Macromolecular Crowding as a Regulator of Gene Transcription

Hiroaki Matsuda,[†] Gregory Garbès Putzel,[†] Vadim Backman,[†] and Igal Szleifer[†‡*]
[†]Department of Biomedical Engineering and Chemistry of Life Processes Institute and [‡]Department of Chemistry, Northwestern University, Evanston, Illinois

ABSTRACT Studies of macromolecular crowding have shown its important effects on molecular transport and interactions in living cells. Less clear is the effect of crowding when its influence is incorporated into a complex network of interactions. Here, we explore the effects of crowding in the cell nucleus on a model of gene transcription as a network of reactions involving transcription factors, RNA polymerases, and DNA binding sites for these proteins. The novelty of our approach is that we determine the effects of crowding on the rates of these reactions using Brownian dynamics and Monte Carlo simulations, allowing us to integrate molecular-scale information, such as the shapes and sizes of each molecular species, into the rate equations of the model. The steady-state cytoplasmic mRNA concentration shows several regimes with qualitatively different dependences on the volume fraction, ϕ, of crowding agents in the nucleus, including a broad range of parameter values where it depends nonmonotonically on ϕ, with maximum mRNA production occurring at a physiologically relevant value. The extent of this crowding dependence can be modulated by a variety of means, suggesting that the transcriptional output of a gene can be regulated jointly by the local level of macromolecular crowding in the nucleus, together with the local concentrations of polymerases and DNA-binding proteins, as well as other properties of the gene's physical environment.

INTRODUCTION

Systems biology aims to understand cellular processes in terms of networks of interactions among molecules. Ideally, this view of biology would be global and yet still rest on a solid reductionist foundation, since each node of a network would represent a chemical reaction that could be isolated and studied in detail. However, it should be appreciated that this basic framework does not fully capture the complexities of life. The reaction represented by each node in a network may, for example, be influenced by the physical environment of some part of the cell.

One form of such a physical influence has received a great deal of attention from theorists and experimentalists alike: macromolecular crowding (1). The crowded nature of cellular environments exerts an important influence on the thermodynamics and kinetics of reactions. The free energy, as well as the rates (2–4), of a reaction will depend on the overall concentration of molecules; these could then play an important role as crowding agents even if they have no specific role in the reaction. Much work has been done to understand the influence of macromolecular crowding on specific types of processes, such as protein-protein binding (5), protein folding and stability (6), and chromatin compaction (7,8). However, little is known about the global influence of crowding on the scale of networks of interactions. From the theoretical side, an important first step in this direction was taken by Morelli et al. (9), who incorporated the influence of crowding on the rates of reactions into simple gene regulatory networks and found that the level of crowding had an important effect on steady-state protein concentrations.

In this work, we study a more complicated model of gene expression, incorporating the fact that in eukaryotes, RNA polymerase is recruited to a gene promoter via its interactions with transcription factors. We explicitly take into account the movement of DNA-binding proteins by the mechanism of facilitated diffusion (10), whereby molecules undergo free diffusion interspersed with periods of one-dimensional diffusion that occurs while the molecule is nonspecifically bound to DNA. Of most importance, we determine the effects of macromolecular crowding on the reaction rates in the model based on explicit assumptions about molecular shapes and sizes. Specifically, the effects of crowding on the diffusion coefficients of molecules are established using Brownian dynamics (BD)simulations, and the crowding-induced contributions to the binding free energies between molecules are calculated from Monte Carlo simulations. We solve the reaction-rate equations of the model in the steady state, including the dependence of the reaction rates on the volume fraction, ϕ, of crowders as determined from the simulations, and explore the effect of crowding in the cell nucleus on the steady-state concentration of cytoplasmic mRNA (see Fig. 1 for an illustration of our approach). Our incorporation of molecular-scale simulation results into the model allows us to correctly assess the order of magnitude of macromolecular crowding effects.

We show that there are several regimes of nuclear crowding dependence of steady-state cytoplasmic mRNA concentration, depending on the concentrations of the reactants as

Submitted September 30, 2013, and accepted for publication February 12, 2014.

*Correspondence: igalsz@northwestern.edu

Hiroaki Matsuda and Gregory Garbès Putzel contributed equally to this work.

Editor: Alan Grodzinsky.

0006-3495/14/04/1801/10 $2.00

http://dx.doi.org/10.1016/j.bpj.2014.02.019

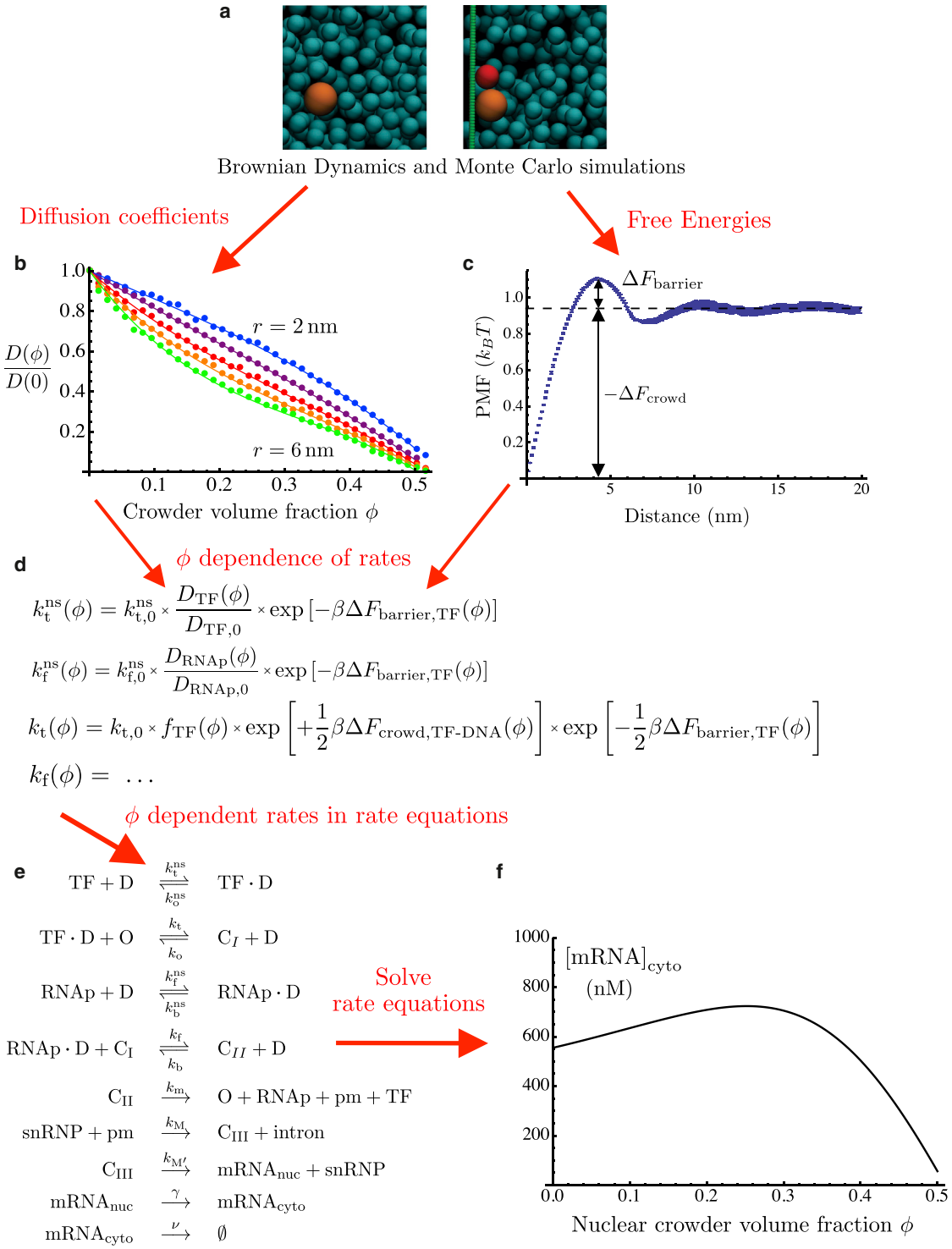

FIGURE 1 Reaction network model of gene expression incorporating molecular simulation data. (*a*) Representative snapshots of BD (*left*) and Monte Carlo (*right*) simulations used to determine diffusion coefficients and free energies, respectively. TF, RNAp, and DNA are represented in red, orange, and green, respectively. Snapshots were made using VMD (34). (*b*) Normalized diffusion coefficients calculated from BD simulations. (*c*) Potential of mean force between TF and DNA, calculated from Monte Carlo simulations at nuclear crowder volume fraction $\phi = 0.25$. (*d*) Representative formulas showing the dependence of reaction rates on the volume fraction, ϕ, of crowders in the nucleus. (*e*) Reactions in the model of gene expression. (*f*) Representative plot of nuclear crowder volume fraction dependence of steady-state cytoplasmic mRNA concentration. To see this figure in color, go online.

well as other parameters of the system, such as specific or nonspecific protein-DNA binding affinities. In many cases, we have found that the mRNA levels depend nonmonotonically on the volume fraction, ϕ, of crowders in the nucleus, reaching a maximum at physiologically relevant values near $\phi \approx 0.3$. This dependence may be accentuated by several means, for example, by decreasing the concentrations of the reactants or by increasing the binding affinity of nonspecific protein-DNA binding. The picture that emerges is one in which concentrations of transcription factors, polymerases, and active gene promoters, combined with the overall level of macromolecular crowding, may exert a quantitatively significant and qualitatively nontrivial influence on the level of expression of genes, with potentially important implications for the regulation of gene expression. In particular, the joint dependence of mRNA output on the level of nuclear crowding and the concentrations of transcription factors and RNA polymerases suggests that in the spatially nonuniform cell nucleus, genes in different locations will experience different effects of macromolecular crowding.

MODEL

The physical modeling of gene regulatory systems as networks of chemical reactions (11) has reached the stage where it is not out of the question to make quantitative comparisons with experiments, at least in the case of bacterial systems (12). As these comparisons become more common it will be increasingly important to incorporate into these models the influence of physical environments occurring in vivo rather than in vitro. In this regard, a key step was taken by Morelli et al. (9) when they considered simple models of gene regulatory networks, taking into account the influence of macromolecular crowding on the reaction rates. Here, we develop a multiscale model of gene expression involving multiple molecular binding events and incorporating microscopic information on the kinetic and thermodynamic effects of crowding. To approach the problem of gene regulation in eukaryotes, where gene regulation involves numerous combinatorial interactions of proteins (13), our model includes both transcription factors and RNA polymerases. The most important element of our approach is our molecular-level treatment of the crowding dependence of the reaction rates. The extent to which crowding slows down diffusion, as well as the way it induces entropic interactions between reactants, is determined from molecular simulations using a simple but consistent choice of shapes and sizes for each molecular species. The overall approach, in which molecular simulations provide the crowding dependence of the parameters in a set of rate equations, is illustrated in Fig. 1.

Rate equations and their steady-state solution

The system of reactions shown in Fig. 1 *e* summarizes the model considered here. In this section, we set out the steady-state equations for these reactions. Molecular-level formulas for the reaction rates will be given in the next section and will allow us to incorporate the effects of macromolecular crowding in the nucleus.

The first two reactions in Fig. 1 *e* represent so-called facilitated diffusion (10). Freely diffusing transcription factors (TFs) may bind transiently and nonspecifically to DNA. While a transcription factor is nonspecifically bound, it diffuses along DNA and may encounter its binding site (O) at a gene promoter, forming a complex we will call C_I. In a similar way, RNA polymerase (RNAp) binds nonspecifically to DNA and specifically to C_I; that is, the TF recruits RNAp to the promoter, forming a transcription-ready complex we call C_{II}. The reactions thus far are all reversible. A transcription-ready complex C_{II} may then undergo transcription initiation with a certain probability, freeing the O, TF, and RNAp and giving rise to a pre-mRNA (pm). The pm undergoes splicing, where it reacts with a small nuclear ribonucleic particle (snRNP) to form Complex III (C_{III}), and an intron, having been spliced out, is released. Further mRNA processing steps are represented by a reaction in which C_{III} gives rise to an mRNA molecule in the nucleus ($mRNA_{nuc}$) as well as the released snRNP. The two final reactions represent the export of mRNA from the nucleus and its degradation in the cytoplasm.

We note that all the steps after the formation of Complex II are modeled as being irreversible. As a consequence, the details of these steps will have no influence on the steady-state level of mRNA production. Indeed, in the steady state,

$$\begin{aligned}\nu[\mathrm{mRNA_{cyto}}] &= \gamma[\mathrm{mRNA_{nuc}}] = k_{M'}[\mathrm{C_{III}}] \\ &= k_M[\mathrm{snRNP}][\mathrm{pm}] = k_m[\mathrm{C_{II}}],\end{aligned} \tag{1}$$

and if we wish to calculate the steady-state concentration of mRNA in the cytoplasm, we need only find the concentration of C_{II}:

$$[\mathrm{mRNA_{cyto}}] = \frac{k_m}{\nu} \times [\mathrm{C_{II}}]. \tag{2}$$

Furthermore, the steady-state flux of mRNA into the cytoplasm is simply

$$\nu \times [\mathrm{mRNA_{cyto}}] = k_\mathrm{m}[\mathrm{C_{II}}]. \tag{3}$$

The steady-state concentration of complex C_{II} can be obtained by numerically solving the coupled equations

$$\frac{d[\mathrm{C_I}]}{dt} = k_\mathrm{t}[\mathrm{TF\cdot D}][\mathrm{O}] - k_\mathrm{o}[\mathrm{C_I}] - k_\mathrm{f}[\mathrm{RNAp\cdot D}][\mathrm{C_I}] + k_\mathrm{b}[\mathrm{C_{II}}] \tag{4}$$

$$\frac{d[\mathrm{C_{II}}]}{dt} = k_\mathrm{f}[\mathrm{RNAp\cdot D}][\mathrm{C_I}] - (k_\mathrm{b} + k_\mathrm{m})[\mathrm{C_{II}}]. \tag{5}$$

To do so, however, the quantities [O], [TF · D], and [RNAp · D] must first be given in terms of the concentrations of complexes C_I and C_{II}:

$$[O] = [O]_{tot} - [C_I] - [C_{II}]$$

$$[TF\cdot D] = \frac{\left([D]_{tot}\Big/K^{ns}_{D,TF}\right)\left([TF]_{tot}-[C_I]-[C_{II}]\right)-\frac{k_m}{k^{ns}_o}[C_{II}]}{1+\left([D]_{tot}\Big/K^{ns}_{D,TF}\right)}$$

$$[RNAp\cdot D] = \frac{\left([D]_{tot}\Big/K^{ns}_{D,RNAp}\right)\left([RNAp]_{tot}-[C_{II}]\right)-\frac{k_m}{k^{ns}_b}[C_{II}]}{1+\left([D]_{tot}\Big/K^{ns}_{D,RNAp}\right)}.$$

The latter two expressions are derived from the steady-state solutions of the reaction-rate equations for the concentrations of [TF·D] and [RNAp·D], respectively. They make use of the total concentrations $[TF]_{tot}$, $[RNAp]_{tot}$, and $[O]_{tot}$, as well as dissociation constants for nonspecific binding. In deriving them, we have made the assumption that the total concentration of DNA basepairs is very large compared to the concentration of TF, so that we will always have $[D] \approx [D]_{tot}$.

To summarize, solving Eqs. 4 and 5 gives us the concentration of transcription-ready complexes C_{II}. We may then use Eq. 2 to compute the concentration of cytoplasmic mRNA or Eq. 3 to compute the rate of mRNA export to the cytoplasm. Having now determined the level of mRNA expression in terms of the reaction rates, our next goal is to use molecular-scale expressions for these rates to incorporate the influence of macromolecular crowding.

Reaction rates: facilitated diffusion

Macromolecular crowding and other molecular-scale physics enter into the model via the reaction rates. In this section, we use the microscopic theory of facilitated diffusion developed by Berg and co-workers (10,14,15) to give expressions for the reaction rates of specific and nonspecific binding processes involving TFs and RNAps. We do not concern ourselves with the rates of the irreversible reactions after pm production, such as splicing and mRNA export, since they do not influence the steady-state results. It is important to keep in mind, however, that they are crucial for determining the dynamics and temporal correlations of the mRNA output (16), which is not a subject of study in this work.

Assuming that nonspecific binding of TF and RNAp to DNA is diffusion-limited, we use for forward rates k_1 and k_3 the expression (10,14,15)

$$k^{ns}_t = \frac{2\pi D_{TF} l}{\ln(\xi/2b)} \quad k^{ns}_f = \frac{2\pi D_{RNAp} l}{\ln(\xi/2b)}. \tag{6}$$

These rates depend on the diffusion coefficients D_{TF} and D_{RNAp} of TF and RNAp, as well as on three length scales, l, b, and ξ. These are, respectively, the length along the DNA of one basepair, the radius of the DNA molecule (viewed as an approximate cylinder), and a correlation length giving the characteristic distance between DNA strands.

The nonspecific dissociation rates follow from the dissociation constants for nonspecific binding, called $K^{ns}_{D,TF}$ and $K^{ns}_{D,RNAp}$:

$$k^{ns}_o = K^{ns}_{D,TF} \times k^{ns}_t \quad k^{ns}_b = K^{ns}_{D,RNAp} \times k^{ns}_f. \tag{7}$$

The association rate constants for specific binding of TF and RNAp are given by the expression derived by Berg et al. (10) for specific protein-DNA binding by facilitated diffusion:

$$\begin{aligned} k_t &= V \times \left(D_{1,TF} \times k^{ns}_o\right)^{1/2}/L \\ k_f &= V \times \left(D_{1,RNAp} \times k^{ns}_b\right)^{1/2}/L \end{aligned} \tag{8}$$

Here, $D_{1,TF}$ and $D_{1,RNAp}$ are the one-dimensional diffusion coefficients of TF and RNAp when these are nonspecifically bound to DNA, and L is one-half of the total length of DNA in the nucleus. The factor V, representing the volume of the nucleus, does not appear in the expression of Berg et al. (10). This is due to the fact that in the reaction equations (Eqs. 4 and 5), we take the forward rates to multiply the product of the volume densities (concentrations) of both reagents in such a way that the association rate constants have the usual dimensions of concentration^{-1} × time^{-1}.

The backward rates for specific binding are determined from (10)

$$\frac{k_o}{k_t} = [D]_{tot} \times \frac{K_{D,TF}}{K^{ns}_{D,TF}} \quad \frac{k_b}{k_f} = [D]_{tot} \times \frac{K_{D,RNAp}}{K^{ns}_{D,RNAp}}. \tag{9}$$

The formulas for the reaction rates given in this section depend on a number of parameters, such as the diffusion coefficients of molecules and their binding affinities, given by dissociation constants. Section 1 of the Supporting Material gives a complete description of all of our choices of the numerical values of these parameters. The resulting numerical values of the reaction rates, which we will use throughout this article unless otherwise mentioned, are given in Table 1. These rates are computed assuming dilute (nuclear crowder volume fraction $\phi = 0$) conditions. In the next section, we will see how the level of crowding influences the kinetics and thermodynamics of the reactions in our model.

Dependence of rates on nuclear crowding level

We now consider the effects of nuclear crowding on the rates of specific and nonspecific binding. It is at this level that microscopic details such as molecular geometries, interactions, and diffusion coefficients enter into our model (see

TABLE 1 Numerical values of model parameters.

Parameter	Description	Value (with $\phi = 0$)
k_t^{ns}	Association rate constant for nonspecific TF-DNA binding	4.9×10^4 mM^{-1}s^{-1}
k_f^{ns}	Association rate constant for nonspecific RNAp-DNA binding	3.6×10^4 mM^{-1}s^{-1}
k_o^{ns}	TF-DNA nonspecific dissociation rate	4.9×10^4 s^{-1}
k_b^{ns}	RNAp-DNA nonspecific dissociation rate	3.6×10^4 s^{-1}
$K_{D,TF}^{ns}$	Dissociation constant for nonspecific TF-DNA binding	1 mM
$K_{D,RNAp}^{ns}$	Dissociation constant for nonspecific RNAp-DNA binding	1 mM
k_t	Association rate constant for TF-promoter (O) binding	0.05 nM^{-1}s^{-1}
k_f	Association rate constant for RNAp-Complex I binding	0.03 nM^{-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 nM
$K_{D,RNAp}$	Dissociation constant for RNAp-O (promoter) binding	1 nM
k_m	Rate of pre-mRNA production	0.02 s^{-1}
γ	Nuclear export rate of mRNA	8×10^{-4} s^{-1}
ν	mRNA degradation rate	3×10^{-4} s^{-1}
$[TF]_{tot}$	Total concentration of TF	30 nM
$[RNAp]_{tot}$	Total concentration of RNAp	30 nM
$[O]_{tot}$	Total concentration of O (promoters)	30 nM
$[D]_{tot}$	Total concentration of DNA basepairs	20 mM

The values given here correspond to the values used in this article unless otherwise stated. See Supporting Material for all details regarding the choice of values.

Fig. 1, *b–d*). Motivated by experimental studies showing dramatic nanostructural differences between the nuclei of cells modeling different stages of carcinogenesis (17), we consider changes in the level of crowding specifically in the cell nucleus.

Consider, for example, the first reaction in Fig. 1 *e*, namely, the nonspecific binding of a TF to DNA, with forward rate constant k_t^{ns} and backward rate k_o^{ns}. Changing the level of crowding affects these rates in several ways (1). First, the reaction rates are reduced because of slower diffusion in a crowded medium. Second, the binding of TF to DNA is enhanced; these two objects have lower excluded volume when in contact than they do when apart, giving rise to an attractive depletion interaction of entropic origin. Finally, the same entropic interaction induces a kinetic barrier (see Fig. 1 *c*) that must be overcome for association or dissociation to proceed. Each of these effects depends on the geometries of the molecules involved, including the crowders. We model RNAps as spheres of radius 5.4 nm, TFs as spheres of radius 4.0 nm, DNA as a cylinder of radius 1 nm, and the crowding agents (crowders) as spheres of radius 3.0 nm. The crowders represent the proteins found in the nucleus, assuming an average molecular mass of 67.7 kDa (18); together with a typical partial specific volume of 0.73 mL/g, this leads to our choice of radius for the spherical crowders.

The reaction rates are proportional to the diffusion coefficients of the TF or RNAp (see Eq. 6). We performed BD simulations of spherical tracer particles of various sizes diffusing among spherical crowders of radius 3 nm (Fig. 1 *a*, *left*). These simulations are described in detail in the Supporting Material (see Fig. 1 *b* for results). They yield the factor $f(\phi)$ by which the diffusion coefficient of a tracer molecule is reduced by the presence of a volume fraction, ϕ, of crowders:

$$f(\phi, r) \equiv \frac{D(\phi, r)}{D(0, r)}, \tag{10}$$

where r is the radius of the diffusing tracer particle (TF or polymerase). These functions are well fit by cubic polynomials in ϕ, the coefficients of which are given in Table S1 of the Supporting Material.

The influence of crowding on the equilibrium of each binding reaction is determined by the contribution of crowding to the free energy of binding,

$$\Delta F(\phi) = \Delta F_{\phi=0} + \Delta F_{\text{crowd}}(\phi). \tag{11}$$

The dissociation constant, K_D, of a reaction by definition varies exponentially with the free energy change. Therefore,

$$K_D(\phi) = K_{D,\phi=0} \times \exp[+\beta \Delta F_{\text{crowd}}(\phi)]. \tag{12}$$

We calculate the crowding-induced contribution, $\Delta F_{\text{crowd}}(\phi)$, to the binding free energy using Monte Carlo simulations (Fig. 1 *a*, *right*; see Supporting Material) in which all the reactants interact via excluded volume. The molecular geometries are as described above, although the DNA (a cylinder of radius 1 nm) is here approximated by a row of overlapping spheres of radius 1 nm, each a distance of 1 nm from the next. From these simulations we obtain the crowder-mediated potential of mean force (PMF) acting between the reactants; an example is shown for TF-DNA binding in Fig. 1 *c*, which shows how ΔF_{crowd}, as well as the crowding-induced free-energy barrier to association, $\Delta F_{\text{barrier}}$, are obtained from the PMF. It was convenient to perform simulations of dissociation rather than association; the fact that the PMF increases as the TF and DNA are pulled apart is a manifestation of the attractive nature of the depletion interaction. When the molecules are in contact, their excluded-volume regions overlap, so that a larger set of positions is available for the crowders, leading to an entropic attraction between the TF and the DNA. The crowding-induced contribution to the free energy of TF binding to DNA is shown in Fig. S2. Likewise, we have performed simulations to calculate the crowding-induced free energy differences occurring upon RNAp-D binding. From the point of view of excluded volume, there is no difference between binding of a TF to specific or nonspecific DNA. However, there is an interesting effect of crowding in the case of specific binding of RNAp. As nonspecifically bound

RNAp slides along DNA and comes into contact with a TF, there is a change in excluded volume, leading to a crowding dependence of the strength of specific binding of RNAp to form complex C_{II}. We have also calculated this free-energy change and the associated free-energy barrier. All crowding-induced free energies, as well as the barriers to association, are well fit by polynomial functions of the volume fraction, ϕ, of crowders (see the Supporting Material).

Based on Eqs. 6 and 10, the full dependence of the rates of nonspecific binding on crowding are now given by

$$k_t^{ns}(\phi) = k_{t,0}^{ns} \times f_{TF}(\phi) \times \exp[-\beta\Delta F_{barrier,TF}(\phi)] \quad (13)$$

$$k_f^{ns}(\phi) = k_{f,0}^{ns} \times f_{RNAp}(\phi) \times \exp\left[-\beta\Delta F_{barrier,RNAp}(\phi)\right] \quad (14)$$

The nonspecific dissociation rates are equal to the association rates multiplied by the appropriate equilibrium dissociation constants (Eq. 7), which themselves depend on ϕ (see Eq. 12).

$$\begin{aligned} k_o^{ns}(\phi) &= K_{D,TF}^{ns}(\phi) \times k_t^{ns}(\phi) \\ &= k_{o,0}^{ns} \times \exp[+\beta\Delta F_{crowd,TF-DNA}(\phi)] \times f_{TF}(\phi) \\ &\quad \times \exp[-\beta\Delta F_{barrier,TF}(\phi)] \end{aligned} \quad (15)$$

$$\begin{aligned} k_b^{ns}(\phi) &= K_{D,RNAp}^{ns}(\phi) \times k_f^{ns}(\phi) \\ &= k_{b,0}^{ns} \times \exp\left[+\beta\Delta F_{crowd,RNAp-DNA}(\phi)\right] \\ &\quad \times f_{RNAp}(\phi) \times \exp\left[-\beta\Delta F_{barrier,RNAp}(\phi)\right] \end{aligned} \quad (16)$$

The microscopic mechanism of facilitated diffusion leads to a complex crowding dependence on the rates of specific association and dissociation. According to Eq. 8, the association rate constants for specific binding depend on ϕ through two sources: the one-dimensional diffusion coefficients, D_1, and the square root of the nonspecific dissociation rates, themselves highly ϕ-dependent, as shown above. We make the plausible assumption that crowding slows one-dimensional diffusion by the same factor as for three-dimensional diffusion. This assumption would be seriously violated if the diffusing objects were much smaller than the crowders; rather, they are larger. Thus,

$$\begin{aligned} D_{1,TF}(\phi) &= D_{1,TF,0} \times f_{TF}(\phi) \quad D_{1,RNAp}(\phi) \\ &= D_{1,RNAp,0} \times f_{RNAp}(\phi), \end{aligned} \quad (17)$$

$$\begin{aligned} k_t(\phi) = k_{t,0} \times f_{TF}(\phi) \times \exp\left[+\frac{1}{2}\beta\Delta F_{crowd,TF-DNA}(\phi)\right] \\ \times \exp\left[-\frac{1}{2}\beta\Delta F_{barrier,TF}(\phi)\right], \end{aligned} \quad (18)$$

and

$$\begin{aligned} k_f(\phi) = k_{f,0} \times f_{RNAp}(\phi) \times \exp\left[+\frac{1}{2}\beta\Delta F_{crowd,RNAp-DNA}(\phi)\right] \\ \times \exp\left[-\frac{1}{2}\beta\Delta F_{barrier,RNAp}(\phi)\right] \\ \times \exp\left[-\beta\Delta F_{barrier,RNAp}^{slide}(\phi)\right]. \end{aligned} \quad (19)$$

Here we have introduced the aforementioned free-energy barrier $\Delta F_{barrier,RNAp}^{slide}$ that must be overcome as RNAp slides along DNA toward TF (that is, toward a C_I complex). Note that the dependence due to the free-energy changes appears under a square root (hence the factors of $1/2$), whereas the dependence due to the factors f does not. This is because the f factors influence the overall specific binding rate through both the one-dimensional diffusion coefficient and the nonspecific dissociation rate (see Eq. 8).

The rates of nonspecific dissociation are the last ones whose ϕ dependence we must determine. From Eq. 9,

$$\begin{aligned} k_o(\phi) &= [D]_{tot} \times \frac{K_{D,TF}(\phi)}{K_{D,TF}^{ns}(\phi)} \times k_t(\phi) \\ &= k_{o,0} \times f_{TF}(\phi) \times \exp\left[+\frac{1}{2}\beta\Delta F_{crowd,TF-DNA}(\phi)\right] \\ &\quad \times \exp\left[-\frac{1}{2}\beta\Delta F_{barrier,TF}(\phi)\right]. \end{aligned} \quad (20)$$

On the other hand, specific binding of RNAp to Complex I involves bringing RNAp into contact with TF while sliding along the DNA. This brings about a change in excluded volume and therefore a crowding-induced free-energy change, as well as a free-energy barrier.

$$\begin{aligned} k_b(\phi) &= [D]_{tot} \times \frac{K_{D,RNAp}(\phi)}{K_{D,RNAp}^{ns}(\phi)} \times k_f(\phi) \\ &= k_{b,0} \times \exp\left[+\beta\Delta F_{crowd,RNAp-TF}^{slide}(\phi)\right] \times f_{RNAp}(\phi) \\ &\quad \times \exp\left[+\frac{1}{2}\beta\Delta F_{crowd,RNAp-DNA}(\phi)\right] \\ &\quad \times \exp\left[-\beta\Delta F_{barrier,RNAp-TF}^{slide}(\phi)\right] \\ &\quad \times \exp\left[-\frac{1}{2}\beta\Delta F_{barrier,RNAp}(\phi)\right]. \end{aligned} \quad (21)$$

In view of Eqs. 2 and 3 for steady-state cytoplasmic mRNA concentration and production rate, respectively, it now remains only to determine the ϕ dependence of the transcription rate, k_m, and the mRNA degradation rate, ν. Each of these processes is of course composed of many complicated subprocesses, as well as being driven by energy consumption. We assume that the rates of these processes are

independent of the crowder volume fraction, ϕ, occurring in the nucleus:

$$k_m(\phi) = k_{m,0} \quad \nu(\phi) = \nu_0.$$

In particular, since we are interested in understanding the effects of changes in the level of crowding specifically in the nucleus, the lack of ϕ independence of ν reflects the fact that mRNA degradation occurs in the cytoplasm. We now have the complete ϕ dependence of the reaction rates needed to calculate the mRNA production level. These dependences involve the effects of slowed diffusion, which we have determined from BD simulations, as well as the crowding-induced free-energy changes, which we have calculated using Monte Carlo simulations. All of these simulations, as well as their quantitative results, are summarized in the Supporting Material.

RESULTS

Equations 4 and 5 were solved numerically, taking into account the previous section's ϕ dependences of the reaction rates, as well as the parameter values from Table 1. Equation 2 then gives the steady-state cytoplasmic mRNA concentration (Fig. 2, *solid black curve*). The cytoplasmic mRNA level shows a distinctly nonmonotonic dependence on the volume fraction, ϕ, of crowders in the nucleus, with a maximum near $\phi = 0.3$, a physiologically relevant value. The figure also shows cytoplasmic mRNA concentrations as a function of ϕ for larger values (10 nM and 100 nM) of the dissociation constant for specific binding of RNAp to C_I. This corresponds to weaker binding, resulting in a lower overall concentration of complex C_{II} and therefore of mRNA. With weaker RNAp binding, the maximum level mRNA production occurs at higher volume fractions.

As the volume fraction ϕ of crowders is increased, the affinities of protein-DNA and protein-protein interactions are enhanced due to the fact that a bound complex has lower excluded volume than free reactants. Thus, if the states of binding of the proteins (TF and RNAp) were at equilibrium, the concentration of complex C_{II} would increase monotonically with the level of crowding. This is illustrated in Fig. 3, which shows the steady-state cytoplasmic mRNA level as a function of ϕ for small values of the transcription rate k_m. For comparison these quantities are shown normalized by their values at $\phi = 0$. In the limit of very small transcription rates k_m, we see the monotonic behavior expected at equilibrium, solely due to the enhancement of binding. Only at very high volume fractions, near $\phi = 0.5$, where the diffusion coefficients vanish (see Fig. 1 *b*), does the mRNA production decrease. This highlights the role of the driven, irreversible process of transcription, whose rate is k_m, in keeping the system out of equilibrium and thus allowing the mRNA production level to depend on the kinetics of diffusion, which slows down as a function of ϕ. Thus, the nonmonotonicity of mRNA concentration as a function of crowder volume fraction, ϕ, is a consequence of the competition of enhanced binding (essentially an equilibrium effect) with the slowing down of diffusion.

The extent of this nonmonotonicity can be modulated by changing various parameters of the system. Fig. 4 shows that the nonmonotonic dependence on ϕ becomes even more pronounced if the reactants (TF, RNAp, and O) are present in smaller concentrations of the order of 3 nM or 0.3 nM, rather than 30 nM. This corresponds to thousands or hundreds of molecules per nucleus, rather than tens of thousands. Fig. 5 shows the effects of lowering the reactant concentrations on the various populations of TFs: free, nonspecifically bound, specifically bound to promoters (C_I), and bound in transcription-ready complexes (C_{II}). This is shown for large reactant concentrations (30 nM; Fig. 5, *upper*) and for lower concentrations (3 nM; Fig. 5, *lower*). In both cases, free TFs make up a very small fraction of the total; moreover, this fraction decreases as a function of crowding due to the enhancement of binding. However, in the case of high concentrations, the fraction of TFs bound in transcription-ready C_{II} complexes is much larger. For this reason, the decrease in freely diffusing TFs contributes little to the concentration of C_{II} (and hence to the mRNA level). In contrast, in the case of low reactant concentrations (Fig. 5, *lower*), the concentration of C_{II} is very small,

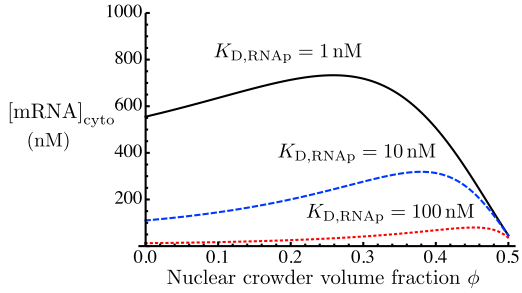

FIGURE 2 Steady-state mRNA concentration as a function of the crowder volume fraction, ϕ, in the nucleus. This is shown for three different values of the dissociation constant $K_{D,RNAp}$ of RNAp binding to C_I. To see this figure in color, go online.

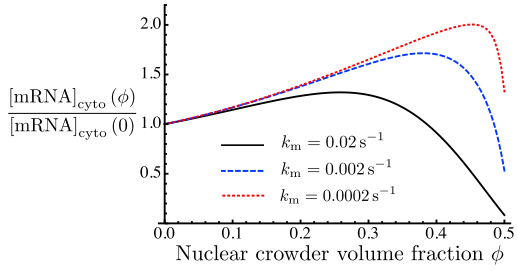

FIGURE 3 Effect of varying transcription rate k_m. The plot shows the fold change in steady-state mRNA concentration as a function of the crowder volume fraction, ϕ, in the nucleus compared to the mRNA concentration at zero crowding. This quantity is shown for three different values of k_m. To see this figure in color, go online.

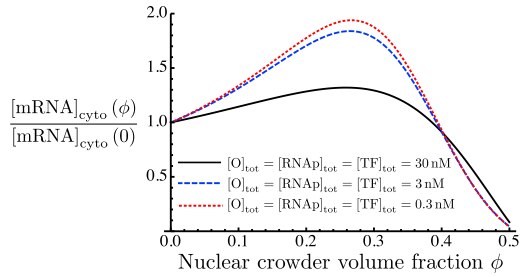

FIGURE 4 Fold change in steady-state mRNA concentration as a function of the crowder volume fraction, ϕ, in the nucleus, compared to the mRNA concentration at zero crowding. Total concentrations of TF, RNAp, and O are equal to 30 nM (*black solid curve*), 3 nM (*blue dashed curve*), and 0.3 nm (*red dotted curve*). The zero-crowding mRNA concentrations $[\mathrm{mRNA}]_{\mathrm{cyto}}(0)$ for these three cases are 555.2 nM, 2.9 nM, and 0.004 nM, respectively. To see this figure in color, go online.

comparable to that of free TFs. As ϕ increases, the decreasing level of free TFs therefore contributes more significantly to the fraction of C_{II} and hence to the cytoplasmic mRNA level.

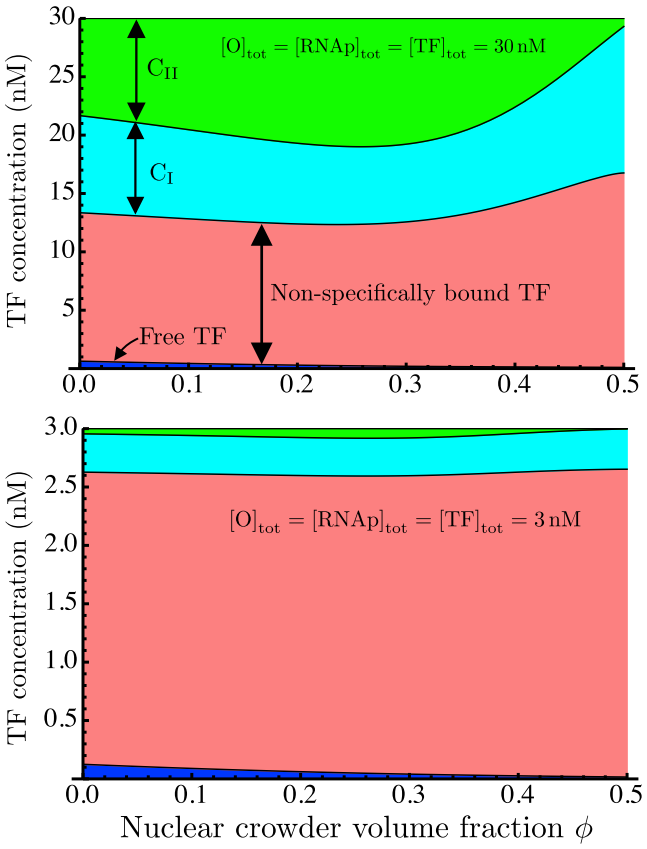

FIGURE 5 Breakdown of total concentration of transcription factors (*TF*) into subpopulations of TFs that are free (*blue*), nonspecifically bound to DNA (*pink*), specifically bound (complex C_I; *cyan*), and bound in transcription-read complexes C_{II} (*green*). This is shown for two different sets of reactant concentrations. (*Upper*) Reactant concentrations of 30 nM. (*Lower*) Reactant concentrations of 3 nM. Note the different *y* axis scales in the two plots. To see this figure in color, go online.

Fig. 6 shows another way in which the dependence of mRNA levels on crowding may be modulated, namely, by changing the strength of the nonspecific protein-DNA association. For the sake of comparison, the cytoplasmic mRNA concentration is plotted relative to its value at $\phi = 0$ for values of the nonspecific dissociation constant both larger (10 mM) and smaller (0.1 mM) than that considered previously. Weaker nonspecific binding (a larger dissociation constant) gives rise to a much weaker ϕ dependence of the mRNA output, whereas stronger nonspecific binding (still much weaker than the specific binding) yields a much more prominent maximum in the mRNA level as a function of crowding. Thus, we find that the strength of this ϕ dependence may be modulated by changes in a variety of quantities, such as concentrations and binding affinities. In reality, these might correspond to changes in local physical conditions surrounding a given gene.

DISCUSSION

In this work, we have used molecular-scale simulations to incorporate the effects of macromolecular crowding into a model of gene transcription as a network of chemical reactions. We made explicit and consistent choices of molecular geometries in determining the effects of crowding on the diffusion coefficients and the binding free energies of all molecular species. The dependence of the diffusion constants on the volume fraction, ϕ, of spherical crowders was determined through BD simulations, whereas the crowding contributions to the free energies of binding were calculated from Monte Carlo simulations. These results were incorporated into expressions for the rates of specific and nonspecific binding involved in the facilitated diffusion of DNA-binding proteins (10). The importance of making a consistent choice of molecular shapes and sizes can be seen from our estimates of $\Delta F_{\mathrm{crowd}}$, the contribution of crowding to the binding free energy, compared to that used by Morelli et al. (9) in their treatment of crowding effects on TF binding (see Fig. 7). In effect, they assumed that the chemical potentials of all molecular species have the

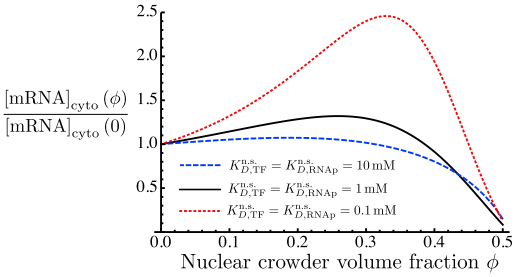

FIGURE 6 Effect of varying the strength of nonspecific protein-DNA binding is shown in plots of the fold change in steady-state mRNA concentration as a function of the nuclear crowder volume fraction, ϕ, for three different values of the dissociation constant for nonspecific protein-DNA binding. To see this figure in color, go online.

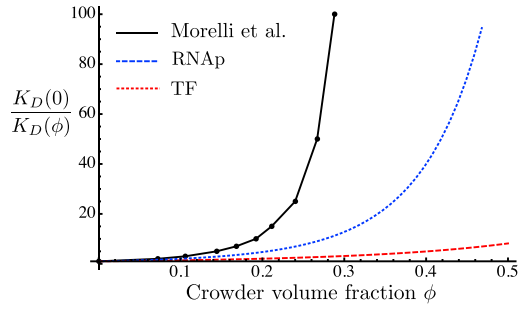

FIGURE 7 Crowding dependence of the inverse of the dissociation constant. The black solid curve represents data from Morelli et al. (9), relating their parameter Γ to the volume fraction, ϕ, of crowders. For this conversion we took a specific volume of 0.96 mL/g used by Minton (27) to fit the osmotic pressure of hemoglobin solutions (35) to a theoretical model for hard spheres. The blue dotted and red dashed curves give the crowding dependence of inverse dissociation constants for RNAp binding and TF binding, respectively, as determined from our Monte Carlo simulation data and Eq. 12. To see this figure in color, go online.

same quantitative dependence on the level of crowding. In a bimolecular association reaction, this corresponds to a very large difference in excluded volume between the free and bound states, equal to the entire excluded volume of a reactant. In contrast, if molecules stick together upon binding, the change in their excluded volume is small compared to the total excluded volume of a molecule. The result is that our crowding-induced free energies are much smaller than those of Morelli et al. (9), showing the importance of a consistent treatment of molecular geometries.

Our results show that in many conditions, the level of cytoplasmic mRNA shows a markedly nonmonotonic dependence on the volume fraction, ϕ, of crowders in the nucleus, frequently reaching a maximum at values near $\phi = 0.3$. This is typical of the levels of crowding typically reported for living cells (19). We have also shown that this crowding dependence may itself be modulated by several means. The location of the maximum in the ϕ dependence may be tuned by changing the binding affinity of RNAp for bound TF (complex C_I), whereas the relative amplitude of the ϕ dependence can be greatly accentuated by decreasing the overall concentrations of reactants (Fig. 4) or by increasing the strength of nonspecific protein-DNA binding.

Morelli et al. (9) also found a nonmonotonic dependence of mRNA levels on the level of crowding under certain conditions, especially when taking account of nonspecific binding of polymerase to DNA. Macromolecular crowding has also been found theoretically to exert a nonmonotonic influence on the accessible surface area of chromatin (7), as well as the timescale of loop formation in polymers (20); these are effects of potential relevance to gene expression levels, and suggest the intriguing possibility that the level of crowding could function as a regulatory control mechanism. Several authors have highlighted the dual influence of crowding, which promotes molecular association while slowing diffusion, and the consequent potential existence of an optimal level of macromolecular crowding (21,22). Here, we emphasize the role of the irreversible process of transcription in keeping the system out of equilibrium, allowing its steady-state properties to depend on the diffusion rates.

There is some evidence that the level of macromolecular crowding is conserved across different mammalian cell lines (23). If the level of macromolecular crowding is under evolutionary and homeostatic control, then, conversely, disregulation of crowding may be associated with faulty regulation of gene expression and with related disease states such as cancer. Thus, the link between macromolecular crowding and transcriptional regulation may shed light on previously observed nanoscale structural differences between healthy and precancerous cells (24), since these differences must be associated with variations in local crowding and concentrations, which in turn influence gene expression.

Experiments by Ge et al. (25) have determined the effects of macromolecular crowding in a cell-free protein expression system, using polyethylene glycol as well as Ficoll as the crowding agent. Notably, they found that the level of mRNA produced by cell-free transcription depends nonmonotonically on the concentration of crowding agents.

Recently, Tan et al. (26) studied the crowding dependence of gene expression in a cell-free system using phage T7 components. Using as crowding agents dextran molecules of two different sizes, they showed that in the case of small dextran molecules the gene expression rate is a nonmonotonic function of the density of crowders. Interestingly, they found that at equal weight/volume percent concentrations of crowders, large dextran molecules generate a much larger crowding effect on gene expression than small dextran molecules. Although Tan et al. state that this size effect is consistent with existing theories of crowding, in fact, considerations of excluded volume predict that at equal crowder volume fractions, and therefore equal weight/volume percent, smaller crowders will induce more significant depletion interactions (4,5,27). The size effect found by Tan et al. (26) highlights the currently unknown relationship between theoretical (hard-sphere) and in vitro (dextran, Ficoll, or polyethylene glycol) models of intracellular crowding.

Notwithstanding the numerous theoretical studies of macromolecular crowding, the role of the physical environment of a gene in determining the level of its transcription is still unclear. With respect to the influence of macromolecular crowding, it will be necessary to go beyond the typical treatment of crowders as inert hard spheres. Several steps have been taken in this direction, for example, by considering crowders with interactions or nonspherical shapes (28,29). However, we must take seriously the fact that the role of the crowding agent in the nucleus is played in part by chromatin rather than by independently mobile molecules. Its physical effects on transcription are perhaps intermediate between those of mobile crowders and the confinement effects of fixed obstacles (30,31). Also, the

structure of chromatin, which determines the accessibility of DNA binding sites to TFs, is predicted to be influenced by macromolecular crowding in a nonmonotonic way (7). Furthermore, to integrate the physical microenvironments of genes into a systems-level treatment of gene expression, one must take into account the spatial compartmentalization of the nucleus, whose origins may in turn be due to entropic crowding effects (32,33). Finally, it is tempting to speculate on the possibility that chromatin territories within the nucleus regulate gene expression by locally controlling, in additional to specific genes, concentrations of reactants and crowders. This local control could lead to gene upregulation or downregulation, according to our predictions, by the proper choice of environment.

SUPPORTING MATERIAL

One table, eight figures, and a detailed description of the model are available at http://www.biophysj.org/biophysj/supplemental/S0006-3495(14)00225-2.

The authors acknowledge useful discussions with Luay Almassalha.

This work was supported by the National Science Foundation under grants EFRI CBET-0937987 and EAGER-124931.

REFERENCES

1. Zhou, H., G. Rivas, and A. Minton. 2008. Macromolecular crowding and confinement: biochemical, biophysical, and potential physiological consequences. *Annu. Rev. Biophys.* 37:375–397.
2. Schnell, S., and T. E. Turner. 2004. Reaction kinetics in intracellular environments with macromolecular crowding: simulations and rate laws. *Prog. Biophys. Mol. Biol.* 85:235–260.
3. Kim, J. S., and A. Yethiraj. 2009. Effect of macromolecular crowding on reaction rates: a computational and theoretical study. *Biophys. J.* 96:1333–1340.
4. Qin, S., L. Cai, and H. X. Zhou. 2012. A method for computing association rate constants of atomistically represented proteins under macromolecular crowding. *Phys. Biol.* 9:066008.
5. Kim, Y. C., R. B. Best, and J. Mittal. 2010. Macromolecular crowding effects on protein-protein binding affinity and specificity. *J. Chem. Phys.* 133:205101.
6. Qin, S., and H. X. Zhou. 2009. Atomistic modeling of macromolecular crowding predicts modest increases in protein folding and binding stability. *Biophys. J.* 97:12–19.
7. Kim, J. S., V. Backman, and I. Szleifer. 2011. Crowding-induced structural alterations of random-loop chromosome model. *Phys. Rev. Lett.* 106:168102.
8. Richter, K., M. Nessling, and P. Lichter. 2007. Experimental evidence for the influence of molecular crowding on nuclear architecture. *J. Cell Sci.* 120:1673–1680.
9. Morelli, M. J., R. J. Allen, and P. R. Wolde. 2011. Effects of macromolecular crowding on genetic networks. *Biophys. J.* 101:2882–2891.
10. Berg, O. G., R. B. Winter, and P. H. von Hippel. 1981. Diffusion-driven mechanisms of protein translocation on nucleic acids. 1. Models and theory. *Biochemistry.* 20:6929–6948.
11. Bintu, L., N. E. Buchler, …, R. Phillips. 2005. Transcriptional regulation by the numbers: models. *Curr. Opin. Genet. Dev.* 15:116–124.
12. Garcia, H. G., and R. Phillips. 2011. Quantitative dissection of the simple repression input-output function. *Proc. Natl. Acad. Sci. USA.* 108:12173–12178.
13. Coulon, A., C. C. Chow, …, D. R. Larson. 2013. Eukaryotic transcriptional dynamics: from single molecules to cell populations. *Nat. Rev. Genet.* 14:572–584.
14. Berg, O. 1978. On diffusion-controlled dissociation. *Chem. Phys.* 31:47–57.
15. Li, G.-W., O. Berg, and J. Elf. 2009. Effects of macromolecular crowding and DNA looping on gene regulation kinetics. *Nat. Phys.* 5: 294–297.
16. Singh, A., and P. Bokes. 2012. Consequences of mRNA transport on stochastic variability in protein levels. *Biophys. J.* 103:1087–1096.
17. Damania, D., H. Subramanian, …, V. Backman. 2010. Role of cytoskeleton in controlling the disorder strength of cellular nanoscale architecture. *Biophys. J.* 99:989–996.
18. Bickmore, W. A., and H. G. Sutherland. 2002. Addressing protein localization within the nucleus. *EMBO J.* 21:1248–1254.
19. Ellis, R. J. 2001. Macromolecular crowding: an important but neglected aspect of the intracellular environment. *Curr. Opin. Struct. Biol.* 11:114–119.
20. Toan, N. M., D. Marenduzzo, …, C. Micheletti. 2006. Depletion effects and loop formation in self-avoiding polymers. *Phys. Rev. Lett.* 97: 178302.
21. Vazquez, A. 2010. Optimal cytoplasmatic density and flux balance model under macromolecular crowding effects. *J. Theor. Biol.* 264: 356–359.
22. Dill, K. A., K. Ghosh, and J. D. Schmit. 2011. Physical limits of cells and proteomes. *Proc. Natl. Acad. Sci. USA.* 108:17876–17882.
23. Guigas, G., C. Kalla, and M. Weiss. 2007. The degree of macromolecular crowding in the cytoplasm and nucleoplasm of mammalian cells is conserved. *FEBS Lett.* 581:5094–5098.
24. Subramanian, H., P. Pradhan, …, V. Backman. 2008. Optical methodology for detecting histologically unapparent nanoscale consequences of genetic alterations in biological cells. *Proc. Natl. Acad. Sci. USA.* 105:20118–20123.
25. Ge, X., D. Luo, and J. Xu. 2011. Cell-free protein expression under macromolecular crowding conditions. *PLoS ONE.* 6:e28707.
26. Tan, C., S. Saurabh, …, P. Leduc. 2013. Molecular crowding shapes gene expression in synthetic cellular nanosystems. *Nat. Nanotechnol.* 8:602–608.
27. Minton, A. P. 1983. The effect of volume occupancy upon the thermodynamic activity of proteins: some biochemical consequences. *Mol. Cell. Biochem.* 55:119–140.
28. Kim, J. S., and A. Yethiraj. 2011. Crowding effects on protein association: effect of interactions between crowding agents. *J. Phys. Chem. B.* 115:347–353.
29. McGuffee, S. R., and A. H. Elcock. 2010. Diffusion, crowding & protein stability in a dynamic molecular model of the bacterial cytoplasm. *PLOS Comput. Biol.* 6:e1000694.
30. Minton, A. P. 1992. Confinement as a determinant of macromolecular structure and reactivity. *Biophys. J.* 63:1090–1100.
31. Saxton, M. J. 1994. Anomalous diffusion due to obstacles: a Monte Carlo study. *Biophys. J.* 66:394–401.
32. Marenduzzo, D., K. Finan, and P. R. Cook. 2006. The depletion attraction: an underappreciated force driving cellular organization. *J. Cell Biol.* 175:681–686.
33. Finan, K., P. R. Cook, and D. Marenduzzo. 2011. Non-specific (entropic) forces as major determinants of the structure of mammalian chromosomes. *Chromosome Res.* 19:53–61.
34. Humphrey, W., A. Dalke, and K. Schulten. 1996. VMD: visual molecular dynamics. *J. Mol. Graph.* 14:33–38, 27–28.
35. Adair, G. 1928. A theory of partial osmotic pressures and membrane equilibria, with special reference to the application of Dalton's law to haemoglobin solutions in the presence of salts. *Proc. R. Soc. Lond. A Math. Phys. Sci.* 120:573–603.

Supplementary Materials

Macromolecular Crowding as a Regulator of Gene Transcription

H. Matsuda, G.G. Putzel, V. Backman, and I. Szleifer

1 Numerical estimates of model parameters

Here we determine numerical estimates of the parameters appearing in the model, namely, the reaction rates and the total concentrations of the reagents. The rates are determined from the microscopic theory of facilitated diffusion developed by Berg and coworkers (1–3). Our estimates are made assuming dilute conditions (nuclear crowder volume fraction $\phi = 0$). We do not concern ourselves with the rates of the irreversible reactions following pre-mRNA production, such as splicing and mRNA export, since they do not influence the steady-state results.

Assuming that non-specific binding of TF and RNAp to DNA is diffusion-limited, we use the following expression (1–3) for the forward rates k_1 and k_3:

$$k_{\mathrm{t}}^{\mathrm{ns}} = \frac{2\pi D_{\mathrm{TF}} l}{\ln(\xi/2b)} \qquad k_{\mathrm{f}}^{\mathrm{ns}} = \frac{2\pi D_{\mathrm{RNAp}} l}{\ln(\xi/2b)} \tag{1}$$

Here D_{TF} and D_{RNAp} are the diffusion coefficients of TF and RNAp, given by the Stokes-Einstein formula

$$D_{\mathrm{TF}} = \frac{k_B T}{6\pi\eta r_{\mathrm{TF}}} \qquad D_{\mathrm{RNAp}} = \frac{k_B T}{6\pi\eta r_{\mathrm{RNAp}}}$$

We assume a temperature of $37^\circ\,\mathrm{C} = 310\,\mathrm{K}$, at which the dynamic viscosity of water is $\eta = 6.5 \times 10^{-4}\,\mathrm{J \cdot s/m^3}$. We are modeling the RNA polymerase as a sphere of radius $r_{\mathrm{RNAp}} = 5.4\,\mathrm{nm}$ and the TF as a sphere of radius $r_{\mathrm{TF}} = 4\,\mathrm{nm}$, giving $D_{\mathrm{TF}} = 8.7 \times 10^{-11}\,\mathrm{m^2/s}$ and $D_{\mathrm{RNAp}} = 6.4 \times 10^{-11}\,\mathrm{m^2/s}$.

The definitions of the variables ξ, b, and l are as follows. ξ is the correlation length for the DNA, that is, the typical distance between neighboring DNA segments. This correlation length will vary depending on the physical arrangement of the DNA into loops, chromatin territories, etc. We choose $\xi = 20\,\mathrm{nm}$, keeping in mind that these rates will vary weakly with ξ; we have

checked that varying ξ in either direction by a factor of 4 does not affect our qualitative results. Two more lengths appear in Eq. 1: the radius b of the DNA molecule, about 1 nm, and the length l along the DNA of one base pair, namely 0.34 nm. Using these values, we have non-specific association rate constants in the absence of crowding:

$$k_{\mathrm{t}}^{\mathrm{ns}} = 4.9 \times 10^4\,\mathrm{mM}^{-1}\mathrm{s}^{-1} \qquad k_{\mathrm{f}}^{\mathrm{ns}} = 3.6 \times 10^4\,\mathrm{mM}^{-1}\mathrm{s}^{-1}$$

The nonspecific dissociation rates follow from the dissociation constants for non-specific binding, called $K_{D,\mathrm{TF}}^{\mathrm{n.s.}}$ and $K_{D,\mathrm{RNAp}}^{\mathrm{n.s.}}$. We take these to be both equal to 1 mM in the absence of crowding, giving

$$k_{\mathrm{o}}^{\mathrm{ns}} = K_{D,\mathrm{TF}}^{\mathrm{n.s.}} \cdot k_{\mathrm{t}}^{\mathrm{ns}} = 4.9 \times 10^4\,\mathrm{s}^{-1} \qquad k_{\mathrm{b}}^{\mathrm{ns}} = K_{D,\mathrm{RNAp}}^{\mathrm{n.s.}} \cdot k_{\mathrm{f}}^{\mathrm{ns}} = 3.6 \times 10^4\,\mathrm{s}^{-1}$$

The association rate constants for specific binding of TF and RNAp are given by the expression derived by Berg et al. (2) for specific protein-DNA binding by facilitated diffusion:

$$k_{\mathrm{t}} = V \cdot (D_{1,\mathrm{TF}} \cdot k_{\mathrm{o}}^{\mathrm{ns}})^{1/2}/L \qquad k_{\mathrm{f}} = V \cdot (D_{1,\mathrm{RNAp}} \cdot k_{\mathrm{b}}^{\mathrm{ns}})^{1/2}/L \tag{2}$$

The factor of V representing the volume of the nucleus does not appear in the expression of Berg et al. (2). This is due to the fact that in the reaction equations Eqs. 4 and 5 we take the forward rates to multiply the product of the *volume* densities (concentrations) of both reagents, in such a way that the association rate constants have the usual dimensions of concentration$^{-1}\times$time^{-1}. We use a typical volume for a HeLa cell of $V = 500\,\mu\mathrm{m}^3$ (4, 5). The length L is one-half of the total length of genomic DNA. A diploid human cell (before S phase) contains 6×10^9 bp. This corresponds to $L = 1\,\mathrm{m}$. The quantities $D_{1,\mathrm{TF}}$ and $D_{1,\mathrm{RNAp}}$ are the one-dimensional diffusion coefficients for motion of TF and RNAp along DNA. Elf et al. (6) report a value of $D_1 = 0.046\,\mu\mathrm{m}^2/s$ for a transcription factor, which is about 100 times smaller than the TF's three-dimensional diffusion coefficient. Assuming that the protein follows a helical path winding around the DNA as has been found experimentally (7), Bagchi et al. (8) derived a formula for the one-dimensional diffusion coefficient of a protein non-specifically bound to DNA:

$$D_1 = \frac{k_B T}{\zeta_{\mathrm{trans}}^{\mathrm{total}}},$$

where

$$\zeta_{\mathrm{trans}}^{\mathrm{total}} = 6\pi\eta R + \left(\frac{2\pi}{10l}\right)^2 \left[8\pi\eta R^3 + 6\pi\eta R (R_{\mathrm{OC}})^2\right]$$

Here R is the radius of the spherical protein while $R_{\rm OC}$ is the distance between the centers of mass of the protein and DNA. The integer 10 appears because this is the number of base pairs (of length $l = 0.34\,\mathrm{nm}$) per turn of DNA. Using the radii of the TF and RNAp we find

$$D_{1,\mathrm{TF}} = \frac{D_{3,\mathrm{TF}}}{157.3} \qquad D_{1,\mathrm{RNAp}} = \frac{D_{3,\mathrm{RNAp}}}{270.4}$$

Using these numerical values, we obtain

$$k_{\rm t} = 0.05\,\mathrm{nM}^{-1}\mathrm{s}^{-1} \qquad k_{\rm f} = 0.03\,\mathrm{nM}^{-1}\mathrm{s}^{-1} \tag{3}$$

The backward rates for specific binding are determined from (2)

$$\frac{k_{\rm o}}{k_{\rm t}} = [D]_{\rm tot} \cdot \frac{K_{D,\mathrm{TF}}}{K^{\rm ns}_{D,\mathrm{TF}}} \qquad \frac{k_{\rm b}}{k_{\rm f}} = [D]_{\rm tot} \cdot \frac{K_{D,\mathrm{RNAp}}}{K^{\rm ns}_{D,\mathrm{RNAp}}} \tag{4}$$

The human genome has about 3×10^9 bp. In a diploid cell, there are therefore about $6 \times 10^9 \approx 10^{-14}$ moles of base pairs in a volume of about $500\,\mu\mathrm{m}^3 = 5 \times 10^{-13}$ L. This gives a concentration of $[\mathrm{D}]_{\rm tot} = 0.02\,\mathrm{M} = 20\,\mathrm{mM}$. Protein-DNA dissociation constants have been found to range widely, with reported values for transcription factor-DNA binding ranging from micromolar to picomolar (9, 10) or even stronger (11), with nanomolar values being common. We have been using dissociation constants of 1 nM for specific binding of TF or RNAp. The resulting values of the dissociation rates are

$$k_{\rm o} = (0.05\,\mathrm{nM}^{-1}\mathrm{s}^{-1}) \cdot (20\,\mathrm{mM}) \cdot \frac{1\,\mathrm{nM}}{1\,\mathrm{mM}} = 1.0\,\mathrm{s}^{-1} \tag{5}$$

$$k_{\rm b} = (0.03\,\mathrm{nM}^{-1}\mathrm{s}^{-1}) \cdot (20\,\mathrm{mM}) \cdot \frac{1\,\mathrm{nM}}{1\,\mathrm{mM}} = 0.6\,\mathrm{s}^{-1} \tag{6}$$

The inverse of the rate k_m represents the time taken by the polymerase (once already bound to the promoter) to initiate transcription, as well as to produce the pre-mRNA transcript. An average initiation time of one minute (12) gives

$$k_{m,0} = 0.02\,\mathrm{s}^{-1}$$

In the model, the step in which nuclear mRNA becomes cytoplasmic mRNA (with rate γ) includes nucleocytoplasmic diffusion to a nuclear pore in addition to directed transport across the nuclear pore. The transport and export of mRNAs has been studied at the single-molecule level in Drosophila by

Mor et al. (13) who report that nucleocytoplasmic diffusion to a nuclear pore complex (NPC) occurs on a timescale of 5–40 minutes, while directed transport across the NPC occurs much more rapidly. Assuming a mean time scale of 20 minutes for an mRNA to exit the nucleus from the site of splicing, we have

$$\gamma_0 = 8 \times 10^{-4}\,\mathrm{s}^{-1}$$

The rate of degradation of mRNA in the cytoplasm varies from one transcript to another(14, 15); we use a typical time scale of about one hour to arrive at

$$\nu_0 = 3 \times 10^{-4}\,\mathrm{s}^{-1}$$

We now must estimate the total concentration (bound as well as free) of promoters, RNA polymerases, and transcription factors. Here a question arises as to how to interpret our model: are we describing a single gene and its specific transcription factors, as did Morelli et al.(16), or rather describing in a coarse-grained way all of the genes at once? In the former case the concentration of promoters in a diploid cell is $2/V$ where V is the volume of the nucleus, while in the latter case we must take into account the total number of active genes. We choose to include all of the active genes at once, to include the possibility that changes in binding affinities (for example, as the level of crowding is changes) for such a significant number of genes might affect the concentration of free RNAp or TF. Experiments in HeLa cells (17) suggest a number of RNAp molecules on the order of 10^4, with numbers of general transcription factors (such as TBP) of the same order of magnitude. Likewise, in a genome with tens of thousands of genes the number of binding sites O is of the same magnitude. Given that the nuclei of these cells have volumes (4, 5) on the order of $V_{\mathrm{nuc}} = 500\,\mu\mathrm{m}^3 = 5 \times 10^{-13}\,\mathrm{L}$ we have

$$[\mathrm{TF}]_{\mathrm{tot}} \approx [\mathrm{RNAp}]_{\mathrm{tot}} \approx [\mathrm{O}]_{\mathrm{tot}} \approx \frac{10^4}{V_{\mathrm{nuc}}} \approx 30\,\mathrm{nM}$$

These are clearly rough estimates, and for many reasons we should consider a wide range of possible concentrations. The distributions of molecules in the nucleus involved in transcription are non-uniform, and we might choose to interpret our concentrations as the local values at the site of transcription (for example, in transcription factories), since the local concentrations are the only relevant ones for the binding reactions. Furthermore, many promoter binding sites will be bound by nucleosomes and inaccessible to DNA-binding proteins, so that the concentration of "active" genes may be reduced compared to the total number of genes.

2 Diffusion Coefficients: Brownian Dynamics simulations

To find the crowding dependence of the diffusion coefficients of TF and RNAp in our model, we performed Brownian Dynamics (BD) simulations of spherical tracer particles of various radii diffusing among spherical crowders of radius 3 nm. These simulations were carried out using the GROMACS package (18). The purely repulsive interaction potential between two spherical particles (either crowders or tracers) of radii r_1 and r_2 was

$$U(r) = 4\epsilon \left(\frac{\sigma}{x + \sigma - r_1 - r_2} \right)^{12}, \tag{7}$$

with $\epsilon = 1\,\mathrm{kJ/mol}$ and $\sigma = 1\,\mathrm{nm}$. According to the Stokes formula, the friction coefficient γ of a sphere is proportional to its radius. In our simulations we have used drag coefficients of $\gamma = r \cdot 30\,\mathrm{amu/ps} = r \cdot 30\,\mathrm{kJ \cdot ps/mol \cdot nm^2}$ where r is the radius of a molecule (crowder or tracer). Our simulations were performed at a temperature of 310 K, although the results of the simulations (diffusion coefficients normalized by their crowder-free values) should be nearly independent of temperature since our interaction potentials are nearly hard-sphere potentials.

For each tracer particle radius we performed simulations at many different levels of crowding by including different numbers of crowders within a $(63\ \mathrm{nm})^3$ simulation box. The number of 3 nm radius crowders in the box varied from 30 to 1140 in steps of 30, corresponding to volume fractions up to about $\phi = 0.5$.

From each simulation trajectory the diffusion coefficient of the tracer particle was determined as one sixth of the average slope of the particle's mean-square displacement over the time interval from $t = 5\,\mathrm{ns}$ to $t = 20\,\mathrm{ns}$. The diffusion coefficients were then normalized to their values in the absence of crowding. The resulting normalized diffusion coefficients are shown in Fig. 1. These sets of simulations were performed for tracer particles of radii 2,4,5, and 6 nm; for tracer particle radius of 3 nm we used simulations with crowders only. The normalized diffusion coefficients were well fit by a cubic polynomial in ϕ, namely

$$\frac{D(\phi, r)}{D(0, r)} \equiv f(\phi, r) = 1 + \alpha \left(\frac{r}{r_{\mathrm{crowd}}} \right) \phi + \beta \left(\frac{r}{r_{\mathrm{crowd}}} \right) \phi^2 + \gamma \left(\frac{r}{r_{\mathrm{crowd}}} \right) \phi^3, \tag{8}$$

here α, β, and γ are fitting parameters, depending only on the ratio of the tracer particle radius to the crowder radius, which are given in Table 1. For

r/r_{crowd}	α	β	γ
2/3	-1.41	0.40	-2.23
3/3	-1.95	1.10	-1.91
4/3	-2.83	3.87	-4.11
5/3	-3.57	6.46	-6.45
6/3	-4.39	9.69	-9.76

Table 1: Coefficients in Eqn. 8 for a cubic fit to $f(\phi)$.

the model discussed in the article, we need normalized diffusion coefficients for tracer particles of radius 4 and 5.4 nm, representing the transcription factor and the RNA polymerase, respectively. For the transcription factor, we use Eqn. 8 above with the fitting parameters for $r = 4$ nm. For the RNAp, we did not have simulations of tracers with radius precisely equal to 5.4 nm, so for the polymerase's diffusion coefficient we have used Eqn. 8 with fitting parameters obtained by interpolating the data of Table 1: namely $\alpha = -3.89$, $\beta = 7.72$, and $\gamma = -7.72$.

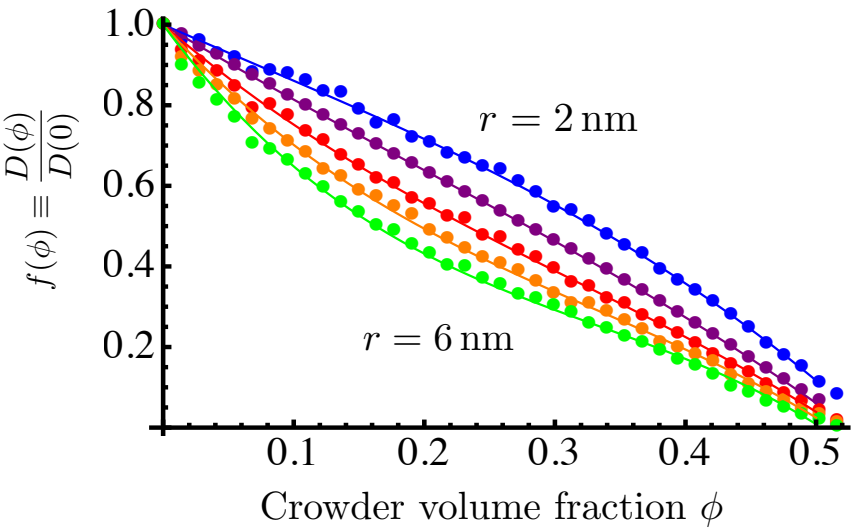

Figure 1: The factor f by which the diffusion coefficient of a tracer molecule is decreased by the presence of a volume fraction ϕ of crowders (radius 3 nm). This is shown for various radii of tracer molecules: 2, 3, 4, 5, and 6 nm (top to bottom). Brownian Dynamics simulations are shown with dots; the curves are polynomial fits given by Eq. 8 with parameters given in Table 1.

3 Crowding-mediated interactions: Monte Carlo simulations

To calculate the crowding-induced contribution $\Delta F_{\rm crowd}(\phi)$ to the binding free energies, as well as the crowding-induced free energy barrier to association $\Delta F_{\rm barrier}(\phi)$, we have performed Monte Carlo simulations, which we discuss now in the context of TF binding to DNA; the case of RNAp binding to DNA is similar. In each simulation, the cubic simulation box of size $(50\ {\rm nm})^3$ contains the TF (a sphere of radius 4 nm) as well as the DNA (a row of 50 overlapping spheres of radius 1 nm, spaced 1 nm apart, approximating a rod of diameter 2 nm) and some number $N_{\rm crowd}$ of crowders. The crowders interact with each other, and with the spheres making up the TF and DNA, as impenetrable hard spheres. The Monte Carlo moves are random small translations of randomly selected crowders. These moves are rejected if they cause any overlap between the crowders or between crowders and the TF or DNA, and accepted otherwise. Every 10 MC moves, a test move is considered which increases the distance between the TF and DNA by 0.1 nm (this distance is the reaction coordinate). Likewise, a test move is also considered which decreases this distance by 0.1 nm. Like the MC moves, the test moves are accepted or rejected based on whether they lead to overlaps between crowders and TF or DNA. Over the course of millions of MC moves, this gives a numerical estimate of the probability $p_{\rm forward}$ of accepting a move that increases the reaction coordinate, as well as the probability $p_{\rm backward}$ of accepting the reverse move. The free energy change in increasing the TF-DNA distance by $\Delta x = 0.1$ nm is then given by (19)

$$\beta \Delta F_{\rm crowd}(x \to x + \Delta x) = \ln \left[\frac{p_{\rm backward}}{p_{\rm forward}} \right], \tag{9}$$

where $\beta = 1/k_B T$. From the free energy changes for each small step of size Δx, we map out the potential of mean force (PMF) between the TF and the DNA (see Fig. 1c of the article). As shown in this figure, the PMF reaches a plateau value which is the free energy required to separate the TF from the DNA against the depletion force caused by the crowders, or equivalently minus the crowder contribution to the free energy of binding. The difference between the maximum of the PMF and the plateau value gives the crowder-mediated free energy barrier $\Delta F_{\rm barrier}$ to association. These sets of simulations were performed for numbers of crowders between 20 and 440 in steps of 20, corresponding to nuclear crowder volume fractions from $\phi =$ 0.018 to 0.4. The resulting free energies and free energy barriers are plotted

in Fig. 2 and Fig. 3 respectively. Likewise, the crowding-induced free energy difference for pulling RNAp away from DNA is shown in Fig. 4 and the corresponding free energy barriers in Fig. 5. Finally, the change in excluded volume as the RNAp slides into contact with the TF along the DNA (forming the complex C_{II}) entails a crowding-induced free energy change, which we have also calculated. This is shown in Fig. 6, with the corresponding free energy barrier plotted in Fig. 7. The free energies determined from the Monte Carlo simulations were well fit by the following expressions.

$$\begin{aligned}
-\beta\Delta F_{\mathrm{crowd,TF}}(\phi) &= 3.2\,\phi + 2.0\,\phi^2 \qquad -\beta\Delta F_{\mathrm{barrier,TF}}(\phi) = 2.5\,\phi^2 \\
-\beta\Delta F_{\mathrm{crowd,RNAp}}(\phi) &= 3.7\,\phi + 2.7\,\phi^2 \qquad -\beta\Delta F_{\mathrm{barrier,RNAp}}(\phi) = 3.1\,\phi^2 \\
-\beta\Delta F^{\mathrm{slide}}_{\mathrm{crowd,RNAp\text{-}TF}}(\phi) &= 2.6\,\phi + 4.6\,\phi^2 \qquad -\beta\Delta F^{\mathrm{slide}}_{\mathrm{barrier,RNAp\text{-}TF}}(\phi) = 0.1\,\phi^2 + 9.2\,\phi^3
\end{aligned}$$

The Monte Carlo simulations were implemented using a C program whose code is available from the authors upon request.

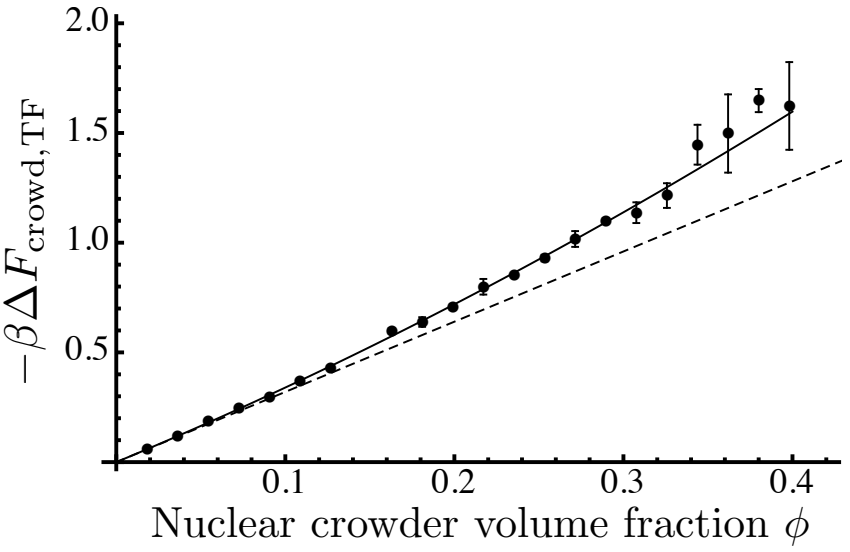

Figure 2: Crowding induced free energy change associated with TF-DNA binding. Symbols show the results of Monte Carlo simulations with error bars given by the standard error of four independent simulations. The solid line shows the polynomial fit $-\beta\Delta F_{\mathrm{crowd,TF}}(\phi) = 3.2\,\phi + 2.0\,\phi^2$ used in computing the results in the paper. The dashed line shows the linear approximation $-\beta\Delta F_{\mathrm{crowd,TF}}(\phi) = 3.2\,\phi$ that can be obtained by exact integration of excluded volume overlaps.

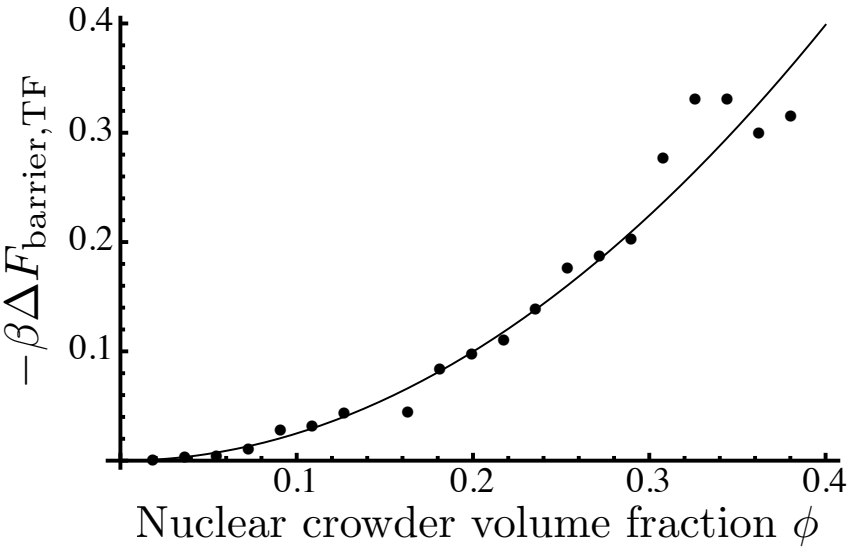

Figure 3: Crowding induced free energy barrier associated with TF-DNA binding. Symbols show the results of Monte Carlo simulations.The solid line shows the polynomial fit $-\beta\Delta F_{\mathrm{barrier,TF}}(\phi) = 2.5\,\phi^2$ used in computing the results in the paper.

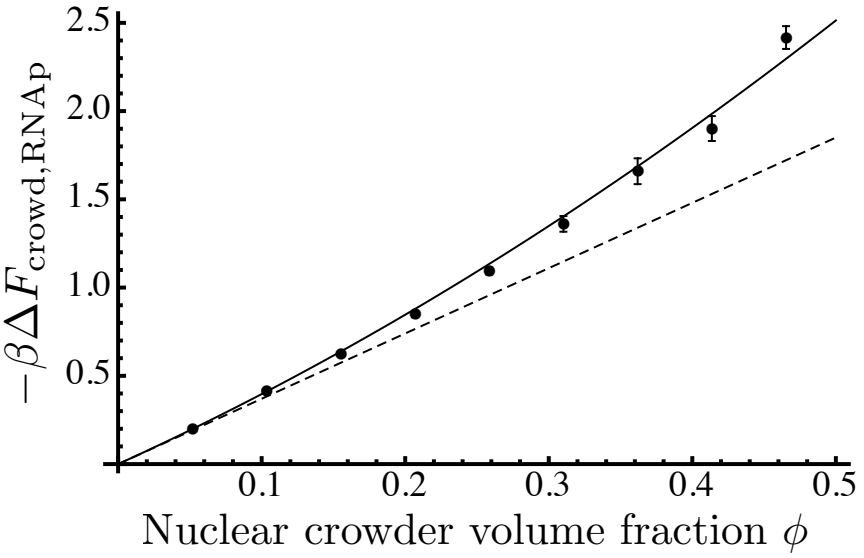

Figure 4: Crowding induced free energy change associated with RNAp-DNA binding. Symbols show the results of Monte Carlo simulations with error bars given by the standard error of ten independent simulations. The solid line shows the polynomial fit $-\beta\Delta F_{\mathrm{crowd,RNAp}}(\phi) = 3.7\,\phi + 2.7\,\phi^2$ used in computing the results in the paper. The dashed line shows the linear approximation $-\beta\Delta F_{\mathrm{crowd,RNAp}}(\phi) = 3.7\,\phi$ that can be obtained by exact integration of excluded volume overlaps.

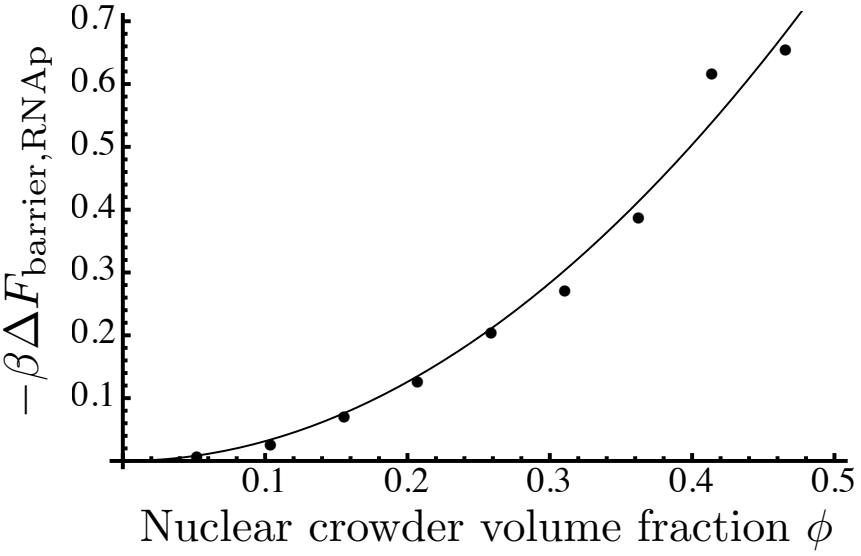

Figure 5: Crowding induced free energy barrier associated with RNAp-DNA binding. Symbols show the results of Monte Carlo simulations.The solid line shows the polynomial fit $-\beta\Delta F_{\text{barrier,TF}}(\phi) = 3.1\,\phi^2$ used in computing the results in the paper.

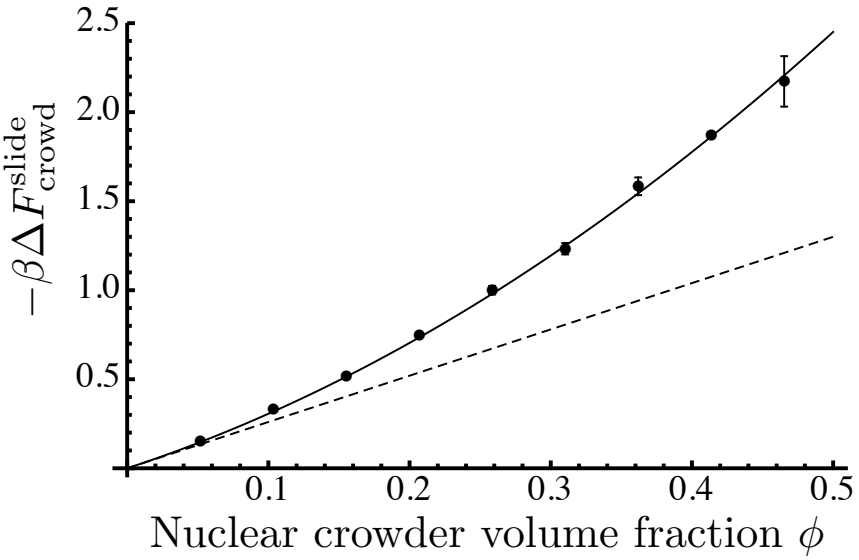

Figure 6: Crowding induced free energy change associated with RNAp "sliding" toward TF along DNA, forming complex C_{II}. Symbols show the results of Monte Carlo simulations with error bars given by the standard error of ten independent simulations. The solid line shows the polynomial fit $-\beta\Delta F^{\rm slide}_{\rm crowd,RNAp\text{-}TF}(\phi) = 2.6\,\phi + 4.6\,\phi^2$ used in computing the results in the paper. The dashed line shows the linear approximation $-\beta\Delta F^{\rm slide}_{\rm crowd,RNAp\text{-}TF}(\phi) = 2.6\,\phi$ that can be obtained by exact integration of excluded volume overlaps.

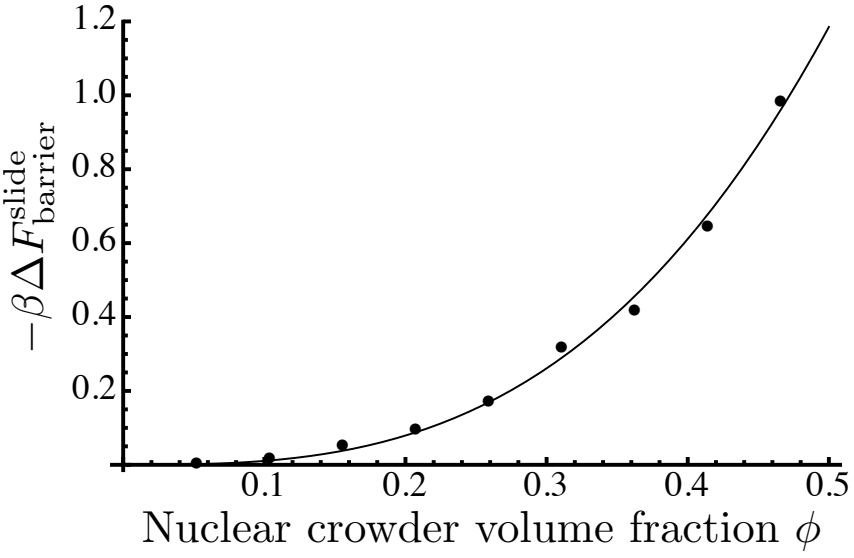

Figure 7: Crowding induced free energy barrier associated with RNAp "sliding" toward TF along DNA, forming complex C_{II}. Symbols show the results of Monte Carlo simulations. The solid line shows the polynomial fit $-\beta\Delta F^{\rm slide}_{\rm barrier,RNAp\text{-}TF}(\phi) = 0.1\,\phi^2 + 9.2\,\phi^3$ used in computing the results in the paper.

4 Effects of Crowder Polydispersity

In order to ensure that our results are not dependent on the strict monodispersity of crowders in the model, we have performed some Monte Carlo simulations to calculate the crowding induced free energy change upon RNAp binding to DNA, using a small level of crowder polydispersity. Specifically, we used crowders with a distribution of three sizes: small (2.7 nm), medium (3.0 nm), and large (3.3 nm). Half of the crowders were of medium size and 25% were small and large. Figure 8 shows that this poly-dispersity does not significantly affect the results.

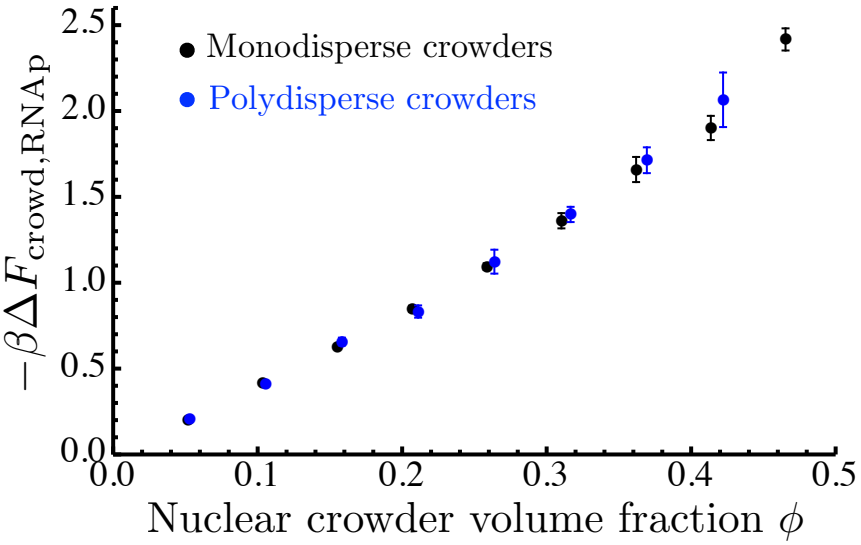

Figure 8: Effects of crowder polydispersity on the crowding induced free energy change upon RNAp binding to DNA.

References

[1] Berg, O., 1978. On Diffusion-Controlled Dissociation. *Chemical Physics* 31:47–57.

[2] Berg, O., R. Winter, and P. von Hippel, 1981. Diffusion-Driven Mechanisms of Protein Translocation on Nucleic Acids. 1. Models and Theory. *Biochemistry* 20:6929–6948.

[3] Li, G.-W., O. Berg, and J. Elf, 2009. Effects of macromolecular crowding and DNA looping on gene regulation kinetics. *Nat Phys* 5:294–297.

[4] Maul, G. G., and L. Deaven, 1977. Quantitative Determination of Nuclear Pore Complexes in Cycling Cells with Differing DNA Content. *J Cell Biol* 73:748–760.

[5] Monier, K., J. G. Armas, S. Etteldorf, P. Ghazal, and K. Sullivan, 2000. Annexation of the inter chromosomal space during viral infection. *Nat Cell Biol* 2:661–665.

[6] Elf, J., G. Li, and X. Xie, 2007. Probing Transcription Factor Dynamics at the Single-Molecule Level in a Living Cell. *Science* 316:1191–1194.

[7] Blainey, P., G. Luo, S. Kou, W. F. Mangel, G. L. Verdine, B. Bagchi, and X. Xie, 2009. Nonspecifically bound proteins spin while diffusing along DNA. *Nat. Struct. Mol. Biol.* 16:1224–1230.

[8] Bagchi, B., P. Blainey, and X. Xie, 2008. Diffusion Constant of a Nonspecifically Bound Protein Undergoing Curvilinear Motion along DNA. *J. Phys. Chem. B* 112:6282–6284.

[9] Geertz, M., D. Shore, and S. Maerkl, 2012. Massively parallel measurements of molecular interaction kinetics on a microfluidic platform. *Proc Natl Acad Sci U S A* 109:16540–16545.

[10] Nalefski, E., E. Nebelitski, J. A. Lloyd, and S. Gullans, 2006. Single-Molecule Detection of Transcription Factor Binding to DNA in Real Time: Specificity, Equilibrium, and Kinetic Parameters. *Biochemistry* 45:13794–13806.

[11] Zabel, U., R. Schreck, and P. Baeuerle, 1991. DNA Binding of Purified Transcription Factor NF-κB. *J Biol Chem* 266:252–260.

[12] Darzacq, X., Y. Shav-Tal, V. de Turris, Y. Brody, S. M. Shenoy, R. D. Phair, and R. H. Singer, 2007. In vivo dynamics of RNA polymerase II transcription. *Nat Struct Mol Biol* 14:796–806.

[13] Mor, A., S. Suliman, R. Ben-Yishay, S. Yunger, Y. Brody, and Y. Shav-Tal, 2010. Dynamics of single mRNP nucleocytoplasmic transport and export through the nuclear pore in living cells. *Nat Cell Biol* 12:543–554.

[14] Yang, E., E. van Nimwegen, M. Zavolan, N. Rajewsky, M. Schroeder, M. Magnasco, and J. E. D. Jr., 2003. Decay Rates of Human mRNAs: Correlation With Functional Characteristics and Sequence Attributes. *Genome Res* 13:1863–1872.

[15] Schwanhäusser, B., D. Busse, N. Li, G. Dittmar, J. Schuchhardt, J. Wolf, W. Chen, and M. Selbach, 2011. Global quantification of mammalian gene expression control. *Nature* 473:337–342.

[16] Morelli, M., R. Allen, and P. ten Wolde, 2011. Effects of Macromolecular Crowding on Genetic Networks. *Biophys J* 101:2882–2891.

[17] Kimura, H., Y. Tao, R. Roeder, and P. Cook, 1999. Quantitation of RNA Polymerase II and Its Transcription Factors in an HeLa Cell: Little Soluble Holoenzyme but Significant Amounts of Polymerases Attached to the Nuclear Substructure. *Mol Cell Biol* 19:5383–5392.

[18] Hess, B., C. Kutzner, D. van der Spoel, and E. Lindahl, 2008. GROMACS 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation. *J Chem Theory Comput* 4:435–447.

[19] Bennett, C. H., 1976. Efficient Estimation of Free Energy Differences from Monte Carlo Data. *J Comput Phys* 22:245–268.