# Systematic localization of Gram-negative bacterial membrane proteins

Anna Sueki[1,2], Frank Stein[1], Mikhail Savitski[1], Joel Selkrig[1] ‡, Athanasios Typas[1]‡

[1] European Molecular Biology Laboratory (EMBL), Genome Biology Unit, Meyerhofstrasse 1, 69117 Heidelberg, Germany
[2] Collaboration for joint PhD degree between EMBL and Heidelberg University, Faculty of Biosciences

‡ to whom correspondence should be addressed:
selkrig@embl.de & typas@embl.de

## Abstract

The molecular architecture and function of the Gram-negative bacterial cell envelope is dictated by protein composition and localization. Proteins that localize to the inner (IM) and outer (OM) membranes 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. We employed multiplexed quantitative mass spectrometry to assess membrane protein localization in a proteome-wide fashion by separating IM and OM vesicles from exponentially growing *E. coli* K-12 cells on a sucrose density gradient. The migration patterns for >1600 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 to the IM (Orfanoudaki and Economou, 2014) and display potential for dual localization to either the IM or cytoplasm, we could allocate 110 to the IM and 206 as soluble 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 them, we were able to find supportive evidence for our localization findings in literature. We anticipate our systems-level analysis of the *E. coli* membrane proteome will serve as a useful reference dataset to query membrane protein localization, as well as provide a novel methodology to rapidly and systematically map membrane protein localization in more poorly characterized Gram-negative species.

# Introduction

The inner and outer membrane (IM and OM) of Gram-negative bacteria carry out fundamental cellular functions crucial for cell viability (Silhavy et al., 2010). The OM directly interfaces with the extracellular environment and provides a formidable physical barrier that excludes the passage of large and hydrophobic compounds (Bos et al., 2007; May and Grabowicz, 2018; Nikaido, 2003). The IM plays a vital role in ensuring selective transport of small compounds into and out of the cell, as well as sensing external cues and transducing information to adaptive transcriptional responses (Jacob-Dubuisson et al., 2018; Kuhn et al., 2017; Luirink et al., 2012). 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 (Rojas et al., 2018; Typas et al., 2011). 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 (Typas and Sourjik, 2015).

Proteins localizing to the IM or OM can be sub-categorized based on 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 multi-protein 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. Lipoproteins 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 (Cowles et al., 2011; Dunstan et al., 2015; Webb et al., 2012), even by traversing through OMPs (Cho et al., 2014; Konovalova et al., 2014; Létoquart et al., 2019).

While most proteins localize to, and function at, either the IM or OM, several trans-envelope protein complexes span both membranes. Some of the larger complexes like flagella (Beeby et al., 2016) and secretion apparatuses (Deng et al., 2017) contain dedicated IM, OM and periplasmic components, whereas smaller ones such as the Tol-Pal complex (Gray et al., 2015; Lloubès et al., 2001; Petiti et al., 2019), the PBP1a/b-LpoA/B peptidoglycan synthase complexes (Egan et al., 2014; Paradis-Bleau et al., 2010; Typas et al., 2010), the efflux pump AcrAB-TolC (Du et al., 2014), the TAM translocation complex (Selkrig et al., 2012) and all TonB-dependent transport complexes (Celia et al., 2016) 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 to 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 (Szewczyk and Collet, 2016). 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 *E. 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 (Almagro Armenteros et al., 2019) 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 (Bagos et al., 2010). Interestingly, periplasmic proteins bear sufficiently distinguishable biophysical, biochemical and structural characteristics when compared to their cytoplasmic counterparts, so that one can accurately discern them with signal-sequencing

agnostic machine-learning predictors (Loos et al., 2019). 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 (Krogh et al., 2001), as well as algorithms to predict β-barrel folding of OMPs such as PRED-TMBB (Bagos et al., 2004) and BOMP (Berven et al., 2004). *E. coli* genome databases, such as STEPdb (Orfanoudaki and Economou, 2014), Ecocyc (Keseler et al., 2017) and Uniprot (The UniProt Consortium, 2019), 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 mis-annotations, as it has been discussed previously (Babu et al., 2018; Orfanoudaki and Economou, 2014). 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 (Bernsel and Daley, 2009; Papanastasiou et al., 2013, 2016; Tsolis and Economou, 2017), 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 (Bantscheff et al., 2012). 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 mis-annotated 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 (Figure 1A). Total membrane vesicles were purified, followed by IM and OM vesicle separation on a sucrose density gradient as previously described (Anwari et al., 2010). The sucrose gradient was then separated into eleven fractions and analyzed by immunoblot and SDS-PAGE. Effective IM and OM vesicle separation was verified using SecG and BamA antiserum (Figure 1B). Fractions 2 to 11 and the total membrane (input sample prior to fractionation) were labelled with 11-plex tandem mass tag (TMT) reagents (Werner et al., 2014) and analyzed and quantified using mass spectrometry (MS) (Figure 1A and S1A). In total, we identified 1605 common proteins across the two biological replicates, and thus proceeded to data normalization and quantification as described in the materials and methods (Figure S1A and Table S1).

To assess protein abundance in each fraction, we calculated the logarithmic fold change ($log_2FC$) of each sucrose gradient fraction relative to the total membrane fraction for each protein. The fractionation pattern across the 10 quantified fractions (fraction 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 (Figure 1B lower). Replicate correlation between independent experiments for all $log_2FC$ values was high ($R = 0.77$; Pearson correlation, $p < 0.001$, Figure S1B). Proteome coverage was analyzed by comparison to protein localization annotation from STEPdb database (Orfanoudaki and Economou, 2014) using Uniprot IDs (summarized and modified as in Table S2). For membrane protein annotation categories, we had an overall 56% coverage (973 out of 1741 proteins) with several categories reaching 70% coverage, whereas non-membrane protein categories did not exceed 25% coverage (Figure 1C). Altogether we could quantitatively assess the sucrose gradient fractionation of 1605 proteins (Table S3), covering most membrane-annotated proteins in STEPdb.

**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 (Figure 1A-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 high, and $log_2$ of f02, f03 and f04 for low) 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 STEPdb database, Table S2). As anticipated, most IM protein categories (i.e. IM-integral, IM-peri and IMLP) displayed a low sucrose gradient ratio, whereas the two OM protein categories, OMPs and OMLPs, showed high sucrose gradient ratios (Figure 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 cut-offs to define IM and OM protein localization using the sucrose gradient ratio (Figure 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 non-specifically 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 (Figure 2A, Table S3). This suggests that the IM-cyto category of proteins referred to as “peripheral IM proteins” in STEPdb and originally described in another study (Papanastasiou et al., 2013), 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 (Figure 1B lower panel after normalization), we asked whether unbiased clustering of the fractionation patterns could be used to distinguish proteins according to their annotated membrane localization. Based on PCA analysis of all identified proteins, we set the number of clusters to 4 for k-means clustering analysis (Figure S2). This resulted in four groups containing the following number of proteins - cluster 1: 111, cluster 2: 650, cluster 3: 680 and cluster 4: 164 (Figure 3 left, Table S3), where each cluster exhibited distinctive fractionation patterns (Figure 3 middle). 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 (Figure 3 right). 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-cluster 1, IM-cluster 2 or soluble-cluster 3 and 4), some proteins were grouped into clusters that conflicted with their STEPdb annotated localization with the IM-cyto group featuring prominently in various clusters (Figure 3 right). 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 carried out two further steps. First, we ran k-means clustering for the dataset where the two replicates were treated separately. Out of this, we identified 140 (out of 1605) proteins whose fractionation patterns between replicates resulted in clustering to different localization clusters (OM-cluster 1, IM-cluster 2 or soluble-cluster 3 and 4)

for the two replicates. We reasoned that this is due to irreproducibility between the replicates and removed these proteins from further analysis. Second, we assessed the similarity of our two methods (k-means vs. sucrose gradient ratio) in assessing protein localization. To do this, we used the thresholds of sucrose gradient ratio for IM and OM defined in Figure 2B. We found a large overlap between the two methods for all three localization categories: IM, OM and soluble (Figure 4A). In total, we identified 1368 proteins (out of 1465 possible) to agree between the two methods, which we further used as our high confidence protein localization dataset. When considering all 1605 proteins, the two methods agreed in 1456 proteins (Figure 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 non-overlap 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 (Figure S3).

As mentioned before, IM-cyto proteins showed a bimodal distribution of sucrose gradient ratio (Figure 2A) and could be found in multiple clusters (2, 3 and 4; Figure 3 right). Having now high-confident protein localization allocations, we could define 110 proteins as IM and 206 as soluble (Figure 4B; see also Figure 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 (Mateus et al., 2018). In agreement with this, IM-cyto proteins categorized here to be soluble had similarly low melting temperatures as cytoplasmic proteins (Figure 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 (Figure S5). Thus, in the experimental conditions we tested, IM-cyto annotated proteins resulted in a mixture of soluble and IM proteins, which through their fractionation patterns, we could allocate their predominant protein localization.

## Identification of potentially mis-annotated proteins

STEPdb 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 dataset 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 1368) high-confidence proteins, including both membrane and soluble, 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. Firstly, we checked the localization annotation in another recent study (Babu et al., 2018). Babu *et al*. 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 in Babu *et al*. We thus conclude that the STEPdb localization for these 11 proteins was likely inaccurate (Table S4). We found corroborating evidence for 12 more such cases in 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 is 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. Firstly, we detected several proteins annotated as solely periplasmic in STEPdb fractionated as membrane proteins in our experiments. Many of them have known interacting membrane partners, which is presumably the reason they co-fractionate with either the IM (EnvC, FdoG, NapG, and RseB) or the OM (LptA) (Figure 4C, Table S4). The situation was similar for periplasmic components of IM ABC transporter complexes (FhuD, PstS, and SapA) which co-fractionated with IM proteins, possibly as a consequence of a direct conditional association with IM proteins upon active transport (Moussatova et al., 2008), but were only annotated as periplasmic in STEPdb (Table S4). In total, there were 19 cases for which STEPdb had incomplete annotation.

295 Conversely, FecB, a known periplasmic component of an ABC transporter complex is
296 annotated both as IM-peri (peripherally associated to IM) and periplasmic in STEPdb,
297 but only identified as soluble in our experimental conditions. In this case, we are
298 failing to detect the IM-association because the transporter is likely inactive in the
299 conditions we probe, and the STEPdb annotation is more accurate (Table S4).
300 Moreover, IM proteins known to form trans-envelope complexes (e.g. TamB and
301 TonB) failed to cluster as either IM or OM proteins in the fractionation experiments.
302 Overall, we could reasonably explain 51 out of the 63 cases where STEPdb and our
303 results disagreed, out of which we could find additional information that supports our
304 localization call (42 proteins) or the original STEPdb annotation (9 proteins)
305 (summarized in Table S4). Overall, these findings demonstrate that our quantitative
306 assessment of protein localization captures accurately the *in vivo* biological state.
307

# Discussion

We quantified the membrane proteome using TMT-labelling 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 immunoblot 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 (Papanastasiou et al., 2013). STEPdb categorizes proteins that peripherally interact with the cytoplasmic face of the IM as a “peripheral IM protein”, which we referred to here for simplicity as IM-cyto (Table S2). Although we detect the majority of IM-cyto proteins (68% of 559 proteins) in the membrane fraction (which is depleted from soluble proteins), most of them are reproducibly assigned as soluble proteins according to their sucrose fractionation pattern (206 out of 316 high-confident calls). This absence of co-fractionation 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 that some of these proteins have conditional, low affinity or transient association with the IM and proteins therein, or a small fraction of the total protein amount is at any given point 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 1368 which were inconsistent with the reported localization annotation in STEPdb. 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 (9). In most cases, sucrose gradient fractionation failed to make the right call when the protein was spanning the envelope or had presumably dual membrane localization. It

342 is likely that the new localization is also correct for most of 12 remaining proteins
343 (Table S4). Thus, our data are helpful for improving protein localization, even for an
344 organism as intensively studied as *E. coli*, which has been subjected to a plethora of
345 targeted and systematic studies and researchers can benefit from carefully curated
346 databases, such as STEPdb.
347
348 In some cases, sucrose gradient fractionation patterns were indicative of protein
349 activity. For example, periplasmic partners of IM ABC transporters often showed
350 sucrose gradient fractionation patterns similar to that of IM proteins (Figure 4C),
351 which we postulate were due to their strong interaction with the cognate IM ABC
352 transporters in the substrate-bound state (Moussatova et al., 2008). In other cases,
353 they behaved as soluble proteins, which likely reflects the inactive state of the ABC
354 transporter. Interestingly, trans-envelope spanning IM proteins, such as TamB and
355 TonB, displayed fractionation patterns similar to soluble proteins. This is presumably
356 due to strong interactions with their OM counterparts, which pull a subpopulation of
357 the protein together with OM vesicles during ultracentrifugation. Consistent with our
358 observations, TonB was previously found in the OM fraction upon sucrose gradient
359 fractionation (Higgs et al., 2002; Letain and Postle, 1997). This suggests that trans-
360 envelope IM proteins can present distinctive properties upon sucrose gradient
361 fractionation. More broadly it implies that sucrose gradient fractionation can provide
362 insights into the activity and mechanical strength of specific envelope complexes
363 during different growth stages and conditions.
364
365 Not only does this work provide a resource of sucrose gradient fractionation for 1605
366 proteins, with 1368 proteins having their cellular localization confidently assigned, the
367 method we present can be used for rapid and systematic characterization of
368 membrane proteomes in different contexts – growth conditions and stages, and
369 under different cellular perturbations. Many membrane proteins are only conditionally
370 expressed (Mateus et al., 2018), whereas other proteins conditionally relocate in and
371 out of membranes (Li and Young, 2012, 2015; Lim et al., 2013). Importantly, our
372 method allows for the systematic mapping of membrane proteomes from other Gram-
373 negative species for which protein localization annotation and transport mechanisms
374 are less (if at all) studied.
375

## ACKNOWLEDGEMENTS

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. JS was supported by fellowships from the EMBL Interdisciplinary Postdoc (EIPOD) program under Marie Skłodowska-Curie Actions COFUND (grant number 291772). AS was supported by the DFG under a grant in the priority program SPP1617 and EMBL.

## DATA AVAILABILITY

The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE (Perez-Riverol et al., 2019) partner repository with the dataset identifier PXD016403.

## CODE AVAILABILITY

The code and pipelines used for data analysis are available upon request.

## DECLARATION OF INTEREST

The authors declare no competing interests.

**Figure 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 ($OD_{578}$ ~0.7), lysed and ultracentrifuged to collect the membrane fraction containing both OM (red) and 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-labelled and analyzed by LC-MS/MS. After data normalization (see Figure S1), 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. Upper panel: Coomassie stained gel. Middle panel: immunoblot analysis for control proteins: BamA for OM, SecG for IM (middle panel). Lower panel: TMT-labelling MS quantification result of the two control proteins (BamA and SecG). $Log_2$ fold-change of each fraction / total membrane fraction for fractions (f02 to f11) for BamA and SecG. Mean with standard deviation is plotted from the two biological replicates.

C) Fraction of proteins identified for each STEPdb localization category. Localization annotation is summarized in Table S2.

**Figure 2: Sucrose gradient ratio separates proteins according to membrane localization.**

A) Smoothened distributions of sucrose gradient ratios for each protein localization category based in STEPdb (as in Figure 1C). Sucrose gradient ratio values were calculated as a difference of the average of high sucrose fractions ($log_2$ of f08, f09, f10) and the average of the low sucrose fractions ($log_2$ of f02, f03, f04) for each protein.

B) Sucrose gradient ratio of three categories (IM, OM and Sol) grouped from A). Cytoplasmic and periplasmic proteins from (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). Dotted lines refer to the 10th and 90th percentiles for IM, OM and soluble proteins for each membrane localization group, respectively. The 90th

431 percentile for IM and the 10th percentile for OM proteins are shown as solid lines as
432 they are used as thresholds to allocate proteins to the three categories.
433
434 **Figure 3: Unbiased K-means clustering of sucrose gradient fractionation**
435 **patterns accurately depicts protein membrane localization**
436 K-means clustering based on log fold-change of each fraction to total membrane
437 sample. Left: heatmap representing each cluster patterns. Middle: fractionation
438 pattern of all proteins in each cluster (grey) and the distribution average (blue). Right:
439 Pie chart representing annotated localization of proteins found in each cluster.
440 Localization annotation as in Figure 1C.
441
442 **Figure 4: Identification of potentially mis-annotated proteins**
443 A) Venn diagrams showing commonly identified proteins for each localization
444 categories (IM, OM and soluble), based on our two quantification methods. Red
445 denotes protein localizations stemming from the sucrose gradient ratios and blue
446 refers to the k-means clustering annotation. 1465 proteins were assigned localization
447 by the different methods. The two methods agreed in the localization of 1368 proteins,
448 which we used as a “high-confidence” localization set.
449 B) Comparison of identified localization with annotated localization for the 1368
450 proteins from panel A. Protein localizations considered not to match with STEPdb
451 annotations are marked with *. The IM-cyto category has a bimodal distribution,
452 likely because proteins could not be confidently allocated to either the IM or the
453 cytoplasm in STEPdb, but can be with the experimental data provided here.
454 C) Schematic representation explaining experimental localization of proteins not or
455 partially matching the STEPdb annotation (besides IM-cyto): 1) Known interacting
456 partner proteins in membranes (Left: LptA interacts with OM protein, LptD/E and
457 fractionates as OM protein as previously reported (Chng et al., 2010), middle: EnvC
458 interacts with IM proteins FtsEX (Yang et al., 2011), similarly for FdoG, NapG, and
459 RseB), right: Rfa proteins in cytoplasm were identified as IM proteins, possibly due to
460 interactions with bona-fide IM Rfa proteins). 2) Periplasmic partner of ABC
461 transporters detected as IM proteins presumably because they bind to the IM
462 components and transporters when transporters are active (Moussatova et al., 2008):
463 this is for example the case for FhuD interacting with FhuBC upon substrate
464 (ferrichrome, orange) binding. 3) Trans-envelope protein interaction is known: TamB

spans the envelope and interacts with OM-located TamA (Selkrig et al., 2012), and TonB with OM-located TonB-dependent transporters, resulting in a misleading sucrose gradient fractionation (distributed equally in all fractions, thus appearing as soluble protein), which in the case of TonB has been previously shown (Higgs et al., 2002; Letain and Postle, 1997). 4) Previously shown dual localization of cytoplasmic protein in OM for Dps and SeqA (d'Alençon et al., 1999; Lacqua et al., 2006; Li et al., 2008). Full list of 63 proteins not matching with STEPdb, with 51 proteins explained can be found in Table S4.

**Figure S1: Quality control of TMT-labelling MS results**

A) Boxplot representation of summed TMT reporter ion signals (signal sums) distribution across sucrose gradient fractions before and after normalization of the two biological replicates (batch-cleaned using the removeBatchEffect function from limma (Ritchie et al., 2015) and then normalized using the vsn package (Huber et al., 2002)). Box boundaries indicate the upper and lower IQR, the median is depicted by the middle boundary and whiskers represent 1.5x IQR.

B) Replicate correlation of $\log_2$ fold-change (logFC) value of each sucrose gradient fraction to total membrane fraction (10 data points per protein, thus total 16050 data points). Pearson's correlation shown as red line ($R = 0.77$, p-value < 2.2e-16).

**Figure S2: PCA analysis of fractionation patterns**

Each dot represents a protein (total 1605 proteins). x-axis for PC1 (65.4% of variability), y-axis for PC2 (19.8% of variability) and color scale represents k-means clustering groups. The data used is in Table S3 (ratio of each fraction to total membrane, average of two replicates).

**Figure S3: Comparison of k-means clustering to sucrose gradient ratio**

Histogram of sucrose gradient ratio for each k-means cluster is shown for all proteins (1605 proteins) in the assay. Number of total protein found in each cluster are: cluster 1 – 111, cluster 2 – 650, cluster 3 – 680, cluster 4 – 164.

**Figure S4: Identification of potentially mis-annotated proteins (all proteins)**

A) Same analysis as Figure 4A for all 1605 proteins without filtering for clustering reproducibility. Venn diagrams showing commonly identified proteins for each

localization categories (IM, OM and soluble), based on our two quantification methods. Red denotes protein localizations stemming from the sucrose gradient ratios and blue refers to the k-means clustering annotation. The two methods agree in the localization of 1456 proteins.

B) Same analysis as Figure 4B for all proteins without filtering for clustering reproducibility. Comparison of identified localization with annotated localization for the 1456 proteins from panel A. Protein localizations considered to be not matching with annotation are highlighted with *.

**Figure S5: Comparison of protein localization with protein melting temperature**

Violin plots encompassing box plots (plotted as in Fig. S1) of protein melting temperatures (Mateus et al., 2018) as a function of protein localization for cytoplasmic, IM and cyto-IM proteins. The latter was split in the two sub-groups we identified in our experiments. Cytoplasm (n=391), IM (n=218, IM-integral, IM-peri and IMLP combined), IM categorized IM-cyto proteins (n=85), and soluble categorized IM-cyto proteins (n=199).

**Supplemental information**

Table S1. TMT-labelling MS results and normalization (signal sum values)

Table S2. Localization annotation used in this study based on STEPdb

Table S3. Membrane ratio and K-means clustering data

Table S4. Protein list non-matching STEPdb annotations & literature search summary

# Materials and Methods

**Bacterial culturing**

The wild-type strain used in this study is *Escherichia coli* K-12 MG1655 Δ(*argF-lac*)U169 *rprA::lacZ* (Majdalani et al., 2002). Bacterial cells were grown for 4 generations in LB-Lennox (referred as LB herein) medium at 37°C with vigorous shaking 200 rpm, and collected for fractionation while still being in exponential growth phase, at an OD (578nm) of 0.6 – 0.8.

**Membrane vesicle isolation and sucrose density fractionation**

Membrane vesicles were isolated and fractionated essentially as previously described (Anwari et al., 2010) with the following deviations. Phosphate Saline Buffer (PBS) was used as the base buffer instead of Tris. After sucrose gradient separation, 1 mL fractions were collected step-wise from the top of the gradient, yielding 11 fractionated samples that were analyzed by Coomassie staining and Western blotting using SDS-PAGE gels as described below. Fractions 2 to 11 (f02-f11), as well as an aliquot of the total input membrane sample (diluted 10 times in $H_2O$), were labeled with 11-plex TMT and subjected to LC-MS/MS.

**Sample preparation and TMT labelling**

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 minutes in an ultrasonic bath, heated for 10 minutes to 80°C, and sonicated again for another 5 minutes. Disulphide bonds were reduced by incubating at 56°C for 30 minutes in 10 mM dithiothreitol (DTT) buffered with 50 mM HEPES (pH = 8.5). Reduced cysteines were alkylated by incubating 30 minutes at room temperature in dark with 20 mM 2-chloroacetamide in 50 mM HEPES buffer pH = 8.5. Samples were prepared for MS using the SP3 protocol (Hughes et al., 2014, 2019). 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 beads, and combining the second elution wash of beads with HEPES buffer. Collected peptides were labelled with TMT10plex Isobaric Label Reagent (ThermoFisher, (Werner et al., 2014)) and with 131C 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 hour at room temperature. The reaction was quenched with 5% hydroxylamine for 15 minutes at room temperature. Then the 10 samples labelled with unique TMT10plex labels were combined into one sample. The combined sample was then cleaned up using OASIS® HLB µElution 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 Germini C18 column (3 µm, 110

Å, 100 x 1.0 mm, Phenomenex). The fractionation was performed as previously described (Reichel et al., 2016). Samples were pooled in 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 (µ-Precolumn C18 PepMap 100, 5µm, 300 µm i.d. x 5 mm, 100 Å) and an analytical column (nanoEase™ M/Z HSS T3 column 75 µm x 250 mm C18, 1.8 µm, 100 Å, Waters). Trapping was carried out with a constant flow of solvent A (0.1% formic acid in water) at 30 µL/min onto the trapping column for 6 minutes. Subsequently, peptides were eluted via the analytical column with a constant flow of 0.3 µL/min with 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 OD x 20 µm ID; 10 µm tip (New Objective) and an applied spray voltage of 2.4 kV. The capillary temperature was set at 275°C. Full mass scan was acquired with mass range 375-1500 m/z in profile mode in the orbitrap with resolution of 120000. The filling time was set at maximum of 50 ms with a limitation of $4x10^5$ ions. Data dependent acquisition (DDA) was performed with the resolution of the Orbitrap set to 30000, with a fill time of 94 ms and a limitation of $1x10^5$ ions. A normalized collision energy of 38 was applied. $MS^2$ data was acquired in profile mode.

**MS data analysis**

IsobarQuant (Franken et al., 2015) and Mascot (v2.2.07, (Perkins et al., 1999)) were used to process the acquired data, which was then searched against a Uniprot *Escherichia coli* proteome database (UP000000625, downloaded on 05/14/2016) containing common contaminants and reversed sequences (The UniProt Consortium, 2019). 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 peptide and protein level was set to 0.01.

**Statistical analysis of MS data**

The protein.txt output files of IsobarQuant (Franken et al., 2015) were processed using the R programming language (R Core Team, 2019). As a quality criterion, only proteins which 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 (Ritchie et al., 2015) and then normalized using the vsn package (Huber et al., 2002). Normalized data was clustered in 4 clusters using the kmeans function of the stat package in R.

**SDS-PAGE and Coomassie staining**

Protein samples were solubilized and reduced by boiling at 95°C for 5 minutes 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 4-20% acrylamide (Teo-Tricine gels from Expedeon, NXG42012) using the running buffer (Run-Blue running buffer: 0.8 M Tricine, 1.2 M Triethanolamine, 2% SDS). Bio-rad systems were used, applying 100 V per chamber. For Coomassie staining, gels were incubated in staining solution (50% methanol, 40% $H_2O$, 10% acetic acid, 1 g Brilliant Blue R250 per 1 L) for 1 hour, and destained with destaining solution (40% ethanol, 10% acetic acid, 50% $H_2O$) 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 PVDF membranes (Merck, IPVH00010), using Western blot transfer buffer (3.03 g Tris, 14.4g glycine, 200 mL methanol per 1 L) for 1.5 hours 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 hour with 5% skim milk in TBST (20 mM Tris, 10 mM NaCl, 0.1% Tween-20), and then incubated with appropriately diluted primary antibodies (α-BamA – 1:10,000, α-SecG – 1:6,000) in 5% skim milk in TBST overnight at 4°C. After three times of 5 minutes washes with TBST, membranes were incubated for 1 hour with secondary α-rabbit antibodies conjugated with horseradish peroxidase (HRP) (GE healthcare, NA934) diluted by 1:10,000 in 5% skim milk in TBST. After these antibody incubations, membranes were washed again three times for 5 minutes with TBST. Proteins were detected by adding ECL substrate (GE Healthcare, RPN2106), then exposing and visualizing using a digital developing machine (ChemiDoc™ Touch Imaging System).

**Databases**

UniProt (The UniProt Consortium, 2019) was used as the source for protein ID and sequences. Protein localization information on STEPdb database (Orfanoudaki and Economou, 2014) was summarized and modified as in Table S2. Modification includes assignment of single localization for proteins with two or more localization annotations.

# References

Almagro Armenteros, J.J., Tsirigos, K.D., Sønderby, C.K., Petersen, T.N., Winther, O., Brunak, S., von Heijne, G., and Nielsen, H. (2019). SignalP 5.0 improves signal peptide predictions using deep neural networks. Nat. Biotechnol. *37*, 420–423.

Anwari, K., Poggio, S., Perry, A., Gatsos, X., Ramarathinam, S.H., Williamson, N.A., Noinaj, N., Buchanan, S., Gabriel, K., Purcell, A.W., et al. (2010). A Modular BAM Complex in the Outer Membrane of the α-Proteobacterium Caulobacter crescentus. PLoS One *5*, e8619.

Babu, M., Bundalovic-Torma, C., Calmettes, C., Phanse, S., Zhang, Q., Jiang, Y., Minic, Z., Kim, S., Mehla, J., Gagarinova, A., et al. (2018). Global landscape of cell envelope protein complexes in Escherichia coli. Nat. Biotechnol. *36*, 103–112.

Bagos, P.G., Liakopoulos, T.D., Spyropoulos, I.C., and Hamodrakas, S.J. (2004). PRED-TMBB: a web server for predicting the topology of beta-barrel outer membrane proteins. Nucleic Acids Res. *32*, W400-4.

Bagos, P.G., Nikolaou, E.P., Liakopoulos, T.D., and Tsirigos, K.D. (2010). Combined prediction of Tat and Sec signal peptides with hidden Markov models. Bioinformatics *26*, 2811–2817.

Bantscheff, M., Lemeer, S., Savitski, M.M., and Kuster, B. (2012). Quantitative mass spectrometry in proteomics: critical review update from 2007 to the present. Anal. Bioanal. Chem. *404*, 939–965.

Beeby, M., Ribardo, D.A., Brennan, C.A., Ruby, E.G., Jensen, G.J., and Hendrixson, D.R. (2016). Diverse high-torque bacterial flagellar motors assemble wider stator rings using a conserved protein scaffold. Proc. Natl. Acad. Sci. U. S. A. *113*, E1917-26.

Bernsel, A., and Daley, D.O. (2009). Exploring the inner membrane proteome of Escherichia coli: which proteins are eluding detection and why? Trends Microbiol. *17*, 444–449.

Berven, F.S., Flikka, K., Jensen, H.B., and Eidhammer, I. (2004). BOMP: a program to predict integral -barrel outer membrane proteins encoded within genomes of

683 Gram-negative bacteria. Nucleic Acids Res. *32*, W394–W399.

684 Bos, M.P., Robert, V., and Tommassen, J. (2007). Biogenesis of the Gram-Negative
685 Bacterial Outer Membrane. Annu. Rev. Microbiol. *61*, 191–214.

686 Celia, H., Noinaj, N., Zakharov, S.D., Bordignon, E., Botos, I., Santamaria, M.,
687 Barnard, T.J., Cramer, W.A., Lloubes, R., and Buchanan, S.K. (2016). Structural
688 insight into the role of the Ton complex in energy transduction. Nature *538*, 60–65.

689 Chng, S.-S., Gronenberg, L.S., and Kahne, D. (2010). Proteins Required for
690 Lipopolysaccharide Assembly in Escherichia coli Form a Transenvelope Complex.
691 Biochemistry *49*, 4565–4567.

692 Cho, S., Szewczyk, J., Pesavento, C., Zietek, M., Banzhaf, M., Roszczenko, P.,
693 Asmar, A., Laloux, G., Hov, A., Leverrier, P., et al. (2014). Detecting Envelope Stress
694 by Monitoring β-Barrel Assembly. Cell *159*, 1652–1664.

695 Cowles, C.E., Li, Y., Semmelhack, M.F., Cristea, I.M., and Silhavy, T.J. (2011). The
696 free and bound forms of Lpp occupy distinct subcellular locations in Escherichia coli.
697 Mol. Microbiol. *79*, 1168–1181.

698 d'Alençon, E., Taghbalout, A., Kern, R., and Kohiyama, M. (1999). Replication cycle
699 dependent association of SeqA to the outer membrane fraction of E. coli. Biochimie
700 *81*, 841–846.

701 Deng, W., Marshall, N.C., Rowland, J.L., McCoy, J.M., Worrall, L.J., Santos, A.S.,
702 Strynadka, N.C.J., and Finlay, B.B. (2017). Assembly, structure, function and
703 regulation of type III secretion systems. Nat. Rev. Microbiol. *15*, 323–337.

704 Du, D., Wang, Z., James, N.R., Voss, J.E., Klimont, E., Ohene-Agyei, T., Venter, H.,
705 Chiu, W., and Luisi, B.F. (2014). Structure of the AcrAB-TolC multidrug efflux pump.
706 Nature *509*, 512–515.

707 Dunstan, R.A., Hay, I.D., Wilksch, J.J., Schittenhelm, R.B., Purcell, A.W., Clark, J.,
708 Costin, A., Ramm, G., Strugnell, R.A., and Lithgow, T. (2015). Assembly of the
709 secretion pores GspD, Wza and CsgG into bacterial outer membranes does not
710 require the Omp85 proteins BamA or TamA. Mol. Microbiol. *97*, 616–629.

711 Egan, A.J.F., Jean, N.L., Koumoutsi, A., Bougault, C.M., Biboy, J., Sassine, J.,
712 Solovyova, A.S., Breukink, E., Typas, A., Vollmer, W., et al. (2014). Outer-membrane
713 lipoprotein LpoB spans the periplasm to stimulate the peptidoglycan synthase PBP1B.

714 Proc. Natl. Acad. Sci. U. S. A. *111*, 8197–8202.

715 Franken, H., Mathieson, T., Childs, D., Sweetman, G.M.A., Werner, T., Tögel, I.,
716 Doce, C., Gade, S., Bantscheff, M., Drewes, G., et al. (2015). Thermal proteome
717 profiling for unbiased identification of direct and indirect drug targets using
718 multiplexed quantitative mass spectrometry. Nat. Protoc. *10*, 1567–1593.

719 Gray, A.N., Egan, A.J.F., Van't Veer, I.L., Verheul, J., Colavin, A., Koumoutsi, A.,
720 Biboy, J., Altelaar, A.F.M., Damen, M.J., Huang, K.C., et al. (2015). Coordination of
721 peptidoglycan synthesis and outer membrane constriction during Escherichia coli cell
722 division. Elife *4*.

723 Higgs, P.I., Letain, T.E., Merriam, K.K., Burke, N.S., Park, H., Kang, C., and Postle, K.
724 (2002). TonB Interacts with Nonreceptor Proteins in the Outer Membrane of
725 Escherichia coli. J. Bacteriol. *184*, 1640–1648.

726 Huber, W., von Heydebreck, A., Sültmann, H., Poustka, A., and Vingron, M. (2002).
727 Variance stabilization applied to microarray data calibration and to the quantification
728 of differential expression. Bioinformatics *18 Suppl 1*, S96-104.

729 Hughes, C.S., Foehr, S., Garfield, D.A., Furlong, E.E., Steinmetz, L.M., and
730 Krijgsveld, J. (2014). Ultrasensitive proteome analysis using paramagnetic bead
731 technology. Mol. Syst. Biol. *10*, 757–757.

732 Hughes, C.S., Moggridge, S., Müller, T., Sorensen, P.H., Morin, G.B., and Krijgsveld,
733 J. (2019). Single-pot, solid-phase-enhanced sample preparation for proteomics
734 experiments. Nat. Protoc. *14*, 68–85.

735 Jacob-Dubuisson, F., Mechaly, A., Betton, J.-M., and Antoine, R. (2018). Structural
736 insights into the signalling mechanisms of two-component systems. Nat. Rev.
737 Microbiol. *16*, 585–593.

738 Keseler, I.M., Mackie, A., Santos-Zavaleta, A., Billington, R., Bonavides-Martínez, C.,
739 Caspi, R., Fulcher, C., Gama-Castro, S., Kothari, A., Krummenacker, M., et al. (2017).
740 The EcoCyc database: reflecting new knowledge about Escherichia coli K-12.
741 Nucleic Acids Res. *45*, D543–D550.

742 Konovalova, A., Perlman, D.H., Cowles, C.E., and Silhavy, T.J. (2014).
743 Transmembrane domain of surface-exposed outer membrane lipoprotein RcsF is
744 threaded through the lumen of -barrel proteins. Proc. Natl. Acad. Sci. *111*, E4350–
745 E4358.

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

Kuhn, A., Koch, H.-G., and Dalbey, R.E. (2017). Targeting and Insertion of Membrane Proteins. EcoSal Plus *7*.

Lacqua, A., Wanner, O., Colangelo, T., Martinotti, M.G., and Landini, P. (2006). Emergence of Biofilm-Forming Subpopulations upon Exposure of Escherichia coli to Environmental Bacteriophages. Appl. Environ. Microbiol. *72*, 956–959.

Letain, T.E., and 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.

Létoquart, J., Rodriguez-Alonso, R., Nguyen, V.S., Louis, G., Calabrese, A.N., Radford, S.E., Cho, S.-H., Remaut, H., and Collet, J.-F. (2019). Structural insight into the formation of lipoprotein-β-barrel complexes by the β-barrel assembly machinery. BioRxiv 823146.

Li, G., and Young, K.D. (2012). Isolation and identification of new inner membrane-associated proteins that localize to cell poles in Escherichia coli. Mol. Microbiol. *84*, 276–295.

Li, G., and Young, K.D. (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.

Li, H., Wang, B.-C., Xu, W.-J., Lin, X.-M., and Peng, X.-X. (2008). Identification and Network of Outer Membrane Proteins Regulating Streptomysin Resistance in Escherichia coli. J. Proteome Res. *7*, 4040–4049.

Lim, B., Miyazaki, R., Neher, S., Siegele, D.A., Ito, K., Walter, P., Akiyama, Y., Yura, T., and Gross, C.A. (2013). Heat shock transcription factor σ32 co-opts the signal recognition particle to regulate protein homeostasis in E. coli. PLoS Biol. *11*, e1001735.

Lloubès, R., Cascales, E., Walburger, A., Bouveret, E., Lazdunski, C., Bernadac, A., and Journet, 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.

778 Loos, M.S., Ramakrishnan, R., Vranken, W., Tsirigotaki, A., Tsare, E.-P., Zorzini, V.,
779 Geyter, J. De, Yuan, B., Tsamardinos, I., Klappa, M., et al. (2019). Structural Basis of
780 the Subcellular Topology Landscape of Escherichia coli. Front. Microbiol. *10*, 1670.

781 Luirink, J., Yu, Z., Wagner, S., and de Gier, J.-W. (2012). Biogenesis of inner
782 membrane proteins in Escherichia coli. Biochim. Biophys. Acta *1817*, 965–976.

783 Majdalani, N., Hernandez, D., and Gottesman, S. (2002). Regulation and mode of
784 action of the second small RNA activator of RpoS translation, RprA. Mol. Microbiol.
785 *46*, 813–826.

786 Mateus, A., Bobonis, J., Kurzawa, N., Stein, F., Helm, D., Hevler, J., Typas, A., and
787 Savitski, M.M. (2018). Thermal proteome profiling in bacteria: probing protein state in
788 vivo. Mol. Syst. Biol. *14*, e8242.

789 May, K.L., and Grabowicz, M. (2018). The bacterial outer membrane is an evolving
790 antibiotic barrier. Proc. Natl. Acad. Sci. *115*, 8852–8854.

791 Moussatova, A., Kandt, C., O'Mara, M.L., and Tieleman, D.P. (2008). ATP-binding
792 cassette transporters in Escherichia coli. Biochim. Biophys. Acta - Biomembr. *1778*,
793 1757–1771.

794 Nikaido, H. (2003). Molecular basis of bacterial outer membrane permeability
795 revisited. Microbiol. Mol. Biol. Rev. *67*, 593–656.

796 Orfanoudaki, G., and Economou, A. (2014). Proteome-wide Subcellular Topologies
797 of E. coli Polypeptides Database (STEPdb). Mol. Cell. Proteomics *13*, 3674–3687.

798 Papanastasiou, M., Orfanoudaki, G., Koukaki, M., Kountourakis, N., Sardis, M.F.,
799 Aivaliotis, M., Karamanou, S., and Economou, A. (2013). The Escherichia coli
800 Peripheral Inner Membrane Proteome. Mol. Cell. Proteomics *12*, 599–610.

801 Papanastasiou, M., Orfanoudaki, G., Kountourakis, N., Koukaki, M., Sardis, M.F.,
802 Aivaliotis, M., Tsolis, K.C., Karamanou, S., and Economou, A. (2016). Rapid label-
803 free quantitative analysis of the E. coli BL21(DE3) inner membrane proteome.
804 Proteomics *16*, 85–97.

805 Paradis-Bleau, C., Markovski, M., Uehara, T., Lupoli, T.J., Walker, S., Kahne, D.E.,
806 and Bernhardt, T.G. (2010). Lipoprotein cofactors located in the outer membrane
807 activate bacterial cell wall polymerases. Cell *143*, 1110–1120.

808 Perez-Riverol, Y., Csordas, A., Bai, J., Bernal-Llinares, M., Hewapathirana, S.,

809 Kundu, D.J., Inuganti, A., Griss, J., Mayer, G., Eisenacher, M., et al. (2019). The
810 PRIDE database and related tools and resources in 2019: improving support for
811 quantification data. Nucleic Acids Res. *47*, D442–D450.

812 Perkins, D.N., Pappin, D.J.C., Creasy, D.M., and Cottrell, J.S. (1999). Probability-
813 based protein identification by searching sequence databases using mass
814 spectrometry data. Electrophoresis *20*, 3551–3567.

815 Petiti, M., Serrano, B., Faure, L., Lloubes, R., Mignot, T., and Duché, D. (2019). Tol
816 Energy-Driven Localization of Pal and Anchoring to the Peptidoglycan Promote
817 Outer-Membrane Constriction. J. Mol. Biol. *431*, 3275–3288.

818 R Core Team (2019). R: A Language and Environment for Statistical Computing.

819 Reichel, M., Liao, Y., Rettel, M., Ragan, C., Evers, M., Alleaume, A.-M., Horos, R.,
820 Hentze, M.W., Preiss, T., and Millar, A.A. (2016). In Planta Determination of the
821 mRNA-Binding Proteome of Arabidopsis Etiolated Seedlings. Plant Cell *28*, 2435–
822 2452.

823 Ritchie, M.E., Phipson, B., Wu, D., Hu, Y., Law, C.W., Shi, W., and Smyth, G.K.
824 (2015). limma powers differential expression analyses for RNA-sequencing and
825 microarray studies. Nucleic Acids Res. *43*, e47.

826 Rojas, E.R., Billings, G., Odermatt, P.D., Auer, G.K., Zhu, L., Miguel, A., Chang, F.,
827 Weibel, D.B., Theriot, J.A., and Huang, K.C. (2018). The outer membrane is an
828 essential load-bearing element in Gram-negative bacteria. Nature *559*, 617–621.

829 Selkrig, J., Mosbahi, K., Webb, C.T., Belousoff, M.J., Perry, A.J., Wells, T.J., Morris,
830 F., Leyton, D.L., Totsika, M., Phan, M.-D., et al. (2012). Discovery of an archetypal
831 protein transport system in bacterial outer membranes. Nat. Struct. Mol. Biol. *19*,
832 506–510.

833 Silhavy, T.J., Kahne, D., and Walker, S. (2010). The Bacterial Cell Envelope. Cold
834 Spring Harb. Perspect. Biol. *2*, a000414–a000414.

835 Szewczyk, J., and Collet, J.-F. (2016). The Journey of Lipoproteins Through the Cell:
836 One Birthplace, Multiple Destinations. Adv. Microb. Physiol. *69*, 1–50.

837 The UniProt Consortium (2019). UniProt: a worldwide hub of protein knowledge.
838 Nucleic Acids Res. *47*, D506–D515.

839 Tsolis, K.C., and Economou, A. (2017). Quantitative Proteomics of the E. coli

Membranome. In Methods in Enzymology, (Elsevier Inc.), pp. 15–36.

Typas, A., and Sourjik, V. (2015). Bacterial protein networks: properties and functions. Nat. Rev. Microbiol. *13*, 559–572.

Typas, A., Banzhaf, M., van den Berg van Saparoea, B., Verheul, J., Biboy, J., Nichols, R.J., Zietek, M., Beilharz, K., Kannenberg, K., von Rechenberg, M., et al. (2010). Regulation of peptidoglycan synthesis by outer-membrane proteins. Cell *143*, 1097–1109.

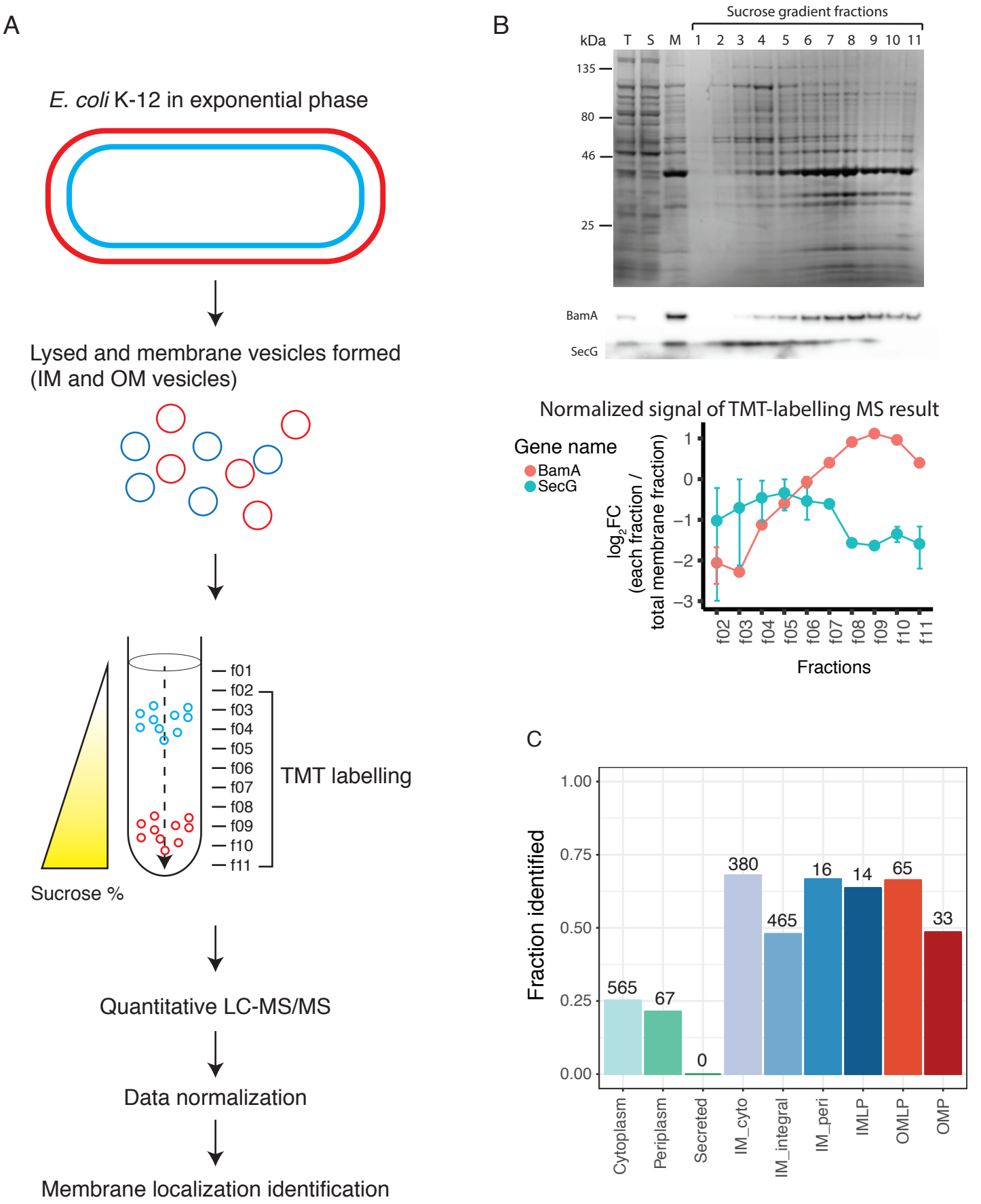
Typas, A., Banzhaf, M., Gross, C.A., and Vollmer, W. (2011). From the regulation of peptidoglycan synthesis to bacterial growth and morphology. Nat. Rev. Microbiol. *10*, 123–136.

Webb, C.T., Selkrig, J., Perry, A.J., Noinaj, N., Buchanan, S.K., and Lithgow, T. (2012). Dynamic association of BAM complex modules includes surface exposure of the lipoprotein BamC. J. Mol. Biol. *422*, 545–555.

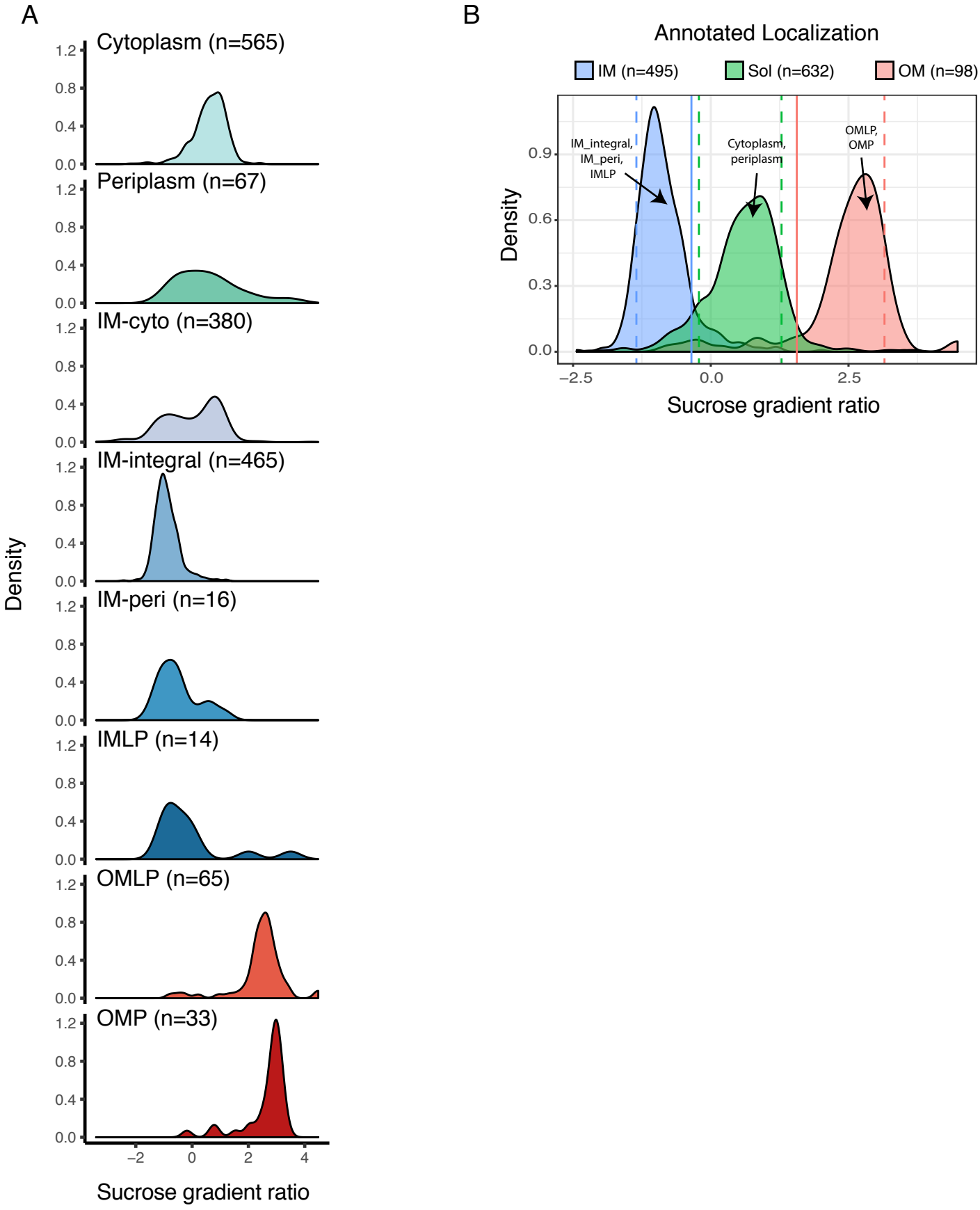
Werner, T., Sweetman, G., Savitski, M.F., Mathieson, T., Bantscheff, M., and Savitski, M.M. (2014). Ion Coalescence of Neutron Encoded TMT 10-Plex Reporter Ions. Anal. Chem. *86*, 3594–3601.

Yang, D.C., Peters, N.T., Parzych, K.R., Uehara, T., Markovski, M., and Bernhardt, T.G. (2011). An ATP-binding cassette transporter-like complex governs cell-wall hydrolysis at the bacterial cytokinetic ring. Proc. Natl. Acad. Sci. *108*, E1052–E1060.

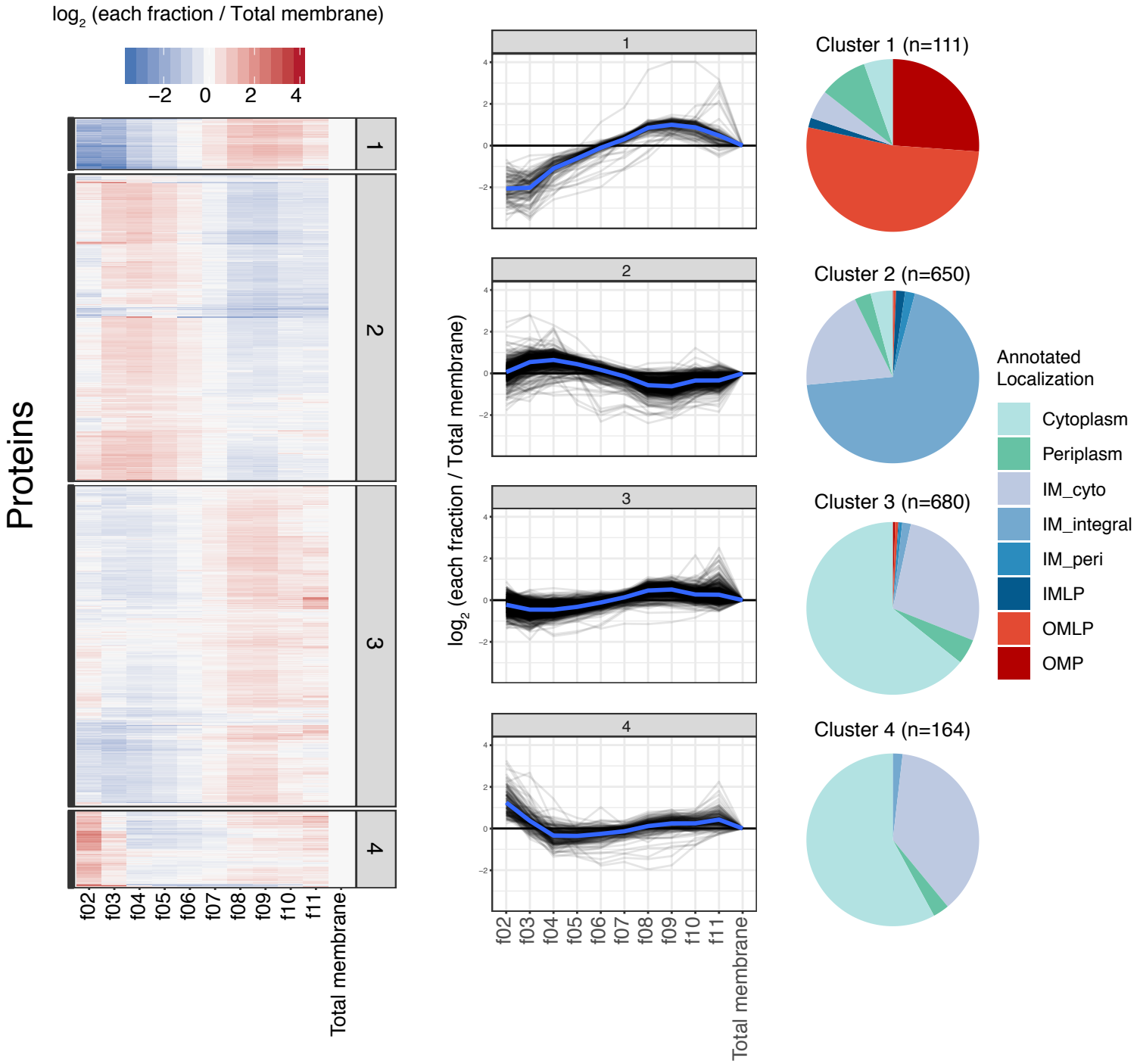
# Figure 1



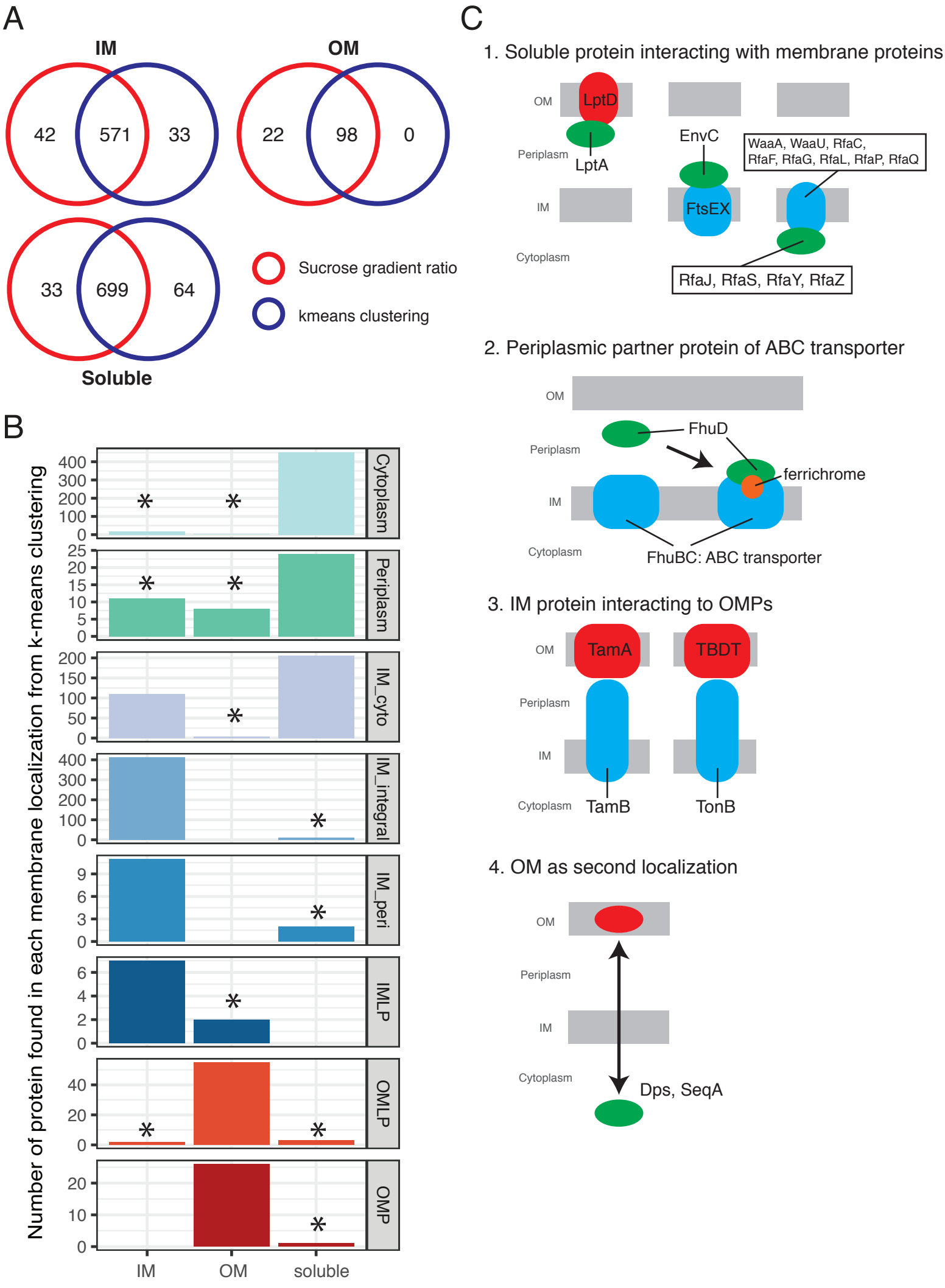
# Figure 2



# Figure 3

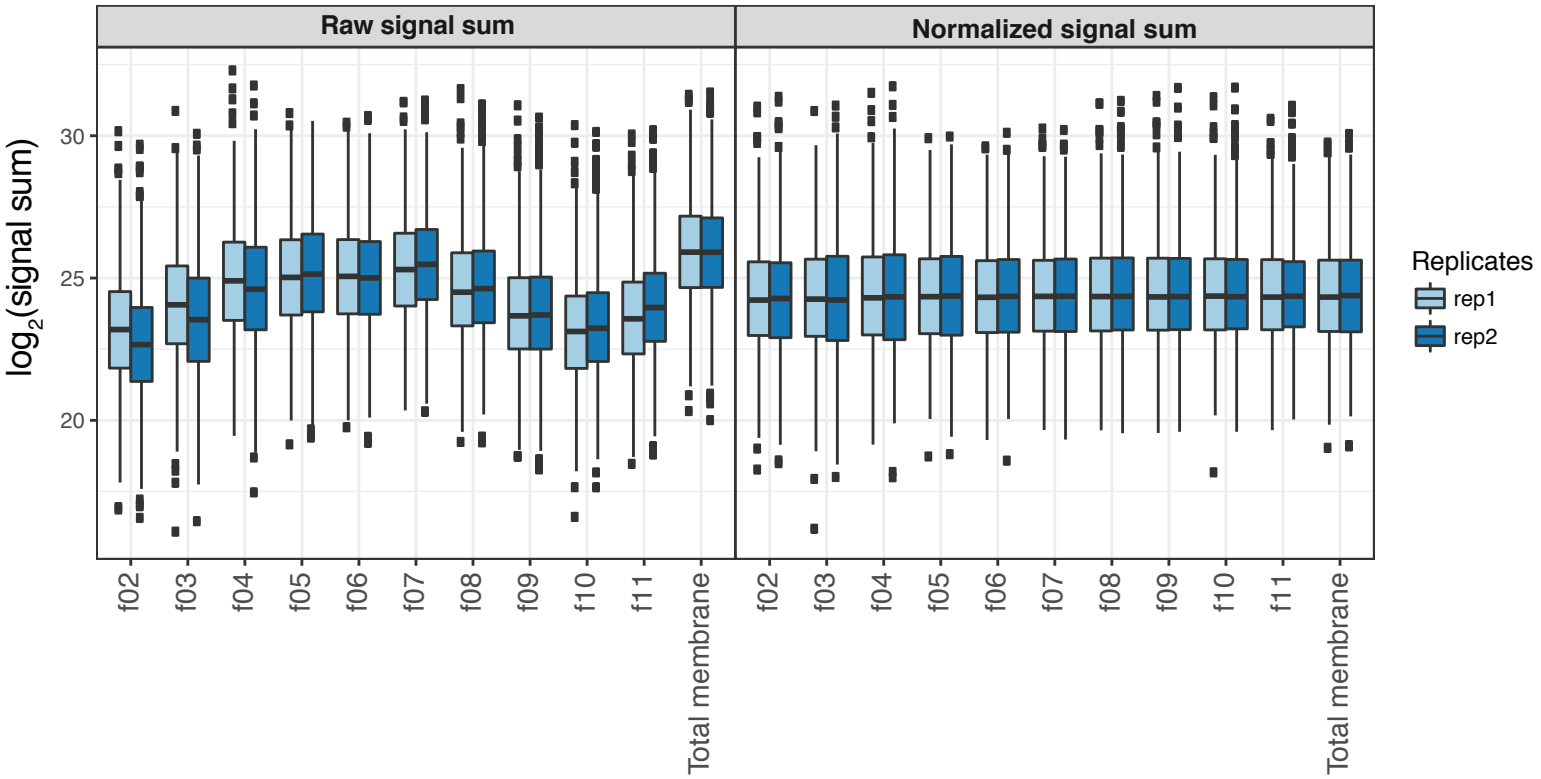


# Figure 4

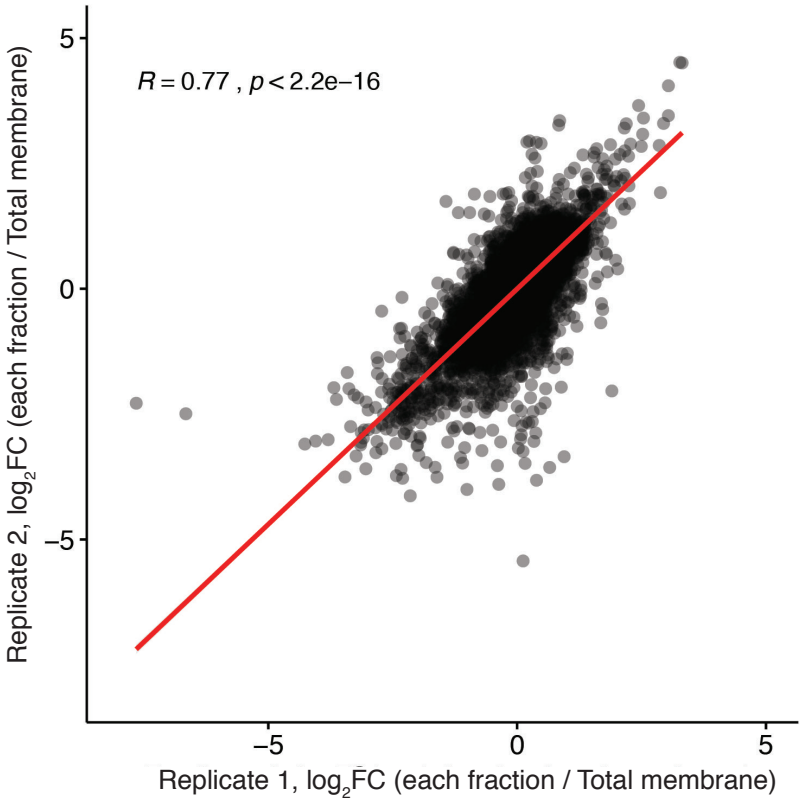


# Figure S1

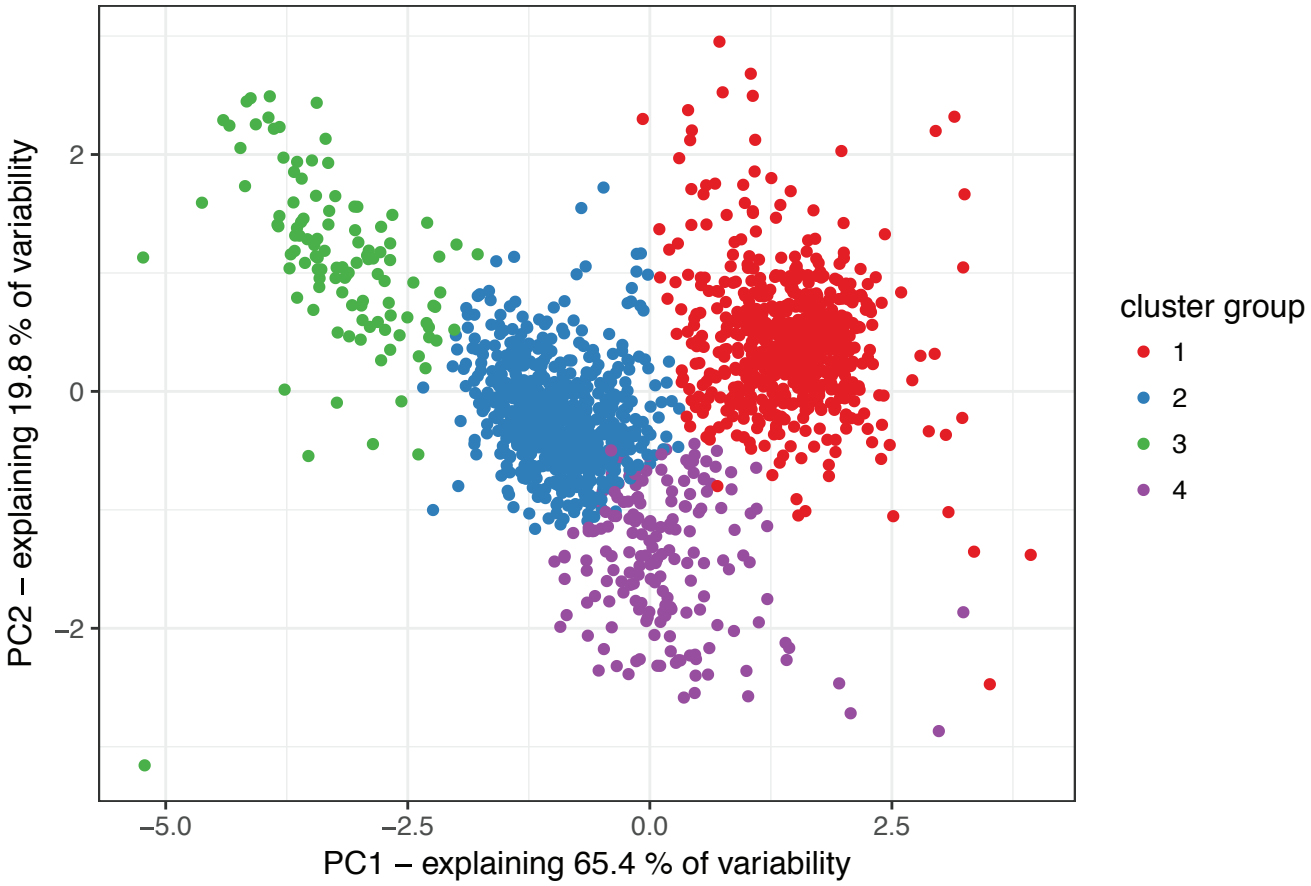
A



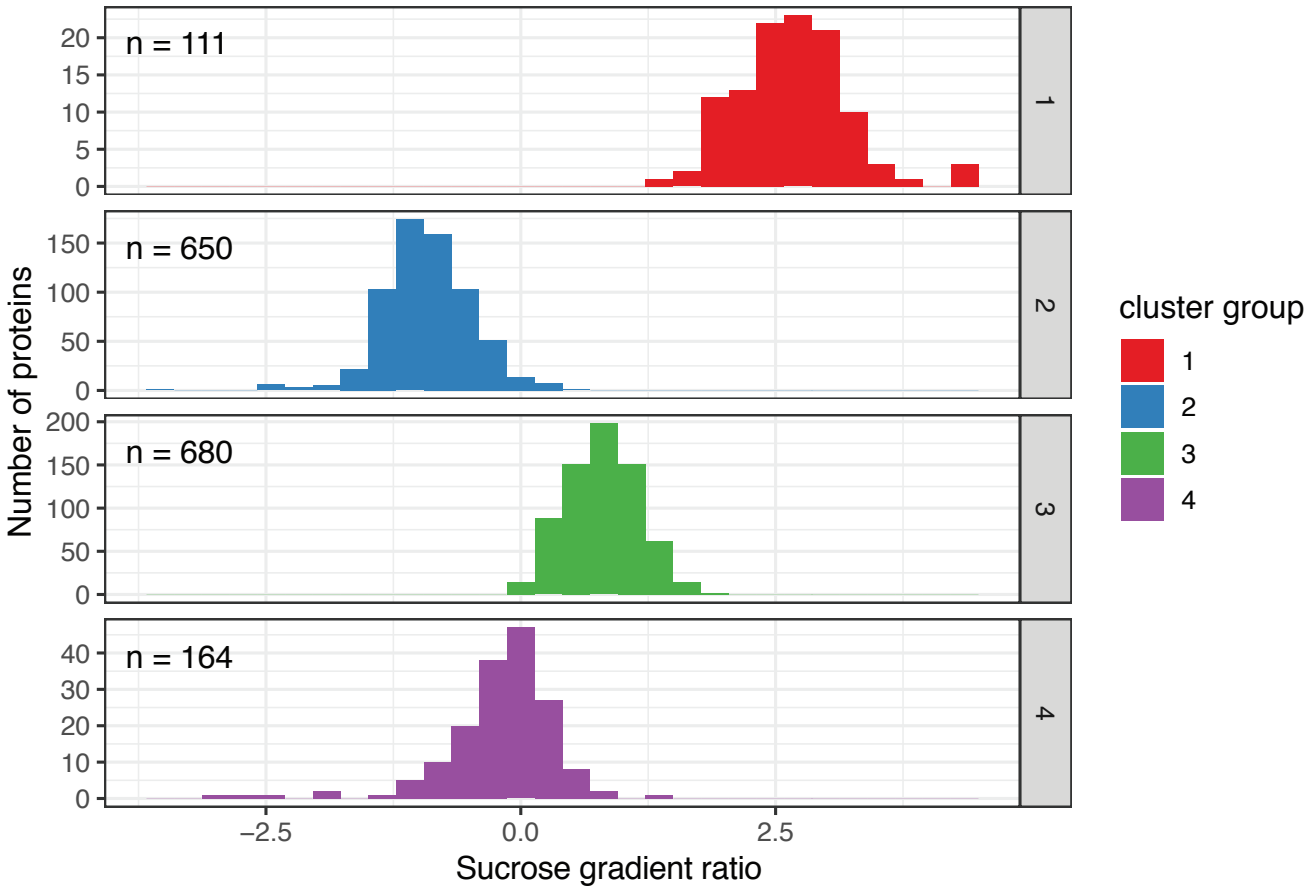

B



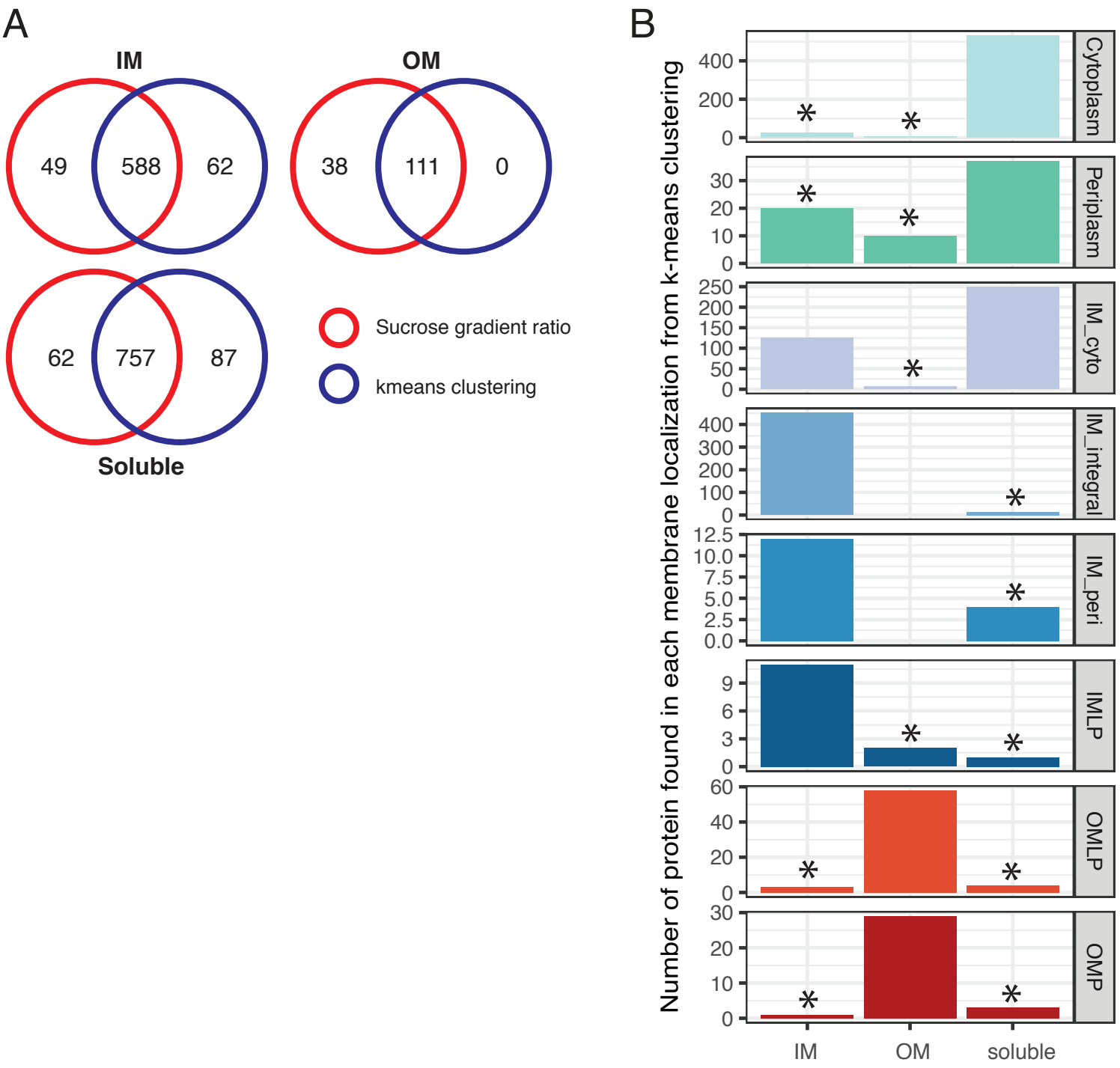
# Figure S2



# Figure S3



# Figure S4



# Figure S5