Multiple-binding-site mechanism explains concentration-dependent unbinding rates of DNA-binding proteins

Charles E. Sing[1], Monica Olvera de la Cruz[1],* and John F. Marko[2],*

[1]Department of Materials Science, Northwestern University, 2220 Cook Dr. Evanston, IL 60208, USA,
[2]Department of Molecular Biosciences, Northwestern University, Evanston, IL 60208, USA and [3]Department of Physics and Astronomy, Northwestern University, Evanston, IL 60208, USA

ABSTRACT

Recent work has demonstrated concentration-dependent unbinding rates of proteins from DNA, using fluorescence visualization of the bacterial nucleoid protein Fis [Graham *et al*. (2011) (Concentration-dependent exchange accelerates turnover of proteins bound to double-stranded DNA. Nucleic Acids Res., 39:2249)]. The physical origin of this concentration-dependence is unexplained. We use a combination of coarse-grained simulation and theory to demonstrate that this behavior can be explained by taking into account the dimeric nature of the protein, which permits partial dissociation and exchange with other proteins in solution. Concentration-dependent unbinding is generated by this simple model, quantitatively explaining experimental data. This effect is likely to play a major role in determining binding lifetimes of proteins *in vivo* where there are very high concentrations of solvated molecules.

INTRODUCTION

The kinetics of DNA–protein interactions control all aspects of gene expression and therefore cell function, and are of strong biological and biophysical interest. Sequence-targeting proteins such as transcription factors are known to have rapid site-targeting abilities, and the underlying dynamics of target search have been intensely investigated (1–8). These mechanisms generally rely on the existence of nonspecific binding, a property of a wide variety of DNA-binding proteins (1,9,10).

However, much less attention has been paid to unbinding of proteins from DNA. Removal and renewal of proteins bound to DNA is equal in importance to binding in determining transcription states and therefore cell behavior. Recent experimental studies of the proteins Fis (9,11–13), NHP6A (10) and others (14,15) have used novel single-molecule methods to probe unbinding dynamics. These *in vitro* experiments directly visualize the presence of binding proteins using fluorescence microscopy or DNA force spectroscopy measurements (16–21), and have demonstrated an unexpected concentration-dependent off-rate k_{off}, of the form (16):

$$k_{\text{off}} = k_0 + k_1\ c \tag{1}$$

where k_0 is the bare zero-concentration reaction constant and k_1 is a coefficient that incorporates a linear dependence of off-rate to concentration c (16). This has been shown to occur for Fis, the main focus of this manuscript, as well as for the DNA-binding proteins HU and NHP6A (16). This suggests that concentration-dependent dissociation may transcend protein-specific secondary structure and be characteristic of more generic tertiary- or quaternary-level features. Similar unbinding behavior has been invoked to understand kinetic data for other DNA-binding proteins (22,23), often in conjunction with other concentration-dependent effects for purposes of characterizing three-body and higher-order interactions (i.e. protein-mediated DNA–DNA interactions) (22,24,25). Analogies can also be drawn to classical reaction-rate descriptions of chemical processes, where interplays exist between kinetics and molecular structure (26).

Interpretation of these experimental results is complicated since similar concentration effects are present even in simple noncovalent systems, necessitating a 'quantitative' physical picture of concentration-dependence in binding interactions. Investigations on concentration-dependent processes have explored the role of geometric constraints present due to conformational characteristics, in conjunction with work on rebinding effects to understand the sequence-specific 'searching' processes that take into account one-dimensional diffusion, DNA hopping and interplay between specific and nonspecific binding

*To whom correspondence should be addressed. Tel: +1 847 467 1276; Email: john-marko@northwestern.edu
Correspondence may also be addressed to Monica Olvera de la Cruz. Email: m-olvera@northwestern.edu

(1–3). Nevertheless, concentration effects are pronounced even for linear chains where such processes are negligible and an alternative 'reaction-limited' scenario is necessary (16).

We note a key observation regarding the molecular structure of the systems of interest; proteins such as Fis (9,12) and HU and many other DNA-binding proteins all exhibit a characteristic dimeric quarternary structure. Fis, the protein of main interest in this article, binds two adjacent major grooves of the DNA molecule via two helix-turn-helix domains. We will show that dimeric binding sites, and in general any binding interaction with multiple chemical bonds, can naturally lead to bulk-concentration-dependent off-rates.

Our approach is to model the Fis-DNA system with minimal molecular detail such that the observed experimental behavior is obtained, using a coarse-grained simulation and theoretical model that retains (i) the dimeric nature of the binding protein, (ii) the identical nature of each monomeric structure such that binding of each is likewise identical and (iii) the monomers are each fundamental binding structures that themselves do not have concentration-dependent off-rates. This is sufficient to generate an unbinding pathway with competitive binding, and can be described using straightforward numerical and analytical theory to yield quantitatively accurate predictions for behavior observed in the experimental literature. Our results show that structural characteristics and properties of the environment may be necessary to describe biomolecule dynamics even in straightforward *in vitro* experiments, and suggest that the principles we explore are likely important to understanding a wide range of receptor–ligand and protein–protein interactions.

MATERIALS AND METHODS

To investigate the binding kinetics and equilibrium behaviors of dimeric protein binders interacting with a DNA substrate, we use a coarse-grained model that constitutes a minimal representation of the geometric situation and the hypothesized binding energy landscape. These simulations are designed to reproduce the experimental procedures in Graham *et al.* (16), which considers fluorescently tagged Fis molecules, which are initially saturated along a DNA chain that is stretched almost completely using a magnetic tweezers force spectroscopy setup. A linear DNA chain is thus represented in a Brownian Dynamics simulation as $n = 50$ beads that are spatially fixed and can bind exclusively and reversibly with dimeric binders that are free to undergo Brownian motion (simulation snapshots shown in Figure 1a). Binders are represented as dimers of beads, which can each bind independently to the DNA with a Monte Carlo update procedure that reproduces the statistics expected from a traditional Bell Model-type reaction (schematically demonstrated in Figure 1b) (27,28). The binding landscape in Figure 1b is determined by the binding barrier $\Delta \tilde{E}_B$ and unbinding barrier $\Delta \tilde{E}_{UB}$, such that there is an energy change on binding of $\Delta \tilde{E}_0 = \Delta \tilde{E}_B - \Delta \tilde{E}_{UB}$ (tildes represent values normalized by $k_B T$). $\Delta \tilde{E}_B = 3.0$ is kept constant, and leads to a binding time τ_D that is equal to the diffusive

time of the bead $\tau_D = \tau_0 e^{\Delta \tilde{E}_B} = 6\pi \eta a^3/(k_B T)$, where η is the solvent viscosity and $1/\tau_0$ is the Monte Carlo update frequency in the simulation. The characteristic unbinding time for a binding interaction is, thus, $\tau_{UB} = \tau_0 e^{\Delta \tilde{E}_{UB}} = \tau_D e^{-\Delta \tilde{E}_0}$. The dimension of the DNA and binder beads have a radius a, and the overall box has a size $200a \times 100a \times 100a$. This setup is shown in Figure 1a, which also demonstrates the time evolution of the system. This time evolution proceeds in a fashion analogous to the experimental case; initially, high concentrations of binders that strongly bind to the chain are used to saturate the DNA chain with bound dimers. Bound dimers are 'tagged' at $t = 0$ and the external concentration of binders is decreased to the measurement value of c. These untagged dimers slowly exchange with the tagged dimers, and once a tagged dimer passes through the periodic boundary conditions it becomes untagged. At long times, the number of tagged dimers decreases as shown in Figure 1a. We provide a detailed methods section in the Supplementary Data (SD) for this manuscript.

RESULTS AND DISCUSSION

Equilibrium behavior

While there are nontrivial dynamical behaviors for the system, the equilibrium behavior of these models is analytically solvable by straightforward-statistical mechanical methods. These computations also provide insight into the origins of the binding dynamics for these systems, which we will demonstrate in the context of correlations.

We plot the equilibrium properties of the binding equilibrium, as the number of occupied binding sites n_B as a function of the binding energy $\Delta \tilde{E}_0$ for a number of concentrations c. This is plotted in Figure 2a and b, which demonstrates n_B versus $\Delta \tilde{E}_0$ for both monomers (a) and dimers (b). This permits comparison between both simulation and theory, but also illustrates the effect of dimerization constraints on the equilibrium structure. In the monomer case, the well-known adsorption $\langle n_B \rangle = N \left(1 + e^{-\Delta \tilde{E}_0 + \tilde{\mu}}\right)^{-1}$ result is plotted as lines in Figure 2a, and fits the data almost exactly. $\tilde{\mu} = \tilde{\mu}_0 - \ln(c)$ is the ideal gas chemical potential with a constant reference chemical potential $\tilde{\mu}_0$. Clearly, the system in Figure 2a reproduces the correct statistical mechanics for the simplified case on varying both c and $\Delta \tilde{E}_{UB}$. A slightly more complicated transfer matrix calculation described in the SD can be used to determine the analogous result for a dimeric system:

$$\langle n_P \rangle = N \left[\frac{e^{-\Delta \tilde{E}_0 + \tilde{\mu}} \left(1 + \frac{\left(1 + e^{-\Delta \tilde{E}_0 + \tilde{\mu}} + 2e^{-\Delta \tilde{E}_0 - 1}\right)}{\sqrt{(1 + e^{-\Delta \tilde{E}_0 + \tilde{\mu}})^2 + 4e^{-2\Delta \tilde{E}_0 + \tilde{\mu} - 1}}} \right)}{1 + e^{-\Delta \tilde{E}_0 + \tilde{\mu}} + \sqrt{(1 + e^{-\Delta \tilde{E}_0 + \tilde{\mu}})^2 + 4e^{-2\Delta \tilde{E}_0 + \tilde{\mu} - 1}}} \right] \quad (2)$$

Above we introduce n_P, the number of bound 'proteins'. This calculation provides the lines demonstrated in Figure 2b, and fits to the data are once again achieved.
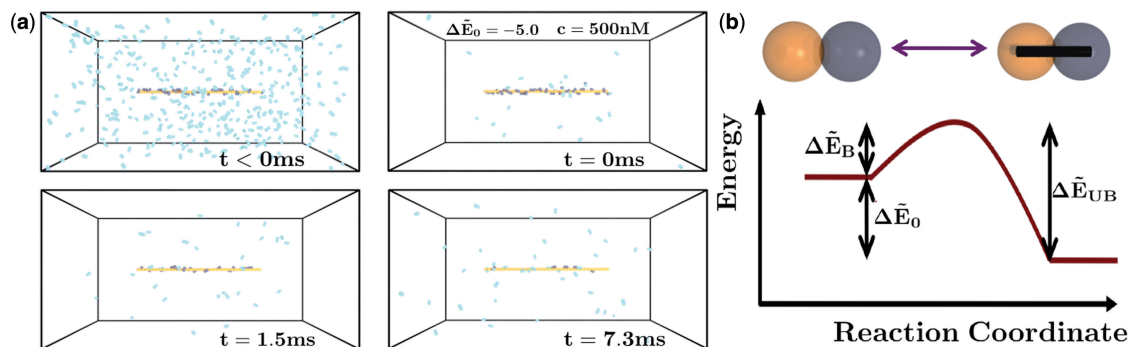
Figure 1. (a) Geometry for simulations of DNA chain (orange) in a box with periodic boundary conditions (indicated by black box). At $t < 0$ ms, high concentrations of binders (cyan) are introduced and irreversibly bind to the chain. At $t = 0$ ms, bound beads are tagged (dark blue) and cyan binders are removed to attain a concentration c. When tagged dimers pass through the boundaries, the monomers become untagged. The total sum of tagged binder monomers is n_B, which decays over time (dark blue binders are less at $t = 7.3$ ms than $t = 1.5$ ms). (b) Phenomenological energy landscape based on the Bell model, where there is an energy barrier to binding given by $\Delta\tilde{E}_B = 3.0$ (to yield binding time on the order of the diffusion time τ_D) and an energy barrier to unbinding $\Delta\tilde{E}_{UB}$. The overall binding energy is given by $\Delta\tilde{E}_0 = \Delta\tilde{E}_B - \Delta\tilde{E}_{UB}$.
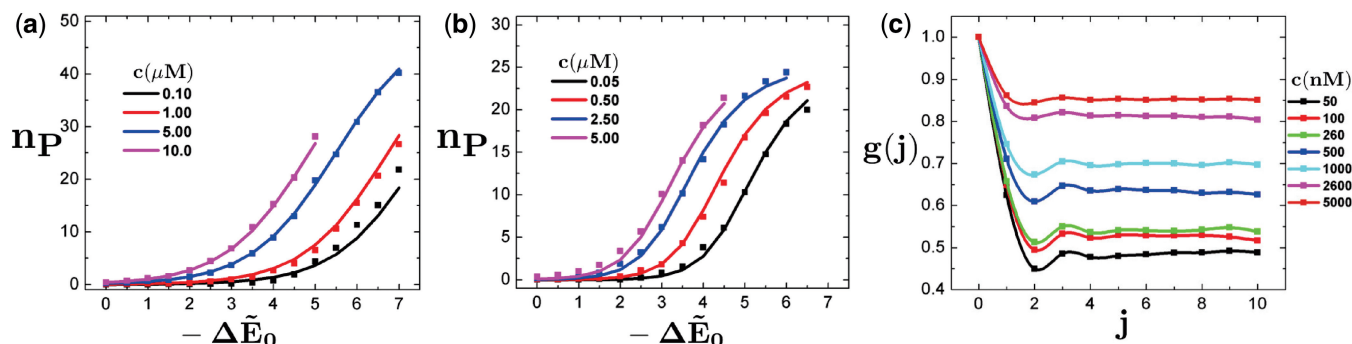

Figure 2. (a) Titration curves demonstrating the equilibrium number of binders n_P along a chain of $N = 50$ binding sites versus the change in energy on binding $\Delta\tilde{E}_0$ for the monomer (a) and dimer (b) situations for a number of different concentrations c. Points represent values determined from simulation, and the lines are the theoretical predictions assuming $\mu_0 = -11.6$ for (a) and $\mu_0 \approx -13.5$ for (b). This demonstrates that the equilibrium behavior of these systems can be described using straightforward statistical mechanics. In (c), the along-the-chain binding correlation function $g(j)$ is plotted versus index distance j. At low concentrations ($c \approx 50$ nm) there is an abundance of binders at $j = 1$ due to the dimeric behavior of the binders, along with a correlation hole at $j = 2$. This oscillation becomes smaller at large c due to the stabilization of singly bound states. $\Delta\tilde{E}_0 = -5.0$.

There is a marked decrease in the location of the apparent transition to a highly bound chain when binders are dimeric, due primarily to binding cooperativity. To further examine the equilibrium structure, we introduce the distinction between singly and doubly bound binders, the former representing states where a binder is only bound at a single monomer with the other monomer unbound and the latter representing the state where both monomers are simultaneously bound.

The singly bound state manifests itself in the along-the-chain binding correlation functions along the DNA substrate. The presence of a binder necessarily limits the binding possibilities of an adjacent binder, which is prevalent in the limit that the binders are primarily doubly bound. We quantify this effect of cooperative binding through the calculation of correlation functions $g(j) = \sum_i \langle n_i n_{i+j} \rangle / \langle n_i^2 \rangle$, which calculates the probability of having an occupied state $n_{i+j} = 1$ that is j binding positions away from an occupied state $n_i = 1$. $g(0) = 1$ at the $j = 0$ position, so this is equivalent to choosing an already bound monomer at $j = 0$ and looking at the probability of having a bound monomer j positions away. This is shown in Figure 2c, which plots $g(j)$ versus j for a

number of concentrations c. At very low concentrations, there is a large likelihood (compared with the large j limit) that there will be binding at the immediately adjacent site. This is due to the dimeric nature of the binders, and simply implies that if one binder bead is bound the adjacent binder bead is likely to be bound at the adjacent site along the DNA.

However, after this binder, there is then a 'correlation hole' where the probability of finding a bound binder is significantly smaller. This is due to the presence of the original binder; the large preference for the original dimer to take positions i and $i+1$ limits the possibilities for a dimer at $i+2$ (it is unlikely to bind at $i+1$ due to the original dimer). This is a manifestation of the well-known 'overlap' effect that provides an anti-cooperative binding behavior for multimeric DNA-binding proteins (29). The resulting oscillations in the correlation function quickly reach their averaged values at $j > 3$. At higher concentrations c, these oscillations are suppressed even at low values of j. $j = 1$ at large c is essentially equivalent to the averaged values for distant $j > 3$ binding sites, suggesting that correlations due to the dimeric structure of the binders are not significant. This suggests the presence of

singly bound states, which can 'fill' the correlation holes with what we call 'stabilized singly bound' states where a singly bound dimer at i is immediately adjacent to other binders occupying $i+1$ and $i-1$ and thus cannot become doubly bound. Evidence for such singly bound states has been seen experimentally in stoichiometric investigations on Fis binding proteins, where overloading of the DNA can occur beyond the saturation of the chain with doubly bound states (11).

This stabilization of singly bound dimers due to an abundance of neighboring dimers drives nontrivial binding dynamics in these systems. Stabilized singly bound states are much more kinetically accessible to unbinding than doubly bound dimers, allowing for a quicker mechanistic pathway toward unbinding. We analyze this aspect by probing the exchange kinetics of such systems in a way that reflects experimental literature.

Exchange kinetics

Experiments on DNA/protein binding using fluorescently tagged proteins to facilitate direct visualization of protein exchange are reproduced in simulation, as described by the 'Simulation Methods' section, which outlines a protocol meant to directly correspond to experimental procedure (16,17). We begin with $t/\tau_D = 2000$ of irreversible binding with a high concentration of binders to the DNA beads. This drives the system into a state where the chain is saturated with binders (mostly doubly bound). The concentration of surrounding (unbound) binders c is immediately dropped to a set number (anywhere from 4 to 400 dimers in the simulation box). The dimers bound to the chain are tagged. The bound dimers remain tagged until they cross the simulation boundary, whereupon they behave as one of the untagged binders. The beads are not removed, for the sake of simulation simplicity; however, in principle this could be done. This means that, in the upcoming theoretical calculations used for comparisons to simulations, c changes as the number of unbound dimers increases during the course of unbinding the initially tagged molecules. Comparisons made to experiment are not affected by this, as the concentration c is controlled by the applied flow that removes excess binders.

The number of tagged beads n_B in the simulation starting from $t = 0$ is plotted in a semilog plot in Figure 4a–c as a function of time t for a number of different concentrations c. The results are averaged over > 30 individual trajectories. There is a decay from $n_B = 50$ as time progresses, corresponding to the unbinding and subsequent diffusion of binders to the boundaries of the system. Three values of $-\Delta\tilde{E}_0$ are considered, and drastic differences are obtained. Importantly, as the value of $-\Delta\tilde{E}_0$ is increased (more binding affinity for the DNA chain), there is a corresponding increase in the effect of external concentration on the system. For small unbinding barriers and low binding affinities $-\Delta\tilde{E}_0 = 3.0$, there is little difference among the various concentrations. This suggests that the rate of unbinding is largely due to the diffusion of binders away from the DNA molecule, which is the rate-limiting step in the process of unbinding and

diffusing out of the system. However, for values of $-\Delta\tilde{E}_0 \geq 5.0$ (Figure 4b and c) the time scale of unbinding drastically decreases on increase of c. We have plotted the decay of these molecules on a semi-log plot ($\ln\langle n_B\rangle$ versus t) that reveals exponential decay at long time scales. This concentration-dependent exchange is concomitant with an overall decrease in the number of binders n_B^* (the asterisk indicates that we consider both tagged and untagged) as the binder-saturated DNA relaxes to equilibrium. We plot this in the Figure 4b inset as the number of binders $\tilde{n}_B^* = (n_B^*(t) - \langle n_B^*\rangle_{eq})/(n_B^*(t=0) - \langle n_B^*\rangle_{eq})$ normalized such that it relaxes from $\tilde{n}_B^* = 1$ to $\tilde{n}_B^* = 0$, and observe a similar concentration-dependence where higher c results in more rapid equilibration.

We note the presence of a small decay that occurs in the first $t < 1$ ms of Figure 4b and the first $t < 5$ ms of Figure 4c beyond the exponential decay at long times. This is due to the initial departure of a minority species of singly bound dimers that were present in the initial setup before the exchange process, and does not significantly affect our results. Rapid equilibration toward an initial state of near-saturation of doubly bound binders is also expected in experiment, since the experimental setup begins in a concentration regime where singly bound states are abundant and is brought into a regime where doubly bound states are known to prevail (11). The relative speed at which this process would occur is so rapid that it is not expected to be captured within the time resolution of the experimental data, and many of the singly bound binders convert directly into doubly bound binders to fill vacancies created from singly bound departures.

We seek a quantitative description of this concentration-dependent exchange process, and aspire to consider its consequences at values that are relevant to experimental values. Even with these relatively straightforward and coarse-grained calculations, simulations at the time scale of ≈ 1000 s are prohibitively computationally expensive. Therefore, we develop theory that accurately describes simulation at short time scales and can be extrapolated to experimental time scales of ≈ 1000 s. To theoretically describe this behavior, we adapt a dynamic scheme to model the evolution of the system during an exchange experiment.

Exchange kinetics theory

The articulation of these states suggests the use of a Master Equation approach, which has been demonstrated to be successful in describing similar situations where discrete changes in state occur on binding and unbinding (30). The Master Equation is a generalized version of a classical rate reaction equation, and is related to other dynamical equations such as the Fokker-Planck and Langevin equations (31). We write the Master Equation as $\frac{\partial\phi_i}{\partial t} = \sum_j k_{ij}\phi_j$ where ϕ_i is the set of states indexed by i and k_{ij} is the generalized rate constant for the motion of j to i. We define the matrix k_{ij} using the state map shown in Figure 3, which articulates a number of states that represent our definition of ϕ_i. $i = 1$ is the situation that the binder is fully bound to the DNA, $i = 2,3$ are states where the binder has separated on one of the monomers,

leaving an open spot that is unbound, $i = 4$ occurs when another binder replaces that open spot, and $i = 5$ occurs when both binders have unbound. The connections between these states form the map in Figure 3, where each arrow corresponds to a value in the definition of a matrix k_{ij}. The end result is a 5×5 matrix that is based on the conceptually relevant motions between states that is based on the probabilities used in the Bell model aspect of the simulation. Three types of transitions occur: for a fixed time interval Δt set by the Monte Carlo aspect of the simulation, unbinding transitions occur with a probability $p_{UB} = e^{-\Delta \tilde{E}_{UB}}$, binding transitions from a half-bound dimer occur with a probability of $p_B = e^{-\Delta \tilde{E}_B}$ and binding transitions with a unbound dimer occur with a probability $p_B^* = C p_B c$, where C is a constant that accounts for the geometric proportionality between the number concentration c and the likelihood that a free dimer will have a binding opportunity over Δt. For our calculations, we use $C \approx 180$; this is related to the volume within which a dimer can be located and be able to bind to the DNA binding site, and is given by $C \sim 4\pi r^3 / (3a^3)$. The ideal value is $r = 3a$ ($C \approx 113$), which is the distance between the binding site center and the dimer center if all three are lined up and adjacent. We find a slightly larger value ($C \approx 180$ corresponding to $r = 3.5a$) provides a better fit to the data, which may be due to excluded volume effects of the chain or details of the association mechanism used in our simulations.

Based on the transition probabilities p_{UB}, p_B and p^*, along with the state map in Figure 3, we write the system evolution matrix k_{ij} as follows:

$$k_{ij} = \frac{1}{\Delta t} \begin{pmatrix} -2p_{UB} & p_B & p_B & 0 & 0 \\ p_{UB} & -p' & 0 & p_{UB} & 0 \\ p_{UB} & 0 & -p' & p_{UB} & 0 \\ 0 & p_B^* & p_B^* & -3p_{UB} & 0 \\ 0 & p_{UB} & p_{UB} & p_{UB} & 0 \end{pmatrix} \quad (3)$$

where nonexistent arrows in Figure 3 are 0, $p' = (p_B + p_{UB} + p_B^*)$, and $k_{ii} = -\sum_{j \neq i} k_{ij}$ ensures that the distribution ϕ_i among states i remains properly normalized. It is possible to numerically use this matrix to calculate the exchange process directly, a process that we elaborate on in the SD. The results of this numerical calculation are shown in Figure 4d–f alongside the simulation data, and demonstrate that there is near-quantitative matching between the two approaches for all the values of c and $\Delta \tilde{E}_0$. This significantly illustrates the effectiveness of this coarse-grained approach in describing what is suspected to be a general phenomenon; a dimeric protein with monomeric portions that themselves have long-lasting binding behaviors will exhibit a concentration-dependent change due to the stabilization of the singly bound state that is put into our theoretical model via the kinetic description of State 4 in Figure 3.

Analytical expression for concentration-dependent rate constants

It is possible to develop an analytical result on incorporation of the above scheme into a conceptual
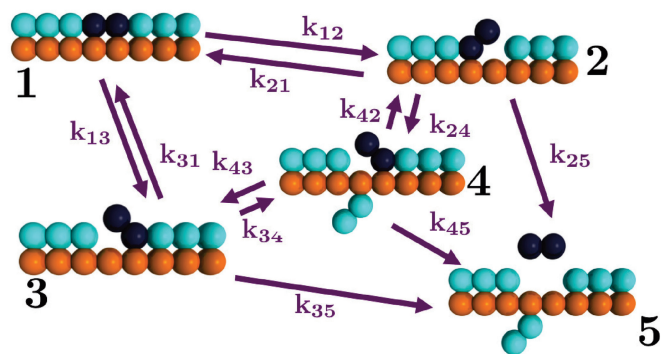

Figure 3. Schematic demonstrating the states used in the theoretical model for the dynamics of dimer-DNA binding. State 1 is a fully bound dimer (dimer is blue, DNA chain is orange, other binders are cyan), which is the initial state of the binders at the onset of an exchange experiment. States 2 and 3 are singly bound dimers with an open and adjacent DNA binding site. In State 4, this adjacent DNA binding site is occupied by a second dimer that has bound from the previously unbound binder population. The dangling side of the binder is therefore not able to rebind and revert to State 1, which is possible in 2 and 3. States 2, 3 and 4 can subsequently unbind their last remaining side to unbind completely from the DNA chain (State 5). Pathways between the States i and j are indicated with appropriate rate constants k_{ij}.

understanding of the numerical results. We define an effective off-rate constant k_{off}, and use an assumption that the 'excited' singly bound states are in equilibrium with the doubly bound state and that overall unbinding results from the unbinding of a singly bound state only. This is therefore succinctly defined as $k_{\text{off}} = \nu_{UB} \langle f_B^* \rangle$ where the value $\langle f_B^* \rangle$ corresponds to the fraction of bound dimers that are singly bound (and can therefore separate from the main chain). There are two different subpopulations within this singly bound fraction, schematically represented in Figure 5: binders that are 'excited' from the doubly bound state and can readily rebind owing to the presence of an adjacent unbound DNA site (denoted with a 0 subscript), and binders that are stabilized in a singly bound state due to the immediately adjacent bound dimers preventing rebinding to an unbound DNA site (denoted with a 1 subscript). The former state occurs in the same fashion regardless of the surrounding concentration, due to its origins as a doubly bound chain that is 'excited' into the singly bound state with a fraction $\langle f_{B,0}^* \rangle = e^{\Delta \tilde{E}_0}$. On unbinding from the singly bound state, the overall pathway represents the bare unbinding constant of the overall binder $k_0 = 2\nu_{UB} \langle f_{B,0}^* \rangle = 2\nu_B e^{2\Delta \tilde{E}_0}$ where the unbinding frequency is simply $\nu_{UB} = \nu_B e^{\tilde{E}_0}$ and the factor of two corresponds to the two different paths to unbinding (left monomer first or right monomer first). The stabilized singly bound state is present in larger quantities, since in equilibrium a pair of singly bound dimers can be included without changing the overall binding energy with the DNA (as occurs in the excited singly bound state). This follows from the creation of a pair as in Figure 5 (essentially equivalent to State 4 in Figure 3), which are now both singly bound and long-lasting. A chemical potential equilibrium must occur between the solution and the singly bound pairs (which have a mixing entropy), which
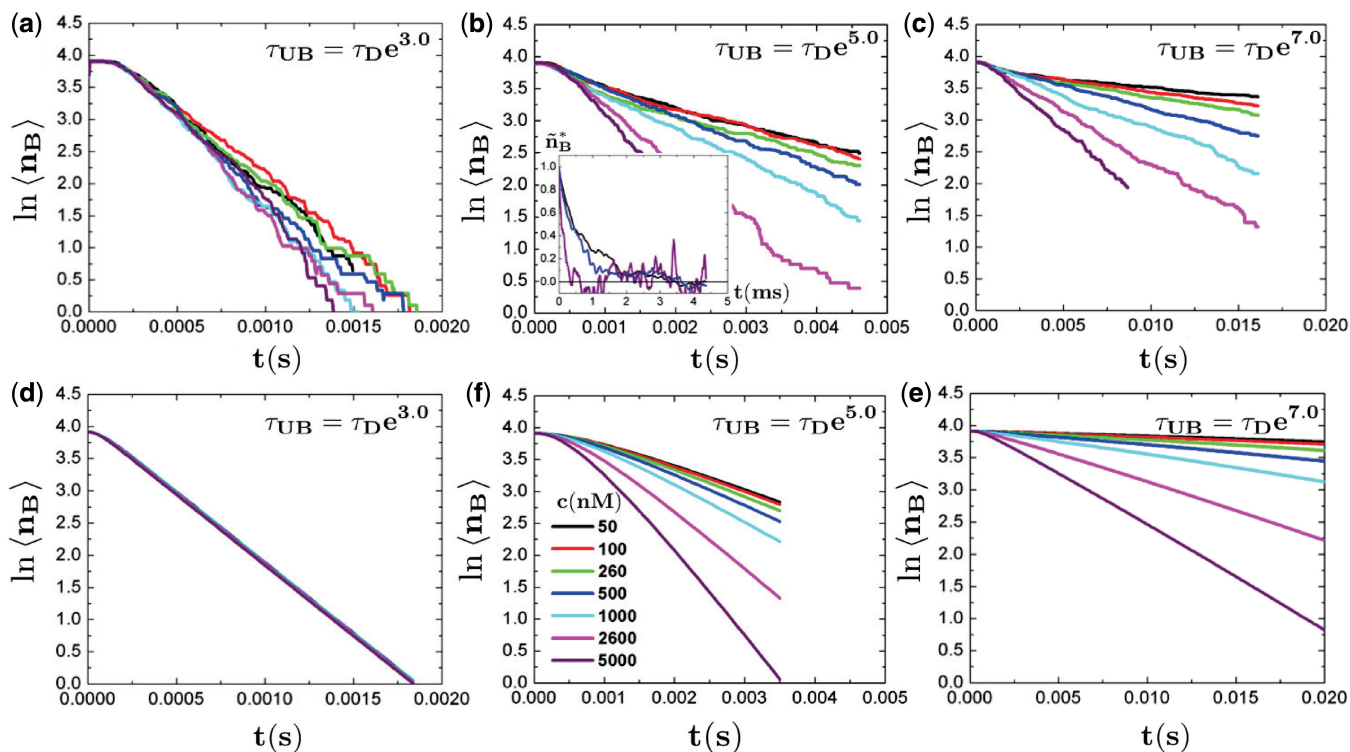
Figure 4. Binder exchange simulation data for (**a**) $\Delta\tilde{E}_0 = -3.0$, (**b**) $\Delta\tilde{E}_0 = -5.0$ and (**c**) $\Delta\tilde{E}_0 = -7.0$. At time $t = 0s$, the DNA chain is saturated with binders (number of occupied sites $n_B = 50$), and the concentration [shown in legend in (b)] is immediately applied to the system. In (a) the exchange is mostly diffusion-limited, and there is not a large difference on changing the binder concentration. In (b) and (c) exchange occurs at different rates depending on the external concentration c of untagged binders. Numerical calculation using the matrix k_{ij} in Equation (3) yields nearly quantitative matching, as shown in (**d–f**) [corresponding to (a–c), respectively]. The inset in (b) plots the normalized overall relaxation of the number of tagged and untagged binders $\tilde{n}_B^*$ as a function of time t in analogy to the main plot in (b). Colors correspond to concentrations in b and demonstrate a concentration-dependent approach to equilibrium ($\tilde{n}_B^* \to 0$) that is concomitant with the concentration-dependent exchange process.
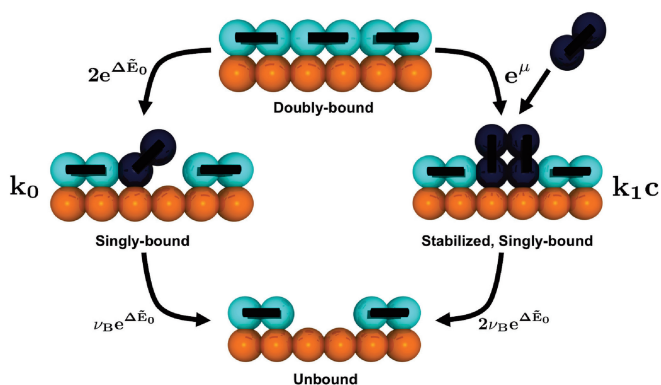

Figure 5. Schematic demonstrating the kinetic pathways used by the theory to develop Equation (4). A doubly bound dimer proceeds either through a concentration-independent pathway (left) to yield $k_0 = 2e^{\Delta\tilde{E}_0} \times v_B e^{\Delta\tilde{E}_0}$, or through the realization of a stable singly bound state (right) to yield $k_1 c = e^{\mu} \times 2v_B e^{\Delta\tilde{E}_0}$.

we write as $\ln\langle\frac{f_{B,1}^*}{2}\rangle = \tilde{\mu}_0 + \ln(c)$ We note the factor of two corresponds to the creation of two singly bound states for each increase in a single molecule bound along the chain, and that there is not a corresponding factor of two in front of the natural log on the left side of the equals sign, as we consider this pair to be located immediately adjacent on

the time scales of interest. Using the fraction of stabilized, singly bound dimers $\langle f_{B,1}^*\rangle$ to calculate its contribution to the effective unbinding rate constant $k_1 c = 2v_B c e^{\Delta\tilde{E}_0 + \tilde{\mu}_0}$ The overall expression for the rate constant is thus given as follows:

$$k_{off} \approx k_0 + k_1 c = 2v_B\left[e^{2\Delta\tilde{E}_0} + ce^{\Delta\tilde{E}_0 + \tilde{\mu}_0}\right] \tag{4}$$

assuming that the two mechanisms described are the most important ones. This is in the form proposed by Graham *et al.* (16) and it leads to a clear physical interpretation that we anticipate will lead to the facile implementation of these results to much more complicated systems. While this picture is (like the more elaborate numerical calculation) kinetic in nature, the key assumption rendering this analytical result is that the stabilized singly bound state is present in a large enough population to be considered in quasi-equilibrium with the 'bath' of doubly bound binders.

Comparison between experimental, analytical and numerical results

We can make a direct comparison between results for existing experimental literature and the aforementioned hypothesis for the binding behavior observed in the analytical and numerical results. We note that in Figure 4,
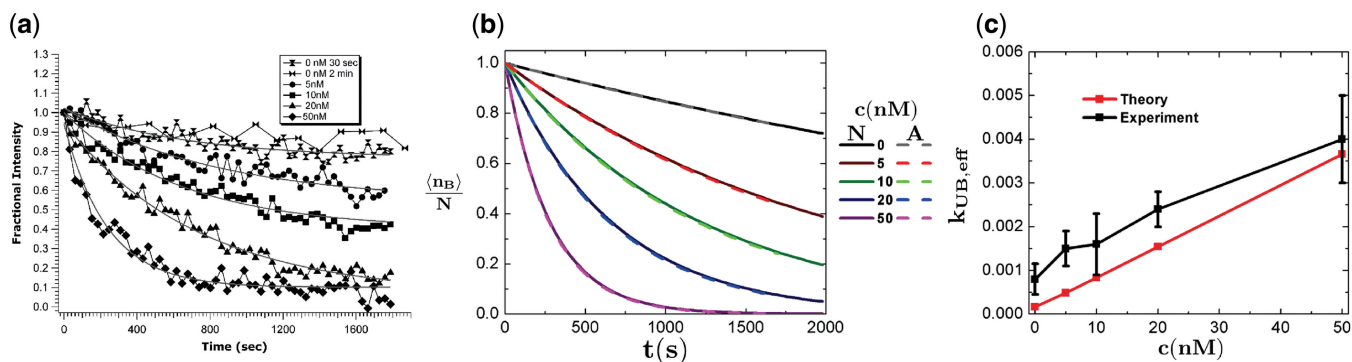
Figure 6. (a) Exchange results from the experimental literature [Graham *et al.* (16)] that demonstrate the concentration-dependent decrease in fluorescence due to a change in the unbinding reaction constant (Used by permission of Oxford University Press). (b) The analogous exchange calculation (n_B versus time t) for a binding energy of $\Delta \tilde{E}_0 = -12.3$, at concentrations relevant to experiments carried out by Graham, *et al.* (16). Both numerical (N) and analytical calculations (A) are performed, with quantitative matching between the two for $\tilde{\mu}_0 = -10.4$ in the analytical result. Qualitative matching of both of these to experimental observations in (a) is observed to occur, suggesting that this model captures much of the physics of concentration-dependent binding dynamics. Effective unbinding rate constants k_{off} (in s^{-1}) can be extracted from these plots using exponential fits, and the results are in (c). Experimental results come from Graham, *et al.* (16). Qualitative matching is demonstrated, and better fits may be obtained on adjustment of both μ_0 and $\Delta \tilde{E}_0$ used in the theory.

near-quantitative matching occurred between numerical and simulation investigations; this established that the appropriate physics are articulated in the numerical calculation such that simulation is reproduced; however, for the time scales of experimental data (≈ 1600 s) only the numerical calculation is computationally expedient. Figure 6 plots the experimental (a), numerical and analytical (both b) results on plots of bound fraction $\langle n_B \rangle / N$ versus time t in seconds. A number of concentrations c in nM are used, and clearly the appropriate trends are observed. Figure 6 demonstrates the predictions for the experimental and theoretical values for k_{off} as a function of concentration c. The slope is k_1 and the intercept is k_0, both values which match qualitatively with the theoretical predictions. Disparities arise in the apparent equilibrium values of the decays observed in the experiment versus the calculation, with the latter decaying to $\langle n_B \rangle / N \to 0$ and the former demonstrating a decay to a finite $\langle n_B \rangle / N > 0$. We attribute this to sequence-dependent effects, and verify that this result is likely due to binding energy inhomogeneity, via analysis in the SD of the case of Gaussian random distributed binding energies (Supplementary Data, Fig. S3). The manifestation of this effect suggests that sequence-dependence may richly affect concentration-dependent off-rate behavior in *in vivo* environments where local binder competition may couple with highly variable and correlated sequence-dependent binding affinities.

CONCLUSION

Experimental work on DNA-binding proteins Fis, HU and NHP6A has recently demonstrated the appearance of concentration-dependent unbinding (16), with similar behaviors observed in other systems (22). This behavior does not occur in usual models of protein–DNA interactions widely used to fit kinetic data where there is by construction a concentration-independent off-rate. This concentration-independent unbinding models a 'spontaneous dissociation' pathway, involving a transition over a

rate-limiting barrier associated with the DNA–protein complex itself and not dependent on other nearby molecules. Here we have shown that simply generalizing such models to dimeric proteins having *two* interaction sites per protein, with each subunit having concentration-independent unbinding kinetics, leads to 'concentration-dependent' off-kinetics of the form $k_{off} = k_0 + k_1 c$ as observed for Fis. This arises from the conformational flexibility of the dimeric protein model, which allows partial dissociation of one side of the protein from its binding site, and subsequent capture of that site by a second protein from solution. The rate at which this can occur is simply controlled by solution phase concentration, as observed in the single-molecule studies (16,22). Our model is able to reproduce rates seen experimentally well, with reasonable choices of microscopic parameters. We note that in the concentration range near to the apparent dissociation point, exchange is the dominant process controlling dissociation, and there is effectively a stabilization of the partially bound state.

Our model is exceedingly simple; undoubtedly more complex behaviors can be generated by addition of additional features. For example, we do not consider the role of binding cooperativity in this model except through the implicit inclusion of overlapping anticooperative effects present in multimeric-binding systems (29). Cooperativity may serve to promote competition effects by favoring the clustering of binders that promotes singly bound dimers. We also consider DNA chains that are essentially saturated; *in vivo* behavior may provide different mechanisms for binding site competition or sequence-dependent effects that are not articulated in this model. These are thus open questions that when investigated may reveal a rich array of competitions that could enhance or suppress concentration-dependent unbinding effects. However, the present article shows that much of the concentration-dependent unbinding observed for Fis can be attributed to its dimeric structure, without invoking more complex causes.

The simplicity of our model suggests that concentration-dependent unbinding effects should be observable for a wide range of molecular interactions. The only essential ingredients are at least two isolated binding elements, and sufficient conformational flexibility of the binding partners so that partial unbinding can allow solvated species the ability to compete with partially unbound ones. We anticipate that the mechanism discussed here for concentration-dependent unbinding may well be applicable to interactions dependent on 'single' binding domains, as long as they involve multiple chemical interactions, as is usual even for protein–DNA and protein–protein interactions involving 'single' binding-domains [which might be considered to be the case for NHP6A (16)]. Furthermore, the mechanism discussed in this article for acceleration of off-rates may be crucial to turning over proteins on DNA, thus increasing the speed of response for changing gene regulation. One could imagine that many DNA-binding proteins have been selected for their ability to be enhanced by this mechanism: a given protein may be stably bound when suitable competitors are not present, but then may be rapidly replaced by another molecule itself selected to be able to invade the former's binding site. We anticipate that *in vivo*, where large concentrations of molecules compete for binding partners, the mechanism for concentration-dependent exchange outlined here may strongly effect the rates of turnover of many molecular species through competitions with effects such as cooperativity, sequence dependence and local crowding. However, we also emphasize that as yet, concentration-dependent unbinding has only been quantitatively studied for a few proteins interacting with DNA; careful case-by-case studies are necessary to determine the generality of this phenomenon.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

FUNDING

REFERENCES

1. Halford,S.E. and Marko,J.F. (2004) How do site-specific DNA-binding proteins find their targets? *Nucleic Acids Res.*, **32**, 3040–3052.
2. Bert,O.G., Winter,R.B. and von Hippel,P.H. (1981) Diffusion-driven mechanisms of protein translocation on nucleic acids. 1. Models and theory. *Biochemistry*, **20**, 6929–6948.
3. Hu,T. and Shklovskii,B.I. (2006) How does a protein search for the specific site on DNA: the role of disorder. *Phys. Rev. E*, **74**, 021903.
4. Gorman,J. and Greene,E.C. (2008) Visualizing one-dimensional diffusion of proteins along DNA. *Nat. Struct. Mol. Biol.*, **15**, 768–774.
5. Lomholt,M.A., van den Broek,B., Kalisch,S.M.J., Wuite,G.J.L. and Metzler,R. (2009) Facilitated diffusion with DNA coiling. *Proc. Natl Acad. Sci. USA*, **106**, 8204–8208.
6. Elf,J., Li,G.-W. and Xie,X.S. (2007) Probing transcription factor dynamics at the single-molecule level in a living cell. *Science*, **316**, 1191–1194.
7. Wang,Y.M., Austin,R.H. and Cox,E.C. (2006) Single molecule measurements of repressor protein 1D diffusion on DNA. *Phys. Rev. Lett.*, **97**, 048302.
8. Blainey,P.C., van Oijen,A.M., Banerjee,A., Verdine,G.L. and Xie,X.S. (2006) A base-excision DNA-repair protein finds intrahelical lesion bases by fast sliding in contact with DNA. *Proc. Natl Acad. Sci. USA*, **103**, 5752–5757.
9. Dorman,C. (2009) Nucleoid-associated proteins and bacterial physiology. *Adv. Appl. Microbiol.*, **67**, 47–64.
10. Stillman,D.J. (2010) Nhp6: A small but powerful effector of chromatin structure in *Saccharomyces cerevisiae. Biochim. Biophys. Acta*, **1799**, 175–180.
11. Skoko,D., Yoo,D., Bai,H., Schnurr,B., Yan,J., McLeod,S.M., Marko,J.F. and Johnson,R.C. (2006) Mechanism of chromosome compaction and looping by the *E. Coli* nucleoid protein Fis. *J. Mol. Biol.*, **364**, 777–798.
12. Shao,Y., Feldman-Cohen,F.S. and Osuna,R. (2008) Functional characterization of the *Escherichia coli* Fis-DNA binding sequence. *J. Mol. Biol.*, **376**, 771–785.
13. Stella,S., Cascio,D. and Johnson,R.C. (2010) The shape of the DNA minor groove directs binding by the DNA-bending protein Fis. *Genes Dev.*, **24**, 814–826.
14. McCauley,M., Hardwidge,P.R., Maher,L.J. III and Williams,M.C. (2005) Dual binding modes for an HMG domain from human HMGB2 on DNA. *Biophys. J.*, **89**, 353–364.
15. Sugimura,S. and Crothers,D.M. (2006) Stepwise binding and bending of DNA by *Escherichia coli* integration host factor. *Proc. Natl Acad. Sci. USA*, **103**, 18510–18514.
16. Graham,J.S., Johnson,R.C. and Marko,J.F. (2011) Concentration-dependent exchange accelerates turnover of proteins bound to double-stranded DNA. *Nucleic Acids Res.*, **39**, 2249–2259.
17. Graham,J.S., Johnson,R.C. and Marko,J.F. (2011) Counting proteins bound to a single DNA molecule. *Biochem. Biophys. Res. Commun.*, **415**, 131–134.
18. Skoko,D., Wong,B., Johnson,R.C. and Marko,J.F. (2004) Micromechanical analysis of the binding of DNA-bending proteins HMGB1, NHP6A, and HU reveals their ability to form highly stable DNA-protein complexes. *Biochemistry*, **43**, 13867–13874.
19. Skoko,D., Yan,J., Johnson,R.C. and Marko,J.F. (2005) Low-force DNA condensation and discontinuous high-force decondensation reveal a loop-stabilizing function of the protein Fis. *Phys. Rev. Lett.*, **95**, 208101.
20. Xiao,B., Johnson,R.C. and Marko,J.F. (2010) Modulation of HU-DNA interactions by salt concentration and applied force. *Nucleic Acids Res.*, **38**, 6176–6185.
21. Micah,M.J., Rueter,E.M., Rouzina,I., Maher,L.J. III and Williams,M.C. (2013) Single-molecule kinetics reveal microscopic mechanism by which high-mobility group B proteins alter DNA flexibility. *Nucleic Acids Res.*, **41**, 167–181.
22. Joshi,C.P., Panda,D., Martell,D.J., Andoy,N.M., Chen,T.-Y., Gaballa,A., Helmann,J.D. and Chen,P. (2012) Direct substitution and assisted dissociation pathways for turning off transcription by a MerR-family metalloregulator. *Proc. Natl Acad. Sci. USA*, **109**, 15121–15126.
23. Ha,T.J. (2013) Single-molecule approaches embrace molecular cohorts. *Cell*, **154**, 723–726.
24. Sidorova,N.Y., Scott,T. and Rau,D.C. (2013) DNA concentration-dependent dissociation of *Eco*RI: Direct transfer or reaction during hopping. *Biophys. J.*, **104**, 1296–1303.
25. Kunzelmann,S., Morris,C., Chavda,A.P., Eccleston,J.F. and Webb,M.R. (2010) Mechanism of interaction between single-stranded DNA binding protein and DNA. *Biochemistry*, **49**, 843–852.

26. Flory,P.J. (1953) *Principles of Polymer Chemistry*. Cornell University Press, Ithaca.
27. Bell,G.I. (1978) Models for specific adhesion of cells to cells. *Science*, **200**, 618–627.
28. Sing,C.E. and Alexander-Katz,A. (2011) Equilibrium structure and dynamics of self-associating single polymers. *Macromolecules*, **44**, 6962–6971.
29. McGhee,J.D. and von Hippel,P.H. (1974) 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**, 469–489.
30. Sing,C.E. and Alexander-Katz,A. (2012) Force spectroscopy of self-associating homopolymers. *Macromolecules*, **45**, 6704–6718.
31. Zwanzig,R. (2001) *Nonequilibrium Statistical Mechanics*. Oxford University Press, New York.