Computational and Structural Biotechnology Journal 20 (2022) 521–533

COMPUTATIONAL AND STRUCTURAL BIOTECHNOLOGY JOURNAL

journal homepage: www.elsevier.com/locate/csbj

Cooperative kinetics of ligand binding to linear polymers

Juan P.G. Villaluenga [a,*], Francisco Javier Cao-García [a,b]

[a] Departamento de Estructura de la Materia, Física Térmica y Electrónica, Universidad Complutense de Madrid, Plaza de Ciencias, 1, 28040 Madrid, Spain
[b] Instituto Madrileño de Estudios Avanzados en Nanociencia, IMDEA Nanociencia, Calle Faraday, 9, 28049 Madrid, Spain

ARTICLE INFO

Article history:
Received 30 September 2021
Received in revised form 27 December 2021
Accepted 30 December 2021
Available online 6 January 2022

Keywords:
Cooperativity
Ligand binding
SSB
DNA
RNA
Polymer
Affinity
Kinetics

ABSTRACT

Ligands change the chemical and mechanical properties of polymers. In particular, single strand binding protein (SSB) non-specifically bounds to single-stranded DNA (ssDNA), modifying the ssDNA stiffness and the DNA replication rate, as recently measured with single-molecule techniques. SSB is a large ligand presenting cooperativity in some of its binding modes. We aim to develop an accurate kinetic model for the cooperative binding kinetics of large ligands. Cooperativity accounts for the changes in the affinity of a ligand to the polymer due to the presence of another bound ligand. Large ligands, attaching to several binding sites, require a detailed counting of the available binding possibilities. This counting has been done by McGhee and von Hippel to obtain the equilibrium state of the ligands-polymer complex. The same procedure allows to obtain the kinetic equations for the cooperative binding of ligands to long polymers, for all ligand sizes. Here, we also derive approximate cooperative kinetic equations in the large ligand limit, at the leading and next-to-leading orders. We found cooperativity is negligible at the leading-order, and appears at the next-to-leading order. Positive cooperativity (increased affinity) can be originated by increased binding affinity or by decreased release affinity, implying different kinetics. Nevertheless, the equilibrium state is independent of the origin of cooperativity and only depends on the overall increase in affinity. Next-to-leading approximation is found to be accurate, particularly for small cooperativity. These results allow to understand and characterize relevant ligand binding processes, as the binding kinetics of SSB to ssDNA, which has been reported to affect the DNA replication rate for several SSB-polymerase pairs.

1. Introduction

Understanding the kinetics of binding of large ligands to long polymers is a very relevant biological problem. The ligand binding to polymers changes their mechanical and chemical properties, plays a relevant role regulating biological functions, shrinks or expands the binding regions of the polymer (by stiffening these regions) and passivates polymer regions for the binding of other units of the ligand, among other effects. The binding may be specific or non-specific depending on whether the ligand binds to the polymer with or without sequence preference. The binding is specific between a basic polypeptide and a polynucleotide, and between an oppositely charged ionic surfactant and a linear polyelectrolyte, among other systems. One of the more attractive problems when the binding is specific involves protein molecules as ligand and DNA as polymer [1–13,43].

Some ligands bind to few sites, such as Thiocoraline binding to dsDNA (double-stranded DNA), human polymerase β binding to ssDNA (single-stranded DNA), binding of lysozyme to chitosans [14–16]. In contrast, 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 [2,5,17]. These ligand bindings have relevant implications, as degrade the polymer (chitosans) or stimulate DNA replication [10,18].

There are relevant systems that exhibit either attractive or repulsive interactions between bound ligands (cooperativity), which perturb the binding process. They include: The binding of ionic surfactants to the charged linear polyelectrolyte changing its physical and chemical properties [19]. The binding of lysozyme to chitosans mediating its degradation [14]. The binding of a variety of proteins and intercalating agents to nucleic acids, relevant for their maintenance [1,9,17,20–25]. In particular, a binding mode of E. Coli SSB to DNA relevant for DNA replication [1–2,26]. In contrast, there are systems in which the ligands do not have mutual interactions (non-cooperative binding), like the binding of oligolisines to nucleic acids, mono- and multivalent

* Corresponding author.
E-mail addresses: jpgarcia@ucm.es (J.P.G. Villaluenga), francao@ucm.es (F.J. Cao-García).

https://doi.org/10.1016/j.csbj.2021.12.043

ions to polyelectrolytes, or polypeptides to polynucleotides; which change the polymer physical and chemical properties [27–29]. Here, we generalize a non-cooperative kinetic model to the cooperative case.

Single-molecule measurements allow very detailed tracking of binding processes, calling for more accurate models. In particular, recently, the binding equilibrium and kinetics of human mitochondrial and E. Coli SSBs to ssDNA have been studied using optical tweezers [5,26]. As SSB is one of the main components of the replisome, the characterization of the SSB to ssDNA is key to understanding the DNA replication [10]. These studies call for accurate kinetic models for high coverage situations. Although the equilibrium for binding large ligands to long polymers has been studied in detail, it is not the case for the kinetics, as discussed below.

The equilibrium state of large ligands bound to macromolecules has been addressed in detail. Many equilibrium models are performed in the framework of statistical mechanical lattice models. In this type of models, the macromolecule is commonly considered as a one-dimensional lattice with a given number of potential binding sites (homogeneously or heterogeneously distributed), and constituted by elementary units (e.g., monomers for polymer, basepairs for DNA), and any bound ligand molecule occupies a certain number of elementary units. Several methods of constructing lattice models of ligand binding to macromolecules have been formulated. There are combinatorial methods that use binomial formulae to derive analytical expressions for the numbers of possible rearrangements of ligands along the macromolecule [16,23,19–21,30–32]. In other type of methods, the system is characterized by a mathematical expression related to the partition function. The elementary units of the macromolecule are associated with different states (e.g., bound/unbound). The states of the whole system can be calculated as different combinations of states of the elementary units. The binding probabilities calculated for the lattice models are then equal to the fraction of molecules in a given state within the statistical ensemble consisting of many identical systems. In a statistical ensemble, each state of the system is characterized by a weight. The partition function is equal to the sum of the weights of all possible states [24,33–35]. The transfer matrix method is based on the construction of transfer matrices for each elementary unit, whose elements contain the probabilities to find the lattice unit in a given state [24,36,37].

These equilibrium studies can be extended to kinetics models [38,39], using using the McGhee and von Hippel detailed binding site counting procedure [30]. Recently, in Ref. [38], we clarified the non-cooperative kinetic equation, equilibrium coverage (McGhee and von Hippel result [30]) and its chemical potential (which matches the Tonks gas chemical potential [40,41]). We also showed the limited validity of a simple coarse-grained kinetic model and of the ideal gas kinetics model [4], which are only valid for low and very low coverages, respectively. Here, we derive the cooperative binding kinetic equation, extending our approach [38] to the cooperative case (where the ligand affinity is modified by the presence of other ligands bound to the polymer), and deriving the approximate kinetic equations at the large ligand size limit at the leading and next-to-leading order.

In Subsection II.A, we briefly summarize the main results of the model for non-cooperative ligand binding [38]. In addition, for non-cooperative binding, we compute the next to leading order in the large ligand approximation (i.e., large number of binding sites per ligand), which is useful for comparison with the cooperative results. In Section II.B, we extend the model for cooperative ligand binding. In section II.C, we extend the results for the more general case in which the cooperativity can affect both the binding and the release rates. Finally, Section III discusses the results.

2. Ligand binding kinetic equation

We derive the kinetic equation for the binding of large ligands to a long polymer using the procedure proposed by McGhee and von Hippel to count the possible binding sites [30]. The polymer is represented by a linear array of N identical binding sites, and when the ligand molecule binds to the polymer covers m sites (i.e., make inaccessible to another ligand). In addition, the polymer is assumed to be infinitely long, i.e., $N \gg m$. The fraction of polymer binding sites covered by the ligand is given by the coverage, $c = nm/N$, where n is the number of ligands bound to the polymer.

2.1. Non-cooperative binding

This subsection reviews previous results for non-cooperative binding [30,38]. We additionally derive the next-to-leading order in large ligand approximation (large m). All these results are useful for comparison with the cooperative binding results presented in the next subsections.

The kinetic equation describing the time variation in the number of ligands bound to the polymer is [38]

$$\frac{dn}{dt} = k_b(n+1)\bar{s} - k_r n, \tag{1}$$

where $\bar{s}$ is the average number of free binding sites per gap, $n+1$ the number of gaps, k_b is the binding kinetic constant, and k_r is the release kinetic constant. The binding kinetic constant k_b generally depends on the free ligand concentration L, and usually linearly $k_b = K_b L$, where K_b represents the association constant. A gap (of size g) is a set of g consecutive free ligand binding sites (free lattice residues). Ligands cannot bind to gaps smaller than its size, $g < m$.

Following the McGhee and von Hippel reasoning, the average number of free binding sites per gap is

$$\bar{s} = \sum_{g=m}^{N} (g-m+1)P_g, \tag{2}$$

where P_g is the probability that any particular gap between two bound ligands is exactly g free polymer residues long, which is given by [30]

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

(See Appendix for the derivation). For an infinite polymer, $N \to \infty$, the average number of free ligand binding sites per polymer [using Eqs. (2) and (3)] turns out to be

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

where we have assumed that $(n+1) \approx n$ (reasonable for an infinite polymer). This result allows to write the kinetic equation in terms of the coverage as

$$\frac{dc}{dt} = k_b m(1-c)\left(\frac{1-c}{1-c+c/m}\right)^{m-1} - k_r c. \tag{5}$$

This equation implies that, if we start with a naked polymer, the coverage will increase until it reaches an equilibrium value where the binding and release terms balance each other. See Fig. 1(Left). The equilibrium coverage can be obtained numerically from Eq. (5) imposing $dc/dt = 0$. It is found that it increases as a function of the ratio between the binding and release rates, $K = k_b/k_r$. See Fig. 2. On the other hand, the dissociation kinetics is modelled using Eq. (5) by setting $k_b = 0$ (simulating setting the ligand con-

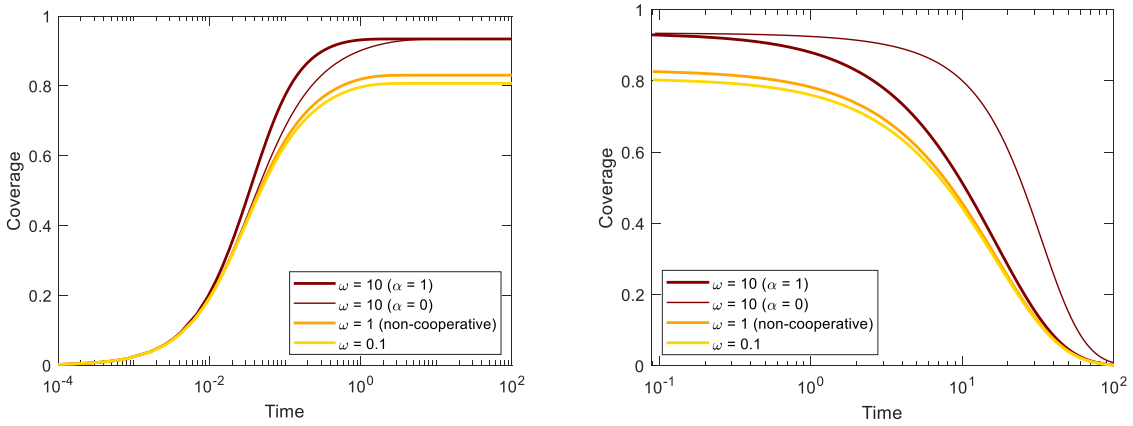

Fig. 1. Attachment and detachment kinetics. Polymer coverage as a function of time compared for non-cooperative ($\omega = 1$) and cooperative ($\omega \neq 1$) cases. (Left) Starting with a naked polymer, ligands binding to m = 30 sites, binding rate k_b = 0.8 s^{-1}, and release rate k_r = 0.06 s^{-1}. (Typical values of SSB binding to ssDNA [5,10,12]). (Right) Starting with the previous equilibrium coverages the detachment dynamics is induced setting $k_b = 0$, simulating setting ligand concentration in solution to zero. The non-cooperative (ω = 1) dynamics is computed from Eq. (5). The cooperative (attractive ω = 10, and repulsive ω = 0.1) dynamics are computed from Eq. (20), and from Eq. (40) for the extreme cases of cooperativity only affecting the binding (α = 1) or only affecting the release ($\alpha = 0$). The activation state parameter α measures the impact of cooperativity on the activation energy of the binding process, as introduced below in Eq. (31) of Section II.C.

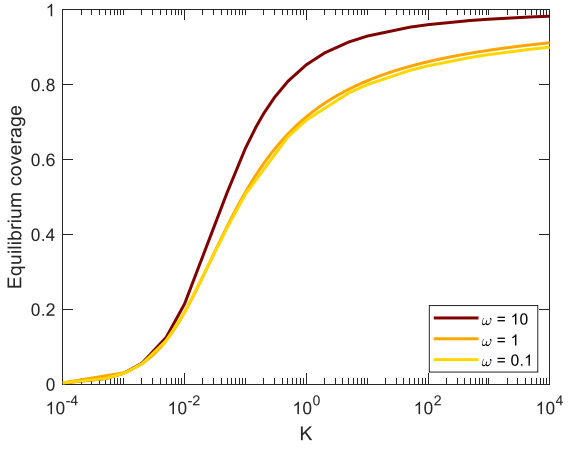

Fig. 2. Equilibrium coverage as a function of the ratio of the binding and release rates, $K = k_b/k_r$, compared for the non-cooperative and cooperative cases for a ligand binding to m = 30 sites. The non-cooperative (ω = 1) equilibrium is computed from the stationary state of Eq. (5), while the cooperative (attractive ω = 10, and repulsive ω = 0.1) equilibria are computed from the stationary state of Eq. (20).

centration in solution to zero). Fig. 1(Right) shows the dissociation kinetics starting with a polymer occupied with the corresponding equilibrium state coverage in each case.

2.1.1. Small ligand binding

2.1.1.1. Single binding ligand (m = 1). When the ligand binds to a single site the kinetic equation, Eq. (5), simplifies to

$$\frac{dc}{dt} = k_b(1-c) - k_r c. \tag{6}$$

The first term provides a coverage increase proportional to the non-covered fraction of the polymer, and the second term decreases the coverage proportionally to the coverage. This results in an equilibrium coverage, c_{eq}, obtained for $dc/dt = 0$,

$$K = \frac{c_{eq}}{1-c_{eq}}, \tag{7}$$

or more explicitly,

$$c_{eq} = \frac{k_b}{k_b+k_r} = \frac{K}{1+K} = \frac{1}{1+\frac{1}{K}} = 1 - \frac{1}{1+K}, \tag{8}$$

where the different ways of expressing the result are useful for comparison with other results. Please, note that Eq. (8) is equivalent to classical Langmuir absorption model [42].

2.1.2. Large ligand binding ($m \gg 1$)

There are relevant cases where the ligand is large and binds to many binding sites, like E. Coli and human mitochondrial SSB that bind 30–70 nucleotides of ssDNA [2,5,17]. For applications to these cases, the large ligand approximation is useful, $m \gg 1$, which accounts to consider that the ligand binds to many binding sites, further simplifying the kinetic equations.

2.1.2.1. Leading order for large ligands ($m \gg 1$). At leading order in the large m limit

$$\left(\frac{1-c}{1-c+c/m}\right)^{m-1} \simeq \exp\left(\frac{-c}{1-c}\right). \tag{9}$$

Then, the kinetic equation becomes

$$\frac{dc}{dt} = k_b m(1-c)\exp\left(\frac{-c}{1-c}\right) - k_r c, \tag{10}$$

which shows that the binding is inhibited by the additional exponential factor (not present in the $m = 1$ case). This factor accounts for the reduction in available binding possibilities, due to the random location of the large ligands that decrease the gaps of size greater than m.

The previous kinetic equation leads to the equilibrium coverage, c_{eq},

$$\ln(Km) = \ln\left(\frac{c_{eq}}{1-c_{eq}}\right) + \frac{c_{eq}}{1-c_{eq}}, \tag{11}$$

or equivalently,

$$c_{eq} = \frac{W(mK)}{1+W(mK)} = \frac{1}{1+\frac{1}{W(mK)}} = 1 - \frac{1}{1+W(mK)}, \tag{12}$$

where $W(z)$ is the so-called Lambert function, which is the solution of the equation $We^W = z$, where z is a complex number. As Lambert function is an increasing function on its argument, this implies that coverage increases for increasing binding mode m provide the ligand K (affinity) stays constant. See Fig. 3. While if the affinity is per site, we have that the affinity for the mode m, K_m, is related to the affinity of the mode $m = 1$, K_1, by $K_m = K_1^m$. In this

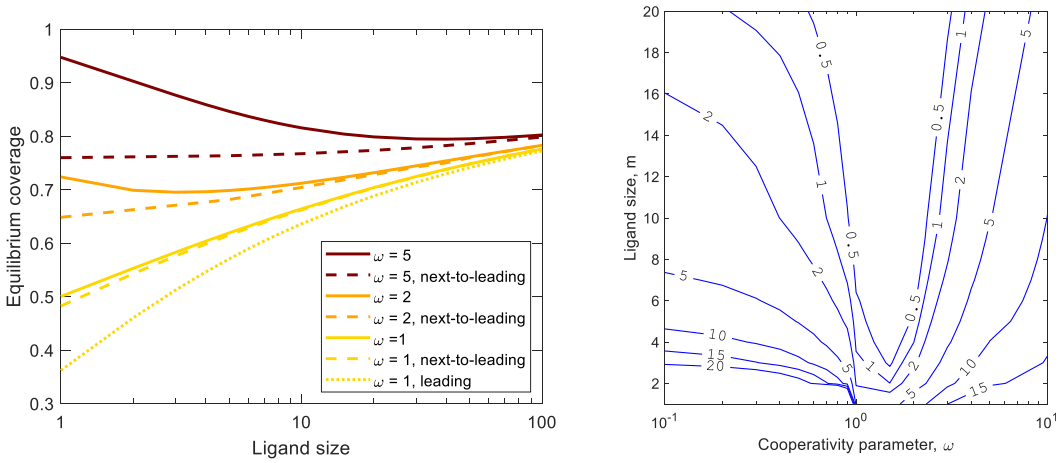

Fig. 3. Comparison of exact and large m approximation results. (Left) Equilibrium coverages as a function of the ligand size m. (Right) Relative error isolines of the next-to-leading order in the large m approximation. The results for the non-cooperative case are computed from the stationary state of Eq. (5) (ω = 1, non-cooperative), Eq. (14) (ω = 1, next-to-leading), and Eq. (10) (ω = 1, leading). The results for the cooperative cases (ω = 5 and ω = 2) are computed from the stationary state of Eq. (40), and their next-to-leading approach from Eq. (53), while their leading approach would coincide with the non-cooperative leading approach results (ω = 1, leading), as derived in the text. The results in this figure are computed using K = 1.

later case, the coverage increases with m, when $K_1 > 1$ (as derived from $(m+1)K_1^{m+1} > mK_1^m$ in the large m limit, while decreases with m, if $K_1 < 1$) (See Fig. 4).

2.1.2.2. Next-to-leading-order for large ligands ($m \gg 1$). Improving the approximation going to next-to-leading order in large m gives a useful expression for later comparison. We have

$$\left(\frac{1-c}{1-c+c/m}\right)^{m-1} \simeq \exp\left(\frac{-c}{1-c}\right)\left(1+\frac{(2-c)c}{2m(1-c)^2}\right). \tag{13}$$

Then, the kinetic equation for the coverage becomes

$$\frac{dc}{dt} = k_b m(1-c)\exp\left(\frac{-c}{1-c}\right)\left(1+\frac{(2-c)c}{2m(1-c)^2}\right) - k_r c. \tag{14}$$

The equilibrium coverage c_{eq} can be obtained from Eq. (14) imposing $dc/dt = 0$,

$$\ln(Km) = \ln\left(\frac{c_{eq}}{1-c_{eq}}\right) + \frac{c_{eq}}{1-c_{eq}} + \ln\left(1-\frac{(2-c_{eq})c_{eq}}{2m(1-c_{eq})^2}\right). \tag{15}$$

The last term provides the next-to-leading order correction in the $1/m$ expansion to the equilibrium coverage. We can compute this correction perturbatively substituting in Eq. (15) the expression

$$c_{eq} = c_{eq}^{(0)} + \frac{c_{eq}^{(1)}}{m}. \tag{16}$$

$c_{eq}^{(0)}$ is the solution for the kinetic equation in a non-cooperative system at the leading order in large m, Eq. (11). $c_{eq}^{(1)}$ is the first-order correction in $1/m$, whose solution is

$$c_{eq}^{(1)} = \frac{\left(2-c_{eq}^{(0)}\right)\left(c_{eq}^{(0)}\right)^2}{2}. \tag{17}$$

This leads to an equilibrium coverage of

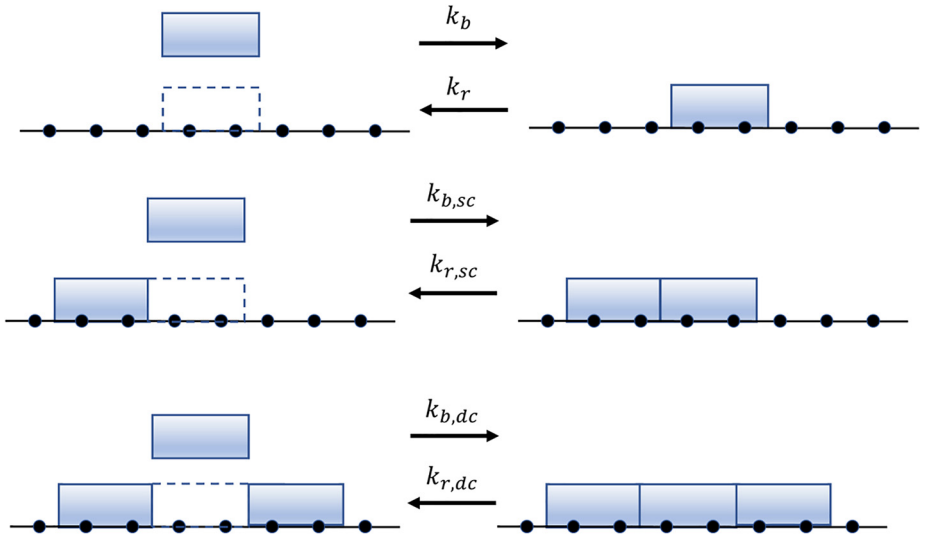

Fig. 4. Binding and release processes to isolated, singly contiguous (sc), and doubly contiguous (dc) sites (respectively).

$$c_{eq} = \left[1 - \frac{1}{1+W(mK)}\right] \cdot \left[1 + \frac{1}{2m}\left\{1 - \frac{1}{[1+W(mK)]^2}\right\}\right], \tag{18}$$

where the second bracket gives the next-to-leading order correction in the large m expansion. As the Lambert function is positive (for positive arguments) the next-to-leading order correction increases the estimate of the equilibrium coverage in a relative amount of the order $1/m$.

2.2. Cooperative binding case

Cooperative effects are present in many systems, where the (direct or indirect) interactions between ligands bound to the polymer are non-negligible [33,36,22–24]. Cooperative effects are known to be present in one of the two main binding modes of E. Coli SSB to DNA [2,17].

To analyze the cooperative effects we follow here the McGhee and von Hippel approach [30,38]. The interaction between ligands is allowed between nearest neighbors when they are bound without intervening free lattice residues. This restriction results in three different types of ligand binding sites: (i) an isolated site to which a ligand binds with a binding constant, k_b; (ii) a singly contiguous site to which a ligand binds with binding constant $k_b\ \omega$; and (iii) a doubly contiguous site to which a ligand binds with binding constant $k_b\ \omega^2$ (see Fig. 4). The cooperativity parameter ω is the equilibrium constant for the process of moving a bound ligand from an isolated site to a singly contiguous site, or from a singly contiguous site to a doubly contiguous site. (See Appendix). For $\omega > 1$, the ligands attract each other, and the binding is positively cooperative; for $\omega < 1$, the ligands repel each other, and the binding is negatively cooperative; for $\omega = 1$, the binding is non-cooperative.

In a cooperative-ligand system, the kinetic equation describing the time variation in the number of ligands bound to the polymer can be written as

$$\frac{dn}{dt} = k_b n\left(\bar{s}_i + \omega\bar{s}_{sc} + \omega^2\bar{s}_{dc}\right) - k_r n, \tag{19}$$

where, as previously noted, we have assumed that $(n+1) \approx n$ for an infinite polymer. We defined $\bar{s}_i$ as the average number of free isolated binding sites per gap, $\bar{s}_{sc}$ as the average number of free singly contiguous binding sites per gap, and $\bar{s}_{dc}$ as the average number of free doubly contiguous binding sites per gap. (See Appendix for the derivation of the corresponding average number of sites). Following McGhee and von Hippel reasoning, the equation describing the time variation in the coverage becomes

$$\frac{dc}{dt} = k_b m(1-c)\left[\frac{(2\omega-1)(1-c)+c/m-R}{2(\omega-1)(1-c)}\right]^{m-1}\left(\frac{1-c-c/m+R}{2(1-c)}\right)^2 - k_r c, \tag{20}$$

where the function R is defined as follows

$$R = \sqrt{(1-c-c/m)^2 + 4\omega(1-c)c/m}. \tag{21}$$

It is noted that the Eq. (20) cannot be applied directly in the case of non-cooperative binding because it diverges when the cooperativity parameter equals unity. However, we verified that, using L'Hôpital's rule and the properties of limits, Eq. (20) becomes Eq. (5) in the limit $\omega \to 1$, as expected.

Fig. 1 shows the evolution of the coverage in a cooperative-ligand system, which has been calculated numerically using Eq. (20) and compares it with the non-cooperative ligand system.

2.2.1. Small ligand binding

2.2.1.1. Single binding ligand ($m = 1$). From Eq. (20), when $m = 1$ the kinetic equation describing the time variation in the coverage turns out to be

$$\frac{dc}{dt} = k_b(1-c)\left(\frac{1-2c+\sqrt{4c(\omega-1)(1-c)+1}}{2(1-c)}\right)^2 - k_r c, \tag{22}$$

where we see that increasing the cooperativity parameter ω increases the binding affinity of the ligand. The equilibrium coverages can be obtained from Eq. (22), by imposing $dc/dt = 0$, as

$$K = \frac{4c_{eq}(1-c_{eq})}{\left[1-2c_{eq}+\sqrt{4c_{eq}(\omega-1)(1-c_{eq})+1}\right]^2} \tag{23}$$

The equilibrium coverage equation for the non-cooperative case, Eq. (7), is consistently recovered for $\omega = 1$.

2.2.2. Large ligand binding ($m \gg 1$)

As previously commented, a large ligand can bind to many binding sites (e.g., SSB to ssDNA). This enables us to approximate the expression assuming $m \gg 1$, for the factors in Eq. (20). We have that

$$\left[\frac{(2\omega-1)(1-c)+c/m-R}{2(\omega-1)(1-c)}\right]^{m-1} \simeq \exp\left(\frac{-c}{1-c}\right)\left[1+\frac{\{2c(\omega-1)+2-c\}c}{2m(1-c)^2}\right], \tag{24}$$

and

$$\left(\frac{1-c-c/m+R}{2(1-c)}\right)^2 \simeq 1 + \frac{2(\omega-1)c}{m(1-c)}. \tag{25}$$

These approximations lead to an approximate kinetic equation for the coverage with large ligands (large m)

$$\frac{dc}{dt} = k_b m(1-c)\exp\left(\frac{-c}{1-c}\right)\left[1+\frac{\{2c(\omega-1)+2-c\}c}{2m(1-c)^2}\right]\left[1+\frac{2(\omega-1)c}{m(1-c)}\right] - k_r c. \tag{26}$$

2.2.2.1. Leading order for large ligands ($m \gg 1$). When only the leading order terms in large m are retained, the cooperative kinetic equation [Eq. (26)] reduces to the non-cooperative leading order equation, Eq. (10). This result implies that the equilibrium coverage is also the same at leading order in large m. Cooperative effects are only present at next-to-leading order in large m. This indicates that isolated binding sites are the dominant contribution for large ligands.

2.2.2.2. Next-to-leading-order for large ligands ($m \gg 1$). We extend here the previous results up to next-to-leading-order in large m. We now conserve first-order terms in $1/m$ (and neglect higher orders) to obtain from Eq. (26) the kinetic equation

$$\frac{dc}{dt} = k_b m(1-c)\exp\left(\frac{-c}{1-c}\right)\left[1+\frac{(2\omega-1)(2-c)c}{2m(1-c)^2}\right] - k_r c. \tag{27}$$

For $\omega = 1$, we consistently recover the analogous non-cooperative result, Eq. (14). For $\omega > 1$, the cooperativity increases the effective binding rate (first term of the right-hand side of the equation), while for $\omega < 1$, the effective binding rate is reduced.

The equilibrium coverage is given by the stationary state of Eq. (27) (i.e., $dc/dt = 0$),

$$\ln(Km) = \ln\left(\frac{c_{eq}}{1-c_{eq}}\right) + \frac{c_{eq}}{1-c_{eq}} + \ln\left(1 - \frac{(2\omega-1)(2-c_{eq})c_{eq}}{2m(1-c_{eq})^2}\right). \quad (28)$$

This equation gives the equilibrium coverage for cooperative large ligands. We solve it perturbatively, using $c_{eq} = c_{eq}^{(0)} + c_{eq}^{(1)}/m$, and found

$$c_{eq}^{(1)} = \left(c_{eq}^{(0)}\right)^2\left(2 - c_{eq}^{(0)}\right)\left(\omega - \frac{1}{2}\right). \quad (29)$$

Thus, the coverage for cooperative large ligands is given by

$$c_{eq} = \left[1 - \frac{1}{1+W(mK)}\right] \cdot \left[1 + \frac{1}{m}\left(\omega - \frac{1}{2}\right)\left\{1 - \frac{1}{[1+W(mK)]^2}\right\}\right]. \quad (30)$$

For $\omega = 1$, the analogous non-cooperative equilibrium equations are recovered, Eq. (15), (17) and (18). Positive cooperativity ($\omega > 1$) increases coverage, while negative cooperativity ($\omega < 1$) decreases coverage, as Eq. (30) shows. This result is consistent with the attractive effect of positive cooperativity because of the resulting increased effective binding energy. (The same argument in the opposite direction holds for negative cooperativity.)

Fig. 3 compares the previous approximations for equilibrium coverage. Interestingly, the next-to-leading order is an excellent approximation even for relatively low values of m. Fig. 3 illustrates that positive cooperativity increases coverage. Note that, for intermediate cooperativity strength (e.g., $\omega = 2$), coverage decreases first and after increases as a function of the ligand bind size m. This slope change is due to the competition of two effects. On the one hand, in the absence of cooperativity, the coverage grows with ligand size m. On the other hand, positive cooperativity ($\omega > 1$) increases the coverage, thanks to the attractive interaction between ligands. However, this cooperativity-induced coverage increase is greater for small ligands, as there are more cooperative interactions per covered monomer (higher density of ligand–ligand interactions).

2.3. General cooperative case

In the previous subsection, we considered cooperative binding and we implicitly assumed that the cooperativity only affected the binding process. However, positive cooperativity can arrive by an enhancement of the binding or by an inhibition of the release (or even both). They correspond, respectively, to an increase in the binding rate or a decrease in the release rate (for the neighbour site). Analogously, negative cooperativity can arise as an inhibition of the binding or an enhancement of the release. In this subsection, we generalize the previous model to cooperative effects on both, binding and release processes.

2.3.1. General kinetic equation

A more general kinetic equation can be derived considering that cooperativity can affect the binding and release rates. Considering the case of singly contiguous ligands, Arrhenius-type equations were proposed for the binding kinetic constant, $k_{b,sc}$, and the release kinetic constant, $k_{r,sc}$,

$$k_{b,sc} = k_b e^{-\frac{\alpha G_\omega}{k_B T}}, \quad (31)$$

$$k_{r,sc} = k_r e^{\frac{(1-\alpha)G_\omega}{k_B T}}, \quad (32)$$

where G_ω is the Gibbs free energy of interaction between adjacent ligands in the polymer. The parameter α, which is named as activation state parameter, measures the lowering of the energy barrier between the binding process and the release process. Dividing the kinetic constants, we obtain the equilibrium constant for the processes

$$K_{sc} = \frac{k_{b,sc}}{k_{r,sc}} = \frac{k_b}{k_r} e^{-\frac{G_\omega}{k_B T}} = K e^{-\frac{G_\omega}{k_B T}}. \quad (33)$$

Similarly, considering doubly contiguous ligands, the kinetic constants are

$$k_{b,dc} = k_b e^{-\frac{2\alpha G_\omega}{k_B T}}, \quad (34)$$

$$k_{r,dc} = k_r e^{\frac{2(1-\alpha)G_\omega}{k_B T}}, \quad (35)$$

and the corresponding equilibrium constant is written as

$$K_{dc} = \frac{k_{b,dc}}{k_{r,dc}} = \frac{k_b}{k_r} e^{-\frac{2G_\omega}{k_B T}} = K e^{-\frac{2G_\omega}{k_B T}}. \quad (36)$$

Please, note that Eqs. (33) and (36) are equivalent to Gibbs free energy isotherm equation. See Fig. 5 for examples of free energy landscapes giving these rates and kinetic constant relations.

As we noted before, the cooperativity parameter ω is the equilibrium constant for the process of moving a bound ligand from an isolated site to a singly contiguous site,

$$\omega = \frac{K_{sc}}{K} = e^{-\frac{G_\omega}{k_B T}}. \quad (37)$$

This allows to restate all the equilibrium constants, and binding and release rates for the cooperative kinetics in terms of their non-cooperative analogs and the model parameters

$$\begin{aligned} &K_{sc} = K\omega;\ K_{dc} = K\omega^2; \\ &k_{b,sc} = k_b\omega^\alpha;\ k_{b,dc} = k_b\omega^{2\alpha}; \\ &k_{r,sc} = k_r\omega^{\alpha-1};\ k_{r,dc} = k_r\omega^{2(\alpha-1)}. \end{aligned} \quad (38)$$

Hence, the time variation in the number of ligands bound to the polymer (kinetic equation) is now given by

$$\frac{dn}{dt} = k_b n\left(\bar{s}_i + \omega^\alpha \bar{s}_{sc} + \omega^{2\alpha}\bar{s}_{dc}\right) - k_r n\left(\bar{l}_i + \omega^{\alpha-1}\bar{l}_{sc} + \omega^{2(\alpha-1)}\bar{l}_{dc}\right), \quad (39)$$

where we define $\bar{l}_i$ as the average fraction of isolated attached ligands in the polymer, $\bar{l}_{sc}$ as the average fraction of attached ligands with a single contiguous ligand in the polymer, and $\bar{l}_{dc}$ as the average fraction of attached ligands with two contiguous ligands in the polymer. (See Appendix for the derivation of the corresponding expressions) The first term of the right-hand side of Eq. (39) is related to the ligand binding process to the polymer, while the second term of the right-hand side is related to the ligand releasing process.

After computing these average fractions ($\bar{s}_i, \bar{s}_{sc}, \bar{s}_{dc}, \bar{l}_i, \bar{l}_{sc}$, and $\bar{l}_{dc}$, see Appendix), we can express the kinetic equation [Eq. (39)] in terms of the conditional probabilities,

$$\frac{dc}{dt} = k_b c\left(\frac{(b_m f)(ff)^{m-1}}{1-(ff)}\left[(ff) + \omega^\alpha(1-(ff))\right]^2\right) - k_r c\left[(b_m f) + \omega^{\alpha-1}(1-(b_m f))\right]^2. \quad (40)$$

The conditional probability (ff) is the probability that, when you are in a free binding site, finding that the next binding site is also free, while $(b_m f)$ is the probability that, when you are in the binding site bound to the rightmost extreme of the ligand, the next binding site is free. See Fig. 6. These two conditional probabilities are related by

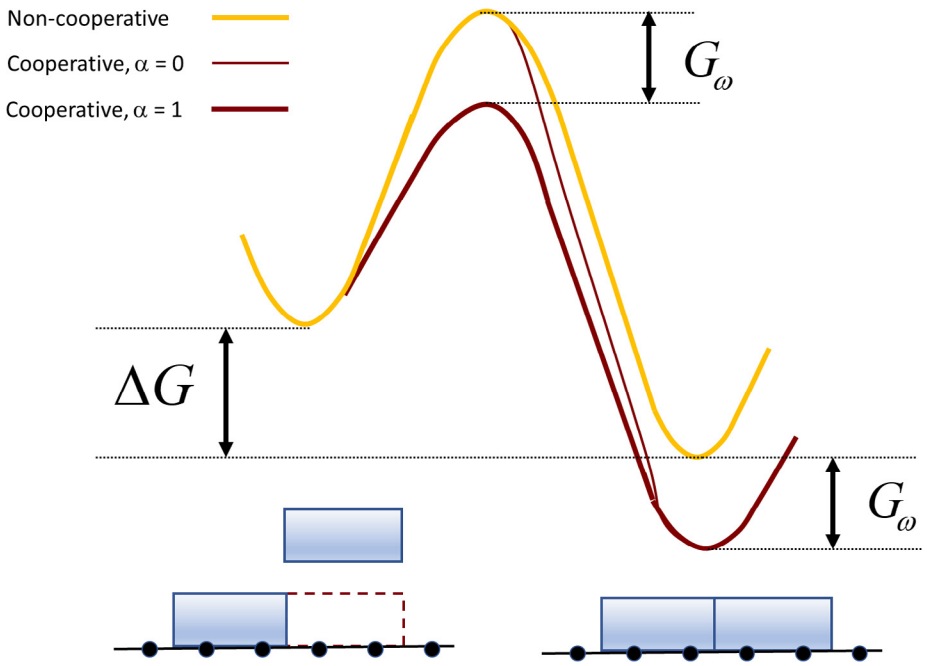

Fig. 5. Free energy landscapes for singly contiguous cooperative binding processes compared to the non-cooperative reaction potential. For cooperative binding, two cases are shown, i.e., enhanced binding ($\alpha = 1$) and inhibited release ($\alpha = 0$). In both cases the free energy of the final state is decreased by an amount G_ω. For enhanced binding ($\alpha = 1$) the activation state also decreases its free energy by an amount G_ω; thus, the energy barrier for the binding process is reduced, while the energy barrier for the release process remains the same. For inhibited release ($\alpha = 0$) the free energy of the activation state does not change; thus, the energy barrier for the binding process is the same as for the non-cooperative case, while the energy barrier for the release process is increased. Both cooperative cases have the same equilibria, as both have the same free energy in the initial and final states, but they differ in the kinetics.

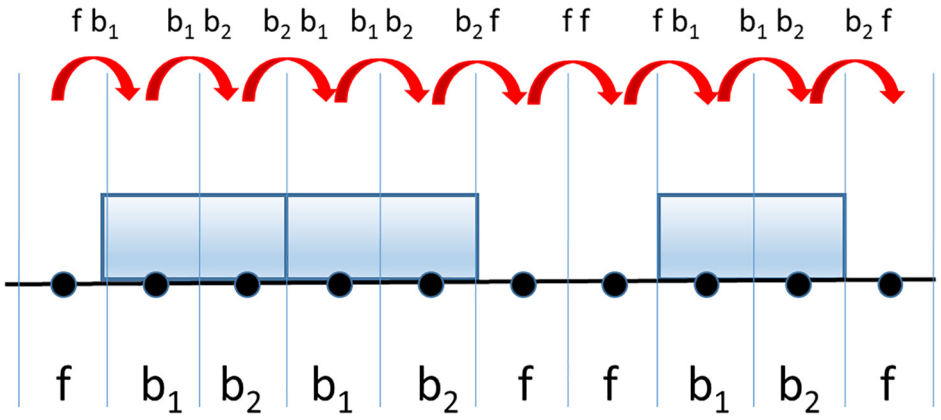

Fig. 6. Schematics of the conditional probabilities used to describe their distribution illustrated for the binding of $n = 3$ ligands with binding mode $m = 2$ bound to a chain of $N = 10$. Figure from Ref. [38].

$$(b_m f) = \frac{(ff)}{\omega - (ff)(\omega - 1)}, \tag{41}$$

(see Appendix), and we have

$$(ff) = \frac{(2\omega - 1)(1 - c) + c/m - R}{2(\omega - 1)(1 - c)}, \tag{42}$$

where R is given by

$$R = \sqrt{(1 - c - c/m)^2 + 4\omega(1 - c)c/m}. \tag{43}$$

[Please note that the denominator of Eq. (41) never becomes zero, $\omega - (ff)(\omega - 1) = \omega[1 - (ff)] + (ff) > 0$. (ff) is a probability between 0 and 1, and both terms never become zero simultaneously. Also note that to recover the non-cooperative case from Eq. (42) we have to follow the procedures described below Eq. (21).]

The general kinetic equation is given by Eq. (40) with the magnitudes given by Eqs. (41) to (43). The stationary state ($dc/dt = 0$) of the kinetic equation gives the equilibrium coverage relation

$$K = \frac{[(1 - \omega)(ff) + \omega](ff)^m}{1 - (ff)}, \tag{44}$$

where (ff) is given by Eq. (42). This implicit expression can be solved numerically to obtain the equilibrium coverage c_{eq}. It is very relevant to note that this result implies that the equilibrium coverage is independent of α, i.e., in the modification of the height of the activation barrier for singly or doubly contiguous binding processes.

The kinetic of two cases is represented in Fig. 1. First, cooperativity due to pure binding enhancement, which corresponds to $\alpha = 1$ and recovers the results presented in previous subsection, Subsection II.B. Second, cooperativity due to pure release inhibition ($\alpha = 0$). Pure binding enhancement has a faster binding kinetics (Fig. 1). Pure release inhibition starts with a coverage kinetics like the non-cooperative case but continues to grow until it reaches the same final equilibrium coverage, as the pure binding enhancement case with the same cooperative free energy. See also Fig. 5 for a schematic representation of the respective free energy landscapes.

2.3.2. Small ligand binding

2.3.2.1. Single binding ligand (m = 1). For $m = 1$, the kinetic equation [Eq. (40)] simplifies to

$$\frac{dc}{dt} = k_b c\left(\frac{(b_1 f)}{1-(ff)}\left[(ff) + \omega^{\alpha}(1-(ff))\right]^2\right) - k_r c\left[(b_1 f) + \omega^{\alpha-1}(1-(b_1 f))\right]^2, \tag{45}$$

with the following conditional probabilities

$$(b_m f) = \frac{2c - 1 + R}{2(\omega - 1)c}, \tag{46}$$

$$(ff) = \frac{2(\omega-1)(1-c) + 1 - R}{2(\omega-1)(1-c)}, \tag{47}$$

where the function R is given by

$$R = \sqrt{(1-2c)^2 + 4\omega c(1-c)}. \tag{48}$$

Analogous equations with a different notation were derived and compared with Monte Carlo simulation in Ref. [39].

The stationary state condition for the equilibrium coverage is given in terms of the conditional probability (ff), for $m = 1$, as

$$K = \frac{[(1-\omega)(ff) + \omega](ff)}{1-(ff)}. \tag{49}$$

This implicit expression can be solved analytically to obtain the equilibrium coverage c_{eq}, as expressed by Eq. (23).

2.3.3. Large ligand binding ($m \gg 1$)

Under the assumption of a ligand binding to many binding sites, $m \gg 1$, the first term of the right-hand side of the general kinetic equation, Eq. (40), can be written as (See Appendix for derivation)

$$k_b c\left(\frac{(b_m f)(ff)^{m-1}}{1-(ff)}\left[(ff) + \omega^{\alpha}(1-(ff))\right]^2\right) = = k_b m\,(1-c)\,\exp\left(\frac{-c}{1-c}\right)\left[1 + c\frac{4(1-c)\omega^{\alpha} - 2 + (2\omega+1)c}{2m(1-c)^2}\right], \tag{50}$$

and the second term of the right-hand side of the Eq. (40) as

$$k_r c\left[(b_m f) + \omega^{\alpha-1}(1-(b_m f))\right]^2 = k_r c\left[1 - \frac{2c\omega(1-\omega^{\alpha-1})}{m(1-c)}\right]. \tag{51}$$

2.3.3.1. Leading order in the large ligand expansion ($m \gg 1$). If only the leading-order ($m \gg 1$) contributions are retained from Eqs. (50) and (51), the general cooperative kinetic equation, Eq. (40), becomes identical to Eq. (10), which is the non-cooperative kinetic equation at leading order in the large m expansion. Therefore, when large size ligands are involved, the equation describing the time variation in the coverage is independent of both the parameter ω and the parameter α. Consequently, the equilibrium coverage can also be expressed by equations (11) or (12), of the non-cooperative case. This generalizes the result presented in Section II.B.2.1, showing that also in the general cooperative case the cooperative effects are absent in the leading order in the large ligand size expansion. Cooperative effects only appear in the effective binding and release rate at the next-to-leading order as we show just below.

2.3.3.2. Next-to-leading-order in the large ligand expansion ($m \gg 1$). Retaining the next-to-leading-order terms of Eqs. (50) and (51), the kinetic equation, Eq. (40), becomes

$$\frac{dc}{dt} = k_b m\,(1-c)\,\exp\left(\frac{-c}{1-c}\right)\left[1 + c\frac{4(1-c)\omega^{\alpha} - 2 + (2\omega+1)c}{2m(1-c)^2}\right] - k_r\, c\left[1 - \frac{2\,c\,\omega\,\left(1-\omega^{\alpha-1}\right)}{m\,(1-c)}\right]. \tag{52}$$

Thus, the equilibrium coverage, c_{eq}, in the limit of large m, is given by

$$\ln(Km) = \ln\left(\frac{c_{eq}}{1-c_{eq}}\right) + \frac{c_{eq}}{1-c_{eq}} + \ln\left(1 - \frac{(2\omega-1)(2-c_{eq})c_{eq}}{2m(1-c_{eq})^2}\right). \tag{53}$$

[See Eq. (A40) of the Appendix and the previous definitions and derivations.] The equilibrium coverage c_{eq} is then independent of the parameter α [consistently with the general result in Eq. (44)], and the equilibrium coverage computed previously for the case of $\alpha = 1$ is recovered, Eq. (28) in Section II.B.2.2. Instead, the kinetic equation, Eq. (52), explicitly depends on the parameter α. This implies that the binding kinetics is faster for enhanced binding, $\alpha = 1$ [recovering Eq. (27)], than for inhibited release, $\alpha = 0$, giving

$$\frac{dc}{dt} = k_b m(1-c)\exp\left(\frac{-c}{1-c}\right)\left[1 + \frac{c(2(1-c) + c(2\omega-1))}{2m(1-c)^2}\right] - k_r c\left[1 - \frac{2c(\omega-1)}{m(1-c)}\right]. \tag{54}$$

The kinetic of these two cases are compared in Fig. 1 using the full kinetic equation, Eq. (40). Fig. 1 illustrates that both reach the same equilibrium coverage in a different time interval. Faster/slower roles are inverted for detachment.

3. Discussion

Making an accurate account of the potential binding sites, we have derived the kinetic equation for the cooperative binding of ligands to long linear polymers. We derived a general kinetic equation considering that cooperativity can affect the binding and release rates. Thus, positive cooperativity emerges by an enhancement of the binding or by an inhibition of the release. Negative cooperativity arises as an inhibition of the binding or an enhancement of the release. We found that positive and negative cooperativity gives different kinetics, but the equilibrium state is independent of the origin of cooperativity and recovers the McGhee and von Hippel's equation for the coverage. In addition, the results show that the equilibrium coverage is increased significantly with increasing positive cooperativity (see Fig. 3). The kinetic equations presented are valid for the global coverage of long polymers. Distant regions of long polymers are nearly independent, making global coverage an effective average over realizations.

For large ligand size, the cooperativity effects are noticeably reduced, and the equilibrium coverages obtained in different scenarios are similar (see Fig. 3). In the large ligand size expansion, the cooperative effects are absent in the leading order for both the general cooperative case and McGhee and von Hippel's approach. Thus, kinetics and equilibrium coverages are independent of cooperativity and activation state parameters. We found that cooperative effects only appear in the effective binding and release rate at the next-to-leading order, which means that iso-

lated binding sites are the dominant contribution for large ligands. Within McGhee and von Hippel's approach, the next-to-leading order contribution shows that positive cooperativity increases coverage, while negative cooperativity decreases coverage. Within our approach, the next-to-leading order contribution shows that the equilibrium coverage is independent of the activation state parameter, but the binding kinetics is faster for enhanced binding, $\alpha = 1$ than for inhibited release, $\alpha = 0$.

The derived cooperative kinetic equation provides the adequate theoretical framework required for the description of the ligand binding processes observed with single-molecule techniques. In particular, recently these techniques have allowed to measure the real-time kinetics of SSB to ssDNA [5,26], including E. Coli SSB, which has a cooperative biding mode. However, already for the non-cooperative case it has been shown [38] that previous simplified kinetic models for these systems [4] are only accurate for low coverage. The key is the detailed counting of potential binding sites. This detailed counting of binding sites was done for the non-cooperative case in Ref. [38], following Ref. [30], allowing to obtain an accurate kinetic equation. Here, we extended this detailed account to obtain an accurate cooperative kinetic equation, allowing more accurate descriptions of cooperative ligand binding processes for medium and high coverage situations. For example, for the typical biding affinities found for SSB to ssDNA, the simplified models predict the full coverage of ssDNA [5,26], while the more accurate models generalized here predict coverage from 80% to 90% (depending on the ligand affinity), i.e., a 10% to 20% error correction. This predicted (and observed) decrease in coverage has also conceptual implications, as allow, in principle binding of other small ligands. Our results from the model presented here stresses that SSB positive cooperativity increases the ssDNA coverage at equilibrium, and during the whole dynamics (provided the cooperative mechanism is enhancement of the binding).

In the application of the present model, the ligand size m is given by the number of occluded monomers. This comprises the bound monomers and the additional occluded monomers due to steric hindrance (between ligands). One ligand can bound in different modes, i.e., occluding different number of monomers, depending on its concentration and salt concentration in the solution. For example, human mitochondrial and E. Coli SSB have been reported to bound, both in a low binding mode ($\sim$ 40 nucleotides) or in high binding mode ($\sim$ 70 nucleotides), depending on the SSB and salt concentrations on solution [5,17]. Understanding this mode selection and the involved mode competition is a relevant open question, which we plan to address in the future extending the kinetic model presented here.

Effects of conformational coiling and steric hindrance of the polymer are beyond the scope of the present study. These effects will reduce the effectively available number of gaps or decrease the binding affinity, and might be relevant for some applications.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgment

We acknowledge Jules Vidal for fruitful discussions on this topic during his stay in UCM. This work was supported by the Spanish Ministry of Economy and Competitiveness [RTI2018-095802-B-I00].

Appendix A

Derivation of the kinetic equation and the equilibrium coverage.

First, a counting convention is adopted. To count the number of free sites in a gap, we start at the right end of one bound ligand, and count by proceeding to the right, one site 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 as the product of the constituent conditional probabilities.

Next, the following notation is adopted. Any bound ligand can be divided into m sections, each one corresponding to the underlying polymer site; we number these sections, 1, 2…, m from left to right. Thus, we have $(m + 1)$ distinguishable types of polymer sites: a free site, labeled f; a site 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, (fb_1) is the probability, given a free site (i.e., a f site), that the left end of a bound ligand (i.e., a b_1 site) 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 none of all of them make sense. For example, there are only two types of polymer sites which can possibly lie to the immediate right of a free site: either another free site or the first part of a bound ligand, i.e., $(ff) + (fb_1) = 1$, implying

$$(fb_1) = 1 - (ff). \tag{A1}$$

Similarly, only a free site 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, $(b_m f) + (b_m b_1) = 1$, which gives

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

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

$$P_g = (b_m f)(ff)^{g-1}(fb_1) = (b_m f)(ff)^{g-1}(1 - (ff)), \tag{A3}$$

where $(b_m f)$ is the probability, having selected the right end of a bound ligand, that the polymer site to the immediate right is free; (ff) is the probability, given a free site, that there is a second free site to the immediate right; and (fb_1) is the probability, given a free site, that the left end of a bound ligand lies to the immediate right. Next, we must obtain the corresponding expressions for the conditional probabilities in terms of the model parameters.

For non-cooperative ligands, if one considers the polymer site immediately to the right of one selected at random, this second site 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 polymer, there are only two ways in which this two-step random selection can be made. Either the first site chosen is free (a f site chosen with probability 1 - c) and has a free site to its right (conditional probability (ff)); or the first site chosen is the right end of a bound ligand (a b_m site chosen with probability c/m) and has a free site to its right (conditional probability $(b_m f)$). Since the overall probability that the second site is free must be independent of the method of random selection, we obtain:

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

From this expression we get

$$(b_m f) = (1 - (ff))\frac{1 - c}{\frac{c}{m}}. \tag{A5}$$

A.1. Non-cooperative case

When ligands bound non-cooperatively neither attract nor repel one another, it can be written that

$$(b_m f) = (ff). \tag{A6}$$

We now combine equations (A5) and (A6) to obtain the following expressions for the conditional probabilities

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

And, using equations (A1) and (A2), we have that

$$(fb_1) = (b_m b_1) = \frac{c/m}{1-c+c/m}. \tag{A8}$$

Then, the resulting expression for P_g is given by

$$P_g = \left(\frac{1-c}{1-c+c/m}\right)^g \left(\frac{c/m}{1-c+c/m}\right). \tag{A9}$$

We next obtain an expression for the average number of free binding sites per gap in terms of P_g, thus $\bar{s} = \sum_{g=m}^{N}(g-m+1)P_g$, by letting N go to infinity,

$$\bar{s} = \sum_{g=m}^{\infty}(g-m+1)(ff)^g(fb_1) = (fb_1)(ff)^{m-1}\frac{(ff)}{[1-(ff)]^2} = \frac{(ff)}{(fb_1)}(ff)^{m-1}. \tag{A10}$$

Substituting Eqs. (A7) and (A8) into Eq. (A10) we obtain Eq. (4).

Generalized cooperative case

The cooperativity parameter is defined as the ratio of the probabilities of the two configurations represented in Fig. 7 [44]. Ligand locations at the left (A) and right (B) are the same for both configurations. The difference between the two configurations is that in the second configuration the two represented ligands are at neighboring locations and have an additional ligand–ligand interaction. This additional ligand–ligand interaction makes Configuration 2 more probable when it is attractive, positive cooperativity.

The probability of Configuration 1 is

$$P_1 = P_A(b_m b_1)(ff)^{x-1}(fb_1)(b_1 b_2)\cdots(b_m f)(ff)^{y-1}P_B, \tag{A11}$$

while the probability of Configuration 2 is

$$P_2 = P_A(b_m b_1)(b_1 b_2)\cdots(b_m f)(ff)^{x+y-1}P_B. \tag{A12}$$

Thus, the cooperativity parameter, defined as the ratio, becomes

$$\omega = \frac{P_2}{P_1} = \frac{(b_m b_1)(ff)}{(b_m f)(fb_1)} = \frac{[1-(b_m f)](ff)}{(b_m f)[1-(ff)]}. \tag{A13}$$

where Eqs. (A1) and (A2) were considered. This implies that in the cooperative case,

$$(b_m f) = \frac{(ff)}{\omega-(ff)(\omega-1)}. \tag{A14}$$

For $\omega > 1$, the ligands attract each other, and the binding is positively cooperative; for $\omega = 1$, the binding is non-cooperative; for $\omega < 1$, the ligands repel each other, and the binding is negatively cooperative.

Combining the last expression of Eq. (A13) with Eq. (A5), and solving the resulting quadratic equation in (ff),

$$m(1-c)[\omega-(ff)(\omega-1)][1-(ff)] - c(ff) = 0, \tag{A15}$$

we get the conditional probability

$$(ff) = \frac{(2\omega-1)(1-c)+c/m-R}{2(\omega-1)(1-c)}, \tag{A16}$$

where R is given by

$$R = \sqrt{(1-c-c/m)^2 + 4\omega(1-c)c/m}. \tag{A17}$$

Using Eqs. (A1), (A2) and (A5), we get the other conditional probabilities:

$$(b_m f) = \frac{c+c/m-1+R}{2(\omega-1)c/m}, \tag{A18}$$

$$(fb_1) = \frac{c-c/m-1+R}{2(\omega-1)(1-c)}, \tag{A19}$$

$$(b_m b_1) = \frac{1-c+(2\omega-1)c/m-R}{2(\omega-1)c/m}. \tag{A20}$$

We next obtain an expression for the average number of each of the three types of free binding sites per gap in terms of P_g, thus $\bar{s} = \sum_{g=m}^{N}(g-m+1)P_g$, by letting N go to infinity. For $g < m$, there are no free binding sites. For $g = m$, there is one doubly binding site, thus

$$\bar{s}_{dc} = 1 \cdot P_m = (b_m f)(ff)^{m-1}(fb_1) = (b_m f)(ff)^{m-1}(1-(ff)), \tag{A21}$$

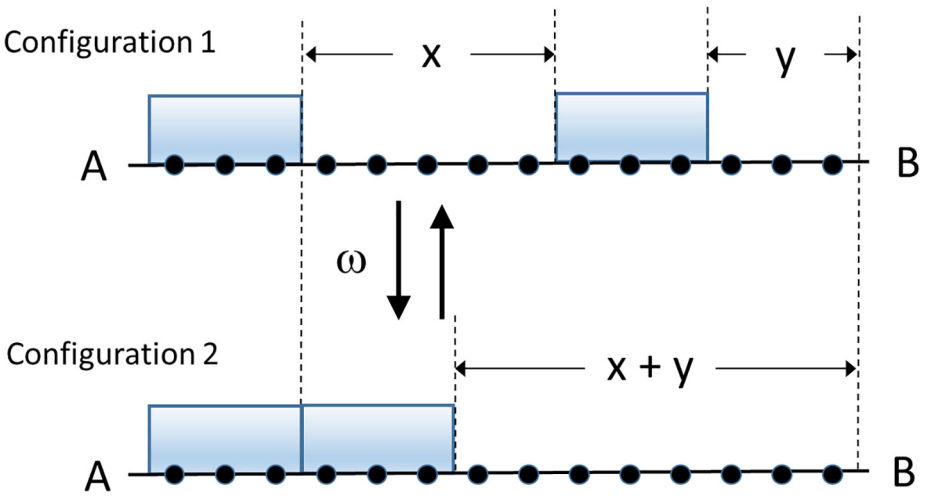

Fig. 7. Definition of the cooperativity parameter, ω, (ligand–ligand interaction parameter) is done in terms of the ratio of the probabilities of the two represented configurations [See Eqs. (A11)–(A13)].

where $\bar{s}_{dc}$ is the average number of free doubly contiguous binding sites per gap. For $g \geq m + 1$, there are two singly contiguous binding site, thus [using Eq. (A3) for the sum

$$\begin{aligned}\bar{s}_{sc} = 2 \cdot \sum_{g=m+1}^{\infty} P_g = & \quad 2(b_m f)(ff)^m (fb_1) \sum_{m=0}^{\infty} (ff)^m = \\ = & \quad 2(b_m f)(fb_1) \tfrac{(ff)^m}{1-(ff)} = 2(b_m f)(ff)^m,\end{aligned} \tag{A22}$$

where $\bar{s}_{sc}$ is the average number of free singly contiguous binding sites per gap. For $g \geq m + 2$, there are $(g - m - 1)$ isolated binding sites per gap, thus [using Eq. (A3) for the sum

$$\begin{aligned}\bar{s}_i = \sum_{g=m+2}^{\infty} (g-m-1)P_g = (b_m f)(ff)^m (fb_1) \sum_{m=1}^{\infty} m(ff)^m = \\ = (b_m f)(fb_1) \tfrac{(ff)^{m+1}}{[1-(ff)]^2} = \tfrac{(b_m f)(ff)^{m+1}}{(fb_1)} = \tfrac{(b_m f)(ff)^{m+1}}{1-(ff)},\end{aligned} \tag{A23}$$

where $\bar{s}_i$ is the average number of free isolated binding sites per gap. Substituting Eqs. (A21), (A22) and (A23) into Eq. (19), using the appropriate expressions for the conditional probabilities, we obtain Eq. (20).

Now, we obtain an expression for the average number of each of the three types of attached ligands in the polymer. We define $\bar{l}_i$ as the average fraction of isolated attached ligands in the polymer, whose expression is given by

$$\bar{l}_i = (fb_1) \frac{1-c}{\frac{c}{m}} (b_m f) = (b_m f)^2. \tag{A24}$$

Next, we define $\bar{l}_{sc}$ as the average fraction of attached ligands with a single contiguous ligand in the polymer,

$$\begin{aligned}\bar{l}_{sc} = (b_m b_1)\ (b_m f) + \quad (fb_1) \tfrac{1-c}{\frac{c}{m}}\ (b_m b_1) = \\ = \quad (b_m b_1)\ (b_m f) + (b_m f)(b_m b_1) = 2(b_m b_1)\ (b_m f).\end{aligned} \tag{A25}$$

Finally, we define $\bar{l}_{dc}$ as the average fraction of attached ligands with two contiguous ligands,

$$\bar{l}_{dc} = (b_m b_1)(b_m b_1) = (b_m b_1)^2. \tag{A26}$$

It is noted that

$$\begin{aligned}\bar{l}_i + \bar{l}_{sc} + \bar{l}_{dc} = (b_m f)^2 + 2(b_m b_1)(b_m f) + (b_m b_1)^2 \\ = [(b_m b_1) + (b_m f)]^2 = 1,\end{aligned} \tag{A27}$$

as expected for consistency.

Hence, the time variation in the number of ligands bound to the polymer, which given by Eq. (39), is written as

$$\frac{dn}{dt} = k_b n S - k_r n L, \tag{A28}$$

where

$$S = \bar{s}_i + \omega^{\alpha} \bar{s}_{sc} + \omega^{2\alpha} \bar{s}_{dc}, \tag{A29}$$

and

$$L = \bar{l}_i + \omega^{\alpha-1} \bar{l}_{sc} + \omega^{2(\alpha-1)} \bar{l}_{dc}. \tag{A30}$$

where α is the so-called activation state parameter. The Eq. (A29) can be expressed in terms of the conditional probabilities, using Eqs. (A21), (A22), and (A23), as

$$\begin{aligned}S = \bar{s}_i + \omega^{\alpha}\ \bar{s}_{sc} + \omega^{2\alpha}\ \bar{s}_{dc} = \frac{(b_m f)(ff)^{m+1}}{(fb_1)} \left[1 + \omega^{\alpha} \frac{(fb_1)}{(ff)}\right]^2 = \\ = \frac{(b_m f)(ff)^{m+1}}{1-(ff)} \left[1 + \omega^{\alpha} \frac{1-(ff)}{(ff)}\right]^2 = \\ = \frac{(b_m f)(ff)^{m-1}}{1-(ff)} [(ff) + \omega^{\alpha}(1-(ff))]^2 \\ = \frac{1-c}{\frac{c}{m}} (ff)^{m-1} [(ff) + \omega^{\alpha}(1-(ff))]^2,\end{aligned} \tag{A31}$$

i.e., it can be expressed in terms of two of the conditional probabilities (ff) and $(b_m f)$, or in terms of the coverage c and the conditional probability (ff).

Moreover, Eq. (A30) can be expressed in terms of the conditional probabilities, using Eqs. (A24), (A25), and (A26), as

$$\begin{aligned}L = \quad & \bar{l}_i + \omega^{\alpha-1}\ \bar{l}_{sc} + \omega^{2(\alpha-1)}\ \bar{l}_{dc} = \\ = \quad & \left[(b_m f) + \omega^{\alpha-1}(b_m b_1)\right]^2 = \left[(b_m f) + \omega^{\alpha-1}(1-(b_m f))\right]^2,\end{aligned} \tag{A32}$$

i.e., it can be stated in terms of the conditional probability $(b_m f)$. This expression clearly shows that for $\omega = 1$ or for $\alpha = 1$, this term is simply equal to 1.

Finally, Eq. (39) can be written in terms of the coverage and the conditional probabilities as

$$\begin{aligned}\tfrac{dc}{dt} = \quad & k_b\ c \left(\tfrac{(b_m f)(ff)^{m-1}}{1-(ff)} [(ff) + \omega^{\alpha}(1-(ff))]^2\right) \\ - \quad & k_r\ c \left[(b_m f) + \omega^{\alpha-1}(1-(b_m f))\right]^2.\end{aligned} \tag{A33}$$

The equilibrium equation in the general case is given by $K = L/S$ (remember that $K = k_b/k_r$). Using the previous expressions for L, S, and Eq. (A14) for $(b_m f)$, the equilibrium equation can be restated in terms of the conditional probability (ff) as

$$K = \frac{[(1-\omega)(ff) + \omega](ff)^m}{1-(ff)}\ . \tag{A34}$$

The conditional probability (ff) for the general case is given by Eq. (A16), but can alternatively expressed as

$$\begin{aligned}(ff) = 1 + \frac{1 - \sqrt{1 + \frac{4(\tilde{r}-1)\delta}{\tilde{r}^{\ 2}}}}{\frac{2\delta}{\tilde{r}}} \\ = 1 + \frac{\left(1 - \sqrt{1 + \frac{4r\delta}{m(1+\frac{r}{m})^2}}\right)(1+\frac{r}{m})}{2\delta}\end{aligned} \tag{A35}$$

where

$$\delta = \omega - 1, \tag{A36}$$

$$\tilde{r} = 1 + \frac{r}{m}, \tag{A37}$$

and

$$r = \frac{c}{1-c}. \tag{A38}$$

A.2.1. Expansion for large m

Under the assumption of $m \gg 1$ we obtain that the conditional probability (ff) is given by the equation

$$(ff) = 1 - \frac{r}{m} + \frac{r^2(1+\delta)}{m^2} - \frac{r^3\left(2\delta^2+3\delta+1\right)}{m^3} + \frac{r^4\left(5\delta^3+10\delta^2+6\delta+1\right)}{m^4} = 1 - \frac{c}{(1-c)m} + \frac{c^2(1+\delta)}{(1-c)^2m^2} - \frac{c^3\left(2\delta^2+3\delta+1\right)}{(1-c)^3m^3} + \frac{c^4\left(5\delta^3+10\delta^2+6\delta+1\right)}{(1-c)^4m^4}, \tag{A39}$$

or equivalently (in terms of ω)

$$(ff) = 1 - \frac{r}{m} + \frac{r^2\omega}{m^2} - \frac{r^3\omega(2\omega-1)}{m^3} + \frac{r^4\omega\left(5\omega^2-5\omega+1\right)}{m^4} = 1 - \frac{c}{(1-c)m} + \frac{c^2\omega}{(1-c)^2m^2} - \frac{c^3\omega(2\omega-1)}{(1-c)^3m^3} + \frac{c^4\omega\left(5\omega^2-5\omega+1\right)}{(1-c)^4m^4}. \tag{A40}$$

Also, in the limit of large m, we obtain that

$$S \simeq \frac{me^{-r}}{r}\left[1 + \frac{\left(2r\omega^{\alpha} + \left(\omega-\frac{1}{2}\right)r^2 - r\right)}{m}\right] = \frac{m(1-c)}{c}\exp\left(\frac{-c}{1-c}\right)\left[1 + c\frac{4(1-c)\omega^{\alpha} - 2 + (2\omega+1)c}{2m(1-c)^2}\right], \tag{A41}$$

and

$$L \simeq 1 - \frac{2r\omega\left(1-\omega^{\alpha-1}\right)}{m} = 1 - \frac{2c\omega\left(1-\omega^{\alpha-1}\right)}{m(1-c)}. \tag{A42}$$

The equilibrium equation, which is given by $K = \frac{L}{S}$, in the limit of large m, becomes

$$K = \frac{re^{r}}{m}\left[1 - \frac{r(r+2)\left(\omega-\frac{1}{2}\right)}{m}\right]. \tag{A43}$$

Using Eq. (A38), this leads to a correction to the coverage $c_{eq} = c_{eq}^{(0)} + \frac{c_{eq}^{(1)}}{m}$ with

$$c_{eq}^{(1)} = \left(c_{eq}^{(0)}\right)^2\left(2 - c_{eq}^{(0)}\right)\left(\omega - \frac{1}{2}\right), \tag{A44}$$

where $c_{eq}^{(0)}$ is the solution for the kinetic equation in a non-cooperative system at the leading order in large m, Eq. (12).

A.2.2. Expansion around $= 0$

We obtain an expansion around $\delta = 0$ of the conditional probability (ff) as

$$(ff) = 1 - \frac{\tilde{r}-1}{\tilde{r}} + \frac{\left(\tilde{r}-1\right)^2\delta}{\tilde{r}^3} - \frac{2\left(\tilde{r}-1\right)^3\delta^2}{\tilde{r}^5} = 1 - \frac{r}{m\left(1+\frac{r}{m}\right)} + \frac{r^2\delta}{m^2\left(1+\frac{r}{m}\right)^3} - \frac{2r^3\delta^2}{m^3\left(1+\frac{r}{m}\right)^5} = \frac{m(-1+c)}{(m-1)c-m} + \frac{m(-1+c)c^2\delta}{((m-1)c-m)^3} + \frac{2m^2(-1+c)^2c^3\delta^2}{((m-1)c-m)^5} = \frac{m(-1+c)}{(m-1)c-m}\left[1 + \frac{c^2\delta}{((m-1)c-m)^2} + \frac{2m(-1+c)c^3\delta^2}{((m-1)c-m)^4}\right]. \tag{A45}$$

In the limit of small $\delta = \omega - 1$, we find that

$$S \simeq \frac{\tilde{r}^{1-m}}{\tilde{r}-1}\left[1 + \frac{\delta\left(\tilde{r}-1\right)\left((m+2\alpha-1)\tilde{r} - m + 1\right)}{\tilde{r}^2}\right] = \frac{m\left(1-\frac{r}{m}\right)^{1-m}}{r}\left[1 + \frac{\delta\, r\left(m+2\alpha-1\right)\left(1+\frac{r}{m}\right) - m + 1\right)}{m\left(1+\frac{r}{m}\right)^2}\right] = \frac{c\left(\frac{1-c}{1-c+\frac{c}{m}}\right)^{m-1}}{m(1-c)}\left[1 - 2\frac{(\omega-1)\left((m-1)\left(\alpha-\frac{1}{2}\right)c - \alpha m\right)c}{m^2\left(1-c+\frac{c}{m}\right)^2}\right], \tag{A46}$$

and

$$L \simeq 1 - 2(1-\alpha)\left(1 - \frac{1}{\tilde{r}}\right)\delta = 1 - \frac{2(1-\alpha)r\delta}{m+r} = 1 - \frac{2mc(1-\alpha)(\omega-1)}{1-c+\frac{c}{m}}. \tag{A47}$$

The equilibrium equation, in the limit of small $\delta = \omega - 1$, becomes

$$K = \left(\tilde{r}-1\right)\tilde{r}^{m-1}\left[1 - \frac{\left(\left(\tilde{r}-1\right)m + \tilde{r} + 1\right)\left(\tilde{r}-1\right)\delta}{\tilde{r}^2}\right]. \tag{A48}$$

This equation shows that a higher α or δ makes the brackets smaller, which is equivalent as increasing K. Thus, we expect higher covertures for $\delta > 0$, i.e., $\omega > 1$, and for $\alpha = 1$.

In addition, the equilibrium equation (in the limit of small $\delta = \omega - 1$) can also be written in terms of r as

$$K = \frac{r}{m}\left(1 + \frac{r}{m}\right)^{m-1}\left[1 - \frac{r((r+2)m+r)\delta}{(m+r)^2}\right] \tag{A49}$$

Appendix B. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.csbj.2021.12.043.

References

[1] Kozlov AG 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 2015;427:763–74.
[2] Antony E, Lohman TM. Dynamics of E. coli single stranded DNA binding (SSB) protein-DNA complexes. Semin Cell Dev Biol 2019;86:102–11. https://doi.org/10.1016/j.semcdb.2018.03.017.
[3] Maffeo C, Aksimentiev A. Molecular mechanism of DNA association with single-stranded DNA binding protein. Nucleic Acids Res 2017;45:12125–39.
[4] Jarillo J, Morín JA, Beltrán-Heredia E, Villaluenga JPG, Ibarra B, Cao FJ, et al. Mechanics, thermodynamics, and kinetics of ligand binding to biopolymers. PLoS ONE 2017;12(4):e0174830.
[5] Morin JA et al. DNA synthesis determines the binding mode of the human mitochondrial single-stranded DNA-binding protein. Nucleic Acids Res 2017;45:7237–48.
[6] Settanni G, Serquera D, Marszalek PE, Paci E, Itzhaki LS. Effects of ligand binding on the mechanical properties of ankyrin repeat protein gankyrin. PLoS Comput Biol 2013;9:e1002864.
[7] Alberts, B. *et al. Molecular Biology of the Cell*. (Garland Science, 2017). doi:10.1201/9781315735368.
[8] Cesconetto EC et al. DNA interaction with Actinomycin D: mechanical measurements reveal the details of the binding data. Phys Chem Chem Phys 2013;15:11070–7.
[9] Siman L et al. Quantitative assessment of the interplay between DNA elasticity and cooperative binding of ligands. Phys Rev Lett 2012;109:248103.
[10] Cerrón F et al. Replicative DNA polymerases promote active displacement of SSB proteins during lagging strand synthesis. Nucleic Acids Res 2019;47:5723–34.
[11] Ciesielski GL et al. Mitochondrial single-stranded DNA-binding proteins stimulate the activity of DNA polymerase γ by organization of the template DNA. J Biol Chem 2015;290:28697–707.
[12] Suksombat S, Khafizov R, Kozlov AG, Lohman TM, Chemla YR. Structural dynamics of E. coli single-stranded DNA binding protein reveal DNA wrapping and unwrapping pathways. Elife 2015;4. https://doi.org/10.7554/eLife.08193.
[13] Jarillo J, Ibarra B, Cao-García FJ. DNA replication: In vitro single-molecule manipulation data analysis and models. Comput Struct Biotechnol J 2021;19:3765–78.
[14] Kristiansen A, Vårum KM, Grasdalen H. 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 1998;1425:137–50.
[15] Camunas-Soler J et al. Single-molecule kinetics and footprinting of DNA bis-intercalation: the paradigmatic case of Thiocoraline. Nucleic Acids Res 2015;43:2767–79.
[16] Rajendran S, Jezewska MJ, Bujalowski W. Human DNA polymerase β recognizes single-stranded DNA using two different binding modes. J Biol Chem 1998;273:31021–31.
[17] Kozlov AG, Shinn MK, Lohman TM. Regulation of Nearest-neighbor cooperative binding of E. coli SSB protein to DNA. Biophys J 2019;1–21. https://doi.org/10.1016/j.bpj.2019.09.047.
[18] Erba E et al. Mode of action at thiocoraline, a natural marine compound with anti-tumour activity. Br J Cancer 1999;80:971–80.
[19] Nishio T, Shimizu T. Model analysis of surfactant-polymer interaction as cooperative ligand binding to linear lattice. Biophys Chem 2005;117:19–25.
[20] Epstein RI. polynucleotide and the ligand. Biophys Chem 1978;8:327–39.

[21] Nechipurenko YD, Gursky GV. Cooperative effects on binding of proteins to DNA. Biophys Chem 1986;24:195–209.
[22] Ghosh S, Mookerjee A. A model for cooperative ligand binding at complementary sites of DNA. Bull Math Biol 1986;48:21–7.
[23] Wolfe AR, Meehan T. 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 1992;223:1063–87.
[24] Teif VB, Rippe K. Statistical-mechanical lattice models for protein-DNA binding in chromatin. J Phys Condens Matter 2010;22(41):414105. https://doi.org/10.1088/0953-8984/22/41/414105.
[25] von Hippel PH, Marcus AH. The many roles of binding cooperativity in the control of DNA replication. Biophys J 2019;117(11):2043–6. https://doi.org/10.1016/j.bpj.2019.10.029.
[26] Naufer, M. N. *et al.* Multiprotein E . coli SSB – ssDNA complex shows both stable binding and rapid dissociation due to interprotein interactions. **49**, 1532–1549 (2021).
[27] Zhang W, Bond JP, Anderson CF, Lohman TM, Record MT. Large electrostatic differences in the binding thermodynamics of a cationic peptide to oligomeric and polymeric DNA. Proc Natl Acad Sci U S A 1996;93:2511–6.
[28] Mascotti DP, Lohman TM. Thermodynamic extent of counterion release upon binding oligolysines to single-stranded nucleic acids. Proc Natl Acad Sci U S A 1990;87:3142–6.
[29] Stigter D, Dill KA. Binding of ionic ligands to polyelectrolytes. Biophys J 1996;71:2064–74.
[30] McGhee JD, von Hippel PH. 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 1974;86:469–89.
[31] Maltsev E, Wattis JAD, Byrne HM. DNA charge neutralization by linear polymers II. Reversible binding. Phys Rev E – Stat Nonlinear, Soft Matter Phys 2006;74:1–15.
[32] Tsodikov OV, Holbrook JA, Shkel IA, Record MT. Analytic binding isotherms describing competitive interactions of a protein ligand with specific and nonspecific sites on the same DNA oligomer. Biophys J 2001;81:1960–9.
[33] Laurila, K., Yli-Harja, O. & Lähdesmäki, H. A protein-protein interaction guided method for competitive transcription factor binding improves target predictions. *Nucleic Acids Res.* **37**, (2009).
[34] Wasson T, Hartemink AJ. An ensemble model of competitive multi-factor binding of the genome. Genome Res 2009;19:2101–12.
[35] Granek JA, Clarke ND. Explicit equilibrium modeling of transcription-factor binding and gene regulation. Genome Biol 2005;6.
[36] Di Cera E, Kong Y. Theory of multivalent binding in one and two-dimensional lattices. Biophys Chem 1996;61:107–24.
[37] Teif, V. B. General transfer matrix formalism to calculate DNA-protein-drug binding in gene regulation: application to OR operator of phage λ. *Nucleic Acids Res.* **35**, (2007).
[38] Villaluenga JPG, Vidal J, Cao-García FJ. Noncooperative thermodynamics and kinetic models of ligand binding to polymers: Connecting McGhee-von Hippel model with the Tonks gas model. Phys Rev E 2020;102:12407.
[39] Epstein IR. Kinetics of nucleic acid–large ligand interactions: Exact Monte Carlo treatment and limitingl cases of reversible binding. Biopolymers 1979;18(8):2037–50.
[40] Tonks L, Taylor JB, Langmuir I, Rev P. The complete equation of state of one, two and three-dimensional gases of hard elastic spheres. J Am Chem Soc 1932;54.
[41] Tonks L. The complete equation of state of one, two and Three-dimensional gases of hard elastic spheres. Phys Rev 1936;50(10):955–63.
[42] Langmuir I. The constitution and fundamental properties of solids and liquids. Part II.-Liquids. J Franklin Inst 1917;184:721.
[43] Clore GM, Gronenborn AM, Davies RW. 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 1982;155:447–66.
[44] Schwarz G. Cooperative Binding to Linear Biopolymers. 1. Fundamental Static and Dynamic Properties. Eur J Biochem 1970;12:442–53. https://doi.org/10.1111/j.1432-1033.1970.tb00871.x.