 618–624.
[68] W. Kühlbrandt, D.N. Wang, Y. Fujiyoshi, Nature 367 (1994) 614–621.
[69] J. Wu, H.R. Kaback, J. Mol. Biol. 270 (1997) 285–293.
[70] J. le Coutre, L.R. Narasimhan, C.K. Patel, H.R. Kaback, Proc. Natl. Acad. Sci. U. S. A. 94 (1997) 10167–10171.
[71] D. Fu, P.C. Maloney, J. Biol. Chem. 272 (1997) 2129–2135.
[72] M. Kates, Techniques in Lipidology, North-Holland Publ., Amsterdam, 1972.
[73] S. Nussberger, K. Dörr, D.N. Wang, W. Kühlbrandt, J. Mol. Biol. 234 (1993) 347–356.
[74] A.R. Ferre-D'Amare, S.K. Burley, Structure 2 (1994) 357–359.
[75] R.M. Garavito, D. Picot, P.J. Loll, J. Bioenerg. Biomembranes 28 (1996) 13–27.
[76] J.P. Rosenbusch, A. Lustig, M. Grabo, M. Zulauf, M. Regenass, Micron 32 (2001) 75–90.
[77] C.K. Engel, L. Chen, G.G. Prive, Biochim. Biophys. Acta 1564 (2002) 47–56.
[78] M.P. Deutcher, in: M.P. Deutcher (Ed.), Methods in Enzymology, vol. 182, Academic Press, San Diego, 1990, pp. 83–89.

[79] D.N. Wang, M.J. Lemieux, J.M. Boulter, in: B. Selinsky (Ed.), Membrane Protein Protocols: Expression, Purification and Crystallization, Humana Press, Totowa, NJ, 2002, in press.
[80] M.C. Maiden, M.C. Jones-Mortimer, P.J. Henderson, J. Biol. Chem. 263 (1988) 8003–8010.
[81] J. Zhuang, G.G. Prive, G.E. Werner, P. Ringler, H.R. Kaback, A. Engel, J. Struct. Biol. 125 (1999) 63–75.
[82] J.-L. Rigaud, B. Pitard, D. Levy, Biochim. Biophys. Acta 1231 (1995) 223–246.
[83] D. Taglicht, E. Padan, S. Schuldiner, J. Biol. Chem. 266 (1991) 11289–11294.
[84] G.E. Martin, N.G. Rutherford, P.J. Henderson, A.R. Walmsley, Biochem. J. 308 (1995) 261–268.
[85] I.C. West, Biochim. Biophys. Acta 1331 (1997) 213–234.
[86] P. Viitanen, J. Newman, D.I. Foster, T.H. Wilson, H.R. Kaback, Methods Enzymol. 125 (1986) 429–452.
[87] M.C. Fann, P.C. Maloney, J. Biol. Chem. 273 (1998) 33735–33740.
[88] D. Iserentant, W. Fiers, Gene 9 (1980) 1–12.
[89] C.R. Wood, M.A. Boss, T.P. Patel, J.S. Emtage, Nucleic Acids Res. 12 (1984) 3937–3950.
[90] F.R. Blattner, G. Plunkett, C.A. Bloch, N.T. Perna, V. Burland, M. Riley, J. Collado-Vides, J.D. Glasner, C.K. Rode, G.F. Mayhew, J. Gregor, N.W. Davis, H.A. Kirkpatrick, M.A. Goeden, D.J. Rose, B. Mau, Y. Shao, Science 277 (1997) 1453–1474.
[91] F. Kunst, N. Ogasawara, I. Moszer, et al., Nature 390 (1997) 249–256.
[92] A. Ruepp, W. Graml, M.L. Santos-Martinez, K.K. Koretke, C. Volker, H.W. Mewes, D. Frishman, S. Stocker, A.N. Lupas, W. Baumeister, Nature 407 (2000) 508–513.
[93] I.T. Paulsen, M.K. Sliwinski, B. Nelissen, A. Goffeau, M.H. Saier, FEBS Lett. 430 (1998) 116–125.
[94] K.P. Shatwell, B.M. Charalambous, T.P. McDonald, P.J. Henderson, J. Bacteriol. 177 (1995) 5379–5380.
[95] F.J. Gunn, C.G. Tate, P.J. Henderson, Mol. Microbiol. 12 (1994) 799–809.
[96] N.M. Sanderson, G.E. Martin, N.G. Rutherford, P.J. Henderson, Biochem. Soc. Trans. 25 (1997) 471S.
[97] H.C. Dent, P.J. Henderson, V.A. Lucas, Biochem. Soc. Trans. 20 (1992) 251S.
[98] E. Bochkareva, A. Seluanov, E. Bibi, A. Girshovich, J. Biol. Chem. 271 (1996) 22256–22261.
[99] T.G. Consler, B.L. Persson, H. Jung, K.H. Zen, K. Jung, G.G. Prive, G.E. Verner, H.R. Kaback, Proc. Natl. Acad. Sci. U. S. A. 90 (1993) 6934–6938.
[100] J.L. Yu, L. Grinius, D.C. Hooper, J. Bacteriol. 184 (2002) 1370–1377.
[101] K. Abe, Z.S. Ruan, P.C. Maloney, J. Biol. Chem. 271 (1996) 6789–6793.
[102] K.I. Racher, R.T. Voegele, E.V. Marshall, D.E. Culham, J.M. Wood, H. Jung, M. Bacon, M.T. Cairns, S.M. Ferguson, W.J. Liang, P.J. Henderson, G. White, F.R. Hallett, Biochemistry 38 (1999) 1676–1684.
[103] E. Tamai, M.C. Fann, T. Tsuchiya, P.C. Maloney, Protein Expr. Purif. 10 (1997) 275–282.
[104] T. Pourcher, S. Leclercq, G. Brandolin, G. Leblanc, Biochemistry 34 (1995) 4412–4420.
[105] I. Mus-Veteau, G. Leblanc, Biochemistry 35 (1996) 12053–12060.
[106] H. Jung, S. Tebbe, R. Schmid, K. Jung, Biochemistry 37 (1998) 11083–11088.
[107] I. Gaillard, D.J. Slotboom, J. Knol, J.S. Lolkema, W.N. Konings, Biochemistry 35 (1996) 6150–6156.
[108] M. Kanamori, H. Kamata, H. Yagisawa, H. Hirata, J. Biochem. (Tokyo) 125 (1999) 454–459.
[109] E. Pinner, E. Padan, S. Schuldiner, J. Biol. Chem. 267 (1992) 11064–11068.