



The 1-Particle-per-k-Nucleotides (1PkN) Elastic Network Model of DNA Dynamics with Sequence-Dependent Geometry

Takeru Kameda¹, Shuhei Isami¹, Yuichi Togashi^{2*}, Hiraku Nishimori^{1,2}, Naoaki Sakamoto^{1,2} and Akinori Awazu^{1,2*}

¹ Department of Mathematical and Life Sciences, Hiroshima University, Hiroshima, Japan, ² Research Center for the Mathematics on Chromatin Live Dynamics, Hiroshima University, Hiroshima, Japan

OPEN ACCESS

Edited by:

Francisco Monroy,
Complutense University of Madrid,
Spain

Reviewed by:

Shoji Takada,
Kyoto University, Japan
Francisco J. Cao,
Complutense University of Madrid,
Spain

*Correspondence:

Yuichi Togashi
togashi@hiroshima-u.ac.jp
Akinori Awazu
awa@hiroshima-u.ac.jp

Specialty section:

This article was submitted to
Biophysics,
a section of the journal
Frontiers in Physiology

Received: 06 December 2016

Accepted: 07 February 2017

Published: 14 March 2017

Citation:

Kameda T, Isami S, Togashi Y,
Nishimori H, Sakamoto N and
Awazu A (2017) The
1-Particle-per-k-Nucleotides (1PkN)
Elastic Network Model of DNA
Dynamics with Sequence-Dependent
Geometry. *Front. Physiol.* 8:103.
doi: 10.3389/fphys.2017.00103

Coarse-grained models of DNA have made important contributions to the determination of the physical properties of genomic DNA, working as a molecular machine for gene regulation. In this study, to analyze the global dynamics of long DNA sequences with consideration of sequence-dependent geometry, we propose elastic network models of DNA where each particle represents k nucleotides (1-particle-per- k -nucleotides, 1PkN). The models were adjusted according to profiles of the anisotropic fluctuations obtained from our previous 1-particle-per-1-nucleotide (1P1N) model, which was proven to reproduce such profiles of all-atom models. We confirmed that the 1P3N and 1P4N models are suitable for the analysis of detailed dynamics such as local twisting motion. The models are intended for the analysis of large structures, e.g., 10-nm fibers in the nucleus, and nucleoids of mitochondrial or phage DNA at low computational costs. As an example, we surveyed the physical characteristics of the whole mitochondrial human and *Plasmodium falciparum* genomes.

Keywords: double-stranded DNA, sequence-dependent geometry, coarse-grained, elastic network model, normal-mode analysis, mitochondrial DNA

INTRODUCTION

Genomic DNA plays many functional roles such as a medium for the storage of genetic information and as a molecular machine involved in gene expression regulation. A recent genome-wide analysis of the intra-nuclear chromosomes of eukaryotes suggested that not only chemical but also sequence-dependent physical properties of DNA predominantly contribute to formation of the genomic structure and gene expression regulation (Shrader and Crothers, 1989; Cacchione et al., 1997; Lowary and Widom, 1998; Widlund et al., 1999; Geggier and Vologodskii, 2010). Specifically, sequence-dependent structures play a key role in the determination of functional properties such as nucleosome positioning in DNA (Freeman et al., 2014b; Awazu, 2017).

Genomic DNA in prokaryotes, viruses, and mitochondria forms compact structures called nucleoids. The structures are shaped through the bending and twisting of DNA, mediated by histone-like proteins such as mitochondrial transcription factor A (TFAM) (Alam et al., 2003; Pohjoismäki et al., 2006; Kaufman et al., 2007; Ngo et al., 2014) and HU proteins (Broyles and Pettijohn, 1986; Kobryn et al., 1999; Guo and Adhya, 2007), with the help of topoisomerases (Brown et al., 1979; Wasserman and Cozzarelli, 1991; Nenortas et al., 1998; Wang, 2002). Each

nucleoid generally contains negative supercoils of DNA, which notably enhance the transcription and replication of genetic loci (Bogenhagen et al., 2014; Farge et al., 2014; Kukat et al., 2015). This observation indicates that the physical processes involving DNA, such as the formation and deformation of nucleoids, perform important functions in the regulation of gene expression in prokaryotes.

Coarse-grained (CG) models of DNA have facilitated investigations of the dynamics of relatively large DNA molecules for long time scales at reduced computational costs. Recently proposed CG models can accurately reproduce some of the properties observed in experiments (Freeman et al., 2014a; Hinckley and de Pablo, 2015; Markegard et al., 2015; Snodin et al., 2015; Singh and Granek, 2016) or all-atom models (Savelyev and Papoian, 2010; Setny and Zacharias, 2013; Naomé et al., 2014), which also depend on solution conditions. However, due to the large degrees of freedom, the computational cost of these models is still not low enough to be applicable to long DNA sequences containing more than 10^4 base pairs, corresponding to a 10-nm fiber with 100 nucleosomes or a mitochondrial nucleoid.

Recently, we devised a simple CG elastic network model, in which each nucleotide is described by one particle, named the 1-particle-per-1-nucleotide (1P1N) model. This model incorporates a sequence-dependent basic structure, and can reproduce the dynamic behavior around the basic structure observed in all-atom models (Isami et al., 2015). In this paper, to further reduce the computational cost, we propose a 1-particle-per- k -nucleotides (1PkN) model, i.e., a CG elastic network model where one particle represents k nucleotides. To construct feasible 1PkN models for $k > 1$, we designed interaction networks to reproduce the behavior of the 1P1N model, and determined appropriate k values. We found that the 1PkN models with $k < 13$ could sufficiently describe the global fluctuation of DNA, whereas the 1P3N and 1P4N models are also applicable to detailed analyses of local fluctuations.

The 1PkN models can deal with sequence-dependent behavior at a more than 10 times reduced computational cost than required for the previous CG models. As a demonstration, we constructed 1P3N models of the entire *Plasmodium falciparum* and human mitochondrial genomes (5,967 and 16,569 bp, respectively), and analyzed their physical properties affecting local bending and twisting, which are important for the formation of nucleoid structures.

MODELS AND METHODS

The 1P1N Coarse-Grained (CG) Model

First, we briefly review our previous CG elastic network model of double-stranded DNA, where each nucleotide is represented by one particle, namely the 1P1N model (Isami et al., 2015).

Similar to other elastic network models, this model depends on a basic structure. To consider sequence-dependent geometry, we constructed a basic structure according to a set of helical parameters (base-step and base-pair parameters) depending on the local nucleotide sequences (typically two consecutive base-pairs). These parameters have been determined by experiments and molecular dynamics simulations (Olson et al., 1998, 2006;

Lankaš et al., 2003; Perez et al., 2008; Morozov et al., 2009; Lavery et al., 2010; Dans et al., 2012; Hospital et al., 2013). Throughout this study, we employed the helical parameters obtained in *in vitro* experiments and from X-ray crystal structure analysis (Table S1) (Olson et al., 1998, 2006; Lankaš et al., 2003; Morozov et al., 2009; Freeman et al., 2014a), which were also used in other recent CG DNA models (Freeman et al., 2014a; Isami et al., 2015). Qualitatively similar results were obtained even if different helical parameter sets were applied (Perez et al., 2008; Lavery et al., 2010; Dans et al., 2012; Hospital et al., 2013) for all analyses described in this paper (data not shown). We used 3DNA (Lu and Olson, 2003) to obtain the coordinate of each atom in the DNA, on the basis of the helical parameters (for detailed methods on the generation of the atom coordinates, see Lu and Olson, 2003).

In the 1P1N model, $\mathbf{x}_i = (x_i, y_i, z_i)$, the position of nucleotide i ($i = 0, 1, 2, \dots$), is given as the coordinate of the C1' carbon of that nucleotide, and m_i , the mass of nucleotide i , is considered to be equal ($m_i = m$). The potential V of the interaction between nucleotides is given as

$$V = \sum_{ij} \frac{C^1 B_{ij}^1}{2} (|\mathbf{x}_i - \mathbf{x}_j| - |\mathbf{x}_i^0 - \mathbf{x}_j^0|)^2, \quad (1)$$

where \mathbf{x}_i^0 is the position of nucleotide i in the basic structure. We assumed that nucleotides i and i^c form a base pair, i.e., the superscript c indicates the corresponding nucleotide in the complementary strand. $B_{ij}^1 = 1$ for $j = i \pm 2, i \pm 1, (i \pm 3)^c, (i \pm 2)^c, (i \pm 1)^c$, and i^c ; and $B_{ij} = 0$ otherwise. C^1 is 7.7 kJ/(Å² mol) (by rescaling $m = 10^{-3}/N_A$ [kg]). Using normal-mode analysis (NMA), we confirmed that this model can accurately reproduce the dynamic parameters of each nucleotide obtained by means of the all-atom model, for the several types of helical parameters mentioned above (Isami et al., 2015).

The 1PkN Coarse-Grained (CG) Model

In this study, we aimed to construct further CG models intended for the analysis of long DNA. Here, we present 1PkN models, where k nucleotides are represented by a single particle. If we were to directly compare such models with the corresponding all-atom models, the computational cost of the analysis would be huge (e.g., a 450-bp model, considered below for the 1P9N case, involves 18,450 atoms), as we would need to assess a statistically sufficient number of sequences. Therefore, we used the 1P1N model as a reference to design and validate these new models.

To construct the 1PkN model, each k particle (i.e., at the center of consecutive k bases) in the 1P1N model was extracted. **Figure 1** shows the cases for $k = 4$ and $k = 5$. More specifically, we assumed a 1PkN model of L bp, consisting of $2 \times \lfloor \frac{L}{k} \rfloor$ particles (where $\lfloor \cdot \rfloor$ indicates the Gauss symbol). The three-dimensional coordinates $\mathbf{q}_n, \mathbf{q}_{n^c}$ and mass m_n, m_{n^c} of the n -th ($n = 0, 1, 2, \dots$) and n^c -th particles in the 1PkN model were expressed as those of the i -th and i^c -th particles (i.e., nucleotides) in the 1P1N model (i.e., $\mathbf{x}_i, \mathbf{x}_{i^c}, m_i$, and m_{i^c}), respectively, where $i = kn + k_0$ and $i^c = (kn + k_0)^c$ (the n -th and n^c -th particles form a base pair). Here, $k_0 = \frac{k-2}{2}$ (k : even) or $\frac{k-1}{2}$ (k : odd) (**Figure 1**).

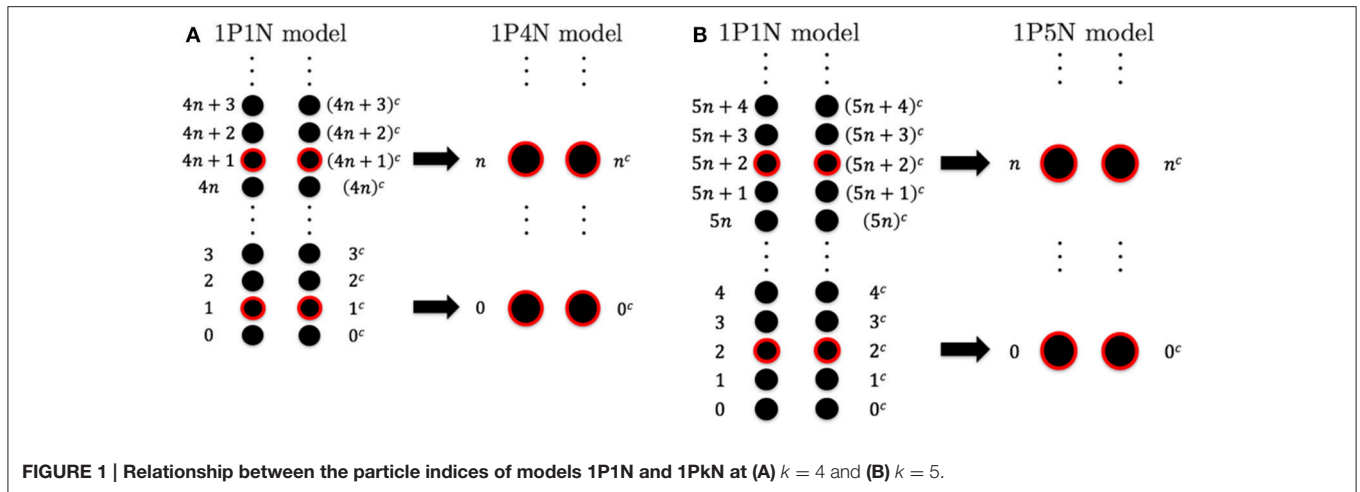


FIGURE 1 | Relationship between the particle indices of models 1P1N and 1PkN at (A) $k = 4$ and (B) $k = 5$.

Then, we define the interaction potential V between the particles as

$$V = \sum_{n,l} \frac{C^k B_{nl}^k}{2} (|\mathbf{q}_n - \mathbf{q}_l| - |\mathbf{q}_n^0 - \mathbf{q}_l^0|)^2, \quad (2)$$

where \mathbf{q}_n^0 is the position of particle n in the basic structure, B_{nl}^k indicates the weight of the connection between particles n and l (Figure 2), and C^k is the connection strength. We determined the appropriate k , C^k , and B_{nl}^k values as shown below.

As mentioned above, the basic structure was constructed according to a set of sequence-dependent helical parameters. Both the 1P1N model and 1PkN ($k > 1$) models incorporate sequence-dependent physical properties through the interaction potential V that depends on the basic structure (\mathbf{x}_i^0 or their subset \mathbf{q}_n^0).

Normal-Mode Analysis (NMA)

To optimize and evaluate the 1PkN model, we compared the structural fluctuations in the 1PkN model with those of the 1P1N model. For this purpose, NMA was used to estimate the mobility of each particle in the 1PkN model. NMA has been applied to elastic network models and is demonstrated to well reproduce the structural properties of biomolecules such as proteins (Tirion, 1996; Atilgan et al., 2001; Bahar and Rader, 2005; Yang et al., 2009; Bahar et al., 2009, 2010; Dykeman and Sankey, 2010). Here, we provide a brief overview of NMA.

We defined $\mathbf{q}(t)$ ($\mathbf{q} = (\mathbf{q}_0, \mathbf{q}_1, \dots, \mathbf{q}_{N-1})$, $\mathbf{q}_n = (x_n, y_n, z_n)$) as a $3N$ -dimensional position vector, and \mathbf{q}^0 as the position vector of the basic structure. Using linear approximation, small deviations $\delta\mathbf{q}(t) = \mathbf{q} - \mathbf{q}^0$ around the basic structure obey the following equation:

$$\delta\mathbf{q}(t) = \sum_{w_k \neq 0} A_k \mathbf{v}^k e^{i w_k t}, \quad (3)$$

where $-(w_k)^2$ and $\mathbf{v}^k = (\mathbf{v}_0^k, \mathbf{v}_1^k, \dots, \mathbf{v}_{N-1}^k)$ ($\mathbf{v}_i^k = (v_{x_i}^k, v_{y_i}^k, v_{z_i}^k)$, and $\sum_i |\mathbf{v}_i^k|^2 = 1$) represent the k -th largest eigenvalue and its

eigenvector of the $3N \times 3N$ Hessian matrix \mathbf{H} , where

$$\mathbf{H} = (H_{nl}), \quad (4)$$

$$H_{nl} = - \left(\frac{\partial^2 V}{\partial q_n \partial q_l} \right)_{\mathbf{q}=\mathbf{q}^0} \quad (5)$$

We presumed that the system is in thermodynamic equilibrium at temperature T . Then, the amplitude A_k is expressed as

$$(A_k)^2 = \frac{k_B T}{(w_k)^2 m}, \quad (6)$$

with the Boltzmann constant k_B . Using this solution, the mean square fluctuation (MSF; i.e., the strength of thermal fluctuation) of the n -th particle and the correlation of motion between particles n and l were obtained as

$$F_n^a = \langle |\delta\mathbf{q}_n|^2 \rangle = \sum_{w_k \neq 0} \frac{k_B T}{(w_k)^2 m} |\mathbf{v}_n^k|^2 \quad (7)$$

and

$$F_{nl}^c = \frac{\langle \delta\mathbf{q}_n \cdot \delta\mathbf{q}_l \rangle}{\sqrt{F_n^a} \sqrt{F_l^a}}, \quad (8)$$

respectively, where $\langle \dots \rangle$ represents the temporal average.

Similarly, for the 1P1N model, the MSF $G_i^a = \langle |\delta\mathbf{q}_i|^2 \rangle$ and the correlation of motion $G_{ij}^c = \langle \delta\mathbf{q}_i \cdot \delta\mathbf{q}_j \rangle / (\sqrt{G_i^a} \sqrt{G_j^a})$ were calculated, where the i -th and j -th particles in the 1P1N model correspond to the n -th and l -th particles in the 1PkN model, respectively.

Note that the particles at the ends of the 1PkN model do not correspond to those at the ends of the 1P1N model. Furthermore, the behavior of the particles at the ends of these models often become specific because there are fewer interacting particles at the ends than in the middle. Therefore, in the estimation of fluctuations, we omitted $\delta\mathbf{q}_n$ for four particles (i.e., two particle pairs) at each end.

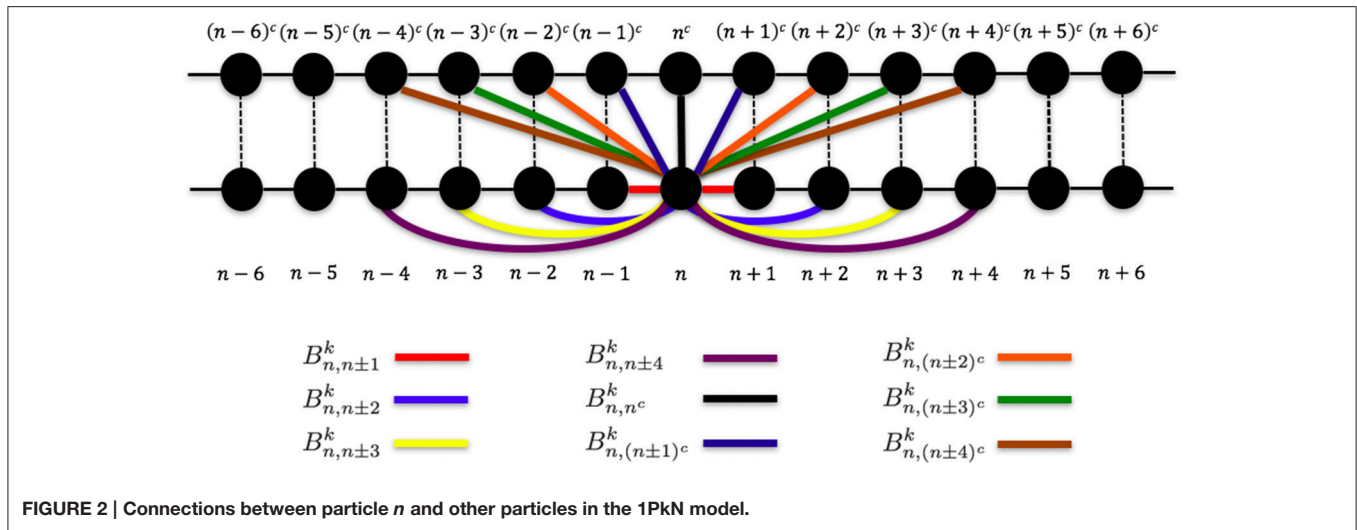


FIGURE 2 | Connections between particle n and other particles in the 1PkN model.

Determination of the Parameters in the 1PkN Model

The 1PkN model above includes the free parameters B_{nl}^k and C^k . We optimized these parameters to fit the overall fluctuation of each particle with that observed in the 1P1N model.

To compare the fluctuations between the 1P1N and 1PkN models, we employed randomly chosen DNA sequences with $50 \times k$ base pairs for each k ; i.e., the number of particle pairs in the 1PkN model was always 50. Pearson's correlation coefficients ρ^a between the MSF F_n^a in the 1PkN and G_i^a in the 1P1N models was used as the performance index. A set of B_{nl}^k that maximizes the average of ρ^a for 500 randomly chosen sequences was determined by a systematic survey (Table 1). For the sake of simplicity, we assumed B_{nl}^k for $l = n \pm 4, n \pm 3, n \pm 2, n \pm 1, (n \pm 4)^c, (n \pm 3)^c, (n \pm 2)^c, (n \pm 1)^c$, and n^c is equal to 0 or 1, and $B_{nl}^k = 0$ otherwise (Figure 2). We confirmed that the results were qualitatively the same even when B_{nl}^k could take intermediate values.

C^k values for $k > 1$ were chosen manually, which alter only the absolute level of the fluctuations and do not affect the correlations we consider below.

Comparison of Anisotropic Fluctuations in the 1P1N and 1PkN Models

As previously demonstrated, the 1P1N model can effectively reproduce the fluctuations of an all-atom model in several typical geometric vector directions (Isami et al., 2015). To evaluate the performance of the 1PkN models, we again adopted anisotropic fluctuations as indices, and compared them between the 1P1N and 1PkN models for randomly chosen DNA sequences. Since the total numbers of modes are different between the 1P1N and 1PkN models for $k > 2$, we cannot compare each of their eigenvectors directly. Thus, we focused on anisotropic MSFs in several representative directions for each particle in the 1P1N and 1PkN models.

The fluctuations in each direction were evaluated as follows. In the 1PkN model, the n -th and n^c -th particles form a base pair.

TABLE 1 | Set of appropriate C^k and B_{nl}^k values for each $k \in \{1, 3, 4, 9, 13\}$ showing the highest correlation of fluctuations between the 1PkN and 1P1N models.

k	1	3	4	9	13
C^k	7.7	7.2	0.7	0.2	0.1
$B_{n,n \pm 1}^k$	1	1	1	1	1
$B_{n,n \pm 2}^k$	1	1	1	1	1
$B_{n,n \pm 3}^k$	0	0	0	1	1
$B_{n,n \pm 4}^k$	0	0	0	0	0
B_{n,n^c}^k	1	1	1	1	1
$B_{n,(n \pm 1)^c}^k$	1	1	1	1	1
$B_{n,(n \pm 2)^c}^k$	1	0	1	1	1
$B_{n,(n \pm 3)^c}^k$	1	0	0	1	1
$B_{n,(n \pm 4)^c}^k$	0	0	0	0	1

For other k values, see Table S2.

We defined the unit vectors

$$\mathbf{b}_n = \frac{\mathbf{q}_{n^c}^0 - \mathbf{q}_n^0}{|\mathbf{q}_{n^c}^0 - \mathbf{q}_n^0|}, \tag{9}$$

$$\mathbf{s}_n = \frac{(\mathbf{q}_{n+1}^0 + \mathbf{q}_{(n+1)^c}^0) - (\mathbf{q}_n^0 + \mathbf{q}_{n^c}^0)}{|(\mathbf{q}_{n+1}^0 + \mathbf{q}_{(n+1)^c}^0) - (\mathbf{q}_n^0 + \mathbf{q}_{n^c}^0)|}, \tag{10}$$

and

$$\mathbf{t}_n = \frac{\mathbf{s}_n \times \mathbf{b}_n}{|\mathbf{s}_n \times \mathbf{b}_n|} \tag{11}$$

to the directions shown in Figure 3. Although \mathbf{b}_n and \mathbf{s}_n are not orthogonal in general, we confirmed that the angles of these vectors were always sufficiently close to $\frac{\pi}{2}$ rad for each n .

Using these vectors, we decomposed the MSFs of the n -th and n^c -th particles of the 1PkN model (i.e., F_n^a and $F_{n^c}^a$) into these three directions to obtain F_n^b ($F_{n^c}^b$), F_n^s ($F_{n^c}^s$), and F_n^t ($F_{n^c}^t$). We also calculated the MSF of relative particle positions DF_n^a of the n -th

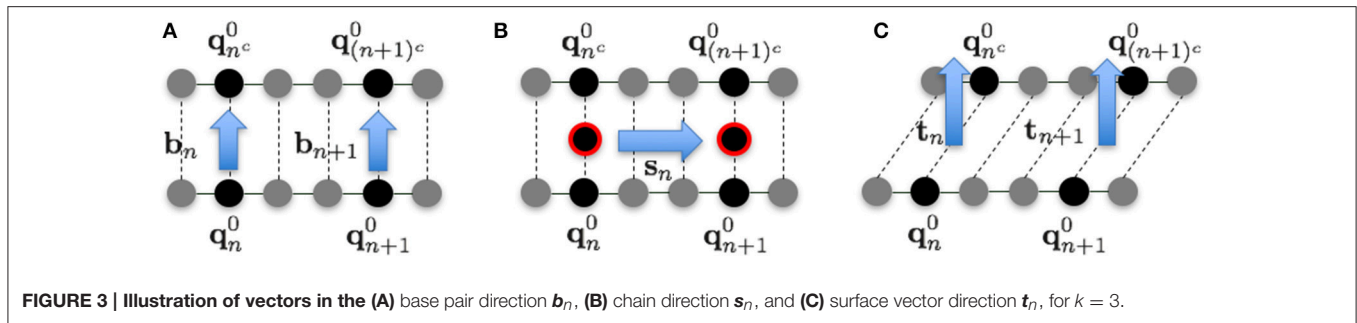


FIGURE 3 | Illustration of vectors in the (A) base pair direction b_n , (B) chain direction s_n , and (C) surface vector direction t_n , for $k = 3$.

particle pair, and its projection DF_n^b , DF_n^s , and DF_n^t to these three directions. Furthermore, we estimated the local flexibility toward the directions of bending and twisting around the n -th particle pair, as denoted by F_n^{bend} and F_n^{twist} , respectively. The definitions of these indices are provided in Table 2.

Moreover, to compare the behavior between the n -th (n^c -th) particle in the 1PkN model and the corresponding i -th (i^c -th) particle in the 1P1N model, their motions must be projected onto the common vectors of the 1PkN model. For this purpose, we adopted $\mathbf{b}_i^1 = \mathbf{b}_n$, $\mathbf{s}_i^1 = \mathbf{s}_n$, and $\mathbf{t}_i^1 = \mathbf{t}_n$ for $i = kn + k_0$ (\mathbf{b}_n , \mathbf{s}_n , and \mathbf{t}_n were obtained from the 1PkN model). Using these vectors, we projected the MSFs of the i -th (i^c -th) particles in the 1P1N model (i.e., G_i^a (G_i^a)) onto the same directions as those of the corresponding n -th (n^c -th) particles in the 1PkN model, to obtain G_i^b (G_i^b), G_i^s (G_i^s), and G_i^t (G_i^t). Similarly, the MSFs of the relative particle positions of the i -th particle pair (DG_i^a , DG_i^b , DG_i^s , and DG_i^t) and the local flexibility values (G_i^{bend} and G_i^{twist}) in the same directions were obtained (Table 2).

Similarities between the fluctuations in the 1PkN model (F_n^\bullet and DF_n^\bullet) and those in the 1P1N model (G_i^\bullet and DG_i^\bullet) for each sequence were evaluated by means of the Pearson's correlation coefficients ρ^\bullet listed in Table 2. The correlations of motion, F_{nl}^c and G_{ij}^c , were also compared using Pearson's correlation coefficient ρ^c .

To compare the absolute magnitude of fluctuations, we define $\langle F_n^\bullet \rangle_n$, $\langle DF_n^\bullet \rangle_n$, $\langle G_i^\bullet \rangle_i$, and $\langle DG_i^\bullet \rangle_i$ as the average of F_n^\bullet , DF_n^\bullet , G_i^\bullet , and DG_i^\bullet , respectively, over the entire sequence except for two particle pairs ($2 \times k$ base pairs) at each end in the 1PkN (1P1N) model.

Models of the Mitochondrial Genome

Using the 1P3N model, we analyzed the dynamics of the mitochondrial DNA of *P. falciparum* (5,967 bp, the shortest genome of all known mitochondrial DNA; NCBI Reference Sequence: NC_002375.1), human (16,569 bp; NCBI Reference Sequence: NC_012920.1), *Schizosaccharomyces pombe* (19,431 bp; NCBI Reference Sequence: NC_001326.1), and *Drosophila melanogaster* (19,524 bp; GenBank accession No.: KJ947872.2), using NMA. As the 3018-th base of human mitochondrial DNA remains unknown (<http://hgdownload.soe.ucsc.edu/goldenPath/hg38/chromosomes/>), we tried all four possible cases with adenine, guanine, cytosine, or thymine at that position.

We assessed the local bending and twisting ability F_n^{bend} and F_n^{twist} , respectively, of the n -th particle pair ($i = 3n + 1$ -th

base pair). The moving average was taken over seven particle pairs: $MF_n^{\text{bend}} = \sum_{m=n}^{n+6} F_m^{\text{bend}}$ and $MF_n^{\text{twist}} = \sum_{m=n}^{n+6} F_m^{\text{twist}}$. The averaging window (21 bp) largely corresponds to the length of the histone-like protein TFAM-binding sequence (22 bp Alam et al., 2003; Pohjoismäki et al., 2006; Kaufman et al., 2007; Ngo et al., 2014).

RESULTS AND DISCUSSION

Validation of the 1PkN Model by Means of Anisotropic Fluctuations

As described in the previous section, we evaluated the 1PkN model using the indices listed in Table 2. We calculated these indices for each DNA sequence. For most k values up to $k = 15$, the profiles of MSFs of the n -th particle (F_n^a , F_n^b , F_n^s , and F_n^t) and the correlations of fluctuations between the n -th and l -th particles (F_{nl}^c) in the 1PkN model were similar to the corresponding values (G_i^a , G_i^b , G_i^s , G_i^t , and G_{ij}^c) obtained in the 1P1N model (Figures 4–7; Figures S1–S16). In particular, good agreement was observed between F_n^a and G_i^a (the overall fluctuation profiles are independent of the value of C^k because C^k influences only the absolute values of fluctuations). For the 500 randomly chosen sequences, ρ^a , ρ^b , ρ^s , ρ^t , and ρ^c showed high average values with low standard deviations for all k values considered (Table 3; details of ρ^\bullet values and the profiles of $G_i^\bullet - F_n^\bullet$ ($DG_i^\bullet - DF_n^\bullet$) are shown in Tables S3–S5, Figures S1–S16).

For all k values considered, the absolute values of ρ^{Db} were always much lower (Table 3, Tables S3–S5). In fact, the amplitude of DF_n^b was shown to be negligible, i.e., much smaller than those of the fluctuations in other directions (Figures S2, S6, S10, S14). From these results, DF_n^s and DF_n^t were presumed to be essential for characterization of the inter-strand motions of double-stranded DNA. Focusing directly on the local bending and twisting flexibility of DNA in detail, F_n^{bend} and F_n^{twist} were thought to be even more important than DF_n^s and DF_n^t . Therefore, we mainly focused on ρ^{bend} and ρ^{twist} for the comparison between the 1PkN and 1P1N models.

For $k = 3$, ρ^{bend} and ρ^{twist} showed particularly high values (Table 3 and Figure 5). Therefore, the 1P3N model with an appropriate B_{nl}^k can effectively reproduce the mechanical properties, including local bending and twisting ability, of the 1P1N model; thus, although indirectly, the 1P3N model was proven to well reproduce such properties of the all-atom model.

TABLE 2 | Indices of fluctuations in the 1PkN and 1P1N models.

Meaning	Definition	Correlation
MSF (mean square fluctuation)	$F_n^a = \langle \delta \mathbf{q}_n ^2 \rangle$, $(F_{nc}^a = \langle \delta \mathbf{q}_{nc} ^2 \rangle)$ $G_i^a = \langle \delta \mathbf{q}_i ^2 \rangle$, $(G_{ic}^a = \langle \delta \mathbf{q}_{ic} ^2 \rangle)$	ρ^a
MSF parallel to \mathbf{b}_n	$F_n^b = \langle \delta \mathbf{q}_n \cdot \mathbf{b}_n ^2 \rangle$, $(F_{nc}^b = \langle \delta \mathbf{q}_{nc} \cdot \mathbf{b}_n ^2 \rangle)$ $G_i^b = \langle \delta \mathbf{q}_i \cdot \mathbf{b}_i ^2 \rangle$, $(G_{ic}^b = \langle \delta \mathbf{q}_{ic} \cdot \mathbf{b}_i ^2 \rangle)$	ρ^b
MSF parallel to \mathbf{s}_n	$F_n^s = \langle \delta \mathbf{q}_n \cdot \mathbf{s}_n ^2 \rangle$, $(F_{nc}^s = \langle \delta \mathbf{q}_{nc} \cdot \mathbf{s}_n ^2 \rangle)$ $G_i^s = \langle \delta \mathbf{q}_i \cdot \mathbf{s}_i ^2 \rangle$, $(G_{ic}^s = \langle \delta \mathbf{q}_{ic} \cdot \mathbf{s}_i ^2 \rangle)$	ρ^s
MSF parallel to \mathbf{t}_n	$F_n^t = \langle \delta \mathbf{q}_n \cdot \mathbf{t}_n ^2 \rangle$, $(F_{nc}^t = \langle \delta \mathbf{q}_{nc} \cdot \mathbf{t}_n ^2 \rangle)$ $G_i^t = \langle \delta \mathbf{q}_i \cdot \mathbf{t}_i ^2 \rangle$, $(G_{ic}^t = \langle \delta \mathbf{q}_{ic} \cdot \mathbf{t}_i ^2 \rangle)$	ρ^t
Inter-strand MSF	$DF_n^a = \langle \delta \mathbf{q}_n - \delta \mathbf{q}_{nc} ^2 \rangle$ $DG_i^a = \langle \delta \mathbf{q}_i - \delta \mathbf{q}_{ic} ^2 \rangle$	ρ^{Da}
Inter-strand MSF parallel to \mathbf{b}_n	$DF_n^b = \langle (\delta \mathbf{q}_n - \delta \mathbf{q}_{nc}) \cdot \mathbf{b}_n ^2 \rangle$ $DG_i^b = \langle (\delta \mathbf{q}_i - \delta \mathbf{q}_{ic}) \cdot \mathbf{b}_i ^2 \rangle$	ρ^{Db}
Inter-strand MSF parallel to \mathbf{s}_n	$DF_n^s = \langle (\delta \mathbf{q}_n - \delta \mathbf{q}_{nc}) \cdot \mathbf{s}_n ^2 \rangle$ $DG_i^s = \langle (\delta \mathbf{q}_i - \delta \mathbf{q}_{ic}) \cdot \mathbf{s}_i ^2 \rangle$	ρ^{Ds}
Inter-strand MSF parallel to \mathbf{t}_n	$DF_n^t = \langle (\delta \mathbf{q}_n - \delta \mathbf{q}_{nc}) \cdot \mathbf{t}_n ^2 \rangle$ $DG_i^t = \langle (\delta \mathbf{q}_i - \delta \mathbf{q}_{ic}) \cdot \mathbf{t}_i ^2 \rangle$	ρ^{Dt}
Local bending ability	$F_n^{\text{bend}} = \langle \delta \mathbf{Q}_{n+1} \cdot \mathbf{R}_{n+1} - \delta \mathbf{Q}_n \cdot \mathbf{R}_n ^2 \rangle$ $G_i^{\text{bend}} = \langle \delta \mathbf{Q}_{i+k} \cdot \mathbf{R}_{i+k} - \delta \mathbf{Q}_i \cdot \mathbf{R}_i ^2 \rangle$	ρ^{bend}
Local twisting ability	$F_n^{\text{twist}} = \langle (\delta \mathbf{q}_{n+1} - \delta \mathbf{q}_{(n+1)c}) \cdot \mathbf{t}_{n+1} - (\delta \mathbf{q}_n - \delta \mathbf{q}_{nc}) \cdot \mathbf{t}_n ^2 \rangle$ $G_i^{\text{twist}} = \langle (\delta \mathbf{q}_{i+k} - \delta \mathbf{q}_{(i+k)c}) \cdot \mathbf{t}_{i+k} - (\delta \mathbf{q}_i - \delta \mathbf{q}_{ic}) \cdot \mathbf{t}_i ^2 \rangle$	ρ^{twist}
Correlation of motion	$F_{nl}^c = \langle \delta \mathbf{q}_n \cdot \delta \mathbf{q}_l \rangle / (\sqrt{F_n^a} \sqrt{F_l^a})$ $G_{ij}^c = \langle \delta \mathbf{q}_i \cdot \delta \mathbf{q}_j \rangle / (\sqrt{G_i^a} \sqrt{G_j^a})$	ρ^c

We defined $\delta \mathbf{Q}_n = \left(\frac{\delta \mathbf{q}_n + \delta \mathbf{q}_{nc}}{2} \right)$, $\delta \mathbf{Q}_i = \left(\frac{\delta \mathbf{q}_i + \delta \mathbf{q}_{ic}}{2} \right)$, $\mathbf{R}_n = (\mathbf{b}_n + \mathbf{t}_n)$, $\mathbf{R}_i = (\mathbf{b}_i + \mathbf{t}_i)$.

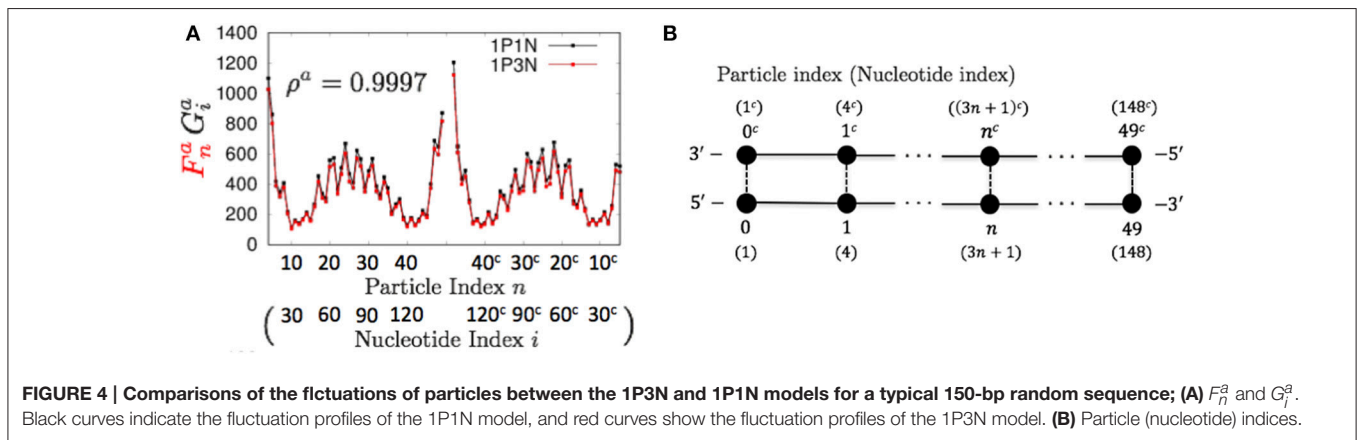


FIGURE 4 | Comparisons of the fluctuations of particles between the 1P3N and 1P1N models for a typical 150-bp random sequence; (A) F_n^a and G_i^a . Black curves indicate the fluctuation profiles of the 1P1N model, and red curves show the fluctuation profiles of the 1P3N model. (B) Particle (nucleotide) indices.

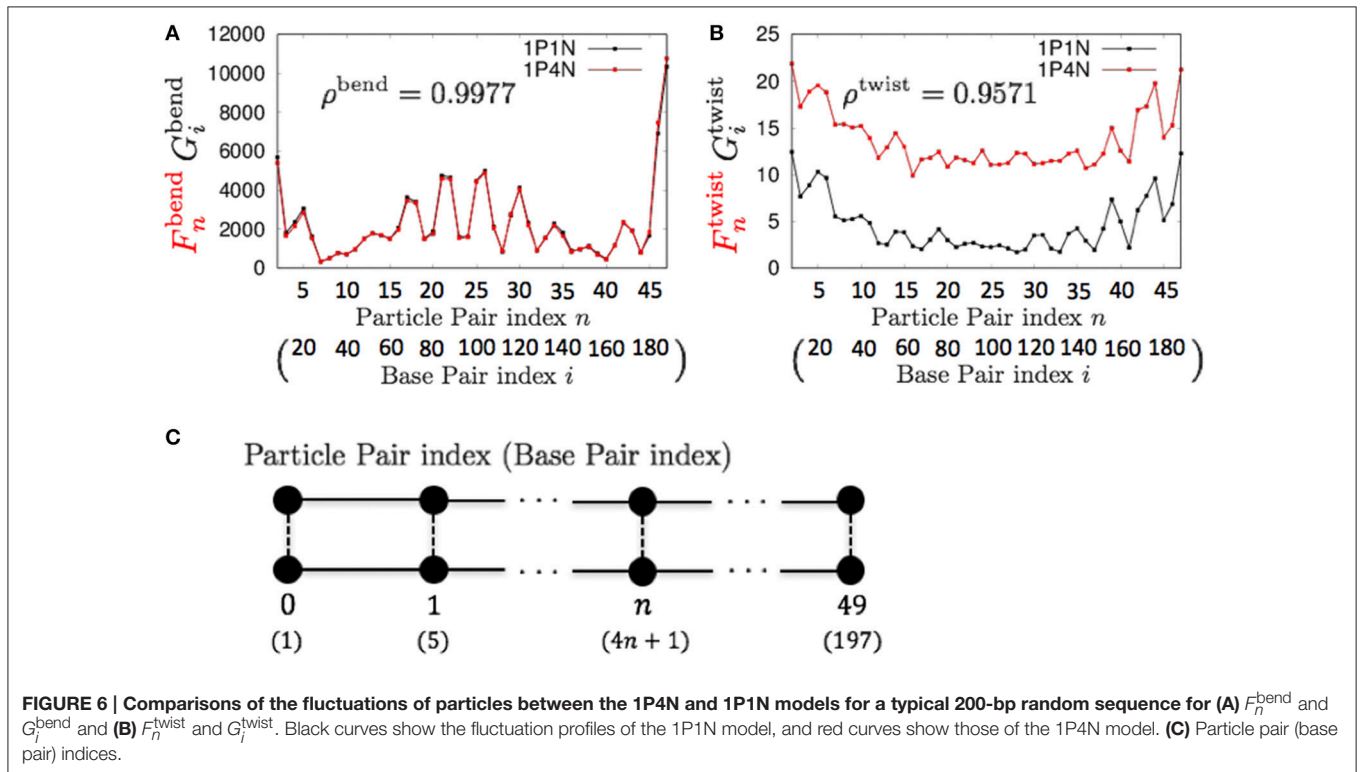
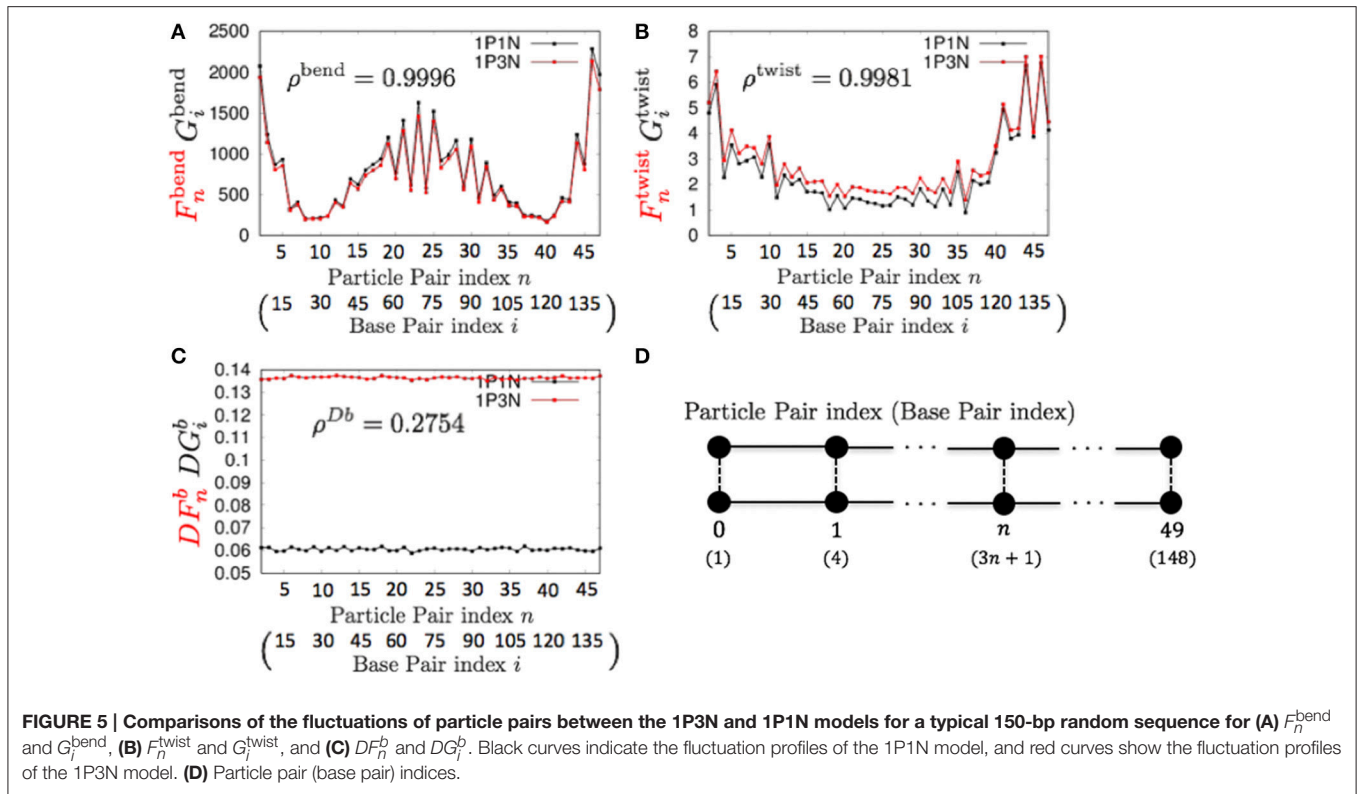
In addition, we confirmed that the result was almost identical for longer (450 bp) sequences (Table 3, Figures S13–S16), suggesting that the high correlations between the 1P3N and 1P1N models were independent of the sequence length.

For $k = 4$, ρ^{bend} and ρ^{twist} were sufficiently high (greater than 0.95), although a non-negligible difference exists between the absolute values of F_n^{twist} and G_i^{twist} (Figure 6). Therefore, if we analyze only the relative patterns of the local bending and twisting flexibility of DNA, the 1P4N model seems to be sufficiently useful.

The local twisting flexibility was not well reproduced for $k > 4$. This reflects the fact that the pitch of B-DNA is 10–11

base pairs, and hence the right-handed double-stranded structure was not retained for $k = 5$ (i.e., 2 particle pairs per pitch) or larger.

These 1P3N and 1P4N models contain 1/3 and 1/4 of the particles, respectively, as compared to the 1P1N model. Although the computational cost of the NMA of these models is much lower (by a factor of 1/10 to 1/30 for $k = 3$) than that of the 1P1N model, the accuracy of the obtained statistical results is almost identical (the computational cost to obtain all the eigenvalues and eigenvectors of an $N \times N$ matrix is generally between $O(N^2)$ and $O(N^3)$, which is the most time-consuming step; hence the reduction is expected to be k^{-2} to k^{-3}).



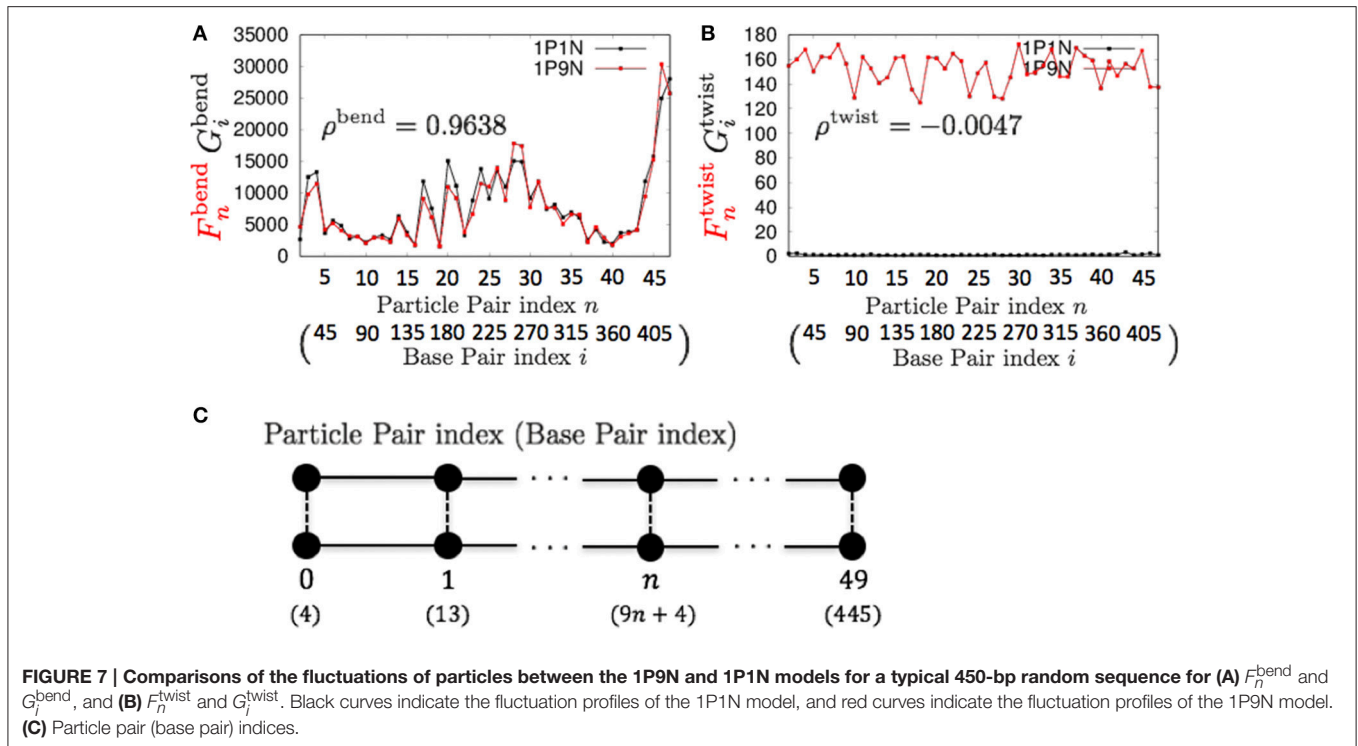


TABLE 3 | Average (Avg.) and standard deviation (S.D.) values of the correlations between the fluctuations of the 1PkN and 1P1N models for 500 randomly chosen $50 \times k$ bp sequences for $k \in \{3, 4, 9, 13\}$.

k (bp)	Fluctuation	3 (150 bp)	4 (200 bp)	9 (450 bp)	13 (650 bp)	3 (450 bp)
Avg. of ρ^a	$F_n^a - G_i^a$	0.9993	0.9967	0.9511	0.8879	0.9998
(S.D. of ρ^a)		(0.0006)	(0.0014)	(0.0320)	(0.0768)	(0.0002)
Avg. of ρ^c	$F_n^c - G_i^c$	0.9998	0.9991	0.9870	0.9712	0.9999
(S.D. of ρ^c)		(0.0001)	(0.0002)	(0.0094)	(0.0207)	(0.0001)
Avg. of ρ^{bend}	$F_n^{\text{bend}} - G_i^{\text{bend}}$	0.9992	0.9980	0.9451	0.8754	0.9998
(S.D. of ρ^{bend})		(0.0005)	(0.0010)	(0.0474)	(0.0946)	(0.0001)
Avg. of ρ^{twist}	$F_n^{\text{twist}} - G_i^{\text{twist}}$	0.9975	0.9549	-0.0028	-0.2917	0.9992
(S.D. of ρ^{twist})		(0.0014)	(0.0163)	(0.1566)	(0.1578)	(0.0005)
Avg. of ρ^{Db}	$F_n^{Db} - G_i^{Db}$	0.2789	0.1908	-0.0695	0.0187	0.2769
(S.D. of ρ^{Db})		(0.1210)	(0.1526)	(0.1514)	(0.1458)	(0.0691)

In addition, in the case of $k = 3$, randomly chosen 450-bp sequences are also inserted.

Therefore, these models can be used for exhaustive analyses of the dynamic features of large DNA molecules. It should also be noted that ρ^a , ρ^b , ρ^s , ρ^t , ρ^{bend} , and ρ^c were higher than 0.9 for all $k < 13$, suggesting that 1PkN models with larger k values (e.g., 1P9N in Figure 7) may be useful for analyses of only the crude sequence-dependent behavior of very long DNA molecules.

Effects of the GC Content on the Magnitude of Fluctuations

Next, to confirm the validity of our 1PkN model, we computed the magnitude of fluctuations in the 1P3N model, $\langle F_n^\bullet \rangle_n$ and $\langle DF_n^\bullet \rangle_n$, and in the 1P1N model, $\langle G_i^\bullet \rangle_i$ and $\langle DG_i^\bullet \rangle_i$, as

representative cases. Here, we used 1000 randomly generated 150-bp sequences, each satisfying the designated GC-content value (rate of GC base pairs in the sequence) 0, 0.1, 0.2, ..., 1.

The MSF exhibited approximately the same magnitude in the 1P3N model and 1P1N model for a GC-content of around 0.5, although in the case of lower GC-contents, $\langle F_n^a \rangle_n$ of the 1P3N model was slightly larger than $\langle G_i^a \rangle_i$ of the 1P1N model (specifically, reaching up to 10% larger) (Figure 8). Most of the profiles of $\langle F_n^\bullet \rangle_n$ and $\langle DF_n^\bullet \rangle_n$ were close to those of $\langle G_i^\bullet \rangle_i$ and $\langle DG_i^\bullet \rangle_i$; however, some cases resulted in considerably different magnitude profiles (Figure 8, Figure S17). Nevertheless, the 1P3N model consistently showed good agreement with the behavior of the 1P1N model.

Application Example: Mitochondrial-Genome Dynamics

As an example of the possible application of the 1P3N model, we analyzed the whole-genome dynamics of mitochondrial DNA. Using the NMA of the 1P3N model constructed for these sequences, the potential for bending and superhelix formation in each local part of these genomes was inferred.

Consistently, there were non-uniform profiles of MF_n^{bend} and MF_n^{twist} detected in *P. falciparum* mitochondrial DNA (Figure 9) and human mitochondrial DNA (Figure 10 and Figures S18–S20). Similar results were obtained for *S. pombe* (19,431 bp) and *D. melanogaster* (19,524 bp) mitochondrial DNA molecules, which are longer than human mitochondrial DNA (Figures S21, S22).

SUMMARY AND CONCLUSIONS

In this study, we developed the 1PkN model, which is more coarse-grained than recent DNA models and is intended for

the analysis of the dynamic properties of long DNA molecules. In particular, we showed that the 1P3N model, where each particle represents three nucleotides, could accurately reproduce the distributions of the anisotropic fluctuations of nucleotides and the flexibility of local regions observed in the 1P1N model. Considering our previous result showing that the 1P1N model well reproduced the fluctuations of all-atom models of DNA (Isami et al., 2015), although indirectly, the 1P3N model was here shown to be in good agreement with all-atom models.

The 1P3N model seems to be useful for detailed analyses of the mechanical properties of long DNA molecules. The 1P4N model appears to be helpful for assessment of the fluctuations and flexibility of local regions as well. Using the 1P3N model, we analyzed the physical properties of whole mitochondrial genomes, even longer than 10^4 bp. From the macroscopic perspective, our models should be connected to further coarse-grained polymer-like models (e.g., Brackley et al., 2013), where the sequence-dependent properties that emerged in our model could be incorporated as the inhomogeneity

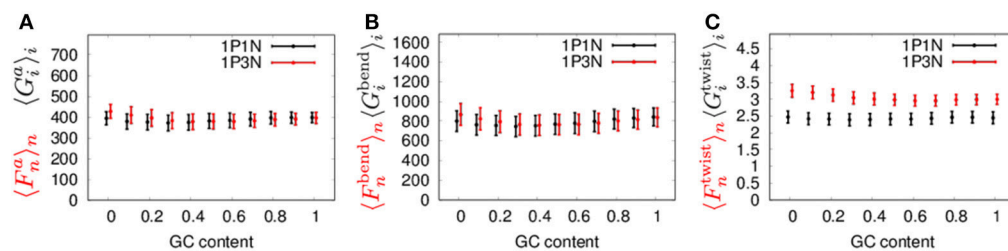


FIGURE 8 | Average \pm standard deviations of (A) $\langle F_n^a \rangle_n$ and $\langle G_i^a \rangle_i$, (B) $\langle F_n^{\text{bend}} \rangle_n$ and $\langle G_i^{\text{bend}} \rangle_i$, and (C) $\langle F_n^{\text{twist}} \rangle_n$ and $\langle G_i^{\text{twist}} \rangle_i$, for 1,000 samples of random 150-bp sequences with an average GC-content = 0, 0.1, 0.2, ... 1.

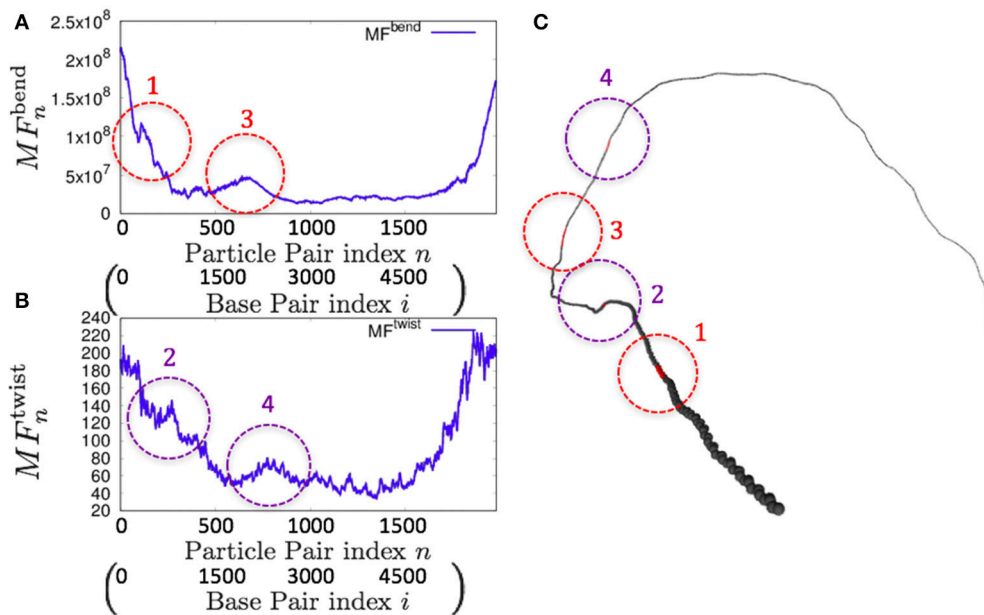
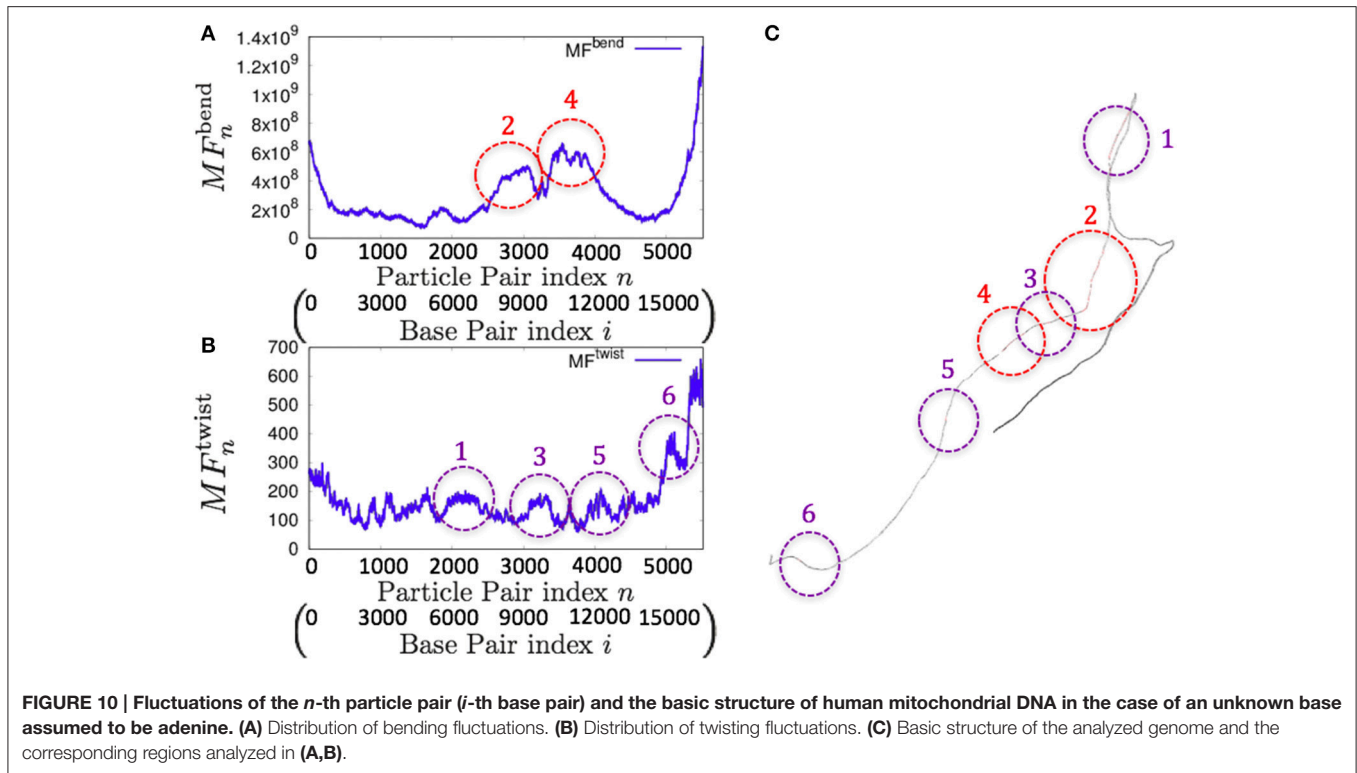


FIGURE 9 | Fluctuations and basic structure of *Plasmodium falciparum* mitochondrial DNA. (A) Distribution of bending fluctuations. (B) Distribution of twisting fluctuations. (C) Basic structure of the analyzed genome and the corresponding regions analyzed in (A,B).



of mechanical parameters such as persistence length. Such multiscale approaches would be beneficial for analyses of huge structures such as whole bacterial genome DNA.

In this study, our analysis depended on the NMA. Moreover, we can perform molecular dynamics simulations of the 1P3N model, as well as the other proposed models, for the analysis of large deformations and interactions with a variety of proteins. These simulations would be useful to elucidate the formation and dynamics of compact nucleoid structures and ~ 30 -nm chromatin structures (Gall, 1966; Adolph, 1980; Langmore and Paulson, 1983; Paulson and Langmore, 1983; Eltsov et al., 2008; Joti et al., 2012) or even larger structures such as the topologically associated domains in the nucleus (Zhu et al., 2007; Grosberg, 2012). Characteristics of DNA also depend on the conditions of the solution, such as temperature and ionic strength (Hamelberg et al., 2000; Mantz et al., 2007; Middleton et al., 2007). Therefore, adjustment of the model to reflect different conditions of solution would be warranted. We are planning to modify the proposed CG model to include the chemical properties of DNA and to analyze chromatin structures using not only NMA but also molecular dynamics simulations.

AUTHOR CONTRIBUTIONS

Conceived and designed the experiments: AA and NS. Performed the experiments: TK and SI. Analyzed the data:

TK, SI, YT, and AA. Wrote the paper: TK, YT, HN, NS, and AA.

FUNDING

This work was partially supported by the Platform Project for Support of Drug Discovery and Life Science Research (Platform for Dynamic Approaches to a Living System) from the Ministry of Education, Culture, Sports, Science and Technology of Japan (MEXT), and the Japan Agency for Medical Research and Development (AMED); and by the Grant-in-Aid for Scientific Research on Innovative Areas, Initiative for High-Dimensional Data-Driven Science through Deepening of Sparse Modeling [grant numbers 4503 and 26120525] of the MEXT. Funding for the open-access fee: AMED.

ACKNOWLEDGMENTS

The authors are grateful to S. Tate for fruitful discussions.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <http://journal.frontiersin.org/article/10.3389/fphys.2017.00103/full#supplementary-material>

REFERENCES

- Adolph, K. W. (1980). Organization of chromosomes in mitotic hela cells. *Exp. Cell Res.* 125, 95–103. doi: 10.1016/0014-4827(80)90193-7
- Alam, T. I., Kanki, T., Muta, T., Ukaji, K., Abe, Y., Nakayama, H., et al. (2003). Human mitochondrial DNA is packaged with tfam. *Nucl. Acids Res.* 31, 1640–1645. doi: 10.1093/nar/gkg251
- Atilgan, A. R., Durell, S. R., Jernigan, R. L., Demirel, M. C., Keskin, O., and Bahar, I. (2001). Anisotropy of fluctuation dynamics of proteins with an elastic network model. *Biophys. J.* 80, 505–515. doi: 10.1016/S0006-3495(01)76033-X
- Awazu, A. (2017). Prediction of nucleosome positioning by the incorporation of frequencies and distributions of three different nucleotide segment lengths into a general pseudo k-tuple nucleotide composition. *Bioinformatics* 33, 42–48. doi: 10.1093/bioinformatics/btw562
- Bahar, I., Lezon, T. R., Bakan, A., and Shrivastava, I. H. (2009). Normal mode analysis of biomolecular structures: functional mechanisms of membrane proteins. *Chem. Rev.* 110, 1463–1497. doi: 10.1021/cr900095e
- Bahar, I., Lezon, T. R., Yang, L.-W., and Eyal, E. (2010). Global dynamics of proteins: bridging between structure and function. *Ann. Rev. Biophys.* 39:23. doi: 10.1146/annurev.biophys.093008.131258
- Bahar, I., and Rader, A. (2005). Coarse-grained normal mode analysis in structural biology. *Curr. Opin. Struct. Biol.* 15, 586–592. doi: 10.1016/j.sbi.2005.08.007
- Bogenhagen, D. F., Martin, D. W., and Koller, A. (2014). Initial steps in rna processing and ribosome assembly occur at mitochondrial dna nucleoids. *Cell Metab.* 19, 618–629. doi: 10.1016/j.cmet.2014.03.013
- Brackley, C. A., Taylor, S., Papanonis, A., Cook, P. R., and Marenduzzo, D. (2013). Nonspecific bridging-induced attraction drives clustering of dna-binding proteins and genome organization. *Proc. Natl. Acad. Sci. U.S.A.* 110, E3605–E3611. doi: 10.1073/pnas.1302950110
- Brown, P. O., Peebles, C. L., and Cozzarelli, N. R. (1979). A topoisomerase from *Escherichia coli* related to DNA gyrase. *Proc. Natl. Acad. Sci. U.S.A.* 76, 6110–6114. doi: 10.1073/pnas.76.12.6110
- Broyles, S. S., and Pettijohn, D. E. (1986). Interaction of the *Escherichia coli* hu protein with DNA: evidence for formation of nucleosome-like structures with altered dna helical pitch. *J. Mol. Biol.* 187, 47–60. doi: 10.1016/0022-2836(86)90405-5
- Cacchione, S., Cerone, M. A., and Savino, M. (1997). *In vitro* low propensity to form nucleosomes of four telomeric sequences. *FEBS Lett.* 400, 37–41. doi: 10.1016/S0014-5793(96)01318-X
- Dans, P. D., Pérez, A., Faustino, I., Lavery, R., and Orozco, M. (2012). Exploring polymorphisms in B-DNA helical conformations. *Nucl. Acids Res.* 40, 10668–10678. doi: 10.1093/nar/gks884
- Dykeman, E. C., and Sankey, O. F. (2010). Normal mode analysis and applications in biological physics. *J. Phys.* 22:423202. doi: 10.1088/0953-8984/22/42/423202
- Eltsov, M., MacLellan, K. M., Maeshima, K., Frangakis, A. S., and Dubochet, J. (2008). Analysis of cryo-electron microscopy images does not support the existence of 30-nm chromatin fibers in mitotic chromosomes *in situ*. *Proc. Natl. Acad. Sci. U.S.A.* 105, 19732–19737. doi: 10.1073/pnas.0810057105
- Farge, G., Mehmedovic, M., Baclayon, M., van den Wildenberg, S. M., Roos, W. H., Gustafsson, C. M., et al. (2014). *In vitro*-reconstituted nucleoids can block mitochondrial DNA replication and transcription. *Cell Rep.* 8, 66–74. doi: 10.1016/j.celrep.2014.05.046
- Freeman, G. S., Hinckley, D. M., Lequieu, J. P., Whitmer, J. K., and de Pablo, J. J. (2014a). Coarse-grained modeling of DNA curvature. *J. Chem. Phys.* 141:165103. doi: 10.1063/1.4897649
- Freeman, G. S., Lequieu, J. P., Hinckley, D. M., Whitmer, J. K., and de Pablo, J. J. (2014b). DNA shape dominates sequence affinity in nucleosome formation. *Phys. Rev. Lett.* 113:168101. doi: 10.1103/PhysRevLett.113.168101
- Gall, J. G. (1966). Chromosome fibers studied by a spreading technique. *Chromosoma* 20, 221–233. doi: 10.1007/BF00335209
- Geggier, S., and Vologodskii, A. (2010). Sequence dependence of dna bending rigidity. *Proc. Natl. Acad. Sci. U.S.A.* 107, 15421–15426. doi: 10.1073/pnas.1004809107
- Grosberg, A. Y. (2012). How two meters of dna fit into a cell nucleus: polymer models with topological constraints and experimental data. *Poly. Sci. Ser. C* 54, 1–10. doi: 10.1134/S1811238212070028
- Guo, F., and Adhya, S. (2007). Spiral structure of *Escherichia coli* hu $\alpha\beta$ provides foundation for DNA supercoiling. *Proc. Natl. Acad. Sci. U.S.A.* 104, 4309–4314. doi: 10.1073/pnas.0611686104
- Hamelberg, D., McFail-Isom, L., Williams, L. D., and Wilson, W. D. (2000). Flexible structure of DNA: ion dependence of minor-groove structure and dynamics. *J. Am. Chem. Soc.* 122, 10513–10520. doi: 10.1021/ja0007071
- Hinckley, D. M., and de Pablo, J. J. (2015). Coarse-grained ions for nucleic acid modeling. *J. Chem. Theory Comput.* 11, 5436–5446. doi: 10.1021/acs.jctc.5b00341
- Hospital, A., Faustino, I., Collepardo-Guevara, R., González, C., Gelpí, J. L., and Orozco, M. (2013). Naflex: a web server for the study of nucleic acid flexibility. *Nucl. Acids Res.* 41, W47–W55. doi: 10.1093/nar/gkt378
- Isami, S., Sakamoto, N., Nishimori, H., and Awazu, A. (2015). Simple elastic network models for exhaustive analysis of long double-stranded dna dynamics with sequence geometry dependence. *PLoS ONE* 10:e0143760. doi: 10.1371/journal.pone.0143760
- Joti, Y., Hikima, T., Nishino, Y., Kamada, F., Hihara, S., Takata, H., et al. (2012). Chromosomes without a 30-nm chromatin fiber. *Nucleus* 3, 404–410. doi: 10.4161/nucl.21222
- Kaufman, B. A., Duricic, N., Mativetsky, J. M., Costantino, S., Hancock, M. A., Grutter, P., et al. (2007). The mitochondrial transcription factor tfam coordinates the assembly of multiple DNA molecules into nucleoid-like structures. *Mol. Biol. Cell* 18, 3225–3236. doi: 10.1091/mbc.E07-05-0404
- Kobryn, K., Lavoie, B. D., and Chaconas, G. (1999). Supercoiling-dependent site-specific binding of HU to naked Mu DNA. *J. Mol. Biol.* 289, 777–784. doi: 10.1006/jmbi.1999.2805
- Kukat, C., Davies, K. M., Wurm, C. A., Spähr, H., Bonekamp, N. A., Kühl, I., et al. (2015). Cross-strand binding of TFAM to a single mtDNA molecule forms the mitochondrial nucleoid. *Proc. Natl. Acad. Sci. U.S.A.* 112, 11288–11293. doi: 10.1073/pnas.1512131112
- Langmore, J. P., and Paulson, J. R. (1983). Low angle x-ray diffraction studies of chromatin structure *in vivo* and in isolated nuclei and metaphase chromosomes. *J. Cell Biol.* 96, 1120–1131. doi: 10.1083/jcb.96.4.1120
- Lankaš, F., Šponer, J., Langowski, J., and Cheatham, T. E. (2003). DNA basepair step deformability inferred from molecular dynamics simulations. *Biophys. J.* 85, 2872–2883. doi: 10.1016/S0006-3495(03)74710-9
- Lavery, R., Zakrzewska, K., Beveridge, D., Bishop, T. C., Case, D. A., Cheatham, T., et al. (2010). A systematic molecular dynamics study of nearest-neighbor effects on base pair and base pair step conformations and fluctuations in B-DNA. *Nucl. Acids Res.* 38, 299–313. doi: 10.1093/nar/gkp834
- Lowary, P. T., and Widom, J. (1998). New dna sequence rules for high affinity binding to histone octamer and sequence-directed nucleosome positioning. *J. Mol. Biol.* 276, 19–42. doi: 10.1006/jmbi.1997.1494
- Lu, X.-J., and Olson, W. K. (2003). 3DNA: a software package for the analysis, rebuilding and visualization of three-dimensional nucleic acid structures. *Nucl. Acids Res.* 31, 5108–5121. doi: 10.1093/nar/gkg680
- Mantz, Y. A., Gervasio, F. L., Laino, T., and Parrinello, M. (2007). Solvent effects on charge spatial extent in dna and implications for transfer. *Phys. Rev. Lett.* 99:058104. doi: 10.1103/PhysRevLett.99.058104
- Markegard, C. B., Fu, I. W., Reddy, K. A., and Nguyen, H. D. (2015). Coarse-grained simulation study of sequence effects on dna hybridization in a concentrated environment. *J. Phys. Chem. B* 119, 1823–1834. doi: 10.1021/jp509857k
- Middleton, C. T., Cohen, B., and Kohler, B. (2007). Solvent and solvent isotope effects on the vibrational cooling dynamics of a DNA base derivative. *J. Phys. Chem. A* 111, 10460–10467. doi: 10.1021/jp0740595
- Morozov, A. V., Fortney, K., Gaykalova, D. A., Studitsky, V. M., Widom, J., and Siggia, E. D. (2009). Using DNA mechanics to predict *in vitro* nucleosome positions and formation energies. *Nucl. Acids Res.* 37, 4707–4722. doi: 10.1093/nar/gkp475
- Naomé, A., Laaksonen, A., and Vercauteren, D. P. (2014). A solvent-mediated coarse-grained model of DNA derived with the systematic newton inversion method. *J. Chem. Theory Comput.* 10, 3541–3549. doi: 10.1021/ct500222s
- Nenortas, E. C., Bodley, A. L., and Shapiro, T. A. (1998). DNA topoisomerases: a new twist for antiparasitic chemotherapy? *Biochim. Biophys. Acta* 1400, 349–354. doi: 10.1016/s0167-4781(98)00146-8

- Ngo, H. B., Lovely, G. A., Phillips, R., and Chan, D. C. (2014). Distinct structural features of tfam drive mitochondrial DNA packaging versus transcriptional activation. *Nat. Commun.* 5:3077. doi: 10.1038/ncomms4077
- Olson, W. K., Colasanti, A. V., Li, Y., Ge, W., Zheng, G., and Zhurkin, V. B. (2006). "DNA simulation benchmarks as revealed by x-ray structures," in *Computational Studies of RNA and DNA*, eds J. Šponer and F. Lankaš (Dordrecht: Springer), 235–257. doi: 10.1007/978-1-4020-4851-3_9
- Olson, W. K., Gorin, A. A., Lu, X.-J., Hock, L. M., and Zhurkin, V. B. (1998). DNA sequence-dependent deformability deduced from protein-DNA crystal complexes. *Proc. Natl. Acad. Sci. U.S.A.* 95, 11163–11168. doi: 10.1073/pnas.95.19.11163
- Paulson, J., and Langmore, J. (1983). Low angle x-ray diffraction studies of hela metaphase chromosomes: effects of histone phosphorylation and chromosome isolation procedure. *J. Cell Biol.* 96, 1132–1137. doi: 10.1083/jcb.96.4.1132
- Perez, A., Lankas, F., Luque, F. J., and Orozco, M. (2008). Towards a molecular dynamics consensus view of B-DNA flexibility. *Nucl. Acids Res.* 36, 2379–2394. doi: 10.1093/nar/gkn082
- Pohjoismäki, J. L., Wanrooij, S., Hyvärinen, A. K., Goffart, S., Holt, I. J., Spelbrink, J. N., et al. (2006). Alterations to the expression level of mitochondrial transcription factor a, TFAM, modify the mode of mitochondrial DNA replication in cultured human cells. *Nucl. Acids Res.* 34, 5815–5828. doi: 10.1093/nar/gkl703
- Savelyev, A., and Papoian, G. A. (2010). Chemically accurate coarse graining of double-stranded DNA. *Proc. Natl. Acad. Sci. U.S.A.* 107, 20340–20345. doi: 10.1073/pnas.1001163107
- Setny, P., and Zacharias, M. (2013). Elastic network models of nucleic acids flexibility. *J. Chem. Theory Comput.* 9, 5460–5470. doi: 10.1021/ct400814n
- Shrader, T. E., and Crothers, D. M. (1989). Artificial nucleosome positioning sequences. *Proc. Natl. Acad. Sci. U.S.A.* 86, 7418–7422. doi: 10.1073/pnas.86.19.7418
- Singh, A. R., and Granek, R. (2016). Sufficient minimal model for dna denaturation: integration of harmonic scalar elasticity and bond energies. *J. Chem. Phys.* 145, 144101. doi: 10.1063/1.4964285
- Snodin, B. E., Randisi, F., Mosayebi, M., Šulc, P., Schreck, J. S., Romano, F., et al. (2015). Introducing improved structural properties and salt dependence into a coarse-grained model of DNA. *J. Chem. Phys.* 142:234901. doi: 10.1063/1.4921957
- Tirion, M. M. (1996). Large amplitude elastic motions in proteins from a single-parameter, atomic analysis. *Phys. Rev. Lett.* 77:1905. doi: 10.1103/PhysRevLett.77.1905
- Wang, J. C. (2002). Cellular roles of dna topoisomerases: a molecular perspective. *Nat. Rev. Mol. Cell Biol.* 3, 430–440. doi: 10.1038/nrm831
- Wasserman, S. A., and Cozzarelli, N. R. (1991). Supercoiled dna-directed knotting by t4 topoisomerase. *J. Biol. Chem.* 266, 20567–20573.
- Widlund, H. R., Kuduvali, P. N., Bengtsson, M., Cao, H., Tullius, T. D., and Kubista, M. (1999). Nucleosome structural features and intrinsic properties of the tataaacgcc repeat sequence. *J. Biol. Chem.* 274, 31847–31852. doi: 10.1074/jbc.274.45.31847
- Yang, L., Song, G., and Jernigan, R. L. (2009). Protein elastic network models and the ranges of cooperativity. *Proc. Natl. Acad. Sci. U.S.A.* 106, 12347–12352. doi: 10.1073/pnas.0902159106
- Zhu, J., Liu, W., Zhou, W., Hu, Y., and He, Y. (2007). A nucleus-encoded topological specificity factor ppmine in physcomitrella patens has conserved function similar to its chloroplast-encoded ancestor. *J. Genet. Genom.* 34, 229–238. doi: 10.1016/S1673-8527(07)60024-1

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2017 Kameda, Isami, Togashi, Nishimori, Sakamoto and Awazu. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.