Consequences of Membrane Protein Overexpression in *Escherichia coli*§

Samuel Wagner‡, Louise Baars‡, A. Jimmy Ytterberg§¶, Anja Klussmeier‡, Claudia S. Wagner‖, Olof Nord**‡‡, Per-Åke Nygren**, Klaas J. van Wijk§, and Jan-Willem de Gier‡$$

Overexpression of membrane proteins is often essential for structural and functional studies, but yields are frequently too low. An understanding of the physiological response to overexpression is needed to improve such yields. Therefore, we analyzed the consequences of overexpression of three different membrane proteins (YidC, YedZ, and Lepl) fused to green fluorescent protein (GFP) in the bacterium *Escherichia coli* and compared this with overexpression of a soluble protein, GST-GFP. Proteomes of total lysates, purified aggregates, and cytoplasmic membranes were analyzed by one- and two-dimensional gel electrophoresis and mass spectrometry complemented with flow cytometry, microscopy, Western blotting, and pulse labeling experiments. Composition and accumulation levels of protein complexes in the cytoplasmic membrane were analyzed with improved two-dimensional blue native PAGE. Overexpression of the three membrane proteins, but not soluble GST-GFP, resulted in accumulation of cytoplasmic aggregates containing the overexpressed proteins, chaperones (DnaK/J and GroEL/S), and soluble proteases (HslUV and ClpXP) as well as many precursors of periplasmic and outer membrane proteins. This was consistent with lowered accumulation levels of secreted proteins in the three membrane protein overexpressors and is likely to be a direct consequence of saturation of the cytoplasmic membrane protein translocation machinery. Importantly accumulation levels of respiratory chain complexes in the cytoplasmic membrane were strongly reduced. Induction of the acetate-phosphotransacetylase pathway for ATP production and a downregulated tricarboxylic acid cycle indicated the activation of the Arc two-component system, which mediates adaptive responses to changing respiratory states. This study provides a basis for designing rational strategies to improve yields of membrane protein overexpression in *E. coli*. Molecular & Cellular Proteomics 6:1527–1550, 2007.

From the ‡Department of Biochemistry and Biophysics, Center for Biomembrane Research, Stockholm University, SE-106 91 Stockholm, Sweden, §Department of Plant Biology, Cornell University, Ithaca, New York 14853, ¶Center for Infectious Medicine, Karolinska University Hospital Huddinge, Karolinska Institutet, SE-141 86 Stockholm, Sweden, and **Department of Molecular Biotechnology, School of Biotechnology, Royal Institute of Technology (KTH), SE-106 91 Stockholm, Sweden

Received, November 14, 2006, and in revised form, April 6, 2007

Published, MCP Papers in Press, April 19, 2007, DOI 10.1074/mcp.M600431-MCP200

In both pro- and eukaryotes 20–30% of all genes encode α-helical transmembrane domain (TMD) proteins, which act in various and often essential capacities (1, 2). Notably, these TMD proteins (hereafter referred to as membrane proteins) play key roles in disease, and they constitute more than half of all known drug targets (e.g. Ref. 3). The natural abundance of membrane proteins is in general too low to conveniently isolate sufficient material for functional and structural studies (4, 5). Therefore, membrane proteins are often obtained through overexpression. The bacterium *Escherichia coli* is the most widely used vehicle for this purpose with overexpressed proteins accumulating in the cytoplasmic membrane (also named inner membrane) or in cytoplasmic inclusion bodies (4). Although membrane proteins can often more easily be expressed in inclusion bodies, their refolding into functional proteins is challenging and often not successful (6). Overexpression of membrane proteins through accumulation in a membrane system avoids this refolding problem but is usually toxic to the organism, thereby severely reducing yields (4). The reasons for this toxicity are not clear; therefore, a better understanding of the physiological response to overexpression is needed to improve such yields through rational design (e.g. through engineering of strains or modifying target proteins). Because optimal protein production conditions cannot be predicted, yield maximization is currently mostly done by “trial and error.” However, green fluorescent protein (GFP)-based methodology developed for *E. coli* now facilitates rapid screening for overexpression in the cytoplasmic membrane and can accelerate the trial and error process (7, 8). Improved prediction of protein overexpression success would be very beneficial but requires an understanding of the physiological response of the cell to overexpression.

It is generally assumed that the overexpressed membrane protein affects integrity of the membrane and thus cell viability, leading to e.g. reduced growth and hampered division (4). In addition, overexpression of membrane proteins may lead to

---

1 The abbreviations used are: TMD, transmembrane domain; BCA, bicinchoninic acid; Lepl, inverted leader peptidase; IPTG, isopropyl β-D-thiogalactopyranoside; 1D, one-dimensional; 2D, two-dimensional; BN, blue native; PMF, peptide mass fingerprinting; SRP, signal recognition particle; GFP, green fluorescent protein; a.u., arbitrary units; TEA, triethanolamine; pta, phosphotransacetylase; bis-Tris, 2-[bis(2-hydroxyethyl)amino]propane-1,3-diol; Tricine, N-[2-hydroxy-1,1-bis(hydroxymethyl)ethyl]glycine; MOWSE, molecular weight search; E2, dihydrolipoyl acetyltransferase; E3, dihydrolipoyl dehydrogenase; Q, ubiquinone.
Consequences of Membrane Protein Overexpression in E. coli

saturation of the protein sorting and translocation machineries, possibly preventing biogenesis of endogenous proteins. Our knowledge of E. coli membrane protein biogenesis is growing rapidly but is far from complete (9). The signal recognition particle pathway (consisting of the signal recognition particle (SRP) and its receptor FtsY) guides a ribosome membrane protein nascent chain complex to the cytoplasmic membrane Sec translocon (10). The ribosome membrane protein nascent chain complex subsequently docks at the Sec translocon. The core of the Sec translocon consists of the integral membrane proteins SecY and SecE, which form a protein-conducting channel (11). SecA is a peripheral subunit of the Sec translocon and is involved in translocation of sizable periplasmic loops of membrane proteins across the membrane, and it is also required for the translocation of secretory proteins (12). TMDs of membrane proteins get trapped in the Sec translocon and move subsequently laterally out from the Sec translocon into the lipid bilayer (9). The cytoplasmic membrane protein YidC may facilitate this process and may mediate the folding of membrane proteins (13, 14). The SecB-dependent pathway targets secretory proteins in a mostly post-translational fashion to the cytoplasmic membrane (15, 16). Because the SRP and SecB pathways converge at the Sec translocon (17) and because SecA engages both membrane and secretory proteins, it is possible that there is competition between sorting of secretory proteins (outer membrane and periplasmic proteins) and the integral membrane proteins of the cytoplasmic membrane.

In addition, relatively little is known about stability, quality control, and degradation of membrane proteins, and it is not known to which extent proteolysis and folding affect overexpression. It is possible that overexpressed proteins are rapidly degraded by endogenous proteases located in the cytosol (such as the ClpP/X/A system) or located in the cytoplasmic membrane (such as FtsH and HtpX) (18–20). This would lead to strongly reduced yields. Besides this, membrane protein overexpression could lead to a general problem of protein homeostasis of the endogenous proteome and lead to the induction of proteolysis and unwanted turnover.

So far, the physiological response to overexpression of membrane proteins in E. coli (or other overexpression “vehicles”) has not been systematically studied, and therefore, this was the objective of this study. Three variants of E. coli membrane proteins (YidC, YedZ, and Lepl) fused to GFP were overexpressed in E. coli BL21(DE3)pLysS from a pET-derived vector (21). Cells were grown aerobically in standard Luria-Bertani broth supplemented with kanamycin (50 μg/ml) and chloramphenicol (30 μg/ml). Overnight cultures were diluted 1:50. 1-liter cultures were grown in Tunair 2.5-liter baffled shaker flasks at 30 °C in an Innova 4330 (New Brunswick Scientific) shaker at 180 rpm. Growth was monitored by measuring the A600 with a Shimadzu UV-1601 spectrophotometer. The pH of the culture medium was monitored with a PHM220 pH meter from Radiometer. For all experiments, protein expression was induced by the addition of 0.4 mM isopropyl β-d-thiogalactopyranoside (IPTG) (final concentration) at an A600 of 0.4–0.5, and cells were harvested 4 h after induction and used for further analysis. Cells with an empty expression vector were used as a control.

Flow Cytometry and Fluorescence Microscopy—Analysis of cells overexpressing GFP fusion proteins and control cells by means of flow cytometry was done using a FACSCalibur (BD Biosciences) instrument. Cultures were diluted in ice-cold PBS to a final concentration of ~10^6 cells/ml immediately after harvesting. A low flow rate was used throughout data collection with an average of 250 events/s. Forward and side scatter acquisition was used for comparison of cell morphology (22), and the cellular accumulation levels of GFP fusion proteins were measured by GFP fluorescence intensity. Cells were incubated on ice for 30 min with a 0.2 μM concentration of the membrane-specific fluorophore FM4-64 (Invitrogen) to compare the amount of membranes per cell, allowing derivation of the relative cell size (23). Data acquisition was performed using CellQuest software (BD Biosciences), and data were analyzed with FloJo software (Tree Star).

For microscopy, cells were mounted on a slide and immobilized in 1% low melting agarose. Microscopy was performed on a Zeiss Axiosplan2 fluorescence microscope equipped with an Orca-ER camera (Hamamatsu). Images were processed with the AxioVision 4.5 software from Zeiss. For analysis of a filamentous growth phenotype, around 700 cells were screened per sample.

Western Blotting Analysis—The expression levels of DnaK, Ffh, Ftsh1, FtsY, GroEL, HtpX, Ibpa/A/B, L5, SecA, SecB, SecE, SecG, and SecY in whole cell lysates or cytoplasmic membranes were monitored by Western blotting analysis. Whole cells (0.025–0.1 A600 unit) and purified cytoplasmic membranes (3–5 μg of protein) were solubilized in Laemmli solubilization buffer and separated by standard SDS-PAGE. Proteins were transferred from the polyacrylamide gel to a PVDF membrane (Millipore). Subsequently membranes were blocked and decorated with antisera to the components listed above as described before (24). Proteins were visualized with secondary horse-radish peroxidase-conjugated antibodies (Bio-Rad) using the ECL system (according to the instructions of the manufacturer, GE Healthcare) and a Fuji LAS 1000-Plus charge-coupled device camera. Blots

Experimental Procedures

**Strain, Plasmids, and Culture Conditions**—Proteins (YidC, YedZ, Lepl, and GST) were overexpressed as GFP fusions in E. coli BL21(DE3)pLysS from a pET-derived vector (21). Cells were grown aerobically in standard Luria-Bertani broth supplemented with kanamycin (50 μg/ml) and chloramphenicol (30 μg/ml). Overnight cultures were diluted 1:50. 1-liter cultures were grown in Tunair 2.5-liter baffled shaker flasks at 30 °C in an Innova 4330 (New Brunswick Scientific) shaker at 180 rpm. Growth was monitored by measuring the A600 with a Shimadzu UV-1601 spectrophotometer. The pH of the culture medium was monitored with a PHM220 pH meter from Radiometer. For all experiments, protein expression was induced by the addition of 0.4 mM isopropyl β-d-thiogalactopyranoside (IPTG) (final concentration) at an A600 of 0.4–0.5, and cells were harvested 4 h after induction and used for further analysis. Cells with an empty expression vector were used as a control.
were quantified using the Image Gauge 3.4 software (Fuji). Experiments were done with independent triplicate samples and were reproducible within 10%.

2D Gel Electrophoresis—2D gel electrophoresis of whole cell lysates was performed as described previously (24). Aggregates (see below) containing 250 μg of protein were solubilized in 7 mL urea, 2 mL thiourea, 1% (w/v) ASB-14, 2 mL tributylphosphine, 5% glycerol, 2% (v/v) IPG buffer for pH 4–7 (GE Healthcare), and bromphenol blue (25). 11-cm-long Immobiline DryStrips, pH 4–7 (GE Healthcare), were used, and isoelectric focusing was performed for 60 kV·h. Aggregated proteins were separated in the second dimension on 8–16% (w/v) IPG buffer for pH 4–7 (GE Healthcare), and bromphenol blue (26), and preparative gels were stained with high sensitivity silver stain (27). Gels used for comparative analysis reached the bottom of the gel. Gels used for MS-based identification of proteins were stained with Coomassie Brilliant Blue R-250 or MS-compatible silver stain (27).

Isolation of Protein Aggregates—Protein aggregates were isolated as described previously (28). 50 mL of culture were used for each aggregate isolation. The protein content of total cells and aggregates was determined with the bicinchoninic acid (BCA) assay according to the instructions of the manufacturer (Pierce). Aggregates were analyzed using three different methods: by SDS-PAGE using 24-cm-long 8–16% acrylamide gradient gels, by Bio-Rad Criterion system 2D gels (see “2D Gel Electrophoresis” above) (both 1D and 2D gels were stained with Coomassie Brilliant Blue R-250 and subjected to MS as described below), and finally by a direct in-solution digest followed by nano-LC-ESI-MS/MS essentially as described before (29).

Isolation of Cytoplasmic Membranes—Cell fractionation was carried out essentially as described before (30) using two subsequent sets of sucrose density gradients. Cells were cultured as described above, harvested at 6000 × g using a Beckman 8,1000 rotor, and washed once with buffer K (50 mM triethanolamine (TEA), 250 mM sucrose, 1 mM EDTA, 1 mM DTT, pH 7.5). The cell pellets were snap frozen in liquid nitrogen and stored at −80 °C. 1000 A450 units of cells were resuspended in 8 mL of buffer K supplemented with 0.1 mg/mL Pefabloc and 5 μg/mL DNase and lysed by two cycles of French pressing (18,000 p.s.i.). The lysate was cleared of unbroken cells by 20-min centrifugation at 8000 × g. The supernatant was applied on top of a two-step sucrose gradient: bottom, 0.8 mL of 55% (w/w) sucrose; top, 5.0 mL of 9% (w/w) sucrose. All sucrose gradients were prepared in buffer M (50 mM TEA, 1 mM EDTA, 1 mM DTT, pH 7.5). The gradients were centrifuged for 2.5 h at 210,000 × g in a Beckman SW 41 rotor, and the membrane fraction was collected from the top of the 55% sucrose layer. This fraction, which contains the total membranes, was diluted 1:3 with buffer M and subjected to a six-step sucrose gradient centrifugation run to obtain pure cytoplasmic membrane fractions. The composition of this second gradient was as follows (from bottom to top): 0.7 mL of 55%, 1.4 mL of 50%, 1.5 mL of 45%, 2.2 mL of 40%, 1.8 mL of 35%, 0.9 mL of 30% (all w/v) sucrose, and 3.3 mL of the sample. The gradients were centrifuged for 15 h at 210,000 × g in a Beckman SW 41 rotor, and the cytoplasmic membrane fraction was collected from the top of the 40% sucrose layer. The protein concentration of the fraction was determined using the BCA assay according to the manufacturer’s instructions (Pierce). The concentrations of the membrane samples were adjusted to 0.5 mg/mL with buffer L (50 mM TEA, 250 mM sucrose, 1 mM DTT, pH 7.5), and aliquoted samples were stored at −80 °C. The cytoplasmic membrane fraction was analyzed by Western blotting, SDS-PAGE using 24-cm-long 8–16% acrylamide gradient gels, and 2D BN-PAGE (see below).

Analysis of Cytoplasmic Membrane Fractions by 2D BN-PAGE—Blue native electrophoresis as described previously (31) was modified to enable the relative quantification of membrane proteomes in the following way (see Fig. 8A). 1-mm-thick first dimension polyacrylamide gels were cast onto GelBond PAG film as recommended by the manufacturer (Cambrex). 5–14% gradient gels were used to resolve proteins and protein complexes between 1000 and 60 kDa. Cytoplasmic membranes prepared as described above were pelleted and subsequently solubilized in buffer containing 750 mM 6-aminocaproic acid, 50 mM bis-Tris-HCl (pH 7.0 at 4 °C), and freshly prepared 0.5% (w/v) n-dodecyl-β-D-maltopyranoside. After removal of unsolubilized material by centrifugation (100,000 × g for 30 min), Serva Blue G was added to a final concentration of 0.5% (w/v), and the samples were loaded onto the first dimension gel. Coomassie-containing cathode buffer was used throughout the run. Electrophoresis of the first dimension was performed at 100–400 V until the dye front reached the end of the gel. For calibration, ferritin (440 and 880 kDa), aldolase (158 kDa), and albumin (66 kDa) (GE Healthcare) were used as molecular mass markers. Lanes cut from the first dimension gel were equilibrated for 15 min in a buffer containing 2% SDS, 5 mM tributylphosphine followed by equilibration for 15 min in 2% SDS, 260 mM iodoacetamide. The lanes were mounted on top of the 1.5-mm-thick second dimension gel by submerging the strips in warm agaro agarose solution (1% (w/v) low melting agarose, 0.5% SDS, bromphenol blue). The samples were separated in the second dimension on 10% Dury (Genomic Sciences) gels (10% acrylamide monomer and 1% bisacrylamide) containing 1 mM Tris-HCl (pH 8.45), 0.1% (w/v) SDS, and 20% (v/v) glycerol. Electrophoresis was performed with a Tricine-SDS buffer system (32) in an Ettan DALTia/clve system (GE Healthcare) at 80 V for ~48 h until the dye front reached the bottom of the gel. Gels were stained with colloidal Coomassie stain (33).

2D Gel Image Analysis and Statistics—Gels were scanned using a GS-800 densitometer from Bio-Rad. Spots in 2D gels were analyzed using the PDQuest software (Bio-Rad). Each comparative standard 2D gel analysis set included four gels, and each comparative 2D BN-PAGE analysis set included three gels. All gels in a set represented independent samples (i.e., samples from different bacterial colonies and cultures), which were subjected to 2D gel and image analysis. Spot quantities were normalized using the “total density in gel image” method to compensate for non-expression-related variations in spot quantities between gels. PDQuest was set to detect differences that were found to be statistically significant using the Student’s t test and a 95% level of confidence, including qualitative differences (“on/off responses”) present in all gels in a group and quantitative differences in at least one replicate group. In analysis sets of 2D BN-PAGE gels, differences were only accepted if the spot intensities were at least 2-fold with a 95% level of confidence. The quantification of saturated spots was approximated using the “con-tour tool” of PDQuest.

Protein Identification by Mass Spectrometry and Bioinformatics—Stained protein spots or bands were excised, washed, and digested with modified trypsin; peptides were extracted manually or automatically (ProGest, Genomic Solutions); and peptides were applied to the MALDI target plates as described previously (34). The mass spectra were obtained automatically by MALDI-TOF MS in reflectron mode (Voyager-DE-STR, PerSeptive Biosystems) followed by automatic internal calibration using tryptic peptides from autodigestion. The spectra were analyzed for monoisotopic peptide peaks (m/z range, 850–5000) using the software MoverZ from Genomic Solutions with a signal to noise ratio threshold of 3.0. Matrix and/or autoproteolytic trypsin fragments were not removed unless otherwise indicated (see below). Spectral annotations (in particular assignments of monoisotopic masses) were verified by manual inspection for a large number of measurements. The resulting peptide mass lists were used to search the Swiss-Prot 48.1 database (release September 27, 2005) for E. coli (7533 sequences) with Mascot (version 2.0; Matrix Science) in automated mode using the following search parameter criteria:

**Molecular & Cellular Proteomics 6.9**
significant protein MOWSE score at $p < 0.05$, no missed cleavages allowed, variable methionine oxidation, fixed carboxymethylation of cysteines, and minimum mass accuracy of 50 ppm. These search result pages were extracted and analyzed by an additional in-house filter applying the following criteria for positive identification: (i) MOWSE score $\geq 50$, (ii) $\geq 4$ matching peptides with an error distribution within $\pm 25$ ppm, and (iii) $\geq 15\%$ sequence coverage. In case of hydrophobic integral membrane proteins on the 2D BN-PAGE gels, the minimum sequence coverage was 10\% but only if the experimental denatured mass corresponded with the predicted denatured mass. The minimum ratio of matched to unmatched peaks accepted was 0.25. False positive rates were less than 1\% as determined by searching with the .pkl list against the *E. coli* database (Swiss-Prot 48.1) mixed with a randomized version of the *E. coli* database generated using a Perl script from Matrix Science.

In a number of cases, the PMF-based identifications of proteins in the aggregates were affected due to the low abundance of proteins and/or the complex nature of the samples analyzed. Trypsin and matrix peaks were rather dominant in these spectra, resulting in a ratio of matched peptides versus unmatched peptides of $<0.25$. Therefore, the respective spectra were re-searched manually with trypsin and matrix peaks subtracted. These identifications are marked in Supplemental Table 2. Matrix and/or autoproteolytic trypsin peaks were identified by processing a gel piece devoid of protein. If indicated, the following peaks were removed: 842.5, 1045.6, 1940.9, 2211.1, and 2283.2 as well as the respective methyl-, formyl-, and acetyl-modified and oxidized forms.

For several critical samples, the peptides were also analyzed by nano-LC-ESI-MS/MS in automated mode on a quadrupole/orthogonal-acceleration TOF tandem mass spectrometer (Q-TOF, Micromass) (see Friso et al. (35) for details). MassLynx 4.0 was used for spectral annotations ($m/z$ range: MS, 350–1750; MS/MS, 50–1750). The resulting peak lists were used to search Swiss-Prot 48.1 database (downloaded locally) automated using Mascot (version 2.0; Matrix Science). When searching Mascot, the maximum precursor and fragment errors were 1.2 and 0.8 Da, respectively. Probability-based MOWSE score, number of matching peptides, and highest peptide score were extracted from the Mascot peptide summary report pages using in-house written software. Minimal criteria for identification were as follows: (i) one matching peptide with a peptide score higher than the minimal significant ($p < 0.05$) individual ion score, (ii) two matching peptides with a score higher than 21, or (iii) three matching peptides with peptide score of 20 or higher. All significant MS/MS identifications by Mascot were manually verified for spectral quality and matching y and b ion series.

The curated and non-redundant Swiss-Prot 48.1 database was used throughout the study to limit redundancy of protein assignments. Syntrophic gene names are given in the supplemental tables. Ambiguous protein assignments for members of multiprotein families were checked manually. Only one ambiguous protein assignment was identified. This is indicated by "*" in the tables (see wzzB in spot 91 in Table II, Fig. 8, and Supplemental Table 3).

**Protein Translocation Assay**—Cells were cultured as described under "Strain, Plasmids, and Culture Conditions." 4 h after induction of protein expression, cells from 1 ml of culture were collected by centrifugation (2 min at 3000 g) and washed in M9-based medium as described previously (36), and washed cells were resuspended in 1 ml of M9-based medium supplemented with 50 $\mu$g/ml kanamycin, 30 $\mu$g/ml chloramphenicol, and 0.4 mm IPTG. Cells were incubated at 30 °C for 10 min, were then labeled with [$^{35}$S]methionine (60 $\mu$Ci/ml, 1 Ci = 37 GBq) for 45 s, and subsequently precipitated in 10% TCA. TCA-precipitated samples were washed with acetone; resuspended

2 Q. Sun and K. J. van Wijk, unpublished data
Hansatech Oxytherm Oxygraph for 2 min. The reaction was started by the addition of 3 A_600 units of cells to 1 ml of PBS at 25 °C. Succinate dehydrogenase and cytochrome oxidase activities of inverted membrane vesicles were determined as described previously (37, 38).

RESULTS

Overexpressed Proteins and Culture Conditions—The widely used E. coli BL21(DE3)pLysS strain served as the organism to follow the response to protein overexpression (39). All proteins were expressed from a pET28a/H11001-derived vector as C-terminal GFP fusions, which facilitated monitoring membrane protein overexpression (21). The GFP variant used for the fusions is optimized for expression and folding in E. coli (40).

Three E. coli membrane proteins that are predominantly overexpressed into the cytoplasmic membrane, YidC, YedZ, and Lepl, were chosen for this study (Fig. 1A) (7). YidC is an integral membrane chaperone, and YedZ is an integral membrane flavocytochrome; each has six TMDs (7, 9, 41, 42). Lepl is a leader peptidase derivative with two TMDs that has been modified such that it inserts with an inverted topology into the cytoplasmic membrane; it is thus a non-functional protein (43). Overexpression of the soluble protein GST was used to distinguish between the general response to overexpression and a possible specific response to membrane protein overexpression. The empty pET28a+-derived GFP fusion vector was used as the control. The experimental outline of the study is shown in a flow chart (Fig. 1B).

Cells of controls and the four GFP transformants were cultured aerobically at 30 °C in Luria-Bertani broth to an A_600 of 0.4–0.5 after which protein expression was induced for 4 h with IPTG. Overexpression of the membrane proteins strongly hampered growth with maximum cell densities reduced by more than 50% compared with the control (A_600 of 0.8–1.3 versus 2.5) and transition to stationary phase after 2 ho induction. In contrast, overexpression of the soluble GST-GFP did not affect growth with maximum cell densities of A_600 of 2.5 (Fig. 2A).

To assess whether the function of YidC-GFP contributed to the toxic effect of overexpression, we also monitored the expression of three non-functional GFP-tagged YidC variants (I361S, Y516S, and ΔTM3) (44). All three YidC-GFP mutants were stably expressed at the same level as the wild-type YidC-GFP protein and had the same effect on growth (data not shown). This indicates that the effect of YidC-GFP overexpression on growth is not due to the specific function of YidC.

Flow Cytometry and Microscopy—The morphology of the cells overexpressing GFP fusions and the yield of GFP fusion proteins in individual cells were studied by flow cytometry and microscopy (Fig. 2, B–F). The GFP yield, measured as geo-
Consequences of Membrane Protein Overexpression in *E. coli*

A.

- αDnaK
- αGroEL
- αlpA/B
- αFlih
- αFtsY
- αSecA
- αSecB

B.

- control
- YidC-GFP
- YedZ-GFP
- LepA-GFP
- GST-GFP

SDS - polyacrylamide gel electrophoresis

 isolectric focusing

MM/ kDa

MM/ kDa
Consequences of Membrane Protein Overexpression in *E. coli*

The metric mean of GFP fluorescence per cell, was highest for soluble GST-GFP (1538 arbitrary units (a.u.)) followed by YidC-GFP (188 a.u.), YedZ-GFP (142 a.u.), and LepI-GFP (138 a.u.), whereas control levels were 2 a.u. Flow cytometry also showed that GFP fusion proteins were homogeneously expressed by the cells of the respective cultures (Fig. 2B). An increase in forward and side scatter of cells overexpressing membrane proteins indicated that these cells were slightly bigger (particularly YidC-GFP) and contained extra internal structures (Fig. 2C). This increased cell size corresponded to an increased amount of membranes as determined by staining with the membrane-specific fluorophore FM4-64 (Fig. 2D). The increase in forward and side scatter and FM4-64 fluorescence was most pronounced for the overexpression of YidC-GFP, thus correlating with GFP expression levels and FM4-64 stains (Fig. 2, C and D).

To characterize the nature of altered cell morphology in more detail, cells were analyzed by fluorescence microscopy (Fig. 2, E and F). A significant proportion of cells overexpressing membrane proteins showed a filamentous phenotype due to incomplete cell division. This observation supported the flow cytometry results, which indicated an increased cell size upon membrane protein overexpression. This effect was most pronounced for YidC-GFP overexpression with 16% of all cells showing a filamentous phenotype compared with only 1% upon GST-GFP overexpression. A filamentous phenotype was defined as a cell length corresponding to at least three non-dividing cells and/or the presence of more than one division septum.

**Analysis of Whole Cell Lysates by Western Blotting**—As a first step in the characterization of the cellular proteomes of the GFP fusion overexpressors, cell lysates were analyzed by Western blotting with antibodies against components of the protein sorting machineries and various stress responses (Fig. 3A). Levels of Ffh and FtsY, core components of the SRP targeting pathway, were unaltered upon membrane protein overexpression. The levels of SecB, involved in the targeting of secretory proteins, were about 20% lower in cells overexpressing the membrane proteins as well as GST-GFP. The levels of SecA, a peripheral subunit of the Sec translocon, were increased by around 50% in cells overexpressing membrane proteins but clearly not in cells overexpressing GST-GFP. This suggests that protein translocation is hampered at the level of the Sec translocon (45). The levels of the chaperone DnaK were around 65% higher after overexpression of membrane and GST-GFP fusion proteins. Accumulation of chaperone GroEL was also increased in all overexpressors but only by 25%. Inclusion body-binding proteins A and B (IbpA/B) are proteins typically found in cytoplasmic aggregates and facilitate disaggregation by CipB (46). IbpA/B levels were practically undetectable in the control but increased 40-fold after overexpression of YidC-GFP and 6–17-fold for the other GFP constructs (Fig. 3A). Taken together, the Western blotting experiments indicate that membrane protein overexpression in the cytoplasmic membrane leads to protein translocation stress and a protein folding problem in the cytoplasm.

**Analysis of Whole Cell Lysates by 2D Gel Electrophoresis**—Total cellular proteomes of cells overexpressing the GFP fusion proteins and control cells were compared by image analysis of 2D gels with denaturing IPG strips (pH 4–7) in the first dimension and Tris-Tricine SDS-PAGE gels in the second dimension (24). The analysis was carried out in four biological replicates. Gels were stained with silver and scanned, and images were subsequently analyzed and compared using the PDQuest software (Bio-Rad). Significance was determined using Student’s t test (for details see “Experimental Procedures”). 2D gels of cells overexpressing the GST-GFP fusion were not included in the final comparative analysis because the large amount of overexpressed soluble GST-GFP protein prevented conclusive analysis. This was not a problem for cells overexpressing membrane proteins because their expression levels were 10-fold lower than the expression levels of GST-GFP and because membrane proteins do not resolve on the 2D gels. Fig. 3B shows a representative gel for the control cells and for cells overexpressing YidC-GFP. The quantitative data for overexpressors of GFP fusion proteins of YidC, YedZ, and LepI are shown as a bar diagram with changes compared with control cells (Fig. 4A). Protein spots were excised, and proteins were digested with trypsin. The resulting peptides were extracted and analyzed by MALDI-TOF MS, and proteins were identified by PMF. Spot statistics and MS data are provided in Supplemental Table 1. The fold changes were typically highest in the YidC-GFP overexpres-
Consequences of Membrane Protein Overexpression in *E. coli*

A.

B.
Consequences of Membrane Protein Overexpression in E. coli

121 spots corresponding to 110 non-redundant proteins were significantly (p < 0.05) changed in the YidC-GFP cells compared with the control; 86 spots decreased, and 35 spots increased. Some proteins were identified in multiple spots located next to each other. 13 spots contained two different proteins with similar pl and molecular weights, and in three cases the identified proteins did not match the molecular weight (Fig. 3B, spots 1, 26, and 52).

Of the 35 proteins with increased accumulation levels, seven were chaperones and proteases: they are the folding mediators DnaK (2.7-fold), GroEL (2.0-fold), the disaggregation chaperones ClpB (4.4-fold) and Ibpa (2.8-fold), and the protease components HslV and HslU (4.3- and 2.6-fold, respectively). Chaperones and proteases are in bold letters in Figs. 3B and 4A. The increased levels of DnaK and GroEL in the 2D gels are in agreement with the Western blotting results presented earlier (Fig. 3A). It should be noted that Ibpa co-localized with the protein UspF in our 2D gels. This explains why the three overexpressors did not show an “on” response of Ibpa as observed by Western blotting. The expression of all these chaperones and proteases is regulated by the transcription factor σ32 (47). The σ32-induced response, better known as the “heat shock response,” is activated in response to protein misfolding in the cytoplasm. The induction of the heat shock response has been described before for the overexpression of cytoplasmic membrane proteins (51).

Further analysis of the 2D gels revealed that upon membrane protein overexpression the levels of several enzymes of the tricarboxylic acid cycle (malate dehydrogenase (Mdh), fumarase A (FumA), fumarase C (FumC), and succinyl-CoA synthetase subunit α (SucD)) were about 3-fold lower than in the control. In contrast, the levels of enzymes involved in the “payoff phase” of glycolysis (glyceraldehyde 3-phosphate dehydrogenase (GapA), phosphoglycerate kinase (Pgk), phosphoglycerate mutase (GpmI), and pyruvate kinase I (PykF)) were increased 2.4–4.6-fold (Figs. 4A and 11). Furthermore levels of the E2 and E3 components of pyruvate dehydrogenase (LpdA and AceF) as well as of acetate kinase (AckA) were all 2–5-fold increased upon membrane protein overexpression. These enzymes catalyze the formation of ATP from pyruvate via acetyl-CoA and acetyl phosphate under anaerobic conditions (52), leading to the production and subsequent secretion of acetate into the culture medium. Indeed a slight acidification of the culture medium was observed upon membrane protein overexpression (Fig. 5A). Notably the extent of the acidification of the medium correlated positively with increased levels of pyruvate dehydrogenase and acetate kinase; YidC-GFP overexpression resulted in the most pronounced effect followed by Lepl-GFP and YedZ-GFP. Also the level of YbeD, involved in the biosynthesis of lipoic acid, which is required as the prosthetic group of pyruvate dehydrogenase, was increased 16-fold (53). Overexpression of the soluble protein GST-GFP did not result in the aforementioned changes and consistently did not affect the pH of the culture medium (Fig. 5A).

25 of the 82 proteins with reduced accumulation levels (30%) were secretory proteins; this is disproportionate because only around 10% of the E. coli genes encode secretory proteins (54). Comparison of the experimental pl and molecular mass with the predicted pl and mass of the precursors and mature (processed) proteins indicated that these proteins were all processed. Hence they must have been translocated across the cytoplasmic membrane. The effect was unique for membrane protein overexpression because GST-GFP overexpression did not result in this phenotype (Fig. 4B). Consistent with the decreased levels of secreted and processed proteins, further inspection of the 2D gels indicated that precursor forms of the abundant secretory proteins OmpA, OmpF, and OppA accumulated in the YidC-GFP and Lepl-GFP overexpressors but not in control cells. Moreover peptides matching to the signal sequence of pre-OmpF were identified by nano-LC-ESI-MS/MS (spot 12 in Fig. 3B, lower panel). Furthermore pulse labeling followed by immunopre-

Fig. 4. Relative quantification of protein spots with differential accumulation levels in whole cell lysates upon membrane protein overexpression. A, relative quantification of protein spots with differential accumulation levels in whole cell lysates upon membrane protein overexpression compared with the empty vector control. Differential protein expression was analyzed using the Student’s t test and a 95% level of confidence (see “Experimental Procedures”). Relative quantifications are based on the average of four independent replicate samples. Proteins were identified by MALDI-TOF MS/PMF. If several proteins were identified in one spot, the first gene name given corresponds to the protein with the most prominent peaks in MALDI-TOF MS spectra and the highest Mascot MOWSE score. Primary gene names are taken from Swiss-Prot (www.expasy.org). Numbers refer to spot positions on the gels in Fig. 3B. Names representing secretory proteins are in italic, and those of chaperones are in bold. A -fold change of 100 indicates a spot that was “on” upon membrane protein overexpression; a -fold change of 0.01 indicates that it was “off.” B, relative spot intensities from 2D gels of whole cell lysates of control cells were plotted against those of YidC-GFP- and GST-GFP-overexpressing cells, respectively. Relative spot intensities represent the average of four independent replicate samples. This relative gels representation circumvents the problem that the overexpressed GST-GFP fusion accounts for a very large fraction of the total protein on 2D gels. Secretory proteins (black triangles) and chaperones/proteases (black squares) whose accumulation levels were significantly changed upon YidC-GFP overexpression and that were identified by MS are highlighted. It is apparent that the abundance of processed forms of many secretory proteins is strongly reduced upon YidC-GFP overexpression, whereas this is not the case upon GST-GFP overexpression. The precursors of some secretory proteins accumulate upon YidC-GFP overexpression (OmpA, OmpF, and OppA).
Consequences of Membrane Protein Overexpression in E. coli

Fig. 5. Membrane protein overexpression affects the pH of the culture and protein secretion. A, the pH of the culture of GFP fusion-overexpressing cells and the control cells was measured every 4 h. Membrane protein overexpression resulted in the slight acidification of the culture medium in contrast to overexpression of the soluble GST-GFP. B, cells overexpressing GFP fusion proteins were pulse-labeled with [35S]methionine, and subsequently OmpA was immunoprecipitated and analyzed as described under “Experimental Procedures.” Membrane protein overexpression led to OmpA precursor accumulation, whereas overexpression of the soluble protein GST-GFP did not.

Precipitation clearly showed accumulation of pre-OmpA upon membrane protein overexpression (Fig. 5B). Precursor accumulation was strongest upon YidC-GFP and LepI-GFP overexpression and absent when the soluble GST-GFP-fusion was expressed. These observations suggest that membrane protein overexpression leads to a bottleneck in protein targeting or translocation of secretory proteins.

Isolation and Characterization of Protein Aggregates—The heat shock response in the overexpressors indicated a significant protein folding problem in the cytoplasm. In addition, flow cytometry experiments suggested the presence of extra internal structures, such as additional internal membranes or protein aggregates, in cells overexpressing membrane proteins. Indeed aggregates could be isolated from cells overexpressing membrane proteins using a Nonidet P-40-based purification (Fig. 6A) (28). Protein aggregates made up around 0.8% of the total cellular protein in cells overexpressing YedZ-GFP and LepI-GFP and around 1.6% in cells overexpressing YidC-GFP. Cells overexpressing GST-GFP accumulated only minor amounts of aggregates and contained predominantly the GST-GFP fusion protein (Fig. 6A).

The protein composition of the aggregates was analyzed by both 1D and 2D gel electrophoresis followed by MALDI-TOF MS and PMF and by a direct in-solution tryptic digestion of the aggregates followed by nano-LC-ESI-MS/MS (Fig. 6, A and B, Table I, and Supplemental Table 2). Interestingly the protein compositions of aggregates isolated from cells overexpressing the different membrane proteins were similar. In total 144 different proteins were identified of which 100 were cytoplasmic, 23 were secretory, and 19 were membrane-associated proteins without any predicted TMDs. Surprisingly integral cytoplasmic membrane proteins were not detected in the 1D gels of aggregates with the exception of the GFP fusion proteins, the FtsH protease adapter protein HflK, and the methyl-accepting chemotaxis protein III (MCP3). HflK and MCP3 have one and two TMDs, respectively, and both have large soluble domains. The near absence of membrane proteins in the aggregates was surprising but was confirmed independently by Western blotting using antisera against abundant cytoplasmic membrane proteins such as the cytochrome bo3 core subunits CybA and CybB and the ATPase core subunit F1C (data not shown).

Estimated from the spot/band intensities in 1D and 2D gels, cytoplasmic chaperones and proteases constituted around 40% of the total aggregated protein, secretory proteins represented around 25%, overexpressed GFP fusion protein was around 20%, and the remainder were mainly established cytoplasmic GroEL/S and DnaK substrates; these proportions appeared true for all three membrane protein overexpressors (Fig. 6C and results not shown) (55, 56). The presence of cytoplasmic proteins in the aggregates pointed to a cytoplasmic localization of the aggregates. This was supported by the presence of the precursor forms of the major outer membrane proteins OmpA, OmpF, and OppA in the 2D gels of aggregates and consistent with the pulse-chase experiments of pre-OmpA.

Analysis of the Cytoplasmic Membrane Proteome by Western Blotting—To monitor the effects of protein overexpression on the composition of the cytoplasmic membrane proteome, cytoplasmic membranes were purified using sucrose gradients and analyzed by Western blotting (Fig. 7). The levels of the Sec translocon components SecY, SecE, and SecG were increased by 10–20% upon membrane protein overexpression. Levels of membrane-associated ribosomal subunit L5 were increased by ~50%, whereas the total concentration in the cell was unchanged (data not shown). In addition, membrane protein overexpression led to an increased association (2.3–4.8-fold) of DnaK with the membrane. Membrane protein overexpression did not significantly affect the levels of FtsH, but HtpX levels were reduced by 50, 30, and 60% in the YidC, YedZ, and LepI overexpressors, respectively. Therefore it is unlikely that HtpX poses a major problem to the stability of overexpressed membrane proteins, whereas the impact of FtsH remains unclear.

Relative Quantification of the E. coli Cytoplasmic Membrane Proteome Using Novel 2D Blue Native PAGE-based Methodology—Because protein bands in 1D gels of complex protein mixtures usually contain more than one protein, staining intensities of these bands cannot be reliably used for quantitative protein analysis. In addition, the use of 1D gel electrophoresis combined with Western blotting is limited by the availability and quality of antibodies. 2D gels using IPG strips in the first dimension are also not suitable for quantitative
membrane protein analysis due to poor recovery and solubility. Although 2D BN-PAGE was used successfully to study the composition and oligomeric state of the cytoplasmic membrane proteome of *E. coli* (31, 57, 58), the current protocols do not appear suitable to facilitate quantitative comparative membrane protein analysis. In particular, the transfer of the first dimension gel slices to the second dimension SDS gels appeared to contribute significantly to gel-gel variability, decreasing the usefulness of 2D BN-PAGE for the relative quantification of membrane proteomes. This most likely explains why 2D BN-PAGE has not been widely used for quantitative purposes so far.

To overcome this limitation, we cast the first dimension BN-PAGE gels on a GelBond PAG film (Cambrex) thus improving the physical strength of the gels. During the polymerization process the gel is covalently linked to the film. This made it possible to quickly transfer the gel slices free of distortion directly onto a second dimension SDS-PAGE gel while it also appeared to reduce diffusion of proteins out of the gel strip during the denaturation and cysteine modification steps preceding SDS-PAGE (Fig. 8A).

The 2D BN-PAGE gels and subsequent image analysis were used to study the effects of overexpression on the cytoplasmic membrane. Examples of the 2D BN-PAGE gels are shown in Figs. 8B and 9. Each analysis set contained independent triplicate samples prepared as described under “Experimental Procedures.” The threshold for acceptance was set at 95% significance and a minimum 2-fold change in accumulation levels (Fig. 10A). Importantly the distribution of the coefficients of variation achieved for the relative quantification of spots using our optimized 2D BN-PAGE gel methodology compared well with results obtained with IPG-based 2D gels for soluble proteins (Fig. 10B).

MS analysis of the 2D BN-PAGE map resulted in the identification of 102 different proteins from 114 spots (Fig. 8B, Table II, and Supplemental Table 3). The protein accumulation patterns were qualitatively very similar for the three overexpressed membrane proteins with overexpression of YidC-GFP generally causing the strongest effects followed by YedZ-GFP and LepI-GFP. Overexpression of GST-GFP hardly affected accumulation levels of membrane proteins (Figs. 9
TABLE I
Identification of proteins in aggregates isolated from membrane protein-overexpressing and control cells

Aggregated proteins were analyzed by 1D and 2D gel electrophoresis followed by MALDI-TOF MS (Fig. 6, A and B) and by in-solution digestion followed by nano-LC-ESI-MS/MS. solut., solution; local., localization; prec., precursor.

| spot 2D | band 1D | in solut. rank | gene name(s) | name | local. substrate | prec. |
|---------|---------|----------------|--------------|------|-----------------|-------|
|         |         | (a)            | (b)          | (c)  | (d)             | (e)   |
| heat-shock proteins |
| 71, 72, 73, 84, 98, 99, 109, 111, 123, 124, 142 | 26 | 9 | lbpA, hslT, htpN | 16 kDa heat shock protein A | c |
| 2, 5, 22, 54, 58, 63, 71, 72, 94, 100 | 26 | 7 | lbpB, hslS, htpE | 16 kDa heat shock protein B | c DnaK |
| 31, 32 | 18 | groL, groEL, mmpA | 60 kDa chaperonin | c |
| 64, 65 | 30 | clpX, lopC | ATP-dependent Clp protease ATP-binding subunit clpX | c Gro, DnaK |
| 86, 102, 103, 113 | 19 | clpP, lopP | ATP-dependent Clp protease proteolytic subunit | c |
| 66, 78 | | hslU, htpI | ATP-dependent hsl protease ATP-binding subunit hslU | c DnaK |
| 140 | | hslV | ATP-dependent protease hslV | c |
| 80, 81, 95 | 9 | clpB, htpM | Chaperone clpB | c |
| | 37.6 | dnaJ, groP | Chaperone protein dnaJ (Heat shock protein J) | c |
| 33, 35 | 4 | dnaK, groP, groF, seq | Chaperone protein dnaK | c |
| 51 | | htpG | Chaperone protein htpG | c |
| 52 | | groB, groES, nodB | GroES protein | c |
| 19, 20 | | bg | Trigger factor | c |
| secretory proteins |
| 98 | marS | Alpha-amylase | p |
| 14, 48.9 | desP, htrA, pdb | Protease desP | p |
| 23.8 | flagA | Flagella basal body P-ring formation protein flagA | p |
| 92 | app | Glucose-1-phosphatase | p |
| 131 | proX, proU | Glycine betaine-binding periplasmic protein | p M |
| 60, 75 | hslJ | Histidine-binding periplasmic protein | p |
| 42 | ybbC | Hypothetical protein ybbC | p |
| 14, 16, 17, 48, 50 | 15 | labB, masB | Maltoporin | om M |
| 65, 89, 95, 120 | 16 | marE | Maltose-binding periplasmic protein | p M |
| 26 | mprA | MprA-interacting protein | p |
| 40, 60 | tsx, nupA | Nuclear-envelope-specific channel-forming protein tax | om |
| 63 | imp, ostA | Organic solvent tolerance protein | om |
| 32 | ybbB | Outer membrane lipoprotein ybbB | om |
| 78, 117, 118, 129, 130, 138 | 21, 22, 23 | ompA, con, tolQ, tut | Outer membrane protein A | om M |
| 3, 9, 10, 11, 12, 13, 15, 28, 44 | 18, 20 | ompF, cmIB, sugF, cry, luxF | Outer membrane protein F | om Q |
| 49, 68, 79, 122 | 22 | tolC, mecB, mukA, relC | Outer membrane protein tolC | om |
| 152, 153 | oppA | Periplasmic oligopeptide-binding protein | p |
| 78, 93, 107 | deqQ, hkoA | Protease deqQ | p |
| 94 | usuA | Protein usuA | p |
| 107 | yciE | Protein yciE | p |
| 67 | yaeC | Putative outer membrane protein yaeC | om |
| 34 | yaeT | Outer membrane protein assembly factor yaeT | om |
| 39 | vacJ | VacJ lipoprotein | om M |
| other proteins |
| 105, 115, 144 | 24 | gpmA, gpmA | 2,3-bisphosphoglycerate-dependent phosphoglycerate mutase | c |
| 27 | sucA | 2-oxoglutarate dehydrogenase E1 component | c |
| 13.7 | rpoJ | 30S ribosomal protein S10 | c |
| 41 | rpsC | 30S ribosomal protein S3 | c |
| 59 | fabH | 2-oxoglutarate-[acyl-carrier-protein] synthase III | c |
| 13 | rplN | 50S ribosomal protein L14 | c |
| 40 | rplD, eryA | 50S ribosomal protein L4 | c |
| 26, 27 | rplE | 50S ribosomal protein L5 | c |
| 141 | rplF | 50S ribosomal protein L9 | c GroEL/S |
| 46, 47 | gnd | 6-phosphogluconate dehydrogenase | c |
TABLE I—continued

| spot 2D | band 1D | in solut. rank | gene name(s) | name | local. | prec. |
|---------|---------|---------------|--------------|------|--------|-------|
| 93.3    | acnA, acn | Acetate hydratase 1 | c        |
| 90      | purA, adeK | Adenylosuccinate synthetase | c        |
| 26      | nudE | ADP compounds hydrolase nudE | c        |
| 61      | arcA, cpxC, dyo, lexA, rsn, seg, sfrA | Aerobic respiration control protein arcA | c, DnaK |
| 96, 121, 130, 137 | ATP synthase alpha chain | cm        |
| 30, 110, 112 | ATP synthase beta chain | cm        |
| 85      | atpD, papB, uncD | ATP synthase epsilon chain | cm        |
| 24      | bfr | Bacterioferritin | c        |
| 7, 34   | lacZ | Beta-galactosidase | c, GroEL/S |
| 1, 23   | bsaS, yoeP | Biofilm regulator bsaS | c        |
| 12      | ftsZ | Cell division protein ftsZ | c        |
| 11      | cscB2, csc | Chloramphenicol acetyltransferase | c        |
| 47.1    | pflA, gluT, icdB | Citrate synthase | c        |
| 51.9    | lpdA, lpd | Dihydrolipoamide dehydrogenase | c, GroEL/S |
| 100, 108 | dps, pncA, vtm | DNA protection during starvation protein | c        |
| 55, 56  | hns, bgY1, cur, drhX, hnsA, hsa, mvaA, osmZ | DNA-binding protein H-NS | c        |
| 28      | tufA, tufB | Elongation factor Tu | cm        |
| 127, 128, 150 | DNA-directed RNA polymerase phase T7 | c        |
| 34.9    | pphC | Electron transport complex protein mfc.C. | c        |
| 38      | mrc | Electron transport complex protein mrc | c        |
| 77, 78, 90, 91, 92 | DNA-binding protein H-NS | c        |
| 90      | eno | Enolase | c        |
| 57      | ftsB | FtsZ, FtsI-like protein 2 | cm, GroEL/S |
| 6, 22   | fliB, flaA, flaB | Flaggellar basal-body rod protein fliB | cm        |
| 16, 21  | fliB, flaA, flaB | Flaggellar hook-associated protein 1 | cm        |
| 8       | fliC, flgT, flgU | Flaggellar hook-associated protein 2 | cm        |
| 108     | wrbA | Flavoprotein wrbA | c        |
| 10      | pflA, pph | Formate acetyltransferase 1 | c        |
| 132     | gatC | Galactoside-1-phosphate 5-dehalogenase | GroEL/S |
| 38.5    | gatD | Galactoside-1-phosphate 5-dehalogenase | c        |
| 13, 13.19 | GFP | Green fluorescent protein | cm        |
| 50.5    | gphK | Glycerol kinase | c, GroEL/S |
| 77, 91, 92 | Glycerol transport ATP-binding protein proV | c        |
| 49.4    | glmA | Glutamine synthetase | cm        |
| 70      | gpmA, ibpA | GTP-binding protein gpmA | c        |
| 43      | hflK, hflA | HflK protein | cm        |
| 45      | ruvB | Holliday junction DNA helicase ruvB | c        |
| 80.1    | mltA, malA | HTH-type transcriptional regulator mltA | c        |
| 36      | ycgV | Hypothetical protein ycgV | c        |
| 92      | yehC | Hypothetical protein yehC | c        |
| 123, 149 | yehW | Hypothetical protein yehW | cm        |
| 145     | yfaF | Hypothetical protein yfaF | c        |
| 126     | yfaA | Hypothetical protein yfaA | c        |
| 84      | yfaB | Hypothetical protein yfaB | c        |
| 37      | yfgM | Hypothetical protein yfgM | cm        |
| 142     | yfcP | Hypothetical protein yfcP | c        |
| 97      | ybgD | Hypothetical protein ybgD | cm        |
| 81, 136 | cscB, cscD | Inner membrane protein cscB | cm        |
| 29      | ldhA, ldhC, lde | Isocitrate dehydrogenase [NADP] | c        |
| 31      | lacI | Lactose operon repressor | c        |
| 83      | lexA, lexAB, spr, lexC, umuA | LexA repressor | c        |
| 15      | lyoD | Lysozyme | c        |
| 11      | maltK | Maltodextrin phosphorylase | c        |
| 47.6    | trg | Methyl-accepting chemotaxis protein III | cm        |
| 25      | nuoB | NADH-quinone oxidoreductase chain B | cm        |
| 63.1    | nuoC, nuoCD, nuoD | NADH-quinone oxidoreductase chain CD | cm        |
| 34      | nikD | Nickel transport ATP-binding protein nikD | cm        |
| 74      | pseA | Phage shock protein A | cm        |
| 133     | phoH | Phosphatase | cm        |
| 119, 134 | ybeE | PhoH-like protein | c        |
TABLE I—continued

| spot 2D | band 1D | in solut. rank | gene name(s) | name | local. | prec. |
|--------|---------|---------------|--------------|------|--------|-------|
|        |         | a             | ppc, glu     | Phosphoenolpyruvate carboxylase | c     |        |
| 9      | 32.9    | b             | ppaA, pps    | Phosphoenolpyruvate synthase    | c     |        |
| 43     | 45      | A             | deoB, dcm, thyR | Phosphoenolpyruvate synthase | c     |        |
|         | 142     | B             | pnp          | Polynucleotide nucleolydyltransferase | c     |        |
| 114    |         |               | yhbG         | Probable ABC transporter ATP-binding protein | c     |        |
| 147    |         |               | phtB         | Probable phospholipase | c     |        |
| 52     |         |               | groS, drpA   | Probable RNA polymerase | c     |        |
| 27     |         |               | ybgI         | Protein ybg | c     |        |
| 38     |         |               | yclE         | Protein yclE | c     |        |
| 71     | 33      |               | gatA         | PTS system, galactose-specific IIA component | c     |        |
| 118    |         |               | manX, gptB, gplE | PTS system, mannose-specific IIA component | c     |        |
| 37.6   |         |               | ybiD         | Putative glycosyl transferase | c     |        |
| 90     | 14      |               | gatZ         | Putative phosphotransferase | c     |        |
| 58.8   |         |               | poxB         | Putative polypeptide | c     |        |
| 44     |         |               | aceE         | Putative polypeptide | c     |        |
| 4      |         |               | rraA, manG   | Regulator of ribonucleolytic activity | A     | c     |
| 14.1   |         |               | scoS         | Regulator of ribonucleolytic activity | A     | c     |
| 29     |         |               | dnaB, groP, gpaA | Replicase DNA helicase | c     |        |
| 89     |         |               | rrd          | Ribonucleolytic | O     |        |
| 37.6   |         |               | mreB, envB, rpoY | Rod shape-determining protein | c     |        |
| 60     | 19.4    |               | minD         | Septum site-determining protein | c     |        |
|        | 12, 13, 17 |               | lipB         | Signal peptidase | I     | cm     |
| 82     |         |               | sbb, extB, iecC | Single-strand binding protein | c     |        |
| 7      |         |               | sspB         | Stringent starvation protein | B     | c     |
| 116    |         |               | sodB         | Superoxide dismutase | Fe    | c     |
| 125    |         |               | gagY         | Tognazza-1,8-bisphosphate aldolase | c     |        |
| 19.4   |         |               | gadE         | GADH | c     |        |
| 148    |         |               | manR         | Transcriptional regulator manR | c     |        |
| 25, 13.3 |         |               | basR, pmrA | Transcriptional regulator protein basR | pmrA | c     |
| 88     |         |               | ompR, kmt, ompB | Transcriptional regulatory protein | ompR | c     |
| 41     |         |               | phoP         | Transcriptional regulatory protein phoP | c     |        |
| 87     |         |               | rfaA, argT | Transcriptional regulatory protein | rfaA | c     |
| 28     |         |               | fadA, fadG | Transketolase | 1     | c     |
| 17     |         |               | treC, digH | Transaldolase-5-phosphate hydrolase | c     |        |
| 48.2   | 36      |               | tnaA         | Tryptophanase | c     |        |
| 151    |         |               | trpS         | Tryptophanase | c     |        |
| 104    |         |               | hisG, hsdB | Type I restriction enzyme | Ecoli | c     |
| 120, 132, 135 |         |               | udg, pmrF, udg | UDP-glucose 5-dehydrogenase | c     |        |
| 139    |         |               | ugpG         | Universal stress protein G | c     |        |
| 146    |         |               | ugd          | Uridine phosphorylase | c     |        |

| empty vector control cells |
|---------------------------|
| 6                         | ompA, con, tolG, tolF | Outer membrane protein A [Precursor] | om |
| 1, 2                      | ompL, cmlB, coa, cry, tolF | Outer membrane protein F [Precursor] | om |
| 3                         | glpB          | 1,4-alpha-glucan branching enzyme | c |
| 5                         | fufA, hufB      | Elongation factor Tu | c | GroEL |
| 4                         | gspA          | Glycogen synthase | c |

a The numbering corresponds to the bands/spots in the 1D/2D gels shown in Fig. 6, A and B. Only prominent bands in the 1D gel are numbered. The identification number of minor bands corresponds to their observed molecular weight.
b The ranking is on the basis of the Mascot MOWSE score of proteins identified by in-solution digestion and nano-LC-ESI-MS/MS (Supplemental Table 2).
c Gene names and synonyms given in the Swiss-Prot E. coli 48.1 database (release September 27, 2005; www.expasy.org).
d Localization according to the Swiss-Prot database. The localization of unknown proteins was predicted using PSORT (77). c, cytoplasm; p, periplasm; om, outer membrane; cm, cytoplasmic membrane.
"Chaperone substrates according to the following references: GroEL/ES; Kerner et al. (56); DnaK; Mogk et al. (55).
M" indicates that MALDI-TOF MS peptide mass fingerprint analysis revealed a peptide mass for a peptide covering parts of a signal sequence. "Q" indicates that this peptide was sequenced by nano-LC-ESI-MS/MS.

1540 Molecular & Cellular Proteomics 6.9
and 10 and Supplemental Table 3). Interestingly the IPTG-induced lactose permease (LacY) was 3-fold increased upon GST-GFP overexpression, whereas no significant increase of LacY was observed upon overexpression of the three different membrane proteins, indicating that its insertion into the cytoplasmic membrane was most likely outcompeted by the overexpressed membrane protein. The overexpression of membrane proteins led to reduced expression levels of a large set of endogenous membrane proteins. From here on, we focus on the results obtained with cytoplasmic membranes isolated from cells overexpressing YidC-GFP. The levels of the following protein complexes were strongly affected (Figs. 9 and 10A): formate dehydrogenase O (FdoG and FdoH; −75%), maltose permease (MalK; −50%), oligopeptide permease (OppB, OppD, and OppF; −75%), glutamine permease (GlnG; −70%), cytochrome d ubiquinol oxidase (CydA; −55%), and succinate dehydrogenase (DhsA and DhsB; −50%).

Enzymatic activity assays monitoring both cytochrome oxidases (bo3 and bd) showed that cytochrome oxidase activity was lowered by 52% in the overexpressor (Fig. 10C). In this respect it should be mentioned that the accumulation levels of the subunits of cytochrome bo3 were reduced just outside the stringent threshold of acceptance in the 2D BN-PAGE gels. Also the capacity of whole cells to consume oxygen was decreased by 62% upon membrane protein overexpression (Fig. 10C). Furthermore enzymatic activity assays showed that succinate dehydrogenase activities were lowered by 43% in the cytoplasmic membranes of the YidC-GFP overexpressor (Fig. 10C). The observed drop in enzymatic activities thus matched the protein accumulation levels determined from the gels. Similar observations were also made for the other membrane protein overexpressors (results not shown).

Overexpression of membrane proteins did not lead to changes in the composition of membrane protein complexes with one notable exception: overexpression of all three membrane proteins resulted in the change of the complex size of endogenous YidC (OxaA) from 109 to 89 kDa. At present, we do not have an explanation for this observation. In keeping with the Western blotting results, membrane protein overexpression led to a 4-fold stronger recruitment of DnaK to the cytoplasmic membrane. It is tempting to speculate that DnaK is attracted to the membrane by misfolded cytoplasmic loops of membrane proteins.

**DISCUSSION**

Overexpression of membrane proteins is a major bottleneck in membrane protein research. *E. coli* is the most widely used vehicle for the production of membrane proteins. To facilitate isolation of membrane proteins it is best to overexpress them in the cytoplasmic membrane rather than in inclusion bodies (6). Unfortunately such membrane-localized accumulation is usually toxic to the cell, thereby severely reducing overexpression yields. Although several reasons can be postulated, the underlying physiology for this toxicity is not understood, thus preventing the design of rational strategies to improve membrane protein overexpression yields. In this study, we characterized *E. coli* cells overexpressing three different membrane protein-GFP fusions in the cytoplasmic membrane. Overexpression of a soluble GFP fusion protein was used to distinguish between general overexpression effects and effects related to overaccumulation in the cytoplasmic membrane. Several complementary proteomics approaches were refined and adapted for this study and complemented by flow cytometry, Western blotting, and pulse labeling. In particular, the development of “backed” BN-PAGE gels greatly improved the relative quantification of membrane protein complexes.

This study shows that the toxicity of membrane protein overexpression is primarily caused by a limiting Sec translocon capacity having a severe impact on both the cell envelope proteome and, surprisingly, also the cytoplasmic proteome. Two effects of the limiting Sec translocon capacity are especially noteworthy: (i) the aggregation of precursors of secretory proteins and of cytoplasmic proteins in the cytoplasm and (ii) the shifted and inefficient energy metabolism likely to be through redox activation of the Arc two-component system. Interestingly although the properties (such as number of TMDs and length) and functions of the overexpressed membrane proteins did not seem to influence the level of toxicity or proteome composition, there was a positive correlation with the amount of overexpressed material per cell. This is encouraging as it suggests that strategies to reduce toxicity and improve overexpression yields may be relatively independent of the type of membrane protein selected for overexpression.

**Fig. 7.** Western blotting analysis of purified cytoplasmic membranes. Cells overexpressing GFP fusion proteins and control cells were cultured, and cytoplasmic membranes were isolated as described under “Experimental Procedures.” Cytoplasmic membranes were separated by means of SDS-PAGE and subsequently subjected to Western blotting analysis with antibodies to SecY, SecE, SecG, L5, DnaK, FtsH, and HtpX.
Our observations are summarized in Fig. 11 and discussed in more detail below.

Toxicity Due to Limiting Capacity of the Sec Translocon—Upon membrane protein overexpression SecA levels were increased. SecA expression is up-regulated when the Sec translocon-dependent translocation of the secretion monitor SecM is hampered (59, 60). Therefore, increased SecA levels indicate that the Sec translocon capacity is not sufficient in cells overexpressing membrane proteins, albeit we did not observe decreased accumulation levels of Sec translocon components under these conditions. Quantitative analysis of the cytoplasmic membrane proteome showed that the levels...
of many endogenous membrane proteins and complexes in the bacterial cytoplasmic membrane were significantly reduced, indicating that the overexpressed membrane protein competes out endogenous membrane proteins. This is nicely illustrated by the observation that overexpression of the soluble protein GST-GFP leads to increased levels of the IPTG-induced LacY transporter in the cytoplasmic membrane, whereas membrane protein overexpression did not (13). Furthermore, the levels of membrane-associated ribosomes were increased upon membrane protein overexpression, whereas total levels of ribosomes in the cells did not change. This is in keeping with previous E. coli studies that showed that high rates of membrane protein translation result in more membrane-associated ribosomes (61, 62). Our data suggest that upon membrane protein overexpression most Sec translocons are engaged in co-translational protein translocation, i.e. the biogenesis of (overexpressed) membrane proteins. As a consequence, only a few Sec translocons will be available for post-translational protein targeting, i.e. the translocation of secretory proteins. Indeed the levels of the processed forms of many secretory proteins were diminished, whereas their precursors accumulated in the cytoplasm as aggregates. It is not likely that the slightly reduced levels of the targeting chaperone SecB account for precursor accumulation upon membrane protein overexpression because upon overexpression of GST-GFP SecB levels were reduced to the same extent, and precursor accumulation was not observed.

Taken together, the insufficient Sec translocon capacity has a severe impact on the composition and as a consequence also the functioning of the cell envelope as is evidenced by hampered cell division and reduced capacity of the respiratory chain. Surprisingly we did not detect up-regulation of markers for envelope stress, such as the periplasmic chaperone Skp or chaperone/protease DegP (data not shown and Refs. 63 and 64), possibly indicating that the signal transduction of cell envelope stress is compromised.

Toxicity Due to Aggregation of Cytoplasmic and Mistargeted Secretory Proteins in the Cytoplasm—We observed...
accumulation of protein aggregates in the cytoplasm of the membrane protein overexpressors but not upon overexpression of soluble GST-GFP. The overexpressed membrane proteins constituted only about 20% of the total aggregated protein. In this respect it should be mentioned that the T7 RNA polymerase-based system used to overexpress the membrane proteins is extremely powerful (65) and may compromise the coupling of transcription, translation, and targeting, leading to the mistargeting and accumulation of some of the overexpressed material in the cytoplasm. This may be the starting point for the formation of protein aggregates. Approximately one-quarter of the total protein in the aggregates was made up of precursor forms of secretory proteins supporting the above mentioned severe secretion defect caused by the
### TABLE II

Mass spectrometry identification of spots in 2D blue native gels of cytoplasmic membranes isolated from BL21(DE3)pLysS

Spots in 2D BN gels of cytoplasmic membranes (Figs. 8B and 9) were excised from gels stained with colloidal Coomassie. Proteins were identified by MS as described under "Experimental Procedures." Proteins belonging to the same complex have the same gray shading. acc., accession number; Local., localization; Nr., number.

| Spot No. (a) | Gene Name (b) | Swiss Prot acc. | Predicted MW (kDa) (c) | Observed MW (kDa) (d) | Predicted MW (kDa) (c) | Observed MW (kDa) (d) | Predicted MW (kDa) (c) | Observed MW (kDa) (d) |
|-------------|---------------|----------------|------------------------|----------------------|------------------------|----------------------|------------------------|----------------------|
| 1           | ywp            | P30295        | 112.0                  | 142.8                | 112.0                  | 142.8                | 112.0                  | 142.8                |
| 2           | kiaH           | P30295        | 112.9                  | 142.8                | 112.9                  | 142.8                | 112.9                  | 142.8                |
| 3           | fha            | P30295        | 112.9                  | 142.8                | 112.9                  | 142.8                | 112.9                  | 142.8                |
| 4           | hfr             | P30295        | 112.9                  | 142.8                | 112.9                  | 142.8                | 112.9                  | 142.8                |

* a The numbering corresponds to the spots in 2D BN gel images in Figs. 8B and 9.
* b Gene name was extracted from the Swiss-Prot database E. coli 48.1 (release September 27, 2005; www.expasy.org). # indicates that the protein assignment was ambiguous.
* c Localization and number of TMDs according to the Swiss-Prot database. The localization of unknown proteins was predicted using PSORT (77).
* d Protein sizes (in kDa) predicted from amino acid sequence. Two sizes are given for secretory proteins: the first size corresponds to the mature form of the protein after signal sequence processing, and the second size corresponds to the precursor form including the signal sequence.
* e Size of proteins calculated from the spot position on the 2D BN gels used for the analysis.
* f Native masses of proteins in the native first dimension calculated from the spot position on the 2D BN gels used for the analysis. It should be kept in mind that membrane protein complexes often run at higher apparent native masses than those calculated.

---

**Consequences of Membrane Protein Overexpression in E. coli**

Molecular & Cellular Proteomics 6.9

1545
Fig. 11. Consequences of membrane protein overexpression in *E. coli*. Typically membrane proteins are targeted via the SRP pathway to the SecYEG translocon where they are inserted co-translationally into the membrane (i–iv) (8); integral membrane (e.g. YidC), cytoplasmic (e.g. DnaK), and periplasmic (e.g. DegP) chaperones may assist the folding of membrane proteins (v). It should be noted that DegP can also...
overexpression of membrane proteins. Analysis of aggregated secretory proteins by the aggregation propensity prediction program Tango showed that the signal sequences of these secretory proteins are excellent seeds for aggregation (66) (data not shown). The remaining proteins in the aggregate preparations were mainly chaperones and proteases, which are most likely to be involved in the disaggregation, refolding, and degradation of the protein aggregates (67–69). The sequestration of chaperones, like DnaK and GroEL/S, to the aggregates reduces their availability and is likely the reason for aggregation of proteins that require these chaperones for proper folding under normal conditions. Aggregation of essential proteins, like the cell division protein MinD or the elongation factor Tu, may contribute significantly to the toxicity of membrane protein overexpression (70). Surprisingly hardly any endogenous membrane proteins could be detected in the aggregates. This is possibly due to efficient degradation by SsrA mRNA-dependent tagging of stalled nascent chains of co-translationally targeted membrane proteins and subsequent turnover by the Clp protease machinery (71).

Taken together, overexpression of membrane proteins in the cytoplasmic membrane severely hampers protein homeostasis in the cytoplasm. This is likely to be another prime, and unexpected, reason for the toxicity of membrane protein overexpression.

Toxicity Due to a Changed and Less Efficient Energy Metabolism—Membrane protein overexpression strongly affected the composition of the cytoplasmic membrane proteome; e.g., levels of the key respiratory chain complexes succinate dehydrogenase and cytochrome bd and bo_3 oxidases were reduced. The state of the respiratory chain in E. coli is monitored by a number of redox sensor systems to maintain redox homeostasis (72). One of these systems is the Arc two-component system, which mediates adaptive responses to changing respiratory conditions by monitoring the redox state of the Q pool. Our proteomics data, in particular the induction of the acetate-pta pathway, indicate that the Arc system is activated upon membrane protein overexpression, which typically occurs when the Q pool is mostly in a reduced state (Fig. 11). Together with the up-regulation of pyruvate dehydrogenase, this may lead to an increased generation of ATP via the acetate-pta pathway resulting in secretion of acetate into the medium. Indeed the pH of the culture medium was lowered upon membrane protein overexpression. Also in keeping with Arc activation were the lowered accumulation levels of the tricarboxylic acid cycle enzymes FumA, FumC, Mdh, and Sdh as well as the glyoxylate bypass protein isocitrate lyase (73). The reduced expression levels of cytochrome bd oxidase, whose expression is induced by ArcA-P, is likely to be caused by hampered biogenesis. The question remains whether reduced Sdh levels were the result of compromised biogenesis or whether its levels are down-regulated due to diminished respiration. Because the E. coli cells in our experiments were grown under fully aerated conditions, oxygen limitation cannot be the reason for Arc induction upon membrane protein overexpression. It is therefore not clear whether the Arc system is induced by a reduced Q pool or whether it is somehow “misled” by effects of membrane protein overexpression. Taken together, ATP, a central and highly demanded commodity for the overproduction of proteins and formation of biomass, is produced inefficiently under the conditions usually used to overexpress membrane proteins.

Future Perspectives—The comparative proteome analysis of aggregates, purified cytoplasmic membranes, and whole cell lysates using native as well as denaturing gels provides important physiological insight into toxicity of membrane protein overexpression in E. coli. Almost all effects were similar for the different overexpressed membrane proteins irrespec-
Consequences of Membrane Protein Overexpression in *E. coli*

tive of their biochemical properties. Thus, it should be possible to design generic strategies to improve membrane protein overexpression yields in *E. coli*. Based on our observations, moderate and timed co-expression of Sec translocon components may alleviate an important bottleneck. This should improve the capacity of membrane protein integration as well as translocation of secretory proteins and thus reduce cell envelope stress and formation of protein aggregates. Membrane protein quality control and degradation may also be promising targets for strain engineering. So far, two main integral membrane proteases have been identified, FtsH and HtpX. It is unlikely that HtpX-mediated proteolysis poses a major problem for membrane protein overexpression as HtpX levels were strongly reduced upon membrane protein overexpression. However, FtsH levels were not significantly affected, and its activity could very well interfere with membrane protein overexpression. It would be interesting to determine whether co-expression of the *Bacillus subtilis* peptide SpoVM, a competitive inhibitor of the *E. coli* FtsH activity, can improve membrane protein overexpression yields (74). Furthermore aggregation of proteins in the cytoplasm appears to affect membrane protein overexpression by reducing the fitness of the overexpression host. Indeed preliminary experiments indicate that by lowering the formation of aggregates in the cytoplasm membrane protein overexpression levels increase considerably.\(^3\) Under the culture conditions used for membrane protein overexpression, redox sensors are likely to be dispensable. Therefore, the deletion or redirection of sensory systems, e.g., Arc, may improve overexpression yields by preventing the misled activation of stress responses upon membrane protein overexpression. In conclusion, our study provides an important platform for designing rational strategies to improve membrane protein overexpression yields.

Acknowledgments—We are indebted to Petra Björk for assistance with microscopy, to Joachim Reimann for assistance with the enzymatic activity assays, to Nadja Meindl-Beinker for helpful discussions about BN-PAGE, and to Ross Dalbey for the YidC mutants. Mirjam Bader Lerch, Mirjam Klepsch, Joen Luirink, Dirk Jan Slotboom, and David Wickström are thanked for critically reading the manuscript.

\(^*\) This work was supported by grants from the Swedish Research Council, the Carl Tryggers Stiftelse, the Marianne and Marcus Wallenberg Foundation, and the European Molecular Biology Organization Young Investigator Programme (to J.-W. d. G.) and a grant from The Swedish Foundation for International Cooperation in Research and Higher Education (STINT) (to J.-W. d. G. and K. J. v. W.). Proteomics infrastructure was supported by a grant from NYSTAR (to K. J. v. W.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

\(\ddagger\) Present address: Affibody AB, Box 20137, SE-16102 Bromma, Sweden.

\(\S\) To whom correspondence should be addressed. Tel.: 46-8-162420; Fax: 46-8-153679; E-mail: degier@dubb.su.se.

REFERENCES

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

2. Krogh, A., Larsson, B., von Heijne, G., and Sonnhammer, E. (2001) Predicting transmembrane protein topology with a hidden Markov model. Application to complete genomes. *J. Mol. Biol.* 305, 567–580

3. Klabunde, T., and Hessler, G. (2002) Drug design strategies for targeting G-protein-coupled receptors. *Chembiochem* 3, 928–944

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

5. Grishamshamer, R. (2006) Understanding recombinant expression of membrane proteins. *Curr. Opin. Biotechnol.* 17, 337–340

6. Drew, D., Fröderberg, L., Baars, L., and de Gier, J. W. (2003) Assembly and overexpression of membrane proteins in Escherichia coli. *Biochim. Biophys. Acta* 1610, 3–10

7. Drew, D., Slotboom, D. J., Friso, G., Reda, T., Geneveaux, P., Rapp, M., Meindl-Beinker, N. M., Lambert, W., Lerch, M., Daley, D. O., Van Wijk, K. J., Hirst, J., Kunji, E., and De Gier, J. W. (2005) A scalable, GFP-based pipeline for membrane protein overexpression screening and purification. *Protein Sci.* 14, 2011–2017

8. 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

9. 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

10. Luirink, J., and Sinning, I. (2004) SRP-mediated protein targeting: structure and function revisited. *Biochim. Biophys. Acta* 1694, 17–35

11. Osborne, A. R., Rapport, T. A., and van den Berg, B. (2005) Protein translocation by the Sec61/SecY channel. *Annu. Rev. Cell Dev. Biol.* 21, 529–550

12. Andersson, H., and von Heijne, G. (1993) Sec dependent and sec independent assembly of E. coli inner membrane proteins: the topological rules depend on chain length. *EMBO J.* 12, 683–691

13. Nagamori, S., Smirnova, I. N., and Kaback, H. R. (2004) Role of YidC in folding of polytopic membrane proteins. *J. Cell Biol.* 165, 53–62

14. van der Laan, N., Nouwen, N. P., and Driessen, A. J. P. (2005) YidC—an evolutionary conserved device for the assembly of energy-transducing membrane protein complexes. *Curr. Opin. Microbiol.* 8, 182–187

15. Randall, L. L., and Hardy, S. J. (2002) SecB, one small chaperone in the complex milieu of the cell. *CMLS Cell. Mol. Life Sci.* 59, 1617–1623

16. Mori, H., and Ito, K. (2001) The Sec protein-translocation pathway. *Trends Microbiol.* 9, 494–500

17. Valent, Q. A., Scotti, P. A., High, S., de Gier, J. W., von Heijne, G., Lentzen, G., Wintemeyer, W., Oudega, B., and Luirink, J. (1998) The Escherichia coli SRP and SecB targeting pathways converge at the translocon. *EMBO J.* 17, 2504–2512

18. Zolkiewski, M. (2006) A camel passes through the eye of a needle: protein unfolding activity of Cip ATPases. *Mol. Microbiol.* 61, 1094–1100

19. Sakoh, M., Ito, K., and Akiyama, Y. (2005) Proteolytic activity of HtpX, a membrane-bound and stress-controlled protease from Escherichia coli. *J. Biol. Chem.* 280, 33305–33310

20. Ito, K., and Akiyama, Y. (2005) Cellular functions, mechanism of action, and regulation of FtsH protease. *Annu. Rev. Microbiol.* 59, 211–231

21. Waldo, G. S., Standish, B. M., Berendzen, J., and Terwilliger, T. C. (1999) Rapid protein-folding assay using green fluorescent protein. *Nat. Biotechnol.* 17, 691–695

22. Hewitt, C. J., and Nebe-Von-Caron, G. (2004) The application of multi-parameter flow cytometry to monitor individual microbial cell physiological state. *Adv. Biochem. Eng. Biotechnol.* 89, 197–223

23. Gunderson, C. W., and Segall, A. M. (2006) DNA repair, a novel antibacterial target: Holliday junction-trapping peptides induce DNA damage and chromosomal segregation defects. *Mol. Microbiol.* 59, 1129–1148

24. Baars, L., Ytterberg, A. J., Drew, D., Wagner, S., Thilo, C., van Wijk, K. J.,...
68. Carrio, M. M., and Villaverde, A. (2003) Role of molecular chaperones in inclusion body formation. FEBS Lett. 537, 215–221
69. Weibezahn, J., Tessarz, P., Schlieker, C., Zahn, R., Maglica, Z., Lee, S., Zentgraf, H., Weber-Ban, E. U., Dougan, D. A., Tsai, F. T., Mogk, A., and Bukau, B. (2004) Thermotolerance requires refolding of aggregated proteins by substrate translocation through the central pore of ClpB. Cell 119, 653–665
70. Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Dat-senko, K. A., Tomita, M., Wanner, B. L., and Mori, H. (2006) Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. Mol. Syst. Biol. 2, 2006.0008
71. Withey, J. H., and Friedman, D. I. (2003) A salvage pathway for protein structures: tmRNA and trans-translation. Annu. Rev. Microbiol. 57, 101–123
72. Green, J., and Paget, M. S. (2004) Bacterial redox sensors. Nat. Rev. Microbiol. 2, 954–966
73. Liu, X., and De Wulf, P. (2004) Probing the ArcA-P modulon of Escherichia coli by whole genome transcriptional analysis and sequence recognition profiling. J. Biol. Chem. 279, 12588–12597
74. Cutting, S., Anderson, M., Lysenko, E., Page, A., Tomoyasu, T., Tatsumoto, K., Tatsuta, T., Kroos, L., and Ogura, T. (1997) SpoVM, a small protein essential to development in Bacillus subtilis, interacts with the ATP-dependent protease FtsH. J. Bacteriol. 179, 5534–5542
75. Wu, T., Malinverni, J., Ruiz, N., Kim, S., Silhavy, T. J., and Kahne, D. (2005) Identification of a multicomponent complex required for outer membrane biogenesis in Escherichia coli. Cell 121, 235–245
76. Spiess, C., Beil, A., and Ehrmann, M. (1999) A temperature-dependent switch from chaperone to protease in a widely conserved heat shock protein. Cell 97, 339–347
77. Nakai, K., and Kanehisa, M. (1991) Expert system for predicting localization sites in Gram-negative bacteria. Proteins: Struct. Funct. Genet. 11, 95–110