ACS SyntheticBiology

pubs.acs.org/synthbio Research Article

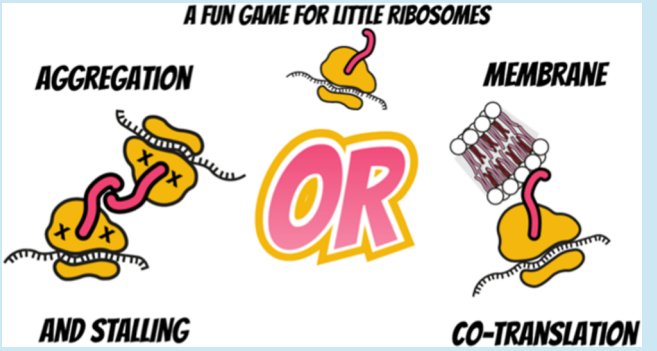
Improving Cell-Free Expression of Model Membrane Proteins by Tuning Ribosome Cotranslational Membrane Association and Nascent Chain Aggregation

Jan Steinkühler,*[,∇] Justin A. Peruzzi,[∇] Antje Krüger, Citlayi G. Villaseñor, Miranda L. Jacobs, Michael C. Jewett, and Neha P. Kamat*

Cite This: *ACS Synth. Biol.* 2024, 13, 129−140 Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Cell-free gene expression (CFE) systems are powerful tools for transcribing and translating genes outside of a living cell. Synthesis of membrane proteins is of particular interest, but their yield in CFE is substantially lower than that for soluble proteins. In this paper, we study the CFE of membrane proteins and develop a quantitative kinetic model. We identify that ribosome stalling during the translation of membrane proteins is a strong predictor of membrane protein synthesis due to aggregation between the ribosome nascent chains. Synthesis can be improved by the addition of lipid membranes, which incorporate protein nascent chains and, therefore, kinetically compete with aggregation. We show that the balance between peptide-membrane association and peptide aggregation rates determines the yield of the synthesized membrane protein. We define a membrane protein expression score that can be used to rationalize the engineering of lipid composition and the N-terminal domain of a native and computationally designed membrane proteins produced through CFE.



KEYWORDS: *cell-free gene expression, membrane proteins, cotranslational synthesis, PURE, biophysics, synthetic biology*

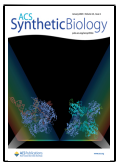
INTRODUCTION

Cell-free gene expression (CFE) systems leverage cellular machinery to transcribe and translate genes outside of a living cell.[1,2] Over the last two decades, CFE systems have grown from a molecular biology tool to a powerful, shelf-stable, and scalable biomanufacturing platform.[3−9] CFE systems have now been used to create wearable biosensors,[10,11] synthetic cells,[12−14] prototype metabolic pathways,[15,16] rapidly screen drug candidates,[17−19] and produce vaccines at the point of care.[20,21] Thus, efforts to expand the capabilities of CFE systems could have a large impact on sustainable biomanufacturing, point-of-use biosensing, and therapeutic production.

One area that has posed a challenge with CFE systems has been the robust expression of membrane proteins. This is because membrane proteins require amphiphilic scaffolds to integrate into, similar to their synthesis in living cells. Because membrane proteins perform critical cellular functions in sensing, signaling, and energy regeneration, their inclusion in CFE systems is critical to expanding the sensing and biomanufacturing capabilities of CFE systems.

To address this need, membranes and membrane mimetics have been included in CFE systems to integrate and improve the expression of membrane proteins.[22,23] Inverted vesicles, formed from cellular membranes during extract preparation, have been used to retain membrane-associated functionality in CFE systems.[6,24] However, the production of native vesicles requires overexpression of membrane components prior to lysis, is limited by the challenges associated with heterologous membrane protein production, and, furthermore, does not allow for tuning of membrane biophysical features, which may affect the final activity of an expressed membrane protein.[25] The ability to directly express membrane proteins into a membrane mimetic in a CFE system could circumvent these challenges. Synthetic membranes, in the form of liposomes and nanodiscs, have been used in CFE systems to improve the expression of membrane proteins.[25−28] Experiments have established that membrane composition, available membrane area, and formation of cotranslational membrane-bound ribosome complexes are crucial for successful CFE of

Received: June 11, 2023
Revised: November 29, 2023
Accepted: December 5, 2023
Published: December 27, 2023



https://doi.org/10.1021/acssynbio.3c00357
ACS Synth. Biol. 2024, 13, 129−140

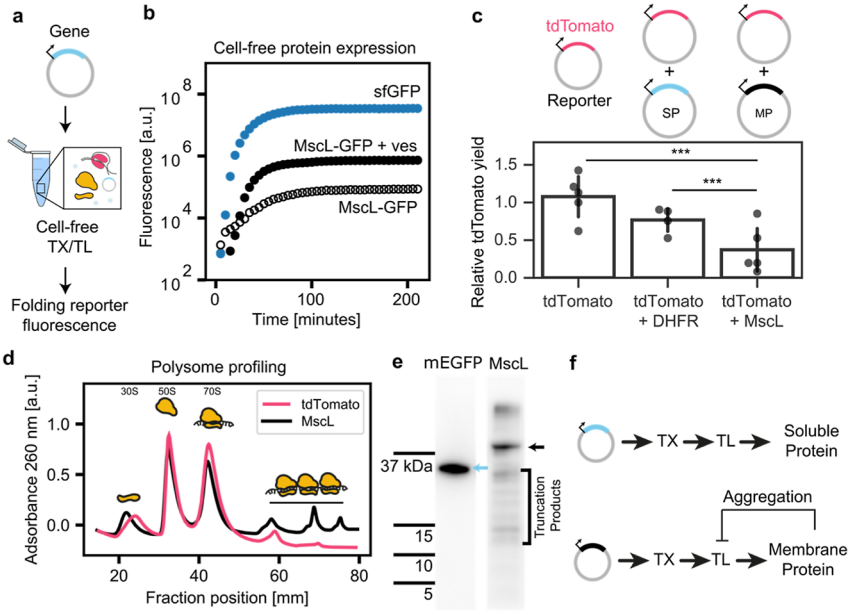


Figure 1. Expression of the membrane protein quenches a CFE reaction. (a) Schematic of the cell-free system and fluorescence assay for protein expression. (b) Representative results for protein expression upon addition of plasmids coding soluble sfGFP or the mechanosensitive channel of large conductance-enhanced GFP fusion (MscL-GFP) with (+ves) or without addition of 100 mM DOPC liposomes. (c) Quenching of CFE is monitored by coexpressing fluorescent tdTomato with a soluble nonfluorescent dihydrofolate reductase (DHFR) (soluble protein, SP) or a nonfluorescent MscL (membrane protein, MP), respectively, and normalized by maximum fluorescence of tdTomato alone. (d) Polysome formation was detected (indicated by black horizontal line) using a sucrose density gradient light (260 nm) adsorption profile after expression of tdTomato or MscL (no liposomes added). (e) Western blot against the N terminus of soluble mEGFP and MscL-GFP demonstrates that more truncation products are produced during the expression of MscL-GFP. Black arrow indicates full length MscL-GFP. (f) Proposed feedback loop that quenches translation by ribosome stalling/aggregation when expressing a membrane protein.

membrane proteins.[28−30] However, each of the properties must be tuned to effectively produce properly folded, functional membrane proteins, limiting the ready adoption of membrane proteins into CFE systems.[25,28,30] The ability to predict optimal reaction conditions for the CFE of membrane proteins could enable efficient membrane protein expression and consequently the rapid expansion of membrane functionality within cell-free systems.

To systematically improve CFE, insight from mechanistic models has proven to be useful for soluble protein expression. Recently, coarse-grained and multiparameter models were used to quantitatively describe CFE systems, including sequence-specific predictions of transcription and translation (TX/TL) kinetics.[31−38] However, to the best of our knowledge, a similar quantitative model for cell-free membrane protein synthesis does not exist. In this paper, we (i) compare the CFE of a soluble protein and membrane protein and (ii) develop a quantitative model to describe cell-free membrane protein synthesis. We then (iii) apply this model to improve the expression of a native and a computationally designed membrane protein by up to 50%.

RESULTS

Expression of Membrane Proteins Reduces the Capacity of CFE Systems to Produce Proteins. As a first step toward deconstructing the expression of membrane proteins in cell-free systems, we quantified CFE activity when expressing a well-folding and soluble model protein, super folder green fluorescent protein (sfGFP), and a model membrane protein, mechanosensitive channel of large conductance (MscL) fused to monomeric enhanced green fluorescent protein (MscL-GFP).[28] We produced proteins from plasmids encoding these proteins using the PURE system, which is based on 31 purified macromolecular components of the cellular transcription and translation machinery plus material and energy resources.[39] Fluorescent protein expression and folding can be monitored by measuring the resulting fluorescent signal as the cell-free reaction proceeds over time (Figure 1a). Previously, we have shown that MscL-GFP fluorescence is positively correlated with full-length protein yield as determined by Western blot densitometry.[28] Apart from the overall fluorescence signal, two other quantities can be extracted from such experiments: the maximum synthesis rate and the lifetime of the reaction. We compared these quantities for soluble protein, sfGFP, and membrane protein, MscL-GFP (blue and open black points Figure 1b). We found that MscL-GFP yield, synthesis rate, and reaction time scale were signficantly lower than those of soluble sfGFP. Addition of lipid material in the form of 100 nm liposomes composed of 1,2-dioleoyl-*sn*-glycero-3-phosphocholine (DOPC) or Brij-58 detergent enhanced some of the overall yield and synthesis rate (filled black points Figures 1b and S1). However, even in the presence of lipid vesicles or detergent, the reactions never came close to the performance of the cell-free system expressing soluble protein. Such large differences cannot be explained by differences in protein molecular weight (MW) or differences in the GFP variants alone. Instead, it appeared that the expression of MscL quenched the CFE reaction and therefore did not use the available resources efficiently. We wanted to understand this effect in more detail.

To investigate the extent to which membrane protein expression inhibits CFE reactions, we monitored the expression of a soluble, fluorescent protein, tdTomato, in the presence of a coexpressed soluble or membrane protein

https://doi.org/10.1021/acssynbio.3c00357
ACS Synth. Biol. 2024, 13, 129−140

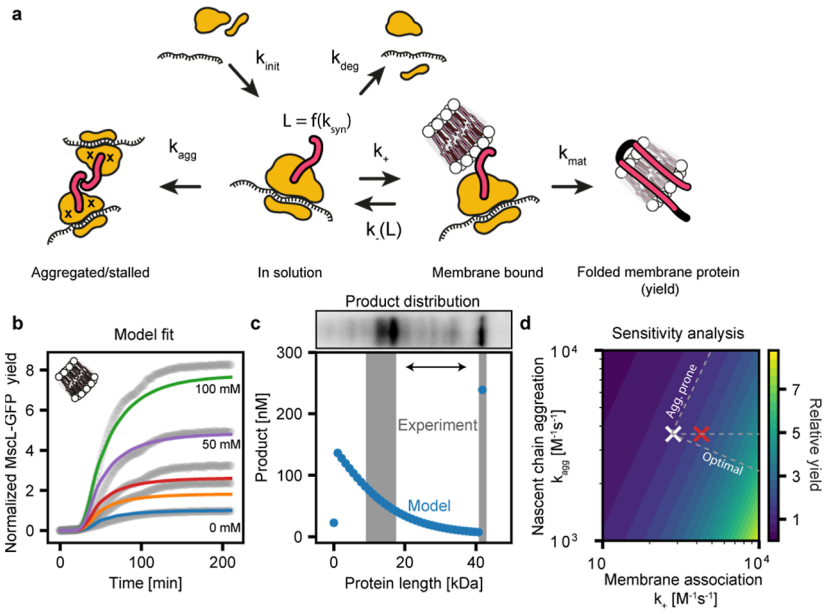


Figure 2. Balance of association of membrane protein nascent peptides with membranes or other ribosome-bound peptides determines the extent of quenching of a cell-free reaction. (a) Overview of the kinetic model showing rate constants for initiation of translation k_{init}, finite reaction lifetime due to resource depletion k_{deg}, aggregation of nascent chains complex k_{agg}, membrane association rate constants k_+ and $k_-(L)$, nascent chain length L that grows with rates k_{syn} and k_{mat}, which are the fluorescence reporter synthesis and maturation rates, respectively. (b) Fit of the model to the data points from Jacobs et al.[28] for MscL-GFP CFE with varying concentrations of liposome lipids (experimental data points are shown in gray, and the best fit model is represented by colored lines) normalized to zero lipid concentration. The experiments measure GFP fluorescence, while the model calculates the amount of full-length GFP; both quantities are referred to as "yield". See the main text for details. (c) (Top) Western blot showing the distribution of protein products at the final time point of a CFE reaction expressing MscL-GFP in the presence of 10 mM DOPC liposomes. MW of aggregation and full-length bands approximately corresponds to the MW defined by the *x*-axis below. Full-length MscL-GFP (46 kDa) is produced in addition to smaller truncation products. (Bottom) Model-derived truncation products are represented in blue, while the gray columns indicate the regions of the gel with the highest detected protein density. Both experimental and model-derived data exhibit a gap in truncation product bands in the 20−40 kDa range, similar to the experimental result (Western blot on top; arrow indicates the protein size range where fewer protein products are detected). (d) Color map of full-length protein yield normalized to best fit values of k_+ and k_{agg} (white cross). Increase of k_+ by hybrid membranes (red cross) and possible trajectories (dashed white lines) result in an increase of protein yield ("optimal" and horizontal line) or no increase ("aggregation prone").

(Figure 1c). Specifically, we compared the expression of fluorescent tdTomato (MW 54 kDa) with either soluble nonfluorescent DHFR (MW 24 kDa) or nonfluorescent MscL (MW 14 kDa) to tdTomato alone. In all three experiments, there were no lipids added to the reaction. We found that coexpression of the soluble DHFR protein only slightly decreased tdTomato yields, signifying minimal competition between the two T7 promoters under the given reaction conditions. In contrast, coexpression of MscL decreased tdTomato yields by about 70% relative to tdTomato expression alone. The capacity of membrane protein coexpression to significantly reduce the expression of a soluble protein suggests that the expression of membrane proteins quenches CFE activity.

We wondered if the observed quenched CFE activity in the presence of coexpressed MscL was caused by the aggregation of misfolded MscL peptides. We reasoned that MscL, which has large hydrophobic peptide segments, should demonstrate a higher propensity for misfolding in the absence of a lipid membrane to insert and fold into. As such, MscL peptides might aggregate during translation, stall the ribosome, and reduce the pool of ribosomes available for translation relative to that of soluble proteins. Accordingly, we measured ribosome aggregation and peptide fragmentation in our cell-free reactions as a function of the expression of MscL or tdTomato. To measure ribosome aggregation, cell-free reactions were quenched on ice and transferred to a sucrose gradient. After ultracentrifugation, differently sized ribosome complexes sediment along the gradient. The presence of RNA material along the sucrose gradient was then measured by light adsorption. When soluble tdTomato was expressed, we obtained the characteristic peaks of 30S, 50S, and 70S, corresponding to the small and large ribosomal subunits and assembled 70S ribosomes (magenta trace, Figure 1d).[40,41] When MscL was expressed, we observed additional peaks that are assigned to polysomes, i.e., multiple ribosomes stalled along an RNA strand (black trace, Figure 1d). Previously, polysome formation was shown to be enhanced by attractive interactions between nascent chain complexes.[42] Because large attractive interactions should be present between the hydrophobic membrane protein residues, the increased interaction of membrane protein nascent chains would be expected to increase polysome formation, as we observed. In addition to polysome formation, stalling of ribosome complexes should give rise to incomplete protein products. Indeed, by probing the N-terminus of soluble GFP and MscL-GFP via Western blot, we observed the presence of truncated protein products only when the membrane protein, MscL was expressed, in contrast to when GFP was expressed (Figures 1e and S2). In summary, our results suggest a negative feedback loop by which expression of a membrane protein stalls or quenches ribosome activity due to aggregation of the ribosome-bound nascent chains in the PURE cell-free system (Figure 1f).

https://doi.org/10.1021/acssynbio.3c00357
ACS Synth. Biol. 2024, 13, 129−140

Table 1. Overview of Model Parameters and Literature Values Obtained from Gonzales et al.[34a]

literature values			fitted values				
k_{syn} [s^{-1}]	k_{deg} [10^{-6} s^{-1}]	[R] [10^{-6} M]	k_{init} [10^6 M^{-1} s^{-1}]	k_{RNA} [10−9 M s^{-1}]	k_+ [10^3 M^{-1} s^{-1}]	k_- [s^{-1}]	k_{agg} [10^3 M^{-1} s^{-1}]
0.33	385	2.4	5 ± 3	0.26 ± 0.04	2.8 ± 0.2	0.8 ± 0.1	3.6 ± 0.4

[a]Standard deviation for fitted values is indicated by ±, and sample size is n = 7.

Kinetic Model Reveals Balance between Membrane Association and Aggregation. We next developed a model to describe different ribosome states during protein expression. We coarse-grained the cell-free synthesis into a series of reactions (see Methods). Specifically, we modeled the synthesis of mRNA as a first-order reaction with the rate constant k_{RNA}. Translation initiation occurs by a second-order reaction between the free ribosome and mRNA with a rate k_{init}. Ribosomes are deactivated with a rate constant k_{deg}, a parameter which captures resource and energy depletion of the PURE system.[33,34] This degradation is what led to the finite activity of the PURE system in previous models (the plateau in Figure 1b). The protein is synthesized with the previously determined rate constant k_{syn}, which leads to a growing nascent chain until the protein is fully synthesized and the ribosome is released from the mRNA. Finally, we modeled the effect of a blocked ribosomal binding site between individual ribosomes by not allowing binding of a ribosome to an occupied mRNA initiation site.

Compared to previous studies, which considered only ribosomes in solution, our model takes two additional states into account. First, we included a bound complex between the ribosome's nascent chain and the membrane, which forms between the translated peptide and the membrane surface with binding/unbinding rate constants k_+ and k_-. In our experiments, the membrane surface was supplied by liposomes, but the model applies in principle to any membrane surface, e.g., lipid nanodiscs or supported lipid bilayers. Second, we considered an aggregated, dysfunctional state, which removes ribosomes from the system with a second-order rate constant k_{agg} due to encounters between RNA-bound ribosomes. In the latter case of aggregation, truncated protein products with lengths of the aggerated nascent chain are produced. We assumed that the binding of the translated peptide to the membrane is reversible and is a function of the length of the synthesized protein. Specifically, we sought to capture the fact that a short, N-terminal segment of a partially synthesized membrane protein will exhibit a lower membrane affinity than a longer translated transmembrane segment of the same membrane protein. To accomplish this effect, we consider k_+ (binding) as constant and k_- (unbinding) decaying exponentially with protein length L, an assumption we will revisit later. Fully synthesized protein folds at a rate k_{mat}, corresponding to the GFP maturation rate, that we can compare to experimental values. This model was formalized as a set of elementary reactions (see Methods).

We proceeded to fit our model to experiments by comparing model-derived results to the fluorescence of the MscL-GFP folding reporter expressed in the presence of vesicles at different lipid concentrations. In our model, the vesicle concentration enters in the binding rate of ribosome-nascent chain complexes, M, to the membrane as $\frac{dM}{dt} = k_+[V][R] - k_-(L)$, where V represents the concentration of vesicles and R is the free ribosome concentration. V is calculated from the vesicle size and total lipid concentration used in experiments. The experiment reports the fluorescence of expressed MscL-GFP normalized to a baseline of McsL-GFP expression with a zero lipid concentration, V = 0. To fit the experimental data, we also normalized the model output to GFP fluorescence at zero lipid concentration and assumed that all fully synthesized MscL-GFP molecules are fluorescent after the maturation time k_{mat}. Considering the model simplifications and typical experimental error, the fit is satisfactory (Figure 2b). To validate our results, we obtained independent repeats of the experimental data to which we fitted our model. In the repeat experiment, we additionally calibrated the GFP fluorescence to absolute concentration values. We find quantitative agreement between the model predictions (parametrized on the older data set) and the repeat experiment, up to a constant error on the order of 200 nM MscL-GFP concentration (Figure S3).

We further established the validity of our model by comparing the calculated fitting parameters (Table 1) to the literature values. Our model's initiation rate lies between two previous estimates of 3×10^3 M^{-1} s^{-1} for the PURExpress system by Doerr et al. and 175×10^6 M^{-1} s^{-1} for in vitro translation in optimized buffer conditions by Rudorf et al.[35,43] Similarly, our model's first-order transcription rate of a membrane protein, k_{RNA}, is within error of the initial transcription rate in the PURE system (0.24×10^{-9} M^{-1} s^{-1}) obtained by measurements of RNA concentration in the PURE system by Gonzales et al.[34] (Supporting Information Note 1). In contrast, our model's membrane association rate constant k_+ is one order of magnitude smaller than what was found using single-molecule experiments between a peptide (GLP-1) and lipid membrane, with a binding rate of 1.0×10^4 M^{-1} s^{-1}.[44] This difference might be due to differences in the peptide sequence between GLP1 and MscL. Additionally, binding of the nascent chain to the membrane requires correct alignment between a ribosomal exit tunnel and a membrane, which reduces the binding rate compared to the free peptide GLP-1. While ribosomal association should impact membrane association, it would not be expected to contribute to the unbinding of the nascent chain. Indeed, the unbinding rate k_- aligns better with the experimental estimate for GLP-1 of 0.8 s^{-1}. Combined, the fit of our model to previously generated data and its consistency with those found in the literature demonstrate that our model adequately describes the synthesis of MscL in the PURE system in relation to varying the lipid concentration.

Next, we investigated the simulation trajectories in more detail. In the model, unbinding of ribosome-nascent chain complexes from the membrane becomes exponentially less likely when the nascent chain sequences are longer than 10 amino acids. Thus, once the nascent chain is about 20 amino acids long, the ribosome-nascent chain complex does not unbind during the CFE lifetime. This means that ribosome-nascent chain complexes either aggregate early or integrate into a membrane cotranslationally, where they are protected from aggregation. Competition between aggregation and integration results in a multimodal distribution of peptide products with a noticeable reduction or absence of products between

https://doi.org/10.1021/acssynbio.3c00357
ACS Synth. Biol. 2024, 13, 129−140

unfinished, aggregated products of 10−20 kDa and full-length protein of 40 kDa, both in the model calculations (Figure 2c) and experiments (blot insert Figure 2c). To further validate that our model captures the difference between full-length, correctly inserted MscL and aggregated product, we conducted sucrose flotation assays after the cell-free reaction has finished. The experiments quantify the amount of MscL-GFP that is aggregated in the pellet and that remains in the soluble fraction in the liposomes (Supporting Information). Here, the experimental value of the GFP fluorescence ratio in top to bottom fractions is 1.06 ± 0.3 (n = 3) and compares very favorably to the simulation result of 1.07, showing that about half of the product is not inserted into liposomes at this lipid concentration. At higher lipid concentrations, the model then predicts that almost all product is found in the liposome fraction. In addition, for higher lipid concentrations, more total product is predicted because less aggregation leads to reduced ribosome stalling and in this way better CFE productivity. Taken together, this data indicates that the fate of each ribosome is determined early, when the synthesized protein is still rather short, by its N-terminal domain binding affinity to the membrane and its aggregation propensity.

To gain quantitative insight into this effect, we varied the rate constants k_+ and k_{agg} in our simulation model (Figure 2d). By calculating the yield of fully synthesized protein relative to the pair of the experimentally determined rates (k_+, k_{agg}), we found that excess aggregation quickly diminishes the yield of the full-length product, while a higher membrane association rate constant increases that yield (changes relative to the white cross in Figure 2d). These two rate constants (k_+ and k_{agg}) can be tuned by changing the molecular components within the CFE reaction, such as the properties of the membrane and protein, allowing for an increased protein yield. For example, changes in membrane composition might change k_+, but they should keep k_{agg} constant (horizontal line in Figure 2d). Recently, we used coarse-grained simulations to show that hybrid polymer/lipid membranes can enhance peptide insertion rates by a factor of 1.5 relative to pure lipid membranes through a generic mechanism based on the generation of membrane packing defects.[45] Through this effect, the association of the nascent-chain ribosome complex to the membrane will be enhanced by the same factor. By increasing k_+ in our model by 1.5, we predicted an increase in MscL yield (red cross Figure 2d). The calculated 27% yield increase is in exact agreement with the experimental result of 28 ± 3% improvement in MscL expression using the same hybrid polymer/lipid membranes. This agreement between the models and experimental data demonstrates the predictive power of our modeling approach, as the polymer/lipid data was not used in the parametrization of the kinetic model.

Apart from changing the membrane composition, the protein sequence could also be altered to increase the membrane protein yield. Our analysis predicts that an N-terminal domain sequence that both lowers aggregation propensity and increases membrane association would increase protein yield very effectively ("optimal" trajectory in Figure 2d). However, sequences with high membrane affinity are often also prone to aggregation. Thus, there is a sequence space that will increase membrane association but increase aggregation propensity to an even greater extent, leading to a constant or even reduction of yield ("aggregation-prone" trajectory in Figure 2d). Taken together, our modeling results suggest the need to balance the membrane affinity and aggregation propensity of a membrane protein N-terminal domain to optimize membrane protein yield for a given membrane protein. Further, protein yield should be systematically enhanced by optimizing N-terminal domain membrane affinity.

Diverse Bacterial N-Terminal Domain Sequences Balance Aggregation and Membrane Association. We wondered if we could design N-terminal peptide sequences in a way that promotes affinity of the peptide for synthetic membranes used in our study while limiting excessive nacent peptide−peptide aggregation. For inspiration of suitable N-terminal sequences, we investigated naturally occurring sequences that have a strong selection pressure against aggregation, which is generally toxic to cells. Initially, we again focused on MscL, which inserts cotranslationally in bacteria, without assistance of the Sec translocon,[46,47] similar to the CFE system studied here. We hypothesized that as different bacterial species have large differences in their membrane composition, there might be a species with membrane compositions that reflects the properties of simple synthetic membranes used in our cell-free reactions. Thus, by fusing a protein domain, which has evolved for optimal folding in membranes similar to our membrane mimetic, to the N-terminus of our protein, we might increase the CFE protein yields. To identify such sequences, we calculated partitioning free energy ΔG_{wm} values from water to a synthetic lipid bilayer interface for the first five residues from the N-termini from a diverse set of bacterial species identified by a consensus motif search (see Methods). Interestingly, the N-terminal helix of MscL is widely conserved between bacterial species, further promoting our investigation into the N-terminal domain sequence space (see Methods and ref 48). The ΔG_{wm} values correspond to partitioning free energies of peptide sequences from water to a zwitterionic phospholipid membrane interface, similar to the DOPC membranes added to the cell-free system.[49] We found a wide range of ΔG_{wm} values between species. Notably, the native *E. coli* N-terminal domain sequence did not exhibit the highest membrane association energy. For our experiments, we chose one sequence with comparable, larger, and smaller ΔG_{wm} values (Figure 3a−c). Additionally, we considered the N-terminal domain of another model membrane protein, which expresses well in PURE cell-free reactions, LacY.[50,51] Finally, we also investigated an artificial polyleucine sequence (LLLL), which would have the most favorable ΔG_{wm} for membrane association. All five sequences were used to construct MscL chimeras between the different N-termini and the remaining *E. coli* MscL-GFP sequence. We then measured the cell-free reaction protein yields by GFP fluorescence. Previous models of CFE have considered sequence-specific changes in the translation initiation rate for soluble protein synthesis. We checked for sequence-specific changes of the ribosome-RNA binding free energy for our chimera constructs and found no correlation to the measured GFP fluorescence (Figure S4). Instead, we found a strong correlation (Pearson's $r = -0.9$, $p < 0.05$) with more favorable membrane association ΔG_{wm} values, suggesting that aggregation, and not translation initiation, is the limiting factor for cell-free membrane protein synthesis in the conditions studied here. Interestingly, the synthetic polylysine construct LLLL fell outside of this correlation (yellow star in Figure 3c), which is known to be very aggregation prone in solution.[52] The limited yield of protein expression observed for LLLL suggests that this peptide promotes protein−protein aggregation,

https://doi.org/10.1021/acssynbio.3c00357
ACS Synth. Biol. 2024, 13, 129−140

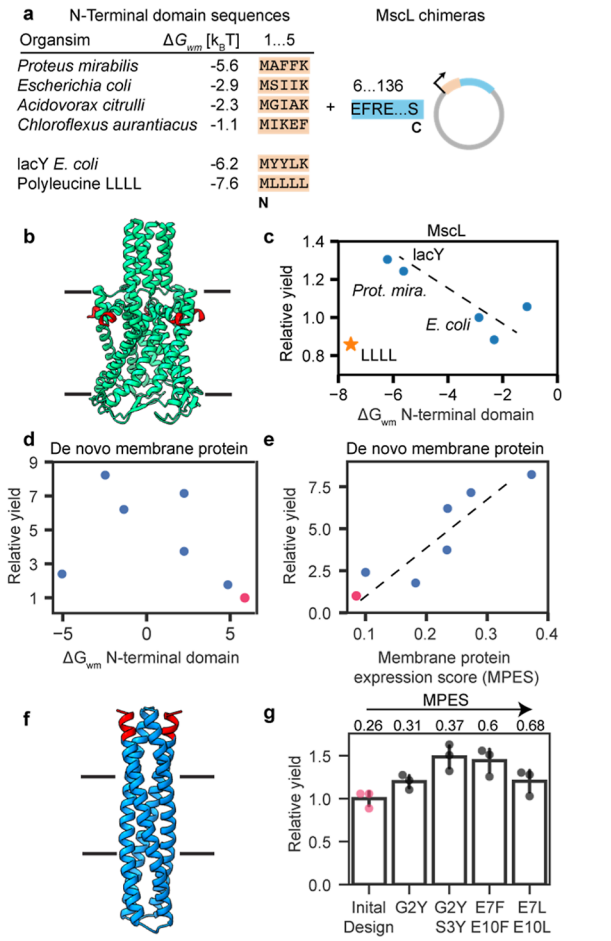


Figure 3. Rational engineering of membrane protein N-terminal domains for the optimization of CFE. (a) Table of selected N-terminal domain sequences with ΔG_{wm} (peptide partitioning free energy from water to lipid bilayer interface) values. The varying N-terminal domains (orange color) were fused to the N-terminus of MscL-GFP (*E. coli*, blue color) to obtain six different MscL-GFP chimeras. (b) Cartoon of MscL inserted in the lipid membrane (black lines). Red indicates the N-terminal domain segment. (c) Resulting GFP fluorescence normalized to *E. coli* MscL (relative yield) of the six MscL-GFP chimeras plotted against the calculated ΔG_{wm} values. (d) Calculated ΔG_{wm} values vs GFP expression determined by fluorescence normalized to the worst expressing de novo-designed protein (red dots) show only weak correlation. (e) Same relative GFP expression data as in panel (d) plotted against the membrane protein expression score (MPES) defined in the main text show a strong correlation (Pearson's $r = -0.9$, $p < 0.05$). (f) Cartoon of de novo protein inserted in the lipid membrane (black lines). Red indicates N-terminal domain segment. (g) Improvement of de novo membrane protein yield by mutations that increase the MPES. MPES score is normalized between 0 and 1, where 1 indicates highest membrane affinity with lowest aggregation propensity. Numbers above bars show calculated MPES score. Each data point represents an independent experiment.

offsetting the large affinity to membrane association. In summary, our results with the cell-derived peptide sequences suggest that aggregation is the limiting factor for the synthesis of membrane proteins in cell-free reactions.

Restoring the Balance Between Membrane Association and Aggregation in N-Termini of De Novo Designed Proteins. The excess aggregation of the synthetic polylysine construct made us curious to see if our insights could be used to improve the expression of de novo membrane proteins. In a previous study, seven different de novo proteins with varying transmembrane domains were synthesized in the PURE system.[53] The design process resulted in seven transmembrane proteins with not only varying transmembrane lengths but also different N-terminal sequences, giving us the opportunity to test the effect of balancing aggregation and membrane affinity. Expression and protein yield, as measured by GFP fluorescence, between the seven designs varied about 8-fold (Figure 3d). Comparison between relative expression and ΔG_{wm} values reveals only a weak correlation, further strengthening our assumption that synthetic N-terminal domain sequences lack the necessary balance between membrane association and aggregation propensity. To assess aggregation propensity, we considered the CamSol solubility score, which is based on a phenomenological amino acid aggregation scale.[54] We scale both ΔG_{wm} and the CamSol solubility score between 1 and 10, where 1 is the largest ΔG_{wm} value and the lowest CamSol score. This analysis gives two values, s_1 (ΔG_{wm}) and s_2 (CamSol), between 1 and 10. We define the MPES as $s_1 \cdot s_2/10$, where 10 would predict the best expression. The correlation between measured expression and score is highly significant (Pearson's $r = 0.9$, $p < 0.01$), showing that synthetic sequences need to consider both aggregation propensity and membrane affinity (Figure 3e). Similarly, the MPES correctly predicted the polyleucine sequence scoring lowest, while the best expression construct scored highest among the MscL chimeras (Figure S5). Motivated by these results, we asked if we could add mutations that increase MPES to improve the expression of otherwise low expressing membrane proteins. We considered the worst-expressing construct, the 20 Å thick transmembrane protein, and generated a single-point mutant G2Y and double mutants (G2Y, S3Y), (E7F, E10F), and (E7L, E10L), which increase MPES (Figure 3f,g). As predicted, all five constructs improved expression, with the best expression by mutants (G2Y, S3Y), which improved expression yield by approximately 50%. Together, these results demonstrate that CFE of membrane proteins can be improved by systematically changing the N-terminal domain sequence and that the MPES score provides a metric to guide sequence design.

DISCUSSION

Here, we have taken a closer look at the biophysical features of cell-free membrane protein expression. By characterizing polysome formation and truncation of protein products for soluble and membrane proteins, we established that the aggregation state of proteins expressed in the PURE system changes with membrane protein expression. Our coarse-grained model emphasizes that increases in membrane protein yield with the addition of a membrane surface are a kinetic effect and not due to saturation of available membrane surface. It also shows that one cannot simply consider CFE of membrane proteins as an equilibrium between membrane-bound and solution-dispersed proteins. Instead, the total yield of protein is strongly influenced by a negative feedback loop of ribosome aggregation and stalling. An important result of our study is the competition between membrane association rate $k_+[V][R]$ and aggregation rate $k_{agg}[R]^2$. At a fixed initial ribosome $[R]$ and vesicle concentration $[V]$, the balance is determined by the ratio of the rate constants k_+ and k_{agg}. Our results are corroborated by experimental results from other groups who found that membrane protein yield is reduced by

https://doi.org/10.1021/acssynbio.3c00357
ACS Synth. Biol. 2024, 13, 129−140

deletion of the N-terminal domain but increased by anchoring of the nascent chain by NTA-His complexes onto the membrane surface.[47,55] Anchoring of expressed proteins to the membrane will increase k_+, while deletion of the N-terminal domain will expose the hydrophobic transmembrane segment to the solute, increasing k_{agg}. Additionally, Eaglesfield et al. have shown that excess aggregation by deletion of the N-terminal domain can be improved by artificial ribosome anchoring to the membrane, fully consistent with our model.[47] Our results are also consistent with results from Harris et al. that have shown large increases in yield of both aggregated and membrane-inserted protein with changes in membrane composition,[30] a result that cannot be explained by equilibrium partitioning models but corroborates our described feedback mechanism.

To improve the membrane protein yield, we examined how the rate constant k_+ could be altered to change membrane association. The rate constant k_+ can be tuned by the molecular properties of the lipid membrane, an effect we have previously studied quantitively, demonstrating that peptide insertion rates into polymer/lipid membranes are increased by the generation of lipid packing defects.[45] Further effects on k_+ might be expected by other changes in membrane composition, e.g., headgroup charge or hydrocarbon chain saturation.[50] Less is known about the effects of membrane composition on the dissociation rate constant k_-. Recently, it was shown that CFE membrane protein yield can be improved by increasing membrane viscosity.[25] Membrane viscosity might lower k_- as the dynamics of protein adsorbed to the interface will be slower. Lower k_- values should lead to larger membrane affinity and, in this way, could improve yield. To understand these effects quantitatively would be useful to obtain more systematic data on peptide or nascent chain unbinding kinetics with membrane composition.

Apart from the contribution of membrane composition, both k_+ and k_- depend on the protein sequence. Unfortunately, no systematic prediction between these rate constants and sequence exists. In principle, molecular dynamic simulation or single-molecule experiments could provide these rates. Instead of relying on these low-throughput methods, we utilized a result from our model that a protein's N-terminal domain will determine success or aggregation early on during protein synthesis. This motivated us, even if the cell-free system is clearly an out of equilibrium system, to approximate this binding step as an equilibrium between the nascent chain complex and membrane surface with a partitioning free energy of ΔG_{wm}. We justify this approximation because the short nascent chain binds and unbinds faster from the membrane than the nascent chain elongation, which gives the system time to sample its equilibrium distribution. If aggregation is avoided, more favorable N-terminal domain ΔG_{wm} values would always predict a higher protein yield, which we indeed observed using MscL-chimeras. These results suggest that evolved sequences from biological organisms have a balanced membrane affinity with aggregation propensity. As expected by this reasoning, a synthetic polylysine sequence, which should have both a high membrane affinity and a high propensity to aggregate, did not improve expression as the most favorable ΔG_{wm} value might suggest. Based on our modeling results, we defined the MPES scale, which allows us to quantify the balance of aggregation vs membrane association by previously determined empirical scales. We believe that our approach, which combines mechanistic insight with rational engineering, could be applied to increase the expression of a wide array of membrane proteins.

Limitations and Possible Extensions of Our Model. Our model made a series of simplifications. For example, translation initiation can be more explicitly modeled as a multistep process, or codon-specific elongation rates might be considered.[35,43] We would expect that by these additions, the predictions for truncated products would become sharper and better resolve the band structure of the Western blots. To account for sequence-specific variations of k_{RNA} and k_{init}, our model could be combined with approaches that consider the thermodynamics of ribosome-RNA binding or transcription initiation rates.[31,32]

Importantly, the main result of the competition between membrane association and aggregation is robust against variations in these parameters. We suggest that the crudest simplification of this work is that a protein of any length at the membrane is protected from aggregation. Instead of the simplified picture of protein maturation with a single rate constant k_{mat}, proteins on the membrane surface insert and fold in a multistep process. For example, force pulling experiments show that multipass proteins fold by insertion of hairpins, divided by barriers on the order of ~10 k_BT.[56] If the lifetimes of unfolded proteins on the membrane surface are comparable, the frequency of collisions between two not-yet-folded proteins on the same or different liposomes leads to additional aggregation pathways (Figure 4). In other words, depending on the stability of the folding pathway, cotranslational insertion does not fully protect membrane proteins from aggregation, as they might interact via intra- or intermembrane interactions during the folding process. This mode of interaction provides a possible explanation for the weaker bands of the truncated product in Figure 2c, which are not captured by the current model. Consistent with these ideas, a seven-pass Rhodopsin was observed to express with a different dependence on membrane composition than the simpler, two-pass MscL studied here and previously.[28] By additional measurements, these short comings can be addressed systematically. In Peruzzi et al., the stability of hydrophobically mismatched de novo proteins was studied in detail using experimental and computational approaches.[53] This study showed that small changes to membrane protein stability by membrane deformation indeed systemically change the protein yield, consistent with the additional aggregation pathways discussed here. Hydrophobic mismatch is straightforward to include quantitatively in our model but would require separate measurements of membrane protein association and insertion

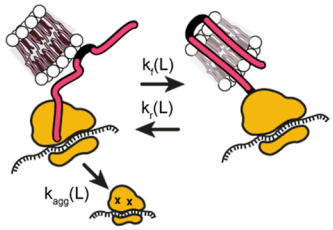


Figure 4. Possible addition considering the stability of the membrane protein folding process at the membrane against aggregation. The protein folding and unfolding pathway, represented by the rates k_f and k_r, might depend on the protein length L. If these rates are significant compared to the protein−protein collision rates, additional aggregation pathways appear.

https://doi.org/10.1021/acssynbio.3c00357
ACS Synth. Biol. 2024, 13, 129−140

kinetics. Even larger changes of the rate constants k_f and k_r with membrane composition, maybe even favoring the unfolded state, would clearly lower protein yield independent of the feedback effects described here.[57] Similarly, variation in translation initiation (in our model, the parameter k_{init}) has a large effect on predicted protein yield, as is well-known for soluble proteins.[32,38] The reason is quite intuitive; if there is little protein synthesis, then there is little aggregation, and indeed translation initiation is then the limiting factor. We want to note that in the present study, k_{init} was kept constant, which reflects that for the studied constructs, the aggregation and stalling effect is more significant compared to variation in translation initiation. It is certainly possible that small deviations from our predictions (e.g. the double mutants in Figure 3g) might be improved by including sequence specific changes of k_{init}.

Another important aspect is the functionality of the synthesized proteins. Previous studies have demonstrated that cell-free systems indeed yield functional proteins,[58] including the MscL and de novo proteins studied here.[28,53] However, the proposed tuning of the N-terminal AA sequence must be undertaken with care so as not to disturb motifs important for protein function. For example, the N-terminal domain is important for the opening of hydrophilic channels.[48] In this study, we did not assess the functionality of the MscL-chimeras and recognize that evaluating protein function after alterations to N-terminal sequences will be important in future studies that use this approach.

Experimentally, we studied the PURE system. We speculate that our main results hold in principle for other CFE systems and maybe even for primitive cells. Clearly, in different CFE systems, additional molecular players might become important. A possible modification of our model would be to allow for additional states that protect from aggregation. For example, chaperones might be present in crude cell extracts, lowering the aggregation rate k_{agg}. Indeed, MscL expresses with high yields in crude cell extracts.[59] To study such processes in detail, it would be interesting to spike the PURE system with chaperones, or similar molecules, and observe their ability to effectively suppress aggregation and thereby improve membrane protein yield.

CONCLUSIONS

We developed a kinetic model for cell-free membrane protein expression. Our results provide a unified framework to understand CFE of membrane proteins and are consistent with a large number of experimental data. In principle, our model might be adjusted for each membrane protein, e.g., by parametrization of sequence-specific aggregation and membrane association rate constants. This approach is limited by its low throughput. Instead, we have used the developed intuition to define a membrane protein expression score that can be readily calculated and demonstrate the benefit that biophysical insight can provide for engineering bottom-up synthetic biology systems.

MATERIALS AND METHODS

Used Plasmids and Templates. MscL-GFP, sfGFP, and tdTomato were prepared as described previously.[20,28] DNA templates for experiments performed in Figure 3, including chimeric MscL and de novo designed proteins, were ordered as gBlocks from Twist Biosciences with a T7 promoter and terminator, as well as a ribosome binding site (Table S1). gBlocks were amplified via PCR and purified using a PureLink PCR Purification Kit (Invitrogen). PCR products were used directly in cell-free reactions.[39]

Assembly of CFE Reactions and Fluorescence Readout. Protein expression was performed with the PURExpress In Vitro Protein Synthesis kit (E6800, NEB), according to the manufacturer's instructions. 30 μL reactions were assembled with a final concentration of 10 mM lipid and 3.3 nM plasmid (or approximately 200 ng). Reactions were allowed to progress at 37 °C for 3 h. All fluorescence experiments were performed on a plate reader (Molecular Devices Spectra Max i3) at 37 °C with a reaction volume of 30 μL. Fluorescence of sfGFP and transmembrane domain-GFP fusion proteins (MscL and de novo-designed proteins) was measured with an excitation of 480 nm and emission at 507 nm. tdTomato was excited at 553 nm, and emission was collected at 581 nm.

Ribosome Sucrose Gradient Experiments. Sucrose gradients were prepared from gradient buffer using a Biocomp Gradient Master as described previously.[40] PURE cell-free reactions of 30 μL with plasmid coding for MscL or tdTomato were prepared and incubated at 37 °C for 1.5 h. Reactions were quenched by being placed on ice. Sucrose gradients were prepared from gradient buffer (20 mM Tris−HCl (pH 7.5 at 4 °C), 100 mM NH_4Cl, and 10 mM $MgCl_2$) with 10 and 40% sucrose in SW41 polyclear centrifuge tubes (Seton Scientific) using a Biocomp Gradient Master and chilled to 4 °C. Cell-free reactions were diluted with 200 μL of gradient buffer and layered onto chilled gradients. The gradients were ultracentrifuged at 41,000 rpm for 3 h at 4 °C [Optima L-80 XP ultracentrifuge (Beckman-Coulter)]. Gradients were analyzed with a Piston Gradient Fractionator (Biocomp) coupled to a Triax FC-2 UV-260/280 flow cell. Traces of 260 nm light adsorption versus elution volumes were obtained for each gradient and adjusted with a blank sucrose sample.

Western Blots. Cell-free expressed proteins were analyzed by Western blot to observe the presence of truncation products. Cell-free expressed protein samples were run on a 16.5% Tricine Mini-PROTEAN Precast Protein Gel to enhance the separation of smaller protein products. Wet transfer was performed onto a PVDF membrane (Bio-Rad) for 45 min at 100 V. Membranes were then blocked for an hour at room temperature in 5% milk in TBST (pH 7.6:50 mM Tris, 150 mM NaCl, HCl to pH 7.6, 0.1% Tween) and incubated for 1 h at room temperature or overnight at 4 °C with the primary solution (anti-Flag, diluted 1:1000 in 5% milk in TBST). Primary antibody solution was decanted, and the membrane was washed three times for 5 min in TBST and subsequently incubated in secondary solution at room temperature for 1 h [HRP-anti-Mouse (CST 7076) diluted 1:3000 in 5% milk in TBST]. Membranes were then washed in TBST, incubated with Clarity Western ECL Substrate (Bio-Rad) for 5 min, and imaged in an Azure Biosystems c280 imager. Uncropped Western blots are shown in Figure S1.

eGFP Calibration Curve. Purified eGFP protein was purchased from ChromoTek (eGFP250, Proteintech Group Inc.) and serially diluted with PURExpress buffer (50 mM HEPES-KOH, pH = 7.6, 100 mM potassium glutamate, 13 mM magnesium acetate, 2 mM spermidine, and 1 mM DTT) to create a calibration curve based on eGFP fluorescence. Fluorescence measurements were performed on a plate reader (Molecular Devices Spectra Max i3) with an excitation of 480

https://doi.org/10.1021/acssynbio.3c00357
ACS Synth. Biol. 2024, 13, 129−140

nm and emission at 507 nm. Other experimental details were as in Jacobs et al.[28]

Motif Search. Annotated MscL homologues were downloaded from UniProt. N-terminal domains between 9 and 20 amino acids were identified by their annotation and extracted, yielding 181 sequences. These sequences were uploaded to the motif discovery tool MEME (v. 5.4.1).[60] The multilevel consensus sequence MSIIKEFR appeared in 43 N-terminal domains for which we calculated partitioning free energies as described below.

Partitioning Free Energy and Solubility Score Calculation. Partitioning free energies were calculated using MPEx (v3.3.0) using the "interfacial scale" with 100% helicity with the setting "No End Groups" at 0 mV bilayer surface charge.[61] Solubility was calculated by the CamSol Web server (http://www-vendruscolo.ch.cam.ac.uk/camsolmethod.html) using the "CamSol intrinsic" setting.

Model Reactions and Fit to Data. RNA synthesis

$$\emptyset \xrightarrow{k_{RNA}} [RNA_0] \quad (1)$$

Ribosome binding RNA + synthesis in solution

$$[RNA_0] + [R] \xrightarrow{k_{init}} [RRNA_0] \quad (2)$$

$$[RRNA_L] \xrightarrow{k_{syn}*} [RRNA_{L+1}], \qquad L < N \quad (3)$$

$$[RRNA_{L=N}] \xrightarrow{k_{syn}*} [P_N] + [RNA_0] + [R] \quad (4)$$

$$[P_N] \xrightarrow{k_{mat}} [GFP] \quad (5)$$

Ribosome binding/unbinding vesicle

$$[RRNA_L] \xleftrightarrow{k+,k-} [VRRNA_L] \quad (6)$$

Co-translational protein synthesis and membrane protein maturation

$$[VRRNA_L] \xrightarrow{k_{syn}*} [VRRNA_{L+1}], \qquad L < N \quad (7)$$

$$[VRRNA_{L=N}] \xrightarrow{k_{syn}*} [VP_N] + [RNA_0] + [R] \quad (8)$$

$$[VP_N] \xrightarrow{k_{mat}} [VGFP] \quad (9)$$

Ribosome degradation/fall off

$$[RRNA_L] \xrightarrow{k_{deg}} [P_L] \quad (10)$$

Ribosome aggregation

$$[RRNA_n] + [RRNA_m] \xrightarrow{k_{agg}} [P_n] + [P_m] \quad (11)$$

We considered the elementary reactions for the synthesis of MscL-GFP with a total length of 1170 nucleotides. For computational efficiency, the synthesis was modeled in steps of 10 amino acids, meaning that $N = \frac{1170}{3}^* \frac{1}{10} = 39$ with a scaled synthesis rate $k_{syn}* = \frac{k_{syn}}{10} = 0.033$ amino acids/s. For a nascent chain of length L, the unbinding rate constant k_- was scaled by exp $(-L^*10)$ to reflect the increase of membrane affinity with protein length. We account for proteins synthesized in solution [GFP] and at the membrane surface [VGFP]. If not described otherwise for comparison with experiments, we calculated the total yield of [GFP] + [VGFP]. For Figure 2c, the truncated protein of length L $[P_L]$ was reported.

The model reactions were implemented in Python (v3.7.4) using the GillesPy2 package (v1.6.7) by numerical integration of the ordinary differential equations defined above. The fit was performed by minimizing the sum of squared differences between model total GFP yield and experimental trajectories for all vesicle concentrations (global fit) using gp_minimize from scikit-optimize (v0.9). The fitting routine gp_minimize was performed by 1000 evaluations of the model function and otherwise default parameter values. Standard deviation of the fitted values was calculated from six gp_minimize runs. Further data processing was done using numpy (v1.21.4).

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acssynbio.3c00357.

Comparison of CFE yield between addition of liposomes and Brij-58 detergent, uncropped Western blots, translation initiation rate calculations, liposome floatation experiments data and methods, and DNA sequences of all constructs (PDF)

■ AUTHOR INFORMATION

Corresponding Authors

Jan Steinkühler − *Department of Biomedical Engineering, Northwestern University, Evanston, Illinois 60208, United States; Bio-Inspired Computation, Kiel University, 24143 Kiel, Germany; Kiel Nano, Surface and Interface Science KiNSIS, Kiel University, 24118 Kiel, Germany*; Email: jst@tf.uni-kiel.de

Neha P. Kamat − *Department of Biomedical Engineering, Northwestern University, Evanston, Illinois 60208, United States; Center for Synthetic Biology and Chemistry of Life Processes Institute, Northwestern University, Evanston, Illinois 60208, United States*; orcid.org/0000-0001-9362-6106; Email: nkamat@northwestern.edu

Authors

Justin A. Peruzzi − *Department of Chemical and Biological Engineering, Center for Synthetic Biology, Northwestern University, Evanston, Illinois 60208, United States*

Antje Krüger − *Department of Chemical and Biological Engineering, Center for Synthetic Biology, Northwestern University, Evanston, Illinois 60208, United States*

Citlayi G. Villaseñor − *Department of Biomedical Engineering, Northwestern University, Evanston, Illinois 60208, United States*

Miranda L. Jacobs − *Department of Biomedical Engineering, Northwestern University, Evanston, Illinois 60208, United States*

Michael C. Jewett − *Department of Chemical and Biological Engineering, Center for Synthetic Biology, Northwestern University, Evanston, Illinois 60208, United States; Center for Synthetic Biology and Chemistry of Life Processes Institute, Northwestern University, Evanston, Illinois 60208, United States; Department of Bioengineering, Stanford University, Stanford, California 94305, United States*; orcid.org/0000-0003-2948-6211

Complete contact information is available at: https://pubs.acs.org/10.1021/acssynbio.3c00357

https://doi.org/10.1021/acssynbio.3c00357
ACS Synth. Biol. 2024, 13, 129−140

Author Contributions

[▽]J.S. and J.A.P. contributed equally. J.S., J.A.P., and N.P.K. designed experiments and wrote the manuscript; J.S. and J.A.P. designed constructs and performed experiments; J.S. conceived the model idea and performed calculations; A.K. and M.C.J. designed and analyzed polysome flotation experiments; A.K. performed polysome flotation experiments; M.L.J. contributed MscL CFE data and plasmids; C.G.V. and N.P.K. designed, performed, and analyzed liposome flotation, calibration, and detergent control experiments. All authors analyzed data, discussed results, and contributed to the manuscript.

Notes

The authors declare the following competing financial interest(s): N.P.K., J.A.P., and J.S. are inventors on a U.S. patent application (US20230277456A1) currently pending, submitted by Northwestern University, that covers organizing cell-free expressed membrane proteins in synthetic membranes.

■ ACKNOWLEDGMENTS

This research was supported in part by the National Science Foundation under Grant no. 1844336 (N.P.K., M.C.J., J.S.) and no. 2145050 (J.A.P.), by the Army Contracting Command no. W911NF-22-2-0246 and W52P1J-21-9-3023 (N.P.K., M.C.J.) and through the computational resources and staff contributions provided for the Quest high-performance computing facility at Northwestern University, which is jointly supported by the Office of the Provost, the Office for Research, and Northwestern University Information Technology. J.A.P. gratefully acknowledges support from the Ryan Fellowship, the International Institute for Nanotechnology at Northwestern University, and an NSF Graduate Research Fellowship. J.S. acknowledges financial support from the Bundesministerium für Bildung und Forschung (BMBF) project NEUZELL (grant no. 031B1333). This work used the Scientific Compute Cluster at GWDG, the joint data center of Max Planck Society for the Advancement of Science (MPG) and University of Göttingen.

■ ABBREVIATION

MscL-GFP, mechanosensitive channel of large conductance fused to monomeric enhanced green fluorescent protein

■ REFERENCES

(1) Silverman, A. D.; Karim, A. S.; Jewett, M. C. Cell-Free Gene Expression: An Expanded Repertoire of Applications. *Nat. Rev. Genet.* **2020**, *21* (3), 151−170.

(2) Carlson, E. D.; Gan, R.; Hodgman, C. E.; Jewett, M. C. Cell-Free Protein Synthesis: Applications Come of Age. *Biotechnol. Adv.* **2012**, *30* (5), 1185−1194.

(3) Blow, J. J.; Laskey, R. A. Initiation of DNA Replication in Nuclei and Purified DNA by a Cell-Free Extract of Xenopus Eggs. *Cell* **1986**, *47* (4), 577−587.

(4) Pardee, K.; Green, A. A.; Ferrante, T.; Cameron, D. E.; DaleyKeyser, A.; Yin, P.; Collins, J. J. Paper-Based Synthetic Gene Networks. *Cell* **2014**, *159* (4), 940−954.

(5) Zawada, J. F.; Yin, G.; Steiner, A. R.; Yang, J.; Naresh, A.; Roy, S. M.; Gold, D. S.; Heinsohn, H. G.; Murray, C. J. Microscale to Manufacturing Scale-up of Cell-Free Cytokine Production—A New Approach for Shortening Protein Production Development Timelines. *Biotechnol. Bioeng.* **2011**, *108* (7), 1570−1578.

(6) Jewett, M. C.; Calhoun, K. A.; Voloshin, A.; Wuu, J. J.; Swartz, J. R. An Integrated Cell-Free Metabolic Platform for Protein Production and Synthetic Biology. *Mol. Syst. Biol.* **2008**, *4* (1), 220.

(7) Pardee, K.; Slomovic, S.; Nguyen, P. Q.; Lee, J. W.; Donghia, N.; Burrill, D.; Ferrante, T.; McSorley, F. R.; Furuta, Y.; Vernet, A.; Lewandowski, M.; Boddy, C. N.; Joshi, N. S.; Collins, J. J. Portable, On-Demand Biomolecular Manufacturing. *Cell* **2016**, *167* (1), 248−259.e12.

(8) Swank, Z.; Laohakunakorn, N.; Maerkl, S. J. Cell-Free Gene-Regulatory Network Engineering with Synthetic Transcription Factors. *Proc. Natl. Acad. Sci. U.S.A.* **2019**, *116* (13), 5892−5901.

(9) Galinis, R.; Stonyte, G.; Kiseliovas, V.; Zilionis, R.; Studer, S.; Hilvert, D.; Janulaitis, A.; Mazutis, L. DNA Nanoparticles for Improved Protein Synthesis In Vitro. *Angew. Chem., Int. Ed.* **2016**, *55* (9), 3120−3123.

(10) Nguyen, P. Q.; Soenksen, L. R.; Donghia, N. M.; Angenent-Mari, N. M.; de Puig, H.; Huang, A.; Lee, R.; Slomovic, S.; Galbersanini, T.; Lansberry, G.; Sallum, H. M.; Zhao, E. M.; Niemi, J. B.; Collins, J. J. Wearable Materials with Embedded Synthetic Biology Sensors for Biomolecule Detection. *Nat. Biotechnol.* **2021**, *39* (11), 1366−1374.

(11) Bracaglia, S.; Ranallo, S.; Ricci, F. Electrochemical Cell-Free Biosensors for Antibody Detection. *Angew. Chem., Int. Ed.* **2023**, *62* (8), No. e202216512.

(12) Garenne, D.; Libchaber, A.; Noireaux, V. Membrane Molecular Crowding Enhances MreB Polymerization to Shape Synthetic Cells from Spheres to Rods. *Proc. Natl. Acad. Sci. U.S.A.* **2020**, *117* (4), 1902−1909.

(13) Noireaux, V.; Maeda, Y. T.; Libchaber, A. Development of an Artificial Cell, from Self-Organization to Computation and Self-Reproduction. *Proc. Natl. Acad. Sci. U.S.A.* **2011**, *108* (9), 3473−3480.

(14) Aufinger, L.; Simmel, F. C. Artificial Gel-Based Organelles for Spatial Organization of Cell-Free Gene Expression Reactions. *Angew. Chem., Int. Ed.* **2018**, *57* (52), 17245−17248.

(15) Liew, F. E.; Nogle, R.; Abdalla, T.; Rasor, B. J.; Canter, C.; Jensen, R. O.; Wang, L.; Strutz, J.; Chirania, P.; De Tissera, S.; Mueller, A. P.; Ruan, Z.; Gao, A.; Tran, L.; Engle, N. L.; Bromley, J. C.; Daniell, J.; Conrado, R.; Tschaplinski, T. J.; Giannone, R. J.; Hettich, R. L.; Karim, A. S.; Simpson, S. D.; Brown, S. D.; Leang, C.; Jewett, M. C.; Köpke, M. Carbon-Negative Production of Acetone and Isopropanol by Gas Fermentation at Industrial Pilot Scale. *Nat. Biotechnol.* **2022**, *40* (3), 335−344.

(16) Karim, A. S.; Dudley, Q. M.; Juminaga, A.; Yuan, Y.; Crowe, S. A.; Heggestad, J. T.; Garg, S.; Abdalla, T.; Grubbe, W. S.; Rasor, B. J.; Coar, D. N.; Torculas, M.; Krein, M.; Liew, F.; Quattlebaum, A.; Jensen, R. O.; Stuart, J. A.; Simpson, S. D.; Köpke, M.; Jewett, M. C. In Vitro Prototyping and Rapid Optimization of Biosynthetic Enzymes for Cell Design. *Nat. Chem. Biol.* **2020**, *16* (8), 912−919.

(17) Hunt, A. C.; Vögeli, B.; Kightlinger, W. K.; Yoesep, D. J.; Krüger, A.; Jewett, M. C. A High-Throughput, Automated, Cell-Free Expression and Screening Platform for Antibody Discovery. *bioRxiv* **2021**.

(18) Hunt, A. C.; Case, J. B.; Park, Y.-J.; Cao, L.; Wu, K.; Walls, A. C.; Liu, Z.; Bowen, J. E.; Yeh, H.-W.; Saini, S.; Helms, L.; Zhao, Y. T.; Hsiang, T.-Y.; Starr, T. N.; Goreshnik, I.; Kozodoy, L.; Carter, L.; Ravichandran, R.; Green, L. B.; Matochko, W. L.; Thomson, C. A.; Vögeli, B.; Krüger, A.; VanBlargan, L. A.; Chen, R. E.; Ying, B.; Bailey, A. L.; Kafai, N. M.; Boyken, S. E.; Ljubetič, A.; Edman, N.; Ueda, G.; Chow, C. M.; Johnson, M.; Addetia, A.; Navarro, M.-J.; Panpradist, N.; Gale, M.; Freedman, B. S.; Bloom, J. D.; Ruohola-Baker, H.; Whelan, S. P. J.; Stewart, L.; Diamond, M. S.; Veesler, D.; Jewett, M. C.; Baker, D. Multivalent Designed Proteins Neutralize SARS-CoV-2 Variants of Concern and Confer Protection against Infection in Mice. *Sci. Transl. Med.* **2022**, *14* (646), No. eabn1252.

(19) Nuti, N.; Rottmann, P.; Stucki, A.; Koch, P.; Panke, S.; Dittrich, P. S. A Multiplexed Cell-Free Assay to Screen for Antimicrobial Peptides in Double Emulsion Droplets. *Angew. Chem., Int. Ed.* **2022**, *61* (13), No. e202114632.

(20) Stark, J. C.; Jaroentomeechai, T.; Moeller, T. D.; Hershewe, J. M.; Warfel, K. F.; Moricz, B. S.; Martini, A. M.; Dubner, R. S.; Hsu, K. J.; Stevenson, T. C.; Jones, B. D.; DeLisa, M. P.; Jewett, M. C. On-

https://doi.org/10.1021/acssynbio.3c00357
ACS Synth. Biol. 2024, 13, 129−140

Demand Biomanufacturing of Protective Conjugate Vaccines. *Sci. Adv.* **2021**, *7* (6), No. eabe9444.
(21) Warfel, K. F.; Williams, A.; Wong, D. A.; Sobol, S. E.; Desai, P.; Li, J.; Chang, Y.-F.; DeLisa, M. P.; Karim, A. S.; Jewett, M. C. A Low-Cost, Thermostable, Cell-Free Protein Synthesis Platform for On-Demand Production of Conjugate Vaccines. *ACS Synth. Biol.* **2023**, *12* (1), 95−107.
(22) Kruyer, N. S.; Sugianto, W.; Tickman, B. I.; Alba Burbano, D.; Noireaux, V.; Carothers, J. M.; Peralta-Yahya, P. Membrane Augmented Cell-Free Systems: A New Frontier in Biotechnology. *ACS Synth. Biol.* **2021**, *10* (4), 670−681.
(23) Yunker, P. J.; Asahara, H.; Hung, K.-C.; Landry, C.; Arriaga, L. R.; Akartuna, I.; Heyman, J.; Chong, S.; Weitz, D. A. One-Pot System for Synthesis, Assembly, and Display of Functional Single-Span Membrane Proteins on Oil-Water Interfaces. *Proc. Natl. Acad. Sci. U.S.A.* **2016**, *113* (3), 608−613.
(24) Hershewe, J. M.; Warfel, K. F.; Iyer, S. M.; Peruzzi, J. A.; Sullivan, C. J.; Roth, E. W.; DeLisa, M. P.; Kamat, N. P.; Jewett, M. C. Improving Cell-Free Glycoprotein Synthesis by Characterizing and Enriching Native Membrane Vesicles. *Nat. Commun.* **2021**, *12* (1), 2363.
(25) Peruzzi, J. A.; Galvez, N. R.; Kamat, N. P. Engineering Transmembrane Signal Transduction in Synthetic Membranes Using Two-Component Systems. *Proc. Natl. Acad. Sci. U.S.A.* **2023**, *120* (19), No. e2218610120.
(26) Huff, H. C.; Maroutsos, D.; Das, A. Lipid Composition and Macromolecular Crowding Effects on CYP2J2-Mediated Drug Metabolism in Nanodiscs. *Protein Sci.* **2019**, *28* (5), 928−940.
(27) Schoborg, J. A.; Hershewe, J. M.; Stark, J. C.; Kightlinger, W.; Kath, J. E.; Jaroentomeechai, T.; Natarajan, A.; DeLisa, M. P.; Jewett, M. C. A Cell-free Platform for Rapid Synthesis and Testing of Active Oligosaccharyltransferases. *Biotechnol. Bioeng.* **2018**, *115* (3), 739−750.
(28) Jacobs, M. L.; Boyd, M. A.; Kamat, N. P. Diblock Copolymers Enhance Folding of a Mechanosensitive Membrane Protein during Cell-Free Expression. *Proc. Natl. Acad. Sci. U.S.A.* **2019**, *116* (10), 4031−4036.
(29) Blackholly, L. R.; Harris, N. J.; Findlay, H. E.; Booth, P. J. Cell-Free Expression to Probe Co-Translational Insertion of an Alpha Helical Membrane Protein. *Front. Mol. Biosci.* **2022**, *9*, 795212.
(30) Harris, N. J.; Reading, E.; Ataka, K.; Grzegorzewski, L.; Charalambous, K.; Liu, X.; Schlesinger, R.; Heberle, J.; Booth, P. J. Structure Formation during Translocon-Unassisted Co-Translational Membrane Protein Folding. *Sci. Rep.* **2017**, *7* (1), 8021.
(31) LaFleur, T. L.; Hossain, A.; Salis, H. M. Automated Model-Predictive Design of Synthetic Promoters to Control Transcriptional Profiles in Bacteria. *Nat. Commun.* **2022**, *13* (1), 5159.
(32) Salis, H. M.; Mirsky, E. A.; Voigt, C. A. Automated Design of Synthetic Ribosome Binding Sites to Control Protein Expression. *Nat. Biotechnol.* **2009**, *27* (10), 946−950.
(33) Stögbauer, T.; Windhager, L.; Zimmer, R.; Rädler, J. O. Experiment and Mathematical Modeling of Gene Expression Dynamics in a Cell-Free System. *Integr. Biol.* **2012**, *4* (5), 494−501.
(34) Gonzales, D. T.; Yandrapalli, N.; Robinson, T.; Zechner, C.; Tang, T.-Y. D. Cell-Free Gene Expression Dynamics in Synthetic Cell Populations. *ACS Synth. Biol.* **2022**, *11* (1), 205−215.
(35) Doerr, A.; de Reus, E.; van Nies, P.; van der Haar, M.; Wei, K.; Kattan, J.; Wahl, A.; Danelon, C. Modelling Cell-Free RNA and Protein Synthesis with Minimal Systems. *Phys. Biol.* **2019**, *16* (2), 025001.
(36) Karzbrun, E.; Shin, J.; Bar-Ziv, R. H.; Noireaux, V. Coarse-Grained Dynamics of Protein Synthesis in a Cell-Free System. *Phys. Rev. Lett.* **2011**, *106* (4), 048104.
(37) Matsuura, T.; Tanimura, N.; Hosoda, K.; Yomo, T.; Shimizu, Y. Reaction Dynamics Analysis of a Reconstituted Escherichia Coli Protein Translation System by Computational Modeling. *Proc. Natl. Acad. Sci. U.S.A.* **2017**, *114* (8), E1336−E1344.
(38) Haberstock, S.; Roos, C.; Hoevels, Y.; Dötsch, V.; Schnapp, G.; Pautsch, A.; Bernhard, F. A Systematic Approach to Increase the Efficiency of Membrane Protein Production in Cell-Free Expression Systems. *Protein Expression Purif.* **2012**, *82* (2), 308−316.
(39) Tuckey, C.; Asahara, H.; Zhou, Y.; Chong, S. Protein Synthesis Using a Reconstituted Cell-Free System. *Curr. Protoc. Mol. Biol.* **2014**, *108*, 16−31.
(40) d'Aquino, A. E.; Azim, T.; Aleksashin, N. A.; Hockenberry, A. J.; Krüger, A.; Jewett, M. C. Mutational Characterization and Mapping of the 70S Ribosome Active Site. *Nucleic Acids Res.* **2020**, *48* (5), 2777−2789.
(41) Krüger, A.; Watkins, A. M.; Wellington-Oguri, R.; Romano, J.; Kofman, C.; DeFoe, A.; Kim, Y.; Anderson-Lee, J.; Fisker, E.; Townley, J.; Participants, E.; d'Aquino, A. E.; Das, R.; Jewett, M. C. Community Science Designed Ribosomes with Beneficial Phenotypes. *Nat Commun.* **2023**, *14*, 961.
(42) Bertolini, M.; Fenzl, K.; Kats, I.; Wruck, F.; Tippmann, F.; Schmitt, J.; Auburger, J. J.; Tans, S.; Bukau, B.; Kramer, G. Interactions Between Nascent Proteins Translated by Adjacent Ribosomes Drive Homomer Assembly. *Science* **2021**, *371* (6524), 57−64.
(43) Rudorf, S.; Thommen, M.; Rodnina, M. V.; Lipowsky, R. Deducing the Kinetics of Protein Synthesis In Vivo from the Transition Rates Measured In Vitro. *PLOS Comput. Biol.* **2014**, *10* (10), No. e1003909.
(44) Fox, C. B.; Wayment, J. R.; Myers, G. A.; Endicott, S. K.; Harris, J. M. Single-Molecule Fluorescence Imaging of Peptide Binding to Supported Lipid Bilayers. *Anal. Chem.* **2009**, *81* (13), 5130−5138.
(45) Steinkühler, J.; Jacobs, M. L.; Boyd, M. A.; Villaseñor, C. G.; Loverde, S. M.; Kamat, N. P. PEO-b-PBD Diblock Copolymers Induce Packing Defects in Lipid/Hybrid Membranes and Improve Insertion Rates of Natively Folded Peptides. *Biomacromolecules* **2022**, *23* (11), 4756−4765.
(46) Neugebauer, S. A.; Baulig, A.; Kuhn, A.; Facey, S. J. Membrane Protein Insertion of Variant MscL Proteins Occurs at YidC and SecYEG of Escherichia Coli. *J. Mol. Biol.* **2012**, *417* (4), 375−386.
(47) Eaglesfield, R.; Madsen, M. A.; Sanyal, S.; Reboud, J.; Amtmann, A. Cotranslational Recruitment of Ribosomes in Protocells Recreates a Translocon-Independent Mechanism of Proteorhodopsin Biogenesis. *iScience* **2021**, *24* (5), 102429.
(48) Bavi, N.; Cortes, D. M.; Cox, C. D.; Rohde, P. R.; Liu, W.; Deitmer, J. W.; Bavi, O.; Strop, P.; Hill, A. P.; Rees, D.; Corry, B.; Perozo, E.; Martinac, B. The Role of MscL Amphipathic N Terminus Indicates a Blueprint for Bilayer-Mediated Gating of Mechanosensitive Channels. *Nat. Commun.* **2016**, *7*, 11984.
(49) Fernández-Vidal, M.; Jayasinghe, S.; Ladokhin, A. S.; White, S. H. Folding Amphipathic Helices into Membranes: Amphiphilicity Trumps Hydrophobicity. *J. Mol. Biol.* **2007**, *370* (3), 459−470.
(50) Harris, N. J.; Booth, P. J. Co-Translational Protein Folding in Lipid Membranes. *Trends Biochem. Sci.* **2019**, *44* (8), 729−730.
(51) Nagamori, S.; Vázquez-Ibar, J. L.; Weinglass, A. B.; Kaback, H. R. In Vitro Synthesis of Lactose Permease to Probe the Mechanism of Membrane Insertion and Folding. *J. Biol. Chem.* **2003**, *278* (17), 14820−14826.
(52) Dorsman, J. C.; Pepers, B.; Langenberg, D.; Kerkdijk, H.; Ijszenga, M.; den Dunnen, J. T.; Roos, R. A. C.; van Ommen, G.-J. B. Strong Aggregation and Increased Toxicity of Polyleucine over Polyglutamine Stretches in Mammalian Cells. *Hum. Mol. Genet.* **2002**, *11* (13), 1487−1496.
(53) Peruzzi, J. A.; Steinkühler, J.; Vu, T. Q.; Gunnels, T. F.; Lu, P.; Baker, D.; Kamat, N. P. Hydrophobic Mismatch Drives Self-Organization of Designer Proteins into Synthetic Membranes. *bioRxiv* **2022**.
(54) Sormanni, P.; Aprile, F. A.; Vendruscolo, M. The CamSol Method of Rational Design of Protein Mutants with Enhanced Solubility. *J. Mol. Biol.* **2015**, *427* (2), 478−490.
(55) Ando, M.; Schikula, S.; Sasaki, Y.; Akiyoshi, K. Proteoliposome Engineering with Cell-Free Membrane Protein Synthesis: Control of Membrane Protein Sorting into Liposomes by Chaperoning Systems. *Adv. Sci.* **2018**, *5* (10), 1800524.

https://doi.org/10.1021/acssynbio.3c00357
ACS Synth. Biol. 2024, 13, 129−140

(56) Choi, H.-K.; Min, D.; Kang, H.; Shon, M. J.; Rah, S.-H.; Kim, H. C.; Jeong, H.; Choi, H.-J.; Bowie, J. U.; Yoon, T.-Y. Watching Helical Membrane Proteins Fold Reveals a Common N- to C-Terminal Folding Pathway. *Science* **2019**, *366* (6469), 1150−1156.
(57) Sanders, M. R.; Findlay, H. E.; Booth, P. J. Lipid Bilayer Composition Modulates the Unfolding Free Energy of a Knotted α-Helical Membrane Protein. *Proc. Natl. Acad. Sci. U.S.A.* **2018**, *115* (8), E1799−E1808.
(58) Findlay, H. E.; Booth, P. J. The Folding, Stability and Function of Lactose Permease Differ in their Dependence on Bilayer Lipid Composition. *Sci. Rep.* **2017**, *7* (1), 13056.
(59) Abdine, A.; Verhoeven, M. A.; Park, K.-H.; Ghazi, A.; Guittet, E.; Berrier, C.; Van Heijenoort, C.; Warschawski, D. E. Structural Study of the Membrane Protein MscL Using Cell-Free Expression and Solid-State NMR. *J. Magn. Reson.* **2010**, *204* (1), 155−159.
(60) Bailey, T. L.; Johnson, J.; Grant, C. E.; Noble, W. S. The MEME Suite. *Nucleic Acids Res.* **2015**, *43* (W1), W39−W49.
(61) Snider, C.; Jayasinghe, S.; Hristova, K.; White, S. H. MPEx: A Tool for Exploring Membrane Proteins. *Protein Sci.* **2009**, *18* (12), 2624−2628.

https://doi.org/10.1021/acssynbio.3c00357
ACS Synth. Biol. 2024, 13, 129−140