



Systematic Localization of *Escherichia coli* Membrane Proteins

Anna Sueki,^{a,b} Frank Stein,^a Mikhail M. Savitski,^a Joel Selkrig,^a Athanasios Typas^a

^aEuropean Molecular Biology Laboratory (EMBL), Genome Biology Unit, Heidelberg, Germany

^bHeidelberg University, Faculty of Biosciences, Heidelberg, Germany

ABSTRACT The molecular architecture and function of the Gram-negative bacterial cell envelope are dictated by protein composition and localization. Proteins that localize to the inner membranes (IM) and outer membranes (OM) of Gram-negative bacteria play critical and distinct roles in cellular physiology; however, approaches to systematically interrogate their distribution across both membranes and the soluble cell fraction are lacking. Here, we employed multiplexed quantitative mass spectrometry using tandem mass tag (TMT) labeling to assess membrane protein localization in a proteome-wide fashion by separating IM and OM vesicles from exponentially growing *Escherichia coli* K-12 cells on a sucrose density gradient. The migration patterns for >1,600 proteins were classified in an unbiased manner, accurately recapitulating decades of knowledge in membrane protein localization in *E. coli*. For 559 proteins that are currently annotated as peripherally associated with the IM (G. Orfanoudaki and A. Economou, *Mol Cell Proteomics* 13:3674–3687, 2014, <https://doi.org/10.1074/mcp.O114.041137>) and that display potential for dual localization to either the IM or cytoplasm, we could allocate 110 proteins to the IM and 206 proteins to the soluble cell fraction based on their fractionation patterns. In addition, we uncovered 63 cases, in which our data disagreed with current localization annotation in protein databases. For 42 of these cases, we were able to find supportive evidence for our localization findings in the literature. We anticipate that our systems-level analysis of the *E. coli* membrane proteome will serve as a useful reference data set to query membrane protein localization, as well as to provide a novel methodology to rapidly and systematically map membrane protein localization in more poorly characterized Gram-negative species.

IMPORTANCE Current knowledge of protein localization, particularly outer membrane proteins, is highly dependent on bioinformatic predictions. To date, no systematic experimental studies have directly compared protein localization spanning the inner and outer membranes of *E. coli*. By combining sucrose density gradient fractionation of inner membrane (IM) and outer membrane (OM) proteins with multiplex quantitative proteomics, we systematically quantified localization patterns for >1,600 proteins, providing high-confidence localization annotations for 1,368 proteins. Of these proteins, we resolve the predominant localization of 316 proteins that currently have dual annotation (cytoplasmic and IM) in protein databases and identify new annotations for 42 additional proteins. Overall, we present a novel quantitative methodology to systematically map membrane proteins in Gram-negative bacteria and use it to unravel the biological complexity of the membrane proteome architecture in *E. coli*.

The inner membranes (IM) and outer membranes (OM) of Gram-negative bacteria carry out fundamental cellular functions crucial for cell viability (1). The OM directly interfaces with the extracellular environment and provides a formidable physical barrier that excludes the passage of large and hydrophobic compounds (2–4). The IM plays a vital role in ensuring selective transport of small compounds into and out of the cell,

Citation Sueki A, Stein F, Savitski MM, Selkrig J, Typas A. 2020. Systematic localization of *Escherichia coli* membrane proteins. *mSystems* 5:e00808-19. <https://doi.org/10.1128/mSystems.00808-19>.

Editor Caroline M. Ajo-Franklin, Lawrence Berkeley National Laboratory

This article followed an open peer review process. The review history can be read [here](#).

Copyright © 2020 Sueki et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](#).

Address correspondence to Joel Selkrig, selkrig@embl.de, or Athanasios Typas, typas@embl.de.

Received 27 November 2019

Accepted 14 February 2020

Published 3 March 2020

producing and conserving energy, as well as sensing external cues and transducing information to adaptive transcriptional responses (5–7). Both membranes facilitate protein translocation from the cytoplasm to the cell envelope and/or the extracellular milieu by dedicated protein machines. Proteins targeted to the IM and OM possess distinct biochemical properties and play fundamental roles in building and maintaining cell envelope integrity. This includes correct assembly of the peptidoglycan layer, which gives bacteria their cell shape and together with the OM defines their mechanical strength (8, 9). Understanding to which membrane proteins are targeted provides important insight on the physical location of their biological activity, which is particularly useful for interrogating proteins of unknown function and investigating modular envelope protein complexes (10).

Proteins localizing to the IM or OM can be subcategorized on the basis of their distinct biophysical properties. IM proteins typically contain α -helical transmembrane domains that mediate their integration into the phospholipid bilayer. These proteins include small-molecule transporters (e.g., ABC transporters for metal ions and sugars) and large multiprotein complexes (e.g., SecYEG translocon and ATP synthase). IM proteins are diverse in their structure, function, and domain localization; some contain structural domains that extend into the cytoplasmic and/or periplasmic space. In the OM, there are two distinct types of proteins: outer membrane proteins (OMPs), which are mostly composed of amphipathic β -strands that form a closed β -barrel structure, and lipoproteins, which are anchored to the OM via an N-terminal lipidated cysteine and typically contain a soluble domain that extends into the periplasmic space. Note that only a small fraction of lipoproteins is retained in the IM, and lipoproteins in both the IM and OM carry out diverse functions in the cell envelope, yet they still remain a largely poorly characterized group of proteins. Although lipoproteins have been traditionally considered to face the periplasmic space, some have been recently shown to reach the cell surface (11–13), even by traversing through OMPs (14–16).

While most proteins localize to, and function at, either the IM or OM, several transenvelope protein complexes span both membranes. Some of the larger complexes like flagella (17) and secretion apparatuses (18) contain dedicated IM, OM, and periplasmic components, whereas smaller ones such as the Tol-Pal complex (19–21), the PBP1a/b-LpoA/B peptidoglycan synthase complexes (22–24), the efflux pump AcrAB-TolC (25), the translocation and assembly module (TAM) (26), and all TonB-dependent transport complexes (27) possess IM or OM components that can span the envelope by reaching their interaction partner in the other membrane. Moreover, in addition to integral membrane proteins, soluble proteins can associate peripherally with membranes via stable or transient protein-protein interactions with integral membrane proteins and/or with the lipid bilayer. In the case of lipoproteins, the attachment to phospholipids is covalent and part of their biosynthesis (28). As many membrane proteins reside within protein complexes that play vital roles for cell envelope integrity, it is important to obtain a blueprint of the bacterial membrane protein composition that experimentally defines which proteins are membrane associated and the membrane to which they are targeted. Furthermore, protein localization is closely linked to protein function, and therefore a key step in ascertaining the mode of action of membrane proteins.

Localization for most proteins in *Escherichia coli* and other Gram-negative bacteria can be predicted quite accurately based on sequence information. Analysis of protein N-terminal signal sequences is commonly used to predict protein localization. For example, SignalP (29) detects signal sequences that target nascent proteins for transport across the IM via the SecYEG/SecA translocon. Similarly, TATFIND identifies substrates of the Tat translocase (30). Interestingly, periplasmic proteins bear sufficiently distinguishable biophysical, biochemical, and structural characteristics compared to their cytoplasmic counterparts, so that one can accurately discern them with signal-sequencing agnostic machine-learning predictors (31). Moreover, several tools exist to predict membrane protein localization and topology, including those that predict transmembrane α -helices and topology of IM-transmembrane proteins such as the

TMHMM server (32), as well as algorithms to predict β -barrel folding of OMPs such as PRED-TMBB (33) and BOMP (34). *E. coli* genome databases, such as STEPdb (35), EcoCyc (36), and UniProt (37), use information from such prediction tools, together with experimental evidence, to assign protein localization. However, the difficulties of assigning protein localization solely based on structure and signal sequence prediction can lead to misannotations, as has been discussed previously (35, 38). Thus, as much of the current knowledge is based on prediction tools, it is of key importance to experimentally verify protein localization to clarify the predictive accuracy of the above-mentioned *in silico* approaches.

Proteomic-based studies of bacterial cell membranes in the past have focused on either the IM or OM proteomes separately (39–42), which precludes definitive statements about the protein allocation across the two membranes and is prone to contamination from abundant soluble proteins. To systematically examine membrane protein localization in an unbiased and systematic manner, we combined sucrose gradient membrane fractionation with quantitative proteomics (43). This allowed us to validate most of the predicted protein localizations, while resolving the protein localization of a large number of proteins with previously ambiguous localization, and uncovering some proteins with unexpected membrane localization, which are currently misannotated in databases.

RESULTS

IM and OM proteome separation and quantification. To systematically assess protein localization within the bacterial cell envelope, we isolated bacterial membrane proteins by harvesting *E. coli* K-12 MG1655 grown to mid-exponential phase (Fig. 1A). Total membrane vesicles were purified, followed by IM and OM vesicle separation on a sucrose density gradient as previously described (44). The sucrose gradient was then separated into 11 fractions and analyzed by immunoblotting and sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE). Effective IM and OM vesicle separation was verified using SecG and BamA antiserum (Fig. 1B). Fractions 2 to 11 (f02 to f11) and the total membrane (input sample prior to fractionation) were labeled with 11-plex tandem mass tag (TMT) reagents (45) and analyzed and quantified using mass spectrometry (MS) (Fig. 1A; see Fig. S1A in the supplemental material). Multiplexing of all 11 fractions into a single MS run is what enables robust comparisons between samples and increases the quantitative power of the fractionation patterns. In total, we identified 1,605 common proteins across the two biological replicates and thus proceeded to data normalization and quantification as described in Materials and Methods (see Fig. S1A and Table S1 in the supplemental material).

To assess protein abundance in each fraction, we calculated the logarithmic fold change (\log_2 FC) of each sucrose gradient fraction relative to the total membrane fraction for each protein. The fractionation pattern across the 10 quantified fractions (fractions 2 to 11) was examined for each protein for each replicate (Table S1). The quantitative MS-based fractionation pattern matched immunoblot data for the control proteins, SecG for the IM and BamA for the OM (Fig. 1B, bottom). Replicate correlation between the two independent experiments for all \log_2 FC values was high ($R = 0.77$; Pearson correlation, $P < 0.001$; Fig. S1B). Proteome coverage was analyzed by comparing to protein localization annotations found in the curated STEPdb database (35) using UniProt identifiers (IDs) (summarized and modified as in Table S2). For membrane protein annotation categories, we had an overall 56% coverage (973 out of the total 1,741 membrane proteins annotated in the STEPdb database for *E. coli*) with several categories reaching 70% coverage, whereas nonmembrane protein categories did not exceed 25% coverage (Fig. 1C). Undetected membrane proteins may be due to lack of expression in our conditions (mid-log, LB, 37°C), generally low protein abundance, small protein size, or MS detection limitations (e.g., due to the peptide ionization properties or size after tryptic digestion). Altogether we could quantitatively assess the sucrose gradient fractionation of 1,605 proteins (Table S3), covering most membrane-annotated proteins in STEPdb.

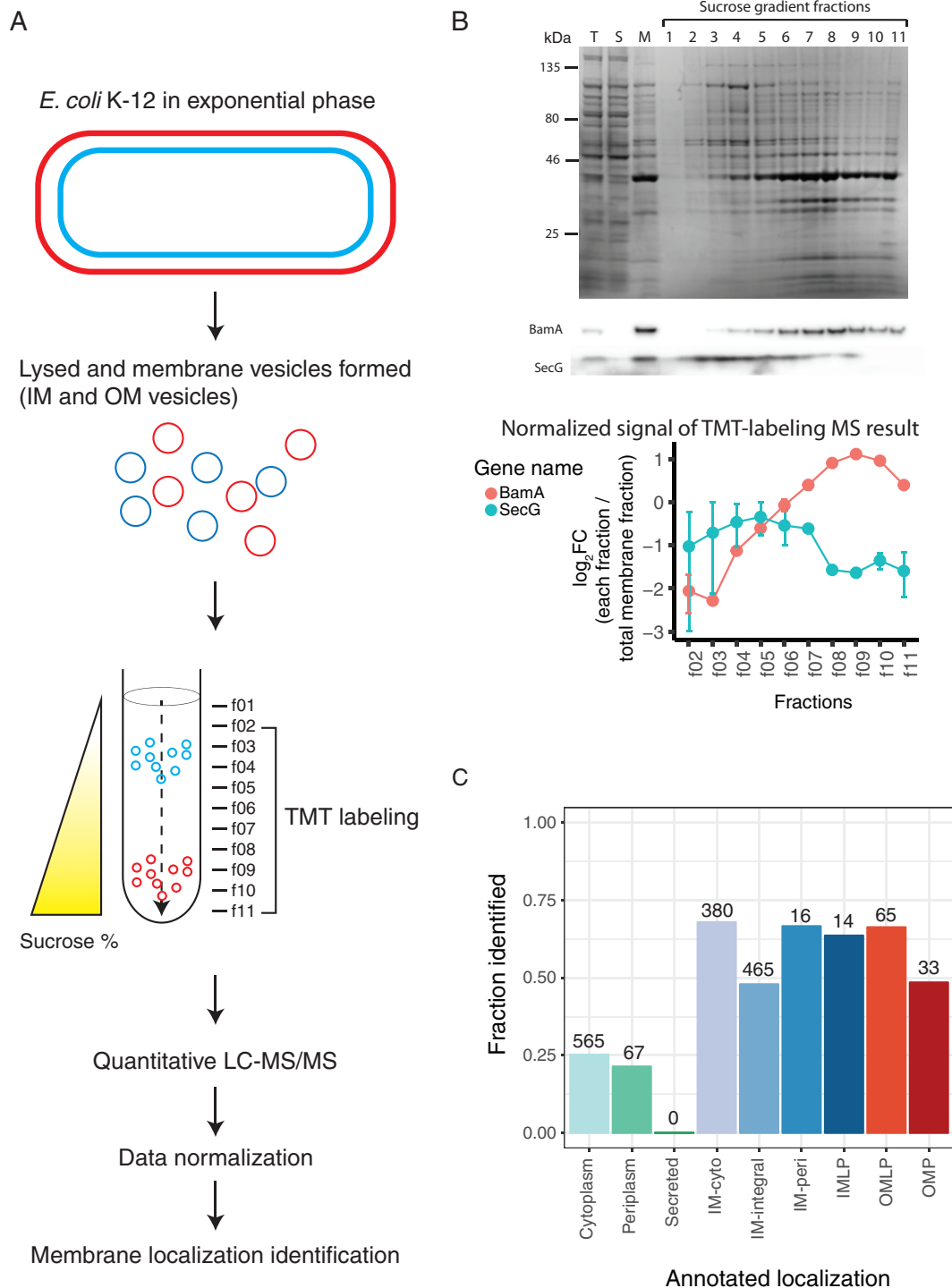


FIG 1 Gram-negative bacterial inner and outer membrane fractionation quality control and membrane proteome coverage. (A) Schematic illustration of the method. *E. coli* cells were harvested in exponential phase (OD₅₇₈ of ~0.7), lysed, and ultracentrifuged to collect the membrane fraction containing both outer membrane (OM) (red) and inner membrane (IM) (blue) vesicles. Total membrane vesicles were separated on a sucrose density gradient, which separates IM from OM vesicles based on their distinct buoyant densities. Samples were collected into 11 fractions (f01 to f11) where f02 to f11 and total membrane sample prior to fractionation were TMT labeled and analyzed by LC-MS/MS. After data normalization (see Fig. S1 in the supplemental material), membrane localization of proteins was defined. (B) SDS-PAGE analysis of sucrose gradient fractionation. T, total cell lysate; S, soluble fraction upon ultracentrifugation; M, total membrane fraction prior to sucrose gradient fractionation. (Top) Coomassie blue-stained gel. (Middle) Immunoblot analysis for control proteins: BamA for OM, SecG for IM. (Bottom) TMT-labeling MS quantification result of the two control proteins (BamA and SecG). The graph shows the log₂ fold change of each fraction/total membrane fraction for fractions (f02 to f11) for BamA and SecG. Mean with standard deviation (error bars) values are plotted from the two biological replicates. (C) Fraction of proteins identified for each STEPdb localization category. Localization annotation is summarized in Table S2.

Systematic assignment of membrane protein localization. Sucrose density gradients are conventionally analyzed by immunoblotting to compare the abundance of a given protein within a high- or low-sucrose-density fraction (Fig. 1A and B). To systematically analyze protein localization, we used the combined averages of the high- and low-sucrose fractions (\log_2 of f08, f09, and f10 for the high-sucrose fraction and \log_2 of f02, f03, and f04 for the low-sucrose fraction) and calculated the difference between the two \log_2 averages, which we referred to as the “sucrose gradient ratio” (Table S3). High values indicate a greater abundance within higher-sucrose-density fractions, as expected for OM proteins. The reverse is true for IM proteins, which exhibit low values due to their enrichment within the low-sucrose-density fractions.

To assess whether our calculated sucrose gradient ratio reflected known protein localization, we grouped these values based on localization annotation (modified from those in the STEPdb database [Table S2]). As anticipated, most IM protein categories (i.e., integral IM proteins [IM-integral], periplasmic proteins peripherally associated with IM [IM-peri], and IM lipoprotein [IMLP]) displayed a low sucrose gradient ratio, whereas the two OM protein categories, OMPs and OM lipoproteins (OMLPs), showed high sucrose gradient ratios (Fig. 2A). This striking concordance with curated annotations and our calculated localization confirms the accuracy of our methodology. We chose the 90th percentile of IM protein distribution (solid blue line) and the 10th percentile of OM protein distribution (solid red line) as cutoffs to define IM and OM protein localization using the sucrose gradient ratio (Fig. 2B), respectively. All proteins that fell between these two cutoffs were classified as soluble proteins.

Interestingly, although soluble proteins were expected contaminants in our experiments, they did not always behave as expected upon sucrose density gradient fractionation. In general, soluble proteins localized in the cytoplasm and periplasm displayed midrange sucrose gradient ratios, which is in agreement with the majority of them being contaminants and nonspecifically associated with either IM or OM vesicles. We noted that the IM-cyto category displayed bimodal characteristics with one peak being consistent with IM localization and another peak that aligned with soluble proteins (Fig. 2A and Table S3). This suggests that the IM-cyto category of proteins referred to as “peripheral IM proteins” in the STEPdb database and originally described in another study (41) consists of a mixture of proteins that have clear preferential localizations either to the cytoplasm or to the IM. We therefore did not use this category for benchmarking our data but rather kept it to later definitively allocate the primary localization of this large group of proteins. Taken together, these data show that quantitative proteomic-based analysis of sucrose gradient-fractionated membrane vesicles can rapidly and systematically localize proteins to the IM or OM.

Unbiased clustering of sucrose gradient fractionation patterns can robustly identify protein membrane localization. As the sucrose density-based fractionation patterns of IM and OM proteins were distinct (Fig. 1B, graph after normalization), we asked whether unbiased clustering of the fractionation patterns could be used to distinguish proteins according to their annotated membrane localization. On the basis of the results of principal-component analysis (PCA) of all identified proteins, we set the number of clusters to 4 for k-means clustering analysis (Fig. S2). This resulted in four groups containing the following number of proteins: cluster 1 with 111 proteins, cluster 2 with 650 proteins, cluster 3 with 680 proteins, and cluster 4 with 164 proteins (Fig. 3A and Table S3), where each cluster exhibited distinctive fractionation patterns (Fig. 3B). We then asked what STEPdb localization annotations are associated with proteins in the four different clusters. Proteins in cluster 1 were strongly enriched for OM proteins, cluster 2 with IM proteins, and cluster 3 and 4 with soluble proteins (Fig. 3C). Thus, k-means clustering successfully grouped proteins based on their membrane localization as a function of their sucrose gradient fractionation pattern. While each cluster was dominated by a single membrane localization annotation (OM for cluster 1, IM for cluster 2, or soluble for clusters 3 and 4), some proteins were grouped into clusters that conflicted with their STEPdb annotated localization, with the IM-cyto group featuring

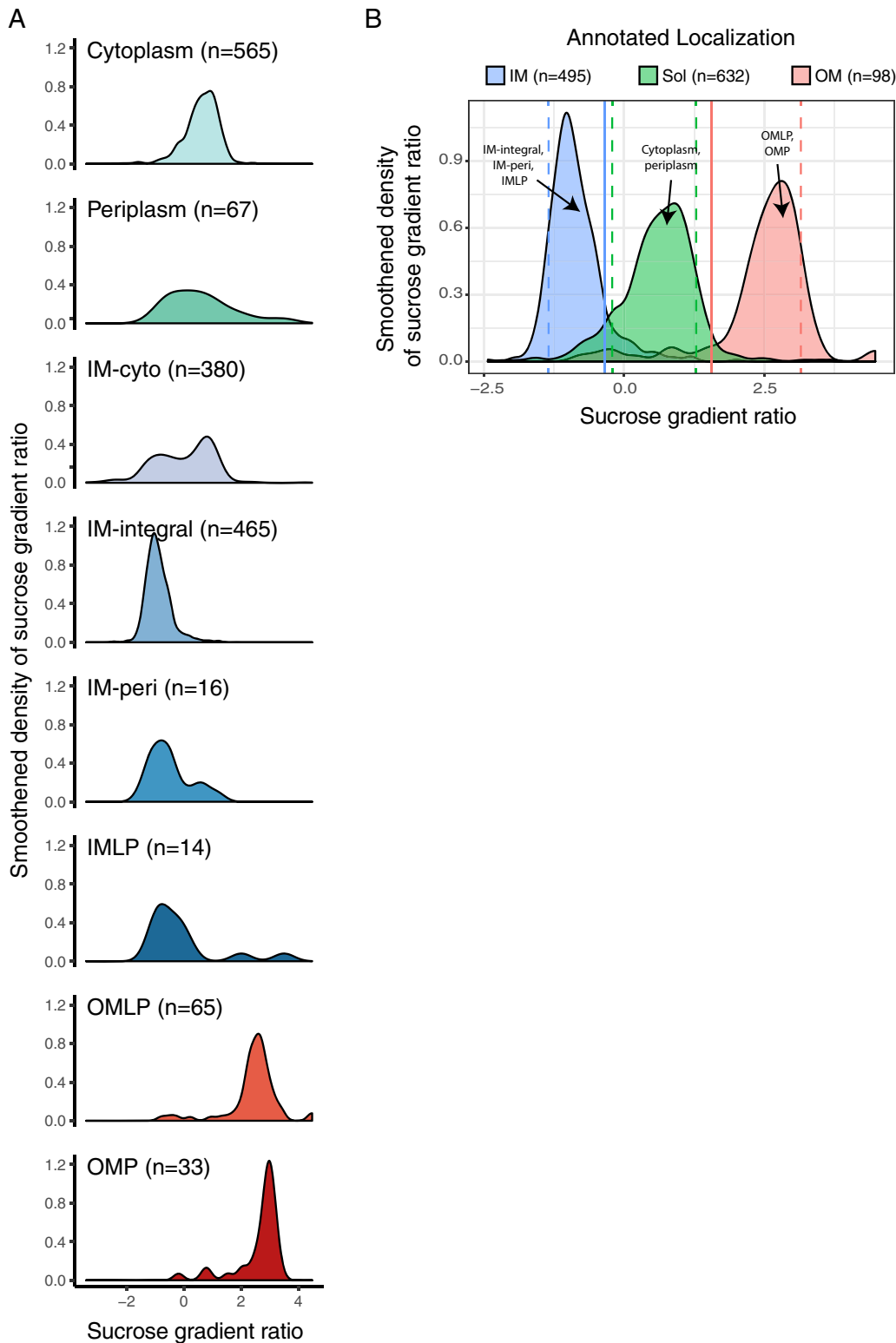


FIG 2 Sucrose gradient ratio separates proteins according to membrane localization. (A) Smoothed distributions of sucrose gradient ratios for each protein localization category in the STEPdb database (as in Fig. 1C). Sucrose gradient ratio values were calculated as a difference of the average of high-sucrose fractions (\log_2 of f08, f09, and f10) and the average of the low-sucrose fractions (\log_2 of f02, f03, and f04) for each protein. (B) Sucrose gradient ratio of three categories (IM, OM, and soluble [Sol]) grouped from panel A. Cytoplasmic and periplasmic proteins from panel A were grouped as soluble proteins ($n = 632$), IM-integral and IM-peri as IM proteins ($n = 495$), and OM and OMLPs as OM proteins ($n = 98$). The dashed lines refer to the 10th and 90th percentiles for IM, OM, and soluble proteins for each membrane localization group. The 90th percentile for IM proteins and the 10th percentile for OM proteins are shown as solid lines, as they are used as thresholds to allocate proteins to the three categories.

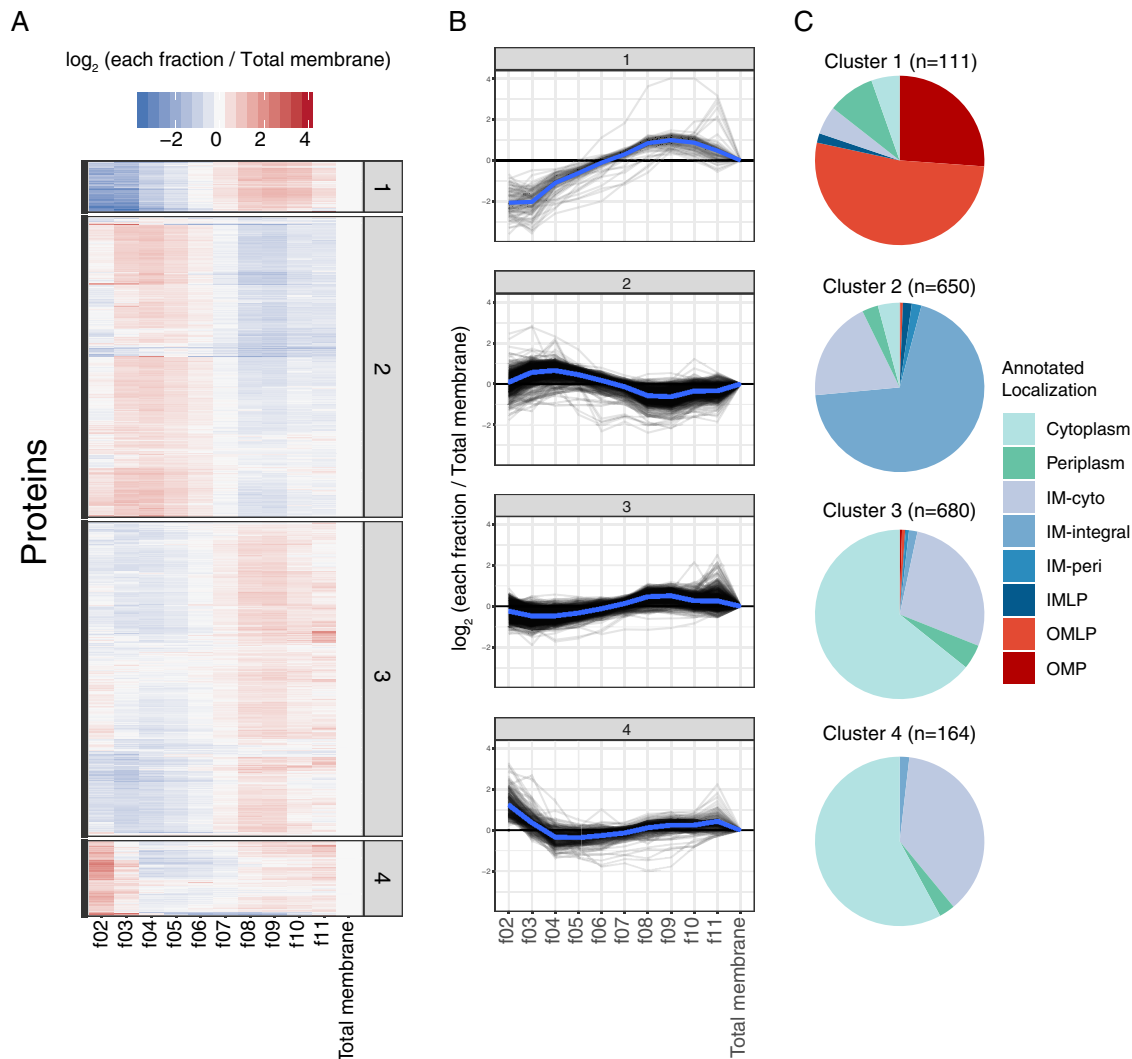


FIG 3 Unbiased k-means clustering of sucrose gradient fractionation patterns accurately depicts protein membrane localization. The k-means clustering was based on \log_2 fold change of each fraction to total membrane sample. (A) Heatmap representing cluster patterns. The cluster number is shown to the right of the heatmap. (B) Fractionation pattern of all proteins in each cluster (gray) and the distribution average (blue). (C) Pie chart representing annotated localization of proteins found in each cluster. The localization annotation is as shown in Fig. 1C.

prominently in various clusters (Fig. 3C). Nevertheless, these results show that systematic and unbiased clustering of sucrose gradient fractionation patterns of the membrane proteome can be used to assign membrane localization.

Filtering and comparison of sucrose gradient ratio and clustering. In order to increase the confidence of our calls for protein localization, we performed two further steps. First, we ran k-means clustering for the data set where the two biological replicates were treated separately. By doing this, we identified 140 (out of 1,605) proteins whose fractionation patterns between biological replicates resulted in clustering to different localization clusters (OM for cluster 1, IM for cluster 2, or soluble for clusters 3 and 4) for the two replicates (Table S3). We reasoned that this is due to irreproducibility between the two biological replicates and removed these proteins from further analysis. Second, we assessed the similarity of our two methods (k-means versus sucrose gradient ratio) in assessing protein localization. To do this, we used the thresholds of sucrose gradient ratio for IM and OM defined in Fig. 2B. We found a large overlap between the two methods for all three localization categories: IM, OM, and soluble (Fig. 4A). In total, we identified 1,368 proteins (out of 1,465 possible) with

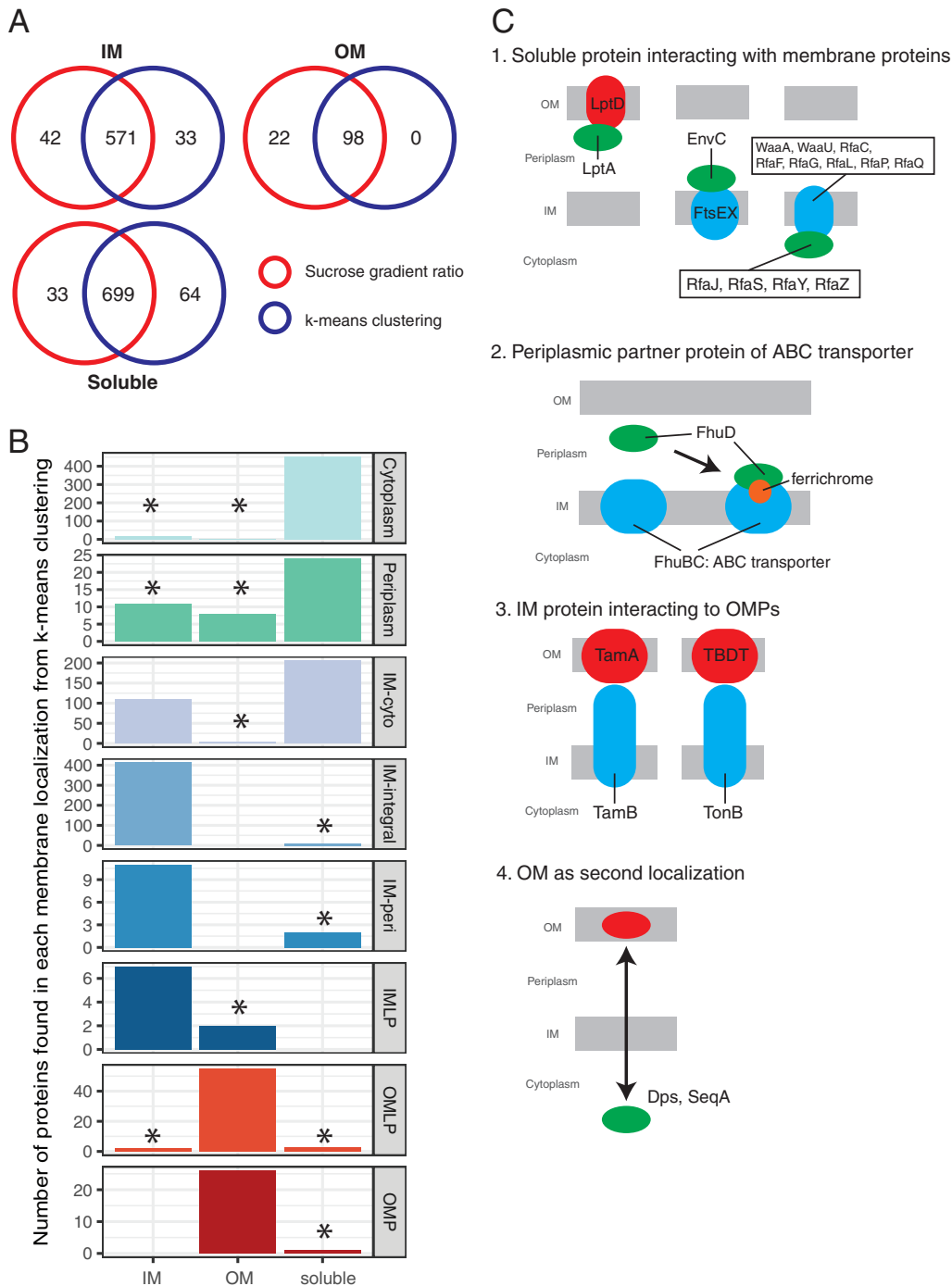


FIG 4 Identification of potentially misannotated proteins. (A) Venn diagrams showing commonly identified proteins for each localization categories (IM, OM, and soluble), based on our two quantification methods. Protein localizations stemming from the sucrose gradient ratios are shown in red, and protein localizations from k-means clustering annotation are shown in blue. A total of 1,465 proteins were assigned localization by the different methods. The two methods agreed in the localization of 1,368 proteins, which we used as a “high-confidence” localization set (Table S3). (B) Comparison of identified localization with annotated localization for the 1,368 proteins from panel A. Protein localizations considered not to match with STEPdb annotations are marked with an asterisk. The IM-cyto category has a bimodal distribution, likely because proteins could not be confidently allocated to either the IM or cytoplasm in the STEPdb database but can be with the experimental data provided here. (C) Schematic representation explaining experimental localization of proteins not matching or partially matching the STEPdb annotation (besides IM-cyto). (Scenario 1) Soluble proteins interacting with membrane proteins. (Left) LptA interacts with OM protein, LptD/E and fractionates as OM protein as previously reported (63). (Middle) EnvC interacts with IM proteins FtsEX (64), similarly for FdoG, NapG, and RseB. (Right) Rfa proteins in cytoplasm were identified as IM proteins, possibly due to interactions with bona-fide IM Rfa proteins. (Scenario 2) Periplasmic partners of ABC transporters were detected as IM proteins presumably because they bind to the IM components and transporters when transporters are active (47). This is for example the case for FhuD interacting with

(Continued on next page)

localization results that agreed between the two methods, which we further used as our high-confidence protein localization data set. When considering all 1,605 proteins, the two methods agreed for 1,456 proteins (Fig. S4A). In general, both quantification methods for protein localization worked well, and in combination provided more confident identification calls (true positive rates are 95% and 48% for overlap and nonoverlap sets, respectively; using STEPdb annotations as true positive). The clustering method worked better than the ratio cutoffs for OM proteins, but on the other hand, cluster 4 seemed to have the most inconsistent calls for protein localization (Fig. S3).

As mentioned before, IM-cyto proteins showed a bimodal sucrose gradient ratio distribution (Fig. 2A) and could be found in multiple clusters (clusters 2, 3, and 4 [Fig. 3C]). Now that we have high-confidence protein localization allocations, we could define 110 proteins as IM and 206 as soluble (Fig. 4B; see also Fig. S4B for all proteins, including lower-confidence ones). This separation is corroborated by the melting temperatures of these two group of proteins. We have previously reported that IM proteins are more thermostable than their cytoplasmic counterparts (46). In agreement with this, IM-cyto proteins categorized here to be soluble had low melting temperatures similar to those of cytoplasmic proteins (Fig. S5). In contrast, IM-cyto proteins categorized as IM proteins had higher melting temperatures, albeit not as high as integral IM proteins, presumably due to their peripheral interaction rather than integral association with the IM (Fig. S5). Thus, with the experimental conditions we tested, IM-cyto annotated proteins resulted in a mixture of soluble and IM proteins, and by using their fractionation patterns, we could allocate their predominant protein localization.

Identification of potentially misannotated proteins. The STEPdb database combines robust computational predictions with a wealth of experimental information to allocate protein localization in *E. coli*, and hence, we used it here as a gold standard data set to benchmark our data and decide on thresholds for making localization calls. In doing so, we noted that a small fraction of our allocations of IM, soluble, and OM proteins conflicted with their STEPdb annotations. Namely, 63 (out of 1,368) high-confidence proteins, including both membrane and soluble proteins, were found in clusters that at least partially conflicted with their corresponding STEPdb localization annotation (Table S4). We manually curated these proteins based on published literature. First, we checked the localization annotation in another recent study (38). Babu et al. (38) summarized different localization annotation databases and primary literature, including STEPdb, generated a score for protein localization, and then annotated protein localization accordingly in their study. Eleven proteins out of the 63 proteins we identified as mismatches agreed with the curated list of Babu et al. (38). Therefore, we conclude that the STEPdb localization for these 11 proteins was likely inaccurate (Table S4). We found corroborating evidence for 12 more such cases in the literature or in other prediction databases. Importantly, in these cases, rather than relying on the combined result from multiple *in silico* prediction algorithms, our data are able to provide the high-confidence experimental evidence needed to verify the localization for these proteins.

We also noted unexpected fractionation patterns for certain proteins upon sucrose density fractionation. First, we detected several proteins annotated as solely periplasmic in the STEPdb database fractionated as membrane proteins in our experiments. Many of these proteins have known interacting membrane partners, which is presumably the reason they cofractionate with either the IM (EnvC, FdoG, NapG, and RseB) or

FIG 4 Legend (Continued)

FhuBC upon substrate (ferrichrome [orange]) binding. (Scenario 3) Transenvelope protein interaction is known. TamB spans the envelope and interacts with OM-located TamA (26) and TonB interacts with OM-located TonB-dependent transporters, resulting in a misleading sucrose gradient fractionation (distributed equally in all fractions, thus appearing as a soluble protein), which in the case of TonB has been previously shown (48, 49). (Scenario 4) Previously shown dual localization of cytoplasmic protein in OM for Dps and SeqA (65–67). A complete list of 63 proteins not matching with the STEPdb annotation, with 51 proteins explained can be found in Table S4.

the OM (LptA) (Fig. 4C and Table S4). The situation was similar for periplasmic components of IM ABC transporter complexes (FhuD, PstS, and SapA) which cofractionated with IM proteins, possibly as a consequence of a direct conditional association with IM proteins upon active transport (47), but were only annotated as periplasmic in STEPdb (Table S4). In total, there were 19 cases for which STEPdb had incomplete annotations.

Conversely, FecB, a known periplasmic component of an ABC transporter complex, is annotated both as IM-peri (peripherally associated with IM) and periplasmic in the STEPdb database but was identified only as soluble under our experimental conditions. In this case, we are failing to detect the IM association because the transporter is likely inactive under the conditions we used and the STEPdb annotation is more accurate (Table S4). Moreover, IM proteins known to form transenvelope complexes (e.g., TamB and TonB) failed to cluster as either IM or OM proteins in the fractionation experiments. Overall, we could reasonably explain 51 out of the 63 cases where STEPdb and our results disagreed, out of which we could find additional information that supports our localization call (42 proteins) or the original STEPdb annotation (9 proteins) (summarized in Table S4). Overall, these findings demonstrate that our quantitative assessment of protein localization captures accurately the *in vivo* biological state.

DISCUSSION

We quantified the membrane proteome using TMT-labeling MS, which allowed us to experimentally identify localization in a systematic and unbiased manner for the majority of membrane proteins in *E. coli*. We verified current knowledge of membrane protein localization for proteins that was determined experimentally and/or predicted bioinformatically. The advantage of this method is that instead of assessing membrane protein localization via conventional immunoblotting of sucrose density gradient fractions, quantitative proteomic approaches can be used to rapidly and quantitatively assess protein localization in an antibody-independent manner.

Comparison of our data with the curated STEPdb annotation revealed high concordance. In addition, our data provided a predominant location for a large part of the *E. coli* membrane proteome referred to as peripherally associated membrane proteins (41). The STEPdb database categorizes proteins that peripherally interact with the cytoplasmic face of the IM as “peripheral IM proteins,” which we referred to here for simplicity as IM-cyto (see Table S2 in the supplemental material). Although we detect the majority of IM-cyto proteins (68% of 559 proteins) in the membrane fraction (which is depleted of soluble proteins), most of them are reproducibly assigned as soluble proteins according to their sucrose fractionation pattern (206 out of 316 high-confidence calls). This absence of cofractionation with the IM proteome suggests that many of these proteins are mainly cytoplasmic in exponentially growing cells in LB, and their previous identification in membrane protein fractions in this study and others is likely because they are recurrent contaminants. We cannot exclude the possibility that some of these proteins have a conditional, low-affinity, or transient association with the IM and proteins therein or that a small fraction of the total protein amount may, at any given point, be associated with the IM. In contrast, about one third of the IM-cyto proteins exhibited clear IM fractionation patterns and thus can be confidently assigned as IM-associated proteins.

We found 63 proteins out of 1,368 which were inconsistent with the reported localization annotation in the STEPdb database. We were able to explain 51 by additional literature data. Those proteins have a wrong or missing annotation in STEPdb (42) or their function/activity makes their sucrose gradient fractionation patterns misleading (8). In most cases, sucrose gradient fractionation failed to make the right call when the protein was spanning the envelope or presumably had dual membrane localization. It is likely that the new localization is also correct for most of 12 remaining proteins (Table S4). Thus, our data are helpful for improving protein localization, even for an organism as intensively studied as *E. coli*, which has been subjected

to a plethora of targeted and systematic studies, and researchers can benefit from carefully curated databases, such as STEPdb.

In some cases, sucrose gradient fractionation patterns were indicative of protein activity. For example, periplasmic partners of IM ABC transporters often showed sucrose gradient fractionation patterns similar to those of IM proteins (Fig. 4C), which we postulate were due to their strong interaction with the cognate IM ABC transporters in their substrate-bound state (47). In other cases, they behaved as soluble proteins, which likely reflects the inactive state of the ABC transporter. Interestingly, transenvelope spanning IM proteins, such as TamB and TonB, displayed fractionation patterns similar to those of soluble proteins. This is presumably due to strong interactions with their OM counterparts, which pull a subpopulation of the protein together with OM vesicles during ultracentrifugation. Consistent with our observations, TonB was previously found in the OM fraction upon sucrose gradient fractionation (48, 49). This suggests that transenvelope IM proteins can present distinctive properties upon sucrose gradient fractionation. More broadly, it implies that sucrose gradient fractionation can provide insights into the activity and mechanical strength of specific envelope complexes during different growth stages and conditions.

Not only does this work provide a resource of sucrose gradient fractionation for 1,605 proteins, with 1,368 proteins having their cellular localization confidently assigned, the method we present can be used for rapid and systematic characterization of membrane proteomes in different contexts—growth conditions and stages, and under different cellular perturbations. Many membrane proteins are only conditionally expressed (46), whereas other proteins conditionally relocate in and out of membranes (50–52). Importantly, our method allows for the systematic mapping of membrane proteomes from other Gram-negative species for which protein localization annotation and transport mechanisms are less (if at all) studied.

MATERIALS AND METHODS

Bacterial culturing. The wild-type strain used in this study is *Escherichia coli* K-12 MG1655 $\Delta(\text{argF-lac})\text{U169 rprA::lacZ}$ (53). Bacterial cells were grown for 4 generations in LB-Lennox (referred as LB herein) medium at 37°C with vigorous shaking at 200 rpm and collected for fractionation while still in the exponential growth phase at an optical density at 578 nm (OD_{578}) of 0.6 to 0.8.

Membrane vesicle isolation and sucrose density fractionation. Membrane vesicles were isolated and fractionated essentially as previously described (44) with the following deviations. Phosphate-buffered saline (PBS) was used as the base buffer instead of Tris. After sucrose gradient separation, 1-ml fractions were collected stepwise from the top of the gradient, yielding 11 fractionated samples that were analyzed by Coomassie blue staining and Western blotting using sodium dodecyl sulfate (SDS)-polyacrylamide gels as described below. Fractions 2 to 11 (f02 to f11), as well as an aliquot of the total input membrane sample (diluted 10 times in H_2O), were labeled with 11-plex tandem mass tag (TMT) and subjected to liquid chromatography coupled to tandem mass spectrometry (LC-MS/MS).

Sample preparation and TMT labeling. Before sample preparation for MS, proteins were solubilized by adding SDS to the samples (final concentration of 1% SDS). Samples were then sonicated for 5 min in an ultrasonic bath, heated for 10 min to 80°C, and sonicated again for another 5 min. Disulfide bonds were reduced by incubating at 56°C for 30 min in 10 mM dithiothreitol (DTT) buffered with 50 mM HEPES (pH 8.5). Reduced cysteines were alkylated by incubating for 30 min at room temperature in the dark with 20 mM 2-chloroacetamide in 50 mM HEPES buffer (pH 8.5). Samples were prepared for MS using the SP3 protocol (54, 55). On bead trypsin (sequencing grade, Promega, V5111) digestion was performed to an enzyme/protein ratio of 1:50 and incubated overnight at 37°C. Digested peptides were then recovered in HEPES buffer by collecting the supernatant after magnet-based separation from the SP3 bead, and combining the second elution wash of beads with HEPES buffer. Collected peptides were labeled with TMT10plex isobaric label reagent (ThermoFisher) (45) and with ^{13}C label (ThermoFisher) according to the manufacturer's instructions as described below. In brief, 0.8 mg of the TMT reagents was dissolved in 42 μl of 100% acetonitrile, and 4 μl of this stock was added to the peptide sample and incubated for 1 h at room temperature. The reaction was quenched with 5% hydroxylamine for 15 min at room temperature. Then the 10 samples labeled with unique TMT10plex labels were combined into one sample. The combined sample was then cleaned up using OASIS HLB $\mu\text{Elution}$ plater (Waters). The samples were separated through an offline high-pH reverse-phase fractionation on an Agilent 1200 Infinity high-performance liquid chromatography system which was equipped with a Gemini C_{18} column (3 μm ; 110 Å; 100 by 1.0 mm; Phenomenex). The fractionation was performed as previously described (56). Samples were pooled into a total of 12 fractions.

Mass spectrometry data acquisition. Chromatography was performed using an UltiMate 3000 RSLC nano LC system (Dionex) fitted with a trapping cartridge ($\mu\text{-Pre}^2$ column C_{18} PepMap 100; 5 μm ; 300- μm inner diameter [i.d.] by 5 mm; 100 Å) and an analytical column (nanoEase M/Z HSS T3 column 75 μm by

250 mm C_{18} ; 1.8 μm ; 100 \AA ; Waters). Trapping was carried out with a constant flow of solvent A (0.1% formic acid in water) at 30 $\mu\text{l}/\text{min}$ onto the trapping column for 6 min. Subsequently, peptides were eluted via the analytical column with a constant flow of 0.3 $\mu\text{l}/\text{min}$ with an increasing percentage of solvent B (0.1% formic acid in acetonitrile) from 2% to 4% in 6 min, from 4% to 8% in 1 min, then 8% to 25% for a further 71 min, and finally from 25% to 40% in another 5 min. The outlet of the analytical column was coupled directly to a Fusion Lumos (Thermo) mass spectrometer using the proxeon nanoflow source in positive ion mode.

Peptides were introduced into the Fusion Lumos via a Pico-Tip Emitter (360- μm outer diameter [o.d.] and 20- μm i.d.; 10 μm tip [New Objective]) and an applied spray voltage of 2.4 kV. The capillary temperature was set at 275°C. A full mass scan was acquired with a mass range of 375 to 1,500 m/z in profile mode in the Orbitrap with resolution of 120,000. The filling time was set at maximum of 50 ms with a limitation of 4×10^5 ions. Data-dependent acquisition (DDA) was performed with the resolution of the Orbitrap set at 30,000, with a fill time of 94 ms and a limitation of 1×10^5 ions. A normalized collision energy of 38 was applied. MS² data were acquired in profile mode.

MS data analysis. IsobarQuant (57) and Mascot (v2.2.07) (58) were used to process the acquired data, which was then searched against a UniProt *Escherichia coli* proteome database (UniProt accession no. UP000000625; downloaded on 14 May 2016) containing common contaminants and reversed sequences (37). The following modifications were included into the search parameters: Carbamidomethyl (C) and TMT10 (K) (fixed modification), Acetyl (Protein N-term), Oxidation (M) and TMT10 (N-term) (variable modifications).

For the full scan (MS1), a mass error tolerance of 10 ppm, and for MS/MS (MS2) spectra of 0.02 Da was set. Further parameters were set: trypsin as protease with an allowance of maximum two missed cleavages; a minimum peptide length of seven amino acids; at least two unique peptides were required for a protein identification. The false-discovery rate on the peptide and protein levels was set at 0.01.

Statistical analysis of MS data. The protein.txt output files of IsobarQuant (57) were processed using the R programming language (59). As a quality criterion, only proteins that were quantified with at least two unique peptides were used. Raw tmt reporter ion signals (signal_sum columns) were first batch cleaned using the removeBatchEffect function from limma (60) and then normalized using the vsn package (61). Normalized data were clustered in four clusters using the kmeans function of the stat package in R.

SDS-PAGE and Coomassie blue staining. Protein samples were solubilized and reduced by boiling at 95°C for 5 min in Laemmli loading buffer (200 mM Tris-HCl [pH 6.8], 8% SDS, 40% glycerol, 400 mM DTT, 0.02% bromophenol blue). Solubilized samples were loaded and separated in gradient gels of 4 to 20% acrylamide (Teo-Tricine gels from Expedeon, NXG42012) using the running buffer (Run-Blue running buffer consisted of 0.8 M Tricine, 1.2 M triethanolamine, and 2% SDS). Bio-Rad systems were used, applying 100 V per chamber. For Coomassie blue staining, gels were incubated in staining solution (50% methanol, 40% H₂O, 10% acetic acid, 1 g Brilliant Blue R250 per 1 liter) for 1 h, and destained with destaining solution (40% ethanol, 10% acetic acid, 50% H₂O) until the desirable signal was achieved. Incubations were performed at room temperature with constant moderate mixing by rocking.

Immunoblot analysis. Proteins were separated on acrylamide gels as described above and transferred to methanol-activated polyvinylidene difluoride (PVDF) membranes (Merck, IPVH00010), using Western blot transfer buffer (3.03 g Tris, 14.4 g glycine, 200 ml methanol [all per 1 liter]) for 1.5 h at 100 V. All the incubation steps from here on were performed with constant moderate agitation by using rocking platforms. Membranes were blocked for 1 h with 5% skim milk in Tris-buffered saline with Tween 20 (TBST) (20 mM Tris, 10 mM NaCl, 0.1% Tween 20) and then incubated with appropriately diluted primary antibodies (anti-BamA diluted 1:10,000 and anti-SecG diluted 1:6,000) in 5% skim milk in TBST overnight at 4°C. After three 5-min washes with TBST, membranes were incubated for 1 h with secondary anti-rabbit antibodies conjugated with horseradish peroxidase (HRP) (catalog no. NA934; GE Health Care) diluted 1:10,000 in 5% skim milk in TBST. After these antibody incubations, membranes were washed again three times for 5 min with TBST. Proteins were detected by adding ECL substrate (catalog no. RPN2106; GE Healthcare), then exposing and visualizing using a digital developing machine (ChemiDoc touch imaging system).

Databases. UniProt (37) was used as the source for protein ID and sequences. Protein localization information on the STEPdb database (35) was summarized and modified as in Table S2 in the supplemental material. Modification includes assignment of single localization for proteins with two or more localization annotations.

Data availability. The mass spectrometry proteomic data have been deposited to the ProteomeXchange Consortium via the PRIDE (62) partner repository with the data set identifier [PXD016403](https://proteomecentral.proteomexchange.org/protein/PXD016403). The code and pipelines used for data analysis are available upon request.

SUPPLEMENTAL MATERIAL

Supplemental material is available online only.

FIG S1, PDF file, 0.6 MB.

FIG S2, PDF file, 1.1 MB.

FIG S3, EPS file, 1.6 MB.

FIG S4, EPS file, 1.7 MB.

FIG S5, EPS file, 1.8 MB.

TABLE S1, XLSX file, 1.4 MB.

TABLE S2, XLSX file, 0.1 MB.

TABLE S3, XLSX file, 0.4 MB.

TABLE S4, XLSX file, 0.02 MB.

ACKNOWLEDGMENTS

We thank Tassos Economou for critically reading the manuscript and providing feedback; members of Typas lab for valuable discussions; H. Tokuda for the SecG antibody; T. Lithgow for the BamA antibody; and members of the EMBL Proteomics Core Facility (PCF), especially Mandy Rettel and Dominic Helm, for assisting with sample preparation and data acquisition.

We acknowledge funding from EMBL for this research. J.S. was supported by fellowships from the EMBL Interdisciplinary Postdoc (EIPOD) program under Marie Skłodowska-Curie Actions COFUND (grant 291772). A.S. was supported by the DFG under a grant in the priority program SPP1617 and EMBL.

This research was conducted by A. Sueki as part of a collaboration for a joint Ph.D. degree from EMBL and Heidelberg University.

We declare that we have no conflicts of interest.

REFERENCES

- Silhavy TJ, Kahne D, Walker S. 2010. The bacterial cell envelope. *Cold Spring Harb Perspect Biol* 2:a000414. <https://doi.org/10.1101/cshperspect.a000414>.
- Bos MP, Robert V, Tommassen J. 2007. Biogenesis of the Gram-negative bacterial outer membrane. *Annu Rev Microbiol* 61:191–214. <https://doi.org/10.1146/annurev.micro.61.080706.093245>.
- Nikaido H. 2003. Molecular basis of bacterial outer membrane permeability revisited. *Microbiol Mol Biol Rev* 67:593–656. <https://doi.org/10.1128/mmr.67.4.593-656.2003>.
- May KL, Grabowicz M. 2018. The bacterial outer membrane is an evolving antibiotic barrier. *Proc Natl Acad Sci U S A* 115:8852–8854. <https://doi.org/10.1073/pnas.1812779115>.
- Luirink J, Yu Z, Wagner S, de Gier J-W. 2012. Biogenesis of inner membrane proteins in *Escherichia coli*. *Biochim Biophys Acta* 1817:965–976. <https://doi.org/10.1016/j.bbabi.2011.12.006>.
- Kuhn A, Koch H-G, Dalbey RE. 7 March 2017, posting date. Targeting and insertion of membrane proteins. *EcoSal Plus* 2016 <https://doi.org/10.1128/ecosalplus.ESP-0012-2016>.
- Jacob-Dubuisson F, Mechaly A, Betton J-M, Antoine R. 2018. Structural insights into the signalling mechanisms of two-component systems. *Nat Rev Microbiol* 16:585–593. <https://doi.org/10.1038/s41579-018-0055-7>.
- Typas A, Banzhaf M, Gross CA, Vollmer W. 2011. From the regulation of peptidoglycan synthesis to bacterial growth and morphology. *Nat Rev Microbiol* 10:123–136. <https://doi.org/10.1038/nrmicro2677>.
- Rojas ER, Billings G, Odermatt PD, Auer GK, Zhu L, Miguel A, Chang F, Weibel DB, Theriot JA, Huang KC. 2018. The outer membrane is an essential load-bearing element in Gram-negative bacteria. *Nature* 559:617–621. <https://doi.org/10.1038/s41586-018-0344-3>.
- Typas A, Sourjik V. 2015. Bacterial protein networks: properties and functions. *Nat Rev Microbiol* 13:559–572. <https://doi.org/10.1038/nrmicro3508>.
- Cowles CE, Li Y, Semmelhack MF, Cristea IM, Silhavy TJ. 2011. The free and bound forms of Lpp occupy distinct subcellular locations in *Escherichia coli*. *Mol Microbiol* 79:1168–1181. <https://doi.org/10.1111/j.1365-2958.2011.07539.x>.
- Dunstan RA, Hay ID, Wilksch JJ, Schittenhelm RB, Purcell AW, Clark J, Costin A, Ramm G, Strugnell RA, Lithgow T. 2015. Assembly of the secretion pores GspD, Wza and CsgG into bacterial outer membranes does not require the Omp85 proteins BamA or TamA. *Mol Microbiol* 97:616–629. <https://doi.org/10.1111/mmi.13055>.
- Webb CT, Selkrig J, Perry AJ, Noinaj N, Buchanan SK, Lithgow T. 2012. Dynamic association of BAM complex modules includes surface exposure of the lipoprotein BamC. *J Mol Biol* 422:545–555. <https://doi.org/10.1016/j.jmb.2012.05.035>.
- Cho S, Szewczyk J, Pesavento C, Zietek M, Banzhaf M, Roszczenko P, Asmar A, Laloux G, Hov A, Leverrier P, Van der Henst C, Vertommen D, Typas A, Collet J-F. 2014. Detecting envelope stress by monitoring β -barrel assembly. *Cell* 159:1652–1664. <https://doi.org/10.1016/j.cell.2014.11.045>.
- Konovalova A, Perlman DH, Cowles CE, Silhavy TJ. 2014. Transmembrane domain of surface-exposed outer membrane lipoprotein RcsF is threaded through the lumen of β -barrel proteins. *Proc Natl Acad Sci U S A* 111:E4350–E4358. <https://doi.org/10.1073/pnas.1417138111>.
- Létoquart J, Rodriguez-Alonso R, Nguyen VS, Louis G, Calabrese AN, Radford SE, Cho S-H, Remaut H, Collet J-F. 2019. Structural insight into the formation of lipoprotein- β -barrel complexes by the β -barrel assembly machinery. *bioRxiv* <https://doi.org/10.1101/823146>.
- Beeby M, Ribardo DA, Brennan CA, Ruby EG, Jensen GJ, Hendrixson DR. 2016. Diverse high-torque bacterial flagellar motors assemble wider stator rings using a conserved protein scaffold. *Proc Natl Acad Sci U S A* 113:E1917–E1926. <https://doi.org/10.1073/pnas.1518952113>.
- Deng W, Marshall NC, Rowland JL, McCoy JM, Worrall LJ, Santos AS, Strynadka NCJ, Finlay BB. 2017. Assembly, structure, function and regulation of type III secretion systems. *Nat Rev Microbiol* 15:323–337. <https://doi.org/10.1038/nrmicro.2017.20>.
- Lloubès R, Cascales E, Walburger A, Bouvet E, Lazdunski C, Bernadac A, Journé L. 2001. The Tol-Pal proteins of the *Escherichia coli* cell envelope: an energized system required for outer membrane integrity? *Res Microbiol* 152:523–529. [https://doi.org/10.1016/s0923-2508\(01\)01226-8](https://doi.org/10.1016/s0923-2508(01)01226-8).
- Gray AN, Egan AJ, van't Veer IL, Verheul J, Colavin A, Koumoutsis A, Biboy J, Altelaar AFM, Damen MJ, Huang KC, Simorre J-P, Breukink E, den Blaauwen T, Typas A, Gross CA, Vollmer W. 2015. Coordination of peptidoglycan synthesis and outer membrane constriction during *Escherichia coli* cell division. *Elife* 4:e07118. <https://doi.org/10.7554/eLife.07118>.
- Petiti M, Serrano B, Faure L, Lloubès R, Mignot T, Duché D. 2019. Tol energy-driven localization of Pal and anchoring to the peptidoglycan promote outer-membrane constriction. *J Mol Biol* 431:3275–3288. <https://doi.org/10.1016/j.jmb.2019.05.039>.
- Typas A, Banzhaf M, van den Berg van Saparoea B, Verheul J, Biboy J, Nichols RJ, Zietek M, Beilharz K, Kannenberg K, von Rechenberg M, Breukink E, den Blaauwen T, Gross CA, Vollmer W. 2010. Regulation of peptidoglycan synthesis by outer-membrane proteins. *Cell* 143:1097–1109. <https://doi.org/10.1016/j.cell.2010.11.038>.
- Paradis-Bleau C, Markovski M, Uehara T, Lupoli TJ, Walker S, Kahne DE, Bernhardt TG. 2010. Lipoprotein cofactors located in the outer membrane activate bacterial cell wall polymerases. *Cell* 143:1110–1120. <https://doi.org/10.1016/j.cell.2010.11.037>.
- Egan AJF, Jean NL, Koumoutsis A, Bougault CM, Biboy J, Sassine J, Solovyova AS, Breukink E, Typas A, Vollmer W, Simorre J-P. 2014. Outer-membrane lipoprotein LpoB spans the periplasm to stimulate the peptidoglycan synthase PBP1B. *Proc Natl Acad Sci U S A* 111:8197–8202. <https://doi.org/10.1073/pnas.1400376111>.
- Du D, Wang Z, James NR, Voss JE, Klimont E, Ohene-Agyei T, Venter H, Chiu W, Luisi BF. 2014. Structure of the AcrAB-TolC multidrug efflux pump. *Nature* 509:512–515. <https://doi.org/10.1038/nature13205>.
- Selkrig J, Mosbahi K, Webb CT, Belousoff MJ, Perry AJ, Wells TJ, Morris F, Leyton DL, Totsika M, Phan M-D, Celik N, Kelly M, Oates C, Hartland EL,

- Robins-Browne RM, Ramarathinam SH, Purcell AW, Schembri MA, Strugnell RA, Henderson IR, Walker D, Lithgow T. 2012. Discovery of an archetypal protein transport system in bacterial outer membranes. *Nat Struct Mol Biol* 19:506–510. <https://doi.org/10.1038/nsmb.2261>.
27. Celia H, Noinaj N, Zakharov SD, Bordignon E, Botos I, Santamaria M, Barnard TJ, Cramer WA, Lloubes R, Buchanan SK. 2016. Structural insight into the role of the Ton complex in energy transduction. *Nature* 538: 60–65. <https://doi.org/10.1038/nature19757>.
 28. Szewczyk J, Collet J-F. 2016. The journey of lipoproteins through the cell: one birthplace, multiple destinations. *Adv Microb Physiol* 69:1–50. <https://doi.org/10.1016/bs.ampbs.2016.07.003>.
 29. Almagro Armenteros JJ, Tsirigos KD, Sønderby CK, Petersen TN, Winther O, Brunak S, von Heijne G, Nielsen H. 2019. SignalP 5.0 improves signal peptide predictions using deep neural networks. *Nat Biotechnol* 37: 420–423. <https://doi.org/10.1038/s41587-019-0036-z>.
 30. Bagos PG, Nikolaou EP, Liakopoulos LD, Tsirigos KD. 2010. Combined prediction of Tat and Sec signal peptides with hidden Markov models. *Bioinformatics* 26:2811–2817. <https://doi.org/10.1093/bioinformatics/btq530>.
 31. Loos MS, Ramakrishnan R, Vranken W, Tsirigotaki A, Tsare E-P, Zorzini V, De Geyter J, Yuan B, Tsamardinos I, Klappa M, Schymkowitz J, Rousseau F, Karamanou S, Economou A. 2019. Structural basis of the subcellular topology landscape of *Escherichia coli*. *Front Microbiol* 10:1670. <https://doi.org/10.3389/fmicb.2019.01670>.
 32. Krogh A, Larsson B, von Heijne G, Sonnhammer EL. 2001. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J Mol Biol* 305:567–580. <https://doi.org/10.1006/jmbi.2000.4315>.
 33. Bagos PG, Liakopoulos TD, Spyropoulos IC, Hamodrakas SJ. 2004. PRED-TMBB: a web server for predicting the topology of beta-barrel outer membrane proteins. *Nucleic Acids Res* 32:W400–W404. <https://doi.org/10.1093/nar/gkh417>.
 34. Berven FS, Flikka K, Jensen HB, Eidhammer I. 2004. BOMP: a program to predict integral β -barrel outer membrane proteins encoded within genomes of Gram-negative bacteria. *Nucleic Acids Res* 32:W394–W399. <https://doi.org/10.1093/nar/gkh351>.
 35. Orfanoudaki G, Economou A. 2014. Proteome-wide subcellular topologies of *E. coli* polypeptides database (STePdb). *Mol Cell Proteomics* 13:3674–3687. <https://doi.org/10.1074/mcp.O114.041137>.
 36. Keseler IM, Mackie A, Santos-Zavaleta A, Billington R, Bonavides-Martínez C, Caspi R, Fulcher C, Gama-Castro S, Kothari A, Krummenacker M, Latendresse M, Muñoz-Rascado L, Ong Q, Paley S, Peralta-Gil M, Subhraveti P, Velázquez-Ramírez DA, Weaver D, Collado-Vides J, Paulsen I, Karp PD. 2017. The EcoCyc database: reflecting new knowledge about *Escherichia coli* K-12. *Nucleic Acids Res* 45:D543–D550. <https://doi.org/10.1093/nar/gkw1003>.
 37. UniProt Consortium. 2019. UniProt: a worldwide hub of protein knowledge. *Nucleic Acids Res* 47:D506–D515. <https://doi.org/10.1093/nar/gky1049>.
 38. Babu M, Bundalovic-Torma C, Calmettes C, Phanse S, Zhang Q, Jiang Y, Minic Z, Kim S, Mehla J, Gagarianova A, Rodionova I, Kumar A, Guo H, Kagan O, Pogoutse O, Aoki H, Deineko V, Caufield JH, Holtzapfle E, Zhang Z, Vastermark A, Pandya Y, Lai CC-L, El Bakkouri M, Hooda Y, Shah N, Burnside D, Hooshyar M, Vlasblom J, Rajagopala SV, Golshani A, Wuchty S, Greenblatt JF, Saier M, Uetz P, Moraes TF, Parkinson J, Emili A. 2018. Global landscape of cell envelope protein complexes in *Escherichia coli*. *Nat Biotechnol* 36:103–112. <https://doi.org/10.1038/nbt.4024>.
 39. Bernsel A, Daley DO. 2009. Exploring the inner membrane proteome of *Escherichia coli*: which proteins are eluding detection and why? *Trends Microbiol* 17:444–449. <https://doi.org/10.1016/j.tim.2009.07.005>.
 40. Tsolis KC, Economou A. 2017. Quantitative proteomics of the *E. coli* membranome. *Methods Enzymol* 586:15–36. <https://doi.org/10.1016/bs.mie.2016.09.026>.
 41. Papanastasiou M, Orfanoudaki G, Koukaki M, Kountourakis N, Sardis MF, Aivaliotis M, Karamanou S, Economou A. 2013. The *Escherichia coli* peripheral inner membrane proteome. *Mol Cell Proteomics* 12:599–610. <https://doi.org/10.1074/mcp.M112.024711>.
 42. Papanastasiou M, Orfanoudaki G, Kountourakis N, Koukaki M, Sardis MF, Aivaliotis M, Tsolis KC, Karamanou S, Economou A. 2016. Rapid label-free quantitative analysis of the *E. coli* BL21(DE3) inner membrane proteome. *Proteomics* 16:85–97. <https://doi.org/10.1002/pmic.201500304>.
 43. Bantscheff M, Lemeer S, Savitski MM, Kuster B. 2012. Quantitative mass spectrometry in proteomics: critical review update from 2007 to the present. *Anal Bioanal Chem* 404:939–965. <https://doi.org/10.1007/s00216-012-6203-4>.
 44. Anwari K, Poggio S, Perry A, Gatsos X, Ramarathinam SH, Williamson NA, Noinaj N, Buchanan S, Gabriel K, Purcell AW, Jacobs-Wagner C, Lithgow T. 2010. A modular BAM complex in the outer membrane of the α -proteobacterium *Caulobacter crescentus*. *PLoS One* 5:e8619. <https://doi.org/10.1371/journal.pone.0008619>.
 45. Werner T, Sweetman G, Savitski MF, Mathieson T, Bantscheff M, Savitski MM. 2014. Ion coalescence of neutron encoded TMT 10-plex reporter ions. *Anal Chem* 86:3594–3601. <https://doi.org/10.1021/ac500140s>.
 46. Mateus A, Bobonis J, Kurzawa N, Stein F, Helm D, Hevler J, Typas A, Savitski MM. 2018. Thermal proteome profiling in bacteria: probing protein state in vivo. *Mol Syst Biol* 14:e8242. <https://doi.org/10.15252/msb.20188242>.
 47. Moussatova A, Kandt C, O'Mara ML, Tieleman DP. 2008. ATP-binding cassette transporters in *Escherichia coli*. *Biochim Biophys Acta* 1778: 1757–1771. <https://doi.org/10.1016/j.bbamem.2008.06.009>.
 48. Higgs PI, Letain TE, Merriam KK, Burke NS, Park H, Kang C, Postle K. 2002. TonB interacts with nonreceptor proteins in the outer membrane of *Escherichia coli*. *J Bacteriol* 184:1640–1648. <https://doi.org/10.1128/jb.184.6.1640-1648.2002>.
 49. Letain TE, Postle K. 1997. TonB protein appears to transduce energy by shuttling between the cytoplasmic membrane and the outer membrane in *Escherichia coli*. *Mol Microbiol* 24:271–283. <https://doi.org/10.1046/j.1365-2958.1997.3331703.x>.
 50. Li G, Young KD. 2012. Isolation and identification of new inner membrane-associated proteins that localize to cell poles in *Escherichia coli*. *Mol Microbiol* 84:276–295. <https://doi.org/10.1111/j.1365-2958.2012.08021.x>.
 51. Lim B, Miyazaki R, Neher S, Siegle DA, Ito K, Walter P, Akiyama Y, Yura T, Gross CA. 2013. Heat shock transcription factor σ 32 co-opts the signal recognition particle to regulate protein homeostasis in *E. coli*. *PLoS Biol* 11:e1001735. <https://doi.org/10.1371/journal.pbio.1001735>.
 52. Li G, Young KD. 2015. A new suite of *tnaA* mutants suggests that *Escherichia coli* tryptophanase is regulated by intracellular sequestration and by occlusion of its active site. *BMC Microbiol* 15:14. <https://doi.org/10.1186/s12866-015-0346-3>.
 53. Majdalani N, Hernandez D, Gottesman S. 2002. Regulation and mode of action of the second small RNA activator of RpoS translation, RprA. *Mol Microbiol* 46:813–826. <https://doi.org/10.1046/j.1365-2958.2002.03203.x>.
 54. Hughes CS, Foehr S, Garfield DA, Furlong EE, Steinmetz LM, Krijgsvelde J. 2014. Ultrasensitive proteome analysis using paramagnetic bead technology. *Mol Syst Biol* 10:757. <https://doi.org/10.15252/msb.20145625>.
 55. Hughes CS, Moggridge S, Müller T, Sorensen PH, Morin GB, Krijgsvelde J. 2019. Single-pot, solid-phase-enhanced sample preparation for proteomics experiments. *Nat Protoc* 14:68–85. <https://doi.org/10.1038/s41596-018-0082-x>.
 56. Reichel M, Liao Y, Rettel M, Ragan C, Evers M, Alleaume A-M, Horos R, Hentze MW, Preiss T, Millar AA. 2016. In planta determination of the mRNA-binding proteome of *Arabidopsis* etiolated seedlings. *Plant Cell* 28:2435–2452. <https://doi.org/10.1105/tpc.16.00562>.
 57. Franken H, Mathieson T, Childs D, Sweetman GMA, Werner T, Tögel I, Doce C, Gade S, Bantscheff M, Drewes G, Reinhard FBM, Huber W, Savitski MM. 2015. Thermal proteome profiling for unbiased identification of direct and indirect drug targets using multiplexed quantitative mass spectrometry. *Nat Protoc* 10:1567–1593. <https://doi.org/10.1038/nprot.2015.101>.
 58. Perkins DN, Pappin DJC, Creasy DM, Cottrell JS. 1999. Probability-based protein identification by searching sequence databases using mass spectrometry data. *Electrophoresis* 20:3551–3567. [https://doi.org/10.1002/\(SICI\)1522-2683\(19991201\)20:18<3551::AID-ELPS3551>3.0.CO;2-2](https://doi.org/10.1002/(SICI)1522-2683(19991201)20:18<3551::AID-ELPS3551>3.0.CO;2-2).
 59. R Core Team. 2019. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.
 60. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, Smyth GK. 2015. limma powers differential expression analyses for RNA-seq and microarray studies. *Nucleic Acids Res* 43:e47. <https://doi.org/10.1093/nar/gkv007>.
 61. Huber W, von Heydebreck A, Sültmann H, Poustka A, Vingron M. 2002. Variance stabilization applied to microarray data calibration and to the quantification of differential expression. *Bioinformatics* 18(Suppl 1): S96–S104. https://doi.org/10.1093/bioinformatics/18.suppl_1.s96.
 62. Perez-Riverol Y, Csordas A, Bai J, Bernal-Llinares M, Hewapathirana S, Kundu DJ, Inuganti A, Griss J, Mayer G, Eisenacher M, Pérez E, Uszkoreit J, Pfeuffer J, Sachsenberg T, Yilmaz Ş, Tiwary S, Cox J, Audain E, Walzer M, Jarnuczak AF, Ternent T, Brazma A, Vizcaíno JA. 2019. The PRIDE

- database and related tools and resources in 2019: improving support for quantification data. *Nucleic Acids Res* 47:D442–D450. <https://doi.org/10.1093/nar/gky1106>.
63. Chng S-S, Gronenberg LS, Kahne D. 2010. Proteins required for lipopolysaccharide assembly in *Escherichia coli* form a transenvelope complex. *Biochemistry* 49:4565–4567. <https://doi.org/10.1021/bi100493e>.
64. Yang DC, Peters NT, Parzych KR, Uehara T, Markovski M, Bernhardt TG. 2011. An ATP-binding cassette transporter-like complex governs cell-wall hydrolysis at the bacterial cytokinetic ring. *Proc Natl Acad Sci U S A* 108:E1052–E1060. <https://doi.org/10.1073/pnas.1107780108>.
65. Lacqua A, Wanner O, Colangelo T, Martinotti MG, Landini P. 2006. Emergence of biofilm-forming subpopulations upon exposure of *Escherichia coli* to environmental bacteriophages. *Appl Environ Microbiol* 72:956–959. <https://doi.org/10.1128/AEM.72.1.956-959.2006>.
66. Li H, Wang B-C, Xu W-J, Lin X-M, Peng X-X. 2008. Identification and network of outer membrane proteins regulating streptomycin resistance in *Escherichia coli*. *J Proteome Res* 7:4040–4049. <https://doi.org/10.1021/pr800310y>.
67. d'Alençon E, Taghbalout A, Kern R, Kohiyama M. 1999. Replication cycle dependent association of SeqA to the outer membrane fraction of *E. coli*. *Biochimie* 81:841–846. [https://doi.org/10.1016/s0300-9084\(99\)00212-6](https://doi.org/10.1016/s0300-9084(99)00212-6).