Noncooperative thermodynamics and kinetic models of ligand binding to polymers: Connecting McGhee–von Hippel model with the Tonks gas model

Juan P. G. Villaluenga [1,*] Jules Vidal,[1] and Francisco Javier Cao-García [1,2,†]

[1]*Departamento de Estructura de la Materia, Física Térmica y Electrónica, Universidad Complutense de Madrid,
Pza. de Ciencias, 1, 28040 Madrid, Spain*

[2]*Instituto Madrileño de Estudios Avanzados en Nanociencia, IMDEA Nanociencia, C/Faraday, 9, 28049 Madrid, Spain*

Ligand binding to polymers modifies the physical and chemical properties of the polymers, leading to physical, chemical, and biological implications. McGhee and von Hippel obtained the equilibrium coverage as a function of the ligand affinity, through the computation of the possible binding sites for the ligand. Here, we complete this theory deriving the kinetic model for the ligand-binding dynamics and the associated equilibrium chemical potential, which turns out to be of the Tonks gas model type. At low coverage, the Tonks chemical potential becomes the Fermi chemical potential and even the ideal gas chemical potential. We also discuss kinetic models associated with these chemical potentials. These results clarify the kinetic models of ligand binding, their relations with the chemical potentials, and their range of validity. Our results highlight the inaccuracy of ideal and simplified kinetic approaches for medium and high coverages.

I. INTRODUCTION

Ligand binding to polymers changes the mechanical and chemical properties of the polymer and plays a very relevant role in regulating biological functions [1–12]. The ligand-binding shrinks or expands the binding region of the polymer and generally stiffens this region [6,13–18]. Additionally, the binding of a ligand frequently passivates a polymer region for the binding of other units of the ligand. Therefore, the binding of a ligand can be seen as the occupation of binding sites. This excluded volume effect implies an effective reduction of the affinity of the ligand for the polymer.

Excluded volume effects are very relevant in chemistry and physics [19,20]. They play a key role in the structure of molecules [21,22], in the properties of gases [23–25], and in the distribution of proteins and lipids in the cell membrane [26–28], to cite some examples. Hard-sphere models, despite their simplicity, allow us to obtain important characteristics of real gases [29–33]. Freely moving chains of concatenated hard spheres allow us to understand the relevant properties of real polymers [21]. The two-dimensional version of the hard-sphere gas, the hard-disk gas, has also been explored [34,35] and explains features of the protein and lipid distribution in membranes [26–28]. The one-dimensional version, the Tonks gas, consists of hard segments located in a line [29]. We have previously shown that the Tonks gas thermodynamics appropriately estimates the coverage of DNA by human mitochondrial single strand binding protein (HmtSSB) [6,36].

A few theoretical approaches for modeling the binding of ligands to macromolecules already exist. McGhee and von Hippel [37] derived a model for predicting the binding of large ligands to a one-dimensional lattice-like macromolecule of infinite length. This result was followed by results for models of finite lattices [38], systems with binding sites not homogeneously distributed throughout the macromolecule [39], systems in which there are interactions between ligands bound to the polymer [40–47], different classes of binding sites present in the macromolecule [48], and systems with the binding of flexible branched oligopolymers [49], to cite some examples.

Recently, we also proposed a simple model to compute the ligand kinetics (based on a coarse-grained binding site counting) [6]. The generalization of this model for ligands with two binding modes provided a qualitative explanation of the transient shortenings observed during HmtSSB binding to DNA [36]. However, we clarify here that this simple model is only valid at low coverage, where it matches the Tonks gas coverage relation. The Tonks gas coverage relation is shown to match the coverage relation derived by McGhee and von Hippel (hereafter GH) [37]. GH proposed a more detailed binding site counting to derive equilibrium coverage for single-mode ligands. We aim here to provide a more complete and rigorous picture of ligand binding to polymer theory addressing the kinetic and the thermodynamics of ligand binding by using the detailed binding site counting GH procedure. We compare the results with the simple coarse-grained and the ideal gas kinetics models and their respective chemical potentials and predicted equilibrium coverages. Here, we focus on systems in which the ligands do not interact between them in the binding process, as is the case of the binding of, for example, oligolisines to nucleic acids [50], mono- and multivalent ions to polyelectrolytes [51], or polypeptides to polynucleotides [52].

Understanding the kinetics of binding of large ligands to long polymers is a very relevant biological problem. The enzymes implicated in the replication and repair of DNA do

*jpgarcia@ucm.es
†francao@fis.ucm.es

FIG. 1. Schematic representation of a chain with some bound ligands. Each four-sphere group could represent E. Coli SSB tetramer binding 60 nucleotides of single-stranded DNA, as happens in one of the E. Coli SSB binding modes [2].

not act on naked single-stranded DNA (ssDNA); they act over these chains covered with proteins. All organisms have proteins that cover ssDNA to protect it from degradation, as single-stranded binding protein (SSB) in bacteria [5,53–55]. Other proteins involved in the DNA protection and repair are the eukaryotic PA and RAD51 protein family and bacterial RecA [28,56]. Increased precision of measurements on the kinetics of these protein-binding kinetics with single-molecule techniques [36,54–56] demand a better theoretical understanding of the correct kinetic models. This better understanding will also benefit the understanding of other ligand-binding dynamics as Thiocoraline binding to double-stranded DNA (dsDNA), with applications in DNA imaging and anticancer pharmacology [57].

In Sec. II A, we obtain the proposed kinetic model for ligand binding by using the McGhee and von Hippel procedure. We show that it reproduces the McGhee and von Hippel result for equilibrium coverage [37], and it also corresponds to the result obtained with the Tonks gas model [6,29]. In Sec. II B, we show that simplification of the proposed model at low coverage leads to a Fermi-like model similar to the model introduced in Ref. [6]. In Sec. II C, we show that further simplification leads to an ideal gas chemical potential. Finally, Sec. III discusses the results.

II. MODEL DERIVATION AND COMPARISON

A. McGhee–von Hippel–Tonks model

We use the procedure proposed by McGhee and von Hippel (GH) to derive a theoretical framework describing the kinetics of the binding of large ligands to a long polymer [37] (see Fig. 1). The resulting model will be called the McGhee-von Hippel-Tonks model (hereinafter denoted the "GHT" model). Within the original GH model, two different equations were derived for the binding density of ligands per macromolecule, i.e., non-cooperative (without interactions between ligands bound to the polymer) and cooperative (there exist interactions between ligands bound to the polymer) ligand binding. Here, we consider that the ligands do not interact between them, and partial binding by the ligand is not allowed. The polymer is represented by a linear array of N identical repeat units (named monomers or polymer residues), and when the ligand molecule binds to the polymer covers m consecutive units. Thus, in the GHT approach, the polymer is modeled as a lattice and the ligand as a single entity binding a fixed number of polymer units. In addition, the polymer is assumed to be infinitely long, implying $N \gg m$. Real polymers are not infinite, and end-effects can be relevant. The error introduced by using the infinite length assumption is estimated to be less than the typical experimental error as long as $N/m > 30$

[38,58]. The fraction of polymer monomers covered by the ligand is given by the coverage $c = nm/N$, where n is the number of ligands bound to the polymer. In the present approach, a polymer residue can only exist in two states, free or bound to the ligand.

Following the words of McGhee and von Hippel, almost every monomer can start a ligand-binding site on a polymer with no bound ligands. This implies that ligands are not constrained to bind only at regularly spaced intervals m monomers apart. Consequently, the actual number of free ligand-binding sites on a naked polymer is $N - m + 1$. The number of potential ligand-binding sites eliminated by binding one ligand can range from 1, if the ligand binds into a gap exactly m residues long, up to $2m - 1$ if it binds to a naked polymer. For a gap g residues long between adjacent bound ligands, the number of binding sites is $g - m + 1$ if $g \geqslant m$, but zero if $g < m$. The average number of free binding sites per gap is

$$\bar{s} = \sum_{g=m}^{N} (g - m + 1)P_g, \tag{1}$$

where P_g is the probability that any particular gap between two bound ligands is exactly g free polymer residues long. Thus, the average number of free ligand-binding sites in the polymer is

$$(n + 1)\bar{s}, \tag{2}$$

where n is the number of ligands bound. Within the requirements of the model, the expression for the conditional probability turns out to be [37]

$$P_g = \left(\frac{1 - c}{1 - c + c/m} \right)^g \left(\frac{c/m}{1 - c + c/m} \right), \tag{3}$$

where c is the coverage (see Appendix I for the derivation). Then, the length of the chain is assumed to be much larger than the number of bound monomers per ligand, $N \gg m$, and that there are many ligands bound, $n \gg 1$. With these assumptions, we get, combining Eqs. (2) and (3) and summing up the series,

$$(n + 1)\bar{s} \approx n\bar{s} = N(1 - c)\left(\frac{1 - c}{1 - c + c/m} \right)^{m-1}, \tag{4}$$

and also an equation for the average number of free binding sites per gap as a function of the coverage,

$$\bar{s} = \frac{(1 - c)}{c/m}\left(\frac{1 - c}{1 - c + c/m} \right)^{m-1}. \tag{5}$$

We now derive a kinetic equation describing the time variation in the number of ligands attached to the polymer:

$$\frac{dn}{dt} = k_b n\bar{s} - k_r n, \tag{6}$$

where k_b is the binding kinetic constant, and k_r is the release kinetic constant. The binding kinetic constant is defined as $k_b = K_b L$, where K_b represents the association constant, and L is the free ligand concentration. If we combine Eqs. (5) and (6), we obtain the GHT model equation for the time evolution
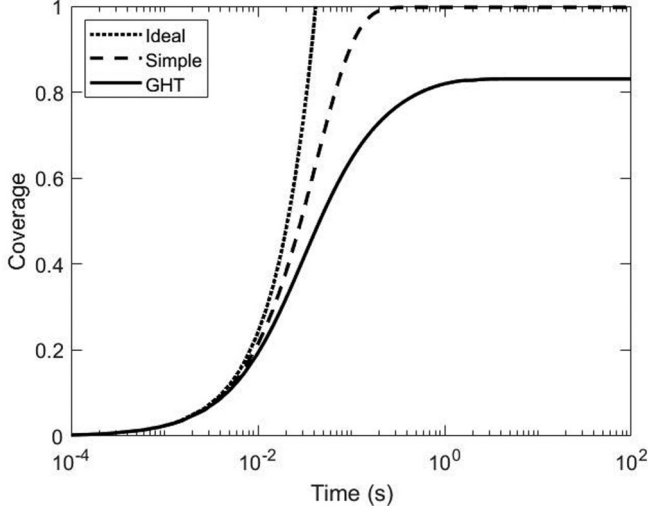
FIG. 2. Comparison of polymer coverage estimated with different models as a function of time. All starting with a naked polymer and with ligands binding to $m = 30$ monomers, with binding rate $k_b = 0.8\,\mathrm{s}^{-1}$, and release rate $k_r = 0.06\,\mathrm{s}^{-1}$. The result for the ideal model is computed from Eq. (18), the result for the simple model is computed from Eq. (15), and the result for the GHT model is computed from Eq. (7).

of the coverage,

$$\frac{dc}{dt} = k_b m(1-c)\left(\frac{1-c}{1-c+c/m}\right)^{m-1} - k_r c. \qquad (7)$$

This evolution equation implies that, if we start with a naked polymer, the coverage will increase until reaching an equilibrium value where the binding and release terms balance each other. The equilibrium coverage can be obtained numerically from Eq. (7) by imposing $dc/dt = 0$. See Fig. 2.

Equation (7) is valid for all values of m (provided $m \ll N$). In the simpler case, when the ligand only binds a monomer, $m = 1$, we recover the intuitive result that the effective binding rate is given by $k_b(1-c)$, i.e., that the effective binding rate is proportional to the fraction of uncovered monomers, $(1-c)$. Some ligands bind few monomers, such as Thiocoraline binding to dsDNA [57], human polymerase β binding to ssDNA [48], or binding of lysozyme to partially acetylated chitosans [39]. Also, there are relevant cases where the ligand is large and binds to many monomers, like E. Coli and human mitochondrial SSB that bind 30–70 nucleotides of ssDNA [2,36,59]. In these cases, we have $m \gg 1$, which implies

$$\left(\frac{1-c}{1-c+c/m}\right)^{m-1} \approx \exp\left(-\frac{c}{1-c}\right), \qquad (8)$$

allowing us to simplify the kinetic equation (7) and leading to the equilibrium coverage equation

$$0 = k_b m(1-c_{eq})\exp\left(-\frac{c_{eq}}{1-c_{eq}}\right) - k_r c_{eq}, \qquad (9)$$

or equivalently

$$\ln\left(\frac{k_b m}{k_r}\right) = \ln\left(\frac{c_{eq}}{1-c_{eq}}\right) + \frac{c_{eq}}{1-c_{eq}}. \qquad (10)$$
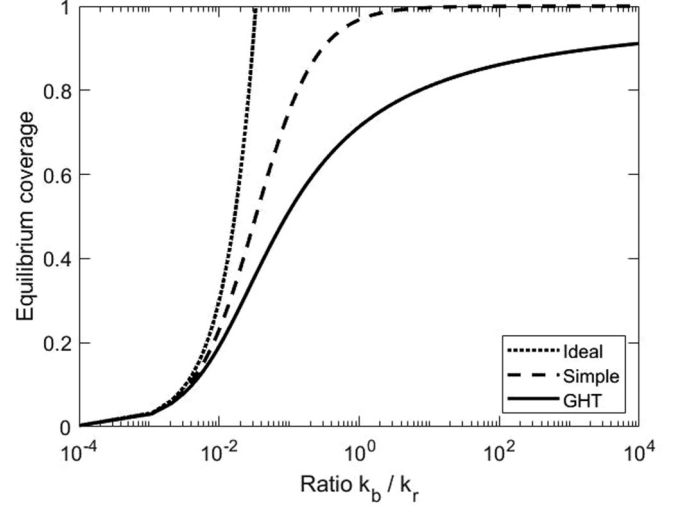

FIG. 3. Comparison of equilibrium coverage estimated with different models as a function of the ratio between the binding rate k_b and the release rate k_r, for a ligand binding to $m = 30$ monomers. The result for the ideal model is computed from the stationary state of Eq. (18), the result for the simple model is computed from the stationary state of Eq. (15), and the result for the GHT model is computed from the stationary state of Eq. (7).

The equilibrium coverage for large m can be computed numerically from Eqs. (9) or (10). Alternatively, the equilibrium coverage can be expressed in terms of the so-called Lambert function W [60,61],

$$c_{eq} = \frac{W\left(\frac{k_b L m}{k_r}\right)}{1 + W\left(\frac{k_b L m}{k_r}\right)}. \qquad (11)$$

The definition of $W(z)$ is that it is the function that solves the equation $W e^W = z$, where z is a complex number.

The equilibrium coverage calculated by using Eq. (11) increases as a function of the ratio between the binding and release rates, as shown in Fig. 3. The increase in the number of monomers occluded by each ligand also increases the equilibrium coverage; see Fig. 4.

The equilibrium coverage in Eq. (10) matches that obtained with a chemical potential of the form

$$\frac{\mu - \mu^o}{k_B T} = \ln\left(\frac{1}{m}\right) + \ln\left(\frac{c}{1-c}\right) + \frac{c}{1-c}, \qquad (12)$$

where μ^o is the chemical potential of a ligand bound to the polymer in a reference state, k_B is the Boltzmann constant, and T is the thermodynamic temperature. See Fig. 5. This expression corresponds to a one-dimensional Tonks gas, which is a one-dimensional generalization of the hard-sphere gas. The right-hand side of Eq. (12) contains the term $\ln(1/m)$, related to the size of the ligand in a monomer unit basis, the ideal gas contribution $\ln(c)$, and the excluded volume corrections given by the remaining terms in the expression.

B. Simple model

In the low coverage range, the last term of the right-hand side of Eq. (12) can be neglected, and a Fermi-like expression
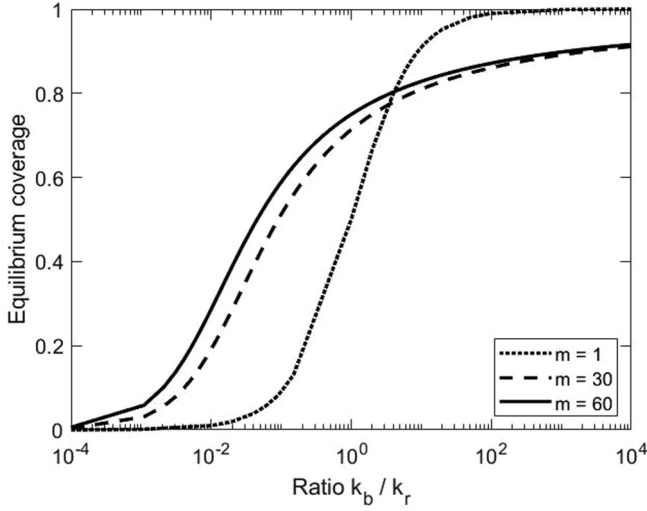
FIG. 4. Influence of the number of occluded monomers by a single ligand in the equilibrium coverage in the GHT model, as a function of the ratio between the binding kinetic constant and the release kinetic constant. The results for $m = 30$ and $m = 60$ are computed from Eq. (11), while the result for $m = 1$ is directly computed from the stationary state of Eq. (7).

can approximate the GHT chemical potential:

$$\frac{\mu - \mu^o}{k_B T} = \ln\left(\frac{1}{m}\right) + \ln\left(\frac{c}{1-c}\right). \qquad (13)$$

Hereinafter, this result is denoted "the simple model." See Fig. 5 for a plot of this chemical potential, which leads to an equilibrium coverage of

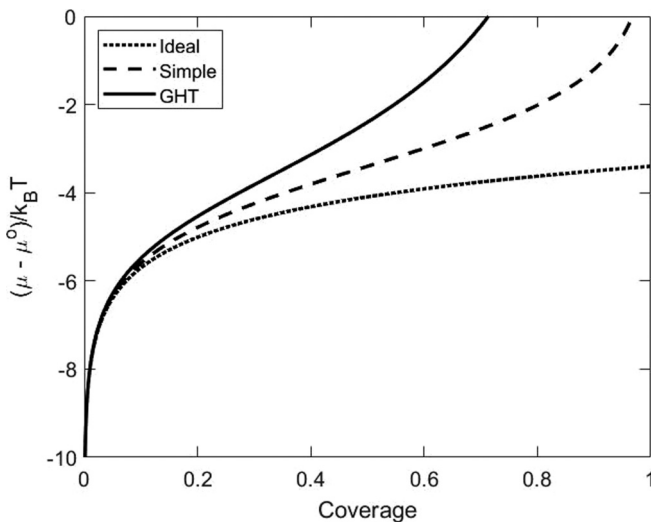$$c_{eq} = \frac{1}{1 + k_r/(mk_b)}. \qquad (14)$$

FIG. 5. Comparison of chemical potentials calculated with different models as a function of the coverage for a ligand covering $m = 30$ monomers. The result for the ideal model is computed from Eq. (16), the result for the simple model is computed from Eq. (13), and the result for the GHT model is computed from Eq. (12).

Figures 2, 3 and 5 show that this approximation is valid for low coverages, i.e., $c < 0.1$.

This equilibrium coverage will also be obtained from a kinetic model of the form

$$\frac{dc}{dt} = k_b m(1 - c) - k_r c. \qquad (15)$$

This kinetic equation can be derived under the assumption that the number of binding sites is given by $N - nm$. Figure 2 shows the evolution of the coverage given by this simplified model and compares it with the GHT model.

A similar description of the process was given by Ref. [6], where the number of binding sites was estimated by the number of naked monomers of the polymer divided by the number of monomers that a single bound ligand occludes, i.e., $(N - nm)/m$. However, in this approach, both the kinetic equation for the coverage and the equilibrium coverage are independent of the number of occluded monomers per ligand, m, in disagreement with the more accurate GHT model derived here.

C. Ideal model

If we consider an ideal lattice-gas model, namely, neglecting the excess contribution due to the excluded volume corrections, the GHT chemical potential in Eq. (12) can be written as follows:

$$\frac{\mu - \mu^o}{k_B T} = \ln\left(\frac{1}{m}\right) + \ln c, \qquad (16)$$

which leads to an equilibrium coverage of

$$c_{eq} = \frac{mk_b}{k_r}. \qquad (17)$$

Hereinafter, this approach is denoted the "ideal model" due to its analogy with the ideal gas model chemical potential (see Figs. 3 and 5.) This result has limited relevance for real systems because the equilibrium coverage is not bound below one, and therefore is clearly unrealistic for high values of binding affinity ($\frac{mk_b}{k_r} > 1$). However, the ideal model is included here because of its theoretical interest. This ideal equilibrium coverage will also be obtained from a kinetic model of the form

$$\frac{dc}{dt} = k_b m - k_r c, \qquad (18)$$

which can also be obtained from the very low coverage limit of Eq. (15) or of Eq. (7). A close inspection of Figs. 2, 3, and 5 shows that its range of validity is even lower than the previous simple model, and it is only applicable at very low coverages, i.e., $c < 0.01$. However, its discussion explains the meaning of the different terms appearing in the kinetic terms, the equilibrium coverage, and the chemical potential of the GHT model introduced here.

III. DISCUSSION

The results presented in this paper show that the detailed counting of ligand-binding possibilities done by the McGhee von Hippel procedure is essential to get an accurate account of the kinetic equation, the equilibrium coverage, or the chemical

potential. The finite-size effects accounted for by this procedure become important already for coverages of the order of 1% to 10% (i.e., $c = 0.01$ to 0.1). Simplified kinetic models as the simple or ideal model presented here for comparison, or other simple models such as the one introduced in Ref. [6], can qualitatively explain some features of the ligand-binding process but fail to give accurate quantitative predictions when coverage effects become important. Additionally, the McGhee-von Hippel-Tonks (GHT) kinetic model presented here is shown to be consistent with the previous result for equilibrium coverage obtained by McGhee and von Hippel [37] and from the Tonks gas chemical potential [6,36]. These clarifications of the kinetic model will provide a relevant tool to understand the available single-molecule data on the binding kinetics of large ligands to long polymers such as, for example, data on SSB binding to ssDNA [2,7,36], histone proteins binding to DNA [43], and Thiocoraline binding to dsDNA [57].

The model presented does not account for the effects of possible diffusion of the ligands along the chain, as they are known to occur in the SSB binding to ssDNA. We performed numerical simulations, including the diffusion of the SSB with the experimentally measured values and found that, in this case, it has subleading effects on the equilibrium coverage.

Another effect not considered in the kinetic model presented here is ligand-ligand interactions, also known as cooperative effects. Cooperative effects, although thought to be negligible in human mitochondrial SSB [36], are known to be present in one of the two main binding modes of E. Coli SSB [2,59]. The inclusion of cooperative effects is a relevant future extension of the ligand-binding kinetic models. They can be taken into account considering the changes in probabilities induced by cooperativity similarly to the procedure in Ref. [37], also extending these results from the equilibrium coverage to the kinetic equations and the chemical potential.

Another relevant extension will be the inclusion of several binding modes of the ligands, as in the case of E. Coli SSB that can be bound to 35 or 65 nucleotides depending on the SSB concentration and the NaCl concentration [59].

Another effect that can be present in real systems is the self-binding of the polymer that can reduce the available binding sites. However, the binding of ligands and its diffusion along the polymer might compete with this self-binding and make them available again. One of the mechanisms for this could be the stiffening of the polymer induced by the ligand binding.

Nucleosome positioning along the DNA is also a relevant case where different additional effects have been reported to be relevant [62–64]: diffusion of the nucleosomes along the DNA, nucleosome cooperativity, DNA unwrapping of nucleosome (i.e., partial binding), and sequence-specific effects.

APPENDIX

The McGhee and von Hippel approach can be considered as a combinatorial method because it uses binomial formulas to derive an analytical expression for the numbers of possible rearrangements of ligands to the one-dimensional polymer lattice.
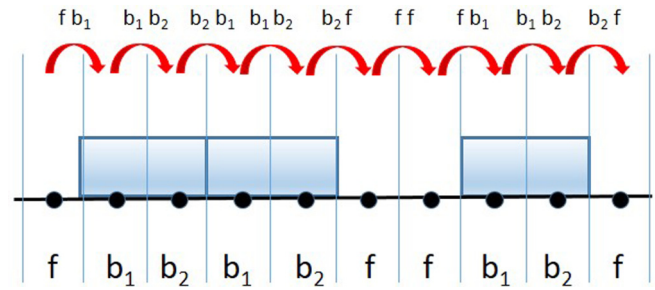

FIG. 6. Schematics of the conditional probabilities used to describe their distribution illustrated for the binding of $n = 3$ ligands with binding mode $m = 2$ bound to a chain of $N = 10$.

First, a counting convention has to be adopted. To count the number of free lattice residues in a gap, we start at the left end of one bound ligand and count by proceeding to the right, one residue at a time, until we reach the left end of the next bound ligand. Using this convention, we can express the overall probability that a given gap is, for example, exactly two free residues long, as the product of the constituent conditional probabilities.

Next, the following notation is adopte: Any bound ligand can be divided into m sections, each one corresponding to the underlying lattice residue; we number these sections, 1, 2, …, m from left to right. Thus we have $(m + 1)$ distinguishable types of lattice residues: a free residue, labeled f; a residue under the number 1 or left end of a bound ligand, labeled b_1; and so on from b_2 up to b_m, the latter representing the right end of a bound ligand. We can thus denote the conditional probabilities used above as a sequence of two such types. For example, $(f b_1)$ is the probability, given a free residue (i.e., an f residue), that the left end of a bound ligand (i.e., a b_1 residue) lies to the immediate right (see Fig. 6).

In principle, there are $(m + 1)^2$ different conditional probabilities that can be expressed by this notation, but not all of them make sense. For example, there are only two types of lattice residues which can lie to the immediate right of a free residue: either another free residue or the first part of a bound ligand, i.e.

$$(f f) + (f b_1) = 1. \tag{A1}$$

Similarly, only a free residue or the left end of a second bound ligand can lie to the immediate right of the right end of a bound ligand, and thus

$$(b_m f) + (b_m b_1) = 1. \tag{A2}$$

Following the McGhee and von Hippel analysis, the probability of a gap having g free lattice residues long is given by

$$P_g = (b_m f)(f f)^{g-1}(f b_1), \tag{A3}$$

where $(b_m f)$ is the probability, having selected the right end of a bound ligand, that the lattice residue to the immediate right is free; $(f f)$ is the probability, given a free residue, that there is a second free residue to the immediate right; and $(f b_1)$ is the probability, given a free residue, that the left end of a bound ligand lies to the immediate right.

If one selects a residue at random, the probability that it is free is $(1 - c)$, where $c = n\,m\,/N$ is the coverage of the chain. Next, we obtain an expression for the conditional probabilities in terms of the model parameters. If one considers the lattice residue immediately to the right of one selected at random, this second residue also, by definition, is selected at random and thus also has a probability of $(1-c)$ of being free. However, due to the nature of the lattice, there are only two ways in which this two-step random selection can be made. Either the first chosen residue is free [a f residue chosen with probability $(1-c)$ and has a free residue to its right—conditional probability (ff)]; or the first chosen residue is the right end of a bound ligand (a b_m residue chosen with probability c/m) and has a free residue to its right [conditional probability $(b_m f)$]. Since the overall probability that the second residue is free must be independent of the method of random selection, we obtain:

$$(1 - c)(ff) + \frac{c}{m}(b_m f) = 1 - c. \tag{A4}$$

Since ligands bound noncooperatively, neither attracting nor repelling one another, we can write

$$(b_m f) = (ff). \tag{A5}$$

We now combine Eqs. (A4) and (A5) with Eqs (A1) and (A2) to obtain the following expressions for the conditional probabilities:

$$(ff) = (b_m f) = \frac{1 - c}{1 - c + c/m}, \tag{A6}$$

and

$$(fb_1) = (b_m b_1) = \frac{c/m}{1 - c + c/m}. \tag{A7}$$

Then, the resulting expression for P_g is given by

$$P_g = \left(\frac{1 - c}{1 - c + c/m} \right)^g \left(\frac{c/m}{1 - c + c/m} \right), \tag{A8}$$

which is the expression used for P_g in Eq. (3).

[1] A. G. Kozlov *et al.*, Intrinsically disordered C-terminal tails of E. coli single-stranded DNA binding protein regulate cooperative binding to single-stranded DNA, J. Mol. Biol. **427**, 763 (2015).

[2] E. Antony and T. M. Lohman, Dynamics of E. coli single stranded DNA binding (SSB) protein-DNA complexes, Semin. Cell Dev. Biol. **86**, 102 (2018).

[3] G. L. Ciesielski *et al.*, Mitochondrial single-stranded DNA-binding proteins stimulate the activity of DNA polymerase γ by organization of the template DNA, J. Biol. Chem. **290**, 28697 (2015).

[4] S. Suksombat, R. Khafizov, A. G. Kozlov, T. M. Lohman, and Y. R. Chemla, Structural dynamics of E. coli single-stranded DNA binding protein reveal DNA wrapping and unwrapping pathways, Elife **4**, e08193 (2015).

[5] C. Maffeo and A. Aksimentiev, Molecular mechanism of DNA association with single-stranded DNA binding protein, Nucleic Acids Res. **45**, 12125 (2017).

[6] J. Jarillo, J. A. Morín, E. Beltrán-Heredia, J. P. G. Villaluenga, B. Ibarra, and F. J. Cao, Mechanics, thermodynamics, and kinetics of ligand binding to biopolymers, PLoS One **12**, e0174830 (2017).

[7] J. A. Morin *et al.*, Mechano-chemical kinetics of DNA replication: Identification of the translocation step of a replicative DNA polymerase, Nucleic Acids Res. **43**, 3643 (2015).

[8] G. Settanni, D. Serquera, P. E. Marszalek, E. Paci, and L. S. Itzhaki, Effects of ligand binding on the mechanical properties of ankyrin repeat protein gankyrin, PLoS Comput. Biol. **9**, e1002864 (2013).

[9] B. Alberts *et al.*, *Molecular Biology of the Cell*, 6th ed. (Garland Science, New York, 2017).

[10] E. C. Cesconetto, F. S. a Junior, F. a P. Crisafuli, O. N. Mesquita, and M. S. Rocha, DNA interaction with Actinomycin D: mechanical measurements reveal the details of the binding data, Phys. Chem. Chem. Phys. **15**, 11070 (2013).

[11] L. Siman, I. S. S. Carrasco, J. K. L. da Silva, M. C. de Oliveira, M. S. Rocha, and O. N. Mesquita, Quantitative Assessment of the Interplay Between DNA Elasticity and Cooperative Binding of Ligands, Phys. Rev. Lett. **109**, 248103 (2012).

[12] F. Cerrón *et al.*, Replicative DNA polymerases promote active displacement of SSB proteins during lagging strand synthesis, Nucleic Acids Res. **47**, 5723 (2019).

[13] A. Bosco, J. Camunas-Soler, and F. Ritort, Elastic properties and secondary structure formation of single-stranded DNA at monovalent and divalent salt conditions, Nucleic Acids Res. **42**, 2064 (2014).

[14] A. Montanari and M. Mézard, Hairpin Formation and Elongation of Biomolecules, Phys. Rev. Lett. **86**, 2178 (2001).

[15] O. A. Saleh, D. B. McIntosh, P. Pincus, and N. Ribeck, Nonlinear Low-Force Elasticity of Single-Stranded DNA Molecules, Phys. Rev. Lett. **102**, 068301 (2009).

[16] D. B. McIntosh and O. A. Saleh, Salt species-dependent electrostatic effects on ssDNA elasticity, Macromolecules **44**, 2328 (2011).

[17] D. B. McIntosh, N. Ribeck, and O. A. Saleh, Detailed scaling analysis of low-force polyelectrolyte elasticity, Phys. Rev. E **80**, 041803 (2009).

[18] M. N. Dessinges, B. Maier, Y. Zhang, M. Peliti, D. Bensimon, and V. Croquette, Stretching Single Stranded DNA, a Model Polyelectrolyte, Phys. Rev. Lett. **89**, 248102 (2002).

[19] R. G. Mortimer, *Physical Chemistry* (Academic Press, Elsevier, 2008).

[20] T. L. Hill, *An Introduction to Statistical Thermodynamics* (Dover Publications, New York, 1986).

[21] P. J. Flory, The configuration of real polymer chains, J. Chem. Phys. **17**, 303 (1949).

[22] P. J. Flory, *Statistical Mechanics of Chain Molecules* (Hanser Publishers, New York, 1989).

[23] J. D. van der Waals, On the continuity of the gaseous and liquid states, Ph.D. thesis, Amsterdam, 1873 (unpublished).

[24] J. A. Barker and D. Henderson, What is liquid? Understanding the states of matter, Rev. Mod. Phys. **48**, 587 (1976).

[25] M. J. Klein, The historical origins of the van der Waals equation, Physica **73**, 28 (1974).

[26] P. Singh, P. Mahata, T. Baumgart, and S. L. Das, Curvature sorting of proteins on a cylindrical lipid membrane tether connected to a reservoir, Phys. Rev. E **85**, 051906 (2012).

[27] S. Aimon, A. Callan-Jones, A. Berthaud, M. Pinot, G. E. S. Toombes, and P. Bassereau, Membrane shape modulates transmembrane protein distribution, Dev. Cell **28**, 212 (2014).

[28] E. Beltrán-Heredia, F.-C. Tsai, S. Salinas-Almaguer, F. J. Cao, P. Bassereau, and F. Monroy, Membrane curvature induces cardiolipin sorting, Commun. Biol. **2**, 225 (2019).

[29] L. Tonks, The complete equation of state of one, two and three-dimensional gases of hard elastic spheres, Phys. Rev. **50**, 955 (1936).

[30] J. K. Percus and G. J. Yevick, Analysis of classical statistical mechanics by means of collective coordinates, Phys. Rev. **110**, 1 (1958).

[31] J. I. Siepmann, I. R. McDonald, and D. Frenkel, Finite-size corrections to the chemical potential, J. Phys. Condenced Matter **4**, 679 (1992).

[32] B. Smit and D. Frenkel, An explicit expression for finite-size corrections to the chemical potential, J. Phys. Condens. Matter **1**, 8659 (1989).

[33] E. Beltrán-Heredia and A. Santos, Fourth virial coefficient of additive hard-sphere mixtures in the Percus-Yevick and hypernetted-chain approximations, J. Chem. Phys. **140**, 134507 (2014).

[34] W.-S. Dai and M. Xie, The equation of state for two-dimensional hard-sphere gases: Hard-sphere gases as ideal gases with multi-core boundaries, Europhys. Lett. **72**, 887 (2006).

[35] S. Luding, Global equation of state of two-dimensional hard-sphere systems, Phys. Rev. E **63**, 042201 (2001).

[36] J. A. Morin *et al.*, DNA synthesis determines the binding mode of the human mitochondrial single-stranded DNA-binding protein, Nucleic Acids Res. **45**, 7237 (2017).

[37] J. D. McGhee and P. H. von Hippel, Theoretical aspects of DNA-protein interactions: co-operative and non-co-operative binding of large ligands to a one-dimensional homogeneous lattice, J. Mol. Biol. **86**, 469 (1974).

[38] R. I. Epstein, Polynucleotide and the ligand, Biophys. Chem. **8**, 327 (1978).

[39] A. Kristiansen, K. M. Vårum, and H. Grasdalen, Quantitative studies of the non-productive binding of lysozyme to partially N-acetylated chitosans. Binding of large ligands to a one-dimensional binary lattice studied by a modified McGhee and von Hippel model, Biochim. Biophys. Acta - Gen. Subj. **1425**, 137 (1998).

[40] A. R. Wolfe and T. Meehan, Use of binding site neighbor-effect parameters to evaluate the interactions between adjacent ligands on a linear lattice. Effects on ligand-lattice association, J. Mol. Biol. **223**, 1063 (1992).

[41] E. Di Cera and Y. Kong, Theory of multivalent binding in one and two-dimensional lattices, Biophys. Chem. **61**, 107 (1996).

[42] K. Laurila, O. Yli-Harja, and H. Lähdesmäki, A protein-protein interaction guided method for competitive transcription factor binding improves target predictions, Nucleic Acids Res. **37**, e146 (2009).

[43] V. B. Teif and K. Rippe, Statistical-mechanical lattice models for protein-DNA binding in chromatin, J. Phys. Condens. Matter **22**, 414105 (2010).

[44] A. V. Morozov, K. Fortney, D. A. Gaykalova, V. M. Studitsky, J. Widom, and E. D. Siggia, Using DNA mechanics to predict in vitro nucleosome positions and formation energies, Nucleic Acids Res. **37**, 4707 (2009).

[45] S. Ghosh and A. Mookerjee, A model for cooperative ligand binding at complementary sites of DNA, Bull. Math. Biol. **48**, 21 (1986).

[46] G. M. Clore, A. M. Gronenborn, and R. W. Davies, Theoretical aspects of specific and non-specific equilibrium binding of proteins to DNA as studied by the nitrocellulose filter binding assay. Co-operative and non-co-operative binding to a one-dimensional lattice, J. Mol. Biol. **155**, 447 (1982).

[47] Y. D. Nechipurenko and G. V. Gursky, Cooperative effects on binding of proteins to DNA, Biophys. Chem. **24**, 195 (1986).

[48] S. Rajendran, M. J. Jezewska, and W. Bujalowski, Human DNA polymerase β recognizes single-stranded DNA using two different binding modes, J. Biol. Chem. **273**, 31021 (1998).

[49] T. Nishio and T. Shimizu, Model analysis of surfactant-polymer interaction as cooperative ligand binding to linear lattice, Biophys. Chem. **117**, 19 (2005).

[50] D. P. Mascotti and T. M. Lohman, Thermodynamic extent of counterion release upon binding oligolysines to single-stranded nucleic acids, Proc. Natl. Acad. Sci. USA **87**, 3142 (1990).

[51] D. Stigter and K. A. Dill, Binding of ionic ligands to polyelectrolytes, Biophys. J. **71**, 2064 (1996).

[52] W. Zhang, J. P. Bond, C. F. Anderson, T. M. Lohman, and M. T. Record, Large electrostatic differences in the binding thermodynamics of a cationic peptide to oligomeric and polymeric DNA, Proc. Natl. Acad. Sci. USA **93**, 2511 (1996).

[53] R. D. Shereda, A. G. Kozlov, T. M. Lohman, M. M. Cox, and J. L. Keck, SSB as an organizer/mobilizer of genome maintenance complexes, Crit. Rev. Biochem. Mol. Biol. **43**, 289 (2008).

[54] S. V. Kuznetsov, A. G. Kozlov, T. M. Lohman, and A. Ansari, Microsecond dynamics of protein-DNA interactions: Direct observation of the wrapping/unwrapping kinetics of single-stranded DNA around the E. coli SSB tetramer, J. Mol. Biol. **359**, 55 (2006).

[55] M. Morse, M. N. Naufer, Y. Feng, L. Chelico, I. Rouzina, and M. C. Williams, HIV restriction factor APOBEC3G binds in multiple steps and conformations to search and deaminate single-stranded DNA, Elife **8**, 1 (2019).

[56] S. Subramanyam, C. D. Kinz-Thompson, R. L. Gonzalez, and M. Spies, Observation and analysis of RAD51 nucleation dynamics at single-monomer resolution, Methods. Enzymol. **600**, 201 (2018).

[57] J. Camunas-Soler, M. Manosas, S. Frutos, J. Tulla-Puche, F. Albericio, and F. Ritort, Single-molecule kinetics and footprinting of DNA bis-intercalation: The paradigmatic case of Thiocoraline, Nucleic Acids Res. **43**, 2767 (2015).

[58] A. Velázquez-Campoy, Ligand binding to one-dimensional lattice-like macromolecules: Analysis of the McGhee-von Hippel theory implemented in isothermal titration calorimetry, Anal. Biochem. **348**, 94 (2006).

[59] A. G. Kozlov, M. K. Shinn, and T. M. Lohman, Regulation of nearest-neighbor cooperative binding of E. coli SSB protein to DNA, Biophys. J. **117**, 2120 (2019).

[60] R. M. Corless, G. H. Gonnet, D. E. G. Hare, D. J. Jeffrey, and D. E. Knuth, On the Lambert *W* function, Adv. Comput. Math. **5**, 329 (1996).

[61] R. M. Corless, G. H. Gonnet, D. E. G. Hare, D. J. Jeffrey, and D. E. Knuth, Lambert's *W* function in Maple, Maple Tech. Newsl. **9**, 12 (1993).

[62] W. Möbius, B. Osberg, A. M. Tsankov, O. J. Rando, and U. Gerland, Toward a unified physical model of nucleosome patterns flanking transcription start sites, Proc. Natl. Acad. Sci. USA **110**, 5719 (2013).

[63] G. Chevereau, L. Palmeira, C. Thermes, A. Arneodo, and C. Vaillant, Thermodynamics of Intragenic Nucleosome Ordering, Phys. Rev. Lett. **103**, 188103 (2009).

[64] J. Riposo and J. Mozziconacci, Nucleosome positioning and nucleosome stacking: Two faces of the same coin, Mol. Biosyst. **8**, 1172 (2012).