Supplemental Information
for:

Physiological Response to
Membrane Protein Overexpression in E. coli

Francesca Gubellini 1, Grégory Verdon 1†, Nathan K. Karpowich 1†, Jonathan D. Luff 1, Grégory Boel 1, Nils Gauthier 1†, Samuel K. Handelman 1†, Sarah Ades 2, and John F. Hunt 1§.

1 Department of Biological Sciences, 702A Fairchild Center, MC2434, Columbia University, New York, NY 10027, USA; and
2 Department of Biochemistry and Molecular Biology, Pennsylvania State University, University Park, PA 16802, USA.

§ To whom correspondence may be addressed. E-mail: jfhunt@biology.columbia.edu. Voice: (212)-854-5443. FAX (212)-865-8246.

† Present addresses: Unité G5 Biologie structurale de la sécrétion bactérienne, Institut Pasteur, 25-28 rue du Docteur Roux, Bâtiment Metchnikoff, 3ème étage, 75015 Paris, France (FG); Department of Physiology and Biophysics, Weill Cornell Medical College, New York, NY 10065, USA (GV); Kimmel Center for Biology and Medicine at the Skirball Institute of Biomolecular Medicine, and Department of Cell Biology, New York University School of Medicine, 540 First Avenue, New York, NY 10016, USA (NKK); National University of Singapore, T-Lab, #05-01, 5A Engineering Drive 1, Singapore 117411 (NG); and Dept. Biomedical Informatics, Mathematical Biology Institute, Jennings Hall Rm. 14-376, Ohio State University, 1735 Neil Avenue, Columbus, OH 43210, USA (SKH).

Running Title: Membrane Protein Overproduction in E. coli.
Supplemental Text

Cellular metabolism and pH homeostasis during protein overexpression. E. coli cells growing in LB experience sequential changes in the expression of genes for metabolic enzymes, which presumably reflect trends in the utilization of metabolic resources (1-3). It is believed that energy is generated first from readily metabolized sugars, next from more complex sugars and single amino acids, and finally from small polypeptides. The organic acids acetate and lactate are generated as a by-product of carbohydrate metabolism and secreted early in the growth program, contributing to mild acidification of the medium, but they are later re-internalized to be utilized as more readily metabolized species are depleted. Their absorption, combined perhaps with secretion of ammonia as a by-product of amino acid metabolism, contributes to mild alkalinization at later times under standard growth conditions in unbuffered LB. The control cells and protein-overexpressing cells in our study generally display similar levels of transcripts for metabolic enzymes, suggesting that the latter are executing the standard program for utilization of metabolic resources even though there is obvious blockage of many of the transcriptional and physiological changes associated with entry into stationary phase. Notably, there are no systematic perturbations in the expression of operons encoding enzymes involved in phospholipid biosynthesis, protein biosynthesis, nucleic acid biosynthesis, carbohydrate or amino acid catabolism, or aerobic or anaerobic respiration. These observations suggest that all of these core metabolic processes are proceeding similarly in protein-overexpressing and control cells and that there are not metabolic barriers to high-level expression of either soluble or membrane proteins for cells growing with standard aeration in LB.

Perturbations are observed in the levels of some transcripts encoding certain specific metabolic enzymes even in the absence of systematic changes in the pathways in which they participate. Noteworthy among these are a reduction in the levels of the transcripts for dipeptide permease (dpp) and acetyl-CoA synthetase (acs). While dpp transcript levels drop, there are no significant changes in those encoding the components of oligopeptide permease (opp genes), which represents the other main pathway for utilization of small peptides. Future research will be required to determine whether the overexpressing cells have depleted dipeptides from the medium more rapidly or whether the differential behavior of the dpp vs. opp systems reflect an alteration in metabolic strategy. While the overexpressing cells show a reduction in the level of the transcript for acs, which plays a major role in acetate utilization at later growth times, there are no reductions in the levels of the transcripts for enzymes participating in lactate re-internalization and utilization. While a blockage in acetate uptake could contribute to blocking alkalinization of the overexpressing cultures, these cultures also show strongly suppressed expression of the genes involved in acid-resistance that generally rise upon the transition into stationary phase. These include the GadX/GadE-regulated genes gadA and gadB that encode glutamic acid decarboxylases that contribute to raising pH by consuming protons to release CO₂. Additional research will be required to understand the factor(s) controlling the altered pH homeostasis in the protein-overexpressing cells. However, this alteration is not an impediment to a high-level expression of either soluble or membrane proteins.
More generally, our results raise questions concerning the manner in which metabolic resources control the developmental program of *E. coli*. Control cells and those overexpressing the highest levels of the induced target proteins display indistinguishable growth properties, despite large differences in gene-expression profile, striking alterations in cytoplasmic morphology in negatively stained thin-section EM images, and significant differences in the pH of the growth medium. These observations suggest the possibility of a complex interplay in which developmental processes significantly influence the utilization of metabolic resources rather than being directly controlled by their availability.
Supplemental Experimental Procedures

**Molecular cloning.** PfuUltra Hotstart DNA polymerase (Agilent Technologies, Santa Clara, CA) was used for gene amplification. Plasmids were propagated during construction in *E. coli* strain DH5α, and protein constructs were confirmed by DNA sequencing (Genewiz Inc., South Plainfield, NJ).

**Cell growth.** Standardized growth methods were used for all experiments including pH measurements, protein expression tests, microscopy studies, reporter-gene assays, and microarray studies. Pre-cultures from a single colony on a fresh LB plate were grown at 37°C overnight in LB medium supplemented with 100 μg/ml ampicillin and 50 μg/ml kanamycin prior to a 1:37 dilution into a fresh aliquot of the same medium. Experiments were conducted on the resulting cultures, which were grown aerobically at 37°C with 250 rpm shaking. The optical density at 600 nm (OD₆₀₀) of these cultures was monitored during all experiments, with cultures diluted to OD₆₀₀ values between 0.05 and 0.5 prior to measurement. Cultures showing evidence of genetic selection producing inconsistency in growth-rate were discarded without further analysis. Protein expression was induced at OD₆₀₀ 0.6-0.8 by addition of 1 mM of isopropyl β-D-1-thiogalactopyranoside (IPTG), and cells were harvested 3 hours after induction. Cell transformations with pBAD plasmids carrying *Ec*MsbA constructs were conducted in the presence of 100 μg/ml ampicillin and 0.5% (w/v) glucose to repress protein expression from the arabinose promoter. For experiments in W3110A and WD2S cells, single colonies were grown in LB supplemented with 100 μg/ml ampicillin, 12 μg/ml tetracycline, and 0.5% (w/v) glucose for 12 hours at room temperature prior to ~10,000-fold dilution in 10 ml of LB containing the same antibiotics and either 0.5% (w/v) glucose or 0.02-0.20% (w/v) L-arabinose. These cultures were grown aerobically in a water-bath at 42°C, with OD₆₀₀ monitored every 30 minutes starting 90 minutes after inoculation; samples for SDS-PAGE analysis were collected 7 hours after inoculation from cultures induced with 0.2% (w/v) L-arabinose.

**Protein expression, solubilization, and purification tests.** Small-scale cultures were grown and induced in LB as described above prior to centrifugation of 8 ml of cells, which were resuspended in 1ml of Lysis Buffer (200 mM NaCl, 5 mM DTT, 20 mM Tris, pH 7.8) and disrupted by sonication with a microtip probe. Unlysed cells were collected by centrifugation for 8 minutes at 8,000 rpm, and the supernatant was centrifuged for 1 hour at 15,000 to produce the supernatant (soluble) and pellet (membrane) fractions used for further analyses. All manipulations after initial cell growth were performed on ice or at 4°C. Detergent solubilization and Ni-NTA batch purification tests were performed on cells from 1 L cultures, which were resuspended in 3 ml of Lysis Buffer prior to lysis by probe sonication. Unlysed cells were collected by centrifugation for 10 minutes at 4,000 rpm, and the resulting supernatant was centrifuged for 30 minutes at 45,000 rpm in a Beckman TLA 100.4 rotor. The pellet from this spin was resuspended in lysis buffer at a protein concentration of 10 mg/ml based on Bradford assay. Samples from each induced culture were divided into three aliquots and agitated overnight at 4°C after addition of 2% (w/v) β-dodecylmaltoside (βDDM), lauryldimethylamine-N-oxide (LDAO), or fos-choline-12 (FC12). The supernatant from a 30-minute, 45,000 rpm spin of this material in a Beckman TLA 100.4 rotor was incubated overnight at 4°C with 100 μl of Ni-NTA agarose resin (QIAexpressionist Manual, Qiagen Inc., Valencia, CA). The resin was
washed with 60 mM imidazole prior to elution with 100 µl of 450 mM imidazole containing either 0.05% (w/v) βDDM 0.05%, 0.1% (w/v) LDAO, or 0.25% FC12.

**Optical and electron microscopy.** For visible microscopy, 500 µl aliquots were collected by centrifugation from cultures grown and induced for 3 hours as described above, and pellets were washed twice with DPBS buffer (GIBCO, Invitrogen, Carlsbad, CA) (2.7mM KCl, 1.5mM KH₂PO₄, 138 mM NaCl, 8mM Na₂HPO₄·7H₂O). Fixation of cells before or after staining gave identical results in initial assays, so replicate assays were conducted using just the latter procedure. The washed cells were resuspended in 1 ml of 200 nM Mitotracker Green (Invitrogen) in DPBS and incubated in the same solution for 15 minutes at room temperature. Stained cells were centrifuged for 2 minutes at 4,000 rpm, washed twice in 1 ml of DPBS, resuspended in 1 ml of 3.7% (v/v) formaldehyde in PBS and incubated 15 minutes at room temperature. Fixed cells were then washed twice in PBS and resuspended in PBS in a final volume of 200 µl (EcGlpT and HP1206*), 300 µl (EcMsbA* and StMsbA) or 400 µl (non-induced and induced empty pQE-60 controls, NBD-EcMsbA, EcEnolase*, EcYojI, and EcYojI**) prior to agitation overnight at 4°C. Immediately before imaging, FM4-64 (Invitrogen) was added at 5 µg/ml. Cells were mounted on freshly prepared, polylysine-coated coverslips and imaged using a Photometrix CoolSNAP camera. For electron microscopy, thin sections were prepared and stained at the Microscopy Image Core Facility at the NYU Langone Medical center using the procedure of Arechaga et al. (4) with minor modifications. Briefly, bacteria were fixed with 2.5% (v/v) glutaraldehyde and 2% (w/v) paraformaldehyde prior to washing with 50 mM cacodylate, pH 7.2. A second fixation step was performed with 2% (w/v) osmium tetroxide, before washing with Kellenberger buffer (5). Cell pellets were embedded in 2% (w/v) agar and then cut and stained in the dark with 0.5% (w/v) uranyl acetate. Samples were dehydrated with alcohol, transferred to mixtures of Epon and propylene oxide, and finally embedded in EMbed 812 (Electron Microscopy Sciences, Hatfield, PA). Thin sections were cut, adsorbed on electron microscope grids and stained with 2% uranyl acetate and lead citrate. Stained grids were then imaged in a Philips CM120 transmission electron microscope (FEI; Eindhoven, The Netherlands) and photographed with a Gatan (4kx2.7k) digital camera (Gatan Inc., Pleasanton, CA). Flagella were visualized in images of whole bacterial cells deposited directly on a carbon grid, stained twice for 3 minutes with 2% (w/v) uranyl acetate, and imaged in a Jeol 100 CX TEM (Jeol Ltd. Tokyo, Japan).

**RNA extraction and microarray analysis.** RNA was extracted from either 4 ml (induced empty pQE-60 control, NBD-EcMsbA, EcEnolase*, EcYojI, and EcYojI**) or 8 ml (EcMsbA*, StMsbA, EcGlpT, and HP1206*) of cultures grown and induced with ITPG for 3 hours as described above. At the end of the induction period, cultures were added to RNAprotect Bacteria Reagent (Qiagen, Valencia, CA) and frozen prior to extraction of RNA with the RNeasy Mini Kit (Qiagen) using protocols 4 and 7 described in the RNAprotect Bacteria Reagent Handbook, including the DNaseI treatment. Following final elution of the purified RNA samples in 50 µl of 10 mM Tris-Cl, pH 8.0, their concentration was measured spectrophotometrically. All showed good 260/280 nm absorbance ratios and were verified not to have significant degradation when visualized on a 0.8% (w/v) RNase-free formaldehyde agarose gel. RNA concentration was normalized between the samples, and cDNAs were synthesized at 42°C using SuperScript II Reverse Transcriptase (Invitrogen). After a 90 minute reaction period
at 42˚ C, the reverse transcriptase was inactivated by a 15-minute incubation at 70˚ C, and the samples were treated for 10 minutes at 37˚ C with 0.02 U/µl RNase H (Invitrogen) and 0.02 U/µl RNase A (EpiCentre, Madison, WI). The resulting cDNAs were purified using the MinElute Purification Kit (Qiagen), fragmented using 0.01U/µl DNase I (EpiCentre) in the presence of 1 mM CaCl₂, and incubated for 10 minutes at 95-99˚ C to inactive the DNase. The resulting DNA samples, which showed primarily fragments 50-200 basepairs in length in agarose-gel electrophoresis, were biotinylated with 0.5 U/µl of Terminal Deoxynucleotidyl Transferase (NEB, Ipswitch, MA) in the presence of 25 µM Biotin-N₆-ddATP (Enzo Life Sciences, New York, NY) and 2.5 mM CoCl₂ for 2 hours at 37° C. At the Gene Expression Center of the University of Wisconsin Biotechnology Center, these samples were hybridized to E. coli 2.0 microarrays (Affymetrix, Santa Clara, CA) using an AFX HybOven480 and automated Fluidics450 Station and then analyzed using a GC 3000 G7 Scanner. Data were extracted from the scanned images using the Affymetrix GeneChip Command Console (GCOS), and the resulting raw data (.cel files) were processed using the RMA (Robust Multi-chip Average) algorithm implemented in the Affymetrix Expression Console, which yielded comparable results. Only data from E.coli K12 genes identified by ‘b’ numbers (6) were considered in the downstream analyses reported in this paper.

**Fisher’s Exact Test for potential transcription-factor involvement.** The mapping of previously characterized E. coli transcription factors to target genes was taken from RegulonDB (7) (http://regulondb.ccg.unam.mx/). A 3-fold change was used as the threshold for inclusion in the over- and under-expressed gene sets, and the probability of the genes controlled by a specific candidate transcription factor (TF) being found at random at the level observed in each of these sets was calculated using the hypergeometric distribution (as described at http://mathworld.wolfram.com/FishersExactTest.html). The null-hypothesis corresponds to the ratio of the count of all genes above the threshold to the total count of all E. coli MG1655 genes (4487). This calculation was performed for all TFs and all sigma factors archived in RegulonDb, separately for the over- and under-expressed gene sets. Correcting for the number of hypotheses evaluated in this way, a 5% confidence level corresponds to a Bonferoni-corrected p-value of 1.99 x 10⁻⁴. Probabilities were calculated from a single microarray dataset, but transcription factors are reported in Table 2 only if they had p-values below 10⁻³ in all replicate datasets and exceeded the Bonferoni-corrected 5% confidence level in at least one.
Supplemental References

1. Baev MV, Baev D, Radek AJ and Campbell JW: Growth of Escherichia coli MG1655 on LB medium: monitoring utilization of sugars, alcohols, and organic acids with transcriptional microarrays. Appl Microbiol Biotechnol 71: 310-6, 2006.

2. Baev MV, Baev D, Radek AJ and Campbell JW: Growth of Escherichia coli MG1655 on LB medium: determining metabolic strategy with transcriptional microarrays. Appl Microbiol Biotechnol 71: 323-8, 2006.

3. Baev MV, Baev D, Radek AJ and Campbell JW: Growth of Escherichia coli MG1655 on LB medium: monitoring utilization of amino acids, peptides, and nucleotides with transcriptional microarrays. Appl Microbiol Biotechnol 71: 317-22, 2006.

4. Arechaga I, Miroux B, Karrassch S, Huijbregts R, de Kruijff B, Runswick MJ and Walker JE: Characterisation of new intracellular membranes in Escherichia coli accompanying large scale over-production of the b subunit of F(1)F(o) ATP synthase. FEBS Lett 482: 215-9, 2000.

5. Kellenberger E and Rytter A: Cell wall and cytoplasmic membrane of Escherichia coli. J Biophys Biochem Cytol 4: 323-6, 1958.

6. Posfai G, Plunkett G, 3rd, Feher T, Frisch D, Keil GM, Umenhoffer K, Kolisnychenko V, Stahl B, Sharma SS, de Arruda M, Burland V, Harcum SW and Blattner FR: Emergent properties of reduced-genome Escherichia coli. Science 312: 1044-6, 2006.

7. Gama-Castro S, Salgado H, Peralta-Gil M, Santos-Zavaleta A, Muniz-Rascado L, Solano-Lira H, Jimenez-Jacinto V, Weiss V, Garcia-Sotelo JS, Lopez-Fuentes A, Porron-Sotelo L, Alquicira-Hernandez S, Medina-Rivera A, Martinez-Flores I, Alquicira-Hernandez K, Martinez-Adame R, Bonavides-Martinez C, Miranda-Rios J, Huerta AM, Mendoza-Vargas A, Collado-Torres L, Taboada B, Vega-Alvarado L, Olivera M, Olivera L, Grande R, Morett E and Collado-Vides J: RegulonDB version 7.0: transcriptional regulation of Escherichia coli K-12 integrated within genetic sensory response units (Gensor Units). Nucleic Acids Res 39: D98-105, 2011.

8. Bendtsen JD, Nielsen H, von Heijne G and Brunak S: Improved prediction of signal peptides: SignalP 3.0. J Mol Biol 340: 783-95, 2004.

9. Nielsen H, Engelbrecht J, Brunak S and von Heijne G: Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. Protein Eng 10: 1-6, 1997.

10. Krogh A, Larsson B, von Heijne G and Sonnhammer EL: Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. J Mol Biol 305: 567-80, 2001.

11. Emanuelsson O, Brunak S, von Heijne G and Nielsen H: Locating proteins in the cell using TargetP, SignalP and related tools. Nat Protoc 2: 953-71, 2007.

12. Doerrler WT, Reedy MC and Raetz CR: An Escherichia coli mutant defective in lipid export. J Biol Chem 276: 11461-4, 2001.
Table S1. Primers and restriction sites used for cloning and mutagenesis.

| Construct cloned into pQE- | Primer | Plasmid a | Fw          | Rv               |
|---------------------------|--------|-----------|-------------|------------------|
| E-Yol and E-Yol²         | pQI60 (Neol–BglII) | CTTGCCATGGAACTTCTTTGACTTTAGTCTGCG | GTAGATCTTTGGCCCTCCGGCAAC |
| E-MshA                   | pQI570 (Sphl–BanII) | ACCTCCATGCATGAAAGATCTCTTCACTG | GCCTGGGATCCCTTGCCCAAAGATTTTCTAG |
| HP1200*                  | pQI60 (Neol–BglII) | CATCCCATGGAAAAAATAAAAAATTTTACTTTAAATTAAC | GTAGATCTTGGAGATGTCCTCTTGTGAG |
| ErGlpT                   | pQI60 (Neol–BglII) | GCAACCTCTGCAAGGCTTGAATTTTAAAACGACCCCAACACGAAGCC | CCGAGATCTGGCTTGATGGCCTTGCGAACAT |
| ND9 E-MshA               | pQI70 (Sphl–BglII) | CCAGCATGCATGGGCACCGTGAAATTCCAGCAATGGC | GCCTGGGATCTCGTGGCCCAAAGATTTTCTAG |
| SNshA                    | pQI70 (Sphl–BanII) | GCATGCAATAGGAAAGATCTCTTACAGTGCGAG | AGATCTCTTGGCCCAAAGATTTTCTAG |

Table S1. Constructs cloned into pBAD-

| Construct cloned into pBAD- | Primer | Plasmid a | Fw          | Rv               |
|----------------------------|--------|-----------|-------------|------------------|
| pBAD-E-MshA                | pBAD-Myc-HisA (Neol–Kpm) | GGCCCAACATGTATACAGGAACAGAGATCTCTTACGCTG | CCGTAGGTACCTATCTGCGGCCAACACGGCATTTTACTG |
| Delta-N4                   | pBAD-Myc-HisA (Neol–Kpm) | GGCCCAACATGTATACAGGAACAGAGATCTCTTACGCTG | CCGTAGGTACCTATCTGCGGCCAACACGGCATTTTACTG |
| msh4 Restoring Mutagenesis³ | pBAD-Myc-HisA | CTACAGGAGAAATACACATGCTAAAGCGACAAAAGATC | AGATCTCTTGGCCCAAAGATTTTCTAG |
| msh4 Restoring Mutagenesis³ | pBAD-Myc-HisA | CTACAGGAGAAATACACATGCTAAAGCGACAAAAGATC | AGATCTCTTGGCCCAAAGATTTTCTAG |

ErMshA-HisTag-Nter ⁴

|   | ¹N   | ²N   | ³N   |
|---|------|------|------|
|   | GCTAACGGAGGAATTACCACTGGCCTATTAA | GGGATTACATGCTCAGCCTCATACATCATTATAAC | GCCATCATGCTCAGCCTCGAGAAGAACAAAAAG |
|   | CAGACAGATCTCTTACGCGAC | AGATCTCTTGGCCCAAAGATTTTCTAG |
|   | GCATCAGGAGAAATACACATGCTAAAGCGACAAAAGATC | AGATCTCTTGGCCCAAAGATTTTCTAG |

Sequencing of glpR

|   | ACCCTGGCGGTCTCTTGGCTG | CGTGGGTTAGGATGGCTT |

a Plasmid used for cloning. Restriction sites used for cloning are reported between parentheses.

b Mutagenesis to eliminate the mutation inserted in msh4 for cloning

³N-terminal His-Tag was added to msh4 inserted in pBAD by mutagenesis. Three mutagenesis PCR were performed using primers 1N, 2N and 3N respectively.
### Table S2

**Table S2. List of genes sharing up- and down-regulation in subsets of strains**

| DOWN-REGULATED | IMPs + *EcEnolase* | IMPs | IMPs except MsbA | Toxic IMPs |
|----------------|--------------------|------|------------------|-----------|
| All strains | micF               | rpoS | ygaA | ygbA |
| ybaT         | yciG               | rmF  | ybaY | dkgA |
| gadC         | yciE               | sodC | yspW | ydcS |
| gadB         | ydeF               | osmB | lyr  | yghX |
| gadD         | yegP               | osmE | csiD | yphA |
| slo           | yehE               | osmC | ddpA | ygbK |
| hdeA         | yhyA               | csgD | ygiG | yejG |
| hdeD         | yhyB               | treA | astD | ydcI |
| gadE         | yhyH               | ihaB | astA | ydcT |
| gadW         | yhyK               | csiE | astC | ydcU |
| ygiG         | yhyL               | ygaF | acaN | ydvV |
| dctR         | yhyM               | ygaG | cysK | yevB |
| mdtE         | yhyN               | ygiF | yglD | yodC |
| mdtF         | yhyO               | yhiD | mlrA | msyB |

| UP-REGULATED | IMPs + *EcEnolase* | IMPs | IMPs except MsbA | Toxic IMPs |
|--------------|--------------------|------|------------------|-----------|
| ecpD         | rpoD               | fimZ | ydfK | hyeA |
| ydeS         | smfD               | htrE | yecT | cadA |
| ydeT         | smfH               | espC | gspC | cadB |
| fhiC         | yctH               | espH | yfZ  | aaeB |
| ynjl         | flgB               | yfr  | yfB  | ycbR |
| yjdN         | flgC               | yahA | yjdM | yebR |
| yjhR         | flgD               | ysdA | ydhA | yaiV |
| flgE         | hspQ               | ycaG | ycsK | yevB |
| flgF         | hppX               | ycb  | ycb  | ycbR |
| flgG         | reuT               | ydl  | ydl  | ycbR |
| flhE         | ribE               | yfd  | yfd  | ycbR |
| flhF         | cspA               | yfa  | yfa  | ycbR |
| flhI         | yipA               | yia  | yia  | ycbR |
| flhJ         | yibG               | yia  | yia  | ycbR |
| flhK         | yicW               | yia  | yia  | ycbR |
| flhL         | yieW               | yia  | yia  | ycbR |
| flhM         | yizW               | yia  | yia  | ycbR |

---

**a** List of genes specifically regulated in the indicated subsets of strains compared to the control strain. Gene function is color coded. Purple: acid response. Red: stationary phase and motility. Blue: unknown function. Cyan: genes with various known functions. Note that some genes put in the unknown function category have had putative functions suggested based on sequence homology (e.g., *prfH* has homology to a biochemically characterized polypeptide chain release factor). However, no orthologous gene or protein has been functionally characterized according to the EcoCyc database (www.ecocyc.org) or a manual literature search.

**b** Transcripts down-regulated in all overexpressing strains.

**c** Transcripts down- and up-regulated in strains overexpressing membrane proteins and also in *EcEnolase* strain.

**d** IMPs stands for Inner Membrane Proteins.

**e** Transcripts down- and up-regulated in strains overexpressing IMPs.

**f** Transcripts up-regulated in strains overexpressing a subset of four IMPs excluding *EcMsbA* and *SmMsbA*.

**g** Transcripts up-regulated in strains overexpressing toxic IMPs.
Fig S1. Predicted locations of cleavable signal peptides and transmembrane α-helices in target IMPs. The primary sequences of the target proteins were analyzed using online bioinformatics resources. The top panel for each target protein shows the output from the program SignalP 3.0 (8, 9) (http://www.cbs.dtu.dk/services/SignalP), which predicts the location of cleavable signal peptides. The bottom panel shows the output from the program TMHMM (10) (http://www.cbs.dtu.dk/services/TMHMM-2.0), which predicts the locations of transmembrane α-helices. The SignalP analysis reports three scores: C estimates the probability of a position being the first in a mature protein after signal-peptide cleavage; S represents the quality of the signal peptide; and Y estimates the quality of the signal-peptidase cleavage site (11). Note that
none of these scores reaches the operative threshold (i.e., values of 0.52, 0.92 and 0.33 for C, S and Y respectively) for any of the target proteins. The TMHMM analysis reports the probabilities for each residue to reside in a transmembrane α-helix (red), in the cytosol (dark blue), or in the periplasm (magenta), based on summation of all possible paths through possible topology models. The schematic at the top of each TMHMM plot shows the most probable structure (using the same color scheme). The predicted locations of the transmembrane α-helices agree well with those observed in the crystal structures of GlpT and MsbA, although TMHMM identifies the final C-terminal α-helix in MsbA with only ~50% probability.
Figure S2. Faster-growing cell populations can spontaneously arise upon induction of toxic protein expression. OD$_{600}$ was used to monitor the growth of *E. coli* MG1655 cells during induction of protein overexpression as described in the Experimental Procedures. This experiment is equivalent to that shown in Fig. 1A in the main text except that the cultures overexpressing *EcGlpT* and *StMsbA* spontaneously generate more rapidly growing cells, presumably due to selection of toxicity-suppressing mutations. Growth curves like these were measured when preparing cells for all physiology experiments reported in this paper, and cultures were discarded rather than analyzed when evidence was observed of toxicity-resistant populations arising in cells containing plasmids expressing toxic proteins.
Figure S3. Western blot analysis of protein expression. Cell fractions were prepared and analyzed as in Fig. 1C. Proteins separated on a 15% SDS-PAGE gel were transferred electrophoretically to a nitrocellulose membrane and Western blotted with an anti-tetra-histidine antibody (Qiagen). The highest band detected in the sample has the proper molecular weight for the full-length C-terminally hexahistidine-tagged protein, while the lower bands are presumably N-terminal truncation products. The EcYojI** construct is not detected here because the C-terminal truncation in this construct eliminates the hexahistidine tag.
Figure S4. Solubilization and Ni-NTA recovery assays. (A) Coomassie-Blue-stained SDS-PAGE gel of total membranes and equivalent volumes of the corresponding detergent extracts. Cultures were grown as in Fig. 1 and harvested three hours after induction of protein expression with 1 mM IPTG. The detergents βDDM, LDAO and FC12 were used at a final concentration of 2% (w/v) during overnight extraction on ice. After clearance of insoluble material by ultracentrifugation, the supernatants were loaded on the gel. The equivalent of 3 ml of induced culture were loaded in each lane for EcMsbA*, 4 ml for EcGlpT, 4.5 ml for EcYojI, and 2 ml for EcYojI**. The EcGlpT protein is difficult to identify in these samples because of its relatively low expression level combined with co-localization with other proteins. (B) Coomassie-Blue-stained SDS-PAGE gel of the high-imidazole eluates from batch Ni-NTA purifications of the detergent extracts. The equivalent of 20 ml of induced culture were loaded in each lane for EcMsbA*, 30 ml for EcGlpT, and 15 ml for EcYojI. The batch purification protocol used here was optimized for parallel processing of multiple samples and gave lower yield than obtained via standard Ni-NTA purification of equivalent samples.
Figure S5. Size distribution of control cells during exponential growth and transition to stationary phase. (A) \( \text{OD}_{600} \) was used to monitor cell growth during the induction procedure, and samples of IPTG-induced empty-vector control cells were taken for analysis at \( \text{OD}_{600} \) values of 1.0 (sample 1), 1.7 (sample 2), and 2.2 (sample 3). Non-induced empty-vector control cells gave equivalent results (data not shown). (B) The length distributions in those samples were measured using DIC optical microscopy as in Fig. 2 in the main text. Mean cell length declines during the standard \textit{E. coli} growth program, dropping from 4.6 \( \mu \text{m} \) at 0.8 \( \text{OD}_{600} \) to 2.5 \( \mu \text{m} \) at 2.5 \( \text{OD}_{600} \).
**Figure S6.** Cellular morphology analyzed by MitoTracker Green fluorescence microscopy. DIC and fluorescence images are shown of the same field of view for each sample. The membrane-permeant lipophilic fluorescent dye MitoTracker Green is commonly used to image intracellular membranes. The brightly stained spheroidal objects near the poles of the NBD-EcMsbA cells correspond to the inclusion bodies that are visible in the DIC images; similar electron-dense objects are observed in the negative-stain EM images of NBD-EcMsbA cells in Fig. 3 in main text. The nature of the intracellular objects stained by Mitotracker Green in some of the other cell populations is unclear. There are no corresponding structures visible in the corresponding DIC images or in the negative-stain EM images of equivalent cells, and their frequency does not correlate with the steady-state level of the overexpressed target protein.
Figure S7. **Negative-stain electron microscopy of whole cells after IPTG induction.** Cells were collected 4 hours after induction, processed and stained with uranyl acetate as described in the Experimental Procedures, and imaged using a Jeol 100 CX transmission electron microscope. The *Ec*GlpT-expressing cells are longer and have more flagella than the matched empty-vector control. Cells expressing the other IMPs showed similar characteristics to the *Ec*GlpT-expressing cells, while cells expressing NBD-*Ec*MsbA or *Ec*Enolase* showed an intermediate phenotype (data not shown).
Figure S8. Correlation plots between pairs of microarray datasets. The Log$_2$ of fold-changes vs. the IPTG-induced empty-vector control are plotted for all MG1655 genes on the Affymetrix 2.0 E. coli array. The top row (panels A-E) shows comparisons of replicate experiments conducted on cells containing the same expression vector (i.e., independent microarray experiments conducted on equivalently prepared cell populations), while the lower rows (panels F-O) show comparisons of experiments conducted on cells containing different expression vectors. The cell populations being compared are indicated by the title at the top of each panel, with the dataset plotted on the ordinate listed first for the cross-sample comparisons. The dotted brown lines in the upper right and lower left quadrants define the regions with genes showing consistent expression changes above the 3-fold linear threshold (i.e., Log$_2$ ≥ 1.58) in the two datasets being compared. The dotted yellow and purple lines define the regions with genes showing inconsistent expression changes in the datasets being compared, as defined by having a change above the 3-fold linear threshold (Log$_2$ ≥ 1.58) in one while showing less than a 1.44-fold linear change (Log$_2$ ≤ 0.5) in the other. Note that there are a minimal number of genes in the inconsistent regions in the experimental replicates in the top row. The middle row compares IMP-overexpressing cells on the ordinate to EcEnolase*-overexpressing cells on the abscissa, while the bottom row compares pairs of IMP-overexpressing cells. There are generally higher correlations between the expression changes in the IMP-overexpressing cells (Fig. 5A in main), especially those expressing the toxic proteins EcGlpT (panel K).
Figure S9. Influence of fliA and glpR-1 mutations on toxicity caused by IMP overexpression. 
OD$_{600}$ was used to monitor growth rate after IPTG induction of protein overexpression in *E. coli* MG1655 cells harboring both a fliA::kan gene disruption and the glpR-1 mutation. Methods were equivalent to those used in Figs. 1A and 7A in the main text. The fliA gene encodes the $\sigma^F$ ($\sigma^{28}$) sigma factor that enhances the expression of genes involved in flagellar biosynthesis. All overexpressing cells show equivalent growth rates either in the presence (shown above) or absence (Fig. 7A in the main text) of the fliA::kan gene disruption in the glpR-1 strain background. However, the cells overexpressing *Ec*GlpT show a consistently higher growth rate in the two strains harboring the glpR-1 mutation compared to wild-type MG1655 (Figs. 1A & 7A in the main text). Therefore, the glpR-1 mutation, which produces constitutive expression of the enzymes involved in glycerol degradation, substantially alleviates the toxicity caused by overexpression of *Ec*GlpT. The strain used in the experiments shown here was constructed using P1 phage to transduce the disrupted fliA::kan gene from the Keio collection into the glpR-1 strain background. Equivalent results were obtained in strain FB20526 from the University of Wisconsin *E. coli* Genome Project, which carries a different disruption of the fliA gene in addition to the same glpR-1 mutation in the MG1655 background (data not shown).
Figure S10. Overexpression properties of different EcMsbA protein constructs in E. coli strains W3110A and WD2S. EcMsbA constructs were expressed at 42° C under control of the arabinose-inducible promoter in the pBAD plasmid. The HisTag-Nter construct carries an N-terminal his-tag. The full-length construct with the native N-terminus (EcMsbA) and the ΔN5 construct have no his-tag. E. coli strain WD2S carries a temperature-sensitive msbA mutation in the W3110A background (12). The Coomassie-Blue-stained SDS-PAGE gel shows samples of OD₆₀₀-normalized total cell extracts from cultures either not induced (ni) or induced (i) with 0.2% (w/v) arabinose to drive expression of the indicated EcMsbA construct from the pBAD plasmid. The arrows indicate the migration positions of the expressed protein constructs. The three non-induced/induced pairs on the left come from E. coli strain W3110A, while the three induced samples on the right come from strain WD2S. Note that the ΔN5 construct is expressed at a substantially reduced level compared to the construct with the native N-terminus. Both the ΔN5 and wild type msbA constructs are expressed at roughly equivalent levels in the two different strains, while the N-terminally tagged construct is produced at a substantially higher level in the temperature-sensitive msbA mutant strain WD2S compared to the parental strain W3110A.