

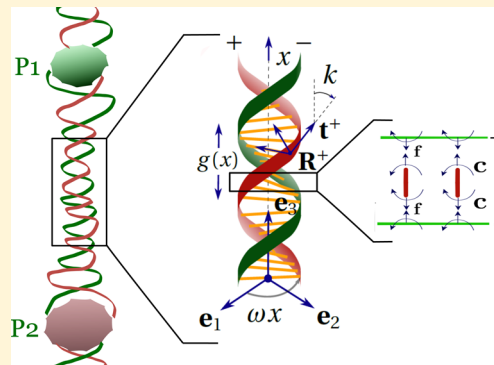
Elasticity as the Basis of Allostery in DNA

Jaspreet Singh¹ and Prashant K. Purohit*

Department of Mechanical Engineering and Applied Mechanics, University of Pennsylvania, Philadelphia, Pennsylvania 19104, United States

Supporting Information

ABSTRACT: Allosteric interactions in DNA are crucial for various biological processes. These interactions are quantified by measuring the change in free energy as a function of the distance between the binding sites for two ligands. Here, we show that trends in the interaction energy of ligands binding to DNA can be explained within an elastic birod model, which accounts for the deformation of each strand as well as the change in stacking energy due to perturbations in position and orientation of the bases caused by the binding of ligands. The strain fields produced by the ligands decay with distance from the binding site. The interaction energy of two ligands decays exponentially with the distance between them and oscillates with the periodicity of the double helix in quantitative agreement with experimental measurements. The trend in the computed interaction energy is similar to that in the perturbation of groove width produced by the binding of a single ligand, which is consistent with molecular simulations. Our analysis provides a new framework to understand allosteric interactions in DNA and can be extended to other rod-like macromolecules whose elasticity plays a role in biological functions.



INTRODUCTION

When a ligand binds to DNA, it induces conformational changes at the binding site, which could propagate to regions tens of base-pairs away, thereby encouraging or inhibiting the binding of a second ligand in those places. Such interactions between two binding agents are called *allosteric* interactions. Our focus here is on a mechanism for allostery based on the elasticity of long molecules. Although we will illustrate our theory using DNA as an example, long-range allosteric interactions have been documented in actin, microtubules, and helical peptide chains. For example, myosin binds to actin filaments leading to the suppression of the formation of cofilin clusters via allosteric signaling.¹ Long-range structural changes induced by taxol binding to microtubules inside a cell prevents cell division, thus making it a potent antitumor agent.² The transfer of chiral stimulus triggered by a binding agent across a helical peptide chain gives the molecule an overall chiral character and is yet another instance of allostery.³ Instances of allostery in DNA have been known for decades.^{4,5} Antitumor drug actinomycin D binds to DNA by intercalating between the adjacent base pairs⁵ containing a guanine base. However, in the presence of daunomycin, another antitumor drug, actinomycin, is observed to bind to poly(dAT) DNA oligomers too. This is due to the allosteric stabilizing influence exerted by the already bound daunomycin molecule near its binding site. We will analyze allostery in ds-DNA because detailed experimental and simulation results are available for it,^{6–8} thus allowing quantitative comparisons with our theory.

We define the allosteric interaction energy $\Delta G = E_{12}^{(0,p)} - E_1^0 - E_2^0$, where $E_{12}^{(0,p)}$ is the free energy of the protein–DNA

complex consisting of two proteins separated by distance p and E_1^0 and E_2^0 are the free energies of the protein–DNA complexes consisting of one protein. Kim et al.⁷ have conducted single-molecule experiments using fluorescence techniques to measure allosteric interaction energy ΔG between two proteins on a DNA oligomer. In their paper,⁷ DNA-binding proteins are categorized as (a) proteins that bend DNA such as LacR and T7-RNAP and (b) proteins that bind to straight DNA such as GRDBD and BamHI. Here, we deal with the latter category.

We use the theory of elastic birods⁹ to develop a mechanical model for investigating protein–DNA interactions. A birod consists of two elastic strands joined by an elastic web. We represent the sugar-phosphate backbone of DNA using the outer strands and the complementary base pairing is modeled using the elastic web. A birod model of DNA has different properties than a homogeneous rod model at short length scales; however, as the length of the birod increases, the elastic properties of both models become indistinguishable.¹⁰

We discuss key features that distinguish our model from the state-of-the-art⁸ worm-like chain model for DNA allostery.

1. **Helical geometry:** Kim et al.¹¹ discovered that the interaction energy ΔG between two proteins on DNA decays exponentially while oscillating with the periodicity of the DNA double helix. It is thereby imperative that we account for the double helical geometry of DNA,

Received: August 2, 2018

Revised: December 14, 2018

Published: December 15, 2018

- which is conveniently incorporated in a birod model,⁹ but is absent in a worm-like chain model of DNA.
2. **Elasticity of base pairs:** Proteins interact with DNA by altering the geometry of the double helix, such as changing the width of major/minor groove.^{6,12} The elasticity of the base pairs, represented by the elastic web in a birod model is essential to accurately model these local deformations.
 3. **Stacking energy:** Stacking energy penalizes the change in orientations of the base pairs with respect to each other. We use a stacking energy quadratic in the twist and stretch of the DNA-double helix.

In an existing model⁸ of allostery, tension in the worm-like polymer chain to which the two proteins are bound plays an important role in the decaying oscillatory behavior of the interaction energy ΔG . However, in the experiments of Kim et al.⁷ and simulations of Dršata et al.,⁶ the oscillatory exponentially decaying allosteric interactions on DNA are present even in the absence of tension. Here, we use an analytical model to show that decaying oscillatory behavior of the interaction energy can arise from the interplay between the double helical geometry and the elasticity of the base pairs.

Our model provides key insights into the structural deformations of the DNA helix, changes in the groove width when a ligand binds to DNA, and the allosteric interaction energy ΔG between two proteins on DNA. We compute the correlations between the displacement variables at the two sites of protein binding and establish their connection to the interaction energy between the two proteins. Our results are in excellent quantitative agreement with the experimental data in Kim et al.⁷ and Kopka et al.¹²

Numerical simulations^{6,13} have been used to propose mechanical models for DNA allostery. These papers describe DNA using three sets of coordinates: intra-base-pair coordinates buckle, propeller, opening, shear, stretch and stagger, inter-base-pair coordinates tilt, roll, twist, shift, slide, and rise, and major and minor groove widths. The DNA-binding protein fixes some (or all) of the degrees of freedom at the site of binding, resulting in deformations away from it. The energy of binding can therefore be computed. The approach is comprehensive, but computationally expensive. Drawing upon the know-how from Dršata et al.,^{6,13} we allow for the bending, twisting, stretching, and shearing of the base pairs. Furthermore, we go beyond their numerical models by considering the mechanics of the outer strands, which, as pointed out later, is crucial to getting the correct twist–stretch coupling for double-stranded DNA.

THEORY

DNA comprises of two helical strands held together via complementary base pairing. When a ligand, such as a protein or a drug, binds to DNA, it exerts forces and moments on the double helix^{14,15} causing deformations at the base-pair level. We use the theory of birods⁹ to investigate these deformations. A birod consists of two elastic strands that interact through an elastic web. This construction makes it suitable for investigating the deformations at the base-pair level in a DNA molecule, which a homogeneous rod model cannot capture.¹⁶ The latter ignores the double-helical structure and the elasticity of the base pairs, both of which are crucial to the problem under consideration. In the following $(\cdot)_x$ denotes $\frac{\partial(\cdot)}{\partial x}$. Lower case letters such as a , r , β^\pm are scalars, bold lower case

letters such as \mathbf{t}^+ , \mathbf{b}^- are vectors, whereas bold upper case letters such as \mathbf{R}^+ , \mathbf{R}_0^+ , \mathbf{Z} are 3×3 tensors.

We assume the phosphate backbones comprising of phosphodiester bonds, represented by outer strands in our model to be inextensible and unshearable. Our assumption of inextensibility can be justified by examining the elastic properties of ssDNA. The stretch modulus of ssDNA, at high ion concentration, is about 1000 pN,¹⁷ which is about twice that of the largest modulus ($K_c = 600$ pN) used in our model. At low ion concentrations, under which the experiments reported in Kim et al. are conducted, ssDNA shows an even stiffer response.¹⁸ The assumption of unshearability is justified if the cross-sectional dimensions are much smaller than the length of a rod,¹⁹ which is indeed true for DNA molecules. Because these backbones consist of consecutive single bonds that allow free rotation about the bond, we assume that they cannot resist twisting moments. The base pairing is represented by the elastic web, which is capable of extending, shearing, bending, and twisting. In addition to the elastic energy, we consider contributions from the stacking energy, which is associated with the change in orientations of the successive base pairs.

In this paper, we consider the proteins that do not bend DNA. Kim et al.⁷ report experimental data for allosteric interaction energy using a DNA-binding protein pair: GRDBD and BamHI, both of which do not bend DNA.^{20,21} Kim et al.⁷ have also experimented with proteins that bend DNA, and they have been theoretically treated elsewhere.^{22,23} We denote the helical strands as \pm ; their positions in the reference state are denoted by \mathbf{r}_0^\pm . We use arclength parameter x to parametrize the double helix (Figure 1). Thus,

$$\begin{aligned}\mathbf{r}_0^+ &= a(\cos \omega x \mathbf{e}_1 + \sin \omega x \mathbf{e}_2) + x \mathbf{e}_3, \\ \mathbf{r}_0^- &= a(\cos(\omega x + \alpha) \mathbf{e}_1 + \sin(\omega x + \alpha) \mathbf{e}_2) + x \mathbf{e}_3\end{aligned}\quad (1)$$

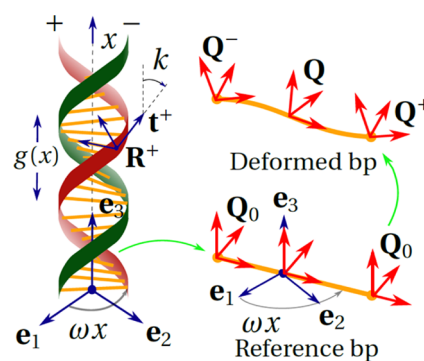


Figure 1. Birod model of DNA. The angle between the tangent \mathbf{t}^+ and \mathbf{e}_3 is k . A base pair in reference and deformed state is shown. The director frames attached to \pm ends of the base pair change from \mathbf{Q}_0 to \mathbf{Q}^\pm , respectively. The rigid rotation of the strand $\mathbf{Q} = (\mathbf{Q}^+ \mathbf{Q}^{-T})^{1/2} \mathbf{Q}^-$ and microrotation $\mathbf{P} = (\mathbf{Q}^+ \mathbf{Q}^{-T})^{1/2}$.

where $a = 1$ nm is the radius of the DNA helix, $p = 3.4$ nm is the pitch, $\omega = \frac{2\pi}{p}$, and α is the phase difference between the helices. Here, we assume $\alpha = \pi$ to make the computations analytically tractable. We consider a deformed configuration where the double helix extends and twists about \mathbf{e}_3 , and its radius and phase angle also change due to binding of ligands. The deformed state of the \pm strands is denoted by $\mathbf{r}^\pm(x)$, where

$$\begin{aligned}
\mathbf{r}^+(x) &= (a+r)(\cos(\omega x + \beta^+) \mathbf{d}_1 + \sin(\omega x + \beta^+) \mathbf{d}_2) \\
&\quad + (x + \int_{-\infty}^x a \xi^+ dx) \mathbf{e}_3, \\
\mathbf{r}^-(x) &= -(a+r)(\cos(\omega x + \beta^-) \mathbf{d}_1 + \sin(\omega x + \beta^-) \mathbf{d}_2) \\
&\quad + (x + \int_{-\infty}^x a \xi^- dx) \mathbf{e}_3
\end{aligned} \quad (2)$$

such that $\mathbf{d}_{1x} = k_3 \mathbf{d}_2$ and $\mathbf{d}_{2x} = -k_3 \mathbf{d}_1$. To gain a better physical insight, we give a visual representation of the deformation described in the above equation in Figure 2. We only show the

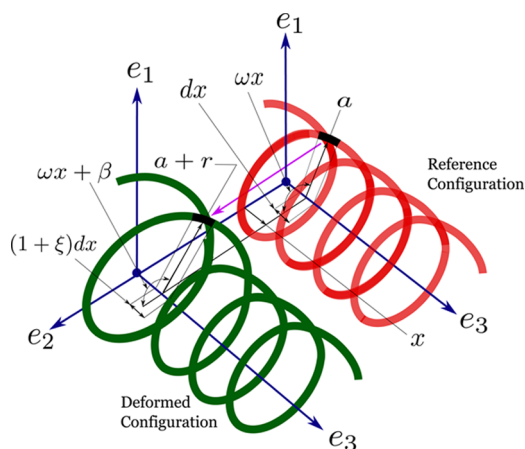


Figure 2. Distortions to the helical geometry of the + strand. We assume that the DNA remains straight after the binding of proteins. Therefore, the axis \mathbf{e}_3 remains undisturbed. The radius of the helix a changes to $a+r$, the phase angle ωx changes to $\omega x + \beta$, and the length of the infinitesimal element dx , shown in black, changes to $(1 + \xi) dx$.

deformation for the + strand for clarity and indicate the strain variables r (change in radius), β (change in phase angle), and ξ (stretch of the center-line). We assume all the displacement and strain parameters r , β^\pm , and ξ^\pm vanish at $x = \pm\infty$ because the deformations caused by the proteins are local. The change in radius r , change in the phase angle β^\pm , stretches ξ^\pm , and the twist k_3 are assumed to be small ($\sim O(\varepsilon)$) such that second-order terms such as r^2 and $\xi\beta^-$ are negligible. However, there could be finite rotations resulting from k_3 .

As shown in Figure 3, the birod consists of two elastic strands joined by an elastic web. We show a straight elastic ladder for easy visualization of the key forces and moments. The + strand exerts a body force \mathbf{l} and a body moment \mathbf{c} on the

– strand via the elastic web. The balance laws for the two outer strands constitute the governing equations for the birod.⁹ The deformation of the elastic web can be calculated once the deformation of the outer strands is known. We need to solve the following balance equations for a helical birod,

$$\begin{aligned}
\mathbf{n}_x^\pm \mp \mathbf{f} + \mathbf{l} &= 0, \\
\mathbf{m}_x^\pm + \mathbf{r}_x^\pm \times \mathbf{n}^\pm + \frac{1}{2}(\mathbf{r}^+ - \mathbf{r}^-) \times \mathbf{f} \mp \mathbf{c} + \mathbf{h} &= 0
\end{aligned} \quad (3)$$

where \mathbf{m}^\pm and \mathbf{n}^\pm are the contact moment and contact force, respectively, in \pm strands. \mathbf{f} and \mathbf{c} are the distributed force and distributed moment exerted by the + strand on the – strand. \mathbf{l} and \mathbf{h} are the body force and body moment exerted by the base pairs onto both \pm strands. In what follows, we use the position vectors for the deformed helix $\mathbf{r}^\pm(x)$ (eq 2) to compute these quantities. The constitutive relations for the forces \mathbf{n}^\pm , moments \mathbf{m}^\pm , and the force \mathbf{l} and moment \mathbf{c} transferred by the web, are discussed in the relevant subsections.

Contact Forces in the Outer Strands (\mathbf{n}^\pm). The outer strands are inextensible, which means $|\mathbf{r}_x^\pm| = |\mathbf{r}_{0x}^\pm|$ yielding

$$\omega^2 r + a\omega(k_3 + \beta_x^\pm) + \xi^\pm = 0 \quad (4)$$

We use the above equation to eliminate ξ^\pm from eq 2. Due to the constraint in eq 4, the contact forces \mathbf{n}^\pm enter as Lagrange multipliers.

Contact Moments in the Outer Strands (\mathbf{m}^\pm). We attach a director frame $\mathbf{R}^\pm = [\mathbf{n}_0^\pm \mathbf{b}_0^\pm \mathbf{t}_0^\pm]$ to each cross section of the \pm strands, where \mathbf{n}_0^\pm , \mathbf{b}_0^\pm , and \mathbf{t}_0^\pm are the normal, binormal, and tangent in the reference state to \pm strand, respectively. \mathbf{n}^\pm , \mathbf{b}_0^\pm , \mathbf{t}_0^\pm , and the curvature in the reference configuration Ω_0^\pm are computed using eq 1 as follows

$$\begin{aligned}
\mathbf{t}_0^\pm &= \frac{\mathbf{r}_{0x}^\pm}{|\mathbf{r}_{0x}^\pm|} = \pm \sin k (-\sin \omega x \mathbf{e}_1 + \cos \omega x \mathbf{e}_2) + \cos k \mathbf{e}_3, \\
\mathbf{n}_0^\pm &= \frac{\mathbf{t}_x^\pm}{|\mathbf{t}_x^\pm|} = \mp (\cos \omega x \mathbf{e}_1 + \sin \omega x \mathbf{e}_2), \\
\mathbf{b}_0^\pm &= \mathbf{t}_0^\pm \times \mathbf{n}_0^\pm \\
&= \mp \cos k (-\sin \omega x \mathbf{e}_1 + \cos \omega x \mathbf{e}_2) + \sin k \mathbf{e}_3, \\
\Omega_0^\pm &= \Omega_0 = (\mathbf{t}_0^\pm \cdot \mathbf{t}_0^\pm)^{1/2} = \omega \sin k
\end{aligned} \quad (5)$$

Similarly, we use eq 2 to compute the Frenet–Serret frame $\mathbf{R}^\pm = [\mathbf{n}^\pm \mathbf{b}^\pm \mathbf{t}^\pm]$ and curvature Ω^\pm in the deformed state. We neglect terms higher than first order, such as $r\beta^+$, ξ^- , $r \sim O(\varepsilon^2)$, and get

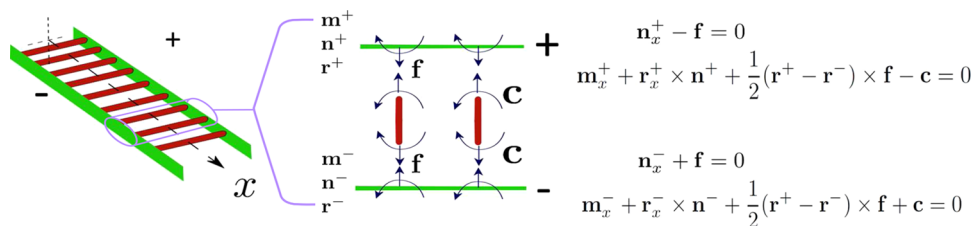


Figure 3. Free-body diagrams that establish the connection between an elastic rod and an elastic birod. We deliberately show a straight ladder instead of helical birod to illustrate the mechanics. An elastic birod comprises two elastic rods + and –. The + strand exerts a moment \mathbf{c} and force \mathbf{f} on the – strand through an elastic web. This transfer of moment and force leads to deformation of the web. In the figure, \mathbf{r}^\pm denotes the position vector for \pm strands and \mathbf{n}^\pm and \mathbf{m}^\pm denote the contact forces and contact moments in \pm strands, respectively. The force and moment balance for + and – strand constitute the governing equations (eq 3) for the elastic birod. For further discussion, see Moakher and Maddocks.⁹

$$\begin{aligned}
\mathbf{R}^\pm &= [\mathbf{n}^\pm \quad \mathbf{b}^\pm \quad \mathbf{t}^\pm] = \mathbf{Z}\mathbf{R}_0^\pm(\mathbf{I} + \Theta^\pm), \\
\mathbf{Z} &= \mathbf{d}_1 \otimes \mathbf{e}_1 + \mathbf{d}_2 \otimes \mathbf{e}_2 + \mathbf{e}_3 \otimes \mathbf{e}_3, \\
\Theta^\pm &= \begin{bmatrix} 0 & -\theta_3^\pm & \theta_2^\pm \\ \theta_3^\pm & 0 & -\theta_1^\pm \\ -\theta_2^\pm & \theta_1^\pm & 0 \end{bmatrix} \\
\theta_1^\pm &= r\omega + a(\beta_x^\pm + k_3), \quad \theta_2^\pm = -r_x \cos k + \beta^\pm \sin k, \\
\theta_3^\pm &= \frac{-\omega r_x - a(\beta_{xx}^\pm + k_{3x})}{\omega \sin k} - \frac{(r_x \cos k - \beta^\pm \sin k) \cos k}{\sin k}
\end{aligned} \quad (6)$$

The bending moment in the outer strands \mathbf{m}^\pm is proportional to the change in curvature $\kappa^\pm = \Omega^\pm - \Omega_0^\pm$ and is directed along the binormal \mathbf{b}^\pm such that $\mathbf{m}^\pm = EI\kappa^\pm\mathbf{b}^\pm$, where EI is the bending modulus of the strand. Note that the twisting moment is zero because the phosphate backbone consists of single bonds that permit free rotations.

Force and Moment Transferred by the Web (\mathbf{f} , \mathbf{c}).

Now, we compute the bending and twisting of the web, which represents base pairing. We attach a director frame \mathbf{Q}_0 to both + and - end of the base pair (Figure 1).

$$\mathbf{Q}_0 = [\mathbf{e}_r \quad \mathbf{e}_\theta \quad \mathbf{e}_3] \quad (7)$$

where $\mathbf{e}_r = \cos \omega x \mathbf{e}_1 + \sin \omega x \mathbf{e}_2$ and $\mathbf{e}_\theta = -\sin \omega x \mathbf{e}_1 + \cos \omega x \mathbf{e}_2$. As the birod deforms, these frames respectively get mapped to \mathbf{Q}^\pm . We compute \mathbf{Q}^\pm using the deformation of \mathbf{R}^\pm from eq 6 keeping in mind that the angles between the columns of \mathbf{R}_0^\pm and \mathbf{Q}_0^\pm should remain constant during deformation implying $(\mathbf{R}_0^{\pm T} \mathbf{Q}_0 = \mathbf{R}^{\pm T} \mathbf{Q})$, thus

$$\mathbf{Q}^\pm = \mathbf{Z}\mathbf{R}_0^\pm(\mathbf{I} + \Theta^\pm)\mathbf{R}_0^{\pm T}\mathbf{Q}_0, \quad \Theta^\pm \sim O(\varepsilon) \quad (8)$$

Now, we can compute the rigid rotation \mathbf{Q} and microrotation \mathbf{P} for each base pair. The microrotation contains information about the “difference” between the rotations \mathbf{Q}^\pm .⁹ This is related to the moment transferred by the base pair \mathbf{c} via an elastic constitutive relation for the web

$$\mathbf{P} = (\mathbf{Q}^+\mathbf{Q}^{-T})^{1/2} = \mathbf{Z}(\mathbf{I} + \Phi^c)\mathbf{Z}^T \quad (9)$$

Here, $\Phi^c = \frac{\mathbf{R}^+\Theta^+\mathbf{R}^{+T} - \mathbf{R}^-\Theta^-\mathbf{R}^{-T}}{2}$ is a skew symmetric tensor. The moment transferred by the base pair is directly proportional to the Gibbs vector of \mathbf{P} . $\boldsymbol{\eta} = \tan \frac{\lambda}{2} \hat{\mathbf{k}}$ is a Gibbs rotation vector for a rotation matrix \mathbf{T} if $\mathbf{T}\hat{\mathbf{k}} = \hat{\mathbf{k}}$ and $1 + 2 \cos \lambda = \text{tr } \mathbf{T}$. In our case, the Gibbs vector of \mathbf{P} is $2\boldsymbol{\eta} = 2\mathbf{Z}\boldsymbol{\eta} = \mathbf{Z}\boldsymbol{\phi}^c$, where $\boldsymbol{\phi}^c$ is the axial vector of skew symmetric tensor Φ^c . Note that in the reference state, $\boldsymbol{\eta}_0 = 0$ because $\mathbf{P}_0 = (\mathbf{Q}_0\mathbf{Q}_0^T)^{1/2} = \mathbf{I}$. The rigid rotation of the base pair $\mathbf{Q} = \mathbf{P}\mathbf{Q}^-$. Here,

$$\mathbf{Q} = \mathbf{Z}(\mathbf{I} + \Phi)\mathbf{Q}_0 \quad (10)$$

and $\Phi = \frac{\mathbf{R}^+\Theta^+\mathbf{R}^{+T} + \mathbf{R}^-\Theta^-\mathbf{R}^{-T}}{2}$ is a skew symmetric matrix. The moment exerted by + strand on the - strand by means of the elastic web, \mathbf{c} , is computed using $\mathbf{c} = \mathbf{Q}\mathbf{H}\mathbf{Q}^T\boldsymbol{\eta}$, where $\mathbf{H} = \text{diag}[H_1, H_2, H_3]$ are the elastic moduli.⁹ Now, we shift our focus to the extension and shear of the web. In the reference configuration, the displacement between the two strands $\mathbf{w}_0 = \frac{\mathbf{r}_0^+ - \mathbf{r}_0^-}{2} = a\mathbf{e}_r$, which, in deformed configuration, changes to $\mathbf{w} = \frac{\mathbf{r}^+ - \mathbf{r}^-}{2}$. The force \mathbf{f} exerted by + strand on the - strand

is computed using $\mathbf{f} = \mathbf{Q}\mathbf{L}(\mathbf{Q}^T\mathbf{w} - \mathbf{Q}_0^T\mathbf{w}_0)$, where $\mathbf{L} = \text{diag}[L_1, L_2, L_3]$ are the elastic moduli, $\beta = \frac{\beta^+ + \beta^-}{2}$ and $\beta^c = \frac{\beta^+ - \beta^-}{2}$.

$$\begin{aligned}
\mathbf{c} &= \mathbf{Q}\mathbf{H}\mathbf{Q}^T\boldsymbol{\eta} \\
&= H_1(-ak_3 - \omega r_x - a\beta_x^\pm)\mathbf{f}_1 + H_2 \frac{(-ak_{3x} - \omega r_{xx} - a\beta_{xx}^\pm)}{\omega} \mathbf{f}_2 \\
&\quad + H_3 \left(\beta^c - \frac{a \cot k}{\omega} \beta_{xx}^c \right) \mathbf{e}_3, \\
\mathbf{f} &= \mathbf{Q}\mathbf{L}(\mathbf{Q}^T\mathbf{w} - \mathbf{Q}_0^T\mathbf{w}_0) \\
&= L_1 r \mathbf{f}_1 + aL_2 \cot k \frac{ak_{3x} + 2\omega r_x + a\beta_{xx}^\pm}{\omega} \mathbf{f}_2 \\
&\quad - a^2 L_3 \frac{\omega^2 \beta^c + \beta_{xx}^c}{\omega} \mathbf{e}_3
\end{aligned} \quad (11)$$

Contributions from the Stacking Energy (\mathbf{l} , \mathbf{h}). We now consider the contributions from the stacking energy. The center line of the double helix \mathbf{e}_3 undergoes both twist k_3 and extension $\xi = \frac{\xi^+ + \xi^-}{2}$. We associate a quadratic stacking energy $E_s = K_c k_3^2 + K_e \left(\frac{\xi^+ + \xi^-}{2} \right)^2$ to penalize this change in the orientation of successive base pairs. Due to this energy, the base pairs exert a body force \mathbf{l} and a body moment \mathbf{h} on both \pm strands, which are given by

$$\mathbf{l} = K_e \left(\frac{\xi^+ + \xi^-}{2} \right) \mathbf{e}_3, \quad \mathbf{h} = K_c k_3 \mathbf{e}_3 \quad (12)$$

RESULTS

Now we have all the ingredients for solving the governing differential equations of a birod. Substituting these quantities in the balance laws (eq 3) gives us a set of 12 differential equations. The complete procedure for solving those equations is in the supplement; however, we highlight crucial points here. It follows from the governing equations that $\beta^+ = \beta^-$ ($= \beta$, say), $n_3^+ = n_1 = n_2 = 0$. $\beta^+ = \beta^-$ implies $\xi^+ = \xi^-$ ($= \xi$, say), thereby reducing 12 equations to 6 equations in 6 unknowns $r, \beta, k_3, n_1^c, n_3$. We look for solutions of the form

$$\begin{aligned}
r(x) &= r_0 e^{-\lambda x}, \quad \beta(x) = \beta_0 e^{-\lambda x}, \quad \xi(x) = \xi_0 e^{-\lambda x}, \\
n_1^c(x) &= n_{10}^c e^{-\lambda x}, \quad n_2^c(x) = n_{20}^c e^{-\lambda x}, \quad n_3(x) = n_{30} e^{-\lambda x}
\end{aligned} \quad (13)$$

We substitute this form into the governing equations (eq 3) and obtain an eigenvalue problem in λ . To make further progress, we need the values of the elastic constants. We use $K_c = 80 \text{ pNnm}^2$, $K_e = 600 \text{ pN}$, $L_1 = L_2 = L_3 = H_1 = H_2 = H_3 = 10 \text{ pN}$. In the supplement, we show that these values yield the correct twist, stretch, and twist-stretch coupling moduli for double-stranded B-DNA.²⁴ Solving for the eigenvalues, λ we get

$$\lambda = \pm \zeta \pm i\omega, \quad \zeta = 0.32 \text{ nm}^{-1} \quad (14)$$

and the solution for the strain parameters $y_1 = r$, $y_2 = k_3$, and $y_3 = \beta$ is of the form

$$\begin{aligned}
y_i(x) &= A_1 \mathbf{V}_1(i) e^{(-\zeta - i\omega)x} + A_2 \mathbf{V}_2(i) e^{(-\zeta + i\omega)x} \\
&\quad + B_1 \mathbf{V}_3(i) e^{(\zeta - i\omega)x} + B_2 \mathbf{V}_4(i) e^{(\zeta + i\omega)x}
\end{aligned} \quad (15)$$

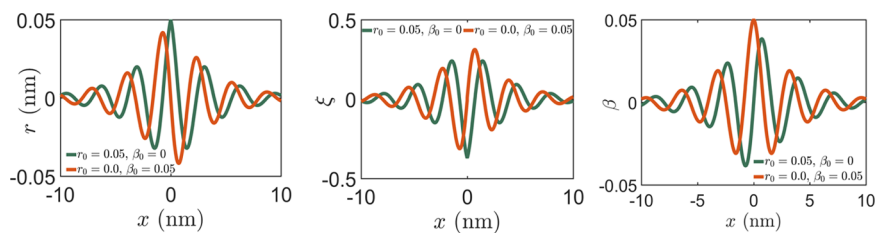


Figure 4. Variation of r , k_3 , ξ , and $\beta^* = \beta^- = \beta$ for a single protein. The red curve corresponds to the boundary conditions $\beta_0 = 0$, $r_0 = 0.05$ nm, and the green curve to $r_0 = 0$, $\beta_0 = 0.05$. Even if $r_0 = 0$, i.e., the protein does not change the radius of the molecule at the binding site, a change in the phase angle $\beta_0 \neq 0$ can cause the radius to change at locations away from the binding site. Similar coupling exists among other strain parameters. The magnitudes r_0 and β_0 chosen here ensure that the change in the groove width, a parameter whose magnitude is known,¹² is in the correct range (3 Å). The decay length is $l_d = \zeta^{-1} \approx 10$ bp, which is close to the one documented in literature.^{7,23}

where $V_j(i)$ is the i th component of the eigenvector corresponding to the eigenvalue j in the exponent. Clearly, the decay length ζ is only a function of the elastic parameters of ds-DNA, in agreement with the conclusion of Kim et al.⁷ Note that the strain parameters are exponentially decaying while oscillating with the period ω of the double helix. We impose the boundary conditions on r and β , remembering that the displacements of the strands must be continuous. For a protein binding at $x = p$

$$\begin{aligned} \text{as } x \rightarrow \pm\infty, \quad r(x), \beta(x) &\rightarrow 0 \\ \text{at } x = p, \quad r(0) = r_0, \quad \beta(0) &= \beta_0 \end{aligned} \quad (16)$$

We present the variation of r , k_3 , and β for a protein binding at $x = 0$ for two different sets of boundary conditions in Figure 4. Notice the sinusoidal correlation between the local deformation of base pairs that is in agreement with earlier work that used Monte Carlo simulations.^{22,23}

We show the deformed shapes of the helices in Figure 5 for three cases: first when one protein binds at $x = 0$, second when two proteins bind at $x = \pm 1.5$ nm, and third when two proteins bind at $x = \pm 3.5$ nm. The boundary condition for each protein is $r_0 = 0.2$ nm, $\beta_0 = 0$. We deliberately choose large values for r_0 and β_0 to distinguish the deformed shape from the reference shape. The large configuration changes near the site of protein

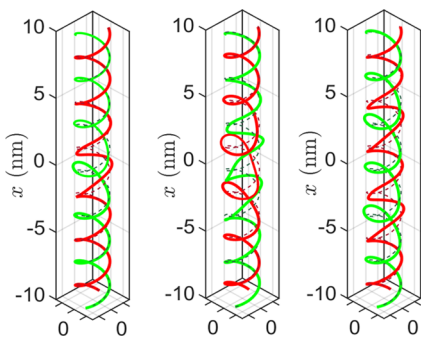


Figure 5. We show the deformed configuration of the double helix, with red and green colors corresponding to + and − strand, respectively. In the first figure, one protein binds at $x = 0$ with $r_0 = 0.2$ nm and $\beta_0 = 0$. The magnitudes are chosen to be artificially large to make the deformations discernible in the figure. In the second figure, two proteins bind at $x = \pm 1.5$ nm. In the third figure, two proteins bind at $x = \pm 3.5$ nm. Notice the overlap of deformations in the second figure, which is absent in the third one. We use eqs 19, 20, and 21 to demonstrate how this overlap leads to an interaction energy between the two proteins. The dotted lines denote the corresponding undeformed configuration.

binding ($x = 0$) decay exponentially with distance. Note, there is a strong overlap in the deformation fields when the distance between two proteins is 3 nm compared to 7 nm. This overlap results in an interaction energy between the two proteins, which we subsequently quantify using eq 18.

We now compute the interaction energy ΔG for two proteins. The energy functional of the double-helical rod is

$$\begin{aligned} E[r, \beta, k_3] = & \frac{1}{2}EI(\kappa^+)^2 + \frac{1}{2}EI(\kappa^-)^2 + \sum_{i=1}^3 \frac{1}{2}(L_i \Delta \mathbf{w}_i^2 \\ & + H_i \hat{\eta}_i^2) + K_c k_3^2 + K_\xi \xi^2 \end{aligned} \quad (17)$$

where $\hat{\eta} = \mathbf{Q}^T \boldsymbol{\eta}$ and $\Delta \mathbf{w} = \mathbf{Q}^T \mathbf{w} - \mathbf{Q}_0^T \mathbf{w}_0$. Consider two proteins, P_1 and P_2 binding at $x = 0$ and $x = p$. The interaction energy ΔG is defined as

$$\Delta G(p) = E_{12}^{(0,p)} - E_1^0 - E_2^p \quad (18)$$

where $E_{12}^{(0,p)} = E[r_{12}, \beta_{12}, (k_3)_{12}]$ is the energy of two proteins binding onto DNA at $x = 0$ and $x = p$, whereas $E_1^0 = E[r_1, \beta_1, (k_3)_1]$ and $E_2^p = E[r_2, \beta_2, (k_3)_2]$ are the energies of a single protein binding at $x = 0$ and $x = p$, respectively. We linearly superimpose the strain fields from each protein (r_1 and r_2 , etc.) to get the resultant strain field (r_{12} , etc.) caused by two proteins simultaneously binding to DNA.

$$r_{12}(x) = r_1(x) + r_2(x - p) \quad (19)$$

We obtain β_{12} and $(k_3)_{12}$ similarly. We compute the interaction energy $\Delta G(p)$ as a function of the distance between two proteins p and plot it in Figure 6 together with experimental data from Kim et al.⁷ In excellent agreement with experiment⁷

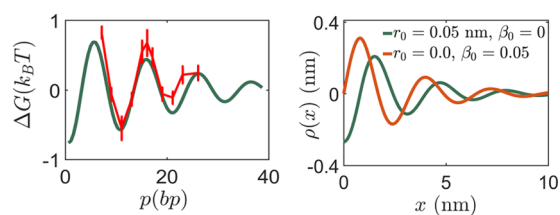


Figure 6. The first figure shows the variation of interaction energy ΔG with distance p between the two proteins P_1 and P_2 . The boundary conditions $r_1 = 0.001$ nm, $\beta_1 = 0.0045$ for P_1 and $r_2 = 0.001$ nm, $\beta_2 = -0.0045$ for P_2 give the best fit to the experimental data for ΔG .⁷ In the second figure, we show the variation of change in groove width $\rho(x) = g(x) - \frac{p}{2}$ when a protein with boundary conditions r_0 , β_0 binds at $x = 0$. The decaying sinusoidal character is documented in previous work.^{7,22} The magnitude of the change in groove width (~ 3 Å) is consistent with estimates in ref 12.

and numerical simulations,²² ΔG decays exponentially while oscillating with the period of the double helix (~ 10 bp).

We justify this variation of interaction energy for a simple case as follows. Consider a strain parameter $\delta(x)$ and the associated quadratic energy potential $\mathcal{E}[\delta(x)] = \int_{-\infty}^{\infty} \frac{\delta^2(x)}{2} dx$. Similar to our strain parameters in eq 15, let us assume $\delta(x) = A e^{-bx} \cos(\mu x)$, then

$$\mathcal{E}[\delta(x)] = \int_{-\infty}^{\infty} \frac{\delta^2(x)}{2} dx = \frac{A^2(2b^2 + \mu^2)}{4b(b^2 + \mu^2)} \quad (20)$$

$\mathcal{E}[\delta(x-p)] = \mathcal{E}[\delta(x)]$. Now, the strain obtained by superposing two strain sources a distance p apart are $\delta_2(x) = \delta(x) + \delta(x-p)$. The energy functional corresponding to $\delta_2(x)$ is

$$\begin{aligned} \mathcal{E}[\delta_2(x)] &= \frac{A^2(2b^2 + \mu^2)}{2b(b^2 + \mu^2)} + A^2 c_1 e^{-bp} \sin(\mu p) \\ &\quad + A^2 c_2 e^{-bp} \cos(\mu p) \\ &= \mathcal{E}[\delta(x)] + \mathcal{E}[\delta(x-p)] + \Delta G \end{aligned} \quad (21)$$

where $c_1 = \frac{b^3}{2b\mu(b^2 + \mu^2)}$ and $c_2 = \frac{\mu(\mu^2 + 2b^2 + pb^3 + pb\mu^2)}{2b\mu(b^2 + \mu^2)}$. It is notable how the decaying sinusoidal behavior of the interaction energy ΔG follows naturally from the functional form of the strain parameters and their eventual superposition. A cartoon illustrating this key point is presented in Figure 7.

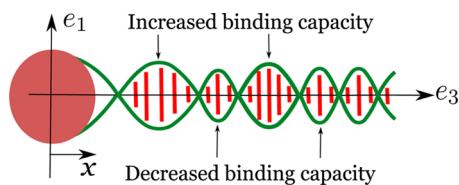


Figure 7. Equation 15 shows that the strain parameters r , β , and ξ decay exponentially while oscillating with the periodicity of the double helix. Let us assume that the protein binding at $x = 0$ increases the radius of the double helix from a to $a + r_0$. This change in radius at $x = 0$ decays exponentially while oscillating with the periodicity of the double helix, away from the binding site. Similar behavior is observed for other strain parameters, β and ξ . Due to this sinusoidal modulation of the geometry, the binding of the second protein is facilitated at some locations, while inhibited at others; this manifests as an exponentially decaying oscillatory behavior observed in the allosteric interaction energy (ΔG).

Next, we focus on the width of the groove because many proteins are known to change the width of the major/minor groove of DNA.^{7,12,25} We define the width of the groove, $g(x)$, as follows (we do not have a major/minor groove because $\alpha = \pi$, for simplicity):

$$g(x) = \mathbf{r}^- \cdot \mathbf{e}_3 \left(x + \frac{\pi}{2\omega} \right) - \mathbf{r}^+ \cdot \mathbf{e}_3 \left(x - \frac{\pi}{2\omega} \right) \quad (22)$$

Note that in the reference configuration, the groove width $g_0 = \frac{\pi}{\omega} = \frac{p}{2}$. We consider a protein binding at $x = 0$ and compute the change in groove width $\rho(x) = g(x) - g_0$ for two sets of boundary conditions, $r_0 = 0$, $\beta_0 = 0.02$ and $r_0 = 0.02$ nm, $\beta_0 = 0$ (see Figure 6). The groove width ρ decays exponentially with increasing distance from the binding site while oscillating with the periodicity of the double helix. This characteristic decaying sinusoidal oscillation is documented in refs 22, 23

and is also observed experimentally.⁷ It has been proposed that this change in groove width could explain the sinusoidally decaying interaction energy (notice the similarity of the two panels in Figure 6) between two proteins bound to DNA because the binding energy of a protein binding to DNA could potentially depend on the groove width.^{22,23} However, we have arrived at the decaying sinusoidal variation of the interaction energy by computing the elastic energy stored in the birod without assuming any connection to the groove width. Thus, we argue that the characteristic variation in groove width and the characteristic variation of the interaction energy have the same underlying cause—the geometry and elasticity of helical DNA at the base-pair level.

To make the above point more concrete, we give another analytical argument. Consider two proteins P1 and P2 binding at $x = 0$ and $x = p$, respectively. Kim et al.⁷ argue that when a protein binds to DNA, it alters the groove width, which leads to $\Delta G \propto \rho_0^2 \rho_p^2 + \rho_p^1 \rho_p^2$, where ρ_p^1 is the change in the groove width caused by protein P1 at $x = p$, and so on. On the other hand, we assume that the protein binds to DNA by fixing the radius r and phase angle β at the binding site. Let us examine whether $\Delta G \propto \alpha_1 r_{12}(p) + \alpha_2 \beta_{12}(p)$, where $r_{12}(p) = r_0^2 + r_p^1 r_p^2$, $\beta_{12}(p) = \beta_0^2 + \beta_p^1 \beta_p^2$ for some constants α_1 and α_2 . Here, r_0^1 is the change in radius caused by protein P1 at $x = 0$. Other quantities (r_0^2 , r_p^1 , r_p^2 , β_0^1 , β_0^2 , β_p^1 , β_p^2) are defined similarly. For simplicity, assume $\alpha_1 = \alpha_2$ and define $q(p) = r_{12}(p) + \beta_{12}(p)$.

We plot $\Delta G_n = \frac{\Delta G}{|\Delta G|_{\max}}$ and $q_n = \frac{q(p)}{|q(p)|_{\max}}$ versus the distance between the two proteins p in Figure 8a. We observe that while the location of peaks and valleys for ΔG_n and $q(p)_n$ coincides, the magnitudes are not identical.

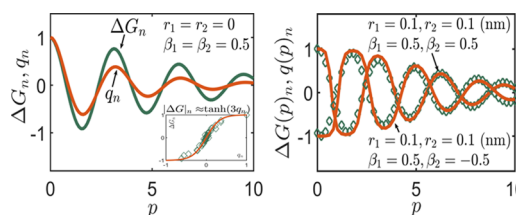


Figure 8. Consider two proteins P1 and P2 binding at $x = 0$ and $x = p$, respectively. The first figure shows the variation of normalized free energy $\Delta G(p)_n = \frac{\Delta G}{|\Delta G|_{\max}}$ and correlation function $q(p)_n = \frac{q(p)}{|q(p)|_{\max}}$ with the distance between the two proteins p . The correlation function $q(p) = (r_0^2 + r_p^1 r_p^2) + (\beta_0^2 + \beta_p^1 \beta_p^2)$ where r_p^1 is the change in radius caused by protein P1 at $x = p$. The boundary conditions for the two proteins are given in the figure. We find that the peaks and valleys of ΔG_n and q_n coincide; however, the magnitudes are not identical. We find that the magnitudes are related as $\Delta G_n \approx \tanh(3q_n)$, as shown in the inset. We test this empirical relation for two different sets of boundary conditions and find a remarkable match. The diamonds denote the free energies computed using eq 18 and the solid line denotes the free energy computed using the normalized correlation function $q(p)_n$. This exercise shows that the correlation functions can be used as a surrogate for free energies.

Our next step is to relate the magnitudes of two quantities ΔG_n and q_n . Assume an empirical relation $\Delta G_n = y(q_n)$. We plot ΔG_n versus q_n in Figure 8 and find that the resultant profile looks akin to $y(x) = \tanh(ax)$, $a \approx 3.0$ gives the best fit. Thus, $\Delta G_n \approx \tanh(3q_n)$. Note that for large values of p ($p > 10$ nm), the correlation function q_n is small, thence $\tanh(3q_n) \approx 3q_n$, and we recover the form similar to that used in Kim et al.⁷

(but with different strain variables) $\Delta G_n \propto q_n$. Note that we used a particular set of boundary conditions to extract the relation $\Delta G_n \approx \tanh(3q_n)$. Now, we test this relation to compute interaction energies for other sets of proteins, which apply different boundary conditions in Figure 8b. We observe a remarkable agreement with the free energies computed using eq 18. Thus, we have shown that the correlation function $q(p)$ can be used as a surrogate for the interaction energy as Kim et al.⁷ did. Evaluating the correlation function $q(p)$ involves measuring displacement variables at two binding sites, which can in turn be related to the free energy using the above scheme.

To conclude, we have demonstrated that the theory of elastic birods can provide useful insights into the allosteric interactions between two proteins binding to a DNA molecule. Our analysis ties together the continuum theory,⁹ experiments,⁷ and numerical simulations.^{6,22} Our computations indicate that the interaction energy (eq 18) for two proteins bound to DNA decays exponentially while oscillating with the period of the DNA double helix. The decay length depends only on the elastic characteristics of the web, whereas the oscillatory behavior is inherited from the underlying double-helical geometry. We have shown that the strong correlation of interaction energy with the changes in groove widths caused by the proteins are rooted in elasticity and geometry of DNA. However, our model suffers from some shortcomings. Existing models^{6,23} rely on numerical simulations, possibly accounting for the stacking energy in a more comprehensive way compared to our approach, which assumes it is quadratic in the twist and stretch of the centerline. Also, these models^{6,23} can account for a wider variety of boundary conditions applied by a protein, owing to more variables describing the DNA structure. Besides, we made various simplifying assumptions such as assuming the outer strands to be inextensible and unshearable. These assumptions could break down close to the binding site where DNA structure might be drastically altered. Some studies¹² shed light on the kinematics of a protein binding to DNA, but an analysis of the deformations at the binding site is beyond the scope of our model. While the results from our model agree with the experimental observations, the outcomes from the existing models in literature such as^{22,23} agree as well. The main strength of our model compared to the existing models is twofold: (i) we account for the mechanics of the outer strands and (ii) our model provides useful insights into the phenomenon at a modest computational cost. Our techniques based on a helical birod model could potentially be applied to other molecules, which have a double helical geometry such as dsRNA, and coiled-coil intermediate filaments.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jpcb.8b07501.

A computation of the elastic constants for ds-DNA using the birod model (ZIP)

A detailed stepwise procedure for the solution to the governing equations (eq 3) (PDF)

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: purohit@seas.upenn.edu.

ORCID

Jaspreet Singh: 0000-0002-7937-0201

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We acknowledge support from NSF through grant number NSF CMMI 1662101 and NIH through grant number NIH R01-HL 135254.

■ REFERENCES

- (1) Ngo, K. X.; Umeki, N.; Kijima, S. T.; Kodera, N.; Ueno, H.; Furutani-Umezumi, N.; Nakajima, J.; Noguchi, T. Q.; Nagasaki, A.; Tokuraku, K.; Uyeda, T. Q. P. Allosteric Regulation by Cooperative Conformational Changes of Actin Filaments Drives Mutually Exclusive Binding with Cofilin and Myosin. *Sci. Rep.* **2016**, 6, No. 35449.
- (2) Mitra, A.; Sept, D. Taxol Allosterically Alters the Dynamics of the Tubulin Dimer and Increases the Flexibility of Microtubules. *Biophys. J.* **2008**, 95, 3252–3258.
- (3) Ousaka, N.; Inai, Y. Transfer of Noncovalent Chiral Information along an Optically Inactive Helical Peptide Chain: Allosteric Control of Asymmetry of the C-terminal Site by External Molecule that Binds to the N-Terminal Site. *J. Phys. Chem. Lett.* **2008**, 74, 1429–1439.
- (4) Ridge, G. S.; Bailly, C.; Graves, D. E.; Waring, M. J. Daunomycin modifies the sequence-selective recognition of DNA by actinomycin. *Nucleic Acids Res.* **1994**, 22, 5241–5246.
- (5) Krugh, T. R.; Young, M. A. Daunorubicin and adriamycin facilitate actinomycin D binding to poly (dA-dT) poly (dA-dT). *Nature* **1977**, 269, 627.
- (6) Dršata, T.; Zgarbova, M.; Spackova, N.; Jurecka, P.; Sponer, J.; Lankas, F. Mechanical Model of DNA Allostery. *J. Phys. Chem. Lett.* **2014**, 5, 3831–3835.
- (7) Kim, S.; Brostromer, E.; Xing, D.; Jin, J.; Chong, S.; Ge, H.; Wang, S.; Gu, C.; Yang, L.; Gao, Y. Q.; et al. Probing Allostery through DNA. *Science* **2013**, 339, 816–819.
- (8) Koslover, E. F.; Spakowitz, A. J. Twist-and Tension-Mediated Elastic Coupling between DNA-Binding Proteins. *Phys. Rev. Lett.* **2009**, 102, No. 178102.
- (9) Moakher, M.; Maddocks, J. H. A Double-Strand Elastic Rod Theory. *Arch. Ration. Mech. Anal.* **2005**, 177, 53–91.
- (10) Wolfe, K. C.; Hastings, W. A.; Dutta, S.; Long, A.; Shapiro, B. A.; Woolf, T. B.; Guthold, M.; Chirikjian, G. S. Multiscale modeling of double-helical DNA and RNA: A unification through lie groups. *J. Phys. Chem. B* **2012**, 116, 8556–8572.
- (11) Kim, S. K.; Nordén, B. Methyl Green: a DNA Major-Groove Binding Drug. *FEBS Lett.* **1993**, 315, 61–64.
- (12) Kopka, M. L.; Yoon, C.; Goodsell, D.; Pjura, P.; Dickerson, R. E. The Molecular Origin of DNA-Drug Specificity in Netropsin and Distamycin. *Proc. Natl. Acad. Sci. USA* **1985**, 82, 1376–1380.
- (13) Dršata, T.; Zgarbova, M.; Jurecka, P.; Sponer, J.; Lankas, F. On the Use of Molecular Dynamics Simulations for Probing Allostery through DNA. *Biophys. J.* **2016**, 110, 874–876.
- (14) Efremov, A. K.; Yan, J. Transfer-Matrix Calculations of the Effects of Tension and Torque Constraints on DNA-Protein Interactions. 2018, arXiv:1802.01437, arXiv.org e-Print archive.
- (15) Wiggins, P. A.; Dame, R. T.; Noom, M. C.; Wuite, G. J. Protein-Mediated Molecular Bridging: a Key Mechanism in Biopolymer Organization. *Biophys. J.* **2009**, 97, 1997–2003.
- (16) Lankaš, F.; Gonzalez, O.; Heffler, L.; Stoll, G.; Moakher, M.; Maddocks, J. H. On the Parameterization of Rigid Base and Basepair Models of DNA from Molecular Dynamics Simulations. *Phys. Chem. Chem. Phys.* **2009**, 11, 10565–10588.
- (17) Rouzina, I.; Bloomfield, V. A. Force-Induced Melting of the DNA Double Helix I. Thermodynamic Analysis. *Biophys. J.* **2001**, 80, 882–893.

- (18) Cocco, S.; Yan, J.; Léger, J.-F.; Chatenay, D.; Marko, J. F. Overstretching and Force-Driven Strand Separation of Double-helix DNA. *Phys. Rev. E* **2004**, *70*, No. 011910.
- (19) Timoshenko, S. P.; Woinowsky-Krieger, S. *Theory of Plates and Shells*; McGraw-hill, 1959.
- (20) Luisi, B. F.; Xu, W.; Otwinowski, Z.; Freedman, L.; Yamamoto, K.; Sigler, P. Crystallographic Analysis of the Interaction of the Glucocorticoid Receptor with DNA. *Nature* **1991**, *352*, 497.
- (21) Newman, M.; Strzelecka, T.; Dorner, L. F.; Schildkraut, I.; Aggarwal, A. K. Structure of Bam HI endonuclease bound to DNA: Partial Folding and Unfolding on DNA Binding. *Science* **1995**, *269*, 656–663.
- (22) Gu, C.; Zhang, J.; Yang, Y. I.; Chen, X.; Ge, H.; Sun, Y.; Su, X.; Yang, L.; Xie, S.; Gao, Y. Q. DNA Structural Correlation in Short and Long Ranges. *J. Phys. Chem. B* **2015**, *119*, 13980–13990.
- (23) Xu, X.; Ge, H.; Gu, C.; Gao, Y. Q.; Wang, S. S.; Thio, B. J. R.; Hynes, J. T.; Xie, X. S.; Cao, J. Modeling Spatial Correlation of DNA Deformation: DNA Allostery in Protein Binding. *J. Phys. Chem. B* **2013**, *117*, 13378–13387.
- (24) Singh, J.; Purohit, P. K. Structural Transitions in Torsionally Constrained DNA and their Dependence on Solution Electrostatics. *Acta Biomater.* **2017**, *55*, 214–225.
- (25) Hancock, S. P.; Ghane, T.; Cascio, D.; Rohs, R.; di Felice, R.; Johnson, R. C. Control of DNA Minor Groove Width and Fis Protein Binding by the Purine 2-amino Group. *Nucleic Acids Res.* **2013**, *41*, 6750–6760.