

## On the continuum formulation for modeling DNA loop formation

Hailong Teng<sup>1,2</sup>, Chung-Hao Lee<sup>1</sup> and Jiun-Shyan Chen\*<sup>1</sup>

<sup>1</sup>*Civil & Environmental Engineering Department, University of California, Los Angeles (UCLA), Los Angeles, USA*

<sup>2</sup>*Livermore Software Technology Company (LSTC), Livermore, USA*

*(Received April 28, 2011, Accepted June 6, 2011)*

**Abstract.** Recent advances in scientific computing enable the full atomistic simulation of DNA molecules. However, there exists length and time scale limitations in molecular dynamics (MD) simulation for large DNA molecules. In this work, a two-level homogenization of DNA molecules is proposed. A wavelet projection method is first introduced to form a coarse-grained DNA molecule represented with superatoms. The coarsened MD model offers a simplified molecular structure for the continuum description of DNA molecules. The coarsened DNA molecular structure is then homogenized into a three-dimensional beam with embedded molecular properties. The methods to determine the elasticity constants in the continuum model are also presented. The proposed continuum model is adopted for the study of mechanical behavior of DNA loop.

**Keywords:** multiscale homogenization; wavelet projection method; coarse-graining; DNA molecules; molecular dynamics.

---

### 1. Introduction

Over the past two decades, rapid development of experimental approaches of single molecule manipulation of DNA molecules has offered researchers the opportunity to control and manipulate DNA molecules and to directly measure the basic physical properties in biological reactions (Cluzel *et al.* 1996, Clausen-Schaumann *et al.* 2000, Smith *et al.* 1992, Smith *et al.* 1996). However, relatively less work has been successfully conducted to fully understand the protein-mediated processes of DNA, such as DNA loop formation, *in vivo* and *in vitro* (Dunn *et al.* 1984, Edelman *et al.* 2003, Krämer *et al.* 1987, Müller *et al.* 1996). Alternatively, mathematical and computational modeling techniques provide the means to study the interactions between DNA and protein/enzymes in DNA loop formation processes.

The DNA loop formation induced by the binding of proteins at DNA sites plays the central role in many biological processes and living organisms to control gene expression, such as transcriptional regulation (RNA synthesis), replication, repression, and recombination (Adhya 1989,

---

\* Corresponding author, Professor, E-mail: [jschen@seas.ucla.edu](mailto:jschen@seas.ucla.edu)

Schleif 1992). Typical examples of protein-mediated DNA looping are the *gal*, *ara*, *lac*, and *deo* operons in *Escherichia (E.) coli*, the  $\lambda$  repressor-operator system, the eukaryotic enhancers, and the transcription of the human  $\beta$ -globin locus (Adhya 1989, Crossley and Orkin 1993, Serfling *et al.* 1985). In the DNA looping process, the structural and mechanical properties of DNAs govern the optimal spacing of operators for functionality. The distance between operators determines the length of DNA loop and it also affects the repression level through the changes in the looping free energy (Coleman and Swigon 2004, Saiz *et al.* 2005).

Wadati and Tsuru (1986) employed the Kirchhoff elastic thin rod model to study the topological and energetic features of the looped DNA with various linking numbers  $Lk$ . This model was also applied by Marko and Siggia (1994) to explain how the competition between the thermal fluctuation and the elastic energy determines the supercoiling radius and pitch. Purohit and Nelson (2006) investigated the supercoiling effect on the protein-mediated DNA loops via the elastic thin rod model, and numerically calculated the looped free energies versus DNA contour lengths at three different levels of DNA supercoiling. The effects of non-uniformly distributed intrinsic curvatures (Garrivier and Fourcade 2000), the base-pair sequence-dependence (Wadati and Tsuru 1986), the salt-dependent mechanical properties (White *et al.* 1999), and the self-contact effects (Chamekh *et al.* 2009, Coleman and Swigon 2004), have been taken into account in this elastic rod model. Balaeff *et al.* (1999, 2004, 2006) adopted the elastic rod model to examine the DNA loop formation during the interaction with the *lac* repressor in *E. coli*, and this computational framework has been applied to the multi-scale studies of protein-DNA complexes (Villa *et al.* 2004).

In this work, we construct a continuum beam formulation with embedded intrinsic molecular properties based on a two-level homogenization by considering a multi-scale wavelet type coarse-graining and a Cauchy-Born type homogenization under the three-dimensional thin rod framework. The content of this paper is organized as follows. The biological aspects of DNA loop formation is introduced in Section 2. The formulation of the proposed hyperelastic continuum model for DNAs based on the two-level hierarchical homogenization is presented in Section 3. In this section, the wavelet projection method for construction of potential functions for the coarse-grained DNA molecule and its homogenization into hyperelastic continuum model based on the Cauchy-Born rule are discussed. In Section 4, we review the three-dimensional thin rod theory which serves as the foundation of our continuum formulation of DNA molecules. In Section 5, numerical studies of the mechanical behavior of DNA molecules, such as the stretching of DNA molecule and DNA loop formation in the *lac* repressor system, are performed. Concluding remarks are outlined in Section 6.

## 2. Biological aspects of DNA loop formation

Protein-mediated DNA loops play a significant role in many biological processes, such as regulation of prokaryotic and eukaryotic gene expression, DNA replication, site-specific DNA recombination, and chromosome segregation (Adhya 1989, Crossley and Orkin 1993, Purohit and Nelson 2006, Schleif 1992). In the genetic system, such as the *lac* repressor in *E. coli*, the transcription can be blocked by the bindings of proteins at two distinct sites of specific gene sequences of DNA molecule which leads to the formation of a DNA loop as illustrated in Fig. 1. The sizes of DNA looping may vary from fifty to several thousands of base pairs (20~700 nm). The two binding sites of the *gal* repressor are at a distance of 113 base pairs (bp); 76 bp (or 385 bp) between binding sites  $O_1$  and  $O_3$  (or between  $O_1$  and  $O_2$ ) in the lactose repressing, as depicted in

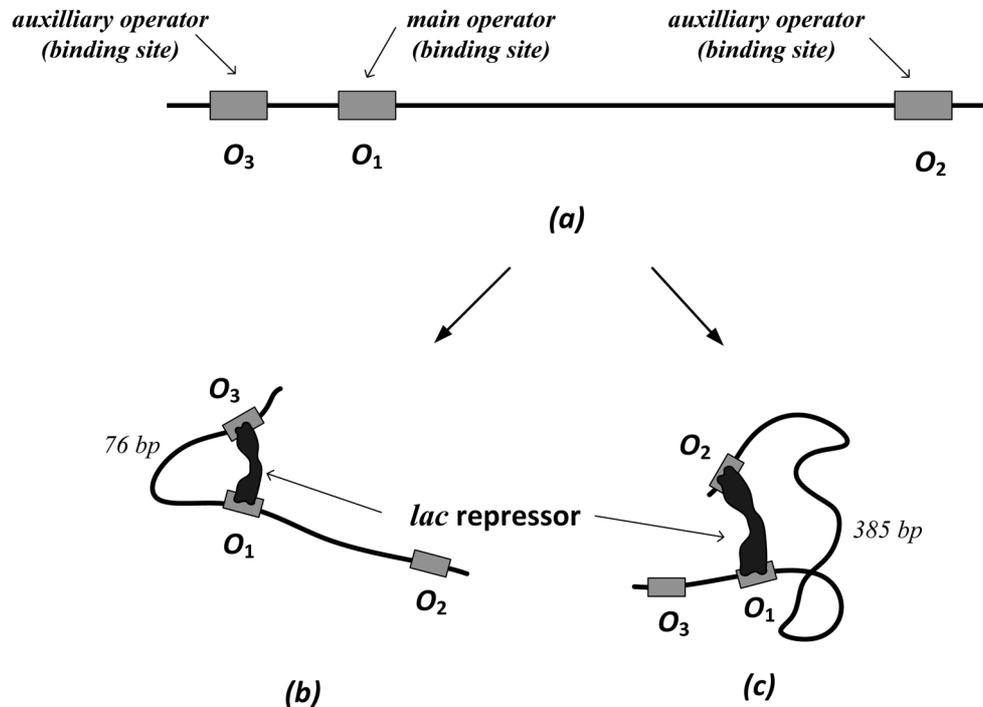


Fig. 1 Schematic representation of the *lac* operon and binding sites of DNA loops. (a) DNA of *E. coli*, shown as thick black segment, and locations of the main ( $O_1$ ) and the two auxiliary ( $O_2$  and  $O_3$ ) operators, (b) and (c) simultaneous binding of the *lac* repressor to the main operator and one of the two auxiliary operators which inhibits the transcription of genes by mRNA.

Figs. 1(b) and 1(c), and the  $\lambda$  repressor switch binding operation can form a DNA loop with a sequence of roughly 2300 bp (Allemand *et al.* 2006). Depending on the specific protein and DNA sequences involved, the DNA loop can affect the transcription by either preventing or enhancing the RNA polymerase binding to the promoter (Schleif 1992). Moreover, the DNA length and its mechanical properties, such as bending and torsional stiffness, have influence on the bio-molecular activities while forming the protein-mediated DNA loops.

### 3. Two-level homogenization of DNA molecules

The proposed homogenization from a atomistic model to a continuum model can be viewed as a discrete to continuum homogenization. The discrete-continuum coupling method has been proposed for extracting constitutive properties from discrete microstructures (Rojek and Oñate 2008) or from microstructures with discontinuities (de Borst *et al.* 2008). The present continuum formulation for DNA molecules is developed based on a two-level homogenization framework. The wavelet based coarse-grained DNA model which has been proposed by Chen *et al.* (2007, 2011) is first introduced as the 1<sup>st</sup> level homogenization (c.f. Figs. 2(a) and 2(b)). The coarse-grained DNA model is then used as the basis model for the construction of a continuum model for the following reasons: (a) the

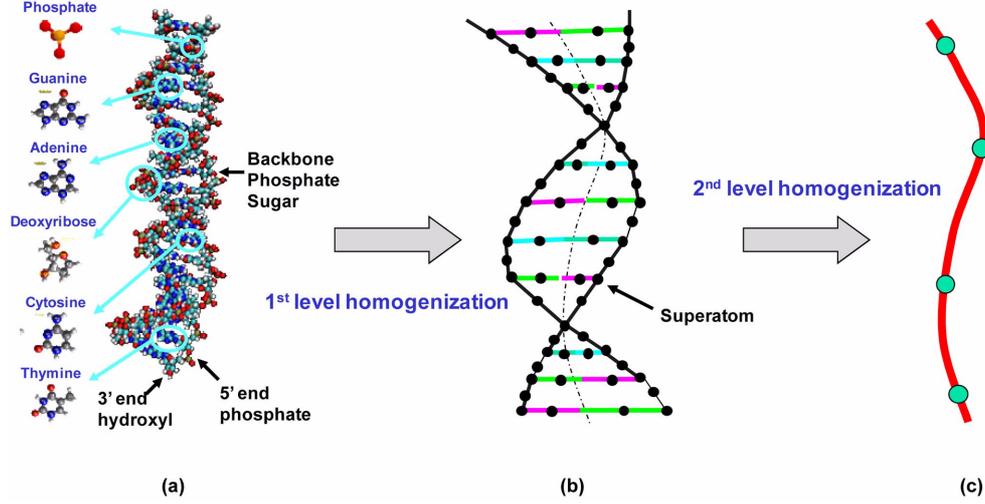


Fig. 2 Schematic of the multi-scale two-level homogenization of DNA molecules, (a) full atomistic DNA model and its basic building blocks, (b) coarse-grained DNA model, and (c) 3D nonlinear elasticity beam model.

coarse-grained model contains much fewer degrees of freedom, (b) the coarsen DNA model is a more regular molecular structure compared to the original molecular structure making it easier to transform to a continuum model, and (c) the potential functions are smoother in the coarse-grained model which allows a coarser temporal discretization as well. In the second level homogenization, as illustrated in Figs. 2(b) and 2(c), the coarse-grained DNA model is utilized to construct a three-dimensional beam for continuum modeling of DNAs.

### 3.1 Overview of multi-scale coarse-graining of DNAs based on the wavelet projection method – 1<sup>st</sup> level homogenization

The multi-scale homogenization based on the wavelet projection method (Mehraeen and Chen 2004) has been applied to construct the coarse-grained DNA model (Chen *et al.* 2007, 2011) as depicted in Figs. 2(a) and 2(b). The key step of the wavelet-based multi-scale homogenization (Chui 1992, Daubechies 1992) is the wavelet projection to transform functions at different scales via the projection matrices,  $\mathbf{P}_j$  and  $\mathbf{Q}_j$

$$\begin{cases} \mathbf{P}_j: V_{j+1} \rightarrow V_j \\ \mathbf{Q}_j: V_{j+1} \rightarrow W_j \end{cases} \quad (1)$$

such that

$$\begin{bmatrix} \mathbf{Q}_j \\ \mathbf{P}_j \end{bmatrix} \mathbf{U}_{j+1} = \begin{bmatrix} \mathbf{U}_{j+1}^{high} \\ \mathbf{U}_{j+1}^{low} \end{bmatrix} \quad (2)$$

where the function  $\mathbf{U}_{j+1}$  at scale  $j+1$  is separated into a low-scale component  $\mathbf{U}_{j+1}^{low} = \mathbf{U}_j$  as the

projection onto  $V_j$  space, and a high-scale component  $U_{j+1}^{high}$  as the projection onto  $W_j$  space, and  $V_j$  and  $W_j$  are the subspaces spanned by the scaling functions  $\varphi_{j,k}(x)$  and wavelet functions  $\psi_{j,k}(x)$ , respectively.

As illustrated in Fig. 2, the coarse-grained *superatom* is chosen to represent the center of mass of each building block. Therefore, via wavelet projection of fine-scale potentials, the effective (coarse-grained) interaction potentials or *force fields*  $\bar{V}$  among these superatoms can be expressed in terms of the valence and non-bonded potentials as follows

$$\bar{V}(\mathbf{r}) = \bar{V}_b(\mathbf{r}) + \bar{V}_{nb}(\mathbf{r}) \quad (3)$$

$$\bar{V}_b(\mathbf{r}) = \sum_{bonds} \bar{k}_b (r_{ij} - \bar{r}_{ij}^{eq})^2 + \sum_{angles} \bar{k}_\theta (\theta_{ijk} - \bar{\theta}_{ijk}^{eq})^2 \quad (4)$$

$$\bar{V}_{nb}(\mathbf{r}) = \sum_i \sum_{j \neq i} \left\{ 4\bar{\varepsilon} \left[ \left( \frac{\bar{\sigma}}{r_{ij}} \right)^{12} - \left( \frac{\bar{\sigma}}{r_{ij}} \right)^6 \right] + \frac{q_i q_j}{\epsilon r_{ij}} \right\} \quad (5)$$

where  $r_{ij} = |\mathbf{r}_i - \mathbf{r}_j|$  is the distance between superatoms  $i$  and  $j$ ,  $\bar{V}_b$  and  $\bar{V}_{nb}$  denotes the effective (coarse-grained) bonded and non-bonded potential functions, respectively,  $\bar{r}_{ij}^{eq}$  and  $\bar{\theta}_{ijk}^{eq}$  are the effective equilibrium distance and angle of the coarse-grained model, respectively,  $\bar{k}_b$  and  $\bar{k}_\theta$  are the effective force constants, and  $\bar{\varepsilon}$  and  $\bar{\sigma}$  are the effective parameters of the van der Waals interaction for the coarse-grained DNA model.

These parameters of the effective force fields for the coarse-grained DNA model are characterized after performing the multi-scale wavelet projection of the fine-scale solutions obtained from the full MD simulation onto a coarse-scale space. The procedures of the wavelet-based hierarchical homogenization for constructing the coarse-grained DNA model are summarized in Table 1. The detailed procedures and results on the characterization of effective potential parameters are not presented in this paper and can be found in Chen *et al.* (2007, 2011).

### 3.2 Continuum description of DNA deformations – 2<sup>nd</sup> level homogenization

In this level of homogenization, we select each turn of DNA double helix as a representative cell, which contains about 10-11 base pairs, and adopt the classical Cauchy-Born rule (Born and Huang 1954, Ericksen 1984) to describe the continuum deformation as

Table 1 Characterization of effective force fields for the coarse-grained model based on the multi-scale wavelet projection method

Step 1	Perform a full MD simulation of DNA stretching in the $NVT$ ensemble*
Step 2	Obtain the probability distributions of distance and angle functions based on the fine-scale simulation results
Step 3	Transform the probability distributions to the fine-scale inter-atomic potentials
Step 4	Homogenize the fine-scale potentials via the multi-scale wavelet projection method
Step 5	Characterize parameters of the effective force fields
Step 6	Repeat Steps 1-5 for different combinations of bonds, bond angles, and non-bonded pairs

\* In the  $NVT$  ensemble, the moles ( $N$ ), volume ( $V$ ), and temperature ( $T$ ) are conserved. The temperature  $T$  at 310 K is considered in the simulation of stretching double-stranded (ds-) DNA.

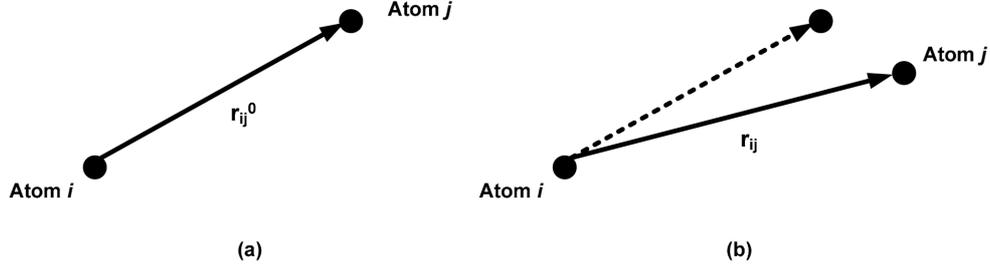


Fig. 3 Schematic of orientations of a bond vector in the (a) initial (undeformed) configuration, and (b) current (deformed) configuration

$$\mathbf{r}_{ij} = \mathbf{F} \cdot \mathbf{r}_{ij}^0 \quad (6)$$

where  $\mathbf{r}_{ij}$  and  $\mathbf{r}_{ij}^0$  are bond vectors connecting superatoms  $i$  and  $j$  in the deformed and undeformed (reference) configurations, respectively,  $\mathbf{F} = \partial \mathbf{x} / \partial \mathbf{X}$  denotes the deformation gradient of a material point in a continuum, and  $\mathbf{X}$  and  $\mathbf{x}$  are the position vectors in the undeformed and the deformed configurations, respectively.

We start with the two-body bond stretching energy given as (Fig. 3)

$$V_{bond} = \sum_{bonds} k_b (r_{ij} - r_{ij}^0)^2 \quad (7)$$

Define a unit vector connecting two atoms

$$\mathbf{N}^{ij} = \frac{\mathbf{r}_{ij}^0}{r_{ij}^0}, \quad \mathbf{r}_{ij}^0 = \mathbf{X}^{(i)} - \mathbf{X}^{(j)}, \quad r_{ij}^0 = |\mathbf{r}_{ij}^0| \quad (8)$$

Similarly,  $\mathbf{r}_{ij} = \mathbf{x}^{(i)} - \mathbf{x}^{(j)}$ ,  $r_{ij} = |\mathbf{r}_{ij}|$ . For simplicity, we omit the superscripts “ $i, j$ ” in the rest of derivation. Under deformation, the stretch ratio of the bond can be expressed by the Green-Lagrangian strain tensor  $\mathbf{E}$  as

$$\frac{r}{r_0} = \sqrt{\mathbf{N} \cdot \mathbf{C} \cdot \mathbf{N}} = \sqrt{1 + 2\mathbf{N} \cdot \mathbf{E} \cdot \mathbf{N}} \quad (9)$$

where  $\mathbf{C} = \mathbf{F}^T \mathbf{F} = \mathbf{I} + 2\mathbf{E}$  is the right-Cauchy deformation tensor. Therefore, the strain energy density for the two-body bond stretching potential  $w_{bond}$  can be expressed as

$$w_{bond} = \frac{1}{V_0} V_{bond} = \frac{1}{V_0} \sum_{bonds} k_b (r_0)^2 (\sqrt{1 + 2\mathbf{N} \cdot \mathbf{E} \cdot \mathbf{N}} - 1)^2 \quad (10)$$

where  $V_0$  is the volume of the representative cell in the undeformed configuration. It can be seen from Eq. (10) that the energy density extracted from the atomistic potential contains the continuum measures such as  $\mathbf{E}$  and  $\mathbf{N}$ , the atomistic potential constant such as  $k_b$ , and the atomistic structure information such as  $r_0$ . Eq. (10) essentially provides a continuum strain energy description with atomistic information included. We then assume a hyperelasticity material behavior for describing the deformation of DNA molecule, where the corresponding stress-strain relationship based on the atomistic potential can be obtained by

$$\mathbf{S}_{bond} = \frac{\partial w_{bond}}{\partial \mathbf{E}} = \frac{1}{V_{0bonds}} \sum k_b(r_0)^2 \left(1 - \frac{1}{1 + 2\mathbf{N} \cdot \mathbf{E} \cdot \mathbf{N}}\right)^2 \mathbf{M} \quad (11)$$

where  $\mathbf{S}_{bond}$  is the 2<sup>nd</sup> Piola-Kirchhoff stress associated with the bond potential, and

$$\mathbf{M} \equiv \begin{bmatrix} N_1^2 & N_1 N_2 & N_1 N_3 \\ N_1 N_2 & N_2^2 & N_2 N_3 \\ N_1 N_3 & N_2 N_3 & N_3^2 \end{bmatrix} \quad (12)$$

Here  $N_i$  is the  $i^{th}$  component of the initial bond orientation  $\mathbf{N}$ .

Similarly, we obtain the continuum descriptions and stress-strain relations associated with three-body and four-body potentials as

3-body potential

$$w_{angle} = \frac{V_{angle}}{V_0} = \frac{1}{V_{0angles}} \sum k_\theta (\theta - \theta_0)^2 \quad (13)$$

$$\mathbf{S}_{angle} = \frac{\partial w_{angle}}{\partial \mathbf{E}} = \frac{2}{V_{0angles}} \sum k_\theta (\theta - \theta_0) \frac{\partial \theta}{\partial \mathbf{E}} \quad (14)$$

4-body potential

$$w_{dihedral} = \frac{V_{dihedral}}{V_0} = \frac{1}{V_{0dihedral}} \sum \frac{V_n}{2} (\phi - \phi_0)^2 \quad (15)$$

$$\mathbf{S}_{dihedral} = \frac{\partial w_{dihedral}}{\partial \mathbf{E}} = \frac{1}{V_{0dihedral}} \sum V_n (\phi - \phi_0) \frac{\partial \phi}{\partial \mathbf{E}} \quad (16)$$

where

$$\frac{\partial \theta}{\partial \mathbf{E}} = -\frac{1}{\sqrt{(\mathbf{N}^{(1)} \cdot \mathbf{C} \cdot \mathbf{N}^{(1)})(\mathbf{N}^{(2)} \cdot \mathbf{C} \cdot \mathbf{N}^{(2)}) - (\mathbf{N}^{(1)} \cdot \mathbf{C} \cdot \mathbf{N}^{(2)})^2}} \left\{ 2\mathbf{N}^{(1)} \mathbf{N}^{(2)T} - \left( \frac{\mathbf{N}^{(1)} \cdot \mathbf{C} \cdot \mathbf{N}^{(2)}}{\mathbf{N}^{(1)} \cdot \mathbf{C} \cdot \mathbf{N}^{(1)}} \right) \mathbf{N}^{(1)} \mathbf{N}^{(1)T} - \left( \frac{\mathbf{N}^{(1)} \cdot \mathbf{C} \cdot \mathbf{N}^{(2)}}{\mathbf{N}^{(2)} \cdot \mathbf{C} \cdot \mathbf{N}^{(2)}} \right) \mathbf{N}^{(2)} \mathbf{N}^{(2)T} \right\} \quad (17)$$

$$\frac{\partial \phi}{\partial \mathbf{E}} = \frac{-2}{\sqrt{(1 - \cos^2 \phi) \Delta}} \left\{ (\mathbf{N}^{(2)} \cdot \mathbf{C} \cdot \mathbf{N}^{(3)}) \mathbf{N}^{(1)} \mathbf{N}^{(2)T} + (\mathbf{N}^{(1)} \cdot \mathbf{C} \cdot \mathbf{N}^{(2)}) \mathbf{N}^{(2)} \mathbf{N}^{(3)T} \right. \\ \left. - (\mathbf{N}^{(2)} \cdot \mathbf{C} \cdot \mathbf{N}^{(2)}) \mathbf{N}^{(1)} \mathbf{N}^{(3)T} + (\mathbf{N}^{(1)} \cdot \mathbf{C} \cdot \mathbf{N}^{(3)}) \mathbf{N}^{(2)} \mathbf{N}^{(2)T} \right.$$

$$\left. - \frac{1}{2\delta} [(\mathbf{N}^{(1)} \cdot \mathbf{C} \cdot \mathbf{N}^{(2)})(\mathbf{N}^{(1)} \cdot \mathbf{C} \cdot \mathbf{N}^{(3)}) - (\mathbf{N}^{(1)} \cdot \mathbf{C} \cdot \mathbf{N}^{(3)})(\mathbf{N}^{(2)} \cdot \mathbf{C} \cdot \mathbf{N}^{(2)})] \frac{\partial \delta}{\partial \mathbf{C}} \right\} \quad (18)$$

$$\begin{aligned}
\delta = & (\mathbf{N}^{(1)} \cdot \mathbf{C} \cdot \mathbf{N}^{(1)})(\mathbf{N}^{(2)} \cdot \mathbf{C} \cdot \mathbf{N}^{(2)})^2(\mathbf{N}^{(3)} \cdot \mathbf{C} \cdot \mathbf{N}^{(3)}) \\
& - (\mathbf{N}^{(1)} \cdot \mathbf{C} \cdot \mathbf{N}^{(1)})(\mathbf{N}^{(2)} \cdot \mathbf{C} \cdot \mathbf{N}^{(2)})(\mathbf{N}^{(2)} \cdot \mathbf{C} \cdot \mathbf{N}^{(3)})^2 \\
& - (\mathbf{N}^{(1)} \cdot \mathbf{C} \cdot \mathbf{N}^{(2)})^2(\mathbf{N}^{(2)} \cdot \mathbf{C} \cdot \mathbf{N}^{(2)})(\mathbf{N}^{(3)} \cdot \mathbf{C} \cdot \mathbf{N}^{(3)}) \\
& + (\mathbf{N}^{(1)} \cdot \mathbf{C} \cdot \mathbf{N}^{(2)})^2(\mathbf{N}^{(2)} \cdot \mathbf{C} \cdot \mathbf{N}^{(3)})^2
\end{aligned} \tag{19}$$

$$\begin{aligned}
\frac{\partial \delta}{\partial \mathbf{C}} = & [(\mathbf{N}^{(2)} \cdot \mathbf{C} \cdot \mathbf{N}^{(2)})^2(\mathbf{N}^{(3)} \cdot \mathbf{C} \cdot \mathbf{N}^{(3)}) - (\mathbf{N}^{(2)} \cdot \mathbf{C} \cdot \mathbf{N}^{(2)})(\mathbf{N}^{(2)} \cdot \mathbf{C} \cdot \mathbf{N}^{(3)})^2] \mathbf{N}^{(1)T} \mathbf{N}^{(1)} \\
& + [2(\mathbf{N}^{(1)} \cdot \mathbf{C} \cdot \mathbf{N}^{(1)})(\mathbf{N}^{(2)} \cdot \mathbf{C} \cdot \mathbf{N}^{(2)})(\mathbf{N}^{(3)} \cdot \mathbf{C} \cdot \mathbf{N}^{(3)}) - \\
& (\mathbf{N}^{(1)} \cdot \mathbf{C} \cdot \mathbf{N}^{(2)})^2(\mathbf{N}^{(3)} \cdot \mathbf{C} \cdot \mathbf{N}^{(3)}) - (\mathbf{N}^{(1)} \cdot \mathbf{C} \cdot \mathbf{N}^{(1)})(\mathbf{N}^{(2)} \cdot \mathbf{C} \cdot \mathbf{N}^{(3)})^2] \mathbf{N}^{(2)T} \mathbf{N}^{(2)} \\
& + [(\mathbf{N}^{(1)} \cdot \mathbf{C} \cdot \mathbf{N}^{(1)})(\mathbf{N}^{(2)} \cdot \mathbf{C} \cdot \mathbf{N}^{(2)})^2 - (\mathbf{N}^{(1)} \cdot \mathbf{C} \cdot \mathbf{N}^{(2)})^2(\mathbf{N}^{(2)} \cdot \mathbf{C} \cdot \mathbf{N}^{(2)})] \mathbf{N}^{(3)T} \mathbf{N}^{(3)} \\
& + [(\mathbf{N}^{(1)} \cdot \mathbf{C} \cdot \mathbf{N}^{(2)})^2(\mathbf{N}^{(2)} \cdot \mathbf{C} \cdot \mathbf{N}^{(3)}) - (\mathbf{N}^{(1)} \cdot \mathbf{C} \cdot \mathbf{N}^{(1)})(\mathbf{N}^{(2)} \cdot \mathbf{C} \cdot \mathbf{N}^{(2)})(\mathbf{N}^{(2)} \cdot \mathbf{C} \cdot \mathbf{N}^{(3)})] \mathbf{N}^{(2)T} \mathbf{N}^{(3)} \\
& + [(\mathbf{N}^{(1)} \cdot \mathbf{C} \cdot \mathbf{N}^{(2)})^2(\mathbf{N}^{(3)} \cdot \mathbf{C} \cdot \mathbf{N}^{(2)}) - (\mathbf{N}^{(1)} \cdot \mathbf{C} \cdot \mathbf{N}^{(1)})(\mathbf{N}^{(2)} \cdot \mathbf{C} \cdot \mathbf{N}^{(2)})(\mathbf{N}^{(3)} \cdot \mathbf{C} \cdot \mathbf{N}^{(2)})] \mathbf{N}^{(3)T} \mathbf{N}^{(2)} \\
& + [(\mathbf{N}^{(1)} \cdot \mathbf{C} \cdot \mathbf{N}^{(2)})(\mathbf{N}^{(2)} \cdot \mathbf{C} \cdot \mathbf{N}^{(3)})^2 - (\mathbf{N}^{(1)} \cdot \mathbf{C} \cdot \mathbf{N}^{(2)})(\mathbf{N}^{(2)} \cdot \mathbf{C} \cdot \mathbf{N}^{(2)})(\mathbf{N}^{(3)} \cdot \mathbf{C} \cdot \mathbf{N}^{(3)})] \mathbf{N}^{(1)T} \mathbf{N}^{(2)} \\
& + [(\mathbf{N}^{(2)} \cdot \mathbf{C} \cdot \mathbf{N}^{(1)})(\mathbf{N}^{(2)} \cdot \mathbf{C} \cdot \mathbf{N}^{(3)})^2 - (\mathbf{N}^{(2)} \cdot \mathbf{C} \cdot \mathbf{N}^{(1)})(\mathbf{N}^{(2)} \cdot \mathbf{C} \cdot \mathbf{N}^{(2)})(\mathbf{N}^{(3)} \cdot \mathbf{C} \cdot \mathbf{N}^{(3)})] \mathbf{N}^{(2)T} \mathbf{N}^{(1)}
\end{aligned} \tag{20}$$

and  $\mathbf{N}^{(m)}$  is the bond orientation of the bond connecting two superatoms  $m + l$  and  $m$ .

The non-bonded interactions are obtained from the grouped interatomic potentials between the groups of atoms in the same molecule and those groups in other molecules. The general form of the non-bonded energy of the atomistic system can be written as

$$E_{nb} = \sum_i \sum_{j>i} V_{nb}(r_{ij}) \tag{21}$$

where  $V_{nb}$  is the non-bonded potential,  $r_{ij}$  is the distance between superatom groups  $i$  and  $j$ . For DNA molecules, the non-bonded interactions include the van der Waals and Coulomb interactions. Consider interactions of two unit cells with volume  $V_0$ , each contains  $n$  superatoms. The continuum energy density for non-bonded interactions is given as (Arroyo and Belytschko 2004).

$$w_{nb}(r) = \left(\frac{n}{V_0}\right)^2 V_{nb}(r) \equiv \rho_{atom}^2 V_{nb}(r) \tag{22}$$

where  $\rho_{atom} = n / V_0$  denotes the number density of superatoms per representative cell volume  $V_0$ , and  $r$  is the distance between the centroids of two groups of superatoms in the deformed configuration. Finally, the total non-bonded energy is computed by the following integration

$$\Pi_{nb} = \frac{1}{2} (\rho_{atom})^2 \int_{\Omega_0} \int_{\Omega_0 - \Omega_x} V_{nb}(r) d\Omega_0 d\Omega_0 \tag{23}$$

where  $\Omega_X$  is the sphere of region centered at  $\mathbf{X}$  with a cutoff radius to exclude the bonded interactions, and  $\Omega_0$  is the domain of the representative cell.

### 3.3 Hyperelastic continuum model for DNA molecules

Based on hyperelasticity, the material response tensor is first obtained by

$$\mathbf{C} = \frac{\partial^2 w}{\partial \mathbf{E} \partial \mathbf{E}} = \frac{\partial \mathbf{S}}{\partial \mathbf{E}} \quad (24)$$

where  $w = w_{bond} + w_{angle} + w_{dihedral} + w_{nb}$  is the DNA energy density function constructed in the previous section. Note that since the contribution of the 4-body dihedral angle potential to both shear and Young's modulus is only 0.1% of that associated with the bond stretching potential, the 4-body potential is ignored in the potential function for the coarse-grained DNA model. By setting  $F_{ij} = \delta_{ij}$  in  $\mathbf{C}$ , the Young's modulus and shear modulus can be extracted. For pure poly(dG)-poly(dC) sequence ds-DNA, we obtained

$$G = 8.7 \times 10^8 Pa \text{ and } E = 5.1 \times 10^8 Pa \quad (25)$$

The above numerically characterized Young's modulus agrees fairly well with the experimental measured quantities, i.e.  $E = 0.5:0.7 \times 10^8 Pa$  by Gevorkian and Khudaverdian (1990). Note that in the current DNA model we do not consider the solution effect and sequence-dependent variation, the predicted Young's modulus should be regarded as an upper bound. From Eq. (25), we obtain the Poisson's ratio as

$$\nu = \frac{E - 2G}{2G} \approx -0.7 \quad (26)$$

This negative Poisson's ratio is also observed experimentally by Manning (1985) and Baumann *et al.* (1997), with the experimentally measured Poisson's ratios of  $-0.7$  and  $-0.4 \sim 0$ , respectively. To gain a better insight into this mechanical property, consider only the 2-body bond stretching potential for a coarse-grained DNA molecule, which gives the following property

$$\frac{G^{bond}}{E^{bond}} = \frac{1/V_0 \sum_{bonds} k_b r_0^2 (N_2 N_3)^2}{1/V_0 \sum_{bonds} k_b r_0^2 N_3^4} = \frac{\sum_{bonds} N_2^2}{\sum_{bonds} N_3^2} \quad (27)$$

According to the coarse-grained DNA helical structure, it can be shown that  $\sum N_2^2 > \sum N_3^2$ , hence  $G^{bond} > E^{bond}$ , and thus a negative Poisson's ratio. The above characterized elasticity constants are adopted in the three-dimensional beam formulation for the continuum modeling of DNAs as introduced in the next section.

## 4. Three-dimensional beam formulation for continuum modeling of DNAs

As described in the previous section, the string of the coarse-grained DNA molecule is to be described by a three-dimensional beam. A rod can be represented by a smooth curve  $C$  in the three-dimensional space parameterized by the arc-length parameter  $S$ . Let the configuration of a rod be

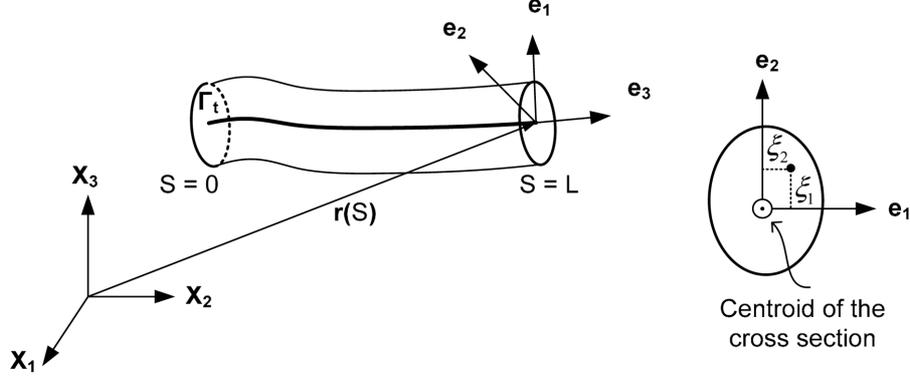


Fig. 4 Geometric representation of 3D elastic rod model

constructed by a varying orthonormal triad  $\{\mathbf{e}_i(S)\}$  ( $\mathbf{e}_1 = \text{normal}$ ,  $\mathbf{e}_2 = \text{binormal}$ ,  $\mathbf{e}_3 = \text{tangent}$ ) as depicted in Fig. 4, with the origins of the frames located at the line of centroid of the rod  $\Gamma_t$ .

The material point at any location of the cross section (on the  $\mathbf{e}_1 - \mathbf{e}_2$  plane) is defined by means of the triplet  $\{\xi_\alpha\}_{\alpha=1}^2$ , where  $S$  is the parametric coordinate of cross section location on the centerline, and  $\{\xi_\alpha\}_{\alpha=1}^2$  represents the position vector of the point away from the centroid on a cross section.

Let  $\mathbf{x}$  be the coordinate of a material point in the deformed configuration. The rod in the deformed configuration is parameterized by  $S$  as

$$\mathbf{x} = \mathbf{r}(S) + \xi_1 \mathbf{e}_1(S) + \xi_2 \mathbf{e}_2(S) = \mathbf{r}(S) + \xi_1 \mathbf{R}(S) \mathbf{E}_1(S) + \xi_2 \mathbf{R}(S) \mathbf{E}_2(S) \quad (28)$$

where  $\mathbf{r}(S)$  is the deformed centerline position vector,  $\mathbf{R}(S)$  is the rotation matrix of the rod's axis which relates the undeformed orthogonal triad  $\{\mathbf{E}_i\}$  to the deformed ones  $\{\mathbf{e}_i(S)\}$ .

The derivative of the orthonormal triad  $\{\mathbf{e}_i(S)\}$  with respect to  $S$  is

$$\mathbf{e}_i'(S) \equiv \frac{d\mathbf{e}_i(S)}{dS} = \mathbf{k}(S) \times \mathbf{e}_i(S) = \mathbf{K}(S) \cdot \mathbf{e}_i(S) \quad (29)$$

where the curvature vector  $\mathbf{k}(S) = k_i(S) \mathbf{e}_i$  contains the centerline curvature components  $k_1(S)$  and  $k_2(S)$ , and the twist component  $k_3(S)$ . The term  $\mathbf{K}$  is the corresponding skew-symmetric tensor which can be expressed in terms of the rotation matrix  $\mathbf{R}$

$$\mathbf{K}(S) = \mathbf{R}'(S) \cdot \mathbf{R}^T(S) \quad (30)$$

In the Reissner rod theory (Reissner 1973), the strain vector  $\boldsymbol{\varepsilon}$  that contains the axial deformation of the centerline and the transverse shear deformations of the cross section can be expressed as

$$\boldsymbol{\varepsilon} = [\gamma_1, \gamma_2, \varepsilon]^T = \mathbf{r}' - \mathbf{e}_3 \quad (31)$$

where  $\mathbf{r}' = d\mathbf{r}/dS$  denotes the derivative of position vector with respect to  $S$ , and  $\gamma_1$ ,  $\gamma_2$ , and  $\varepsilon$  are the shear strains and axial strain of the rod, respectively.

The curvature vector  $\mathbf{k}$  is related to the three Euler angles  $\boldsymbol{\mathcal{G}} = [\theta, \phi, \psi]^T$  by

$$\mathbf{k} - \mathbf{k}_0 = [k_1 - k_1^0, k_2 - k_2^0, k_3 - k_3^0]^T = \boldsymbol{\mathcal{G}}' = [\theta', \phi', \psi']^T \quad (32)$$

where  $k_1^0, k_2^0$  and  $k_3^0$  are the initial curvatures and the initial twist of the rod, respectively.

For an isotropic material with homogeneous circular cross section, the force and moment vectors can then be written as

$$\mathbf{f} = GA\gamma_1\mathbf{e}_1 + GA\gamma_2\mathbf{e}_2 + EA\varepsilon\mathbf{e}_3 \quad (33)$$

$$\mathbf{m} = EI(k_1 - k_1^0)\mathbf{e}_1 + EI(k_2 - k_2^0)\mathbf{e}_2 + GJ(k_3 - k_3^0)\mathbf{e}_3 \quad (34)$$

where  $EA, GA, EI$ , and  $GJ$  are the axial, shearing, bending, and torsional rigidities, respectively,  $E$  and  $G$  are the Young's and shear modulus characterized from the coarse-grained DNA model, respectively, and  $I$  and  $J$  are the moment of inertia and polar moment of inertia, respectively. Eqs. (33) and (34) imply that the strain energy of the elastic rod has the following quadratic form

$$\begin{aligned} \Pi_{int} &= \frac{1}{2} \int_0^L EI(k_1 - k_1^0)^2 + EI(k_2 - k_2^0)^2 + GJ(k_3 - k_3^0)^2 + EA\varepsilon^2 + GA\gamma_1^2 + GA\gamma_2^2 dS \\ &= \frac{1}{2} \int_0^L (\mathbf{k} - \mathbf{k}_0)^T \mathbf{D}_k (\mathbf{k} - \mathbf{k}_0) dS + \frac{1}{2} \int_0^L \boldsymbol{\varepsilon}^T \mathbf{D}_\varepsilon \boldsymbol{\varepsilon} dS \end{aligned} \quad (35)$$

where  $\mathbf{D}_k = \text{diag}(EI, EI, GJ)$ ,  $\mathbf{D}_\varepsilon = \text{diag}(EA, GA, GA)$ , and  $L$  is the contour length of the rod. The work done by the external loading is expressed by

$$\Pi_{ext} = \int_0^L \bar{\mathbf{f}} \cdot \mathbf{d} + \bar{\mathbf{m}} \cdot (\mathbf{k} - \mathbf{k}_0) dS + \sum_{i=1}^n [\mathbf{p}(S_i) \cdot \mathbf{d}(S_i) + \mathbf{t}(S_i) \cdot (\mathbf{k}(S_i) - \mathbf{k}_0)] \quad (36)$$

where  $\mathbf{d} = \mathbf{r} - \mathbf{r}_0$  is the vector of centerline displacements,  $\mathbf{r}_0$  is the undeformed position vector,  $\mathbf{p}(S_i)$  and  $\mathbf{t}(S_i)$  are the concentrated force and moment vectors, respectively, acting on the cross section  $S_i$ , and  $\bar{\mathbf{f}}$  and  $\bar{\mathbf{m}}$  are the distributed force and moment vectors, respectively.

In the present study, we introduce a three-dimensional beam model whose total energy is given by

$$\Pi = \Pi_{int} - \Pi_{ext} \quad (37)$$

where the internal energy and external work are defined in Eqs. (35) and (36), respectively. The elasticity constants are obtained from the homogenized continuum model discussed in Section 3. Given appropriate boundary and loading conditions, a finite element discrete equation can be obtained by taking the stationary condition of Eq. (37), which has the form

$$\begin{aligned} \delta\Pi = 0 &= \int_0^L \delta\mathbf{k}^T \mathbf{D}_k (\mathbf{k} - \mathbf{k}_0) dS + \int_0^L \delta\boldsymbol{\varepsilon}^T \mathbf{D}_\varepsilon \boldsymbol{\varepsilon} dS \\ &- \int_0^L \delta\mathbf{d}^T \bar{\mathbf{f}} + \delta\mathbf{k}^T \bar{\mathbf{m}} dS - \sum_{i=1}^n [\delta\mathbf{d}^T \mathbf{p}(S_i) + \delta\mathbf{k}^T \mathbf{t}(S_i)] \end{aligned} \quad (38)$$

Eq. (38) is generally nonlinear, so standard linearization with the Newton-Raphson method is adopted for solving the incremental displacement and rotation vectors. The finite element approximations of the trial and test functions are introduced as follows

$$\delta\mathbf{d}^h = \sum_{l=1}^{NEN} N_l(S) \delta\mathbf{d}_l$$

$$\begin{aligned}\delta \mathcal{G}^h &= \sum_{I=1}^{NEN} \hat{N}_I(S) \delta \mathcal{G}_I \\ \delta \boldsymbol{\varepsilon}^h &= \delta \mathbf{d}'^h - \delta \mathcal{G}^h \times \mathbf{d}'^h \\ \delta \mathbf{k}^h &= \delta \mathcal{G}'^h\end{aligned}\quad (39)$$

where superscript prime denotes the derivative with respect to  $S$ ,  $N_I$  and  $\hat{N}_I$  are the shape functions associated with node  $I$  for the approximations of the displacement vector and rotational vector, respectively, and  $NEN$  is the number of nodes in each element. In the continuum modeling of DNA molecules as presented in Section 5, a linear approximation with  $C^0$  continuity is introduced for both translational and rotational degrees of freedom, and selective reduced integration, in which reduced integration is used for the shear and membrane energy while full integration is employed for the bending energy, is adopted to resolve the shear and membrane locking resulting from the equal-order interpolations (Hughes and Liu 1981a, 1981b).

## 5. Applications

### 5.1 Simulation of DNA stretching

A segment of ds-DNA molecule with pure G–C sequences as shown in Fig. 5 is considered. The Verlet velocity time integration algorithm with 1 fs time step size and the AMBER (Cornell *et al.* 1996) force field are used in the MD simulation of stretching the DNA molecule. The DNA model is first relaxed to yield the thermal equilibrium position at the temperature of 310 K. We then impose a slightly increased length on the DNA and perform energy minimization accordingly. This allows the DNA to be stretched gradually in equilibrium. The energy-extension and force-extension curves are shown in Figs. 6(a) and 6(b), respectively. The numerical result of force-extension is consistent with the single molecule experiments on DNA molecules under extension (Cluzel *et al.* 1996, Smith *et al.* 1992, 1996). Fig. 7(a) illustrates the progressive deformed configurations during

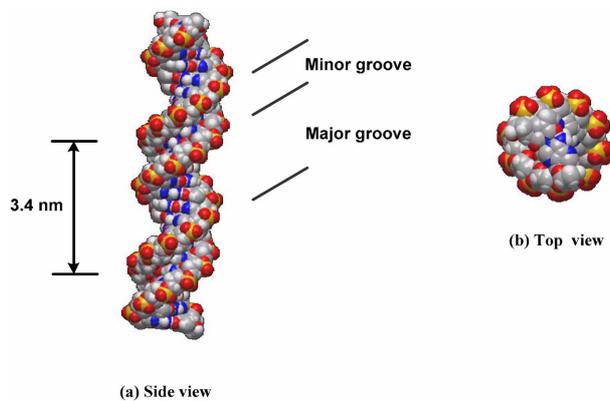


Fig. 5 Molecular structure of the G-C rich double-stranded DNA molecule

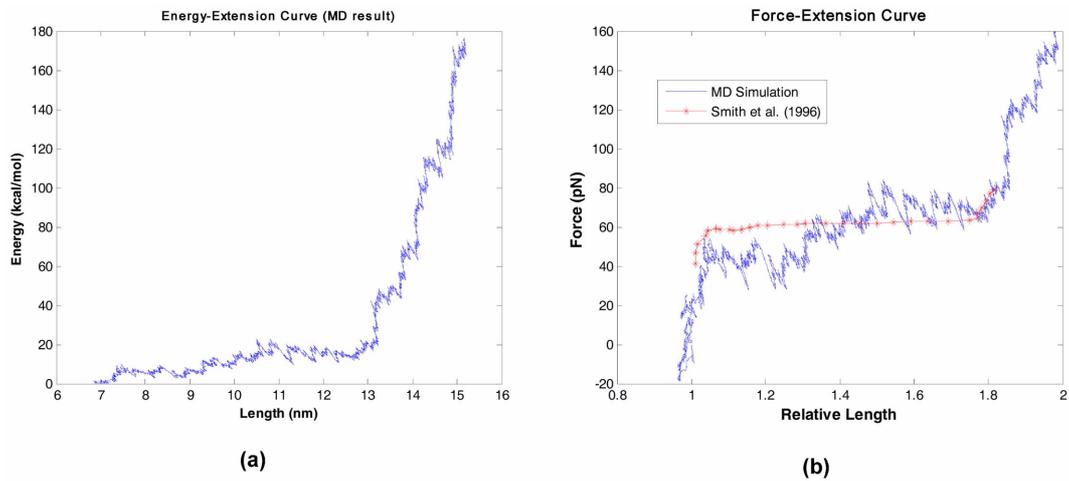


Fig. 6 Full MD simulation results of stretching ds-DNA: (a) energy versus length curve, and (b) force versus relative length curves (MD simulation in blue and experiment by Smith *et al.* in red cross (Smith *et al.* 1996))

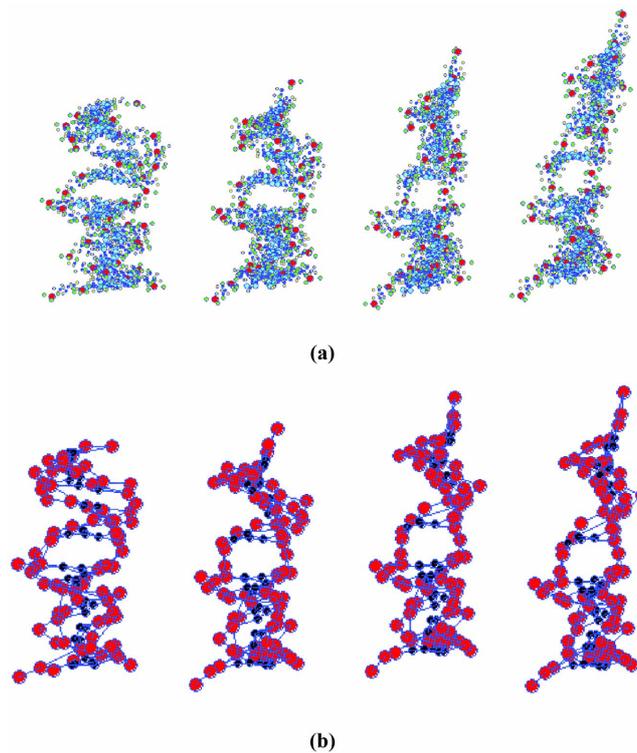


Fig. 7 Progressive DNA configurations of (a) full atomistic DNA model, and (b) coarse-grained DNA model

the stretching process, and the results show that the ds-DNA is unzipped and leads to a significant DNA molecular structural change. This numerically observed ladder form which corresponds to the

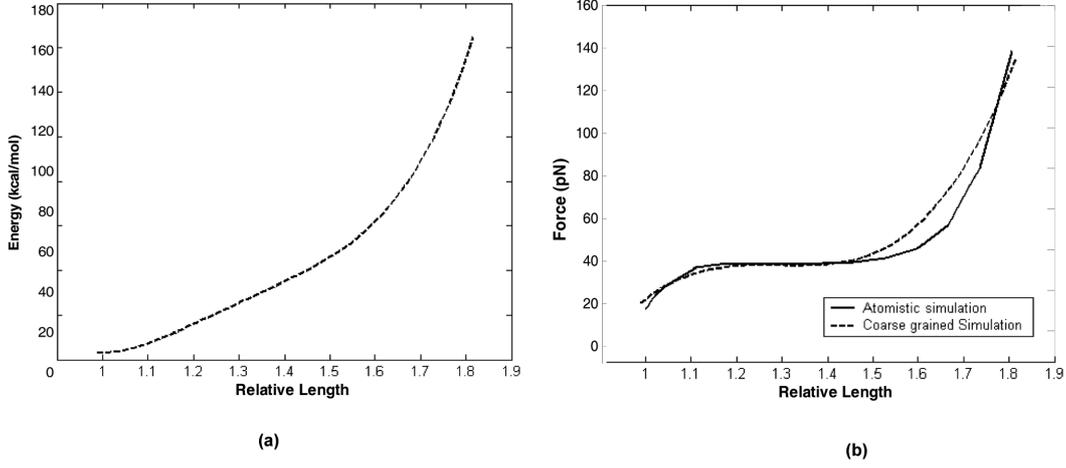


Fig. 8 (a) Predicted energy versus extension curve of the coarse-grained DNA model, and (b) the corresponding predicted force versus extension curve. (Note that full MD results are fit by polynomial functions.)

overstretching transition region in the force-extension curve agrees with the experimental observation (Cluzel *et al.* 1996, Smith *et al.* 1996).

For comparison between the fine-scale and the coarse-grained simulation results, in the following study the solution effect is not considered in the DNA molecular model. The coarse-grained model is constructed following the procedures introduced in Section 3.1 where the corresponding force fields of the coarse-grained DNA model are determined by performing the multi-scale wavelet projection. The corresponding energy-extension and force-extension curves and the deformed configurations are compared in Figs. 8(a), 8(b) and 7(b), respectively, where good results in the coarse-grained DNA model are shown. The homogenized coarse-grained model allows a large time step size and reduces large numbers of degrees of freedom leading to three orders of CPU speed-up.

Next, we model the DNA stretching process by using the proposed continuum beam model through the multi-scale two-level homogenization as illustrated in Sections 3 and 4. To take into account the molecular structural change from B-DNA to S-DNA in the overstretching transition, a continuum strain energy function is proposed as

$$W = \alpha(s)W^1 + (1 - \alpha(s))W^2 \quad (40)$$

where  $s$  is the stretch ratio,  $\alpha(s)$  is a transition function with  $\alpha = 1$  for  $s \leq 1.2$  and  $\alpha < 1$  for  $s > 1.2$ ,  $W^1$  is the strain energy function derived from the coarse-grained B-DNA model, and  $W^2$  is a phenomenological strain energy function representing the strain hardening behavior of the S-form DNA after overstretching transition (Gent 1996).

$$W^2(I_1) = -\frac{G}{2}J_m \ln\left(1 - \frac{I_1 - 3}{J_m}\right) \quad (41)$$

where  $G$  is the shear modulus,  $I_1$  is the principal invariant, and  $J_m > 0$  is a parameter taking into account the limiting molecule extensibility. The values of  $J_m = 1.88$  (Ogden *et al.* 2008) and  $G = 8.7 \times 10^8$  Pa are employed in the continuum modeling. Fig. 9 shows the effectiveness of the proposed continuum model.

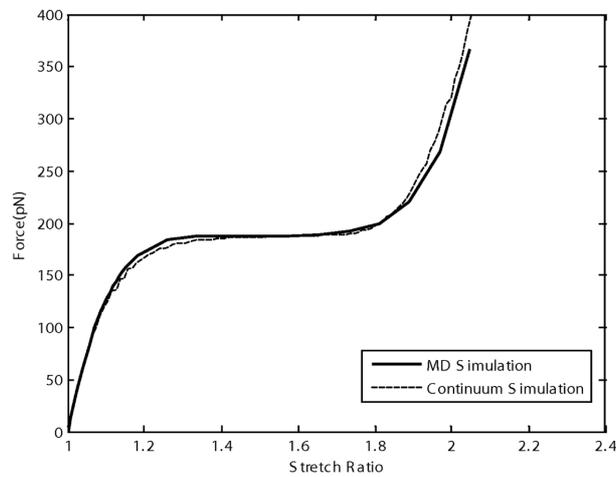


Fig. 9 Comparison of force-extension curve between MD simulation and the continuum model

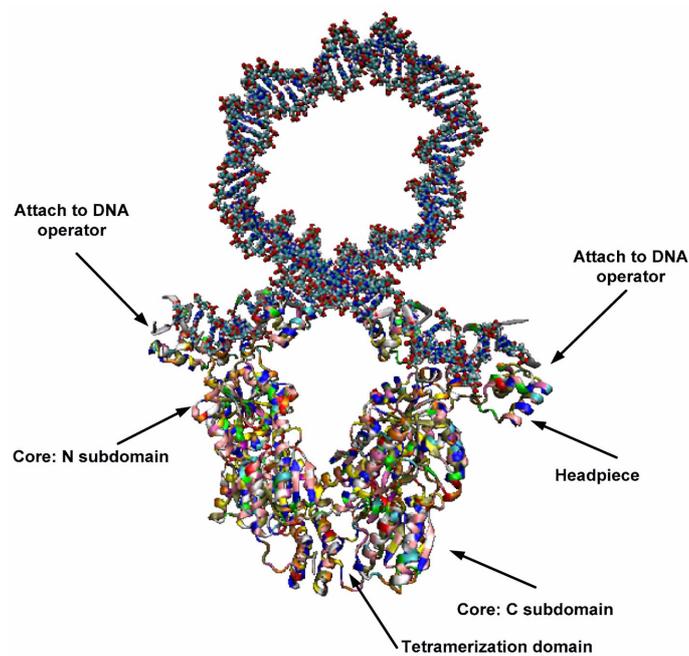


Fig. 10 Crystal structure of the *lac* repressor binding to DNA molecule<sup>4</sup>

### 5.2 Modeling of DNA loop formation

Loop formation of DNA molecule is an important protein-DNA interaction process in many biological systems. We apply the proposed three-dimensional beam to the simulation of the DNA

<sup>4</sup>LBI pdb file for the lac repressor is obtained from Protein Data Bank.

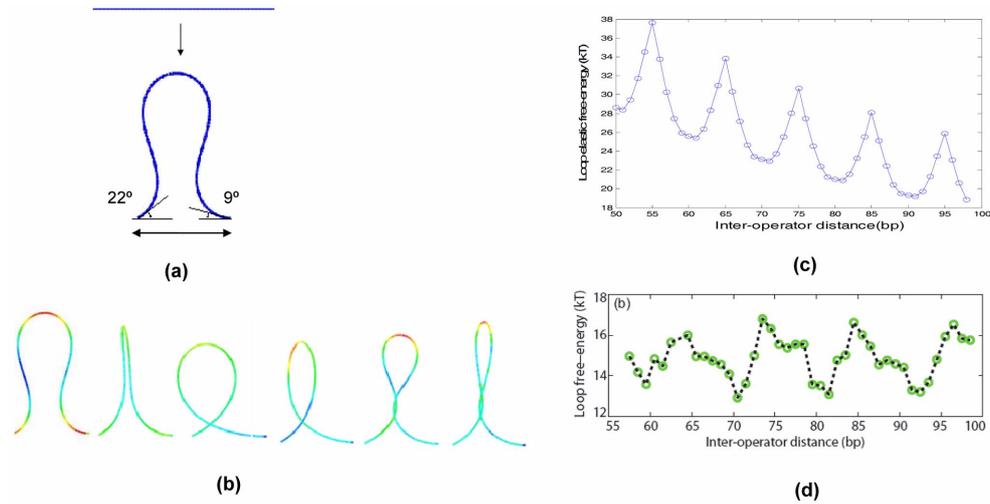


Fig. 11 Continuum modeling of DNA loop formation (a) boundary conditions for the continuum beam model, (b) configurations of DNA looping, (c) numerically calculated elastic energy versus DNA length, and (d) experimentally measured loop free-energy versus DNA length (Purohit and Nelson 2006)

looping formed by a sequence-specific DNA-binding protein, the *lac* repressor (*LacI*), which inhibits the gene coding in the metabolism of lactose in bacteria. The *lac* repressor functions as a tetramer, where the four subunits form a V-shaped molecule as depicted in Fig. 10. The *lac* repressor protein has three distinct regions: (a) a core domain that binds to the lactose and other similar molecules, which is divided into *N* and *C* subdomains, (b) a tetramerization domain joining four monomers in an alpha-helix bundle, and (c) a headpiece domain in which two *LacI* proteins bind to the DNA.

The *LacI*-DNA system in solution contains millions of atoms with time scales spanning from microseconds to milliseconds, and a full atomistic simulation is practically unaffordable to investigate the DNA looping behaviors. Here, the proposed continuum descriptions are employed to simulate the configuration and to compute the free energy of the DNA looping processes.

In this study, we construct a three-dimensional nonlinear elasticity beam model of the DNA molecule composed of 76 base pairs long with repeating G-C sequences. The elasticity constants are extracted from the multi-scale atomistic-continuum model introduced in Section 3.3, and the DNA is modeled by 100 Hughes-Liu beam elements (Hughes and Liu 1981) in DYNA3D (Whirley and Engelmann 1993). The *lac* repressor is treated as a rigid coupler and boundary conditions of the beam are obtained from the crystal structure of the *lac* repressor DNA complex (Lewis *et al.* 1996). In the simulation, the two ends of the beam are first moved to their predefined positions as illustrated in Fig. 11(a), and the two ends are bent to satisfy the above-mentioned boundary conditions. The right end of the beam is then rotated around its axial axis to investigate the DNA loop process.

The simulated configurations shown in Fig. 11(b) are consistent with Balaeff *et al.*'s work (Balaeff *et al.* 1999). The calculated elastic energy presented in Fig. 11(c) agrees reasonably well with the experimentally measured free energy of the DNA looping (Purohit and Nelson 2006). The predicted periodicity of the free energy, which is consistent with the helical periodicity of DNA molecules, is qualitatively captured by the continuum beam model. Since the protein-DNA interaction effect is not considered in the numerical modeling, it results in the deviation of the calculated elastic energy

from the experimentally measured free energy for short DNA molecules. This will be improved in the future by including the protein-DNA interactions while treating the *lac* repressor protein as flexible bodies.

## 6. Conclusions

In this work, a continuum beam formulation of DNAs with embedded molecular properties has been developed through the proposed two-level homogenization process. In the first level homogenization, a wavelet projection approach was introduced to project the fine scale potential function to a coarser scale for the superatoms representing the coarse-grained DNA molecules. The coarsened molecules are then used as the reference in the second level homogenization using continuum mechanics deformation measures to describe the motion of the coarsen molecules. Under this construction, the strain energy density functions and stress-strain relationships were derived from the coarse-grained DNA model based on a potential equivalence. The proposed multi-scale DNA model enables systematic characterization of the fundamental mechanical properties of DNAs and offers an effective computational framework for continuum modeling of DNA molecules. The numerically characterized elasticity constants of DNA molecules agree well with the experimental measurements. The calculated negative Poisson's ratio has been shown as the consequence of the unique geometry of DNA molecules. These materials constants were also adopted in the three-dimensional beam formulation for the applications of continuum modeling of DNAs.

In this study, stretching of DNA molecules using the full MD simulation, coarse-grained simulation, and continuum modeling have been performed and compared. The predicted force-extension relation and the overall deformation response have been shown to have good agreement between the homogenized continuum and the fine-scale MD models. Combined with the significant reduction of the degrees of freedom in the spatial and temporal domains, the continuum DNA model is demonstrated to be effective for modeling DNA loop formation. The numerically predicted loop formation energies agree reasonably with the experimentally measured data for long DNA molecules. Nevertheless, solution effects and DNA-protein interactions need to be considered for further improvement of the current formulation to capture the more complicated responses in the interaction between short DNAs and the *lac* repressor systems. The proposed computational framework has been shown to be effective for modeling DNA-protein interactions, and it also provides the means to better understand the protein-mediated biological processes *in vivo* and *in vitro*.

## Acknowledgements

The support of UCLA Dissertation Year Fellowship to the second author is greatly acknowledged.

## References

- Adhya, S. (1989), "Multipartite genetic control elements: communication by DNA loop", *Ann. Rev. Genetics*, **23**(1), 227-250.

- Allemand, J.F., Cocco, S., Douarche, N. and Lia, G. (2006), "Loops in DNA: an overview of experimental and theoretical approaches", *Eur. Phys. J. E*, **19**, 293-302.
- Arroyo, M. and Belytschko, T. (2004), "Finite element methods for the non-linear mechanics of crystalline sheets and nanotubes", *Int. J. Numer. Meth. Eng.*, **59**, 419-456.
- Balaeff, A., Koudella, C.R., Mahadevan, L. and Schulten, K. (2004), "Modelling DNA loops using continuum and statistical mechanics", *Philos. T. R. Soc. A*, **362**, 1355-1371.
- Balaeff, A., Mahadevan, L. and Schulten, K. (1999), "Elastic rod model of a DNA loop in the *lac* operon", *Phys Rev. Lett.*, **83**, 4900-4903.
- Balaeff, A., Mahadevan, L. and Schulten, K. (2006), "Modeling DNA loops using the theory of elasticity", *Phys Rev. E*, **73**, 031919.
- Baumann, C.G., Smith, S.B., Bloomfield, V.A. and Bustamante, C. (1997), "Ionic effects on the elasticity of single DNA molecules", *Proc. Natl. Acad. Sci. USA*, **94**, 6185-6190.
- Born, M. and Huang, K. (1954), *Dynamical theory of crystal lattices*, Oxford University Press, New York.
- Chamekh, M., Mani-Aouadi, S. and Moakher, M. (2009), "Modeling and numerical treatment of elastic rods with frictionless self-contact", *Comput. Method. Appl. M.*, **198**, 3751-3764.
- Chen, J.S., Lee, C.H., Teng, H. and Wang, H. (2011), *Atomistic to continuum modeling of DNA molecules*, Advances in Soft Matter Materials, Springer.
- Chen, J.S., Teng, H. and Nakano, A. (2007), "Wavelet-based multi-scale coarse graining approach for DNA molecules", *Finite Elem. Anal. Des.*, **43**, 346-360.
- Chui, C.K. (1992), *An introduction to wavelets*, Academic Press, New York.
- Clausen-Schaumann, H., Rief, M., Tolksdorf, C. and Gaub, H.E. (2000), "Mechanical stability of single DNA molecules", *Biophys. J.*, **78**, 1997-2007.
- Cluzel, P., Lebrun, A., Heller, C., Lavery, R., Viovy, J.L., Chatenay, D. and Caron, F. (1996), "DNA: an extensible molecule", *Science*, **271**, 792-794.
- Coleman, B.D. and Swigon, D. (2004), "Theory of self-contact in Kirchhoff rods with applications to supercoiling of knotted and unknotted DNA plasmids", *Philos. T. R. Soc. A*, **362**, 1281-1299.
- Cornell, W.D., Cieplak, P., Bayly, C.I., Gould, I.R., Merz, K.M., Ferguson, D.M., Spellmeyer, D.C., Fox, T., Caldwell, J.M. and Kollman P.A. (1996), "A second generation force field for the simulation of proteins, nucleic acids, and organic molecules", *J. Am. Chem. Soc.*, **117**, 2309-2309.
- Crossley, M. and Orkin, S.H. (1993), "Regulation of the  $\beta$ -globin locus", *Curr. Opin. Genet. Dev.*, **3**, 232-237.
- Daubechies, I. (1992), "Ten lectures on wavelets", *CBMS-NSF Series in Applied Mathematics*, **61**, SIAM, Philadelphia, PA.
- de Borst, R., Réthoré J., Abellan, M.A. (2008), "Two-scale approaches for fracture in fluid-saturated porous media", *Interact. Multiscale Mech.*, **1**(1), 83-101.
- Dunn, T.M., Hahn, S., Ogden, S., and Schleif, R. F. (1984), "An operator at - 280 base pairs that is required for repression of araBAD operon promoter: addition of DNA helical turns between the operator and promoter cyclically hinders repression", *Proc. Natl. Acad. Sci. USA*, **81**, 5017-5020.
- Edelman, L.M., Cheong, R. and Kahn, J.D. (2003), "Fluorescence resonance energy transfer over  $\sim$  130 basepairs in hyperstable *lac* repressor-DNA loops", *Biophys. J.*, **84**, 1131-1145.
- Ericksen, J. (1984), *The Cauchy-Born hypothesis for crystals*, in Phase transformation and material instabilities in solid [Ed. Gurtin, M.], 61-77, Academic Press, New York.
- Garrivier, D. and Fourcade, B. (2000), "Twisting DNA with variable intrinsic curvature", *Europhys. Lett.*, **49**, 390-395.
- Gent, A.N. (1996), "A new constitutive relation for rubber", *Rubber Chem. Technol.*, **69**, 59-61.
- Gevorkian, S.G. and Khudaverdian, E.E. (1990), "Mechanical properties of DNA films", *Biopolymers*, **30**, 279-285.
- Hughes, T.J.R. and Liu, W.K. (1981a), "Nonlinear finite element analysis of shells: Part I Three-dimensional shells", *Comput. Method. Appl. M.*, **26**, 331-362.
- Hughes, T.J.R. and Liu, W.K. (1981b), "Nonlinear finite element analysis of shells: Part II Two-dimensional shells", *Comput. Method. Appl. M.*, **27**, 167-181.
- Krämer, H., Niemöller, M., Amouyal, M., Revet, R., von Wilcken-Bergmann, B. and Müller-Hill, B. (1987), "*lac* repressor forms loops with linear DNA carrying two suitably spaced *lac* operators", *EMBO J.*, **6**, 1481-1491.

- Lewis, M., Chang, G., Horton, N.C., Kercher, M.A., Pace, H.C., Schumacher, M.A., Brennan, R.G. and Lu, P.Z. (1996), "Crystal structure of the lactose operon repressor and its complexes with DNA and inducer", *Science*, **271**, 1247-1254.
- Manning, G.S. (1985), "Packaged DNA-an elastic model", *Cell Biophys.*, **7**, 57-89.
- Marko, J.F. and Siggia, E.D. (1994), "Fluctuations and supercoiling of DNA", *Science*, **265**, 506-508.
- Mehraeen, S. and Chen, J.S. (2004), "Wavelet-based multi-scale projection method in homogenization of heterogeneous media", *Finite Elem. Anal. Des.*, **40**, 1665-1679.
- Müller, J., Oehler, S. and 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.
- Ogden, R.W., Saccomandi, G. and Sgura, I. (2008), "Phenomenological modeling of DNA overstretching", *Quant Biol, Biomol*, arXiv:0802.3323v1.
- Purohit, P.K. and Nelson, P.C. (2006), "Effect of supercoiling on formation of protein-mediated DNA loops", *Phys. Rev. E*, **74**, 061907.
- Reissner, E. (1973), "On one-dimensional, large-displacement, finite-strain beam-theory", *Stud. Appl. Math.*, **52**, 87-95.
- Rojek, J., Oñate, E. (2008), "Multiscale analysis using a coupled discrete/finite element model", *Interact. Multiscale Mech.*, **1**(1), 1-31.
- Saiz, L., Rubi, M. and Vilar, J.M.G. (2005) "Inferring the *in vivo* looping properties of DNA", *Proc. Natl. Acad. Sci.*, **102**, 17642-17645.
- Schleif, R. (1992), "DNA looping", *Ann. Rev. Biochem.*, **61**, 199-223.
- Serfling, E., Jasin, M. and Schaffner, W. (1985), "Enhancers and eukaryotic gene transcription", *Trends Genet.*, **1**, 224-230.
- Smith, S.B., Cui, Y. and Bustamante, C. (1996), "Overstretching B-DNA: the elastic response of individual double-stranded and single-stranded DNA molecules", *Science*, **271**, 795-799.
- Smith, S.B., Finzi, L. and Bustamante, C. (1992), "Direct mechanical measurement of the elasticity of single DNA molecules by using magnetic beads", *Science*, **258**, 1122-1126.
- Villa, E., Balaeff, A., Mahadevan, L. and Schulten, K. (2004), "Multiscale method for simulating protein-DNA complexes", *Multiscale. Model. Simul.*, **2**, 527-553.
- Wadati, M. and Tsuru, H. (1986), "Elastic model of looped DNA", *Physica*, **21D**, 213-226.
- Whirley, R.G. and Engelmann, B.E. (1993), *DYNA3D: A nonlinear, explicit, three-dimensional finite element code for solid and structural mechanics, user manual and theoretical manual*.
- White, J.H., Lund, R.A. and Bauer, W.R. (1999), "Effect of salt-dependent stiffness on the conformation of a stressed DNA loop containing initially coplanar bends", *Biopolymers*, **49**, 605-619.