Supporting Material for "*In vivo* kinetics of segregation and polar retention of MS2-GFP-RNA complexes in *Escherichia coli*"

Abhishekh Gupta[†,‡], Jason Lloyd-Price[†,‡], Ramakanth NeeliVenkata[†], Samuel M.D. Oliveira[†], and Andre S. Ribeiro[†]

† Laboratory of Biosystem Dynamics, Department of Signal Processing, Tampere University of Technology, FI-33101 Tampere, Finland.

‡ Equal contributions

Measurements in M63 Media

We imaged cells grown in M63 media at 37°C for 2 hours, following induction of the target RNA and of the tagging MS2-GFP proteins (one hour before), using the same experimental and analysis protocols as in the experiments in LB described in the main text. The spatial distribution of newly-produced complexes, corresponding to Fig. 2 *A* in the main text for LB, is presented in Fig. S1.

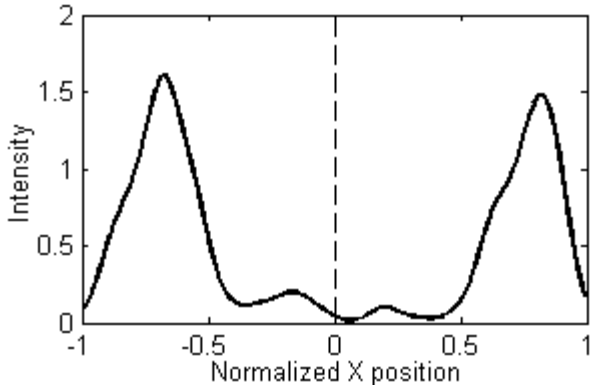

FIGURE S1 KDE of the spatial distribution of the fluorescence intensity (in arbitrary units) of complexes along the major axis of the cells, extracted from all cells and time points (*black line*, bandwidth 0.05). Data is extracted from cells that inherited no complexes but produced one or more. The old pole is at +1 and the new pole is at -1. All cells were born during the measurement period. The dashed vertical line represents the cell center. Measurements are from 63 cells grown in M63 media at 37°C. The fraction of complexes observed in the older half of the cells was 0.45 which is statistically indistinguishable from an unbiased partitioning of complexes (p-value of the binomial test with *N* equal to the number of observed cells is 0.45).

The folded, spatial distribution of all complexes, corresponding to Figs. 2, *B* and *C* in the main text for LB, is presented in Fig. S2. Also shown are the results from the 'region detection' method.

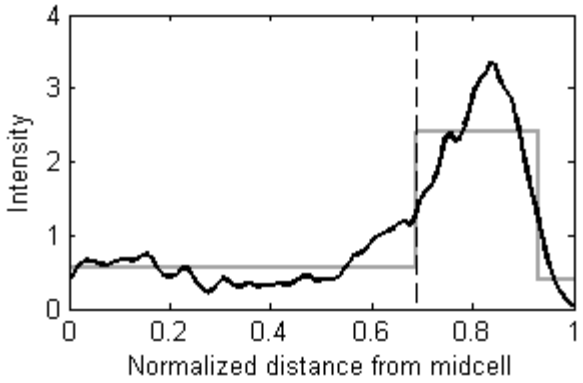

FIGURE S2 KDE of the spatial distribution of the fluorescence intensity (in arbitrary units) of complexes along the major axis of the cells, extracted from all cells and time points (*black line*, bandwidth 0.05). Complex positions were normalized by half the cell length. Also shown is the fit to a piecewise-constant probability density function by maximum likelihood (*gray line*). All cells were born during the measurement period. The vertical dashed lines represent the detected separation points between the midcell and poles. Measurements are from 221 cells grown in M63 media at 37°C, with separation point detected at 0.69.

2D Spatial distribution of complexes

We obtained the KDE of the 2D distributions of complexes from all time points in both temperature conditions. Results are shown in Figs. S3, *A* and *B*, for cells at 37°C and 24°C in LB, respectively.

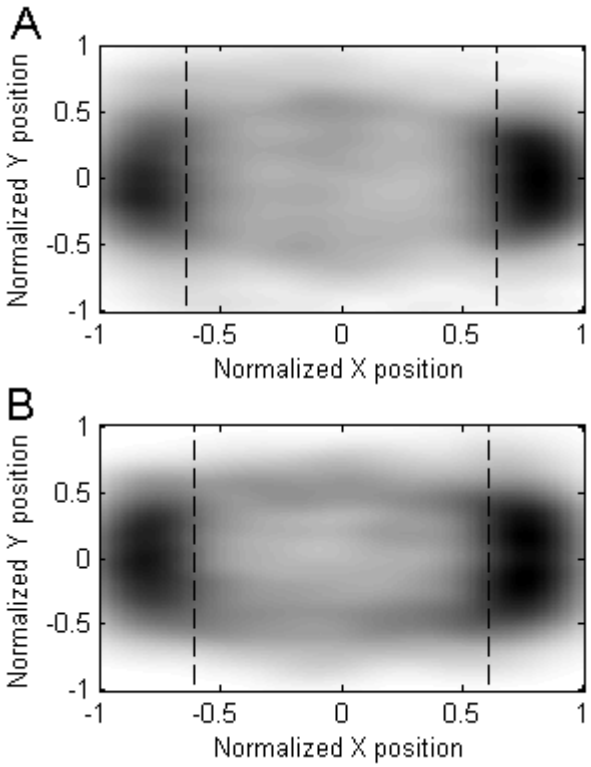

FIGURE S3 KDE of the 2D spatial distributions of complexes from all time points. All cells were born during the measurement period. Measurements are from (*A*) 531 cells at 37°C and (*B*) 372 cells at 24°C, both grown in LB media. Separation points (obtained from Figs. 2, *B* and *C*) are also shown.

Models of long-term spatial distributions of large molecules in the cytoplasm of *E. coli*

A cell is modeled as a 1-dimensional space, which is divided into N homogeneous sub-volumes, indexed from $[1, N]$. The motion of the complexes is modeled along the major cell axis by unimolecular reactions following the Reaction-Diffusion Master Equation (1). Collisions between complexes are not modelled. We define $\vec{a}(x)$ as the propensity of the forward reaction (modeling the motion of a complex from position x to position x+1) and $\overleftarrow{a}(x)$ as the propensity of the reverse reaction (from x to x-1).

These two propensities account for the combined effects of the cell geometry (rod shape and pole caps) and of the nucleoid on the displacements of the complexes. Let $\mathbf{P}(t)$ be the N-by-1 vector describing the probability of observing a complex in each sub-volume at time t. $\mathbf{P}(t)$ therefore evolves according to the following master equation, in matrix-vector form:

$$\frac{d\mathbf{P}(t)}{dt} = \mathbf{A}\mathbf{P}(t)$$

where

$$\mathbf{A}=\begin{bmatrix} -\vec{a}(1) & \overleftarrow{a}(2) & 0 & & 0 & 0 \\ \vec{a}(1) & -\overleftarrow{a}(2)-\vec{a}(2) & \overleftarrow{a}(3) & \cdots & 0 & 0 \\ 0 & \vec{a}(2) & -\overleftarrow{a}(3)-\vec{a}(3) & & 0 & 0 \\ & \vdots & & \ddots & & \vdots \\ 0 & 0 & 0 & & -\overleftarrow{a}(N-1)-\vec{a}(N-1) & \overleftarrow{a}(N) \\ 0 & 0 & 0 & \cdots & \vec{a}(N-1) & -\overleftarrow{a}(N) \end{bmatrix}$$

Since a complex can travel from any sub-volume in the cell to any other sub-volume, given enough time, the system is ergodic. Therefore, as $t\to\infty$, $\mathbf{P}(t)$ will converge to a unique solution, $\mathbf{P}_\infty$. This solution can be found by solving the linear system of equations $\mathbf{0}=\mathbf{A}\mathbf{P}_\infty$, with the constraint that the total probability must sum to 1. As this is the long-term spatial distribution of the complexes predicted by the model, this was the distribution we fit to the measurements.

In a model not accounting for the caps of the cells, the propensities of the forward and reverse diffusion reactions would be proportional to the diffusion constant of the complexes, *D*:

$$\vec{a}(x)=\overleftarrow{a}(x)=\frac{N^2D}{2}$$

To account for the rod shape, i.e. a cylinder capped with two half-spheres, the length of the cell was parameterized by $B\in[0,1]$, the normalized distance from midcell at which the cap begins. The forward propensities were attenuated by $\phi(x)$, the ratio between the areas of the cross sections of the cell (denoted $S(x)$) at adjacent positions. As such, $\overleftarrow{a}(x)$ remains the same and $\vec{a}(x)$ becomes:

$$\vec{a}(x)=\frac{N^2D}{2}\phi(x)$$

where

$$\phi(x)=\frac{S(x+1)}{S(x)}$$

$$S(x)=\begin{cases}\pi, & c(x)<B \\ \pi\left[1-\left(\dfrac{c(x)-B}{1-B}\right)^2\right], & c(x)\geq B\end{cases}$$

$$c(x)=\frac{x-0.5}{N}$$

Here, $c(x)$ translates the index of a sub-volume into the normalized distance from the midcell to the center of the sub-volume. In this case, $B = 1$ recovers the cylindrical cell from above, and $B = 0$ produces a spherical cell.

The effects of a nucleoid are introduced in the above model by adding a Gaussian function to $\vec{a}(x)$ while subtracting it from $\overleftarrow{a}(x)$. This anisotropy was parameterized with center $\mu \in [0,1]$, standard deviation σ, and height h. Specifically:

$$\vec{a}(x) = \frac{N^2 D}{2}\left[\phi(x) + h \cdot \exp\left\{-\frac{(c(x)-\mu)^2}{2\sigma^2}\right\}\right]$$

$$\overleftarrow{a}(x) = \frac{N^2 D}{2}\left[1 - h \cdot \exp\left\{-\frac{(c(x)-\mu)^2}{2\sigma^2}\right\}\right]$$

To fit the models to the measurements, we use the Earth-Mover's metric (2), otherwise known as the first Wasserstein metric (3), defined as:

$$W(F,G) = \int_{-\infty}^{\infty} |F(x) - G(x)| dx$$

where F and G are the cumulative distribution functions of the model and the measurements.

To obtain the fraction of complexes moving towards the pole from sub-volume x, we first initialized the model with all probability in sub-volume x, denoted $\mathbf{P}^x(0)$, and numerically integrated the system over one minute using the Matlab function `ode23s` to obtain the probability distribution at $t = 1$, $\mathbf{P}^x(1)$. This fraction was then calculated as:

$$\frac{1}{2}\mathbf{P}_x^x(1) + \sum_{i=x+1}^{N}\mathbf{P}_i^x(1)$$

Spatial distribution without anisotropy

We constructed a 1-dimensional model with the forward and backwards propensities of diffusion events set to be equal, and inversely proportional to the observed spatial distribution (results in Fig. S4). When both propensities are equal, the probability that a complex will travel in one direction is 0.5, and thus there is no velocity anisotropy. When the propensities are inversely proportional to the observed distribution, in the long term, the complexes tend to linger in the areas where they were observed with high probability. Thus, the long-term spatial distribution is exactly as observed (the lines are indistinguishable in Fig. S4 *B*), while producing a negligible anisotropy in the predicted displacement distribution (Fig. S4 *A*).

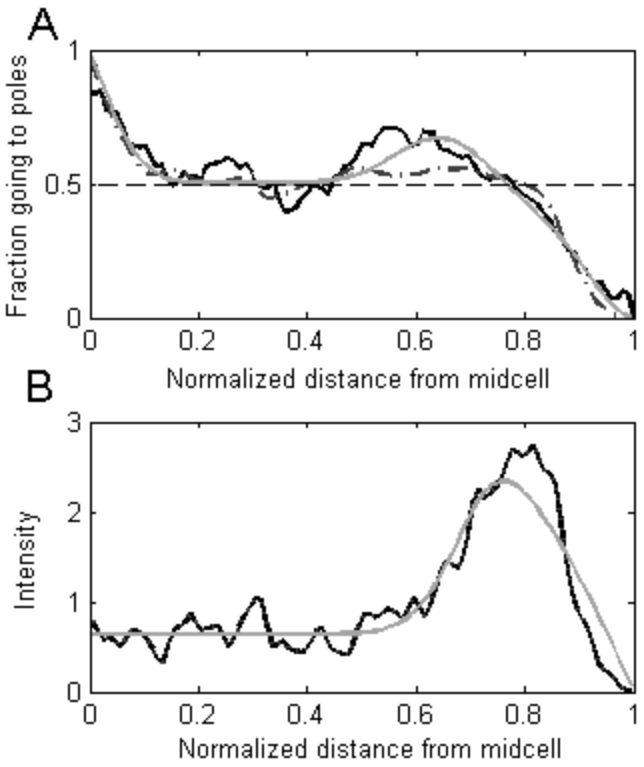

FIGURE S4 (*A*) Predicted fraction of complexes travelling towards the pole from each position along the major axis. (*B*) Long-term spatial distribution of complexes. Results are from a model with a localized anisotropy as in Fig. 8 (*gray lines*), a model with heterogeneous speeds (*dashed lines*), and the measurements at 24°C (*black lines*). Note that the dashed line in (*B*) is superimposed by the black line.

SUPPORTING REFERENCES

1. Gardiner, C.W., K.J. McNeil, D.F. Walls, and I.S. Matheson. 1976. Correlations in Stochastic Theories of Chemical Reactions. *J. Statiscal Phys.* 14: 307.

2. Rubner, Y., C. Tomasi, and L.J. Guibas. 2000. The Earth Mover's Distance as a Metric for Image Retrieval. *Int. J. Comput. Vis.* 40: 99–121.

3. Dobrushin, R.L. 1970. Prescribing a system of random variables by conditional distributions. *Theory Probab. Its Appl.* 15: 458.