Supporting Information

Korkmazhan et al. 10.1073/pnas.1700941114

TASEP Average Occupancy

The occupancy variable τ_i for each site where $i = 1 \dots L$ takes the value 1 if site i is occupied and 0 if it is not. For TASEP with open boundaries, the average occupancy of each site is given exactly by:

$$\langle \tau_i \rangle = \sum_{p=0}^{L-i-1} \frac{2p!}{p!(p+1)!} \frac{S_{L-1-p}}{S_L} + \frac{S_{i-1}}{S_L} \sum_{p=2}^{L-i+1} \frac{(p-1)(2L-2i-p)!}{(L-i)!(L-i+1-p)!} \beta^{-p} \quad \text{[S1]}$$

for $i \leq L-1$ and:

$$\langle \tau_L \rangle = \frac{1}{\beta} \frac{S_{L-1}}{S_L} \quad \text{[S2]}$$

for $i = L$ (12). Here S_n is given by:

$$S_n = \sum_{j=1}^{n} \frac{j(2n-1-j)!}{n!(n-j)!} \frac{(\gamma/\beta)^{j+1} - (\gamma/\alpha)^{j+1}}{(\gamma/\beta) - (\gamma/\alpha)}. \quad \text{[S3]}$$

Distribution of Residence Times Near the Membrane

Here we obtain the distribution $\Phi(t)$ of residence times t spent by an mRNA near the membrane, under the assumption that an mRNA is anchored to the membrane whenever there are ribosomes in the PSR. $\Phi(t)$ is therefore the distribution of times from an entry of a ribosome to an empty PSR to a complete evacuation of the PSR.

At very small α, when the density of ribosomes in the mRNA is very low, a ribosome that enters the empty PSR leaves it at the other end before any other ribosome enters, and $\Phi(t)$ is given approximately by a Gamma distribution:

$$\Gamma(\ell, t) = \gamma^{\ell} \frac{t^{\ell-1}}{(\ell-1)!} e^{-\gamma t}.$$

At higher values of α, it becomes likely that when a ribosome finishes translating the mRNA there are other ribosomes in the PSR behind it. To address such cases, we let $\upsilon(k)$ be the probability that there is a gap of k sites between a ribosome at the last site and the one behind it (namely, that the latter is positioned at site $L-k-1$). This distribution is given approximately by:

$$\upsilon(k) \simeq \alpha(1-\alpha)^k \beta^{-1} \frac{\beta-\alpha}{1-2\alpha} + \alpha^k (1-\alpha) \beta^{-1} \left(1 - \frac{\beta-\alpha}{1-2\alpha}\right), \quad \text{[S4]}$$

which has been obtained exactly for $\alpha < \gamma/2$ and $\alpha < \beta$ in the large L limit (50). With this, the probability $\Upsilon(k)$ of a gap equal to or smaller than k is given approximately by:

$$\Upsilon(k) \simeq \frac{\alpha^{k+1}(1-\beta-\alpha) + (1-\alpha)^{k+1}(\beta-\alpha) + 2\alpha - 1}{(2\alpha-1)\beta}. \quad \text{[S5]}$$

When the ribosome at the last site detaches from the mRNA, it leaves behind it a nonempty PSR if the gap to the following ribosome is smaller than ℓ, which occurs with probability $\Upsilon_{\text{PSR}} \equiv \Upsilon(\ell-1)$ (Fig. S1*A*). This probability, which can be written as $\left\langle \tau_L \left(1 - \prod_{i=m+1}^{L-1} (1-\tau_i)\right) \right\rangle / \langle \tau_L \rangle$, can also be estimated by the mean-field approximation $1 - \prod_{i=m+1}^{L-1} (1 - \langle \tau_i \rangle)$ using the exact expression for $\langle \tau_i \rangle$ (12, 13), which has the advantage that it extends to all α. The two approximations are validated by their clear agreement with the results of Monte-Carlo simulations (Fig. S1*B*).

To calculate the probability $\Phi(t)$ of a residence time t, we sum over the conditional probabilities given the number of proteins produced during this time or equivalently over the number of ribosomes traveling through the PSR. This leads to:

$$\frac{\Phi(t)}{1-\Upsilon_{\text{PSR}}} = \Gamma(\ell, t) + \sum_{k=0}^{\ell-1} \upsilon(k) \left[\int dt' \Gamma(k+1, t') \Gamma(\ell, t-t') + \sum_{k'=0}^{\ell-1} \upsilon(k') \left[\iint dt'\, dt'' \Gamma(k+1, t') \Gamma(k'+1, t'') \times \Gamma(\ell, t-t'-t'') + \cdots \right] \right].$$

As above, the motion of the most advanced ribosome is uninterrupted, and its exit time distribution is given exactly by a Gamma distribution. The denominator on the left-hand side accounts for the normalization from the probability of having the PSR empty once a ribosome exits the mRNA. The first term represents one ribosome moving over the whole PSR and exiting in a time t; the second term represents one ribosome moving over the whole PSR in a time $t-t'$ and another ribosome $k+1$ sites behind the exiting one moving k steps in a time t' to exit and so on, summed over all possible ks that lie in the PSR weighed by the probability of having a gap of size k. The expression takes into account all possible ways in which different numbers of ribosomes can exit before having an empty PSR.

To reduce the computational complexity of this expression, we ignore the fluctuations in the gap size and assume that all gaps take the mean value $g = \sum_k k \upsilon(k)$, yielding:

$$\frac{\Phi(t)}{1-\Upsilon_{\text{PSR}}} \simeq \Gamma(\ell, t) + \Upsilon_{\text{PSR}} \int_0^t dt'\, \Gamma(g+1, t') \Gamma(\ell, t-t') + \Upsilon_{\text{PSR}}^2 \iint dt'\, dt'' \Gamma(g+1, t') \Gamma(g+1, t'') \Gamma(\ell, t-t'-t'') + \cdots$$

We calculate g approximately using Eq. **S4**. By Laplace transforming every term through the linearity of the transform and using the fact that convolutions of probability distributions transform to products of the distributions in the Laplace domain, we can sum the Laplace transform of the whole expression to infinity to get Eq. 1:

$$\tilde{\Phi}(s) = \frac{(1-\Upsilon_{\text{PSR}}) \gamma^{\ell} (\gamma+s)^{-\ell}}{1 - \Upsilon_{\text{PSR}} \gamma^{g+1} (\gamma+s)^{-(g+1)}}. \quad \text{[S6]}$$

Finally, we take the inverse Laplace transform of $\tilde{\Phi}(s)$ numerically (52) to obtain $\Phi(t)$. We confirmed the validity of this analysis by comparing with results of Monte-Carlo simulations (Fig. S2*A*).

Distribution of Residence Times Away from the Membrane

We now turn to calculate the probability distribution for the residence time away from the membrane $\Theta(t)$. To do this, we consider again the most forward ribosome on the lattice, at the time when the PSR has just been evacuated. If this ribosome is k sites away from the PSR, corresponding to a gap of $\ell + k - 1$ with the

present ribosome, the time t before it reaches the PSR follows a Gamma distribution $\Gamma(k,t)$. Thus:

$$\Theta(t) = \frac{1}{1-\Upsilon_{\mathrm{PSR}}}\left[\sum_{k=1}^{m}\Gamma(k,t)\upsilon(\ell+k-1) + \upsilon(m+\ell)\,\alpha e^{-\alpha t}\left(\frac{\gamma}{\gamma-\alpha}\right)^{m+1} \times\left(1-\frac{\int_{(\gamma-\alpha)t}^{\infty}t^{m}e^{-t}dt}{m!}\right)\right]. \quad \text{[S7]}$$

The second term in the sum accounts for the case where upon exit of the translating ribosome, the mRNA remains unbound by any ribosome. In this case, arrival of a ribosome to the PSR involves translation initiation, namely entry to the first site of the lattice with rate α and translation of the entire SRR. While an exact expression is available for the full correlation functions of the TASEP, we use below their mean-field approximation to streamline the computation. We verified the accuracy of this approximation by comparison with Monte-Carlo simulations (Fig. S2*B*).

Suggested Experimental Validation of Model Predictions

Our model predicts several possible effects of the translation initiation rate α and the structure of the mRNA on the spatial distribution of mRNAs and membrane-bound proteins as well as on localization-dependent lifetime of the mRNA. Here we briefly consider possible experimental tests of these predictions.

In Fig. 4 we predict a nonlinear dependence of mRNA enrichment to the membrane on the translation initiation rate. This effect can be quantified for any gene encoding a membrane-bound protein from a single-gene operon. The translation initiation rate of this gene can then be perturbed either genetically, by generating a library of strains carrying different mutations to the RBS (53) or by chemical induction of a regulatory small RNA. The effect of initiation efficiency on the spatial distribution of the mRNA can then be measured using smFISH and superresolution microscopy. A similar approach can be taken to test the prediction that different genes that compose a polycistronic mRNA act collaboratively to drive it toward the membrane.

Our analysis suggests that mRNA residence near the membrane may increase the likelihood of protein colocalization in the membrane. Many membrane protein clusters are encoded on polycistronic transcripts and can serve to test this prediction. This can be done by measuring the effect of expressing a tagged variant of one of the genes from a separate transcript. Of particular interest is the case where this gene is a related outer-membrane protein. Particular examples in *E. coli* include FecA, an outer-membrane protein that shares an operon with the inner-membrane ABC transporter proteins FecBCDE, and the curli transport system, which includes inner- and outer-membrane components expressed from the *csg* operon. Genes in both of these operons are regulated posttranscriptionally by the small RNAs OmrA and OmrB, facilitating testing of the effect of translation initiation (54).

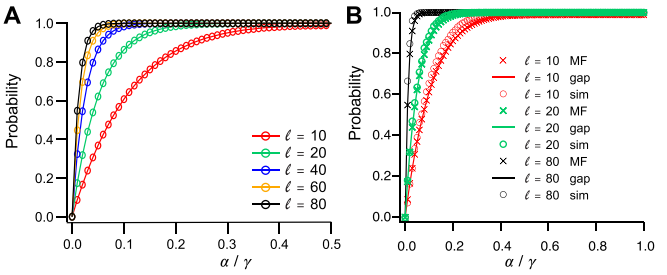

Fig. S1. The probability Υ_{PSR} that when a ribosome ends translation and releases the mRNA, other ribosomes are found in the PSR, keeping the mRNA near the membrane. (*A*) This probability as calculated from Eq. **S5** with $k=\ell-1$. (*B*) Comparison of mean field (MF), gap distribution (gap), and simulation (sim) results for different ℓ.

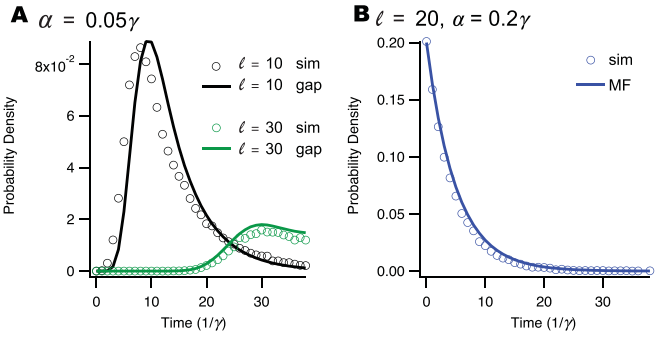

Fig. S2. Comparison with simulation results for residence time distributions. (*A*) Distribution of times spent near the membrane: gap distribution (gap) results for $\alpha=0.05\gamma$ and different ℓ, compared with simulation (sim). (*B*) Distribution of times spent away from the membrane: mean field (MF) result for $\alpha=0.2\gamma$ and $\ell=20$, compared with simulation (sim).

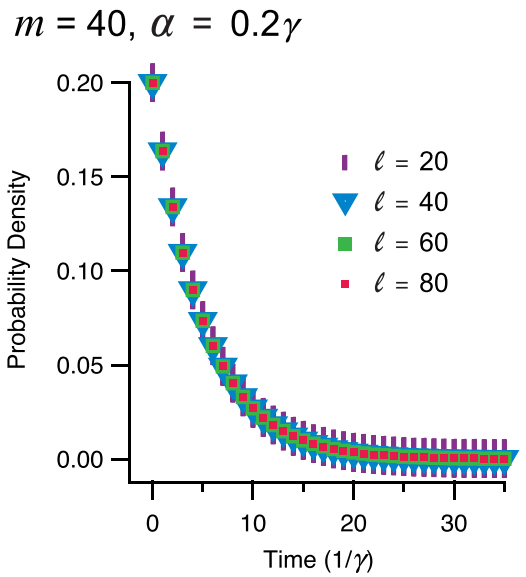

Fig. S3. Effect of PSR size on distribution of residence times away from the membrane. With a constant SSR length $m = 40$, the probability distribution for the time spent away from the membrane is insensitive to the length of the PSR.

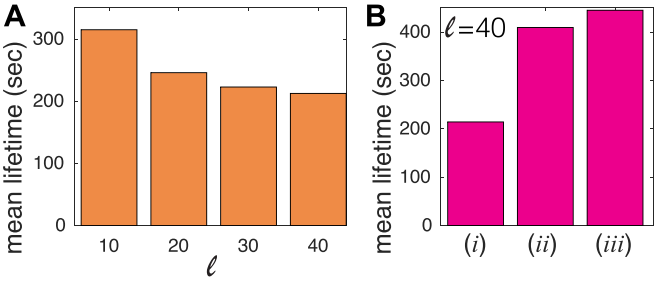

Fig. S4. Possible effect of mRNA localization on its stability. Here it is assumed that the half-life of mRNA near the membrane is 3 min and mRNA away from the membrane is 10. The distributions $\Phi(t)$ and $\Theta(t)$ are used to simulate the lifetime of mRNAs, from synthesis in the cytoplasm to degradation. (*A*) When the mRNA is more likely to stay near the membrane because its PSR is longer (Fig. 2*C*), its decay is accelerated. (*B*) When mRNA decay follows a multistep process, the mean lifetime of an mRNA may depend on the residence time distribution and not just on its mean. In *E. coli*, diffusion-limited search time can take around 50 s per searching molecule. In some cases, the number of available components can be in the range of 10 to 100, such that the waiting time for each step is on the order of 0.5 to 5 s. This time is on the same order as the residence time of an mRNA near the membrane for short ℓ or small α (Fig. 2), meaning that in these cases an mRNA will sometimes not stay near the membrane long enough for multiple processes to occur. In other cases, including stoichiometrically acting small RNAs, the number of available components may be even lower, and the effect of the multistep process can be noticeable even when the mRNA spends longer times near the membrane. To demonstrate this situation, we assume that mRNA decay near the membrane follows a Gamma distribution with parameters $a = 3$ and $b = 1$, which has the same mean half-life as in *A*. We then assume that half-life of mRNAs near the membrane is 214 s, as for $\ell = 40$ in *A*, but follows different distributions: (*i*) an exponential distribution; (*ii*) the distribution obtained in Fig. 2, which is approximately a Gamma distribution with $a = 4.22$ and $b = 22.3$; and (*iii*) a Gamma distribution with $a = 42.2$ and $b = 2.23$, which peaks at higher times.