PROCEEDINGS A

rspa.royalsocietypublishing.org

Research

Cite this article: Singh J, Purohit PK. 2018 Allosteric interactions in a birod model of DNA. *Proc. R. Soc. A* **474**: 20180136.
http://dx.doi.org/10.1098/rspa.2018.0136

Received: 26 February 2018
Accepted: 5 September 2018

Subject Areas:
mechanics, biomechanics, biophysics

Keywords:
allostery, birod, DNA

Author for correspondence:
Prashant K. Purohit
e-mail: purohit@seas.upenn.edu

Electronic supplementary material is available online at https://dx.doi.org/10.6084/m9.figshare.c.4238600.

THE ROYAL SOCIETY PUBLISHING

Allosteric interactions in a birod model of DNA

Jaspreet Singh and Prashant K. Purohit

Department of Mechanical Engineering and Applied Mechanics, University of Pennsylvania, Philadelphia, PA 19104, USA

JS, 0000-0002-7937-0201

Allosteric interactions between molecules bound to DNA at distant locations have been known for a long time. The phenomenon has been studied via experiments and numerical simulations, but a comprehensive understanding grounded in a theory of DNA elasticity remains a challenge. Here, we quantify allosteric interactions between two entities bound to DNA by using the theory of birods. We recognize that molecules bound to DNA cause local deformations that can be captured in a birod model which consists of two elastic strands interacting via an elastic web representing the basepairs. We show that the displacement field caused by bound entities decays exponentially with distance from the binding site. We compute the interaction energy between two proteins on DNA as a function of distance between them and find that it decays exponentially while oscillating with the periodicity of the double helix, in excellent agreement with experiments. The decay length of the interaction energy can be determined in terms of the mechanical properties of the strands and the webbing in our birod model, and it varies with the GC content of the DNA. Our model provides a framework for viewing allosteric interactions in DNA within the ambit of configurational forces of continuum elasticity.

1. Introduction

Configurational forces that describe the interaction between defects in an elastic solid are those that depend explicitly on the positions of the defects [1,2]. For example, two parallel screw dislocations at a distance a from each other interact with a configurational force per unit length proportional to $1/a$ or an energy per unit length proportional to $\log a$ [3]. Similarly, the

interaction energy of a point defect located at distance a from an edge dislocation varies as $1/a$. Just as defects produce local elastic fields in a solid, proteins binding to DNA also deform it locally. Since DNA behaves like an elastic rod at scales of a few tens of nanometres [4], we expect that if two proteins bind to DNA separated by a distance a, then the deformation fields created by them will overlap and lead to an interaction energy which depends on a in a clearly quantifiable way. There is compelling experimental evidence for such an interaction, some of which has been extracted by connecting the interaction energy with the kinetics of protein binding/unbinding. In spirit, this is similar to continuum elasticity in which configurational forces often determine defect dynamics through a kinetic law [1,2]. Kim *et al.* [5] have exploited this connection of interaction energies to kinetics to show that gene expression, which depends on RNA polymerase binding affinity to DNA in live bacteria, is a function of the proximity of LacR and T7 RNA polymerase bound to DNA. Similarly, the IHF protein affects RNA polymerase activity in *Escherichia coli* DNA [6]. Again, the binding of the drug distamycin to calf thymus DNA has been shown to be cooperative, i.e. if one drug molecule binds to the DNA, then it becomes energetically favourable for other drug molecules to bind [7]. Similarly, binding of the Hox transcription factor to DNA contributes nearly $1.5\,\mathrm{kcal\,mol^{-1}}$ to binding the Exd transcription factor [8]. These effects are called *allosteric* interactions on DNA. Our goal in this paper is to quantify interaction energies between proteins binding to DNA as a function of the distance a separating them and the boundary conditions imposed by the proteins on the DNA. We will apply our methods to the quantitative experimental results of Kim *et al.* [5], who measure allosteric effects on gene expression as well as transcription factor affinity to DNA.

In the experiments of Kim *et al.* [5] one end of a DNA molecule is attached to the passivated surface of a flow cell and binding sites are provided for two specific proteins to bind. The length of the DNA between these binding sites, a, is increased in 1 basepair (bp) increments between 7 bp and 45 bp. First, one type of fluorescently labelled protein (call it A) is flowed into the cell so that it binds to the DNA. Then, the second protein (call it B) is flowed in at a specific concentration. The dissociation times of the fluorescent protein are then monitored as a function of a. This dissociation time depends on the free energy change ΔG of the DNA + two protein complexes from the state when the two proteins are bound to that when protein A is unbound. Now, in general, the free energy ΔG of the ternary complex formed by the DNA and proteins A and B consists of three parts [5],

$$\Delta G = \Delta G_A + \Delta G_B + \Delta\Delta G_{AB}(a), \tag{1.1}$$

where ΔG_A and ΔG_B are the free energy changes caused by binding of A and B, respectively, to the DNA. These are constants. The last term $\Delta\Delta G_{AB}(a)$ is the portion of the free energy change that accounts for the interaction of the two proteins bound to the DNA while being separated by a distance a. The off-rate of A, which is affected by this term, is plotted as a function of a in Kim *et al.* [5], and it is found that it oscillates with a period of 10–11 bp with the amplitude of oscillation decreasing as a function of a. Similar curves for a free energy as a function of separation between protein binding sites on DNA have been obtained experimentally for the binding of the *lac* repressor to DNA [9,10]. It has been shown that these free energy profiles can be reproduced by modelling DNA as an elastic rod which is forced into forming a loop due to stereo-specific binding of the *lac* repressor monomers [11], which come together due to thermal fluctuations. However, Kim *et al.* [5] have ruled out DNA loop formation by careful experimental design and choice of DNA-binding proteins. They have also found that the form of the curve is independent of ionic strength (ruling out electrostatic interactions between A and B), but dependent on modifications of the linker DNA. Kim *et al.* [5] infer that this implies $\Delta\Delta G_{AB}(a)$ largely depends on DNA mechanical properties. However, as yet there is no analytical description of how the interaction energy $\Delta\Delta G_{AB}(a)$ depends on the DNA mechanical properties.

Allosteric effects and their relation to protein–DNA interactions have been studied using molecular dynamic (MD) simulations [12,13]. Gu *et al.* [12] have studied various kinds of deformations which include shift, roll, rise, twist, slide and tilt of the DNA bases. They observed

a sinusoidal correlation in the major groove widths similar to the one observed by Kim *et al.* [5]. Furthermore, Gu *et al.* point out that the presence of GC-rich sequences dampens the allosteric effects, which is what Kim *et al.* observe experimentally. Major groove widths have also been implicated in the MD simulations of Hancock *et al.* [13], who show how bound proteins alter this quantity. By contrast, our approach in this paper is based on elastic energy considerations and could complement the analysis of major groove widths as an indicator of allostery in DNA.

Our goal in this paper is to quantitatively describe allosteric interactions using the *birod* model of DNA of Moakher & Maddocks [14], who originally derived it to study DNA melting. This birod model is a double-stranded rod theory in which addition to the standard variables of a Cosserat rod theory (i.e. the centreline of the rod cross section $\mathbf{r}(s,t)$ and a material frame $[\mathbf{d}_1(s,t)\ \mathbf{d}_2(s,t)\ \mathbf{d}_3(s,t)]$), there are two micro-structural variables—$\mathbf{w}(s,t)$, a micro-displacement measuring the change in distance between the two strands, and $\mathbf{P}(s,t)$, a micro-rotation measuring the change in orientation of one strand relative to the other. Fortunately, the forces conjugate to these micro-structural variables obey balance laws that look similar to the balance of forces and moments equations of a standard Cosserat rod. They are coupled to the macroscopic balance equations for the centreline of the rod through distributed body forces and moments. Moakher & Maddocks [14] have provided hyper-elastic constitutive laws for these micro-structural variables that are based on *quadratic* energies.

The theory of birods has been successfully used by Lessinnes and co-workers [15,16] to study the growth and evolution of two filaments elastically bound to each other. These authors have demonstrated the utility of the theory in accurately modelling biological structures across multiple length scales—from tissues and arteries to growth of roots and stems in plants. Manning *et al.* [17] have used both a discrete basepair model and a corresponding continuum rod model to study the cyclization of short DNA molecules (150 bp). The results obtained from both these approaches match remarkably well. However, for shorter length scales (approx. 16 bp) Lankaš *et al.* [18] have assayed the merit of the assumptions of rigid bases versus rigid basepairs to estimate the stiffness parameters of a DNA oligomer and found that the simulated data are closely consistent with the assumption of rigid bases, but not rigid basepairs. Following this line of thought Petkeviciute [19] constructed a parameter set to model sequence-dependent equilibrium probability distribution for rigid base configurations of a DNA oligomer. These developments necessitate the inclusion of elasticity of basepairs via the webbing in an elastic birod [14] to accurately model the local deformations caused by proteins at small length scales.

An effort to account for the elasticity of basepairs has been made by Dršata *et al.* [20], who have applied it to the DNA–protein system investigated by Kim *et al.* [5]. The model by Dršata *et al.* does not require any input elastic constants but uses parameters (elasticity in terms of DNA helical variables) extracted from MD simulations of free DNA. With this model, they explained and interpreted the data by Kim *et al.* (Dršata *et al.* [21]), including the periodicity and decay of the allosteric coupling between separate protein binding sites along DNA. In the work of Dršata *et al.* [20], the structure of a DNA molecule is described using three sets of coordinates: intra-basepair coordinates—buckle, propeller, opening, shear, stretch and stagger; inter-basepair coordinates—tilt, roll, twist, shift, slide and rise; and major and minor groove widths. Any conformational perturbations to the structure of DNA caused by a binding ligand are penalized using an energy functional quadratic in the aforementioned coordinates. In this framework, the binding ligand interacts with the DNA molecule by altering the various degrees of freedom (inter-basepair, intra-basepair and major/minor grooves).

In contrast to Dršata *et al.* [20], we use a theory of elastic birods [14] to study the problem of protein–DNA interactions. The phosphate backbones are represented by the outer strands of the birod, while the basepairs are represented using an elastic web. We allow basepairs to bend, twist, shear and stretch elastically. The outer strands, on the other hand, are assumed to be inextensible rods similar to the worm-like chains of polymer elasticity [22]. The binding ligand causes a change in the geometry of the DNA double helix by bending its centreline and changing its radius.

2. Strategy to compute interaction energy

In this section, we give a concise blueprint of our strategy to compute the interaction energy for two proteins binding to DNA. We assume elastic deformations throughout. When a protein binds to DNA, it causes local bending and twisting. We assume that the resulting twist and curvatures are small. These curvatures could possibly add up to produce large displacements and rotations. The two phosphate backbones of DNA constitute the helical outer strands, which are out of phase by a phase angle $\alpha = 2.1$ radians. We assume these backbones to be inextensible. These outer strands consist of sugar–phosphate single bonds. Thus, we assume that they cannot support twisting moments. The inextensibility of the outer strands is a strong geometrical constraint which induces a change in the radius and phase angle between the two helices when a protein causes local deformations. We assume that these changes are small and of the same order as the curvatures.

We give a stepwise procedure to do the calculation and in the following sections we label each step. We give a pictorial representation for this procedure in the electronic supplementary material, figure S1.

(i) We begin by assuming a form of displacement for each of the outer strands, which are assumed to be inextensible and unshearable.
(ii) We then use this displacement to calculate the tangent, normal and binormal to the deformed configuration of the outer strands, thereby obtaining the rotation matrix attached to the deformed configuration of the outer strands.
(iii) Once we get the deformation and rotation of the outer strands, we use these to calculate the extension, shear and rotation of the web.
(iv) At this point, we are in a position to substitute these quantities into the balance laws for the birod. We, then, seek non-zero solutions to the resulting system of differential equations. This leads to an eigenvalue problem.
(v) In the next step, we apply the boundary conditions to evaluate the constants in the solution to the eigenvalue problem.
(vi) We carry out this process first when there is a single protein binding onto the DNA, and second when there are two proteins binding.
(vii) Finally, we subtract the two energies obtained in the previous step to get our energy of interaction. We find that it takes the form of a decaying exponential oscillating with the periodicity of the underlying DNA helix.

We demonstrate the above procedure for a straight non-helical birod (called a ladder) in the electronic supplementary material, §S1. Our calculations show that the interaction energy for two defects on a ladder decays exponentially with the distance between them. In the next section, we focus on how the geometry of the DNA helix leads to an interaction energy which decays exponentially while oscillating with the periodicity of the helix.

3. Interaction energy for two DNA-binding proteins

(a) Step 1: deformation of the outer strands

DNA consists of two helical strands with radius $b = 1$ nm and pitch $p = 3.4$ nm, out of phase by $\alpha = 2.1$ radians, wrapped around a common axis as shown in figure 1. We follow the notation used by Moakher & Maddocks [14] and refer to the two strands as $\pm$. The undeformed state of the outer strands denoted by $\mathbf{r}_0^{\pm}(x)$ is a helix with a constant radius and pitch. We choose to parametrize both the curves by arclength parameter x. Here, $\omega = 2\pi/p$ and k is the characteristic angle of the helix such that $\tan k = 2\pi b/p = \omega b$,

$$\mathbf{r}_0^{+} = b(\cos \omega x\, \mathbf{e}_1 + \sin \omega x\, \mathbf{e}_2) + x\, \mathbf{e}_3$$

and

$$\mathbf{r}_0^{-} = b(\cos(\omega x + \alpha)\, \mathbf{e}_1 + \sin(\omega x + \alpha)\, \mathbf{e}_2) + x\, \mathbf{e}_3.$$

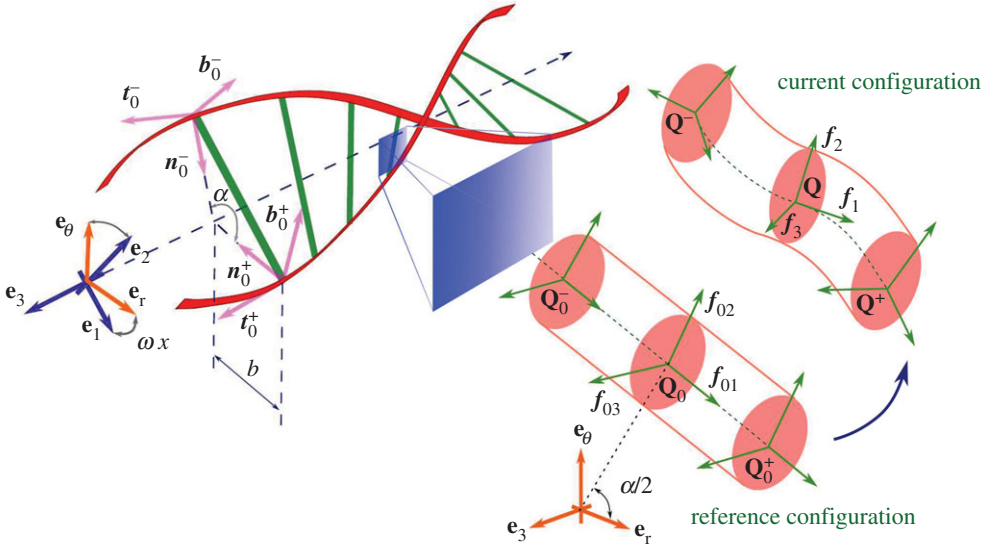

Figure 1. A DNA molecule as a double-helical elastic birod is shown on the left. The phosphate backbones are represented by outer strands while the complementary basepairing is represented by the elastic web. The phase angle between the two helices is $\alpha = 2.1$ radians. Here $\mathbf{R}^+ = [\boldsymbol{n}_0^+\ \boldsymbol{b}_0^+\ \boldsymbol{t}_0^+]$ and $\mathbf{R}^- = [\boldsymbol{n}_0^-\ \boldsymbol{b}_0^-\ \boldsymbol{t}_0^-]$ are the Frenet–Serret frames attached to the + and − strands, respectively. Basepairs in the reference and current configuration are shown on the right. $\mathbf{Q}_0^+ = \mathbf{Q}_0^- = \mathbf{Q}_0$ in the reference configuration. In the current configuration, the rigid rotation of the basepair is quantified by $\mathbf{Q} = \mathbf{Z}(\mathbf{1} + \boldsymbol{\Phi})\mathbf{Q}_0$ (equation (3.22)) and the elastic moment $\boldsymbol{c}$ is related to the Gibbs rotation vector of $\mathbf{P} = (\mathbf{Q}^+\mathbf{Q}^{-\mathrm{T}})^{1/2}$ (equation (3.18)). (Online version in colour.)

Let us now focus on the two strands separately. The calculations for the + strand are given in this section while the results for the − strand are given in the electronic supplementary material. We posit a form of displacement wherein the radius of the helix changes and its axis is allowed to take arbitrary shapes within the ambit of the assumptions specified in §2. Here $[\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3]$ denotes the standard spatial reference frame and $\mathbf{e}_3$ is along the common axis of the two helices ± in the *reference* configuration. This common axis in the *deformed* configuration is defined by the set of orthogonal directors $[\mathbf{d}_1(x), \mathbf{d}_2(x), \mathbf{d}_3(x)]$. The displacement fields which define the undeformed and deformed configurations are

$$\left.\begin{aligned} \mathbf{r}_0^+ &= b(\cos\omega x\ \mathbf{e}_1 + \sin\omega x\ \mathbf{e}_2) + x\ \mathbf{e}_3 \\ \text{and}\qquad \mathbf{r}^+(x) &= (b+r)(\cos(\omega x + \beta^+)\ \mathbf{d}_1 + \sin(\omega x + \beta^+)\ \mathbf{d}_2) + \int_0^x \mathrm{d}x(1 + b\xi)\ \mathbf{d}_3, \end{aligned}\right\} \tag{3.1}$$

where

$$r = r(x), \quad \beta^+ = \beta^+(x), \quad \xi = \xi(x).$$

Here r is the change in the radius of the helix, β^+ is the change in the phase of the + strand and ξ can be considered as a stretching of the axis of the helix.

Let $\mathbf{Z}$ be a second-order orthogonal tensor which relates the directors of the deformed centreline $\mathbf{d}_i$ to those of the undeformed one $\mathbf{e}_i$, $i = 1, 2, 3$. As stated in §2, the curvatures (k_1, k_2, k_3) associated with the deformation of the centreline are assumed to be small, nonetheless these could aggregate to potentially produce large rotations. The orthogonal tensor $\mathbf{Z}$ operates as follows:

$$\mathbf{d}_i = \mathbf{Z}\ \mathbf{e}_i, \quad \mathbf{Z} = \sum_{i=1}^{3} \mathbf{d}_i \otimes \mathbf{e}_i, \quad i = 1, 2, 3 \tag{3.2}$$

and

$$\left.\begin{aligned} &\mathbf{d}_{ix} = \boldsymbol{\kappa} \times \mathbf{d}_i, \quad \text{where } \boldsymbol{\kappa} = k_1\,\mathbf{d}_1 + k_2\,\mathbf{d}_2 + k_3\,\mathbf{d}_3 \\ \text{and} \quad &\mathbf{d}_{1x} = k_3\,\mathbf{d}_2 - k_2\,\mathbf{d}_3, \quad \mathbf{d}_{2x} = k_1\,\mathbf{d}_3 - k_3\,\mathbf{d}_2, \quad \mathbf{d}_{3x} = k_2\,\mathbf{d}_1 - k_1\,\mathbf{d}_2. \end{aligned}\right\} \tag{3.3}$$

In the above equations, we assume that

$$r^+(x), k_1(x), k_2(x), k_3(x), \zeta(x), \beta^+(x) \sim O(\varepsilon).$$

Thus, in the treatment henceforth, any product terms such as r^2 or ξk_3 are $O(\varepsilon^2)$ and are neglected.

(b) Step 2: rotation of strands

We proceed in a standard way by attaching a Frenet–Serret director frame consisting of normal, binormal and tangent to each cross section of the strand as shown in figure 1. We denote it by $\mathbf{R}_0^+(x)$ in the reference configuration,

$$\left.\begin{aligned} &\mathbf{R}_0^+ = [\boldsymbol{n}_0^+ \quad \boldsymbol{b}_0^+ \quad \boldsymbol{t}^+], \\ &\boldsymbol{n}_0^+ = -\cos\omega x\,\mathbf{e}_1 - \sin\omega x\,\mathbf{e}_2, \\ &\boldsymbol{b}_0^+ = -\cos k(-\sin\omega x\,\mathbf{e}_1 + \cos\omega x\,\mathbf{e}_2) + \sin k\,\mathbf{e}_3 \\ \text{and} \quad &\boldsymbol{t}_0^+ = \sin k(-\sin\omega x\,\mathbf{e}_1 + \cos\omega x\,\mathbf{e}_2) + \cos k\,\mathbf{e}_3. \end{aligned}\right\} \tag{3.4}$$

For the sake of brevity, we use

$$(\cos\omega x\,\mathbf{d}_1 + \sin\omega x\,\mathbf{d}_2) = \mathbf{f}_1^+, \quad (-\sin\omega x\,\mathbf{d}_1 + \cos\omega x\,\mathbf{d}_2) = \mathbf{f}_2^+, \quad \mathbf{d}_3 = \mathbf{f}_3^+.$$

As the strand deforms, the frame $\mathbf{R}_0^+$ evolves into $\mathbf{R}^+(x)$, which consists of normal, binormal and tangent to the deformed configuration of the strand. Our next step is to calculate the tangent vector to the deformed configuration. We differentiate equation (3.1) to obtain

$$\begin{aligned} \mathbf{r}_x^+ = {} & (r_x - b\omega\beta^+)(\cos\omega x\,\mathbf{d}_1 + \sin\omega x\,\mathbf{d}_2) + (b\omega + \omega r + b\beta_x^+ + bk_3)(-\sin\omega x\,\mathbf{d}_1 + \cos\omega x\,\mathbf{d}_2) \\ & + (1 + b\xi - bk_2\cos\omega x + bk_1\sin\omega x)\,\mathbf{d}_3. \end{aligned} \tag{3.5}$$

We assume the strand to be inextensible and unshearable. This means

$$\begin{aligned} |\mathbf{r}_x^+|^2 = {} & 1 + \omega^2 b^2 + 2b(\omega^2 r + b\omega\beta_x^+ + b\omega k_3 + \xi - k_2\cos\omega x + k_1\sin\omega x) \\ & + O(\varepsilon^2) = |\mathbf{r}_{0x}^+|^2 = 1 + b^2\omega^2, \end{aligned}$$

which leads us to the inextensibility condition,

$$\xi - k_2\cos\omega x + k_1\sin\omega x = -\omega^2 r - b\omega(k_3 + \beta_x^+). \tag{3.6}$$

We will subsequently use this equation to impose boundary conditions. We substitute equation (3.6) into equation (3.5) to get

$$\mathbf{r}_x^+ = (r_x - b\omega\beta^+)\,\mathbf{f}_1^+ + (b\omega + \omega r + b\beta_x^+ + bk_3)\,\mathbf{f}_2^+ + (1 - b\omega^2 r - b^2\omega(k_3 + \beta_x^+))\,\mathbf{f}_3^+. \tag{3.7}$$

Now, we need to find the director frame for the strand in the deformed configuration. We start by calculating the tangent vector

$$\begin{aligned} \boldsymbol{t}^+ = {} & \frac{\mathbf{r}_x^+}{|\mathbf{r}_{0x}^+|} \\ = {} & (r_x\cos k - \beta^+\sin k)\,\mathbf{f}_1^+ + (\sin k + \omega r\cos k + b(\beta_x^+ + k_3)\cos k)\,\mathbf{f}_2^+ \\ & + (\cos k - \omega r\sin k - b(\beta_x^+ + k_3)\sin k)\,\mathbf{f}_3^+ \\ = {} & \mathbf{Z}(\boldsymbol{t}_0^+ - (r_x\cos k - \beta^+\sin k)\boldsymbol{n}_0^+ - (\omega r + b(\beta_x^+ + k_3))\boldsymbol{b}_0^+). \end{aligned} \tag{3.8}$$

We differentiate the tangent vector to calculate the normal in the deformed configuration

$$\begin{aligned}\boldsymbol{t}_x^+ = &(-\omega\sin k + (r_{xx}+\xi)\cos k - (\beta_x + k_3)\sin k)\,\mathbf{f}_1^+ \\ &+ (2\omega\cos k r_x + -\omega\beta^+\sin k + b\cos k(\beta_{xx}^+ + k_{3x}) - f\cos k)\,\mathbf{f}_2^+ \\ &+ (f - \omega r_x - b(\beta_{xx}^+ + k_{3x}))\sin k\,\mathbf{f}_3^+ + O(\varepsilon^2).\end{aligned} \tag{3.9}$$

We can use the above expression to calculate the curvature Ω^+ for the strand. We find that this is equal to the sum of the original curvature ($\omega\sin k$) and the one induced by the process of deformation κ^+. Hence,

$$\left.\begin{aligned}&\Omega^+ = (\boldsymbol{t}_x^+ . \boldsymbol{t}_x^+)^{1/2} = \omega\sin k - (r_{xx}+\xi)\cos k + (\beta_x^+ + k_3)\sin k \\ \text{and}\quad &\kappa^+ = \Omega^+ - \omega\sin k = -(r_{xx}+\xi)\cos k + (\beta_x^+ + k_3)\sin k.\end{aligned}\right\} \tag{3.10}$$

The bending moment $\boldsymbol{m}^+$ in the strand is proportional to κ^+,

$$\boldsymbol{m}^+ = EI\kappa^+\boldsymbol{b}^+ = EI\kappa^+(-\cos k\,\mathbf{f}_2^+ + \sin k\,\mathbf{f}_3^+) \tag{3.11}$$

Also, the normal is

$$\begin{aligned}\boldsymbol{n}^+ &= \frac{1}{\Omega^+}\boldsymbol{t}^+ = -\mathbf{f}_1^+ + \frac{1}{\sin k}(r_x\sin k - \beta^+\sin k)\,\mathbf{f}_2^+ + \frac{f - \omega r_x - b(\beta_{xx}^+ + k_{3x})}{\omega\sin k}(-\cos k\,\mathbf{f}_2 + \sin k\,\mathbf{f}_3^+) \\ &= \mathbf{Z}\left(\boldsymbol{n}_0^+ + (r_x\cos k - \beta^+\sin k)\boldsymbol{t}_0^+ + \left(-\frac{(r_x\cos k - \beta^+\sin k)\cos k}{\sin k} + \frac{g}{\omega\sin k}\right)\boldsymbol{b}_0^+\right),\end{aligned} \tag{3.12}$$

where

$$g(x) = f(x) - \omega r_x - b(\beta_{xx}^+ + k_{3x}), \quad f(x) = k_1\cos\omega x + k_2\sin\omega x.$$

Using the above deformed orthogonal frame attached to each cross section

$$\mathbf{R}^+ = [\boldsymbol{n}^+ \quad \boldsymbol{b}^+ \quad \boldsymbol{t}^+] = \mathbf{Z}\mathbf{R}_0^+(\mathbf{1} + \boldsymbol{\Theta}^+), \tag{3.13}$$

where $\boldsymbol{\Theta}^+$ is a skew symmetric tensor and $\mathbf{Z} = \sum_{i=1}^{3}\mathbf{d}_i \times \mathbf{e}_i$ as defined in equation (3.2)

$$\left.\begin{aligned}&\boldsymbol{\Theta}^+ = \begin{bmatrix} 0 & -\theta_3^+ & \theta_2^+ \\ \theta_3^+ & 0 & -\theta_1^+ \\ -\theta_2^+ & \theta_1^+ & 0 \end{bmatrix}, \\ \text{in which}\quad &\theta_1^+ = r\omega + b(\beta_x^+ + k_3), \quad \theta_2^+ = -r_x\cos k + \beta^+\sin k, \\ &\theta_3^+ = \frac{g}{\omega\sin k} - \frac{(r_x\cos k - \beta^+\sin k)\cos k}{\sin k}.\end{aligned}\right\} \tag{3.14}$$

We can derive all the above quantities $\mathbf{r}^-$, $\mathbf{R}^-$ and κ^-, etc., for the $-$ strand too. We give the relevant expressions for these quantities in the electronic supplementary material.

(c) Step 3: mechanics of basepairing

The sugar–phosphate backbones of the DNA molecule are tied together by means of complementary basepairing. We model the basepairing by elastic rods capable of extension, shear, bending and twisting. We attach the orthogonal frame $\mathbf{Q}_0 = [\mathbf{f}_{01}\ \mathbf{f}_{02}\ \mathbf{f}_{03}]$ to the strands such that $\mathbf{f}_{01}$ is a unit vector pointing from the $-$ strand to the $+$ strand in the reference configuration as shown in figure 1. Thus,

$$\left.\begin{aligned}&\mathbf{Q_0} = [\mathbf{f}_{01} \quad \mathbf{f}_{02} \quad \mathbf{f}_{03}], \\ &\mathbf{f}_{01} = \sin\left(\omega x + \frac{\alpha}{2}\right)\mathbf{e}_1 - \cos\left(\omega x + \frac{\alpha}{2}\right)\mathbf{e}_2 \\ \text{and}\quad &\mathbf{f}_{02} = \cos\left(\omega x + \frac{\alpha}{2}\right)\mathbf{e}_1 + \sin\left(\omega x + \frac{\alpha}{2}\right)\mathbf{e}_2, \quad \mathbf{f}_{03} = \mathbf{e}_3.\end{aligned}\right\} \tag{3.15}$$

We denote the two ends of the rod in the web as $\pm$ such that the $+$ end lies on the $+$ strand and the $-$ end lies on the $-$ strand. The deformation of the web is completely determined

by the displacement $(\mathbf{r}^+(x), \mathbf{r}^-(x))$ and rotation $(\mathbf{R}^+(x), \mathbf{R}^-(x))$ of its ends. As the outer strands undergo the deformation prescribed by equation (3.1), the web itself undergoes various kinds of deformation. We describe the rotation of the web via a rigid rotation and a micro-rotation [14]. The micro-rotation encapsulates the information about the difference in rotation of the two ends of the web. We calculate the mechanical quantities associated with the extension and bending of the web in two separate sections below.

(i) Bending and twisting of the web

Our objective in this section is to calculate the micro-rotation tensor $\mathbf{P}$. We attach a copy of $\mathbf{Q}_0$, say $\mathbf{Q}_0^{\pm}$, on the + and − end of every spoke in the reference configuration. $\mathbf{Q}_0^{\pm}$ change to $\mathbf{Q}^{\pm}$ in the current configuration. The 'difference' between $\mathbf{Q}_0^+$ and $\mathbf{Q}_0^-$ gives the bending and torsion of the web while the 'average' of $\mathbf{Q}_0^+$ and $\mathbf{Q}_0^-$ gives the rigid rotation of the web. We relate $\mathbf{Q}^{\pm}$ to the rotations of ± strands $\mathbf{R}(x)^{\pm}$. The angles between the columns of $\mathbf{Q}_0^+$ and $\mathbf{R}_0^+$ should remain the same during the deformation, which translates into the following condition:

$$\left.\begin{aligned} &\mathbf{R}_0^{+\mathrm{T}}\mathbf{Q}_0 = \mathbf{R}^{+\mathrm{T}}\mathbf{Q}^+, \\ &\mathbf{Q}^+ = \mathbf{R}^+\mathbf{R}_0^{+\mathrm{T}}\mathbf{Q}_0 = \mathbf{Z}\mathbf{R}_0^+(\mathbf{1}+\boldsymbol{\Theta}^+)\mathbf{R}_0^{+\mathrm{T}}\mathbf{Q}_0 \\ \text{and}\quad &\mathbf{Q}^- = \mathbf{Z}\mathbf{R}_0^-(\mathbf{1}+\boldsymbol{\Theta}^-)\mathbf{R}_0^{-\mathrm{T}}\mathbf{Q}_0. \end{aligned}\right\} \tag{3.16}$$

We are now in a position to calculate the micro-rotation $\mathbf{P}$ responsible for generating elastic moment in the web. Let the micro-rotation tensor in the reference configuration be $\mathbf{P}_0$, which changes to $\mathbf{P}$ during deformation. We use an expression for $\mathbf{P}/\mathbf{P}_0$ given in Moakher & Maddocks [14],

$$\left.\begin{aligned} &\mathbf{P}_0^2 = \mathbf{Q}_0^+\mathbf{Q}_0^{-\mathrm{T}} = \mathbf{I} \\ \text{and}\quad &\mathbf{P}^2 = \mathbf{Q}^+\mathbf{Q}^{-\mathrm{T}} = \mathbf{Z}\mathbf{R}_0^+(\mathbf{1}+\boldsymbol{\Theta}^+)\mathbf{R}_0^{+\mathrm{T}}\mathbf{Q}_0\mathbf{Q}_0^{\mathrm{T}}\mathbf{R}^-(\mathbf{1}-\boldsymbol{\Theta}^-)\mathbf{R}_0^{-\mathrm{T}}\mathbf{Z}^{\mathrm{T}} \\ &\quad = \mathbf{Z}(\mathbf{1}+\mathbf{R}^+\boldsymbol{\Theta}^+\mathbf{R}^{+\mathrm{T}} - \mathbf{R}^-\boldsymbol{\Theta}^-\mathbf{R}^{-\mathrm{T}})\mathbf{Z}^{\mathrm{T}}. \end{aligned}\right\} \tag{3.17}$$

This gives

$$\mathbf{P}_0 = \mathbf{I}, \quad \mathbf{P} \approx \mathbf{Z}\left(\mathbf{1} + \frac{\mathbf{R}^+\boldsymbol{\Theta}^+\mathbf{R}^{+\mathrm{T}} - \mathbf{R}^-\boldsymbol{\Theta}^-\mathbf{R}^{-\mathrm{T}}}{2}\right)\mathbf{Z}^{\mathrm{T}} = \mathbf{Z}(\mathbf{1}+\boldsymbol{\Phi}^{\boldsymbol{c}})\mathbf{Z}^{\mathrm{T}}. \tag{3.18}$$

Note that $\boldsymbol{\Phi}^{\boldsymbol{c}}$ is a skew symmetric tensor. The next step is to calculate the Gibbs rotation vector of $\mathbf{P}$ [14]. The Gibbs rotation vector $\bar{t}$ of a rotation matrix $\mathbf{T}$ is defined as $\bar{t} = \tan(\theta/2)\mathbf{k}$ such that $\mathrm{tr}\,\mathbf{P} = 1 + 2\cos\theta$ and $\mathbf{k}$ is a unit vector such that $\mathbf{T}\mathbf{k} = \mathbf{k}$. Consider $\bar{\mathbf{P}} = \mathbf{1} + \boldsymbol{\Phi}^{\boldsymbol{c}}$ where $\boldsymbol{\Phi}^{\boldsymbol{c}} \sim O(\varepsilon)$. The axis of the infinitesimal rotation $\bar{\mathbf{P}}$ is the axial vector of $\boldsymbol{\Phi}^{\boldsymbol{c}}$. Hence,

$$\bar{\mathbf{P}}\boldsymbol{\phi}^{\boldsymbol{c}} = (\mathbf{1}+\boldsymbol{\Phi}^{\boldsymbol{c}})\boldsymbol{\phi}^{\boldsymbol{c}} = \boldsymbol{\phi}^{\boldsymbol{c}}, \quad \text{which gives } \mathbf{k} = \frac{\boldsymbol{\phi}^{\boldsymbol{c}}}{|\boldsymbol{\phi}^{\boldsymbol{c}}|}. \tag{3.19}$$

We cannot calculate the magnitude of the rotation by taking $\mathrm{tr}\,\bar{\mathbf{P}}$, since it gives $1 + 2\cos\theta = 3$, which implies $\theta = 0$. We consider the following limit:

$$1 + \boldsymbol{\Phi}^{\boldsymbol{c}} = \lim_{\phi_1^c \to 0}\lim_{\phi_2^c \to 0}\lim_{\phi_3^c \to 0} \mathbf{R}_1(\phi_1^c)\mathbf{R}_2(\phi_2^c)\mathbf{R}_3(\phi_3^c). \tag{3.20}$$

Now we take the trace of the r.h.s. and get $\theta = |\boldsymbol{\phi}^{\boldsymbol{c}}|$. Hence, the Gibbs rotation vector of $\bar{\mathbf{P}}$, $\bar{\boldsymbol{\eta}}$ is given as

$$2\bar{\boldsymbol{\eta}} = 2\tan\frac{\theta}{2}\mathbf{k} \approx |\boldsymbol{\phi}^{\boldsymbol{c}}|\frac{\boldsymbol{\phi}^{\boldsymbol{c}}}{|\boldsymbol{\phi}^{\boldsymbol{c}}|} = \boldsymbol{\phi}^{\boldsymbol{c}}. \tag{3.21}$$

The Gibbs rotation vector of $\mathbf{P}$ is simply $\boldsymbol{\eta} = \mathbf{Z}\bar{\boldsymbol{\eta}}$. Note that in the undeformed state $\boldsymbol{\eta}_0 = \bar{\boldsymbol{\eta}}_0 = 0$. We now proceed to calculate the rigid rotation of the spoke $\mathbf{Q}$,

$$\mathbf{Q} = \mathbf{P}\mathbf{Q}^- = \mathbf{Z}\left(\mathbf{1} + \frac{\mathbf{R}^+\boldsymbol{\Theta}^+\mathbf{R}^{+\mathrm{T}} + \mathbf{R}^-\boldsymbol{\Theta}^-\mathbf{R}^{-\mathrm{T}}}{2}\right)\mathbf{Q}_0 = \mathbf{Z}(\mathbf{1}+\boldsymbol{\Phi})\mathbf{Q}_0 = \mathbf{Z}(\mathbf{1}+\boldsymbol{\Phi})\mathbf{Q}_0. \tag{3.22}$$

Here $\boldsymbol{\eta} \sim O(\varepsilon)$. Now, the micro-moment $\boldsymbol{c}$ is related linearly to the $\boldsymbol{\eta}$ via an elastic tensor $\mathbf{H}$,

$$\boldsymbol{c} = \mathbf{Q}\bar{\mathbf{H}}[\mathbf{Q}^{\mathrm{T}}\boldsymbol{\eta} - \mathbf{Q}_0^{\mathrm{T}}\boldsymbol{\eta}_0] + O(\varepsilon^2) \approx \mathbf{Z}\mathbf{Q}_0\bar{\mathbf{H}}\mathbf{Q}_0^{\mathrm{T}}\bar{\boldsymbol{\eta}}. \tag{3.23}$$

For further reference, let

$$\hat{\boldsymbol{\xi}} = \mathbf{Q}_0^{\mathrm{T}}\bar{\eta}. \tag{3.24}$$

(ii) Extension of the web

The distance between the two strands is $\boldsymbol{w} = (\mathbf{r}^+ - \mathbf{r}^-)/2$ and in the undeformed configuration $\mathbf{w}_0 = (\mathbf{r}_0^+ - \mathbf{r}^-{}_0)/2$. By direct calculation, we observe

$$\left.\begin{aligned} \mathbf{w}_0 &= b\sin\frac{\alpha}{2}\left(\sin\left(\omega x + \frac{\alpha}{2}\right)\,\mathbf{e}_1 - \cos\left(\omega x + \frac{\alpha}{2}\right)\,\mathbf{e}_2\right) \\ \text{and}\qquad \boldsymbol{w} &= \left(b\sin\frac{\alpha}{2} + w_1\right)\left(\sin\left(\omega x + \frac{\alpha}{2}\right)\,\mathbf{d}_1 - \cos\left(\omega x + \frac{\alpha}{2}\right)\,\mathbf{d}_2\right) \\ &\quad + w_2\left(\cos\left(\omega x + \frac{\alpha}{2}\right)\,\mathbf{d}_1 + \sin\left(\omega x + \frac{\alpha}{2}\right)\,\mathbf{d}_2\right), \end{aligned}\right\} \tag{3.25}$$

where

$$w_1 = \frac{r + r^-}{2}\sin\frac{\alpha}{2} - b\,\frac{\beta^+ - \beta^-}{2}\cos\frac{\alpha}{2} \quad \text{and} \quad w_2 = \frac{r - r^-}{2}\cos\frac{\alpha}{2} + b\,\frac{\beta^+ + \beta^-}{2}\sin\frac{\alpha}{2}.$$

The force exerted by the + strand on the − strand $\boldsymbol{f}$ is given by

$$\boldsymbol{f} = \mathbf{Q}\bar{\mathbf{L}}[\mathbf{Q}^{\mathrm{T}}\boldsymbol{w} - \mathbf{Q}_0^{\mathrm{T}}\boldsymbol{w}_0], \tag{3.26}$$

where $\bar{\mathbf{L}}$ is a tensor of mechanical properties of the web. This force $\boldsymbol{f}$ causes the web to extend and shear. For further reference let

$$\hat{\mathbf{w}} = \mathbf{Q}^{\mathrm{T}}\boldsymbol{w} - \mathbf{Q}_0^{\mathrm{T}}\boldsymbol{w}_0. \tag{3.27}$$

(iii) Stacking energy

DNA consists of consecutive basepairs stacked on top of each other in a regular fashion. The resistance to external forces and moments not only comes from the elastic deformation of the strands and the webbing, but also from the change in alignment of the basepairs. We call the energy associated with this change in the bases' position and spatial orientation 'stacking energy'. Stacking energy plays a critical role in various phenomena such as melting of DNA [23,24]. We prescribe a form of free energy which is quadratic in the twist k_3 and stretch ξ,

$$F_{\mathrm{int}} = K_c k_3^2 + K_e \xi^2. \tag{3.28}$$

There are other sophisticated expressions for the stacking energy [24], but we use the quadratic form for two reasons: (i) the non-quadratic terms in the energy of [24] account for effects such as basepair severing which are crucial to DNA melting, which does not occur in our problem and (ii) a quadratic energy keeps our problem linear. This interaction energy results in a distributed body force $\boldsymbol{l}$ and distributed body moment $\boldsymbol{h}$ on the strands

$$\boldsymbol{h} = K_c k_{3x}\,\mathbf{d}_3 \quad \text{and} \quad \boldsymbol{l} = K_e \xi_x\,\mathbf{d}_3. \tag{3.29}$$

(d) Step 4: governing equations

We are now in a position to solve the governing equations for the mechanics of our helical birod. These equations consist of the balance of linear momentum and angular momentum for both the strands. In the balance equations (3.30) and (3.31):

— $\boldsymbol{m}^{\pm} = EI\kappa^{\pm}$ (equation (3.11)) denotes the elastic moment in the ± strand. $\boldsymbol{n}^{\pm}$ are the contact forces for which there is no constitutive relation since the outer strands are assumed to be inextensible and unshearable.

— $\boldsymbol{f}$ and $\boldsymbol{c}$ are the distributed force and moment, respectively, exerted by the + strand on the − strand.
— $\boldsymbol{l}$ and $\boldsymbol{h}$ are the distributed force and moment exerted by basepairs on the + and − strand.

The balance equations are

$$\boldsymbol{n}_x^+ - \boldsymbol{f} + \boldsymbol{l} = 0, \tag{3.30a}$$

$$\boldsymbol{n}_x^- + \boldsymbol{f} + \boldsymbol{l} = 0, \tag{3.30b}$$

$$\boldsymbol{m}_x^+ + \mathbf{r}_x^+ \times \boldsymbol{n}^+ + \tfrac{1}{2}(\mathbf{r}^+ - \mathbf{r}^-) \times \boldsymbol{f} - \boldsymbol{c} + \boldsymbol{h} = 0 \tag{3.31a}$$

and

$$\boldsymbol{m}_x^- + \mathbf{r}^-{}_x \times \boldsymbol{n}^- + \tfrac{1}{2}(\mathbf{r}^+ - \mathbf{r}^-) \times \boldsymbol{f} + \boldsymbol{c} + \boldsymbol{h} = 0. \tag{3.31b}$$

Let $[\mathbf{f}_1\ \mathbf{f}_2\ \mathbf{f}_3] = \mathbf{Z}\mathbf{Q}_0$. This gives

$$\left.\begin{aligned} \mathbf{f}_1 &= \left(\sin\left(\omega x + \frac{\alpha}{2}\right)\mathbf{d}_1 - \cos\left(\omega x + \frac{\alpha}{2}\right)\mathbf{d}_2\right) \\ \text{and}\quad \mathbf{f}_2 &= \left(\cos\left(\omega x + \frac{\alpha}{2}\right)\mathbf{d}_1 + \sin\left(\omega x + \frac{\alpha}{2}\right)\mathbf{d}_2\right), \quad \mathbf{f}_3 = \mathbf{d}_3. \end{aligned}\right\} \tag{3.32}$$

We decompose the forces $\boldsymbol{n}^+ = (\boldsymbol{n} + \boldsymbol{n}^c) \sim O(\varepsilon)$ and $\boldsymbol{n}^- = (\boldsymbol{n} - \boldsymbol{n}^c) \sim O(\varepsilon)$; $\boldsymbol{n} = n_1\mathbf{f}_1 + n_2\mathbf{f}_2 + n_3\mathbf{f}_3$ and $\boldsymbol{n}^c = n_1^c\mathbf{f}_1 + n_2^c\mathbf{f}_2 + n_3^c\mathbf{f}_3$. Now, $\boldsymbol{n}_x = (n_{1x} - \omega n_2)\ \mathbf{f}_1 + (n_{2x} + \omega n_1)\ \mathbf{f}_2 + n_{3x}\mathbf{f}_3 + O(\varepsilon^2)$. Similarly for $\boldsymbol{n}_x^c$. We use $\boldsymbol{c} = c_1\mathbf{f}_1 + c_2\mathbf{f}_2 + c_3\mathbf{f}_3$ and $\boldsymbol{f} = f_1\mathbf{f}_1 + f_2\mathbf{f}_2 + f_3\mathbf{f}_3$ from equations (3.26) and (3.23). Then, the balance equations become:

$$\left.\begin{aligned}
&n_{1x} - \omega n_2 = 0,\\
&n_{2x} + \omega n_1 = 0,\\
&n_{3x} + K_e\xi_x = 0,\\
&n_{1x}^c - \omega n_2^c - f_1 = 0,\\
&n_{2x}^c + \omega n_1^c - f_2 = 0,\\
&n_{3x}^c - f_3 = 0,\\
&EI\cos k\left[(\kappa_x^+ + \kappa_x^-)\cos\frac{\alpha}{2} + (\kappa^+ - \kappa^-)\omega\sin\frac{\alpha}{2}\right] - 2n_2 + 2a\omega n_3^c\sin\frac{\alpha}{2} = 0,\\
&EI\cos k\left[(\kappa_x^- - \kappa_x^+)\sin\frac{\alpha}{2} + (\kappa^- + \kappa^+)\omega\cos\frac{\alpha}{2}\right] + 2n_1 + 2a\omega n_3\cos\frac{\alpha}{2} - 2af_3\sin\frac{\alpha}{2} = 0,\\
&EI\sin k(\kappa_x^+ + \kappa_x^-) + 2af_2\sin\frac{\alpha}{2} - 2sn_2\cos\frac{\alpha}{2} - 2a\omega n_1^c\sin\frac{\alpha}{2} + 2K_ck_{3x} = 0,\\
&EI\cos k\left[(\kappa_x^+ - \kappa_x^-)\cos\frac{\alpha}{2} + (\kappa^+ + \kappa^-)\omega\sin\frac{\alpha}{2}\right] + 2a\omega n_3\sin\frac{\alpha}{2} - 2n_2^c - 2c_1 = 0,\\
&EI\cos k\left[-(\kappa_x^- + \kappa_x^+)\sin\frac{\alpha}{2} + (\kappa^+ - \kappa^-)\omega\cos\frac{\alpha}{2}\right] + 2\omega n_3^c\cos\frac{\alpha}{2} + 2n_1^c - 2c_2 = 0\\
\text{and}\quad &EI\sin k(\kappa_x^+ - \kappa_x^-) - 2a\omega n_2^c\cos\frac{\alpha}{2} - 2a\omega n_1\sin\frac{\alpha}{2} - 2c_3 = 0.
\end{aligned}\right\} \tag{3.33}$$

We have 12 differential equations in the 12 unknowns $(r, f, \xi, k_3, \beta^+, \beta^-, n_1^c, n_2^c, n_3^c, n_1, n_2, n_3)$. We substitute the following ansatz into the equations:

$$y = y_0\,\mathrm{e}^{-\lambda x} \quad \text{where } y \text{ could be } r(x), f(x), \xi(x), k_3(x), \beta^+(x), \beta^-(x), n_1^c, n_2^c, n_3^c, n_1, n_2, n_3. \tag{3.34}$$

This results in an eigenvalue problem. We find 23 eigenvalues, but retain only six for reasons explained in the electronic supplementary material. Let those six eigenvalues be $\pm\lambda, \pm\mu, \pm\delta$ and the corresponding eigenvectors $\mathbf{v}_{\pm\lambda}$ and $\mathbf{v}_{\pm\mu}$. Let

$$\mathbf{v}(x) = [r(x)\, f(x)\, \xi(x)\, k_3(x)\, \beta^+(x)\, \beta^-(x)\, n_1^c(x)\, n_2^c(x)\, n_3^c(x)\, n_3(x)\, n_1(x)\, n_2(x)]^{\mathrm{T}}.$$

Hence,

$$\mathbf{v}(x) = p_1 \mathrm{e}^{-\lambda x}\mathbf{v}_\lambda + p_2 \mathrm{e}^{\lambda x}\mathbf{v}_{-\lambda} + p_3 \mathrm{e}^{-\mu x}\mathbf{v}_\mu + p_4 \mathrm{e}^{\mu x}\mathbf{v}_{-\mu} + p_5 \mathrm{e}^{-\delta x}\mathbf{v}_\delta + p_6 \mathrm{e}^{\delta x}\mathbf{v}_{-\delta}. \tag{3.35}$$

Here, p_1, p_2, p_3, p_4, p_5 and p_6 are the constants which are determined using boundary conditions.

(e) Step 5: boundary conditions

We assume that the impact of a protein binding to DNA is twofold: (i) the protein fixes the curvatures at the binding site as in [25,27] and (ii) the protein causes a change in the radius of the DNA helix [5] as shown in the inset of figure 5*b*. Thus, we apply boundary conditions on the curvatures k_1, k_2 and the change in radius r of the DNA helix. We discuss two cases: (i) when one protein binds to the DNA and (ii) when two proteins bind to it.

(i) *One protein*. Let us assume that the protein binds at $x = 0$. The boundary conditions for this case are:

$$\left.\begin{array}{lll} \text{At} & x = 0, & k_1(x) = k_{10}, \quad k_2(x) = k_{20}, \quad r(x) = r_0. \\ \text{As} & x \to \pm\infty, & k_1(x),\ k_2(x),\ r(x) \to 0. \end{array}\right\} \tag{3.36}$$

The second boundary condition says that the DNA is straight far away from the protein and that the perturbation in DNA radius occurs only in the vicinity of the bound protein.

(ii) *Two proteins*. Let us assume that the two proteins bind at $x = 0$ and $x = a$, respectively. We divide our domain into three parts $-\infty < x < 0$, $0 < x < a$ and $a < x < \infty$, each of which has different boundary conditions attached to it.

$$\left.\begin{array}{ll} \text{Region 1}: x \in (-\infty, 0) & \\ \text{as } x \to -\infty, & k_1(x), k_2(x), r(x) \to 0, \\ \text{at } x = 0, & k_1(x) = k_{11},\ k_2(x) = k_{12},\ r(x) = r_1. \\ \text{Region 2}: x \in (0, a) & \\ \text{at } x = 0, & k_1(x) = k_{11},\ k_2(x) = k_{12},\ r(x) = r_1, \\ \text{at } x = a, & k_1(x) = k_{21},\ k_2(x) = k_{22},\ r(x) = r_2. \\ \text{Region 3}: x \in (a, \infty) & \\ \text{at } x = a, & k_1(x) = k_{21},\ k_2(x) = k_{22},\ r(x) = r_2, \\ \text{as } x \to \infty, & k_1(x), k_2(x), r(x) \to 0. \end{array}\right\} \tag{3.37}$$

(f) Step 6: energy of the birod

We assume small elastic deformations throughout, hence the resulting energy is quadratic in the strain variables. The elastic energy has contributions from the bending of the outer strands (equation (3.11)), the extension, bending and twisting of the web (equations (3.24) and (3.25)) and the stacking energy (equation (3.29)).

$$E = \int_{-\infty}^{\infty} \left[\frac{1}{2}EI\kappa^{+2} + \frac{1}{2}EI\kappa^{-2} + \frac{1}{2}\hat{\mathbf{w}}.\mathbf{L}\hat{\mathbf{w}} + \frac{1}{2}\hat{\boldsymbol{\zeta}}.\mathbf{H}\hat{\boldsymbol{\zeta}} + K_e\xi^2 + K_c k_3^2\right] \mathrm{d}x. \tag{3.38}$$

We are especially interested in the interaction energy ΔG, which is the elastic energy of interactions between the two proteins,

$$\Delta G = E_a^2 - E_0^1 - E_a^1, \tag{3.39}$$

where E_a^2 is the energy of two proteins bound to DNA, one at $x = 0$ and other at $x = a$, and E_a^1 and E_0^1 are the elastic energies corresponding to a single protein binding at $x = a$ and $x = 0$, respectively.

4. Elastic constants

Our model has nine elastic constants $L_1, L_2, L_3, H_1, H_2, H_3, K_c, K_e, EI$. The experimental values for these constants are not known. In order to get some idea about the magnitude of the elastic constants we calculate the extensional modulus, torsional modulus and twist–stretch coupling modulus for a double-stranded DNA within our birod model. The explicit calculation is presented in the electronic supplementary material. We choose

$$\left.\begin{aligned} &K_c = 80\,\mathrm{pN\,nm^2}, \quad K_e = 600\,\mathrm{pN}, \quad H_1 = 15\,\mathrm{pN}, \quad H_2 = 10\,\mathrm{pN},\\ &H_3 = 20\,\mathrm{pN}, \quad L_1 = 50\,\mathrm{pN\,nm^{-1}}\\ \text{and}\quad &L_2 = 250\,\mathrm{pN\,nm^{-1}}, \quad L_3 = 30\,\mathrm{pN\,nm^{-1}}, \quad EI = 65\,\mathrm{pN\,nm^2}. \end{aligned}\right\} \tag{4.1}$$

This choice of elastic constants gives the extensional modulus $S \approx 1245\,\mathrm{pN}$, torsional modulus $C \approx 490\,\mathrm{pN\,nm^2}$ and twist–stretch coupling modulus $g \approx -90\,\mathrm{pN\,nm}$, which are close to actual values for double-stranded DNA [4] measured in experiments. We point out that this choice of elastic constants is not unique, nonetheless we use them to make further calculations.

When we substitute these constants into the governing equations (equation (3.33)) and solve the eigenvalue problem involving λ, we get the following eigenvalues:

$$\lambda_1 = -0.68,\ \lambda_2 = -0.42,\ \lambda_3 = -0.36,\ \lambda_4 = 0.36,\ \lambda_5 = 0.42,\ \lambda_6 = 0.68. \quad \text{Units: } \mathrm{nm^{-1}} \tag{4.2}$$

Other eigenvalues are either very large ($\to \pm\infty$), very small (~ 0) or purely imaginary. Purely imaginary and zero eigenvalues when substituted in $\mathrm{e}^{\lambda x}$ give a sinusoidal and a constant function, respectively, which do not decay to zero as $x \to \pm\infty$. As mentioned in §3e, the curvatures k_1, k_2 and change in radius r must go to zero at $\pm\infty$. Thus, zero or purely imaginary eigenvalues cannot satisfy our boundary conditions, and are, therefore, not useful. We refer the reader to the electronic supplementary material for further discussion on the choice of eigenvalues.

Consider a situation in which two proteins bind DNA, one at $x = 0$ and the other at $x = a$. In the region $a < x < \infty$, the solution equation (3.35) consists of only negative eigenvalues. There are three negative eigenvalues $\lambda_{2,3,4}$ and consequently three unknown constants. We have three boundary conditions on k_1, k_2 and r at $x = a$ to determine those constants. Similarly in the region $-\infty < x < 0$, the solution consists of only positive eigenvalues $\lambda_{7,8,9}$, so the constants can again be evaluated from three boundary conditions. We use this scheme to evaluate the strain parameters which we substitute into the expression for the elastic energy functional equation (3.38). Note that the dominant eigenvalue $\pm 0.36\,\mathrm{nm^{-1}}$ corresponds to a decay length of 2.8 nm ($\approx$10 bp), which is what Kim *et al.* [5] report in their experiments.

5. Results

The experimental evidence for allosteric interactions when two proteins bind to DNA is documented in Kim *et al.* [5]. Many earlier papers have also described allostery in DNA, but Kim *et al.* present exquisite quantitative details which call for a quantitative explanation.

To unravel the physics behind these allosteric interactions, we begin by examining the case when one protein binds to DNA. As discussed in §2d, the strain variables $(r, \zeta, \beta^{\pm}, k_{1,2,3})$ are linear combinations of decaying exponentials. For instance, consider $k_3(x)$ for a protein binding at $x = 0$:

$$\left.\begin{aligned} &k_3(x) = p_1 \mathbf{v}_{-\lambda}(4)\,\mathrm{e}^{\lambda x} + p_2 \mathbf{v}_{-\mu}(4)\,\mathrm{e}^{\mu x} + p_3 \mathbf{v}_{-\delta}(4)\,\mathrm{e}^{\delta x} \quad x < 0\\ \text{and}\quad &k_3(x) = q_1 \mathbf{v}_{\lambda}(4)\,\mathrm{e}^{-\lambda x} + q_2 \mathbf{v}_{\mu}(4)\,\mathrm{e}^{-\mu x} + q_3 \mathbf{v}_{\delta}(4)\,\mathrm{e}^{-\delta x} \quad x > 0, \end{aligned}\right\} \tag{5.1}$$

where $\lambda = 0.36\,\mathrm{nm^{-1}}$, $\mu = 0.42\,\mathrm{nm^{-1}}$ and $\delta = 0.68\,\mathrm{nm^{-1}}$. $\mathbf{v}_{\pm\lambda}$, $\mathbf{v}_{\pm\mu}$ and $\mathbf{v}_{\pm\delta}$ are the eigenvectors associated with eigenvalues $\pm\lambda$, $\pm\mu$ and $\pm\delta$, respectively. The constants p_i and q_i $(i = 1, 2, 3)$ are evaluated using the boundary conditions at $x = 0$. It is not difficult to see that the strain variables decay to zero as $x \to \pm\infty$. We can replace k_3 in the above equation by other strain variables $(r, \xi, \beta^{\pm})$ and recover similar behaviour. We discuss a few characteristics of the variation of the

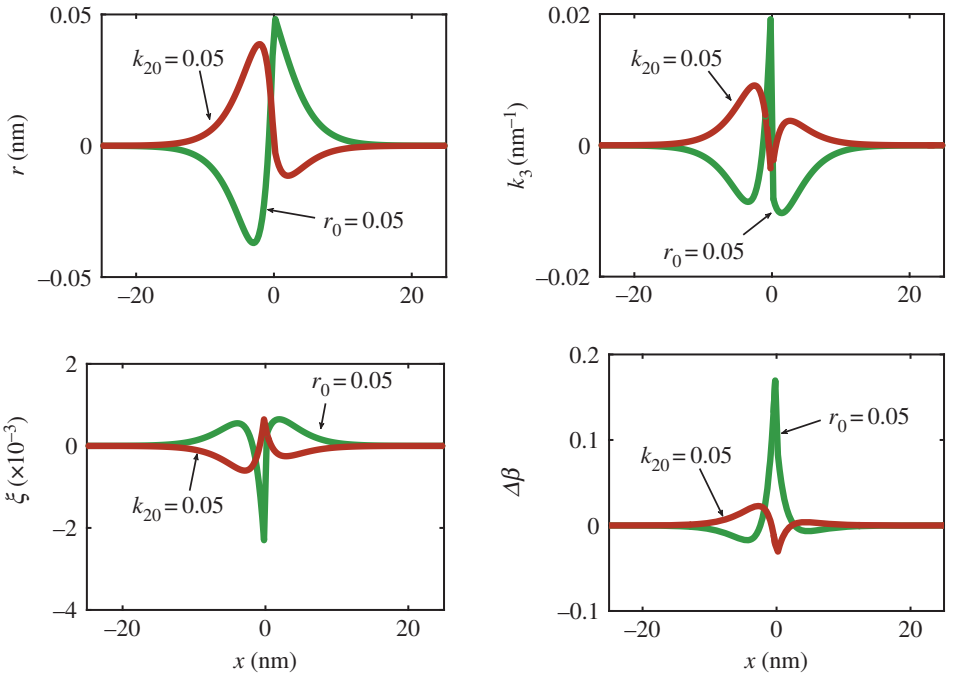

Figure 2. Variation of strain variables for a single protein. We plot the change in radius r, twist k_3, stretch of centreline ξ and change in phase angle $\Delta\beta = \beta^+ - \beta^-$ for the double helix. The red curve corresponds to the boundary conditions $k_{10} = r_0 = 0$ and $k_{20} = 0.1\,\mathrm{nm}^{-1}$ at $x = 0$ and the green curve corresponds to $k_{10} = k_{20} = 0$ and $r_0 = 0.05$ nm at $x = 0$. The asymmetry of the double helix (there is a major and minor groove in DNA) arising from the phase angle $\alpha = 2.1$ radian gives the curves a slight asymmetry about the site of protein binding. The curves are exactly symmetric about the site of protein binding if we choose phase angle $\alpha = \pi$ radians (which results in no major and minor groove) as shown in the electronic supplementary material. (Online version in colour.)

strain parameters as functions of position. The results are plotted in figures 2 and 3. The strain parameters $(r, k_3, \beta^{\pm})$ decay exponentially with distance from the site of protein binding. The curvatures exhibit an exponentially decaying sinusoidal character with a period of 11 bp. This periodic decay of the curvatures manifests itself as sinusoidal variations in the interaction energy. We find that these plots are slightly asymmetric about $x = 0$. We attribute this to the structural asymmetry in the right-handed double helix with phase angle $\alpha = 2.1$ radians. If we choose the phase angle $\alpha = \pi$ radians instead, we find that the plots are exactly symmetric about the site of protein binding as shown in the electronic supplementary material. We now consider the case when two proteins bind to DNA, one at $x = 0$ and the other at $x = a$. We proceed in a similar manner as above and express the strain profiles as linear combinations of exponentials

$$\left.\begin{aligned}
&\text{Case 1} \quad k_3(x) = p_1\mathbf{v}_{-\lambda}(4)\,\mathrm{e}^{\lambda x} + p_2\mathbf{v}_{-\mu}(4)\,\mathrm{e}^{\mu x} + p_3\mathbf{v}_{-\delta}(4)\,\mathrm{e}^{\delta x} && x < 0,\\
&\text{Case 2} \quad k_3(x) = m_1\mathbf{v}_{\lambda}(4)\,\mathrm{e}^{-\lambda x} + m_2\mathbf{v}_{\mu}(4)\,\mathrm{e}^{-\mu x} + m_3\mathbf{v}_{-\lambda}(4)\,\mathrm{e}^{\lambda x}\\
&\qquad\qquad + m_4\mathbf{v}_{-\mu}(4)\,\mathrm{e}^{\mu x} + m_5\mathbf{v}_{\delta}(4)\,\mathrm{e}^{-\delta x} + m_6\mathbf{v}_{-\delta}(4)\,\mathrm{e}^{\delta x} && 0 < x < a\\
\text{and} \quad &\text{Case 3} \quad k_3(x) = q_1\mathbf{v}_{\lambda}(4)(3)\,\mathrm{e}^{-\lambda x} + q_2\mathbf{v}_{\mu}(4)\,\mathrm{e}^{-\mu x} + q_3\mathbf{v}_{\delta}(4)\,\mathrm{e}^{-\delta x} && x > a.
\end{aligned}\right\} \tag{5.2}$$

The constants p_i and q_i $(i = 1, 2, 3)$ are determined by three boundary conditions (on k_1, k_2 and r) at $x = 0$ and $x = a$, respectively. The constants m_j $(j = 1, 2, 3, 4, 5, 6)$ are determined by six boundary conditions at $x = 0$ and $x = a$. The behaviour of the strain variables for two proteins is similar to that for one protein as shown in figure 4. When two proteins are separated by a large distance $a > 10 \times 3.4\,\mathrm{nm}$ (i.e. more than 10 helical turns of DNA), the strain profile looks

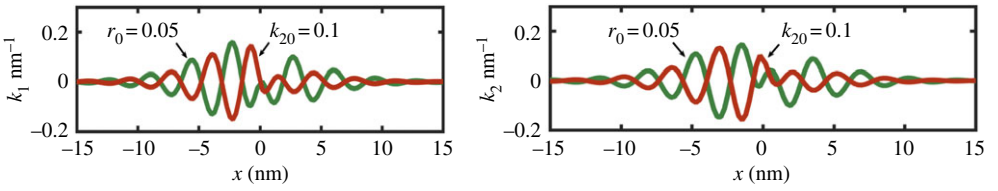

Figure 3. Variation of curvatures k_1 and k_2 for a single protein. The red curve corresponds to the boundary conditions $k_{10} = r_0 = 0$ and $k_{20} = 0.1\,\text{nm}^{-1}$ at $x = 0$ and the green curve corresponds to $k_{10} = k_{20} = 0$ and $r_0 = 0.05$ nm at $x = 0$. We find that the curvature decays exponentially and oscillates with a period ≈11 bp. (Online version in colour.)

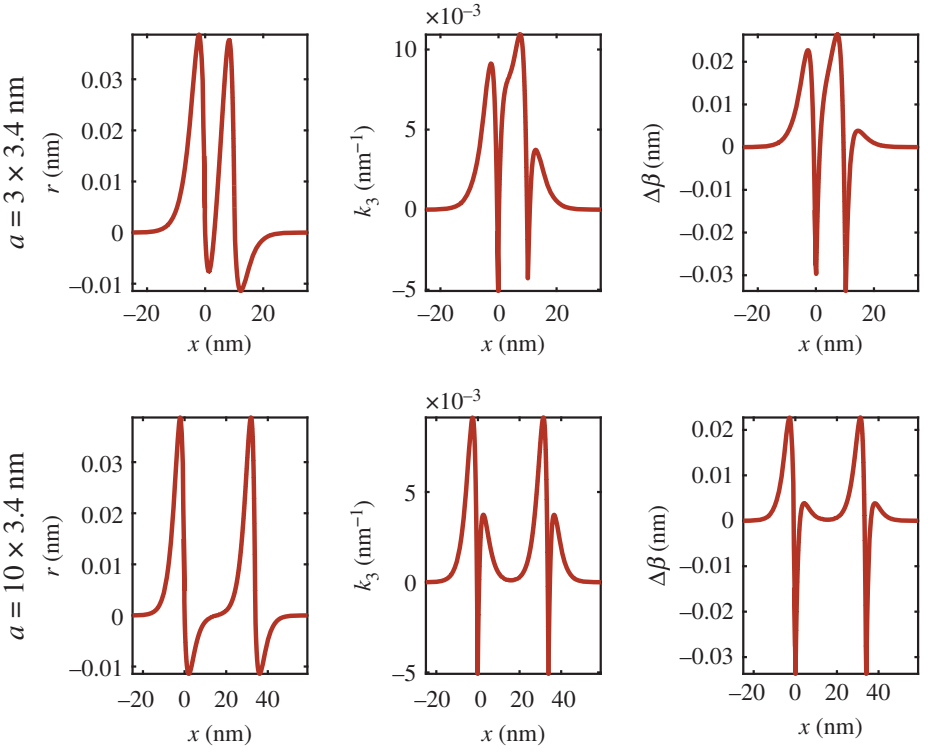

Figure 4. Variation of r, k_3 and $\Delta\beta$ for two proteins. Here a is the distance between the sites of protein binding. The strain variables decay exponentially away from the site of protein binding. When the distance between the proteins is large (10 × 3.4 nm), the profile looks like a concatenation of two solutions for a single protein. (Online version in colour.)

like a concatenation of the profiles of two proteins binding separately. Their strain fields do not interact at such distances, thus there is little interaction energy. When the distance decreases, the strain fields of the two proteins overlap, and this is responsible for the interaction energy. As discussed in the electronic supplementary material, two defects on a straight ladder interact via an interaction energy that decays exponentially with the distance between them. Now, we focus on the double-helical birod and examine the behaviour of different boundary conditions on the interaction energy ΔG in figure 5. We assume for simplicity that both proteins apply the same boundary conditions on the DNA, the exact numerical values are given in the figure. If we choose the change in radius $r_0 = 0$ and apply the boundary conditions only on the two curvatures k_1, k_2, the interaction energy decays exponentially while varying sinusoidally with a period of $5.5 \approx 11/2$ bp. This case corresponds to proteins that bend DNA as shown in the inset of figure 5*b*. On the other hand, if the curvatures k_1, k_2 are zero while the change in radius r_0 is non-zero,

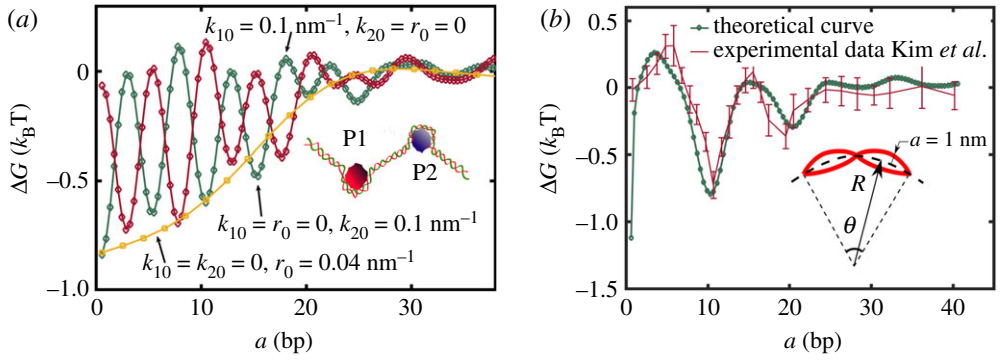

Figure 5. We plot the interaction energy between two proteins (equation (3.39)). In (*a*), we plot the behaviour of ΔG for various boundary conditions. If the boundary conditions are specified on the curvatures, we get an exponentially decaying profile oscillating with 5–6 bp ($\approx \frac{11}{2}$ bp). The oscillatory behaviour arises from the periodic geometry of DNA. In (*b*), the experimental data reproduced for comparison are from Kim *et al* [5]. We use $k_{11} = k_{21} \approx 0.02\,\mathrm{nm}^{-1}$, $k_{12} = k_{22} = 0.05\,\mathrm{nm}^{-1}$, $r_1 = -r_2 = 0.02$ nm. The inset in (*a*) shows a protein–DNA complex in which the proteins locally bend DNA. The inset in (*b*) shows that bending a DNA oligomer leads to widening of the groove on one side while narrowing it on the other. We find that the change in the groove width is approximately 2 A which is close to the values reported by Kopka *et al.* [26] (0.5–2 A). (Online version in colour.)

we get an exponentially decaying profile devoid of any oscillatory character, which is similar to the results for the ladder. The exponentially decaying component originates from the elasticity of the web, and the sinusoidal behaviour comes from the double-helical structure of DNA. From this exercise, we conclude that in order to get a sinusoidally varying interaction energy a protein must change the local curvature in the DNA, a mere change in radius of the DNA is not sufficient to give rise to the interaction energy profiles observed in experiments.

In our model, the magnitude of the interaction energy increases monotonically with an increase in the magnitude of the changes in curvatures or radius caused by the two proteins. Thus, by systematically varying the boundary conditions imposed by the proteins we can establish agreement of our theoretical results for ΔG with the experimental values documented by Kim *et al.* [5]. This is done in figure 5*b*. The values of the curvatures that give the best fit to the experimental data are $k_{11} = k_{21} = 0.02\,\mathrm{nm}^{-1}$, $k_{12} = k_{22} = 0.05\,\mathrm{nm}^{-1}$ and $r_1 = -r_2 = 0.02$ nm. The magnitude of the curvature $|k| = \sqrt{k_1^2 + k_2^2} \approx 0.055\,\mathrm{nm}^{-1}$, which gives a radius of curvature in figure 5*b* (inset) $R = 1/k \approx 18$ nm. Assuming the centreline does not extend and the pitch of the DNA helix is 3.4 nm, we get $\theta = 3.4/18 \approx 0.18$ rad. The radius of the DNA molecule is $a = 1$ nm, thus the change in the groove width is approximately given by $\Delta = (R + a)\theta - R\theta = 1 \times 0.18\,\mathrm{nm} \approx 2$ A. This value is close to the one reported by Kopka *et al.* [26] (0.5–2 A). Hence, our choice of curvature boundary conditions is reasonable; it is, however, not unique and it is coupled with the choice of stiffnesses of the webbing in our birod model. Be that as it may, our exercise above demonstrates that a birod model can capture the dependence of interaction energy on the distance between proteins bound to DNA. Calibration of the model and faithfully connecting it to experiment will require deeper analysis, and, perhaps also, computation. It is important to note here that an elastic model of DNA allostery, consistently parametrized from atomic-resolution MD simulations, has already been proposed by Dršata *et al.* [20], which was later extended to describe the experiment of Kim *et al.* (Dršata *et al.* [21]). Our model is similar to the one presented by Dršata *et al.* in certain aspects:

(i) Quadratic energy: Dršata *et al.* [20] use a quadratic energy to penalize deformations, which is similar to our approach. This is appropriate for small deformations as assumed in our work and that of Dršata *et al.* The expression for the interaction energy in Dršata *et al.* [21] (eqn. (14) in their supplemental information) is identical to ours (equation (3.39)).

(ii) Boundary conditions: in our paper, a protein interacts with the DNA helix by imposing boundary conditions on the curvatures and the radius of the double helix at the binding site. In Geggier *et al.*, a protein interacts with DNA by changing the width of the minor groove. These approaches are equivalent since extending the minor/major groove at the binding site causes bending of DNA.

Our model differs from Geggier *et al.* [20] on the following key points:

(i) Effect of outer strands: the basepair centric model of Geggier *et al.* [20] ignores the effect of the mechanics of the phosphate backbones. We model the backbones as inextensible rods (or worm-like chains of polymer elasticity [22]) capable of bending. Elasticity of the outer strands is crucial for twist–stretch coupling observed in DNA (see the electronic supplementary material, §S3).
(ii) Stacking energy: we use a stacking energy quadratic in the twist and stretch of the centreline to account for the change in the orientations of successive basepairs relative to each other. Our approach is simplistic compared with the holistic approach of Geggier *et al.* [20], where stacking energy has quadratic contributions from the six inter-basepair degrees of freedom.
(iii) Boundary conditions: one of the drawbacks of our model is its inability to account for the nature of protein–DNA contacts. Geggier *et al.* [21] point out that while some proteins such as *Bam*HI show tight protein–DNA contacts, thereby constraining all inter- and intra-basepair degrees of freedom, others such as GRDBD interact rather loosely by altering only the major groove width. This flexibility is absent from our current model.

The period of the interaction energy in figure 5*a* is approximately 5.5 bp while that in figure 5*b* is 11 bp as in the experiment. Why? Note that the strain variables in a two-protein complex shown in figure 6*b* are a function of both the parameter x and the distance between the two proteins a. We fix x (=2 nm from protein P_1) and focus on the dependence on a. We assume that both the proteins apply identical boundary conditions. If the proteins do not cause any change in the radius such that $r_0 = 0$, then the strain parameters involved in the elastic energy (equation (3.38)) $\propto e^{-\Gamma x}\psi(\omega a)$, where $\psi(\omega a)$ is a sinusoidal function oscillating with a period 11 bp, and the elastic energy of the two-protein complex $\propto (e^{-\Gamma x}\psi(\omega a))^2$ oscillates with a period 5.5 bp. On the other hand, when the protein causes both a change in radius r_0 and a change in curvature k_{20}, the strain variables are $\propto (e^{-\Gamma_1 a}\psi(\omega a) + e^{-\Gamma_2 x})$ and the elastic energy of the two-protein complex $\propto (e^{-\Gamma_1 a}\psi(\omega a) + e^{-\Gamma_2 a})^2$ oscillates with a period of 11 bp due to the cross term $e^{-(\Gamma_1+\Gamma_2)a}\psi(\omega a)$. We plot the interaction energy $\Delta G(a)$ between the two proteins constituting the protein complex in figure 6*b* and verify the periods for respective boundary conditions, which resolves the apparent discrepancy in the periods in figure 5*a*,*b*. As a final application of our birod model, we examine the sequence dependence of the allosteric interaction energy ΔG. While there is overwhelming qualitative evidence, both experimental [5] and numerical [12], showing that AT-rich sequences exhibit stronger allosteric interactions than GC-rich ones, a theoretical explanation is still lacking. Stronger interactions are associated with longer decay lengths. Using our theory, we can find the dependence of the decay length on the elastic constants of the web. Since, AT basepairs consist of two hydrogen bonds, the corresponding elastic constants for the web are expected to be lower than GC basepairs, which consist of three hydrogen bonds. In an attempt to simulate such a scenario, we replace the elastic constants for the web (K_c, K_e, L_i, H_i $i = 1, 2, 3$) in equation (4.1) with ($\chi K_c, \chi K_e, \chi L_i, \chi H_i$ $i = 1, 2, 3$) while keeping EI fixed, and vary the parameter χ in the range $0.5 \leq \chi \leq 1$. We define a measure of the decay length l_d to be the inverse of the eigenvalue having the least non-zero magnitude, obtained in equation (4.2). For instance, if $\chi = 1$, decay length $l_d = 1/0.34\,\text{nm} \approx 10\,\text{bp}$. We plot the variation of l_d with χ in figure 7. We find that the decay length increases with the decrease in elastic constants of the web. We plot $\log l_d$ versus $\log \chi$ and deduce that $l_d \sim 1/\chi^{2/3}$.

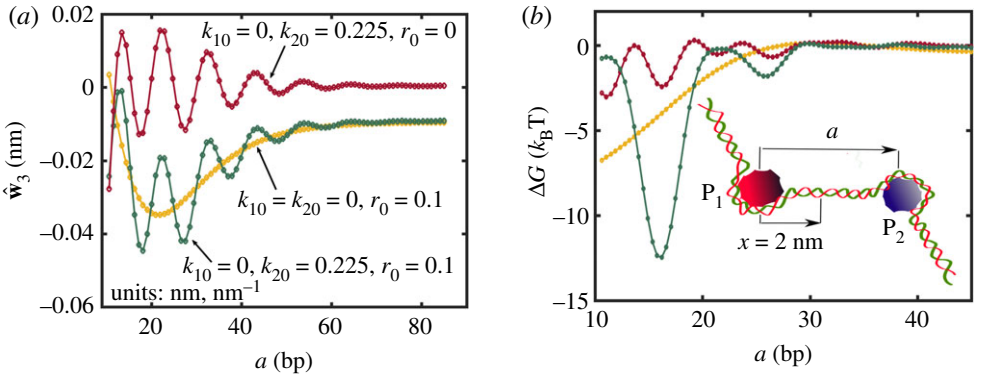

Figure 6. The inset in (*b*) shows a two-protein complex. The boundary conditions are identical for both the proteins $k_{11} = k_{21} = k_{10}, k_{12} = k_{22} = k_{20}, r_1 = r_2 = r_0$; the legend in (*a*) contains the exact numerical values. For (*b*), the legend is the same as in (*a*). We examine the behaviour of $\hat{\mathbf{w}}_3(x = 2\text{ nm}, a)$ (equation (3.38)) as a function of the distance between the two proteins *a* for these boundary conditions. The strain variables oscillate with a period of 11 bp. We observe that, in the case of $r_0 = 0$, the strain parameter $\hat{\mathbf{w}}_3(x = 2\text{ nm}, a)$ decays as $e^{-\Gamma a}\psi(\omega a)$, where $\psi(\omega a)$ is a sinusoidal function, hence the combined energy of a two-protein complex which is proportional to $(e^{-\Gamma a}\psi(\omega a))^2$ oscillates with a period of 5.5 bp (period of $\sin^2 x$ is half that of $\sin x$). If $k_{10} = k_{20} = 0$ the decay is exponential. If $r_0 \neq 0$ and k_{10} or $k_{20} \neq 0$, $\hat{\mathbf{w}}_3(x = 2\text{ nm}, a) \sim (e^{-\Gamma_1 a}\psi(\omega a) + e^{-\Gamma_2 a})$ and the energy of the two-protein complex, which is proportional to $(e^{-\Gamma_1 a}\psi(\omega a) + e^{-\Gamma_2 a})^2$, oscillates with a period of 11 bp. The behaviour of the other strain variables in equation (3.38) is similar. We plot the interaction energy $\Delta G(a)$ in (*b*) for the boundary conditions indicated in the legend of (*a*) and use it to verify the period we predict using this argument. (Online version in colour.)

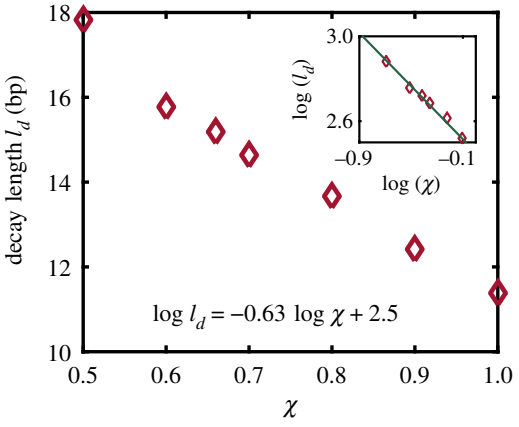

Figure 7. Decay length l_d is defined as the inverse of the eigenvalue with the least non-zero magnitude, for $\chi = 1$, $l_d = 1/|\lambda|_{\min} = 1/0.34 \approx 9$ bp. χ is meant to account for the reduction in the elastic constants for AT basepairs compared with GC basepairs. The elastic constants for the web are $(\chi K_c, \chi K_e, \chi H_1, \chi H_2, \chi H_3, \chi L_1, \chi L_2, \chi L_3)$, equation (4.1) gives the numerical values for $\chi = 1$. We find that the decay length increases with a decrease in elastic constants for the web, thus AT-rich DNA sequences are expected to have higher decay lengths. Qualitative experimental and numerical evidence in support of the above conclusion is documented in [5,12], respectively. The inset shows how we extracted the power law $l_d \sim \chi^{-2/3}$. (Online version in colour.)

6. Conclusion

Kim *et al.* [5] have presented compelling quantitative evidence for allosteric interactions between two proteins bound to DNA at distant locations. They showed that the interaction energy for two proteins separated by distance *a* on DNA is a decaying exponential oscillating with a period of 11 bp. Various attempts to numerically simulate the allosteric interactions have been made

[12,13] and have associated the oscillating interaction energy with the major groove width in the double-helical structure of DNA. We approach the problem from a purely mechanical standpoint. We conjecture that the local deformation field in DNA caused by a bound protein is similar to that produced by a defect in an elastic solid. We begin by computing the interaction energy for two defects on a ladder and find that it decays exponentially with the distance between them. We, then, proceed to replicate the same calculation for DNA by modelling it as a double-helical birod [14]. We assume that the outer phosphate backbones represented by $\pm$ strands to be inextensible and unshearable while the basepairs are capable of elastic extension, shear, bending and twisting. We assume a general form of displacement for these strands (equation (3.1)) which we use to calculate the micro-displacement and micro-rotation for the basepairs. We, then, use these expressions to solve the governing equations for our birod. A crucial factor in our treatment is the boundary conditions. We follow Kwiecinski *et al.* [25], Kim *et al.* [5] and Liang & Purohit [27] and impose boundary conditions on the curvatures and the radius of the DNA double helix. The question of what kind of boundary conditions a protein could possibly apply is not yet comprehensively addressed in the literature and is not the central issue of this study either. Rather, our message is that, after solving the governing equations and plugging in boundary conditions, we recover the exponentially decaying profile that oscillates with a period of 11 bp. We end by examining the sequence dependence of allosteric interactions and show that AT-rich sequences exhibit stronger interactions than GC-rich sequences.

Even though our birod model does surprisingly well by capturing the dependence of interaction energy on distance, there are many important caveats that we must point out. First, we do not expect our birod model to be accurate near the site of protein binding. The deformations near the binding site could be large enough that a linear elastic theory may not be applicable. Our assumptions that the outer strands are inextensible and the web is elastic could also break down in the vicinity of the binding site. Second, we have little knowledge of the elastic constants of the web. We have assumed some stiffness parameters for the web that gave the right experimentally verified moduli for the DNA, but there could have been another set of parameters that would have given similar results. One may have to appeal to molecular simulations [18–20,28] to get these parameters. Third, the boundary conditions applied by the proteins on the DNA are not clear. One may have to look for guidance from molecular simulations or protein–DNA co-crystal structures to get a clearer picture [21]. Finally, we have not accounted for fluctuations or entropic interactions in our model. This is partly justifiable because the length of DNA between two protein binding sites for which significant allosteric interactions are observed is often much smaller than the persistence length of the DNA. However, a rigorous calculation should be done to verify this assumption. In spite of these shortcomings, our model could provide a starting point for analysing allosteric interactions in DNA within the broad framework of configurational forces in elastic solids.

Authors' contributions. The research was conceived by P.K.P. and was carried out and written jointly by J.S. and P.K.P.

Competing interests. We declare no competing interests.

Funding. We acknowledge partial support for this work from a United States National Science Foundation grant no. NSF CMMI 1662101 and from the National Institutes of Health via grant no. NIH R01-HL 135254.

Acknowledgements. We acknowledge insightful discussion with Yujie Sun, who is one of the authors in Kim *et al.* [5]. We also thank John H. Maddocks for pointing us to literature regarding the material parameters entering the birod model.

References

1. Phillips R. 2001 *Crystals, defects and microstructures, modeling across scales*. Cambridge, UK: Cambridge University Press.
2. Gurtin ME. 2008 *Configurational forces as basic concepts of continuum physics*, vol. 137. Berlin, Germany: Springer Science and Business Media.
3. Weertman J, Weertman JR. 1992 *Elementary dislocation theory*. New York, NY: Oxford University Press.

4. Singh J, Purohit PK. 2017 Structural transitions in torsionally constrained DNA and their dependence on solution electrostatics. *Acta Biomater.* **55**, 214–225. (doi:10.1016/j.actbio.2017.03.052)
5. Kim S *et al.* 2013 Probing allostery through DNA. *Science* **339**, 816–819. (doi:10.1126/science.1229223)
6. Parekh BS, Hatfield GW. 1996 Transcriptional activation by protein-induced DNA bending: evidence for a DNA structural transmission model. *Proc. Natl Acad. Sci. USA* **93**, 1173–1177. (doi:10.1073/pnas.93.3.1173)
7. Hogan M, Dattagupta N, Crothers DM. 1979 Transmission of allosteric effects in DNA. *Nature* **278**, 521–524. (doi:10.1038/278521a0)
8. Moretti R, Donato LJ, Brezinski ML, Stafford RL, Hoff H, Thorson JS, Dervan PB, Ansari AZ. 2008 Targeted chemical wedges reveal the role of allosteric DNA modulation in protein-DNA assembly. *JACS Chem. Biol.* **3**, 220–229. (doi:10.1021/cb700258r)
9. Müller J, Oehler S, Müller-Hill B. 1996 Repression of lac promoter as a function of distance, phase and quality of an auxiliary lac operator. *J. Mol. Biol.* **257**, 21–29. (doi:10.1006/jmbi.1996.0143)
10. Becker NA, Kahn JD, Maher III LJ. 2005 Bacterial repression loops require enhanced DNA flexibility. *J. Mol. Biol.* **349**, 716–730. (doi:10.1016/j.jmb.2005.04.035)
11. Purohit PK, Nelson PC. 2006 Effect of supercoiling on formation of protein-mediated DNA loops. *Phys. Rev. E* **74**, 061907. (doi:10.1103/PhysRevE.74.061907)
12. Gu C *et al.* 2015 DNA structural correlation in short and long ranges. *J. Phys. Chem. B* **119**, 13 980–13 990. (doi:10.1021/acs.jpcb.5b06217)
13. Hancock SP, Ghane T, Cascio D, Rohs R, Di Felice R, Johnson RC. 2013 Control of DNA minor groove width and Fis protein binding by the purine 2-amino group. *Nucleic Acids Res.* **41**, 6750–6760. (doi:10.1093/nar/gkt357)
14. Moakher M, Maddocks JH. 2005 A double-strand elastic rod theory. *Arch. Ration. Mech. Anal.* **177**, 53–91. (doi:10.1007/s00205-005-0360-y)
15. Moulton DE, Lessinnes TH, Goriely A. 2013 Morphoelastic rods. Part I: A single growing elastic rod. *J. Mech. Phys. Solids* **61**, 398–427. (doi:10.1016/j.jmps.2012.09.017)
16. Lessinnes T, Moulton DE, Goriely A. 2017 Morphoelastic rods. Part II: Growing birods. *J. Mech. Phys. Solids* **100**, 147–196. (doi:10.1016/j.jmps.2015.07.008)
17. Manning RS, Maddocks JH, Kahn JD. 1996 A continuum rod model of sequence dependent DNA structure. *J. Chem. Phys.* **105**, 5626–5646. (doi:10.1063/1.472373)
18. Lankaš F, Gonzalez O, Heffler LM, Stoll G, Moakher M, Maddocks JH. 2009 On the parameterization of rigid base and basepair models of DNA from molecular dynamics simulations. *Phys. Chem. Chem. Phys.* **11**, 10 565–10 588. (doi:10.1039/b919565n)
19. Petkeviciute D. 2012 A DNA coarse-grain rigid base model and parameter estimation from molecular dynamics simulations. PhD thesis, École Polytechnique Fédérale de Lausanne, Lausanne, Switzerland.
20. Dršata T, Zgarbová M, Špačková N, Jurečka P, Šponer J, Lankaš F. 2014 Mechanical model of DNA allostery. *J. Phys. Chem. Lett.* **5**, 3831–3835. (doi:10.1021/jz501826q)
21. Dršata T, Zgarbová M, Jurečka P, Šponer J, Lankaš F. 2016 On the use of molecular dynamics simulations for probing allostery through DNA. *Biophys. J.* **110**, 874–876. (doi:10.1016/j.bpj.2015.12.039)
22. Marko JF, Siggia ED. 1995 Stretching DNA. *Macromolecules* **28**, 8759–8770. (doi:10.1021/ma00130a008)
23. Peyrard M, Bishop AR. 1989 Statistical mechanics of a nonlinear model for DNA denaturation. *Phys. Rev. Lett.* **62**, 2755. (doi:10.1103/PhysRevLett.62.2755)
24. Dauxois T, Peyrard M, Bishop AR. 1993 Entropy-driven DNA denaturation. *Phys. Rev. E* **47**, R44. (doi:10.1103/PhysRevE.47.R44)
25. Kwiecinski J, Chapman SJ, Goriely A. 2017 Self-assembly of a filament by curvature-inducing proteins. *Physica D* **344**, 68–80. (doi:10.1016/j.physd.2016.12.001)
26. Kopka ML, Yoon C, Goodsell D, Pjura P, Dickerson RE. 1985 The molecular origin of DNA-drug specificity in netropsin and distamycin. *Proc. Natl Acad. Sci. USA* **82**, 1376–1380. (doi:10.1073/pnas.82.5.1376)
27. Liang X, Purohit PK. 2018 A method to compute elastic and entropic interactions of membrane inclusions. *Extreme Mech. Lett.* **18**, 29–35. (doi:10.1016/j.eml.2017.10.003)
28. Olson WK. 1996 Simulating DNA at low resolution. *Curr. Opin. Struct. Biol.* **6**, 242–256. (doi:10.1016/S0959-440X(96)80082-0)