Supporting Material

Mechanical properties of base-modified DNA are not strictly determined by base stacking or electrostatic interactions

Justin P. Peters,[†] Lauren S. Mogil,[†] Micah J. McCauley,[‡] Mark C. Williams,[‡] and L. James Maher, III[†*]

[†]Department of Biochemistry and Molecular Biology
Mayo Clinic College of Medicine
200 First St. SW, Rochester, MN 55905, USA
[‡]Department of Physics
Northeastern University
111 Dana Research Center, Boston, MA 02115, USA
[*]Corresponding Author. E-mail: maher@mayo.edu.

S.1 DNA Characterization

DNA Constructs

Cloned intrinsically straight ∼200-bp sequences were created using unique 5-bp direct repeats to eliminate long-range sequence-directed curvature (1) and were the kind gift of A. Vologodskii. Figure S.1 shows the longest construct in a series with flanking *Nar*I sites (2), designated pJ1506.

Plasmid pJ1506

```
  1  CGGTGATGAC GGTGAAAACC TCTGACACAT GCAGCTCCCG GAGACGGTCA CAGCTTGTCT GTAAGCGGAT GCCGGGAGCA
     GCCACTACTG CCACTTTTGG AGACTGTGTA CGTCGAGGGC CTCTGCCAGT GTCGAACAGA CATTCGCCTA CGGCCCTCGT
     LJM-4762-- ---------- ---------- ---------- ---------- ---------- ---------- ----------
     LJM-4762-- ---------- ---------- ---------- ---------- ---------- ---------- ----------

 81  GACAAGCCCG TCAGGGCGCG TCAGCGGGTG TTGGCGGGTG TCGGGGCTGG CTTAACTATG CGGCATCAGA GCAGATTGTA
     CTGTTCGGGC AGTCCCGCGC AGTCGCCCAC AACCGCCCAC AGCCCCGACC GAATTGATAC GCCGTAGTCT CGTCTAACAT
     ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
     ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------

161  CTGAGAGTGC ACCATATGCG GTGTGAAATA CCGCACAGAT GCGTAAGGAG AAAATACCGC ATCAGGCGCC ATTCGCCATT
     GACTCTCACG TGGTATACGC CACACTTTAT GGCGTGTCTA CGCATTCCTC TTTTATGGCG TAGTCCGCGG TAAGCGGTAA
     ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
     ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------

241  CAGGCTGCGC AACTGTTGGG AAGGGCGATC GGTGCGGGCC TCTTCGCTAT TACGCCAGCT GGCGAAAGGG GGATGTGCTG
     GTCCGACGCG TTGACAACCC TTCCCGCTAG CCACGCCCGG AGAAGCGATA ATGCGGTCGA CCGCTTTCCC CCTACACGAC
     ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
     ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------

321  CAAGGCGATT AAGTTGGGTA ACGCCAGGGT TTTCCCAGTC ACGACGTTGT AAAACGACGG CCAGTGAATT CGAGCTCGGT
     GTTCCGCTAA TTCAACCCAT TGCGGTCCCA AAAGGGTCAG TGCTGCAACA TTTTGCTGCC GGTCACTTAA GCTCGAGCCA
     ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
     ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
                     LJM-3222-------- ---------- ---------- ---------- ---------- ----------

401  ACCCGGGGAT CCTCTCGCGG CGCCCGCGAC TCGAGCCTAG CCTATGACAT GACACGTTAC GTTAGTCGAG TCGATCAGAT
     TGGGCCCCTA GGAGAGCGCC GCGGGCGCTG AGCTCGGATC GGATACTGTA CTGTGCAATG CAATCAGCTC AGCTAGTCTA
     ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
     ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
     ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------

481  CAGACGCTAC GCTAGCTGAG CTGACTGTAC TGTATGCAAT GCAACCTCAC CTCAGGACAG GACACGTGAC GTGATGCTAT
     GTCTGCGATG CGATCGACTC GACTGACATG ACATACGTTA CGTTGGAGTG GAGTCCTGTC CTGTGCACTG CACTACGATA
     tweezers-- ---------- ---------- ---------- ---------- ---------- ---------- ----------
     AFM------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
     cyclization---------- ---------- ---------- ---------- ---------- ---------- ----------

561  GCTACCAGAC CAGCTGCACT GCAGACTGGA CTGACGCTAC GCTATCGCAT CGCAGATGAG ATGAAGCCGG GCGCCGCCAT
     CGATGGTCTG GTCGACGTGA CGTCTGACCT GACTGCGATG CGATAGCGTA GCGTCTACTC TACTTCGGCC CGCGGCGGTA
     ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
     ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
     ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------

641  GGTCATAGCT GTTTCCTGTG TGAAATTGTT ATCCGCTCAC AATTCCACAC AACATACGAG CCGGAAGCAT AAAGTGTAAA
     CCAGTATCGA CAAAGGACAC ACTTTAACAA TAGGCGAGTG TTAAGGTGTG TTGTATGCTC GGCCTTCGTA TTTCACATTT
     ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
     ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
     ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------

721  GCCTGGGGTG CCTAATGAGT GAGCTAACTC ACATTAATTG CGTTGCGCTC ACTGCCCGCT TTCCAGTCGG GAAACCTGTC
     CGGACCCCAC GGATTACTCA CTCGATTGAG TGTAATTAAC GCAACGCGAG TGACGGGCGA AAGGTCAGCC CTTTGGACAG
     ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
     ---------- ---------- ------LJM-3223
     ---------- ---------- ------LJM-3223
```

```
 801  GTGCCAGCTG CATTAATGAA TCGGCCAACG CGCGGGGAGA GGCGGTTTGC GTATTGGGCG CTCTTCCGCT TCCTCGCTCA
      CACGGTCGAC GTAATTACTT AGCCGGTTGC GCGCCCCTCT CCGCCAAACG CATAACCCGC GAGAAGGCGA AGGAGCGAGT
      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------

 881  CTGACTCGCT GCGCTCGGTC GTTCGGCTGC GGCGAGCGGT ATCAGCTCAC TCAAAGGCGG TAATACGGTT ATCCACAGAA
      GACTGAGCGA CGCGAGCCAG CAAGCCGACG CCGCTCGCCA TAGTCGAGTG AGTTTCCGCC ATTATGCCAA TAGGTGTCTT
      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------

 961  TCAGGGGATA ACGCAGGAAA GAACATGTGA GCAAAAGGCC AGCAAAAGGC CAGGAACCGT AAAAAGGCCG CGTTGCTGGC
      AGTCCCCTAT TGCGTCCTTT CTTGTACACT CGTTTTCCGG TCGTTTTCCG GTCCTTGGCA TTTTTCCGGC GCAACGACCG
      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------

1041  GTTTTTCCAT AGGCTCCGCC CCCCTGACGA GCATCACAAA AATCGACGCT CAAGTCAGAG GTGGCGAAAC CCGACAGGAC
      CAAAAAGGTA TCCGAGGCGG GGGGACTGCT CGTAGTGTTT TTAGCTGCGA GTTCAGTCTC CACCGCTTTG GGCTGTCCTG
      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------

1121  TATAAAGATA CCAGGCGTTT CCCCCTGGAA GCTCCCTCGT GCGCTCTCCT GTTCCGACCC TGCCGCTTAC CGGATACCTG
      ATATTTCTAT GGTCCGCAAA GGGGGACCTT CGAGGGAGCA CGCGAGAGGA CAAGGCTGGG ACGGCGAATG GCCTATGGAC
      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------

1201  TCCGCCTTTC TCCCTTCGGG AAGCGTGGCG CTTTCTCATA GCTCACGCTG TAGGTATCTC AGTTCGGTGT AGGTCGTTCG
      AGGCGGAAAG AGGGAAGCCC TTCGCACCGC GAAAGAGTAT CGAGTGCGAC ATCCATAGAG TCAAGCCACA TCCAGCAAGC
      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------

1281  CTCCAAGCTG GGCTGTGTGC ACGAACCCCC CGTTCAGCCC GACCGCTGCG CCTTATCCGG TAACTATCGT CTTGAGTCCA
      GAGGTTCGAC CCGACACACG TGCTTGGGGG GCAAGTCGGG CTGGCGACGC GGAATAGGCC ATTGATAGCA GAACTCAGGT
      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------

1361  ACCCGGTAAG ACACGACTTA TCGCCACTGG CAGCAGCCAC TGGTAACAGG ATTAGCAGAG CGAGGTATGT AGGCGGTGCT
      TGGGCCATTC TGTGCTGAAT AGCGGTGACC GTCGTCGGTG ACCATTGTCC TAATCGTCTC GCTCCATACA TCCGCCACGA
      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------

1441  ACAGAGTTCT TGAAGTGGTG GCCTAACTAC GGCTACACTA GAAGGACAGT ATTTGGTATC TGCGCTCTGC TGAAGCCAGT
      TGTCTCAAGA ACTTCACCAC CGGATTGATG CCGATGTGAT CTTCCTGTCA TAAACCATAG ACGCGAGACG ACTTCGGTCA
      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------

1521  TACCTTCGGA AAAAGAGTTG GTAGCTCTTG ATCCGGCAAA CAAACCACCG CTGGTAGCGG TGGTTTTTTT GTTTGCAAGC
      ATGGAAGCCT TTTTCTCAAC CATCGAGAAC TAGGCCGTTT GTTTGGTGGC GACCATCGCC ACCAAAAAAA CAAACGTTCG
      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------

1601  AGCAGATTAC GCGCAGAAAA AAAGGATCTC AAGAAGATCC TTTGATCTTT TCTACGGGGT CTGACGCTCA GTGGAACGAA
      TCGTCTAATG CGCGTCTTTT TTTCCTAGAG TTCTTCTAGG AAACTAGAAA AGATGCCCCA GACTGCGAGT CACCTTGCTT
      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------

1681  AACTCACGTT AAGGGATTTT GGTCATGAGA TTATCAAAAA GGATCTTCAC CTAGATCCTT TTAAATTAAA AATGAAGTTT
      TTGAGTGCAA TTCCCTAAAA CCAGTACTCT AATAGTTTTT CCTAGAAGTG GATCTAGGAA AATTTAATTT TTACTTCAAA
      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------

1761  TAAATCAATC TAAAGTATAT ATGAGTAAAC TTGGTCTGAC AGTTACCAAT GCTTAATCAG TGAGGCACCT ATCTCAGCGA
      ATTTAGTTAG ATTTCATATA TACTCATTTG AACCAGACTG TCAATGGTTA CGAATTAGTC ACTCCGTGGA TAGAGTCGCT
      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------

1841  TCTGTCTATT TCGTTCATCC ATAGTTGCCT GACTCCCCGT CGTGTAGATA ACTACGATAC GGGAGGGCTT ACCATCTGGC
      AGACAGATAA AGCAAGTAGG TATCAACGGA CTGAGGGGCA GCACATCTAT TGATGCTATG CCCTCCCGAA TGGTAGACCG
      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------

1921  CCCAGTGCTG CAATGATACC GCGAGACCCA CGCTCACCGG CTCCAGATTT ATCAGCAATA AACCAGCCAG CCGGAAGGGC
      GGGTCACGAC GTTACTATGG CGCTCTGGGT GCGAGTGGCC GAGGTCTAAA TAGTCGTTAT TTGGTCGGTC GGCCTTCCCG
      ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------

2001  CGAGCGCAGA AGTGGTCCTG CAACTTTATC CGCCTCCATC C
      GCTCGCGTCT TCACCAGGAC GTTGAAATAG GCGGAGGTAG G
      ---------- ---------- ---------- ----LJM-4763
```

Figure S.1: Plasmid pJ1506. Forward (LJM-3222, LJM-4762) and reverse (LJM-3223, LJM-4763) primers are used to PCR amplify fragments for tweezers (cyan), AFM (magenta), and cyclization (orange), all of which share central *Nar*I restriction sites (bold red) flanking the 5-bp direct repeats.

Thermal Denaturation

Stability of Inosine-substituted DNA Duplex

Thermal denaturation experiments monitored SYBR Green I (Invitrogen) fluorescence using an ICycler thermocycler (BioRad) over a temperature range of 50–100°C, measurements collected every 0.1°C with a temperature slope of 24°C/h. DNA fragments 418 base pairs in length with inosine completely replacing guanosine were PCR amplified from pJ1506 using conditions described in Materials and Methods (main manuscript). Samples (20 μL) contained 100 ng DNA (∼30 nM) in 10 mM sodium cacodylate, pH 6.6 with 10 mM NaCl and either 0.06X, 0.08X, 0.1X, or 0.2X SYBR Green I (Invitrogen).

The melting temperature (T_m) and change in standard free energy (ΔG°), enthalpy (ΔH°), and entropy (ΔS°) of a melting reaction were determined using the classic van't Hoff equation assuming a two-state transition model as described previously (2). This analysis assumes that the van't Hoff plots are absolutely linear, that ΔH° and ΔS° are temperature independent, and that there is no change in heat capacity for the melting transition, i.e. $\Delta C_p = 0$. Slight curvature in van't Hoff plots due to $\Delta C_p \neq 0$ can lead to significant errors in graphically evaluated thermodynamic parameters. However, since the enthalpies and entropies derived in this graphical manner are correlated, the relative error of ΔG° is typically smaller than the relative errors of ΔH° and ΔS°.

To eliminate the effects of potential binding affinity differences of the dye, each experiment was performed in triplicate for a given dye concentration, and dye-free values of the parameters were determined by linear extrapolation to zero dye concentration. The current and previously reported (2) parameter estimates in Table S.1 are the mean and standard deviation from three independent repeats of this extrapolation procedure.

Table S.1: Estimates of thermodynamic parameters from thermal denaturation of inosine-substitued DNA. The reported values of melting temperature (T_m) and change in standard free energy (ΔG°), enthalpy (ΔH°), and entropy (ΔS°) of the dissociation reaction are the mean and standard deviation from three independent repeats. Data collected for inosine-substituted DNA are compared to data previously reported (2) for natural and diaminopurine-substituted DNA.

DNA	T_m (°C)	ΔH° (kcal mol^{-1})	ΔS° (cal mol^{-1} K^{-1})	ΔG°_{37} (kcal mol^{-1})
natural	87.1 ± 0.5	400 ± 50	1100 ± 130	56 ± 7
diaminopurine	93.0 ± 0.5	490 ± 30	1350 ± 80	76 ± 5
inosine	72.8 ± 2.3	370 ± 30	1060 ± 90	45 ± 4

Inosine Free Energy of Stacking

The stacking ability of an unpaired nucleotide (X) was evaluated from "dangling end" thermodynamic measurements of the self-complementary DNA oligonucleotide 5′-XCGCGCG (2). The oligo with X = 2′-deoxyinosine (I) was purchased from TriLink BioTechnologies. Thermal denaturation experiments monitored SYBR Green I (Invitrogen) fluorescence using an ICycler thermocycler (BioRad) over a temperature range of 20–80°C, measurements collected every 0.1°C with a temperature slope of 24°C/h. Samples (20 μL) contained 300–900 ng DNA (∼5–25 μM) in 10 mM sodium phosphate, pH 7.0 with 1 M NaCl and either 0.05X, 0.1X, 0.15X, or 0.2X SYBR Green I (Invitrogen). The thermal denaturation data were processed following a two-state transition model as described previously (2). Briefly, the two states are referred to as "folded" (fully associated) and

"unfolded" (fully dissociated) for simplicity. The data were converted from fluorescence (F) as a function of temperature (T) to fraction folded (θ) as a function of temperature

$$\theta(T) = \frac{F(T) - u(T)}{f(T) - u(T)} \tag{S.1}$$

where the unfolded (u) and folded (f) baselines were determined from the data by linear regression. Within the analysis interval ($0.15 < \theta < 0.85$) the melting temperature (T_m) was evaluated from $\theta = 0.5$. To eliminate the effects of potential binding affinity differences of the dye, each experiment was performed in triplicate for a given dye concentration, and dye-free values of T_m were determined by linear extrapolation to zero dye concentration. Values for the thermodynamic parameters (assumed independent of temperature) were determined by fitting (linear regression analysis) plots of T_m^{-1} as a function of the natural logarithm of DNA concentration $\ln(C_t)$

$$\frac{1}{T_m} = \frac{R}{\Delta H^\circ} \ln C_t + \frac{\Delta S^\circ}{\Delta H^\circ} \tag{S.2}$$

where R is the ideal gas constant (details about this variation of the van't Hoff equation have been discussed elsewhere) (2). Finally, ΔG° was determined from the Gibbs free energy equation

$$\Delta G^\circ = \Delta H^\circ - T\Delta S^\circ \tag{S.3}$$

Figure S.2 compares the new data for I to data previously reported (2).

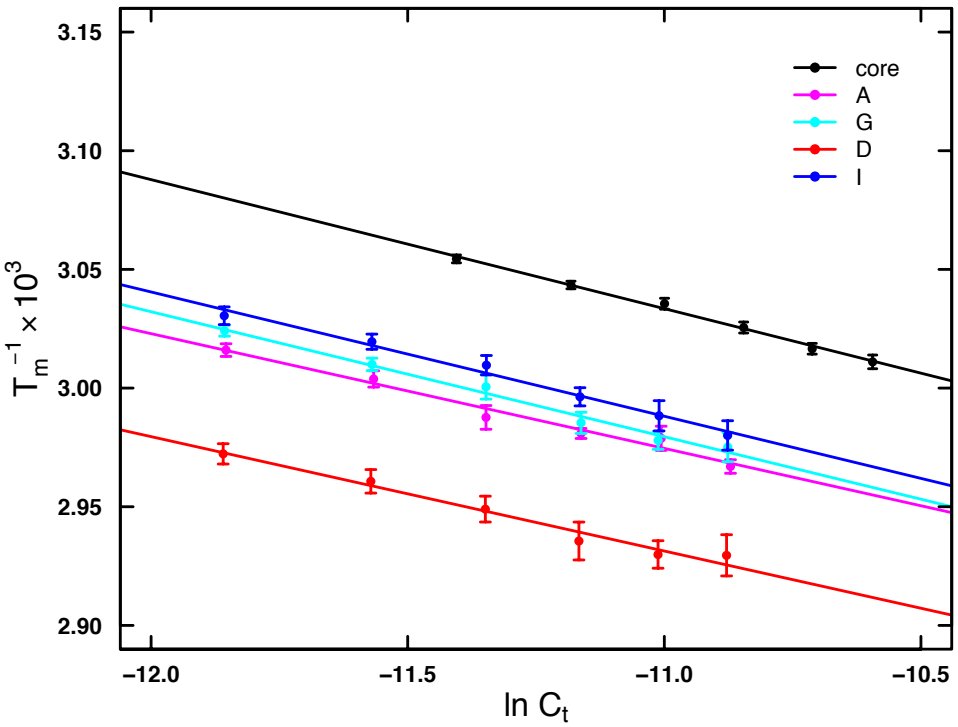


Figure S.2: van't Hoff plots of dangling DNA nucleotides. Data collected for I are compared to data previously reported (2).

The current and previously reported (2) parameter values in Table S.2 are the mean and standard deviation from three independent repeats. Stabilization of the core sequence duplex from base stacking is reported as both ΔT_m and $\Delta\Delta G^\circ$. While I stabilized the core hexamer duplex, it displayed the least increase in T_m, only 5.1°C, and least favorable free energy of stacking ($\Delta\Delta G^\circ_{37}$), only 0.7 kcal/mol. Thus, in dangling end experiments I is a poorer stacker than the G it replaces. As reported previously (2), D is a better stacker than the A it replaces, and the overall stacking order for these nucleotides is D > A > G > I.

Table S.2: Free energy of stacking ($\Delta\Delta G^\circ_{37}$) for inosine (I) determined from dangling end thermal denaturation. Data collected for I are compared to data previously reported (2).

X	T_m (°C) for 5 μM DNA	ΔT_m (°C)	ΔH° (kcal/mol)	ΔS° (cal/mol-K)	ΔG°_{37} (kcal/mol)	$\Delta\Delta G^\circ_{37}$ (kcal/mol)
core	49.5 ± 0.4	—	37 ± 2	89 ± 6	8.9 ± 0.1	—
A	56.6 ± 0.3	7.1 ± 0.5	41 ± 3	101 ± 10	10.0 ± 0.2	1.1 ± 0.2
G	55.5 ± 0.4	6.0 ± 0.6	38 ± 4	92 ± 13	9.7 ± 0.3	0.8 ± 0.3
D	61.3 ± 0.6	11.8 ± 0.7	42 ± 7	101 ± 22	10.6 ± 0.6	1.7 ± 0.6
I	54.6 ± 0.5	5.1 ± 0.6	38 ± 4	93 ± 13	9.6 ± 0.3	0.7 ± 0.3

S.2 Atomic Force Microscopy (AFM)

Background

Biopolymers like DNA can be visualized using atomic force microscopy (AFM), with the goal of deriving mechanical properties (e.g. persistence length). Since AFM is a surface technique, however, the way in which the molecules adhere to the surface must be considered. The high charge density of DNA allows for its immobilization onto a planar two-dimensional surface by virtue of ionic interactions between DNA phosphates and surface charges. The adsorption of DNA onto (negatively-charged) mica surfaces is typically promoted using either divalent cations or polyamines. Two modes of DNA adsorption have been proposed (3): i) strong adsorption without equilibration (i.e. kinetic trapping) leading to a conformation reflecting a projection of the DNA conformation in solution onto a plane and ii) weak adsorption so that the DNA molecule has time to freely equilibrate on the surface before immobilization in a particular two-dimensional equilibrium conformation. These processes diminish the number of possible DNA configurations; therefore, any quantitative description of the apparent DNA conformation confined to a plane requires two-dimensional (2D) reformulation of the existing models for three-dimensional (3D) conformations. In particular, the wormlike chain (WLC) model has proven to be an effective theory of DNA mechanics that can successfully describe experiments on single DNA molecules in two or three dimensions (4).

Image Processing and Data Analysis

Intrinsically straight DNA fragments 753 bp in length and containing the desired base substitutions were prepared for AFM studies. The 3D image in Figure S.3 shows 2D equilibrium conformations of

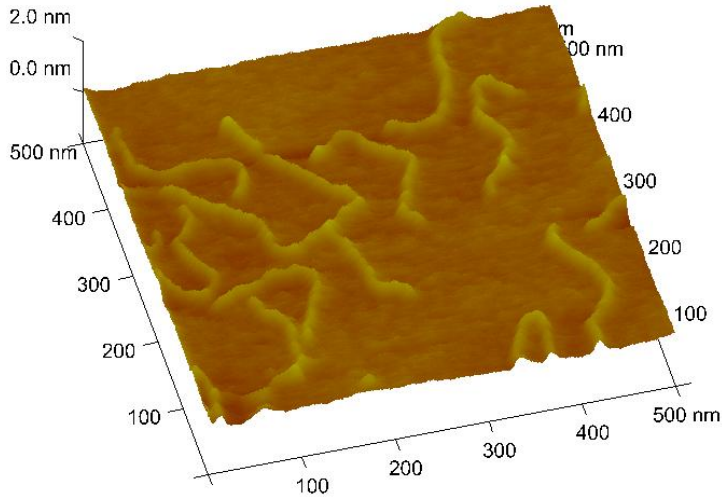
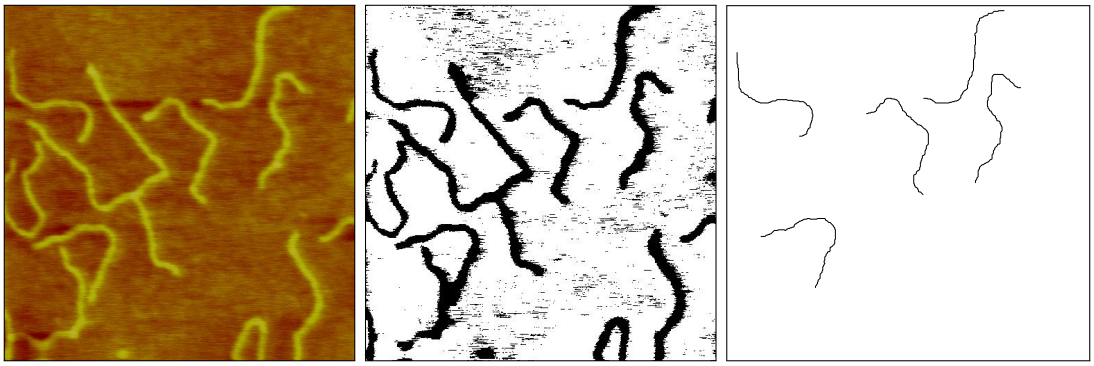


Figure S.3: AFM image.

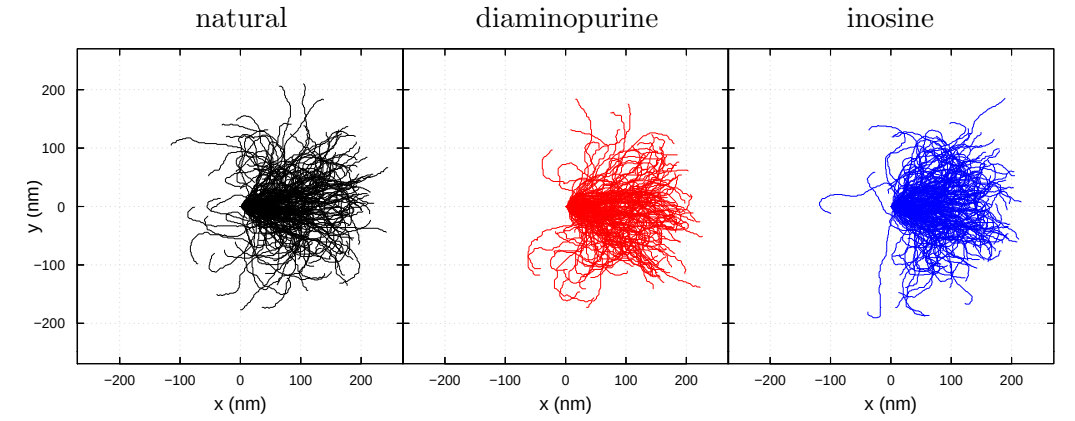
these DNA fragments when subjected to thermal fluctuations. From a large set of images collected for each DNA variant, the DNA molecules were digitized into pixel skeletons using an algorithm previously described (5) and illustrated in Figure S.4.



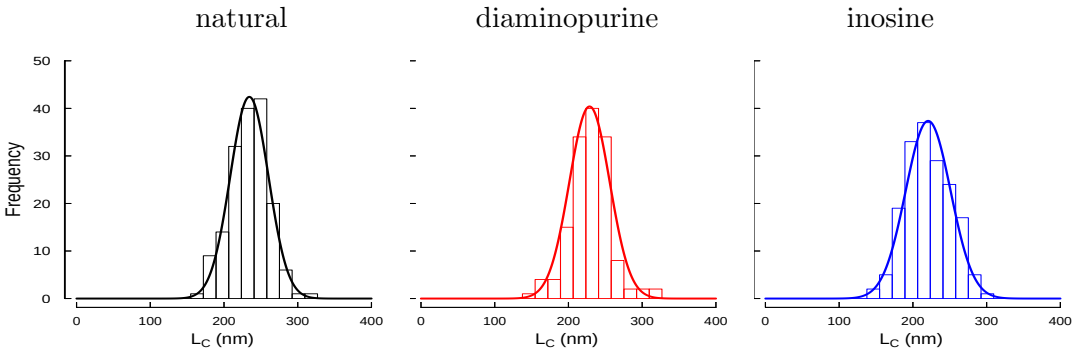
(a) (b) (c)

Figure S.4: Image processing. AFM images (a) were transformed into bi-color maps (b) with a threshold. Pixels above the threshold were then iteratively eroded from the edge of the DNA segments, if the removal of a pixel did not severe the segment. This thinning process was repeated until no more pixels could be removed, leaving behind DNA skeletons only one pixel wide (c) for later analysis.

The generated sets of pixel coordinates (skeletons) were aligned along the positive x-axis in Figure S.5(a), graphically illustrating the bending stiffness. Next, the skeletons were analyzed to determine the distribution of DNA contour lengths using the (n_e, n_o, n_c)-based corner chain estimator (6). When the next pixel in the DNA skeleton only has one coordinate (x or y) different from the previous pixel, the segment between two pixels is considered even. If both coordinates (x and y) are different from the previous pixel, the segment is considered odd. If moving from



(a)



(b)

Figure S.5: Analysis of DNA pixel skeletons.

one pixel to the next there is an odd to even or even to odd transition, the segment is treated as a corner. Therefore, (n_{e}, n_{o}, n_{c}) represents the number of even, odd, and corner segments in the DNA skeleton. For the corner chain estimator (6), the contour length is given by

$$L_C = 0.980n_{\mathrm{e}} + 1.406n_{\mathrm{o}} - 0.091n_{\mathrm{c}} \tag{S.4}$$

The values of L_C appear normally distributed, as shown in Figure S.5(b), with mean and standard deviation reported in Table S.3 for the indicated number of skeletons, N. Only DNA molecules satisfying the following criteria were included in the data set for analysis: i) both ends were visible and the molecule was not in touch with (or crossed over by) any other molecules and ii) the estimated contour length L_C fell within the range 100–400 nm (bounds chosen given the expected contour length of ∼250 nm for natural DNA). Although the detection and thinning of the molecules was automated, a human supervisor could reject erroneously segmented skeletons or those not meeting the above criteria during interactive steps.

Table S.3: Summary of DNA pixel skeletons

DNA variant	N	L_C (nm)	h (Å/bp)
natural	166	234 ± 26	3.11 ± 0.34
diaminopurine	146	229 ± 27	3.04 ± 0.36
inosine	172	221 ± 30	2.93 ± 0.40

The value for DNA helical rise (h) was determined by dividing the average estimated contour length by the expected total number of base pairs, Table S.3. As reported in previous studies (3, 6, 7, 8) estimates of helical rise determined using tapping-mode AFM in air (e.g. $h = 3.11 \pm 0.34$ Å/bp for natural DNA) underestimate the 3.38 Å/bp value measured by crystallography.

After detection of DNA skeletons, the trajectories of DNA centerlines were extracted automatically (but with human supervision) using a routine developed with the particular goal of analyzing local bend angles (9), as shown in Figure S.6. This initial analysis revealed the importance of step

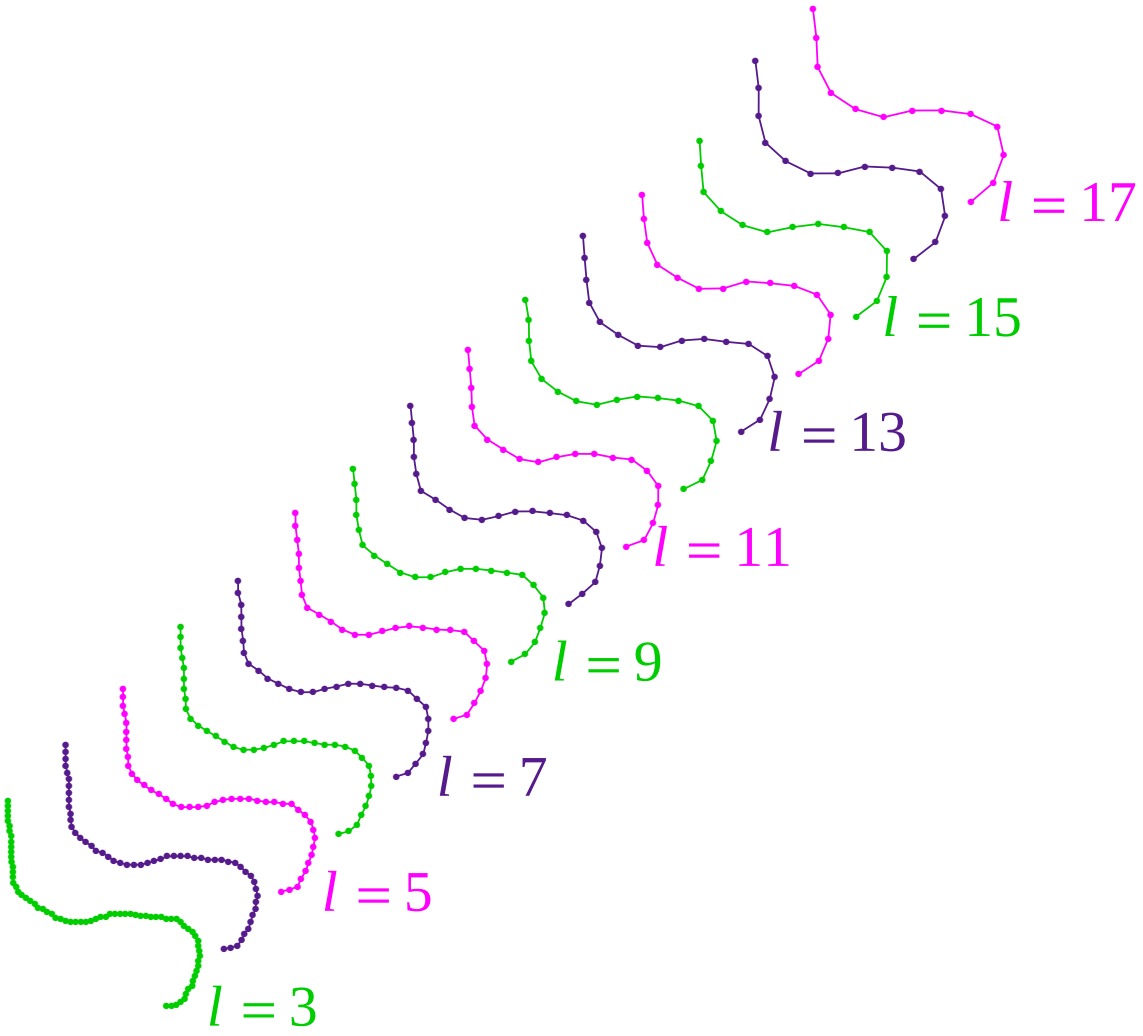


Figure S.6: Generating DNA centerline representations. Using the algorithm of Wiggins *et al.* (9) along with start and end points[1] defined from the pixel skeletons, (x, y) coordinate pairs separated by step size l were generated for each DNA molecule using each of $l = 3, 4, 5, \ldots, 17$ nm.

sizes in the range 4–9 nm so that subsequent coordinate representations were generated for each of $l = 4, 4.25, 4.5, \ldots, 9$ nm. From a set of coordinates (which can be thought of as a representation of the DNA axis) we computed statistical quantities such as the the Euclidean distance between

[1]The true end point actually lies within a circular region centered on the pixel end point with radius equal to one percent of the chosen step size.

the first and the last points of a DNA representation with contour length separation distance d (where d is an obligate multiple of l), the so called end-end distance $R(d)$.[2] Extra precaution was taken during sampling to avoid intrinsic correlations. It is known that reusing data (i.e. computing a quantity for each separation distance along a DNA contour) yields highly correlated points (7). To avoid this, each DNA representation was randomly dividend into shorter length segments such that no piece was used twice and the maximal segment length was restricted to 80% of the expected contour length (∼200 nm). This yielded a logarithmic distribution of the segment lengths, Figure S.7. Additionally, non-overlapping logarithmic sampling of the DNA representations allowed

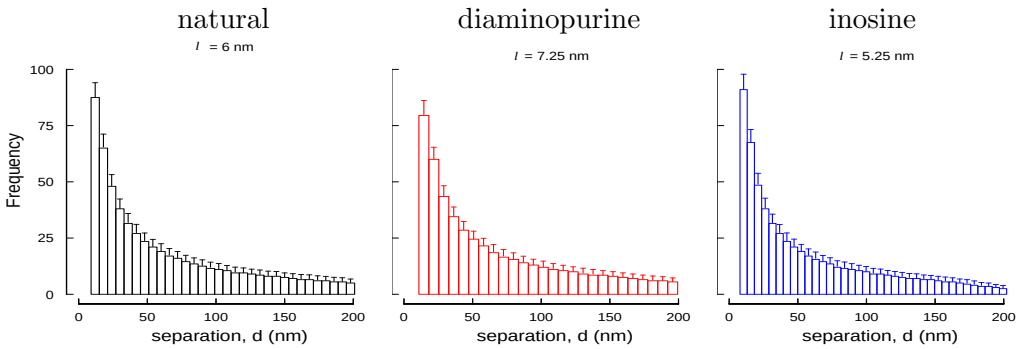


Figure S.7: Non-overlapping logarithmic sampling of DNA representations generated with the indicated step size l.

for estimation of the statistical reliability of the measurements. The error bars in Figure S.7 (and Figure S.9 – Figure S.16) show one standard deviation computed from 1000 different random draws from the same set of DNA representations. For the quantities that are quotients we applied error propagation to determine the standard deviation.

The measured quantities were then fit with corresponding equations from wormlike chain (WLC) theory. The WLC model captures the entropic elastic behavior of a semi-flexible biopolymer through idealization as an intrinsically straight, inextensible elastic rod (10). In fact, the WLC model is the continuous limit of the discrete freely rotating chain, as illustrated in Figure S.8. For a chain of contour length L ($L \equiv nl$ for the discrete chain of n segments of length l), the end-end distance vector $\mathbf{R}$ is given from the unit tangent vector $\mathbf{u}(s)$ at arc length position s by

$$\int_0^L \mathbf{u}(s)\, ds = \mathbf{R} \tag{S.5}$$

In the WLC model, bending deformations induced by thermal fluctuations (k_B is the Boltzmann constant and T is the absolute temperature) are attributed to a classical (obeying Hooke's law) elastic energy function

$$\frac{E_{\mathrm{WLC}}}{k_B T} = \frac{P\theta^2}{2d} \tag{S.6}$$

where the elastic bend constant P controls the exponential decay of the orientational correlation between unit tangent vectors separated by curvilinear (arc length) distance d with angle θ between them (Figure S.8)

$$\langle \mathbf{u}(s') \cdot \mathbf{u}(s'+d) \rangle = e^{-\frac{d}{P}} \tag{S.7}$$

[2]Clearly, $R \equiv R(d)$ is a function of arc length separation d, but only R is written for simplicity. Similarly, $\theta(d)$ and other length and angle moments will be assumed implicit functions of d.

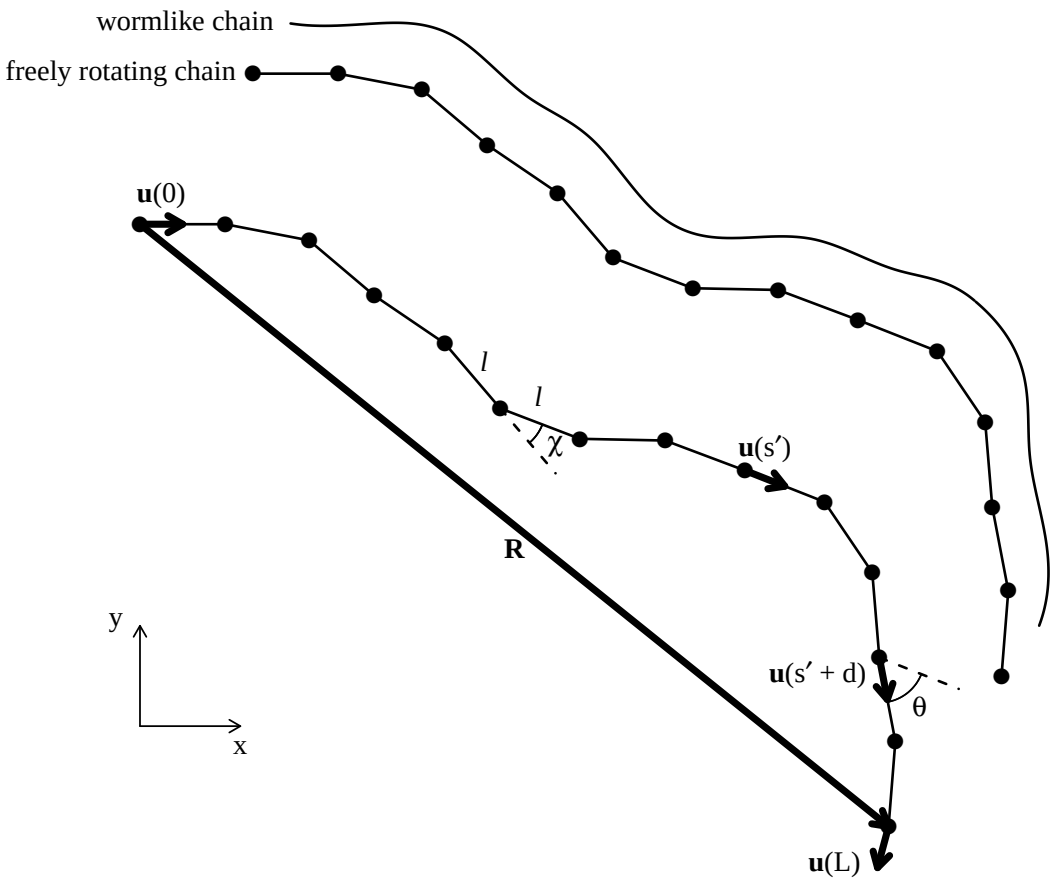


Figure S.8: Comparison of the continuous wormlike chain (WLC) with the discrete freely rotating chain (FRC) and illustration of the geometric quantities discussed in the text.

and the $\langle\ \rangle$ notation stands for averaging over thermal realizations. Equation S.7 can be reinterpreted using the definition of scalar projection as

$$\langle \cos\theta \rangle = e^{-\frac{d}{P}} \tag{S.8}$$

The elastic bend constant P is referred to as the *persistence length* (length through which the memory of the initial orientation of the chain persists) and quantifies the chain's resistance to bending. The bending energy function E_{WLC} is called classical (or harmonic) because it is a quadratic function of the bending strain variable θ. The probability distribution of θ (angular distribution function) is given by Boltzmann statistics

$$G(\theta; d, P) = \frac{e^{-\frac{P\theta^2}{2d}}}{\int_{-\infty}^{\infty} e^{-\frac{P(\theta')^2}{2d}} d\theta'} = \sqrt{\frac{P}{2\pi d}}\ e^{-\frac{P\theta^2}{2d}} \tag{S.9}$$

Thus, the harmonic energy function makes the WLC angular distribution Gaussian. However, even if E is chosen to be non-harmonic, the angular distribution will approach a Gaussian form at large separations d because the iterated convolution of any distribution with itself converges to a Gaussian form. That is, even if non-harmonic elastic behavior is present, it will be hidden on long length scales by thermal fluctuations. The prediction that $G(\theta)$ is normally distributed can

be explored by examining the moments of θ, which are given by

$$\langle \theta^m \rangle = \int_{-\infty}^{\infty} \theta^m G(\theta)\, d\theta \tag{S.10}$$

All odd moments are zero (regardless of dimensionality of the molecules)

$$\langle \theta \rangle = \langle \theta^3 \rangle = \langle \theta^5 \rangle = \ldots = 0 \tag{S.11}$$

as Figure S.9 demonstrates

$$\frac{\langle \theta \rangle}{\langle \theta^2 \rangle^{\frac{1}{2}}} \approx \frac{\langle \theta^3 \rangle}{\langle \theta^2 \rangle^{\frac{3}{2}}} \approx \frac{\langle \theta^5 \rangle}{\langle \theta^2 \rangle^{\frac{5}{2}}} \approx \frac{\langle \theta^7 \rangle}{\langle \theta^2 \rangle^{\frac{7}{2}}} \approx 0 \tag{S.12}$$

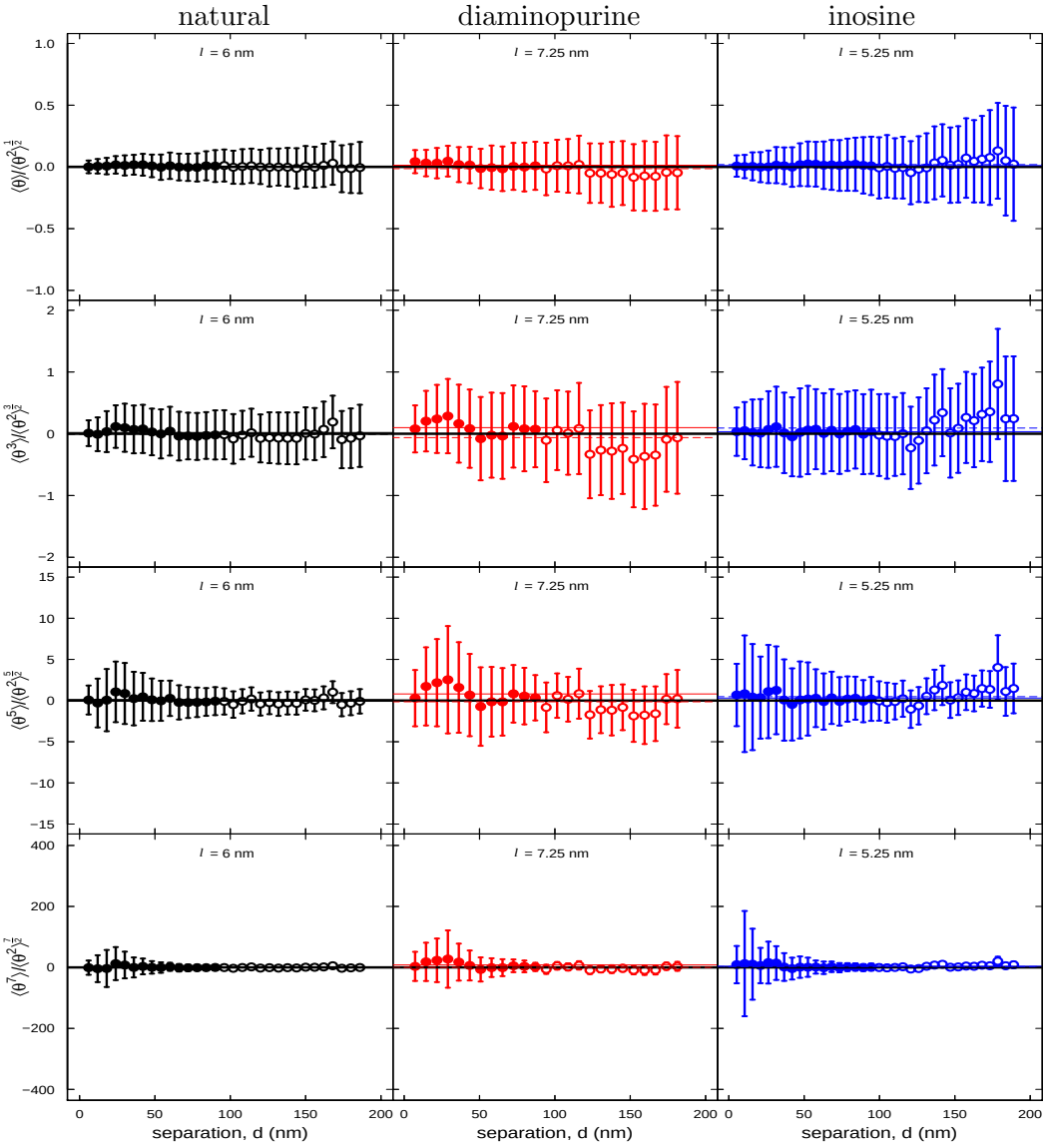


Figure S.9: First four odd moments of θ. The mean of the filled symbols is indicated with a solid line, while the mean of all symbols (filled and open) is indicated with a dashed line.

The even moments give non-zero predictions and each is a function of the three dimensional persistence length (i.e. $P \equiv P_{3D}$).[3] Consider

$$\langle \theta^2 \rangle = \frac{2d}{P} \tag{S.13}$$

This equation has an identical mathematical form in 2D

$$\langle \theta^2 \rangle_{2D} = \frac{2d}{P_{2D}} \tag{S.14}$$

Due to the loss of one degree of freedom moving 3D $\rightarrow$ 2D, there exists a scaling relationship between the two persistence lengths, so that the substitution $P_{2D} = 2P$ gives this 2D conformational statistic in terms of the desired 3D persistence length

$$\langle \theta^2 \rangle_{2D} = \frac{d}{P} \tag{S.15}$$

A similar argument applies to the fourth moment

$$\langle \theta^4 \rangle_{2D} = 3\frac{d^2}{P^2} \tag{S.16}$$

and sixth moment

$$\langle \theta^6 \rangle_{2D} = 15\frac{d^3}{P^3} \tag{S.17}$$

Persistence length estimates were obtained from these functions $f(d;P)$ by employing a weighted nonlinear least squares method, i.e. minimizing the cost function $C(P)$ with respect to the free parameter P

$$\chi^2 = \min_P C(P) = \min_P \sum_{d=0}^{l_{\max}} \left(\frac{\mu_d - f(d;P)}{\sigma_d} \right)^2 \tag{S.18}$$

where μ_d and σ_d are experimentally determined mean values and standard deviations, respectively, from sampling and $f(d;P)$ are theoretical predictions of the WLC model. We restrict the fitting to $0 < d < l_{\max}$, with $l_{\max} \approx 2P$ ($\sim$100 nm). Owing to self-avoiding interactions, fitting over the entire range can yield an overestimation of P (7). Consistent with previous reports (3), self-avoiding interactions affect θ at shorter separations than R. The fitted value of P (and its uncertainty) along with the average persistence length value from from 10 distinct estimates, P_{ave}, are shown in Figure S.10 for the first three even moments of θ. Each of the additional 7 estimates will be discussed forthwith.

The ratio of the fourth moment to the square of the second moment, the so-called kurtosis (k), satisfies the following

$$k = \frac{\langle \theta^4 \rangle}{\langle \theta^2 \rangle^2} = 3 \tag{S.19}$$

which along with the ratio

$$\frac{\langle \theta^6 \rangle}{\langle \theta^2 \rangle^3} = 15 \tag{S.20}$$

serve as additional checks of the normality prediction, as shown in Figure S.11.

[3] As a convention, the subscript will be omitted when referring to the 3D case.

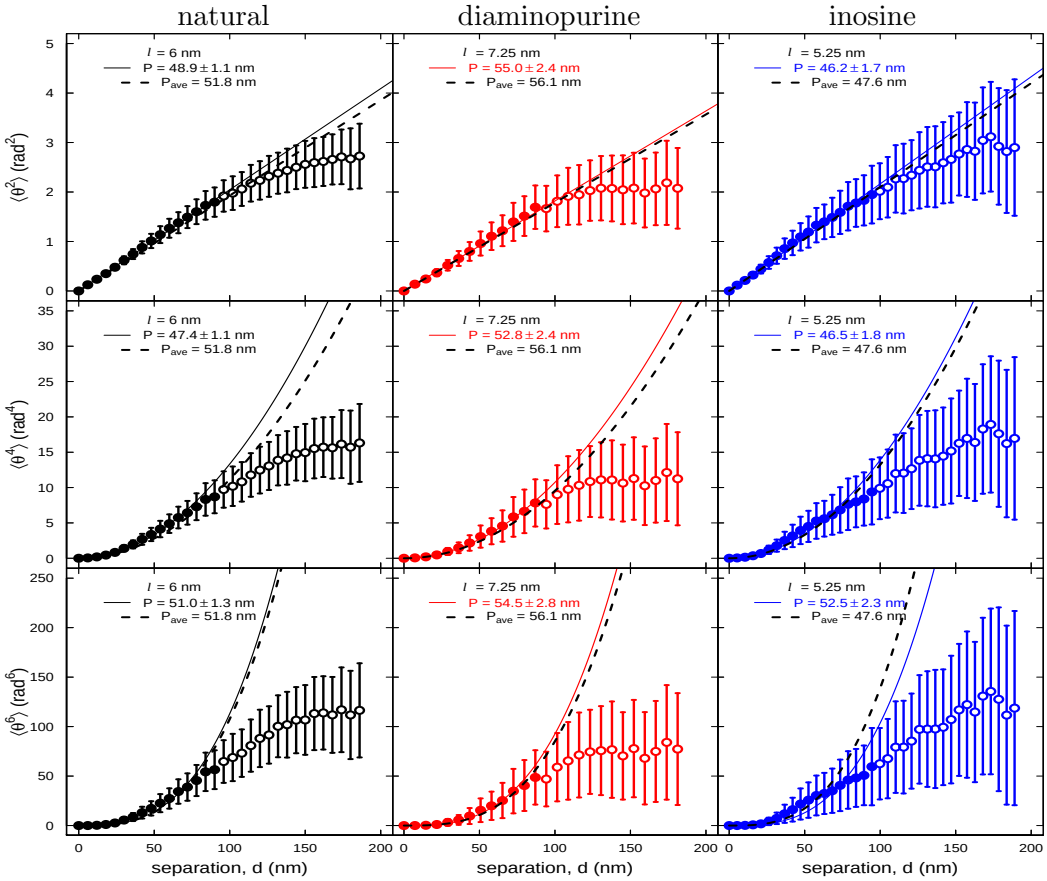


Figure S.10: The second, fourth, and sixth moments of θ yield estimates of P.

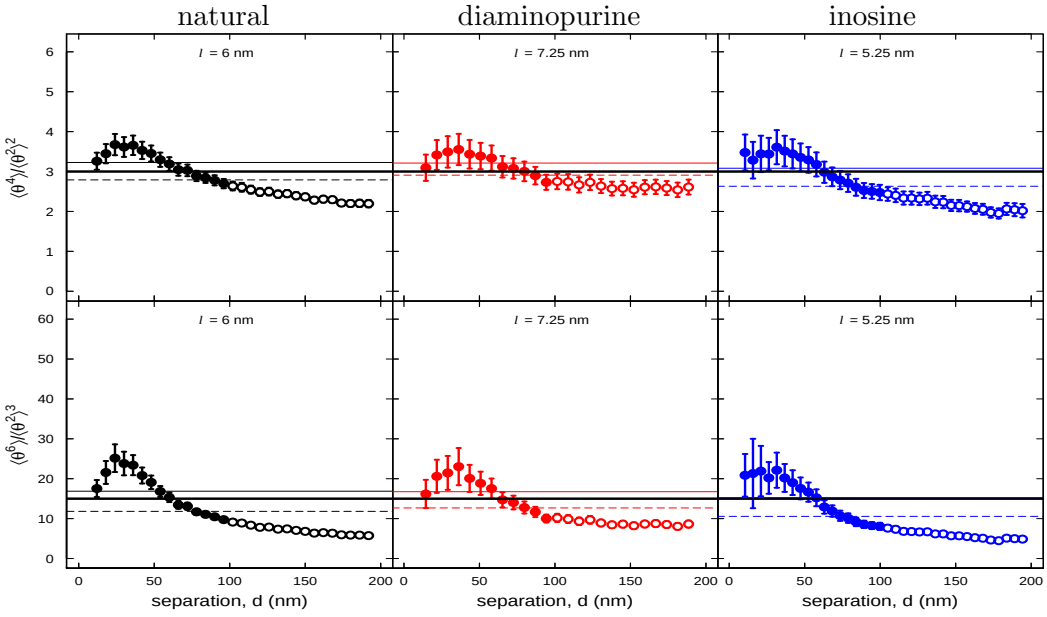


Figure S.11: Normality prediction. See Figure S.9 legend.

Recall the prediction of Equation S.8 (adjusted for dimensionality)

$$\langle \cos\theta \rangle_{2D} = e^{-\frac{d}{(2P)}} \tag{S.21}$$

The fitted value of P is shown in Figure S.12.

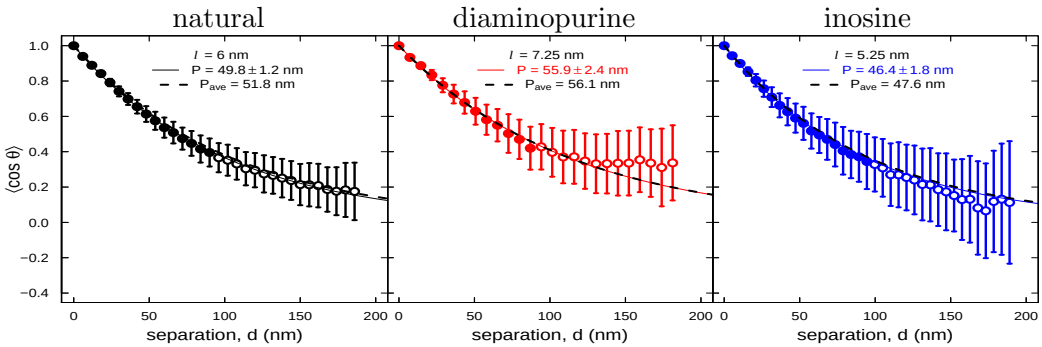


Figure S.12: Estimate of P from $\cos\theta$.

The WLC model also makes predictions about projection of the end-end distance vector $\mathbf{R}$ onto the initial tangent vector $\mathbf{u}_0$

$$\langle \mathbf{R}\cdot\mathbf{u}_0\rangle_{2D} = \int_0^d \langle \mathbf{u}(s)\cdot\mathbf{u}_0\rangle\, ds = \int_0^d e^{-\frac{s}{2P}}\, ds = 2P\left(1-e^{-\frac{d}{2P}}\right) \tag{S.22}$$

and the second moment of R

$$\begin{aligned}
\langle R^2\rangle_{2D} &= \int_0^d\int_0^d \langle \mathbf{u}(s)\cdot\mathbf{u}(s')\rangle\, ds\, ds' \\
&= \int_0^d\int_0^d e^{-\frac{|s-s'|}{2P}}\, ds\, ds' \\
&= \int_0^d 2\int_0^{s'} e^{-\frac{s-s'}{2P}}\, ds\, ds' \\
&= 4dP\left[1-\frac{2P}{d}\left(1-e^{-\frac{d}{2P}}\right)\right]
\end{aligned} \tag{S.23}$$

The same result is reached by observing that $d\langle R^2\rangle_{2D} = 2\langle \mathbf{R}\cdot\mathbf{u}_0\rangle_{2D}\, ds$

$$\langle R^2\rangle_{2D} = \int_0^d 2\langle \mathbf{R}\cdot\mathbf{u}_0\rangle_{2D}\, ds = \int_0^d 4P\left(1-e^{-\frac{s}{2P}}\right) ds = 4dP\left[1-\frac{2P}{d}\left(1-e^{-\frac{d}{2P}}\right)\right] \tag{S.24}$$

Unlike $\langle \mathbf{R}\cdot\mathbf{u}_0\rangle$ and the second moment of R, the fourth moment of R does not possess the same mathematical form in 3D as in 2D, so that this conformational statistic has been shown (7) to provide a reliable means of determining the dimensionality of the deposited molecules

$$\langle R^4\rangle_{2D} = 32d^2P^2 - 240dP^3 + 696P^4 - \frac{320}{3}dP^3e^{-\frac{d}{2P}} - \frac{6272}{9}P^4e^{-\frac{d}{2P}} + \frac{8}{9}P^4e^{-\frac{2d}{P}} \tag{S.25}$$

Additionally, the normalized difference $(\langle R^4\rangle - \langle R^2\rangle^2)/d^4$ does not vanish for large separations. Fits using these equations are shown in Figure S.13. These last two fits provide confirmation that the imaged DNA molecules were immobilized in 2D equilibrium confirmations.

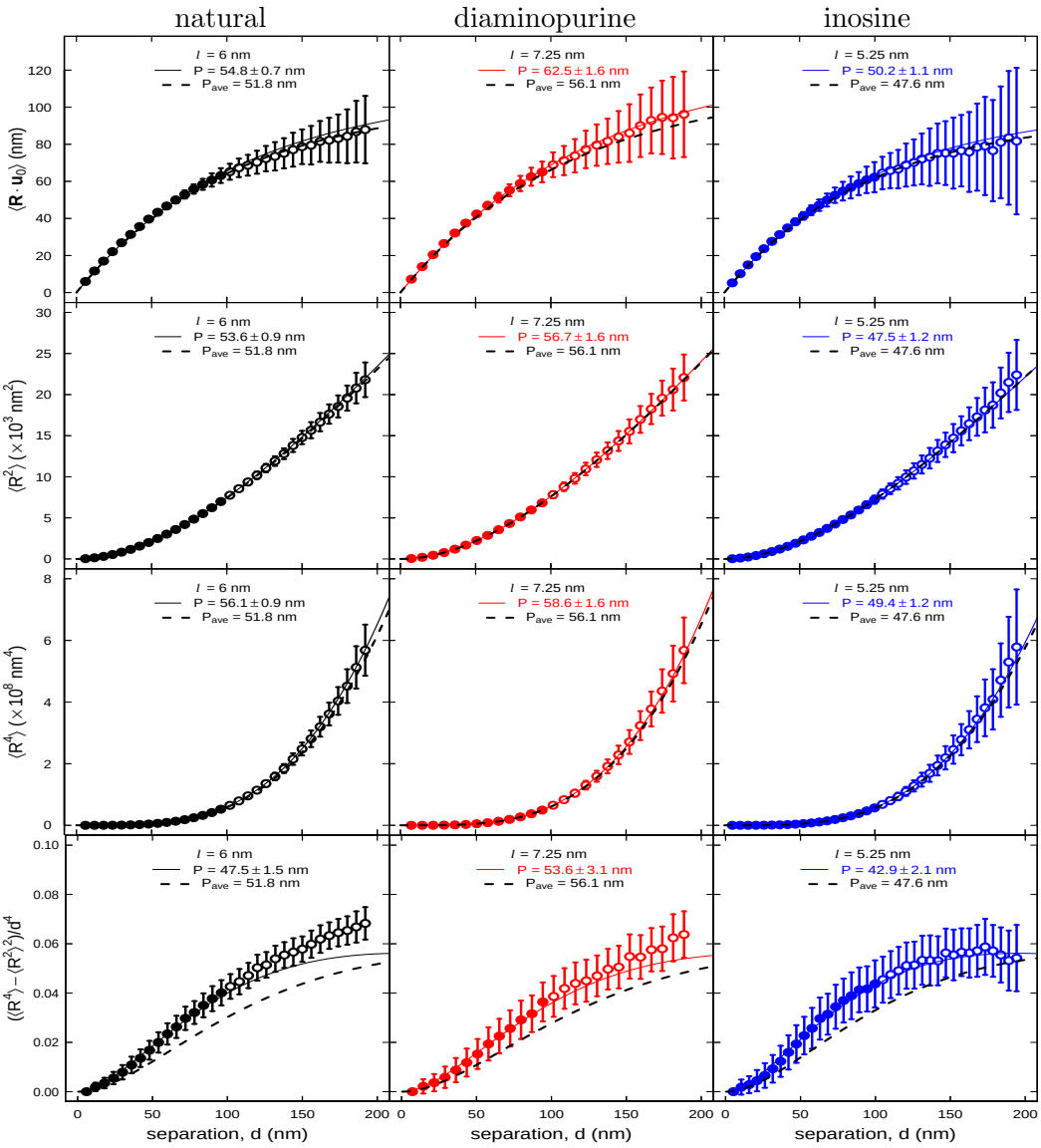


Figure S.13: Estimates of P from the moments of R.

We now revisit the WLC angular distribution (Equation S.9). Binning of the measured angles is necessary to calculate the histogram $-\ln G(\theta; d, P)$ (7). For a bin size of $\Delta\theta$ around a bin center of θ, the probability is evaluated using the error function

$$\mathrm{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^{\pi} e^{-t^2}\, dt \tag{S.26}$$

so that

$$G(\theta; d, P) = \mathrm{erf}\left(\sqrt{\frac{P}{2d}}\left(\theta + \frac{\Delta\theta}{2}\right)\right) - \mathrm{erf}\left(\sqrt{\frac{P}{2d}}\left(\theta - \frac{\Delta\theta}{2}\right)\right) \tag{S.27}$$

The negative logarithm of this equation was used for the fitting shown in Figure S.14.

Finally, Figure S.8 shows that the angle between adjacent segments is denoted as χ (a special case of θ when $d = l$). Equation S.21 is still applicable

$$\langle \cos\chi \rangle_{2\mathrm{D}} = e^{-\frac{l}{2P}} \tag{S.28}$$

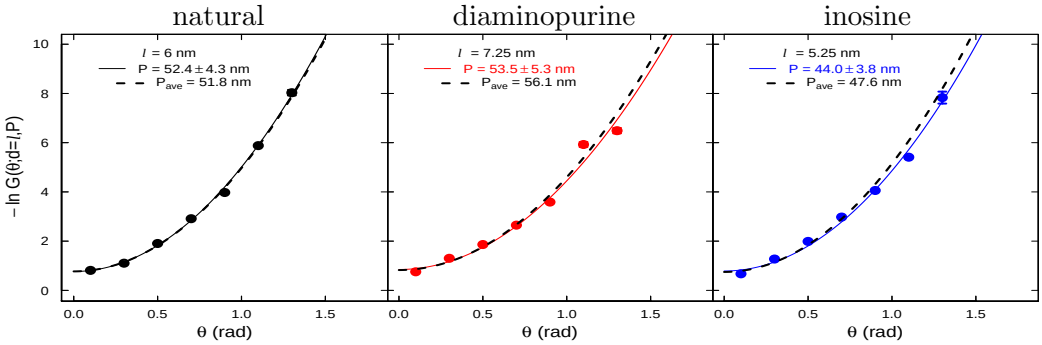


Figure S.14: Estimate of P from $-\ln G$.

and can be rearranged to yield

$$P = \frac{-l}{2\ln(\langle\cos\chi\rangle_{2D})} \tag{S.29}$$

Abels *et al.* (11) looked at multiple segment lengths l and showed that the value of $\langle\cos\chi\rangle$ peaks at a particular value of l and decreases at lower and higher l. The authors reasoned that as l decreased below the pixel resolution for the AFM setup, discretization of χ angles led to an underestimation of $\langle\cos\chi\rangle$. Additionally, they proposed an undercount of the number of large and small χ angles as l became large enough to average over them, similarly leading to underestimation of $\langle\cos\chi\rangle$. Therefore, the quantity $\langle\cos\chi\rangle_{\max}$ provides a reliable means for determining the optimal segment length for analysis and provides an additional estimate of P at this optimal segment length. This is illustrated in Figure S.15 for step sizes l ranging from 3 nm to 17 nm and Figure S.16 for step

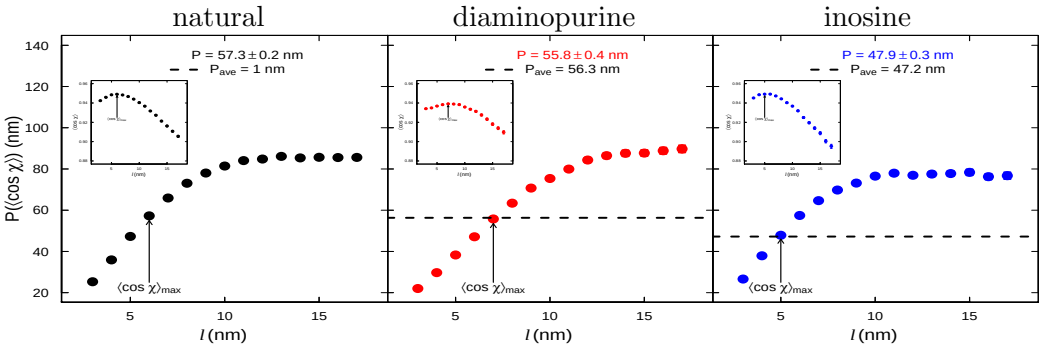


Figure S.15: Estimate of $P(\langle\cos\chi\rangle_{\max})$ when l ranges from 3 nm to 17 nm.

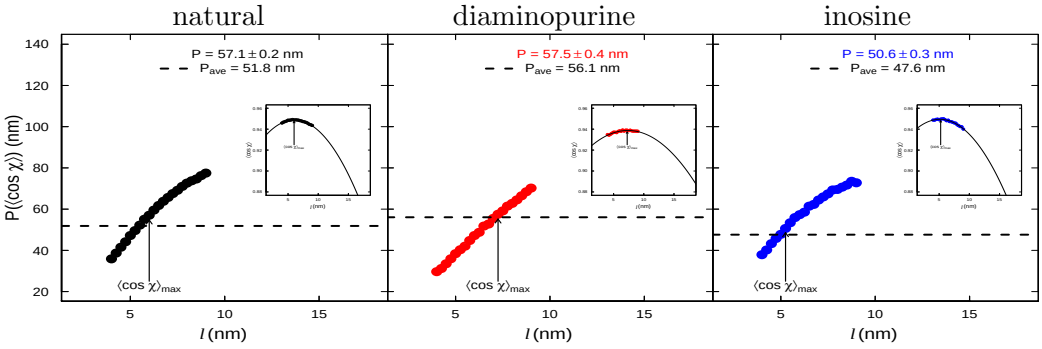


Figure S.16: Estimate of $P(\langle\cos\chi\rangle_{\max})$ when l ranges from 4 nm to 9 nm.

Table S.4: Summary of P estimates for step sizes l ranging from 3 nm to 17 nm. The optimal l is indicated in gray.

l (nm)	$P \pm$ SEM (nm)									$P \pm$ SD (nm)	
	$\langle\cos\theta\rangle$	$\langle\theta^2\rangle$	$\langle\theta^4\rangle$	$\langle\theta^6\rangle$	$\langle\mathbf{R}\cdot\mathbf{u}_0\rangle$	$\langle R^2\rangle$	$\langle R^4\rangle$	$\langle R^4\rangle - \langle R^2\rangle^2$	$-\ln G(\theta)$	$\langle\cos\chi\rangle$	average
natural											
3	40.0 ± 0.9	40.9 ± 0.8	44.9 ± 1.0	51.0 ± 1.3	46.7 ± 0.8	47.4 ± 0.9	48.7 ± 0.9	49.4 ± 1.5	25.0 ± 2.2	25.3 ± 0.0	41.9 ± 9.5
4	43.7 ± 1.0	43.8 ± 0.9	46.0 ± 1.1	51.2 ± 1.4	49.9 ± 0.8	49.9 ± 0.9	51.6 ± 0.9	48.2 ± 1.5	35.1 ± 2.9	35.9 ± 0.1	45.5 ± 6.0
5	47.1 ± 1.1	46.7 ± 1.0	47.2 ± 1.1	52.5 ± 1.4	52.3 ± 0.7	52.1 ± 0.9	54.1 ± 0.9	48.2 ± 1.4	41.4 ± 2.5	47.3 ± 0.1	48.9 ± 3.8
6	50.2 ± 1.2	49.3 ± 1.1	47.9 ± 1.2	51.9 ± 1.4	55.1 ± 0.7	53.8 ± 0.9	56.3 ± 0.9	48.0 ± 1.5	50.6 ± 4.3	57.3 ± 0.2	52.0 ± 3.4
7	52.7 ± 1.3	51.7 ± 1.2	49.9 ± 1.2	52.6 ± 1.4	56.8 ± 0.7	55.3 ± 0.9	58.0 ± 0.9	48.4 ± 1.5	53.0 ± 3.5	65.9 ± 0.3	54.4 ± 5.0
8	54.8 ± 1.4	53.6 ± 1.3	50.7 ± 1.3	52.3 ± 1.4	58.7 ± 0.8	56.6 ± 0.9	59.4 ± 0.9	48.0 ± 1.5	59.0 ± 4.0	73.1 ± 0.4	56.6 ± 6.9
9	56.6 ± 1.4	55.2 ± 1.4	51.7 ± 1.4	53.4 ± 1.4	60.2 ± 0.8	57.7 ± 0.9	60.7 ± 0.9	48.0 ± 1.4	58.2 ± 4.6	78.0 ± 0.5	58.0 ± 8.1
10	57.7 ± 1.5	56.2 ± 1.5	51.7 ± 1.4	53.2 ± 1.4	61.2 ± 0.8	58.2 ± 0.9	61.5 ± 1.0	47.5 ± 1.4	61.2 ± 5.1	81.4 ± 0.5	59.0 ± 9.1
11	58.4 ± 1.6	56.0 ± 1.6	49.9 ± 1.4	53.0 ± 1.5	63.6 ± 0.9	59.5 ± 1.0	63.1 ± 1.0	47.4 ± 1.4	64.5 ± 5.6	84.1 ± 0.6	60.0 ± 10.3
12	60.8 ± 1.7	59.2 ± 1.7	53.7 ± 1.6	53.4 ± 1.6	65.8 ± 0.9	60.4 ± 1.0	64.1 ± 1.1	48.0 ± 1.5	63.4 ± 6.2	84.8 ± 0.7	61.4 ± 9.9
13	60.0 ± 1.7	58.1 ± 1.7	52.6 ± 1.5	54.0 ± 1.5	65.2 ± 0.9	60.1 ± 1.0	63.9 ± 1.1	47.1 ± 1.4	64.7 ± 6.8	86.1 ± 0.7	61.2 ± 10.5
14	60.8 ± 1.8	58.0 ± 1.7	50.8 ± 1.5	53.2 ± 1.6	68.9 ± 1.0	62.0 ± 1.1	65.8 ± 1.1	47.6 ± 1.4	63.1 ± 7.4	85.4 ± 0.8	61.6 ± 10.7
15	60.5 ± 1.8	58.2 ± 1.7	52.1 ± 1.6	54.2 ± 1.6	68.5 ± 1.1	61.6 ± 1.1	65.9 ± 1.2	46.4 ± 1.4	64.1 ± 8.0	85.7 ± 0.9	61.7 ± 10.8
16	62.0 ± 2.0	59.6 ± 1.9	52.0 ± 1.7	52.8 ± 1.7	72.9 ± 1.2	63.8 ± 1.3	68.3 ± 1.3	47.5 ± 1.5	65.8 ± 8.5	85.6 ± 0.9	63.0 ± 11.2
17	62.8 ± 2.0	60.2 ± 1.9	52.7 ± 1.7	54.1 ± 1.7	73.0 ± 1.2	63.9 ± 1.3	68.3 ± 1.3	47.5 ± 1.5	65.8 ± 9.2	85.6 ± 0.9	63.4 ± 11
diaminopurine											
3	41.1 ± 1.7	42.6 ± 1.6	47.4 ± 2.1	53.4 ± 2.6	49.3 ± 1.8	47.3 ± 1.7	47.8 ± 1.6	55.9 ± 3.1	21.2 ± 2.2	22.0 ± 0.1	42.8 ± 12
4	44.6 ± 1.8	45.0 ± 1.7	47.5 ± 2.1	53.0 ± 2.7	52.4 ± 1.7	50.5 ± 1.7	51.4 ± 1.6	55.2 ± 3.1	27.5 ± 2.4	29.7 ± 0.1	45.7 ± 9.6
5	48.5 ± 2.0	48.4 ± 1.9	50.0 ± 2.2	54.9 ± 2.7	55.7 ± 1.5	52.3 ± 1.5	53.6 ± 1.5	53.6 ± 2.9	39.3 ± 3.6	38.3 ± 0.2	49.5 ± 6.2
6	52.2 ± 2.2	51.6 ± 2.1	50.9 ± 2.2	54.3 ± 2.7	59.2 ± 1.5	55.0 ± 1.6	56.3 ± 1.6	55.0 ± 3.2	48.0 ± 4.3	47.1 ± 0.3	53.0 ± 3.7
7	55.6 ± 2.4	54.7 ± 2.3	52.9 ± 2.4	55.5 ± 2.8	62.8 ± 1.6	57.7 ± 1.7	59.5 ± 1.7	54.9 ± 3.1	53.7 ± 5.1	55.8 ± 0.4	56.3 ± 3.0
8	58.7 ± 2.6	57.5 ± 2.6	54.4 ± 2.6	55.5 ± 2.9	65.1 ± 1.6	58.9 ± 1.7	60.9 ± 1.7	54.5 ± 3.1	52.7 ± 4.1	63.4 ± 0.5	58.2 ± 4.1
9	61.3 ± 2.8	60.2 ± 2.7	57.3 ± 2.7	58.0 ± 2.9	66.8 ± 1.6	60.1 ± 1.6	62.6 ± 1.7	53.6 ± 3.0	59.3 ± 6.6	70.7 ± 0.7	61.0 ± 4.9
10	63.8 ± 3.0	62.4 ± 3.0	57.7 ± 3.0	57.9 ± 3.2	69.5 ± 1.7	62.3 ± 1.8	64.8 ± 1.8	55.3 ± 3.1	61.2 ± 5.1	75.4 ± 0.8	63.0 ± 6.0
11	65.8 ± 3.1	64.5 ± 3.1	60.2 ± 3.3	59.4 ± 3.4	71.3 ± 1.7	63.0 ± 1.8	65.7 ± 1.8	55.2 ± 3.1	67.6 ± 8.1	80.0 ± 1.0	65.3 ± 6.9
12	66.9 ± 3.4	65.5 ± 3.4	59.4 ± 3.4	57.1 ± 3.4	72.9 ± 1.8	64.4 ± 1.9	67.3 ± 1.9	55.0 ± 3.2	71.3 ± 8.9	84.3 ± 1.2	66.4 ± 8.6
13	67.2 ± 3.4	65.7 ± 3.4	60.1 ± 3.4	59.3 ± 3.6	74.4 ± 1.9	65.5 ± 2.0	68.8 ± 2.1	54.9 ± 3.1	71.5 ± 9.8	86.5 ± 1.3	67.4 ± 8.9
14	69.4 ± 3.6	68.1 ± 3.6	64.7 ± 3.6	65.4 ± 3.8	77.5 ± 2.0	67.2 ± 2.0	70.5 ± 2.1	55.8 ± 3.1	73.7 ± 10.6	87.6 ± 1.4	70.0 ± 8.5
15	67.7 ± 3.7	66.0 ± 3.6	60.3 ± 3.4	59.7 ± 3.5	76.8 ± 2.1	66.8 ± 2.2	70.2 ± 2.2	54.8 ± 3.1	68.6 ± 7.9	87.7 ± 1.5	67.9 ± 9.3
16	68.5 ± 3.8	66.5 ± 3.8	59.1 ± 3.6	58.3 ± 3.8	82.7 ± 2.5	69.7 ± 2.5	73.5 ± 2.5	55.0 ± 3.3	69.4 ± 8.5	88.8 ± 1.5	69.2 ± 10.6
17	69.7 ± 4.0	67.0 ± 4.0	58.1 ± 3.7	59.2 ± 3.7	81.3 ± 2.5	69.5 ± 2.5	73.8 ± 2.5	53.4 ± 3.0	69.9 ± 9.1	89.7 ± 1.7	69.1 ± 10.9
inosine											
3	39.5 ± 1.5	40.3 ± 1.4	43.9 ± 1.6	51.1 ± 2.2	46.7 ± 1.4	44.4 ± 1.3	45.6 ± 1.3	45.3 ± 2.3	26.1 ± 2.2	26.6 ± 0.1	40.9 ± 8.3
4	43.6 ± 1.7	43.7 ± 1.6	45.4 ± 1.8	51.3 ± 2.2	49.1 ± 1.3	46.3 ± 1.3	48.0 ± 1.3	43.7 ± 2.3	34.1 ± 2.9	37.9 ± 0.2	44.3 ± 5.1
5	46.7 ± 1.8	46.5 ± 1.7	47.2 ± 1.8	52.9 ± 2.3	50.6 ± 1.1	47.7 ± 1.2	49.7 ± 1.2	43.5 ± 2.2	39.6 ± 3.6	47.9 ± 0.3	47.2 ± 3.7
6	49.2 ± 2.0	48.6 ± 1.9	47.2 ± 1.9	51.4 ± 2.3	52.6 ± 1.1	49.4 ± 1.2	51.5 ± 1.3	44.1 ± 2.3	50.1 ± 4.3	57.5 ± 0.4	50.2 ± 3.5
7	51.7 ± 2.1	51.0 ± 2.0	49.5 ± 2.0	52.7 ± 2.3	54.2 ± 1.1	50.2 ± 1.2	52.6 ± 1.3	43.8 ± 2.2	57.4 ± 5.1	64.7 ± 0.5	52.8 ± 5.4
8	53.4 ± 2.3	52.6 ± 2.2	50.0 ± 2.2	52.7 ± 2.5	56.7 ± 1.2	52.0 ± 1.3	54.6 ± 1.4	44.0 ± 2.3	60.4 ± 5.8	69.8 ± 0.6	54.6 ± 6.8
9	54.0 ± 2.3	53.3 ± 2.3	51.3 ± 2.2	54.0 ± 2.4	57.7 ± 1.3	52.9 ± 1.4	55.4 ± 1.4	44.5 ± 2.3	62.1 ± 6.6	73.2 ± 0.7	55.8 ± 7.6
10	55.5 ± 2.5	54.7 ± 2.4	53.3 ± 2.3	55.8 ± 2.5	58.7 ± 1.3	53.8 ± 1.4	56.6 ± 1.5	44.8 ± 2.3	65.9 ± 7.4	76.5 ± 0.8	57.6 ± 8.4
11	56.4 ± 2.6	55.5 ± 2.5	53.5 ± 2.3	55.5 ± 2.5	60.3 ± 1.3	54.4 ± 1.4	57.3 ± 1.5	44.9 ± 2.3	63.9 ± 8.2	78.0 ± 1.0	58.0 ± 8.6
12	56.4 ± 2.7	54.6 ± 2.6	49.5 ± 2.4	52.9 ± 2.7	62.1 ± 1.5	54.7 ± 1.5	57.7 ± 1.6	43.9 ± 2.3	62.5 ± 6.2	77.0 ± 1.0	57.1 ± 8.9
13	56.0 ± 2.8	54.1 ± 2.7	50.2 ± 2.5	55.1 ± 2.8	62.7 ± 1.6	55.3 ± 1.7	58.6 ± 1.7	42.6 ± 2.2	63.0 ± 6.8	77.5 ± 1.1	57.5 ± 9.2
14	58.2 ± 2.8	57.0 ± 2.8	53.7 ± 2.6	55.0 ± 2.7	65.6 ± 1.6	56.8 ± 1.6	60.1 ± 1.7	45.1 ± 2.4	64.1 ± 7.4	77.8 ± 1.2	59.3 ± 8.6
15	57.3 ± 2.9	55.6 ± 2.8	51.4 ± 2.6	55.5 ± 2.8	64.8 ± 1.7	56.3 ± 1.7	59.8 ± 1.8	43.5 ± 2.2	61.8 ± 8.0	78.4 ± 1.4	58.4 ± 9.1
16	57.8 ± 3.1	55.3 ± 3.0	50.4 ± 2.7	54.1 ± 3.0	67.6 ± 1.9	57.2 ± 1.9	60.9 ± 1.9	43.0 ± 2.3	63.1 ± 8.6	76.3 ± 1.4	58.6 ± 9.2
17	56.9 ± 3.1	54.6 ± 2.9	50.6 ± 2.7	55.6 ± 3.0	67.2 ± 1.9	57.5 ± 1.9	61.2 ± 2.0	43.0 ± 2.3	60.7 ± 9.3	76.7 ± 1.6	58.4 ± 9.1

sizes l ranging from 4 nm to 9 nm.

The complete analysis for step sizes l ranging from 3 nm to 17 nm is summarized in Table S.4. Several trends can be observed from these data. First, all but one of the ten WLC predictions strongly depends on choice of step size l. Generally, the estimates of P are strictly increasing functions of increasing l. In some cases, the dependence on l is so dramatic that P increases 400% from l=3 nm to l=17 nm. In contrast, the estimate of P from the $\langle R^4 \rangle - \langle R^2 \rangle^2$ prediction is a decreasing function of increasing l. However, the maximal observed decrease in P is less than 5%. Overall, this estimator appears to be the most reliable: it is the most insensitive to choice of step size and it provides a means of assessing the dimensionality of the molecules being imaged. If choosing only one of the ten WLC predictions to estimate P, the $\langle R^4 \rangle - \langle R^2 \rangle^2$ prediction would certainly be the first choice.

This initial analysis revealed the importance of identifying the proper step size to obtain the best possible estimates of P. In particular, we determined that the optimal step size l_{optimal} occurred between 4 nm to 9 nm. We worried that discretization of l might prevent us from identifying the true (non-integral) value of l_{optimal}, with possible negative ramifications. Therefore, we next examined step sizes from 4 nm to 9 nm at every quarter nanometer. This analysis is summarized in Table S.5. Importantly, the estimates of P from the fractional optimal step sizes were consistently within one half nanometer of the previous estimates. This is well within the uncertainty from fitting. These results indicated that discretization of l (to integral values) had little effect on the estimates of P and rendered a more exact determination of l_{optimal} unnecessary.

In order to better determine which physical and thermodynamic feature(s) of base-substituted DNA polymers explain their mechanical properties, we applied AFM experiments to seven thymine variants with functional groups that occupy the major groove (main manuscript Figure 7) that have been previously characterized (2). The ten estimates of P at l_{optimal} for the thymine variants (numbered **1**-**7**) are given in Table S.6. Finally, all of the compiled AFM data, including the average and standard deviation of the ten estimates, are summarized in Table 3 of the main manuscript.

Table S.5: Summary of P estimates for step sizes l ranging from 4 nm to 9 nm. The optimal l is indicated in gray.

l (nm)	$P \pm$ SEM (nm)									$P \pm$ SD (nm)	
	$\langle\cos\theta\rangle$	$\langle\theta^2\rangle$	$\langle\theta^4\rangle$	$\langle\theta^6\rangle$	$\langle\mathbf{R}\cdot\mathbf{u}_0\rangle$	$\langle R^2\rangle$	$\langle R^4\rangle$	$\langle R^4\rangle - \langle R^2\rangle^2$	$-\ln G(\theta)$	$\langle\cos\chi\rangle$	average
natural											
4	43.4 ± 1.0	43.4 ± 0.9	44.3 ± 1.0	49.6 ± 1.2	49.4 ± 0.7	49.6 ± 0.8	51.3 ± 0.9	48.0 ± 1.5	35.1 ± 2.9	35.9 ± 0.1	45.0 ± 5.8
4.25	44.2 ± 1.0	44.1 ± 0.9	45.2 ± 1.0	50.7 ± 1.2	50.2 ± 0.7	50.5 ± 0.8	52.4 ± 0.9	48.2 ± 1.4	34.1 ± 2.1	38.7 ± 0.1	45.8 ± 5.8
4.5	45.1 ± 1.0	44.9 ± 1.0	45.5 ± 1.0	51.0 ± 1.2	50.6 ± 0.7	50.7 ± 0.8	52.6 ± 0.8	47.5 ± 1.4	38.5 ± 3.2	41.5 ± 0.1	46.8 ± 4.6
4.75	45.8 ± 1.0	45.4 ± 0.9	45.8 ± 1.0	51.1 ± 1.2	51.3 ± 0.7	51.2 ± 0.8	53.3 ± 0.8	47.4 ± 1.4	41.0 ± 3.4	44.2 ± 0.1	47.6 ± 3.9
5	46.7 ± 1.1	46.2 ± 1.0	46.2 ± 1.0	51.4 ± 1.3	51.9 ± 0.7	51.7 ± 0.8	53.9 ± 0.9	47.0 ± 1.4	43.6 ± 3.6	47.2 ± 0.1	48.6 ± 3.4
5.25	47.6 ± 1.1	47.1 ± 1.0	46.8 ± 1.1	51.9 ± 1.3	52.4 ± 0.7	52.3 ± 0.8	54.5 ± 0.9	47.4 ± 1.4	41.2 ± 2.6	49.6 ± 0.2	49.1 ± 3.9
5.5	48.4 ± 1.1	47.7 ± 1.0	47.1 ± 1.1	51.6 ± 1.3	53.2 ± 0.7	52.7 ± 0.8	55.1 ± 0.9	47.2 ± 1.4	42.6 ± 2.8	52.2 ± 0.2	49.8 ± 3.8
5.75	49.4 ± 1.1	48.6 ± 1.1	47.4 ± 1.1	51.5 ± 1.3	54.1 ± 0.7	53.2 ± 0.8	55.7 ± 0.9	47.3 ± 1.4	45.9 ± 2.9	55.1 ± 0.2	50.8 ± 3.6
6	49.8 ± 1.2	48.9 ± 1.1	47.4 ± 1.1	51.0 ± 1.3	54.8 ± 0.7	53.6 ± 0.9	56.1 ± 0.9	47.5 ± 1.5	52.4 ± 4.3	57.1 ± 0.2	51.8 ± 3.5
6.25	50.4 ± 1.2	49.5 ± 1.1	48.2 ± 1.1	52.1 ± 1.3	54.7 ± 0.7	53.8 ± 0.8	56.3 ± 0.9	47.1 ± 1.4	50.0 ± 3.1	59.5 ± 0.2	52.2 ± 3.9
6.5	50.9 ± 1.2	49.9 ± 1.2	48.2 ± 1.2	51.8 ± 1.4	55.6 ± 0.7	54.1 ± 0.9	56.8 ± 0.9	47.0 ± 1.4	53.8 ± 4.7	61.8 ± 0.3	53.0 ± 4.4
6.75	51.7 ± 1.2	50.7 ± 1.2	48.5 ± 1.2	51.0 ± 1.3	56.3 ± 0.7	54.6 ± 0.9	57.2 ± 0.9	47.4 ± 1.5	57.0 ± 4.9	63.8 ± 0.3	53.8 ± 4.9
7	52.4 ± 1.3	51.4 ± 1.2	49.3 ± 1.2	52.3 ± 1.4	56.7 ± 0.7	55.0 ± 0.9	57.7 ± 0.9	47.3 ± 1.4	52.5 ± 3.5	65.7 ± 0.3	54.0 ± 5.2
7.25	52.8 ± 1.3	51.6 ± 1.3	48.9 ± 1.2	51.2 ± 1.4	57.4 ± 0.8	55.3 ± 0.9	58.1 ± 0.9	47.2 ± 1.4	52.9 ± 3.7	67.6 ± 0.3	54.3 ± 5.8
7.5	53.2 ± 1.3	52.0 ± 1.3	49.4 ± 1.3	52.2 ± 1.4	57.6 ± 0.8	55.6 ± 0.9	58.5 ± 0.9	46.9 ± 1.4	55.5 ± 3.8	69.2 ± 0.3	55.0 ± 6.1
7.75	54.0 ± 1.3	52.7 ± 1.3	50.0 ± 1.3	52.8 ± 1.4	57.4 ± 0.7	55.6 ± 0.9	58.6 ± 0.9	46.9 ± 1.4	55.2 ± 3.9	71.0 ± 0.4	55.4 ± 6.5
8	54.5 ± 1.4	53.2 ± 1.3	50.3 ± 1.3	52.1 ± 1.4	58.6 ± 0.8	56.2 ± 0.9	59.3 ± 0.9	47.0 ± 1.4	58.0 ± 4.0	72.6 ± 0.4	56.2 ± 7.0
8.25	54.4 ± 1.3	53.1 ± 1.3	50.6 ± 1.3	52.8 ± 1.4	58.3 ± 0.8	56.3 ± 0.9	59.3 ± 0.9	46.8 ± 1.4	56.4 ± 4.2	73.8 ± 0.4	56.2 ± 7.2
8.5	55.1 ± 1.4	53.8 ± 1.4	50.4 ± 1.3	51.5 ± 1.4	59.4 ± 0.8	56.5 ± 0.9	59.6 ± 0.9	47.0 ± 1.5	57.2 ± 4.3	74.6 ± 0.4	56.5 ± 7.5
8.75	55.6 ± 1.4	54.2 ± 1.4	50.4 ± 1.3	51.8 ± 1.4	59.9 ± 0.8	57.0 ± 0.9	60.2 ± 1.0	46.8 ± 1.4	57.7 ± 4.5	76.3 ± 0.4	57.0 ± 8.0
9	55.8 ± 1.4	54.4 ± 1.4	51.1 ± 1.3	52.9 ± 1.4	59.9 ± 0.8	57.1 ± 0.9	60.3 ± 0.9	46.8 ± 1.4	57.4 ± 4.6	77.5 ± 0.4	57.3 ± 8.2
diaminopurine											
4	45.1 ± 1.9	45.7 ± 1.8	48.9 ± 2.2	54.1 ± 2.8	53.2 ± 1.7	50.9 ± 1.7	51.6 ± 1.6	57.1 ± 3.3	29.3 ± 2.0	29.7 ± 0.1	46.6 ± 9.7
4.25	45.4 ± 1.8	45.9 ± 1.7	48.8 ± 2.1	53.6 ± 2.6	53.2 ± 1.6	50.8 ± 1.6	51.6 ± 1.6	55.5 ± 3.1	30.4 ± 2.1	31.2 ± 0.1	46.6 ± 8.9
4.5	46.4 ± 1.9	46.6 ± 1.8	48.9 ± 2.1	54.0 ± 2.6	53.7 ± 1.6	51.2 ± 1.6	52.1 ± 1.6	55.0 ± 3.0	33.2 ± 3.3	33.5 ± 0.1	47.5 ± 8.0
4.75	47.8 ± 1.9	47.9 ± 1.8	50.0 ± 2.2	55.0 ± 2.8	55.2 ± 1.6	51.9 ± 1.6	53.1 ± 1.5	53.9 ± 3.0	32.9 ± 2.4	35.9 ± 0.2	48.4 ± 7.8
5	49.1 ± 2.0	49.0 ± 1.9	49.9 ± 2.1	54.6 ± 2.5	56.9 ± 1.6	53.1 ± 1.6	54.3 ± 1.6	55.4 ± 3.1	38.8 ± 3.6	38.4 ± 0.2	49.9 ± 6.6
5.25	49.6 ± 2.1	49.6 ± 1.9	50.8 ± 2.2	55.6 ± 2.7	56.9 ± 1.6	53.7 ± 1.6	54.9 ± 1.6	56.2 ± 3.1	41.7 ± 3.8	40.3 ± 0.2	50.9 ± 5.9
5.5	50.3 ± 2.1	50.1 ± 2.0	50.8 ± 2.3	54.8 ± 2.8	57.7 ± 1.6	53.7 ± 1.6	55.0 ± 1.6	54.6 ± 3.0	41.6 ± 4.0	42.1 ± 0.2	51.1 ± 5.4
5.75	50.9 ± 2.1	50.4 ± 2.0	50.7 ± 2.2	54.5 ± 2.7	58.0 ± 1.5	53.8 ± 1.5	55.5 ± 1.6	52.9 ± 2.9	46.4 ± 4.1	44.8 ± 0.3	51.8 ± 4.0
6	52.5 ± 2.2	51.9 ± 2.1	51.0 ± 2.2	54.3 ± 2.7	59.8 ± 1.6	55.4 ± 1.6	56.9 ± 1.6	55.1 ± 3.1	47.3 ± 6.7	47.2 ± 0.3	53.1 ± 4.0
6.25	53.0 ± 2.2	52.5 ± 2.1	51.6 ± 2.2	55.1 ± 2.6	60.0 ± 1.5	55.1 ± 1.5	56.7 ± 1.6	54.4 ± 3.0	47.1 ± 4.5	48.8 ± 0.3	53.4 ± 3.8
6.5	54.1 ± 2.3	53.4 ± 2.2	52.2 ± 2.3	54.9 ± 2.7	61.1 ± 1.6	56.5 ± 1.6	58.2 ± 1.6	54.9 ± 3.1	50.8 ± 4.7	51.8 ± 0.3	54.8 ± 3.1
6.75	55.0 ± 2.4	54.2 ± 2.3	52.6 ± 2.4	54.7 ± 2.8	61.6 ± 1.6	56.4 ± 1.6	58.0 ± 1.6	55.1 ± 3.2	47.9 ± 3.4	52.8 ± 0.4	54.8 ± 3.6
7	55.2 ± 2.4	54.3 ± 2.3	52.6 ± 2.3	55.2 ± 2.7	62.1 ± 1.6	56.9 ± 1.6	58.6 ± 1.6	54.2 ± 3.1	52.7 ± 5.1	55.5 ± 0.4	55.7 ± 2.9
7.25	55.9 ± 2.4	55.0 ± 2.4	52.8 ± 2.4	54.5 ± 2.8	62.5 ± 1.6	56.7 ± 1.6	58.6 ± 1.6	53.6 ± 3.1	53.5 ± 5.3	57.5 ± 0.4	56.1 ± 2.9
7.5	56.1 ± 2.5	55.1 ± 2.4	52.9 ± 2.4	55.1 ± 2.7	62.7 ± 1.6	57.3 ± 1.6	59.2 ± 1.6	53.9 ± 3.0	53.0 ± 5.5	59.2 ± 0.4	56.4 ± 3.2
7.75	57.7 ± 2.6	56.7 ± 2.5	54.4 ± 2.6	57.0 ± 3.0	64.2 ± 1.6	58.4 ± 1.6	60.4 ± 1.7	54.9 ± 3.1	57.8 ± 5.6	61.5 ± 0.5	58.3 ± 3.0
8	58.6 ± 2.6	57.4 ± 2.6	54.3 ± 2.6	55.2 ± 2.8	64.4 ± 1.6	57.7 ± 1.6	59.8 ± 1.6	53.4 ± 3.0	58.5 ± 5.8	62.8 ± 0.5	58.2 ± 3.5
8.25	59.1 ± 2.6	58.0 ± 2.5	55.4 ± 2.6	56.8 ± 2.8	65.1 ± 1.6	58.5 ± 1.6	60.9 ± 1.7	53.4 ± 2.9	52.5 ± 4.2	64.5 ± 0.6	58.4 ± 4.2
8.5	59.7 ± 2.7	58.6 ± 2.6	55.5 ± 2.7	55.9 ± 2.9	66.5 ± 1.6	59.6 ± 1.7	61.6 ± 1.7	55.2 ± 3.2	58.9 ± 6.2	66.5 ± 0.6	59.8 ± 4.1
8.75	59.7 ± 2.7	58.6 ± 2.6	55.5 ± 2.7	56.1 ± 2.9	66.4 ± 1.6	59.5 ± 1.7	61.7 ± 1.7	54.1 ± 3.0	58.8 ± 6.4	68.4 ± 0.6	59.9 ± 4.6
9	60.8 ± 2.7	59.6 ± 2.7	56.7 ± 2.7	57.5 ± 2.9	66.8 ± 1.6	59.8 ± 1.6	62.3 ± 1.7	53.4 ± 2.9	59.9 ± 6.6	70.2 ± 0.7	60.7 ± 4.9
inosine											
4	43.0 ± 1.6	43.1 ± 1.5	44.8 ± 1.7	50.9 ± 2.2	48.6 ± 1.2	45.8 ± 1.2	47.3 ± 1.2	44.1 ± 2.2	34.1 ± 2.9	37.9 ± 0.2	44.0 ± 5.0
4.25	43.7 ± 1.7	43.7 ± 1.6	45.3 ± 1.8	52.1 ± 2.3	48.7 ± 1.2	46.1 ± 1.2	47.9 ± 1.2	43.2 ± 2.2	35.1 ± 3.1	40.2 ± 0.2	44.6 ± 4.7
4.5	44.7 ± 1.7	44.5 ± 1.6	45.6 ± 1.8	52.2 ± 2.3	49.3 ± 1.2	46.5 ± 1.2	48.3 ± 1.2	42.8 ± 2.2	33.2 ± 2.3	43.3 ± 0.2	45.0 ± 5.1
4.75	44.9 ± 1.7	44.7 ± 1.6	46.3 ± 1.9	52.7 ± 2.4	49.5 ± 1.2	46.9 ± 1.2	48.7 ± 1.2	42.7 ± 2.2	39.2 ± 3.4	45.8 ± 0.2	46.1 ± 3.7
5	46.0 ± 1.8	45.7 ± 1.7	46.4 ± 1.9	53.0 ± 2.4	49.7 ± 1.1	46.8 ± 1.2	48.7 ± 1.2	42.2 ± 2.1	38.9 ± 3.6	47.8 ± 0.3	46.5 ± 3.9
5.25	46.4 ± 1.8	46.2 ± 1.7	46.5 ± 1.8	52.5 ± 2.3	50.2 ± 1.1	47.5 ± 1.2	49.4 ± 1.2	42.9 ± 2.1	44.0 ± 3.8	50.6 ± 0.3	47.6 ± 3.0
5.5	48.0 ± 1.8	47.5 ± 1.8	47.0 ± 1.8	52.2 ± 2.2	51.2 ± 1.1	47.9 ± 1.1	49.8 ± 1.2	43.6 ± 2.2	44.9 ± 4.0	53.4 ± 0.3	48.5 ± 3.1
5.75	48.4 ± 1.9	47.9 ± 1.8	47.2 ± 1.8	52.2 ± 2.2	51.7 ± 1.1	48.4 ± 1.2	50.3 ± 1.2	43.7 ± 2.2	46.1 ± 6.4	55.9 ± 0.3	49.2 ± 3.5
6	48.7 ± 1.9	48.1 ± 1.8	47.2 ± 1.9	51.7 ± 2.3	52.3 ± 1.1	49.1 ± 1.2	51.1 ± 1.3	44.1 ± 2.3	49.9 ± 4.3	57.3 ± 0.4	50.0 ± 3.5
6.25	49.2 ± 1.9	48.6 ± 1.8	47.8 ± 1.9	52.7 ± 2.3	52.3 ± 1.1	48.9 ± 1.2	50.9 ± 1.2	43.7 ± 2.2	51.5 ± 4.5	58.5 ± 0.4	50.4 ± 3.9
6.5	49.8 ± 2.0	49.3 ± 1.9	48.2 ± 1.9	52.1 ± 2.3	52.9 ± 1.1	49.1 ± 1.2	51.3 ± 1.2	43.5 ± 2.2	60.2 ± 7.1	61.4 ± 0.4	51.8 ± 5.4
6.75	50.6 ± 2.0	50.0 ± 2.0	48.5 ± 2.0	51.8 ± 2.3	53.8 ± 1.1	49.7 ± 1.2	51.9 ± 1.3	43.9 ± 2.3	56.6 ± 7.5	62.3 ± 0.5	51.9 ± 4.9
7	51.2 ± 2.1	50.6 ± 2.0	49.4 ± 2.0	52.8 ± 2.3	53.7 ± 1.1	49.6 ± 1.2	51.8 ± 1.2	43.8 ± 2.2	57.5 ± 5.1	64.4 ± 0.5	52.5 ± 5.4
7.25	51.8 ± 2.1	51.0 ± 2.1	48.3 ± 2.1	51.8 ± 2.4	54.7 ± 1.1	49.9 ± 1.2	52.4 ± 1.3	43.0 ± 2.2	59.4 ± 5.2	65.9 ± 0.5	52.8 ± 6.2
7.5	52.2 ± 2.1	51.5 ± 2.0	49.7 ± 2.0	52.6 ± 2.3	54.4 ± 1.1	50.3 ± 1.2	52.7 ± 1.3	43.4 ± 2.2	58.8 ± 8.4	67.4 ± 0.5	53.3 ± 6.3
7.75	52.7 ± 2.2	51.9 ± 2.1	49.9 ± 2.0	53.5 ± 2.3	54.6 ± 1.1	50.5 ± 1.2	52.9 ± 1.3	43.5 ± 2.2	58.3 ± 8.7	69.3 ± 0.6	53.7 ± 6.7
8	53.5 ± 2.2	52.7 ± 2.2	50.3 ± 2.2	52.7 ± 2.4	55.7 ± 1.1	51.0 ± 1.2	53.5 ± 1.3	43.8 ± 2.3	61.7 ± 5.8	69.6 ± 0.6	54.4 ± 6.9
8.25	53.3 ± 2.2	52.6 ± 2.1	50.5 ± 2.1	53.5 ± 2.4	56.0 ± 1.1	51.5 ± 1.2	53.7 ± 1.3	44.4 ± 2.2	61.5 ± 6.0	70.7 ± 0.6	54.8 ± 7.0
8.5	54.2 ± 2.3	53.4 ± 2.2	50.8 ± 2.2	52.6 ± 2.4	57.1 ± 1.2	51.9 ± 1.3	54.2 ± 1.3	44.6 ± 2.4	60.3 ± 6.2	71.7 ± 0.7	55.1 ± 7.1
8.75	55.5 ± 2.3	54.8 ± 2.3	52.1 ± 2.3	53.5 ± 2.4	56.9 ± 1.1	52.0 ± 1.2	54.4 ± 1.3	44.9 ± 2.3	67.6 ± 6.3	73.5 ± 0.7	56.5 ± 8.2
9	53.6 ± 2.3	52.9 ± 2.2	51.2 ± 2.2	54.3 ± 2.4	56.8 ± 1.2	51.7 ± 1.3	54.1 ± 1.4	44.2 ± 2.3	61.8 ± 6.6	72.8 ± 0.7	55.3 ± 7.6

Table S.6: Summary of P estimates at optimal step size l_{optimal} for the thymine variants.

DNA variant	$P \pm \text{SEM}$ (nm)									$P \pm \text{SD}$ (nm)	l_{optimal} (nm)
	$\langle \cos\theta \rangle$	$\langle \theta^2 \rangle$	$\langle \theta^4 \rangle$	$\langle \theta^6 \rangle$	$\langle \mathbf{R} \cdot \mathbf{u}_0 \rangle$	$\langle R^2 \rangle$	$\langle R^4 \rangle$	$\langle R^4 \rangle - \langle R^2 \rangle^2$	$-\ln G(\theta)$	$\langle \cos\chi \rangle$	
1	53.5 ± 2.3	53.3 ± 2.2	53.1 ± 2.3	54.5 ± 2.5	52.8 ± 1.1	50.3 ± 1.2	51.8 ± 1.2	51.1 ± 3.0	54.7 ± 6.0	58.2 ± 0.3	5.5
2	60.1 ± 3.1	58.0 ± 3.0	52.8 ± 2.9	53.8 ± 3.0	65.7 ± 1.9	59.5 ± 1.9	62.3 ± 2.0	54.0 ± 3.6	55.4 ± 7.5	57.0 ± 0.4	6.75
3	54.3 ± 2.0	53.7 ± 1.9	51.9 ± 2.0	54.2 ± 2.3	55.9 ± 1.1	53.7 ± 1.2	55.4 ± 1.2	52.5 ± 2.6	54.9 ± 6.3	60.8 ± 0.4	5.75
4	57.8 ± 3.3	56.6 ± 3.2	53.5 ± 3.0	55.0 ± 3.3	61.6 ± 2.0	56.4 ± 1.9	58.8 ± 2.0	52.4 ± 3.9	53.0 ± 6.3	59.6 ± 0.5	5.75
5	53.5 ± 2.1	52.1 ± 2.0	50.0 ± 2.1	54.3 ± 2.4	59.8 ± 1.5	56.5 ± 1.5	59.0 ± 1.6	50.3 ± 2.5	44.0 ± 3.8	49.8 ± 0.3	5.25
6	45.5 ± 1.9	45.6 ± 1.8	46.4 ± 1.9	49.9 ± 2.1	46.6 ± 1.0	44.7 ± 1.1	45.2 ± 1.1	48.5 ± 2.8	49.0 ± 6.1	51.2 ± 0.3	5.5
7	79.6 ± 3.7	78.7 ± 3.7	75.2 ± 3.9	74.3 ± 4.5	85.8 ± 2.1	80.5 ± 2.3	82.7 ± 2.4	78.2 ± 5.2	81.0 ± 8.4	84.9 ± 0.7	7.75

Supporting References

1. Vologodskaia, M. Y. and Vologodskii, A. V. 2002. Contribution of the intrinsic curvature to measured DNA persistence length. *J. Mol. Biol.* 317:205–213.

2. Peters, J. P., Yelgaonkar, S. P., Srivatsan, S. G., Tor, Y., and Maher III, L. J. 2013. Mechanical properties of DNA-like polymers. *Nucleic Acids Res.* 41:10593–604.

3. Rivetti, C., Guthold, M., and Bustamante, C. 1996. Scanning force microscopy of DNA deposited onto mica: equilibration versus kinetic trapping studied by statistical polymer chain analysis. *J. Mol. Biol.* 264:919–932.

4. Neuman, K. C. and Nagy, A. 2008. Single-molecule force spectroscopy: optical tweezers, magnetic tweezers and atomic force microscopy. *Nat. Methods.* 5:491–505.

5. Wang, H., Dodd, I. B., Dunlap, D. D., Shearwin, K. E., and Finzi, L. 2013. Single molecule analysis of DNA wrapping and looping by a circular 14mer wheel of the bacteriophage 186 CI repressor. *Nucleic Acids Res.* 41:5746–5756.

6. Rivetti, C. and Codeluppi, S. 2001. Accurate length determination of DNA molecules visualized by atomic force microscopy: evidence for a partial B-to A-form transition on mica.. *Ultramicroscopy.* 87:55.

7. Faas, F., Rieger, B., Van Vliet, L., and Cherny, D. 2009. DNA deformations near charged surfaces: electron and atomic force microscopy views. *Biophys. J.* 97:1148–1157.

8. Moukhtar, J., Faivre-Moskalenko, C., Milani, P., Audit, B., Vaillant, C., Fontaine, E., Mongelard, F., Lavorel, G., St-Jean, P., and Bouvet, P. 2010. Effect of genomic long-range correlations on DNA persistence length: from theory to single molecule experiments. *J. Phys. Chem. B.* 114:5125–5143.

9. Wiggins, P. A., van der Heijden, T., Moreno-Herrero, F., Spakowitz, A. J., Phillips, R., Widom, J., Dekker, C., and Nelson, P. C. 2006. High flexibility of DNA on short length scales probed by atomic force microscopy. *Nat. Nanotechnol.* 1:137–141.

10. Vologodskii, A. and Frank-Kamenetskii, M. D. 2013. SURVEY AND SUMMARY Strong bending of the DNA double helix. *Nucleic Acids Res.* 41:6785–6792.

11. Abels, J. A., Moreno-Herrero, F., van der Heijden, T., Dekker, C., and Dekker, N. H. 2005. Single-molecule measurements of the persistence length of double-stranded RNA. *Biophys. J.* 88:2737–2744.