Published in final edited form as:

Phys Chem Chem Phys. 2010 February 14; 12(6): 1399–1406. doi:10.1039/b916183j.

How stiff is DNA?

Guohui Zheng, Luke Czapla, A. R. Srinivasan, and Wilma K. Olson

Department of Chemistry & Chemical Biology, BioMaPS Institute for Quantitative Biology, Rutgers, the State University of New Jersey, Piscataway, NJ, 08854, USA

Wilma K. Olson: wilma.olson@rutgers.edu

Abstract

The natural stiffness of DNA, which contributes to the interactions of the many proteins involved in its biological processing and packaging, also plays an important role in modern nanotechnology. Here we report new Monte-Carlo simulations of deformable DNA molecules of potential utility in understanding the behavior of the long, double-helical polymer in the tight confines of a cell and in the design of novel nanomaterials and molecular devices. We directly determine the fluctuations in end-to-end extension associated with the conventional elastic-rod representation of DNA and with more realistic models that take account of the precise deformability of base-pair steps. Notably, the variance of end-to-end distance shows a quadratic increase with chain length in short chains of both types. We also consider the contributions to chain extension from the chemical linkages used to attach small molecular probes to DNA. The distribution of computed distances is sensitive to the intrinsic structure and allowed deformations of the tether. Surprisingly, the enhancement in end-to-end variance associated with the presence of the probe depends upon chain length, even when the probe is rigidly connected to DNA. We find that the elastic rod model of DNA in combination with a slightly fluctuating tether accounts satisfactorily for the distributions of end-to-end distances extracted from the small-angle X-ray scattering of gold nanocrystals covalently linked to the ends of short DNAs. There is no need to introduce additional structural fluctuations to reproduce the measured uptake in end-to-end fluctuations with chain length.

Introduction

The mechanical properties of DNA play a key role in its biological processing, determining how the long, thin, double-helical molecule responds to the binding of proteins and functions in confined spaces within a cell. Spectroscopic tools developed over the years to measure the distances between small, covalently linked chemical labels — *e.g.*, fluorescent dyes^{1–3} and nitroxide spin labels^{4–6} — provide some of the best available estimates of the natural structure and deformability of DNA in solution. The observed intramolecular distances between these probes mirror the known helical pathway of DNA, but the fluctuations in distances detected in duplexes of a few helical turns substantially exceed those expected from the classic helical wormlike chain model⁷ used to characterize the polymeric properties of DNA.

Recent studies of the small-angle X-ray scattering between gold nanocrystals attached to opposing ends of short DNA duplexes (Fig. 1A) reveal smaller variations in the distances between the tethered probes.^{8,9} The variation in distance with chain length, however, is greater than the uptake in end-to-end fluctuations expected from the simple geometric model

used to interpret the data. The apparent discrepancy — attributed to intrinsic stretching fluctuations in DNA appreciably larger than those deduced from either single-molecule force-extension measurements^{10–12} or analyses of high-resolution structures^{13,14} — has stimulated our interest in this system. Oversimplified models of polymeric behavior can sometimes be misleading.¹⁵ Interpretation of the observed properties of even a short, chemically labeled duplex requires a model that conforms closely and in an identifiable manner with the structural and deformational characteristics of both the DNA and the tethered probes. A simple model with direct control of the structural components can offer useful insights into the molecular system.

Here we investigate the subtle relationship between the local elastic properties of DNA, the fluctuations of tethered gold nanocrystals, and the overall configurational properties of short DNA duplexes of the type recently characterized by small-angle X-ray scattering. We explore the system directly by combining Gaussian sampling¹⁶ of the likely spatial arrangements of the DNA base-pair steps with Metropolis-Monte-Carlo simulations¹⁷ of movements in the tether. We take advantage of the multiple ‘time-step’ Monte-Carlo approach pioneered by Berne and co-workers¹⁸ to treat these two very different types of molecular movement. We examine the chain-length-dependent fluctuations in end-to-end extension associated with the twisted wormlike chain behavior of double-helical DNA. We also consider the contributions of rigid and flexible tethers to the distances between gold nanocrystals on the ends of short, fluctuating duplexes. Finally, we compare the predicted spread of distances of different DNA-tether models with the observed fluctuations and present our findings in the context of the experimental data.

Methods

DNA model

DNA is modeled at the level of base-pair steps in terms of six rigid-body parameters: three angular variables termed tilt, roll, and twist and three variables called shift, slide, and rise with dimensions of distance.¹⁹ A configuration of DNA is defined by the set of parameters at each base-pair step and is said to be relaxed when all parameters adopt their preferred equilibrium values.

The potential governing the fluctuations in base-pair steps is assumed to follow a quadratic expression of the form:

$$\Psi = \frac{1}{2} \sum_{i=1}^6 \sum_{j=1}^6 f_{ij} \Delta\theta_i \Delta\theta_j, \quad (1)$$

where the $\Delta\theta_i$ are deviations of the base-pair-step parameters θ_i from their intrinsic value θ_i^0 , and the f_{ij} are ‘stiffness’ constants. Local sequence-dependent structure and deformability in DNA can be incorporated in the θ_i^0 and f_{ij} .¹³

If the $\Delta\theta_i$ at base-pair step n are collected in the 6×1 vector $\Delta\Theta_n$ and the f_{ij} in the 6×6 force-constant matrix $\mathbf{F}_n$, eqn (1) takes the form: 1

$$\Psi_n = \frac{1}{2} \Delta\Theta_n^T \mathbf{F}_n \Delta\Theta_n, \quad (2)$$

with the total deformation energy U of DNA equal to the sum of the Ψ_n over all N base-pair steps:

$$U = \sum_{n=1}^N \Psi_n. \quad (3)$$

Gaussian sampling

We take advantage of the quadratic form of the energy in eqn (1) and the assumption that the base-pair steps fluctuate independently of one another to collect a Boltzmann distribution of dimeric states. We achieve this, as described elsewhere,¹⁶ by diagonalizing $\mathbf{F}$ and sampling linear combinations of base-pair-step parameters along the principal axes of dimeric deformation.

We consider several simple models of DNA. We first treat the double helix as an ideal, inextensible, naturally straight molecule with an intrinsic helical repeat of 10.5 bp/turn. The tilt and roll angles are accordingly null and the twist is $\sim 34.3^\circ$ in the rest state

($\theta_1^0 = \theta_2^0 = 0; \theta_3^0 = 34.3^\circ$). The translational parameters are ‘fixed’ at their intrinsic values

($\theta_4^0 = \theta_5^0 = 0; \theta_6^0 = 3.4^\circ$) by the assignment of large force constants. The root-mean-square

fluctuations in tilt are equated to those in roll, *i.e.*, $\langle \Delta\theta_1^2 \rangle^{1/2} = \langle \Delta\theta_2^2 \rangle^{1/2}$, so that bending is isotropic, and assigned values of 4.84° corresponding to a persistence length

$a = 2\Delta s / (\langle \Delta\theta_1^2 \rangle + \langle \Delta\theta_2^2 \rangle)$ of nearly 500 Å (if Δs , the per residue base-pair displacement, is taken as 3.4 Å). The fluctuations in twist are assumed to be independent of the bending deformations so that the model corresponds to the classic twisted wormlike chain

representation of DNA.⁷ The assumed fluctuations in twist $\langle \Delta\theta_3^2 \rangle^{1/2} = 4.09^\circ$ correspond to a

global twisting constant $C = k_B T / \langle \Delta\theta_3^2 \rangle$ somewhat larger in magnitude than the global bending constant A , *i.e.*, $C/A = 1.4$, where $A = ak_B T$. This choice of C is compatible with measurements of the equilibrium topoisomer distributions of DNA minicircles and the fluorescence depolarization anisotropy of ethidium bromide molecules intercalated in DNA minicircles.^{20,21}

We also consider more realistic representation that incorporate the known conformational properties of the DNA base-pair steps in knowledge-based elastic expressions of form of eqn. (1). The latter models allow for well-known features of DNA deformability such as anisotropic bending,²² the coupling of bending and shearing deformations,²³ chain extensibility,²⁴ *etc.*, as well as the subtle differences in deformability among different base-pair steps. Thus, local chain units can stretch as well as bend and twist. The force constants, which are derived from the covariance of step parameters in high-resolution structures,¹³ are scaled such that a mixed-sequence chain, with all 16 base-pair steps equally weighted,¹⁴ has the same persistence length as the ideal DNA model. Simulated sequences with a high proportion of pyrimidine-purine steps are more deformable and those with a high proportion of purine-pyrimidine or purine-purine steps are stiffer than the ideal and mixed-sequence chains. For simplicity, we ignore the small, sequence-dependence differences in intrinsic step parameters and the effects of adjacent nucleotides, which have almost no effect on the extension of short DNA chains.

DNA reconstruction

Recovery of atomic information from the DNA base-pair-step parameters is essential for understanding and visualizing the modeled fluctuations in double-helical structure. Moreover, the computational treatment of tethered gold nanocrystals requires knowledge of the coordinates of the points to which the labels are attached. We thus make use of the *rebuild* algorithm from the 3DNA software package^{25,26} to construct atomic models of DNA from the rigid-body parameters. We ignore potential fluctuations in base-pair geometry, assuming that the four base pairs — A·T, T·A, G·C, C·G — adopt standard Watson-Crick arrangements.²⁷

Generation of an atomic-level model necessitates the transformation of the coordinate frame on each base pair (Fig. 1C) into the global DNA reference frame. This is achieved using a serial product of matrices $\mathbf{A}_n$ that incorporate the 3×1 displacement vector $\mathbf{r}_n$ and the 3×3 rotation matrix $\mathbf{T}_{n,n+1}$, which relate coordinate frames on successive base pairs ($n, n+1$):

$$\mathbf{A}_{1:N} = \mathbf{A}_1 \mathbf{A}_2 \cdots \mathbf{A}_{N-1} \mathbf{A}_N, \quad (4)$$

where

$$\mathbf{A}_n = \begin{bmatrix} \mathbf{T}_{n,n+1} & \mathbf{r}_n \\ \mathbf{0} & 1 \end{bmatrix}. \quad (5)$$

The dependence of $\mathbf{T}_{n,n+1}$ and $\mathbf{r}_n$ on the base-pair-step parameters Θ_n follows the formulation introduced by Zhurkin *et al.*²² and subsequently developed by El Hassan and Calladine.²⁸

Determination of the atomic coordinates of the base pairs requires the additional

transformation of the coordinate vector of each atom $\mathbf{v}_s = \begin{bmatrix} x_s & y_s & z_s \end{bmatrix}^T$ from the standard base-pair reference frame to the global DNA frame. For example, the coordinates of atoms on base-pair $n+1$ can be expressed in the frame of base-pair n by the following transformation:

$$\mathbf{v}_n = \mathbf{T}_{n,n+1} \mathbf{v}_s + \mathbf{r}_n. \quad (6)$$

DNA end-to-end distance and contour length

The DNA end-to-end vector $\mathbf{r}_{1-N}$, which joins the centers of the first and last base pairs, is accumulated in the global generator matrix $\mathbf{A}_{1-N}$ described in eqn (4):

$$\mathbf{r}_{1:N} = \begin{bmatrix} \mathbf{I}_3 & \mathbf{0} \end{bmatrix} \mathbf{A}_{1:N} \begin{bmatrix} \mathbf{0} \\ 1 \end{bmatrix}. \quad (7)$$

Here $\mathbf{I}_3$ is the identity matrix of order three and the $\mathbf{0}$'s are null matrices of orders necessary to fill the 3×4 premultiplication and 4×1 postmultiplication vectors. The DNA end-to-end distance r_{DNA} is the magnitude of $\mathbf{r}_{1-N}$. The variance in the DNA end-to-end distance $\langle \delta r_{\text{DNA}}^2 \rangle$ is given by the standard difference of averages, $\langle \delta r_{\text{DNA}}^2 \rangle = \langle r_{\text{DNA}}^2 \rangle - \langle r_{\text{DNA}} \rangle^2$.

The DNA contour length L_{DNA} is the sum of the distances between sequential base pairs:

$$L_{\text{DNA}} = \sum_{n=1}^N |\mathbf{r}_n|, \quad (8)$$

where $\mathbf{r}_n$ is the displacement vector stored in the generator matrix associated with base-pair step n . The variance in the contour length $\langle \delta L_{\text{DNA}}^2 \rangle$ is obtained, like that for $\langle \delta r_{\text{DNA}}^2 \rangle$, from the mean-square and average values of L_{DNA} .

The end-to-end distance and contour length are identical if the DNA is perfectly straight. Twisting and stretching a straight DNA do not affect this equality, but bending and shearing lead to differences between the two measurements. Thus, if gold nanocrystals are tethered to the ends of DNA along the lines discussed below, the distance between the centers of the gold particles r_{Au} and the nanocrystal-DNA contour length L_{Au} will differ, given that the tether may bend and may not lie along the DNA helical axis in its equilibrium rest state.

Tether model

The gold nanocrystals tethered to the 3'-ends of DNA in recent small-angle X-ray-scattering experiments^{8,9} are small spherical constructs (~ 75 atoms) attached, via sulfur, to a three-carbon thiol that is connected in turn to DNA through a phosphodiester linkage (Fig. 1B). The spatial positions of the nanocrystals with respect to the DNA bases thus depend upon the internal coordinates (bond lengths, valence angles, and dihedral angles) of both the tether and the sugar-phosphate backbone.

The Cartesian coordinates of the tether are determined with a simple build-up procedure that starts with the approximate coordinates of three successive sugar atoms (C2', C3', O3') generated in the reconstruction of DNA from base-pair-step parameters.^{25,26} Given these coordinates ($\mathbf{v}_{n-2}$, $\mathbf{v}_{n-1}$, $\mathbf{v}_n$), the spatial position $\mathbf{v}_{n+1}$ of a fourth atom $n+1$ can be determined from knowledge of (i) the length b of the chemical bond that joins atom n to atom $n+1$, (ii) the magnitude of the valence angle θ formed by atoms $n-1$, n , and $n+1$, and (iii) the value of the dihedral angle ϕ described by atoms $n-2$, $n-1$, n , and $n+1$.

The coordinates of successive atoms are obtained by iteration of the following procedure. First, the components of $\mathbf{v}_{n+1}$ are defined by the expression:

$$\mathbf{v}_{n+1} = \mathbf{v}_n + \mathbf{R}_{n-1,n} \cdot \mathbf{b}_{n,n+1}, \quad (9)$$

where $\mathbf{R}_{n-1,n}$ is a 3×3 matrix that converts a local reference frame associated with atoms $n-2$, $n-1$, and n into the global frame of the molecule and $\mathbf{b}_{n,n+1}$ is a representation of the bond vector between atoms n and $n+1$ in the assumed local frame.

The components of $\mathbf{R}_{n-1,n}$ are given by:

$$\mathbf{R}_{n-1,n} = \begin{bmatrix} \mathbf{x}_{n-1,n} & \mathbf{y}_{n-1,n} & \mathbf{z}_{n-1,n} \end{bmatrix}, \quad (10)$$

where $\mathbf{z}_{n-1,n}$ is a unit vector along the bond that connects atoms $n-1$ and n , *i.e.*, $\mathbf{z}_{n-1,n} = (\mathbf{v}_n - \mathbf{v}_{n-1})/|\mathbf{v}_n - \mathbf{v}_{n-1}|$, $\mathbf{y}_{n-1,n}$ is the unit normal to the plane containing atoms $n-2$, $n-1$, and n , *i.e.*, $\mathbf{y}_{n-1,n} = (\mathbf{z}_{n-2,n-1} \times \mathbf{z}_{n-1,n})/|\mathbf{z}_{n-2,n-1} \times \mathbf{z}_{n-1,n}|$, and $\mathbf{x}_{n-1,n}$ is defined by the right-handed rule, *i.e.*, $\mathbf{x}_{n-1,n} = \mathbf{y}_{n-1,n} \times \mathbf{z}_{n-1,n}$.

The components of $\mathbf{b}_{n,n+1}$ in the local coordinate frame are given by the product:

$$\mathbf{b}_{n,n+1} = \mathbf{R}_z(\phi) \mathbf{R}_y(\pi - \theta) \mathbf{b}, \quad (11)$$

where $\mathbf{R}_u(\zeta)$ is a matrix describing the rotation of a vector through angle ζ about axis $\mathbf{u} = [u_1 \ u_2 \ u_3]$, $\mathbf{z} = [0, 0, 1]$ and $\mathbf{y} = [0, 1, 0]$ are the chosen axes of rotation, and $\mathbf{b} = [0, 0, b]$ is the representation of the bond between atoms n and $n+1$ in a local frame associated with atoms $n-1$, n , and $n+1$. The elements of $\mathbf{R}_u(\zeta)$ in this expression follow the standard definition:²⁹

$$r_{v\mu} = (1 - \cos\zeta) u_v u_\mu - \sin\zeta \sum_{\kappa} \varepsilon_{v\mu\kappa} u_\kappa + \cos\zeta \delta_{v\mu}, \quad (12)$$

where $\delta_{v\mu}$ is the Kronecker delta, *i.e.*, $\delta_{v\mu} = 1$ when $v = \mu$, $\delta_{v\mu} = 0$ when $v \neq \mu$, and $\varepsilon_{v\mu\kappa} = \pm 1$ when v, μ, κ is an even or odd permutation of 1, 2, 3, respectively, and vanishes otherwise.

DNA-tether interactions

We allow the tether to undergo small conformational fluctuations and large structural rearrangements via random and specific variations in backbone dihedral angles. The potential V associated with these changes is given by a standard summation of torsional and nonbonded terms:³⁰

$$V = \sum_{\text{dihedrals}} \frac{K_\varphi}{F} [1 + \cos(n\varphi - \gamma)] + \sum_{i < j} \left[\frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^6} \right], \quad (13)$$

which is evaluated over all pairwise combinations of movable particles, including the gold nanoassembly.

The local moves of the tether also include fluctuations, consistent with experiment,⁸ in the virtual distance b_{S-Au} between the sulphur atoms on the tether and the centers of the gold nanocrystals. We assume the stretching energy to be quadratic and assign an elastic constant

k_b equal to $(kT/2) \left\langle \delta b_{SGAu}^2 \right\rangle^{-1}$, where $\left\langle \delta b_{SGAu}^2 \right\rangle^{1/2}$ is the observed root-mean-square deviation in the virtual-bond distance, k the Boltzmann constant, and T the temperature in Kelvin.

We compute the total non-bonded interaction E between the DNA and tethers using a Lennard-Jones potential over all atom pairs on the two fragments, and an electrostatic potential between the gold nanocrystals and DNA phosphate groups.

$$E = \sum_{\{m,n\}} \left[\frac{A_{mn}}{r_{mn}^{12}} - \frac{B_{mn}}{r_{mn}^6} \right] + \sum_{\{Au,P\}} \frac{q_{Au} q_P}{\varepsilon(r_{AuP} - r_{Au})}, \quad (14)$$

The net negative charge on the nanocrystal⁸ (here taken to be -0.2 esu) is located at the center of the spherical gold assembly (an extended atom of radius $r_{Au} = 7$ Å) and that on DNA on the P atoms with a value (-0.24 esu) in accordance with the predictions of counterion condensation theory for a B DNA polyelectrolyte in monovalent salt solution.³¹

The atomic parameters used in eqns (13–14) to describe the non-bonded interactions of DNA and tether atoms are taken from the AMBER 10 force field.³² The effects of solvent on electrostatic interactions are treated implicitly with the dielectric constant ϵ assigned a value of 80.

Monte-Carlo simulation

We simulate the system in three stages using a multiple ‘time-step’ Monte-Carlo approach.¹⁸ First, we generate a random configuration of DNA using Gaussian sampling at each base-pair step. This is a straightforward process, which allows for fast rearrangement of the base pairs.¹⁶ Second, based on the sampled DNA configuration, we simulate the motions of the tethers, which are rooted in the DNA, using the Metropolis-Monte-Carlo method¹⁷ in combination with the energy term in eqn (13). This step is repeated several times after each move of the first type, so that the tethers undergo sufficient rearrangement. Finally, we accept or reject the configuration generated in the first two stages of computation by comparing the DNA-linker interactions obtained with eqn (14) with that present before the simulated move, again using the Metropolis algorithm.

Results and discussion

Global fluctuations of DNA

We start by examining the fluctuations in end-to-end extension associated with the twisted wormlike chain behavior of short double-helical DNA. We investigate a series of unlabeled molecules of the same chain lengths (10, 15, 20, 25, 30, 35 bp) considered in recent small-angle X-ray-scattering studies.^{8,9} Interpretation of the solution properties of the chemically labeled duplexes used in these and related experiments^{1–6} requires knowledge of the DNA motions as well as any effects of the tethered labels.

We apply two different models of DNA deformability: the first an ideal, inextensible, twisted wormlike chain with the intrinsic structure and elastic parameters described in Methods and the second a naturally straight, mixed-sequence chain subject to the fluctuations in base-pair steps seen in high-resolution structures and scaled to yield a persistence length of ~ 500 Å.^{13,14}

As expected, the average end-to-end distances $\langle r_{\text{DNA}} \rangle$ (points connected by dashed lines in Fig. 2A) are slightly smaller than the mean contour lengths $\langle L_{\text{DNA}} \rangle$ (points connected by solid lines) of the fluctuating DNA molecules. The differences between $\langle L_{\text{DNA}} \rangle$ and $\langle r_{\text{DNA}} \rangle$ are smaller for the inextensible, ideal chain (filled-in circles) compared to the ‘realistic’ knowledge-based model (open squares) that allows for the displacement (primarily shearing) of adjacent base pairs. The nearly identical values of $\langle r_{\text{DNA}} \rangle$ for the two types of chains reflect the similar persistence lengths in the models.

In contrast to published expectations,⁸ the variance in DNA end-to-end distance shows a quadratic dependence on chain length, with consistently greater fluctuations in global structure for the more ‘realistic’ model compared to the ideal chain (Fig. 2B). The uptake of radial fluctuations differs in longer chains (see Fig. S1 in the Electronic Supplementary Information). The dependence of $\langle \delta r_{\text{DNA}}^2 \rangle$ on chain length is roughly linear over the range 500–1600 bp and levels off to a constant value at the very long chain lengths (~ 2500 bp) where the simulated double helix is known to exhibit random-coil behavior.³³ The variance in contour length $\langle \delta L_{\text{DNA}}^2 \rangle$ shows a linear dependence on chain length, with the desired near-zero slope for the simulated, inextensible model and a slope of $0.08 \text{ Å}^2/\text{bp}$ for the mixed-sequence chain.

Effects of rigid tethers

We next consider the contributions of two kinds of rigid tethers to the end-to-end distances between gold nanocrystals attached to a short (15-bp), fluctuating DNA duplex. Here the DNA is modeled as an ideal, inextensible, twisted wormlike chain and the linkers are fixed in one of two different rigid states: an extended form with dihedral angles ($\varepsilon = 180^\circ$, $\zeta = -90^\circ$, $\alpha = 180^\circ$, $\beta = 180^\circ$, $\gamma = 180^\circ$, $\eta = 180^\circ$) selected such that the centers of the nanocrystals are close to the DNA helical axis in the equilibrium rest state and a kinked arrangement with dihedral angles ($\varepsilon = 180^\circ$, $\zeta = 180^\circ$, $\alpha = -60^\circ$, $\beta = 180^\circ$, $\gamma = 180^\circ$, $\eta = 180^\circ$) chosen so that centers of the nanocrystals are far from the helical axis in the rest state.

Although the DNA undergoes the same motions in both cases, the distributions of end-to-end distances $W(r_{\text{Au}})$ differ significantly. The separation between nanocrystals is much larger but the range of distances adopted by the extended linker (Figs. 3B,E) is much narrower than that adopted by the kinked tether (Figs. 3C,F). The distribution of extended tethers resembles that of DNA alone (Figs. 3A,D), *i.e.*, similar shapes. Because the projections of the nanocrystal centers on the terminal base-pair planes roughly coincide with the origins of the base-pair frames, the added chain extension tends to widen the arc of sampled points without significantly altering the highly skewed shape of the end-to-end distribution. The roughly sevenfold increase in radial variance with added chain extension, *i.e.*, $\langle \delta r_{\text{Au}}^2 \rangle$ vs. $\langle \delta r_{\text{DNA}}^2 \rangle$, is not quite as rapid as that illustrated in Fig. 2. The extended linkers add ~ 15.5 Å, or the equivalent of 5 rigid base-pair steps to the simulated DNA chain. The rearrangement of sampled points associated with the kinked linker has a drastic effect on the end-to-end distribution, increasing the variance by nearly two orders of magnitude compared to that of DNA alone. The shape of the latter distribution more closely resembles a Gaussian distribution with relatively symmetric tails on either side of the most probable end-to-end separation.

In order to gain a better understanding of how the tethered nanocrystals, although rigidly attached to DNA, contribute to the variance of the system as a whole, we studied the interdependence of the DNA end-to-end distances r_{DNA} and nanocrystal end-to-end distances r_{Au} . The two types of distances are highly correlated when the nanocrystals are bound to extended tethers (Fig. 4A) and uncorrelated when bound to the kinked tethers (Fig. 4B). The spread of data supports the qualitative rationale presented above.

In contrast to the nanocrystal centers attached to DNA via extended linkers, which build up symmetrically on the two ends of the fluctuating duplex, those attached via kinked linkers accumulate on one side of the molecule (Figs. 3B,C). Thus, the distances between chemical probes attached in the former manner will depend primarily on helical displacement, whereas the separation of probes tethered via kinked linkers will also reflect the helical twist. Furthermore, DNA with extended linkers should show a monotonic increase of end-to-end distance with chain length, whereas those with kinked linkers should exhibit non-monotonic variation. Interestingly, the observed distances between gold nanocrystals attached to short DNA duplexes of increasing chain length show minimal deviation from linearity, but the uptake of variance follows the zig-zag behavior expected of a slightly offset linker.⁸

We also examined the global bending of DNA measured by the positions of terminal base pairs and nanocrystal centers with respect to the DNA center, *i.e.*, the origin of the reference frame on the central base pair. The bending angles Γ_{DNA} described by the DNA base-pair points are roughly equivalent to and linearly correlated with those associated with the nanocrystals linked by extended tethers (Fig. 4C). The values of Γ_{Au} span a substantially

wider range of values when the nanocrystals are attached to the kinked tethers (Fig. 4D) and show no correlation with the bending of DNA alone.

Interestingly, the enhancement in variance brought about by the attachment of rigid tethers varies with DNA chain length (Fig. 5). That is, the difference $\langle \delta r_{\text{Au}}^2 \rangle - \langle \delta r_{\text{DNA}}^2 \rangle$ grows with chain length despite the fixed arrangement of the nanoparticles with respect to terminal base pairs. The tethers thus contribute different levels of intrinsic variance to the distances between gold nanocrystals, contrary to the assumptions⁸ that the contribution of the tether to the variance is fixed and that the variance in end-to-end length is directly proportional to chain length. Although the computed variation in $\langle \delta r_{\text{Au}}^2 \rangle - \langle \delta r_{\text{DNA}}^2 \rangle$ depends in part on the long-range electrostatic interactions between gold centers and DNA included in the simulations, the chain-length-dependent growth in the difference persists in neutral systems. The observed values reflect the dependence of the end-to-end distance of the tethered assembly on numerous factors, including the DNA end-to-end distance, the tether lengths, the angles between tethers and DNA, and the ‘torsion’ of the tethers with respect to the DNA axis.

Simulated distributions of end-to-end distances

The introduction of slight flexibility in the extended tethers brings the computed distributions of end-to-end distances $W(r_{\text{Au}})$ in reasonable agreement with those extracted from the small-angle X-ray scattering of gold nanocrystals (Table 1).^{8,9} The ‘stiff’ tethers used in these simulations have a single degree of conformational freedom: the dihedral angle η immediately preceding the thiol-nanocrystal linkage (Fig. 1B), which fluctuates in an energy well about its *trans* rest state. The choice of torsional parameters introduced in eqn (13) ($K_\phi = 1.2$, $F = 1$, $n = 1$) restricts the sampled angular states to values in the range $180 \pm 50^\circ$. The ‘flexible’ tethers introduced in other calculations allow for fluctuations of the same magnitude in all seven (ϵ , ζ , α , β , γ , η) dihedral angles of the tether.

The average distances $\langle r_{\text{Au}} \rangle$ between nanocrystals linked via ‘stiff’ tethers to either ideal, inextensible or mixed-sequence chains account for the experimentally reported data (Table 1). The predicted spread of distances $\langle \delta r_{\text{Au}}^2 \rangle$, although consistent with the observed nonlinear increase in range with chain length, slightly overestimates the fluctuations observed in 10–15-bp chains and somewhat underestimates the measured variation in 30–35-bp chains, *i.e.*, the dependence of variance on chain length. Variation of the persistence length within the wide range of values (450–490 Å)^{34,35} used to account for the solution properties of long DNA and/or modifications of the treatment of the tether can improve the match with experiment.

The predicted range of separation distances increases if the sequence-dependent deformability of DNA base-pair steps is incorporated in the calculations, *i.e.*, each of the dimers in the double helix obeys a characteristic set of force constants.¹³ The greater range of local distortions in the base-pair steps incorporated in the model has very little, if any, effect on the average distances between gold nanoparticles. The elastic constants of the dimers in the specific sequences, however, are lower, on average, than those governing the deformations of the mixed-sequence step, where the contributions of all 16 dinucleotides are equally weighted.¹⁴ These differences in local structural mobility underlie the enhanced variance of the simulated duplexes. In particular, the lateral shearing of adjacent base pairs along their long axes, *i.e.*, fluctuations in Slide, and the coupling of these motions with the preferential bending of DNA about the same axis (Roll), soften the apparent Young’s modulus of ‘real’ sequence-dependent vs. ideal DNA.²⁴ Moreover, omission of the fluctuations in Slide from the ‘real’ model reduces the variation in the distances between

gold nanocenters substantially. The contributions of lateral shearing to the extension of DNA suggest the underlying structural basis of the published rationalization⁸ of the distance fluctuations of short end-labeled DNA in terms of enhanced stretching.

Enhanced fluctuations in the extended tether also increase the range of distances between gold nanocrystals attached to ideal DNA chains. The mean offset of the probes from the helical axis (~ 14.0 Å), associated with the broader range of states sampled by the flexible tether with respect to a rigid helix, gives rise to a strong non-monotonic (sinusoidal-like) dependence of the end-to-end variance with chain length. The offset is smaller and the non-monotonic behavior accordingly less pronounced in the other models considered here (~ 8.5 Å for the ‘stiff’ tethers linked to the ends of ideal, mixed-sequence, and sequence-specific chains). For the reasons noted above, the variance is greater when the probes lie on opposing faces of the fluctuating helix.

Finally, the distributions of distances between nanocrystals with ‘stiff’ tethers at the ends of short, ideal, inextensible DNA chains are slightly skewed from ideal Gaussian curves (Fig. 6). That is, the modified duplexes tend to shorten rather than lengthen with respect to their most probable extension. The skewness becomes more pronounced with increase in chain length in rough correspondence with experiment (where secondary peaks of shorter chain extension appear in plots of relative abundance derived from the scattering data). The distributions associated with mixed-sequence DNA chains containing ‘stiff’ tethers roughly coincide with those shown for the ideal chains. The simulated curves, however, widen and shift to lower values of r_{Au} when sequence-dependent deformability is considered. The incorporation of ‘flexible’ tethers on ideal DNA models similarly broadens and shifts the end-to-end distributions while concomitantly enhancing the propensity of the chains to shorten. (See the plotted curves in Figs. S2–S4 in the Electronic Supplementary Information.)

Conclusions

The physical properties of DNA depend upon chain length. The dimensions of chains of a few hundred base pairs are typical of a wormlike coil that bends smoothly and gradually into compact forms.⁷ Because the deformations in three-dimensional structure used to account for this behavior are quite limited, short chain fragments are often modeled as rigid rods. As demonstrated herein, this oversimplification misses the key contribution of the natural dimeric flexibility of DNA to the end-to-end properties of chains of only a few helical turns. In particular, there is no need to posit enhanced cylindrical stretching fluctuations as the source of the recently measured quadratic dependence of the variance in DNA end-to-end distance on chain length.⁸ Mixed-sequence chains with the more restricted levels of stretching deduced from single-molecule experiments^{10–12} and analyses of high-resolution structures^{13,14} also show a quadratic increase in end-to-end variance with chain length (Fig. 2). Indeed, even ideal, inextensible DNA chains limited to isotropic bending and twisting fluctuations exhibit such behavior.

As chain length increases, the deformability in the added base-pair steps opens the range of three-dimensional forms available to the DNA helix. If the ends of the chain are separated by a sufficient number of intervening residues, the duplex exhibits ideal Gaussian (random coil) behavior. Thus, the variance in end-to-end distances is linear in longer chains (over the range of chain lengths where the twisted wormlike coil model is normally fitted to DNA properties) and levels off to a constant when the chain is very long.

Given the restrictions on local base-pair structure, the tethers used to attach chemical probes to short DNA may contribute to the detected dispersion of chain ‘ends’. For example, the

distribution of end-to-end distances is narrower (Fig. 3) and more closely correlated (Fig. 4) with the distances between terminal base-pair centers for probes that are directed along rather than perpendicular to the helical axis. The spin-labeled sugar-phosphate backbones probed in electron paramagnetic resonance studies^{4–6} fall in the latter category. The distributions of intramolecular distances extracted from such experiments thus reflect their relative helical positioning as well as any distortions imposed by chemical modification of the double-helical structure.

Surprisingly, the contributions of even perfectly ‘stiff’ tethers to the end-to-end dispersion of DNA chain ‘ends’ are nonlinear. That is, the difference in the variance in the distances between chemical probes and the variance in the distances between terminal base pairs increases with chain length (Fig. 5).

Fluctuations in the tether conformation also affect the distances and dispersion of chain ‘ends’ (Table 1). The average distance between chemical probes decreases and the dispersion increases if the probe lies close to the DNA helical axis in its equilibrium rest state. The distance between probes, however, may increase upon tether deformation if the probe lies far from the DNA axis in its rest state. The greater variance in DNA chain extension detected in fluorescence resonance energy transfer experiments³ compared to small-angle X-ray-scattering^{8,9} studies may thus reflect the enhanced flexibility of the longer tethers attached to fluorescent dyes compared to those used to link gold nanocrystals to DNA. The enhancement in ‘end’-to-‘end’ variance and differences in average ‘end’-to-‘end’ extension may be even greater if the tethers undergo large rearrangements between different conformational forms.

The sequence-dependent deformability of DNA base-pair steps may also contribute to the measured dispersion of chain ‘ends’ (Table 1). Like the effects of tether deformations, these effects are nonlinear and become more pronounced at longer chain lengths.

Thus, one can account for the observed distances between the ‘ends’ of short DNA chains in terms of the normal physical properties of the double helix and the chemical linkers used to attach various molecular probes (Fig. 6). The ranges of measured distances provide useful benchmarks for all-atom simulations of DNA. The coarse-grained approach used here does not differentiate among the many ways to fit the measured distances. Rather this work highlights the overlooked contribution of small room-temperature fluctuations on the configurational properties of short DNA duplexes and the importance of small conformational deformations in the interpretation of spectroscopic measurements of DNA chain extension.

Finally, a technical comment,³⁶ which addresses the effect of linker offset on the observed distances between gold nanocrystals, and a response from the authors,³⁷ which includes new data supporting the stretching arguments used originally to account for the build-up of the distances between tethered nanocrystals, appeared after this work was completed. The fluctuations in Slide incorporated in our ‘realistic’ models of short DNAs suggest the underlying structural basis for this phenomenological interpretation.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The U.S. Public Health Service (research grants GM20861 and GM34809, instrumentation grant RR022375) has generously supported this work. G.Z. acknowledges support of a predoctoral traineeship from the U.S. Public Health Service (New Interdisciplinary Research Workforce Training Grant DK071502).

References

- Cooper JP, Hagerman PJ. *Biochemistry*. 1990; 29:9261–9268. [PubMed: 2271593]
- Clegg RM, Murchie AIH, Zechel A, Lilley DMJ. *Proc Natl Acad Sci, USA*. 1993; 90:2994–2998. [PubMed: 8464916]
- Laurence TA, Kong X, Jager M, Weiss S. *Proc Natl Acad Sci, USA*. 2005; 102:17348–17353. [PubMed: 16287971]
- Schiemann O, Piton N, Mu Y, Stock G, Engels JW, Prisner TF. *J Am Chem Soc*. 2004; 126:5722–5729. [PubMed: 15125665]
- Cai Q, Kusnetzow AK, Hubbell WL, Haworth IS, Gacho GPC, Van Eps N, Hideg K, Chambers EJ, Qin PZ. *Nucleic Acids Res*. 2006; 34:4722–4730. [PubMed: 16966338]
- Ward R, Keeble DJ, El-Mkami H, Norman DG. *Chem BioChem*. 2007; 8:1957–1964.
- Shimada J, Yamakawa H. *Macromolecules*. 1984; 17:689.
- Mathew-Fenn RS, Das R, Harbury PAB. *Science*. 2008; 322:446–449. [PubMed: 18927394]
- Mathew-Fenn RS, Das R, Silverman JA, Walker PA, Harbury PAB. *PLoS One*. 2008; 3:e3229. [PubMed: 18927606]
- Smith SB, Cui Y, Bustamante C. *Science*. 1996; 271:795–799. [PubMed: 8628994]
- Wang MD, Yin H, Landick R, Gelles J, Block SM. *Biophys J*. 1997; 72:1335–1346. [PubMed: 9138579]
- Gore J, Bryant Z, Nollmann M, Le MU, Cozzarelli NR, Bustamante C. *Nature Chemical Biology*. 2006; 8:836–839.
- Olson WK, Gorin AA, Lu XJ, Hock LM, Zhurkin VB. *Proc Natl Acad Sci, USA*. 1998; 95:11163–11168. [PubMed: 9736707]
- Olson, WK.; Colasanti, AV.; Czapla, L.; Zheng, G. *Coarse-Graining of Condensed Phase and Biomolecular Systems*. Voth, GA., editor. Taylor and Francis Group, LLC; 2008. p. 205–223.
- Flory, PJ. *Statistical Mechanics of Chain Molecules*. Interscience Publishers; New York: 1969.
- Czapla L, Swigon D, Olson WK. *J Chem Theor Comp*. 2006; 2:685–695.
- Metropolis NA, Rosenbluth AW, Rosenbluth MN, Teller H, Teller E. *J Chem Phys*. 1953; 21:1087–1092.
- Hetenyi B, Bernacki K, Berne BJ. *J Chem Phys*. 2002; 117:8203–8207.
- Dickerson RE, Bansal M, Calladine CR, Diekmann S, Hunter WN, Kennard O, von Kitzing E, Lavery R, Nelson HCM, Olson WK, Saenger W, Shakked Z, Sklenar H, Soumpasis DM, Tung CS, Wang AHJ, Zhurkin VB. *J Mol Biol*. 1989; 208:787–791.
- Horowitz DS, Wang JC. *J Mol Biol*. 1984; 173:75–91. [PubMed: 6321743]
- Heath PJ, Clendenning JB, Fujimoto BS, Schurr JM. *J Mol Biol*. 1996; 260:718–730. [PubMed: 8709150]
- Zhurkin VB, Lysov YP, Ivanov VI. *Nucleic Acids Res*. 1979; 6:1081–1096. [PubMed: 440969]
- Tolstorukov MY, Colasanti AV, McCandlish D, Olson WK, Zhurkin VB. 2007 submitted.
- Matsumoto A, Olson WK. *Biophys J*. 2002; 83:22–41. [PubMed: 12080098]
- Lu XJ, Olson WK. *Nucleic Acids Res*. 2003; 31:5108–5121. [PubMed: 12930962]
- Lu XJ, Olson WK. *Nature Protocols*. 2008; 3:1213–1227.
- Olson WK, Bansal M, Burley SK, Dickerson RE, Gerstein M, Harvey SC, Heinemann U, Lu XJ, Neidle S, Shakked Z, Sklenar H, Suzuki M, Tung CS, Westhof E, Wolberger C, Berman HM. *J Mol Biol*. 2001; 313:229–237. [PubMed: 11601858]
- El Hassan MA, Calladine CR. *J Mol Biol*. 1995; 251:648–664. [PubMed: 7666417]

29. Jeffreys, H.; Jeffreys, BS. *Methods of Mathematical Physics*. Cambridge University Press; Cambridge, UK: 1946.
30. Fox T, Kollman PA. *J Phys Chem B*. 1998; 102:8070–8079.
31. Manning GS. *Quart Rev Biophys*. 1978:179–246.
32. Case, DA.; Darden, TA.; Cheatham, TE., III; Simmerling, CL.; Wang, J.; Duke, RE.; Luo, R.; Crowley, M.; Walker, RC.; Zhang, W.; Merz, KM.; Wang, B.; Hayik, S.; Roitberg, A.; Seabra, G.; Kolossvary, I.; Wong, KF.; Paesani, F.; Vanicek, J.; Wu, X.; Brozell, SR.; Steinbrecher, T.; Gohlke, H.; Yang, L.; Tan, C.; Mongan, J.; Hornak, V.; Cui, G.; Mathews, DH.; Seetin, MG.; Sagui, C.; Babin, V.; Kollman, PA. *AMBER 10*. University of California; San Francisco, San Francisco, CA: 2008.
33. Olson WK. *Biopolymers*. 1979; 18:1213–1233. [PubMed: 435613]
34. Hagerman PJ. *Biopolymers*. 1981; 20:1503–1535. [PubMed: 7023566]
35. Vologodskaya M, Vologodskii A. *J Mol Biol*. 2002; 317:205–213. [PubMed: 11902837]
36. Becker NB, Everaers R. 2009; 325:538.
37. Mathew-Fenn RS, Das R, Fenn TD, Schneiders M, Harbury PAB. *Science*. 2009; 325:538. [PubMed: 19644093]

Fig. 1.

(A) Atomic-level representations of a 15-bp DNA duplex with gold nanocrystals (large spheres) attached via short tethers to the 3'-ends of complementary strands. (B) Close-up of the tether highlighting the chemistry and dihedral angles of the linkage. Color-coding denotes atom type: carbon (pink); nitrogen (blue); oxygen (green); sulfur (yellow); phosphorus (violet). (C) Reference frame, called the 'middle' frame,²⁵ associated with a DNA base pair. Antiparallel directions of complementary strands are denoted by arrows.

Fig. 2.

Chain-length dependence of (A) the average end-to-end distances and contour lengths (points connected respectively by dashed and solid lines) and (B) the associated variances for two types of DNA — an ideal, inextensible, twisted wormlike chain (filled circles) and a mixed-sequence chain, which is naturally straight in its equilibrium rest state and subject to the deformational properties characteristic of high-resolution structures (open squares). Both models are naturally straight with 10.5 bp/turn and elastic constants scaled to yield a persistence length of ~ 500 Å.¹⁴

Fig. 3. Effects of tethers on the end-to-end variance of a 15-bp ideal, inextensible DNA duplex. Scatterplots depict the simulated positions of the centers of (A) terminal base pairs (blue dots) and (B, C) rigidly tethered gold (Au) nanocrystals (orange dots) with respect to the equilibrium structure of DNA. Tethers adopt (B) fully extended and (C) kinked forms. Normalized distributions of the corresponding (D) DNA...DNA and (E, F) Au...Au end-to-end distances (r_{DNA} and r_{Au}) that stem from base-pair-step deformations.

Fig. 4.

Scatter plots of the covariance of DNA and nanocrystal end-to-end distances — r_{DNA} and r_{Au} , for the (A) extended and (B) kinked tethers considered respectively in Figs. 3B/E and C/F and the covariance (C, D) of the corresponding global bending angles — Γ_{DNA} and Γ_{Au} for the same chains.

Fig. 5. Chain-length dependence of (A) the variance of the DNA-DNA and Au-Au end-to-end distances $\langle \delta r_{\text{DNA}}^2 \rangle$ and $\langle \delta r_{\text{Au}}^2 \rangle$ (filled and open circles, respectively) of an ideal, inextensible DNA duplex with gold nanocrystals configured along the same lines as Fig. 3B and (B) the differences between the two measurements.

Fig. 6. Simulated probability density distributions of the distances r_{Au} between gold nanocrystals attached via ‘stiff’, extended tethers to the ends of ideal, inextensible DNA duplexes of 10 bp (red), 15 bp (green), 20 bp (black), 25 bp (cyan), 30 bp (magenta), and 35 bp (blue). In contrast to the perfectly rigid tethers considered in Figs. 3–5, the system modeled here incorporates small fluctuations in the dihedral angle $\eta = 180 \pm 50^\circ$ and the length $b_{S-Au} = 7 \pm 1$ Å of the virtual bond between sulphur and the center of the gold nanocrystal. See Table 1 for the means and variances of these normalized profiles.

Table 1

Observed vs. computed distances and fluctuations between gold nanocrystals attached to DNA chains.

N	Exp	I+s	M+s	S+s	I+f
$\langle r_{Au} \rangle$					
10	55.7±0.3	53.9±0.5	53.3±0.6	52.5±0.6	48.5±0.8
15	69.7±0.4	69.6±0.6	69.5±0.4	67.4 ± 0.5	60.5±2.7
20	86±0.4	85.2±0.4	85.7±0.4	83.8±0.6	78.5±1.9
25	101±0.5	101.1±1.0	101.2±0.5	98.3±0.6	92.6±2.5
30	119.1±0.6	117.5±0.3	117.1±0.6	113.0±0.7	110.8±1.7
35	131.3±0.7	132.3±0.6	133.0±0.6	127.5±0.7	124.6±1.9
$\langle \delta r_{Au}^2 \rangle$					
10	8.5±0.6	17.5±2.4	20.4±2.9	22.8±3.7	27.4±6.4
15	16.5±1.1	22.1±1.5	21.8±1.0	27.4±1.3	54.0±21.5
20	21.6±1.4	22.9±3.1	24.2±2.1	32.0±3.3	41.8±7.6
25	30.0±2.0	28.4±3.6	30.6±3.1	51.2±5.4	57.7±16.8
30	41.1±2.7	31.8±1.3	33.7±2.4	59.7±3.0	44.0±8.2
35	50.9±3.4	42.2±3.7	41.6±3.5	86.9±4.3	70.6±16.3

Abbreviations: N: DNA chain length (bps); Exp: experimental observation; I+s: ideal, inextensible DNA with 'stiff' tethers; M+s: mixed-sequence DNA with 'stiff' tethers; S+s: sequence-dependent DNA with 'stiff' tethers; I+f: ideal, inextensible DNA with 'flexible' tethers.