Biophysical Journal, Volume *118*

Supplemental Information

Dynamic Crowding Regulates Transcription

Anne R. Shim, Rikkert J. Nap, Kai Huang, Luay M. Almassalha, Hiroaki Matusda, Vadim Backman, and Igal Szleifer

Supporting Material

Dynamic crowding regulates transcription

Anne R. Shim[1,2], Rikkert J. Nap[1,2], Kai Huang[1,2], Luay M. Almassalha[1], Hiroaki Matusda[1], Vadim Backman[1,2], Igal Szleifer[1,2,3]

[1]Department of Biomedical Engineering, Northwestern University, Evanston, IL, 60208, USA; [2]Chemistry of Life Processes Institute, Northwestern University, Evanston, IL, 60208, USA; [3]Department of Chemistry, Northwestern University, Evanston, IL, 60208, USA

Steady-state Crowding and Volume Fractions

Cellular kinetics are halted above $\phi \approx 0.5$ (1, 2); therefore, we consider crowding volume fractions within the range of ϕ=0-0.5. We begin all calculations with a time period of constant crowder volume fraction, $\phi = 0.3$, to verify that our results are due to crowding kinetics, not steady-state crowding. As relative mRNA production is constant during this time interval, we can assume that all subsequent changes in mRNA production are the result of crowding kinetics.

Copy number to gene concentration derivation:

The concentration of gene(s) is calculated as $\frac{Copy\ Number}{Cell\ Volume}$.

To convert to nM,

$$[C] = \frac{Copy\ Number}{Cell\ Volume\ x\ Avogadro's\ number\ x\ 10^9\ nanomoles\ per\ mole}.$$

The lowest copy number is that of a single gene with 1 copy per chromosome, or a copy number of 2. The highest copy number we consider is the entire population of genes. There is some debate over the total number of genes in the human genome; however, the number of protein coding genes have been estimated as roughly 20,000 (3). Each of these genes have between 2 and 100s of copies. Therefore, we estimate total gene copy number to be on the order of magnitude of 10^4 or 10^5. Likewise, there is a large range in nuclear volume, depending on cell type, disease state, etc. Typical cells have diameters between 2-10 μm, or volumes of 4.19-523.6 μm^3. Therefore, we consider [C] less than 1 and up to 51.

Reaction Rate Equations and Coefficients Explained

Each reaction rate equation (explained below in Section 3) is modified by the nuclear nanoenvironment by two competing, crowding-induced effects. Crowding alters the diffusion of proteins and contributes to the free energy of the nucleus, which determines the likelihood of protein binding. These effects were quantified and calculated by simulations, described below in Section 1 and Section 2, respectively.

1. **Brownian Dynamics Simulations of Diffusion Coefficients:** Diffusion coefficients were calculated for transcriptional elements (Table S1) using Brownian Dynamics (BD) simulations. These simulations are carried out using GROMACS version 2016.4. Each simulation was comprised of one spherical tracer particle (representing one of the transcriptional elements from Table S1) diffusing through an environment of spherical crowders (r = 3 nm). Complex sizes were determined theoretically, as explained in Section 3, and all other transcriptional element sizes were determined from literature. The results of BD simulated diffusion for spherical tracer particles of r=2-6 nm were previously published in Matsuda *et. al (1)*. These results, along with the data for particles of r = 18.38 nm are shown in Figure S9.

Table S1. Size of transcriptional elements

Particle	Radius
Transcription Factor	4 nm (4)
RNA Polymerase	5.4 nm (4)
Small Nuclear Ribonucleic Particle	6.25 nm (5)
Complex 3	7.25 nm
Complex 2	10.4 nm
mRNA	18.8 nm (6)

The number of crowders in each simulation ranged from 30 to 1140 in steps of 30 in a simulation box of $(63\ \text{nm})^3$. Spherical tracer particles freely diffused throughout the crowders and the mean-square displacement was calculated from t = 5 ns to t = 20 ns. The average slope of the mean-square displacement was calculated and the diffusion coefficient was determined to be one sixth of the average slope. The diffusion coefficients were fit by a cubic polynomial and normalized to diffusion at zero crowding. This cubic polynomial is:

$$\frac{D(\phi)}{D(0)} = 1 + \alpha\phi + \beta\phi^2 + \gamma\phi^3,$$

with the coefficients for previously unpublished tracer particles given in Table S3.

Table S2. Coefficients for a cubic fit of diffusion by particle radius

Particle	Radius	α	β	γ
mRNA	18.8 nm	-5.114	9.7329	-5.239

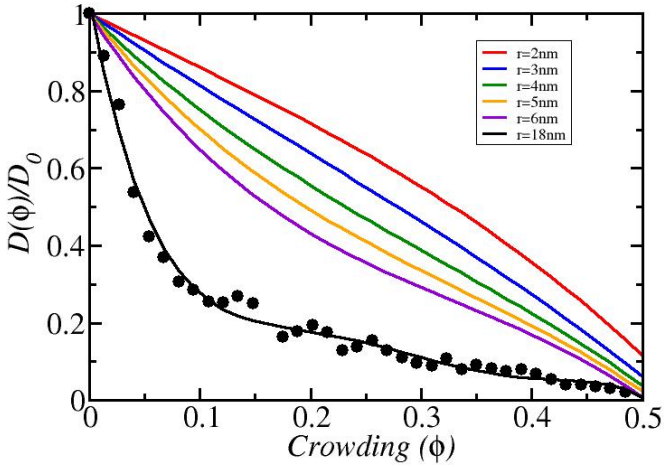

Fig S1. The diffusion coefficient of a tracer molecule is decreased by the presence of higher volume fractions of crowders (r = 3 nm). Previously published results (red–purple top to bottom represent r = 2-6 nm) represent protein diffusion. Black line and data points are r = 18.8 nm, which represents mRNA.

2. **Monte Carlo Simulations of Binding Free Energy Contributions:** The crowding-induced contributions to free energy were calculated with Monte Carlo (MC) simulations, written in C. MC simulations calculated $\Delta F_{crowd}(\phi)$, the free energy released during binding, and $\Delta F_{barrier}(\phi)$, the free energy barrier to association. Simulations were comprised of a spherical protein, a row of 50 overlapping spheres to represent DNA (r = 1 nm, spaced 1 nm apart), and crowders (r = 3 nm) in a $(50\ \text{nm})^3$ simulation box. Spherical proteins were the same size as those used in BD simulations (Table S2). MC moves were small, random translations of randomly selected crowders. Every 10 MC moves, a test move was considered, which both increases and decreases the distance between the protein and the DNA by 0.1 nm. The move is rejected if it causes an overlap between any species in the system and is accepted otherwise. Over the course of millions of MC moves, the probability of accepting a move that increases the distance between proteins, $p_{forward}$, or decreases the distance between proteins, $p_{backward}$, is calculated. This allows us to calculate the free energy change:

$$\beta\Delta F_{crowd}(x \to x + \Delta x) = \ln\left[\frac{p_{backward}}{p_{forward}}\right],$$

where $\beta = 1/k_BT$. $\beta\Delta F_{crowd}$ is fit to a polynomial. To calculate $\Delta F_{barrier}(\phi)$, the potential of mean force (PMF) is mapped between Δx and free energy. The PMF reaches a plateau value, which is the free energy required to separate the protein from the DNA against the depletion force caused by the crowders. This is equivalent to the negative contribution of crowding to the free energy of binding. $\Delta F_{barrier}(\phi)$ is calculated by subtracting this plateau value from the maximum of the PMF. A representative PMF is shown for species in Figures S7-S9 below.

$e^{\beta\Delta F_{crowd}}$ and $e^{\beta\Delta F_{barrier}}$ are input into the dynamical model. The free energy change was calculated for C_{II}, C_{III}, and snRNP (Figures S10-S12). Each MC simulation was repeated 3 times and averaged. The average change in free energy over these simulations was fit to a polynomial.

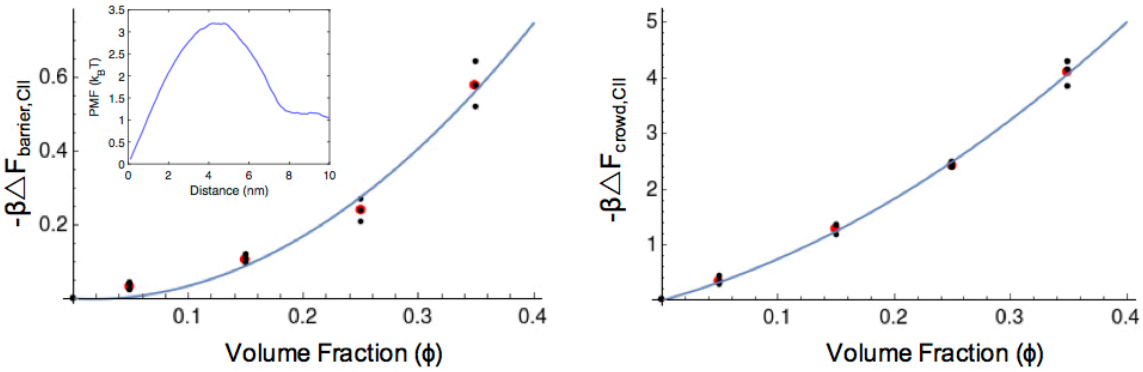

Fig S2. Crowding induced free energy barrier (left) and crowding induced change in free energy (right) associated with C_{II} complex formation or dissociation. Black symbols show the results of Monte Carlo simulations and red symbols are the average of the Monte Carlo simulation results. The solid line is the polynomial fit.

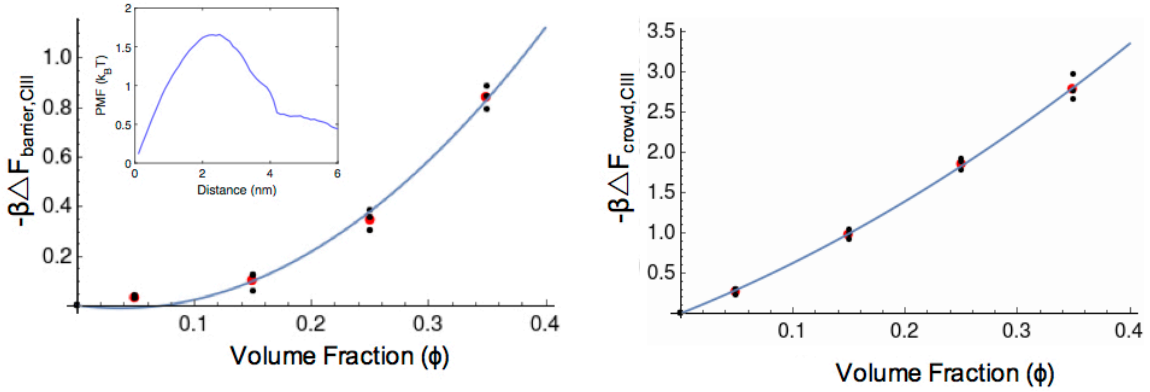

Fig S3. Crowding induced free energy barrier (left) and crowding induced change in free energy (right) associated with C_{III} complex formation or dissociation. Black symbols show the results of Monte Carlo simulations and red symbols are the average of the Monte Carlo simulation results. The solid line is the polynomial fit.

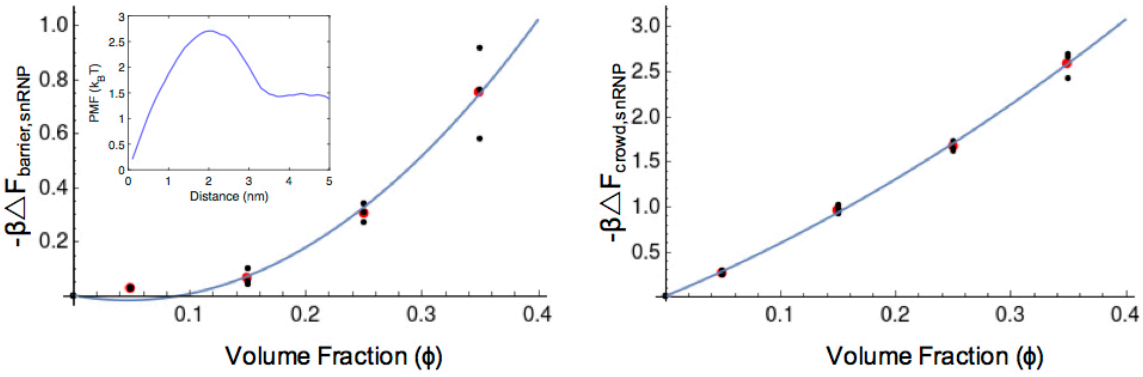

Fig S4. Crowding induced free energy barrier (left) and crowding induced change in free energy (right) associated with snRNP complex formation or dissociation. Black symbols show the results of Monte Carlo simulations and red symbols are the average of the Monte Carlo simulation results. The solid line is the polynomial fit.

3. **Reaction Rate Equations and Coefficients Explained:** The transcription cascade is described by reaction rate equations, whose coefficients describe their dependence on the physical environment (crowding). Also included in the coefficients are microscopic details, such as molecular geometries, interactions, and diffusion coefficients. All reaction rate coefficients are described after their respective equations. The following are true for all coefficients:
    - Crowding induced effects are based on the following geometries: RNA polymerase r = 5.4 nm, transcription factor r = 4 nm, DNA r = 1 nm, crowders r = 3 nm, with average molecular weight 67.7 kDa and specific volume 0.73 mL/g.
    - Based on Brownian Dynamics simulations (Section 1), the diffusion coefficient is reduced by $f(\phi, r) \equiv \frac{D(\phi,r)}{D(0,r)}$ and will be represented in equations as $f_{TF}(\phi), f_{RNA_p}(\phi), f_{snRNP}(\phi), or\ f_{mRNA}(\phi)$.
    - Based on Monte Carlo simulations (Section 2), the crowding-induced contribution to the binding free energy leads to $\Delta F = \Delta F_{\phi=0} + \Delta F_{crowd}(\phi)$, which influences the dissociation constant, K_D by: $K_D(\phi) = K_{D,\phi=0} e^{\beta \Delta F_{crowd}(\phi)}$.

Reaction rate equations adapted from Matsuda *et. al* (1).

Reversible reactions

Nonspecific reactions: For the rate of association of TF to DNA (k_t^{ns}) and RNAp to DNA (k_f^{ns}), crowding affects the diffusion of the transcriptional proteins and the kinetic barrier to association between proteins and DNA:

$$k_t^{ns}(\phi) = k_{t,0}^{ns} \cdot f_{TF}(\phi) \cdot e^{-\beta \Delta F_{barrier,TF}(\phi)},$$
$$k_f^{ns}(\phi) = k_{f,0}^{ns} \cdot f_{RNA_p}(\phi) \cdot e^{-\beta \Delta F_{barrier,RNA_p}(\phi)}.$$

The rate of dissociation of transcriptional proteins and DNA is equal to the rate of association multiplied by the dissociation constant. The dissociation constant depends on the change in free energy that occurs during dissociation:

$$K_{D,TF}^{ns}(\phi) = K_{D,TF,\phi=0}^{ns} \cdot e^{-\beta \Delta F_{crowd,TF \cdot DNA}(\phi)},$$
$$K_{D,RNA_p}^{ns}(\phi) = K_{D,RNA_p,\phi=0}^{ns} \cdot e^{-\beta \Delta F_{crowd,RNA_p \cdot DNA}(\phi)}.$$

Therefore, the rate of nonspecific dissociation is given by:

$$k_o^{ns}(\phi) = k_{o,0}^{ns} \cdot f_{TF}(\phi) \cdot e^{-\beta \Delta F_{crowd,TF \cdot DNA}(\phi)} \cdot e^{-\beta \Delta F_{barrier,TF}(\phi)},$$
$$k_b^{ns}(\phi) = k_{b,0}^{ns} \cdot f_{RNA_p}(\phi) \cdot e^{-\beta \Delta F_{barrier,RNA_p}(\phi)} \cdot e^{-\beta \Delta F_{crowd,RNA_p \cdot DNA}(\phi)}.$$

Specific reactions: The rate of specific association is due to facilitated diffusion, adapted from Berg *et. al* (7).

Specific association depends on: $k_t = V \frac{(D_{1,TF} \cdot k_o^{ns})^{\frac{1}{2}}}{L}$, where

V = volume of the nucleus
$D_{1,TF}$= one dimensional diffusion coefficient
k_o^{ns} = nonspecific dissociation
L = half of the total length of DNA.

Based on this relation, specific association is proportional to f_{TF} (both $D_{1,TF}$ and k_o^{ns} are proportional to f_{TF} and are each taken to the ½ power) and proportional to the free energy barrier to association by a factor of $e^{1/2}$. Therefore, the forward rates of specific association are:

$$k_t(\phi) = k_{t,0} \cdot f_{TF}(\phi) \cdot e^{\frac{1}{2}\beta \Delta F_{crowd,TF \cdot DNA}(\phi)} \cdot e^{-\frac{1}{2}\beta \Delta F_{barrier,TF}(\phi)},$$
$$k_f(\phi) = k_{f,0} \cdot f_{RNA_p}(\phi) \cdot e^{-\frac{1}{2}\beta \Delta F_{barrier,RNA_p}(\phi)} \cdot e^{\frac{1}{2}\beta \Delta F_{crowd,RNA_p \cdot DNA}(\phi)} \cdot e^{-\beta \Delta F_{barrier,RNA_p}^{slide}(\phi)}.$$

$e^{-\beta \Delta F_{barrier,RNA_p}^{slide}(\phi)}$ is the entropic gain in free energy due to C_{II} formation.

The rate of specific dissociation is given by:

$$\frac{K_{dis}}{K_{as}} = [D]_{tot}\, x\, \frac{K_D}{K_D^{ns}}\,.$$

Here, $[D]_{tot}\, x\, \frac{K_D}{K_D^{ns}}$ becomes $K_{dis,0}$ and the rate of specific dissociation becomes:

$$k_o(\phi) = k_{o,0} \cdot f_{TF}(\phi) \cdot e^{\frac{1}{2}\beta\Delta F_{crowd,TF\cdot DNA}(\phi)} \cdot e^{-\frac{1}{2}\beta\Delta F_{barrier,TF}(\phi)},$$

$$k_b(\phi) = k_{b,0} \cdot f_{RNA_p}(\phi) \cdot e^{-\frac{1}{2}\beta\Delta F_{barrier,RNA_p}(\phi)} \cdot e^{\beta\Delta F^{slide}_{crowd,RNA_p\cdot TF}(\phi)} \cdot e^{-\beta\Delta F^{slide}_{barrier,RNA_p}(\phi)}.$$

$e^{\beta\Delta F^{slide}_{crowd,RNA_p\cdot TF}(\phi)}$ is the crowding-induced free energy change due to RNAp and TF contact.

Reaction Rate Coefficients at Zero Crowding

The original reaction rate coefficients (i.e. $k_{t,0}^{ns}$, $k_{t,0}$, etc.) were determined by the relations developed by Berg *et. al (7)* and take into account geometric information such as the length of one base pair, the distance between DNA strands, the radius of DNA etc. The original values of these reaction rates are explained in Matsuda *et. al* (1).

These rate coefficients do not incorporate crowding effects for the rates of irreversible reactions. With the addition of crowding effects into these reactions, the rate coefficients are now non-static and become a range of values, dependent on the volume fraction of crowders. Therefore, we set the reaction rate coefficients from Matsuda *et. al* as the average reaction rate coefficient of the crowding-dependent coefficient from ϕ=0-0.5. Each initial reaction rate coefficient for the irreversible reactions is determined by the relations:

$$k_{m,0} = k^2_{m,matsuda}/(\frac{1}{0.5-0} * \int_0^{0.5} k_{m,matsuda} * e^{-\beta\Delta F_{barrier,TF}} d\phi)\,,$$

$$k_{M,0} = k^2_{M,matsuda}/(\frac{1}{0.5-0} * \int_0^{0.5} k_{M,matsuda} * e^{-\beta\Delta F_{barrier,snRNP}} d\phi),$$

$$k_{M\prime,0} = k^2_{M\prime,matsuda}/(\frac{1}{0.5-0} * \int_0^{0.5} k_{M\prime,matsuda} * e^{-\beta\Delta F_{barrier,CIII}} d\phi),$$

$$\gamma_0 = \gamma^2_{0,matsuda}/(\frac{1}{0.5-0} * \int_0^{0.5} \gamma_{0,matsuda} * f_{mRNA}\, d\phi).$$

Reaction Rate Equations Not Adapted from Matsuda *et. al* (1).

Irreversible Reactions

The rate of transcription (k_m) is influenced by the barrier to dissociation of C_{II}:

$$k_m(\phi) = k_{m,0} \cdot e^{-\beta\Delta F_{barrier,C_{II}}(\phi)}.$$

Subsequently, pre-mRNA processing occurs in two steps. First, the association of snRNP and pre-mRNA (k_M) depends on the diffusion of snRNP, as well as the barrier of association between snRNP and pre-mRNA:

$$k_M(\phi) = k_{M,0} \cdot f_{snRNP}(\phi) \cdot e^{-\beta\Delta F_{barrier,snRNP}(\phi)}.$$

After association, processing of pre-mRNA to mRNA ($k_{M'}$) depends on the barrier to dissociation of C_{III}:

$$k_{M\prime}(\phi) = k_{M\prime,0} \cdot e^{-\beta\Delta F_{barrier,C_{III}}(\phi)}.$$

As the final step, mRNA in the nucleus diffuses to the cytoplasm (γ), which depends only on the crowding effects on diffusion of mRNA:

$$\gamma(\phi) = \gamma_0 \cdot f_{mRNA}(\phi).$$

Reaction Rate Coefficients at Zero Crowding

The original reaction rate coefficients were determined by the relations developed by Berg *et. al (7)* and take into account the following geometries:

Complex 2: C_{II} is a complex of DNA (r=1 nm), RNA polymerase (r=5.4nm), and a transcription factor (r=4nm). Therefore, this complex would have an average radius of 10.4 nm.

Small Nuclear Ribonucleic Particle: There are many types of snRNPs, so we based our study on the U1 snRNP, which is well characterized (5). The U1 snRNP is a complex with a radius of roughly 6.25 nm. Data for r=6nm was used to represent snRNPs.

Complex 3: C_{III} is a complex of a small nuclear ribonucleic particle (r=6.25 nm) and mRNA (r=1 nm for linear mRNA). Therefore, this complex would have an average radius of 7.25 nm.

mRNA: mRNA transcripts vary in size. The radius of gyration of mRNA can be as low as 16.8-20.8 nm; therefore, we used r=18.8nm as an average radius for mRNA in solution.

Table S3: Numerical values of model parameters

Parameter	Description	Value (with ϕ=0)
k_t^{ns}	Association rate constant for nonspecific TF-DNA binding	4.9 x 10^4 $mM^{-1}s^{-1}$
k_f^{ns}	Association rate constant for nonspecific RNAp-DNA binding	3.6 x 10^4 $mM^{-1}s^{-1}$
k_o^{ns}	TF-DNA nonspecific dissociation rate	4.9 x 10^4 s^{-1}
k_b^{ns}	RNAp-DNA nonspecific dissociation rate	3.6 x 10^4 s^{-1}
$K_{D,TF}^{ns}$	Dissociation constant for nonspecific TF-DNA binding	1 mM
K_{D,RNA_p}^{ns}	Dissociation constant for nonspecific RNAp-DNA binding	1 mM
$K_{D,TF}$	Dissociation constant for specific TF-DNA binding	1.0 x 10^{-6} mM
K_{D,RNA_p}	Dissociation constant for specific RNAp-DNA binding	1.0 x 10^{-6} mM
k_t	Association rate constant for TF-promoter (O) binding	5.0 x 10^{-4} $mM^{-1}s^{-1}$
k_f	Association rate constant for RNAp-Complex I binding	3.0 x 10^{-4} $mM^{-1}s^{-1}$
k_o	TF-promoter (O) dissociation rate	1.0 s^{-1}
k_b	RNAp-Complex I dissociation rate	0.6 s^{-1}
$K_{D,TF}$	Dissociation constant for TF-O (promoter) binding	1.0 x 10^{-6} mM
K_{D,RNA_p}	Dissociation constant for RNAp-O (promoter) binding	1.0 x 10^{-6} mM
k_m	Rate of pre-mRNA production	0.02 s^{-1}
k_M	Intron splicing rate	64.5 $mM^{-1}s^{-1}$
$k_{M'}$	mRNA creation in the nucleus	.004 s^{-1}
γ	Nuclear exportation rate of mRNA	2 x 10^{-3} s^{-1}
ν	mRNA degradation rate	3 x 10^{-4} s^{-1}
$[DNA]$	Concentration of DNA	20 mM
$[RNA_p]$	Concentration of RNA polymerases	3.0 x 10^{-6} mM
$[O]$	Concentration of promoters	3.0 x 10^{-6} mM
$[TF]$	Concentration of transcription factors	3.0 x 10^{-6} mM

Supporting Figures

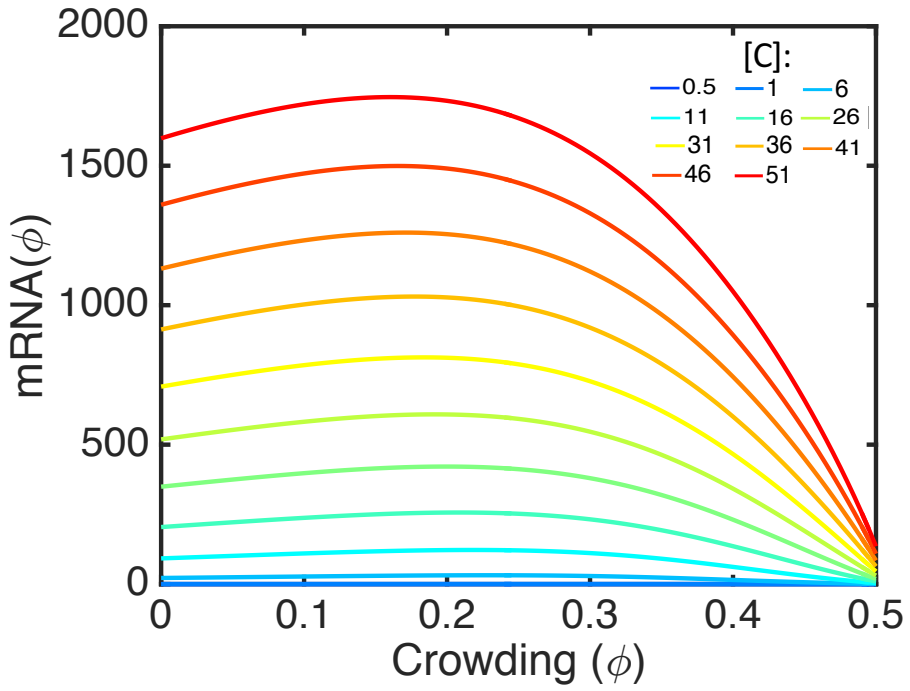

Fig S5. Genes with higher concentrations have more overall expression than genes with lower concentrations. As gene concentration increases, genes become relatively less sensitive to crowding; the high expression at $\boldsymbol{\phi = 0}$ causes relatively lower response in expression relative to $\boldsymbol{\phi = 0}$.

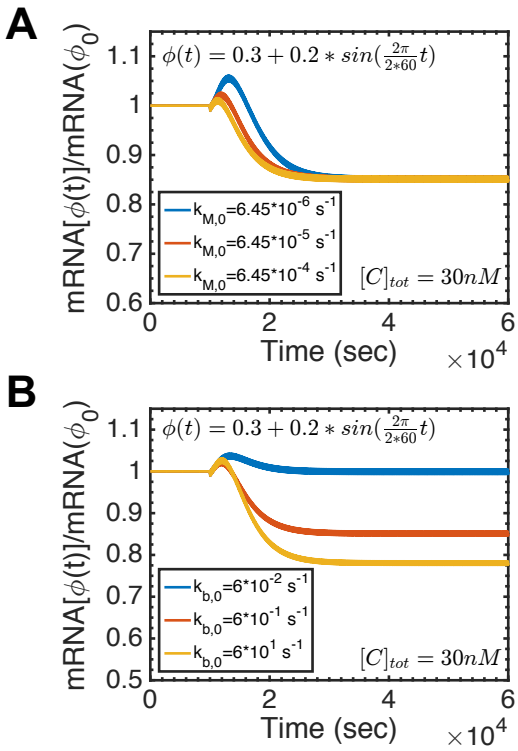

Fig. S6. Changing the initial rate coefficients of the chemical reactions confirms that the reaction rates k_M, $k_{M'}$, and γ, the intermediate, irreversible, crowding-dependent reaction rates, do not affect the level of expression at oscillating stability. For example, changing k_M alters the short-term expression during the transition state (A). This independence from k_M, $k_{M'}$, and γ was analytically determined under steady-state conditions. At steady-state, $\upsilon[mRNA_{cyto}] = \gamma[mRNA_{nuc}] = k_{M'}[C_{III}] = k_M[snRNP][pm] = k_m[C_{II}]$. Therefore, $[mRNA_{cyto}] = \frac{k_m}{\upsilon}[C_{II}]$ and k_M, $k_{M'}$, and γ do not affect $[mRNA_{cyto}]$. However, no such analytical solution exists when not at steady state. Conversely, mRNA expression at oscillating stability is determined by the remaining, crowding-dependent reaction rates. For example, k_b changes the expression level at oscillating stability by up to 30%, but has a much less dramatic effect on short-term expression (B).

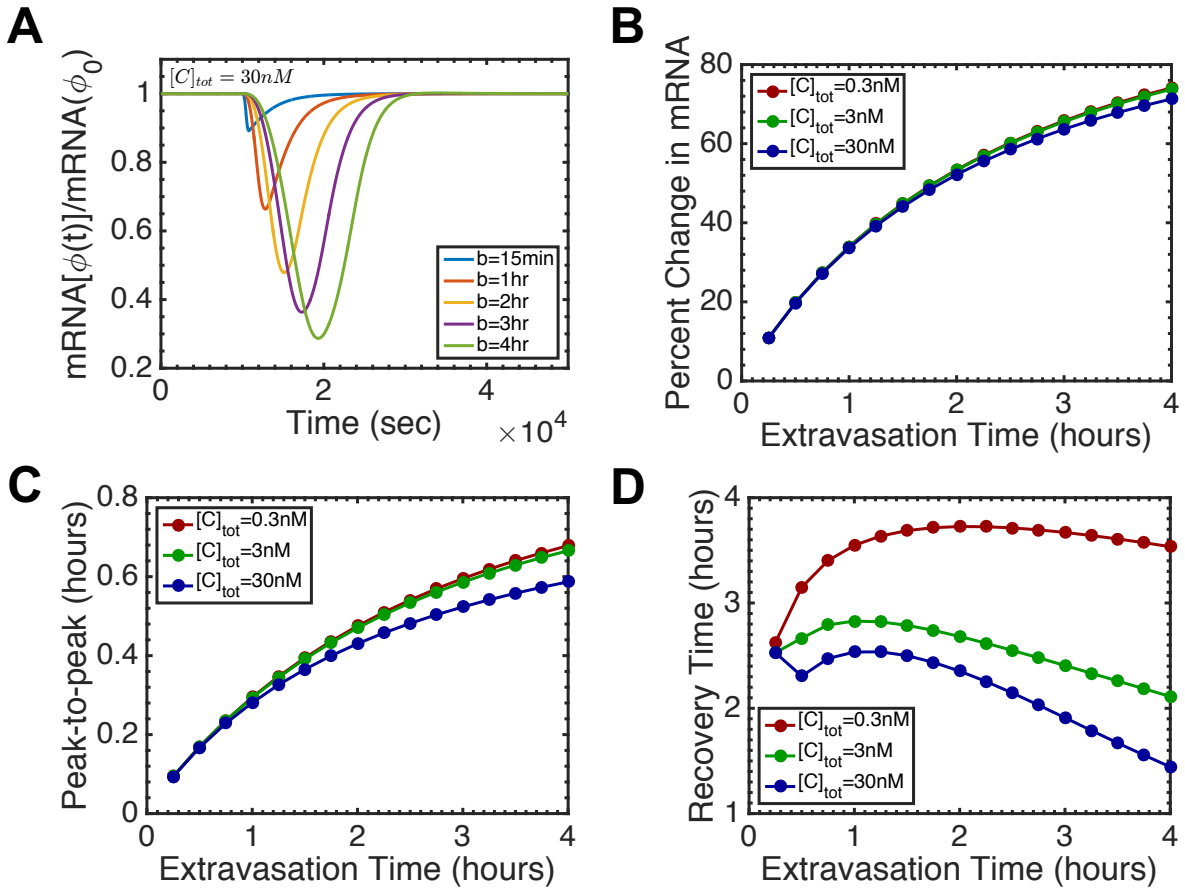

Fig. S7: The force of extravasation (A=0.2, $\phi_0 = 0.3$) causes universal downregulation of gene expression, for all types of genes and lengths of extravasation (A, B). However, as extravasation time increases, the percent change in mRNA becomes less sensitive to lengthening the extravasation time (B). This is independent of gene concentration, as all gene concentrations have approximately the same change in mRNA at each extravasation time. The time between maximum compression and minimal expression (peak-to-peak) is also insensitive to gene concentration for short extravasations (C). However, the peak-to-peak distance for genes with high concentration deviates at extravasation times greater than one hour. In contrast, the recovery time (the time from the end of extravasation to the return of expression to within 99% of the original expression level) is highly sensitive to gene concentration (D). Genes with low concentration require up to two-fold higher recovery times than genes with high concentration.

Supporting References

1. Matsuda, H., G. G. Putzel, V. Backman, and I. Szleifer. 2014. Macromolecular Crowding as a Regulator of Gene Transcription. Biophysical Journal 106:1801-1810.
2. Kang, H., Y.-G. Yoon, D. Thirumalai, and C. Hyeon. 2015. Confinement-Induced Glassy Dynamics in a Model for Chromosome Organization. Physical Review Letters 115(19).
3. Willyard, C. 2018. New human gene tally reignites debate. Nature 558:354-355.
4. Maeshima, K., K. Kaizu, S. Tamura, T. Nozaki, T. Kokubo, and K. Takahashi. 2015. The physical size of transcription factors is key to transcriptional regulation in chromatin domains. Journal of Physics: Condensed Matter 27(064116).
5. Stark, H., P. Dube, R. Lührmann, and B. Kastner. 2001. Arrangement of RNA and proteins in the spliceosomal U1 small nuclear ribonucleoprotein particle. Nature 409:539-542.
6. Milo, R., P. Jorgensen, U. Moran, G. Weber, and M. Springer. 2010. BioNumbers—the database of key numbers in molecular and cell biology. Nucleic Acids Research 38(D750-D753).
7. Berg, O. G., R. B. Winter, and P. H. v. Hippel. 1981. Diffusion-driven mechanisms of protein translocation on nucleic acids: Model and theory. Biochemistry 20:6929-6948.