Temperature dependence of the DNA double helix at the nanoscale: structure, elasticity and fluctuations

Sam Meyer, Daniel Jost, Nikos Theodorakopoulos, Michel Peyrard, Richard Lavery and Ralf Everaers

Supporting Material

I. Dataset

Trinuc	Occ
A**A**A	2
A**A**C	2
C**A**A	3
G**A**T	2
T**A**G	3
T**A**T	2
A**C**A	2
C**C**C	9
G**C**G	9
T**C**T	2

Tetranuc	Occ
A**AA**C	2
C**AA**A	3
A**AC**A	2
T**AG**A	3
G**AT**A	2
A**CA**A	2
G**CG**C	5
A**GA**T	2
C**GC**G	4
G**GG**G	9
A**TA**G	2

Table S1: Number of occurrences of trinucleotide (intra parameters) and tetranucleotide (step parameter) sequences. We analyzed these sequences separately, rather than the central mono- or dinucleotides where the Gaussian approximation is not valid (see text and Fig. S3).

II. Methods

Analysis of the trajectory DNA conformations were analyzed by the program Curves+ (1), which computes a full set of helical, backbone and groove parameters. Curves+ uses the commonly agreed "Tsukuba" reference frame for the description of each base (2), and the Cambridge convention for the names and signs of all helical parameters (3). We focused on the internal base-pair and the base-pair step deformations. Starting from the time series of these parameters as provided by Curves+ (and the associated trajectory analysis program Canal), subsequent analysis (Boltzmann inversion, covariance and matrix inversions, error estimates, Monte Carlo generation of "artificial data", linear regressions, plots) was developed in Python, with the use of the Numpy/Scipy libraries (4) and the MatPlotLib library for plotting (5).

We checked for the absence of any melting or end-fraying of the oligomers, using the intra parameters obtained at the different positions and temperatures (Fig. S2), especially the highest (325K and 350K). Only two cases showed aberrant values, which were attributed to partial melting and eliminated. We also checked that the distributions are monomodal and stable at all temperatures (see Fig. S1), and we compute the *covariance matrix*. By inverting the latter, we get the *stiffness matrix* (Eq. 2), which will be used to estimate the enthalpic and entropic contributions according to Eq. 6.

A comparison between the distributions obtained with different occurrences of given sequence fragments shows that the statistical power is sufficient to resolve the relatively small temperature-induced variations, except in some cases at the lowest temperatures where kinetics are slower. To quantify this effect, we systematically estimate statistical confidence intervals for the computed quantities (see next section). Because the runs at the lower temperatures have larger error bars, their relative weight in the subsequent modeling is proportionnally reduced: to compensate this effect, we conducted more simulations for this range of temperatures.

Linear fitting of the simulation data To estimate the errors on the computed stiffness, covariance and equilibrium values, we used the *block averaging method* (7). This is a generalization of the simple idea that the statistical errors can be estimated by comparing the values obtained by splitting a trajectory into shorter parts ("blocks"), and computing the quantity of interest separately for each block.

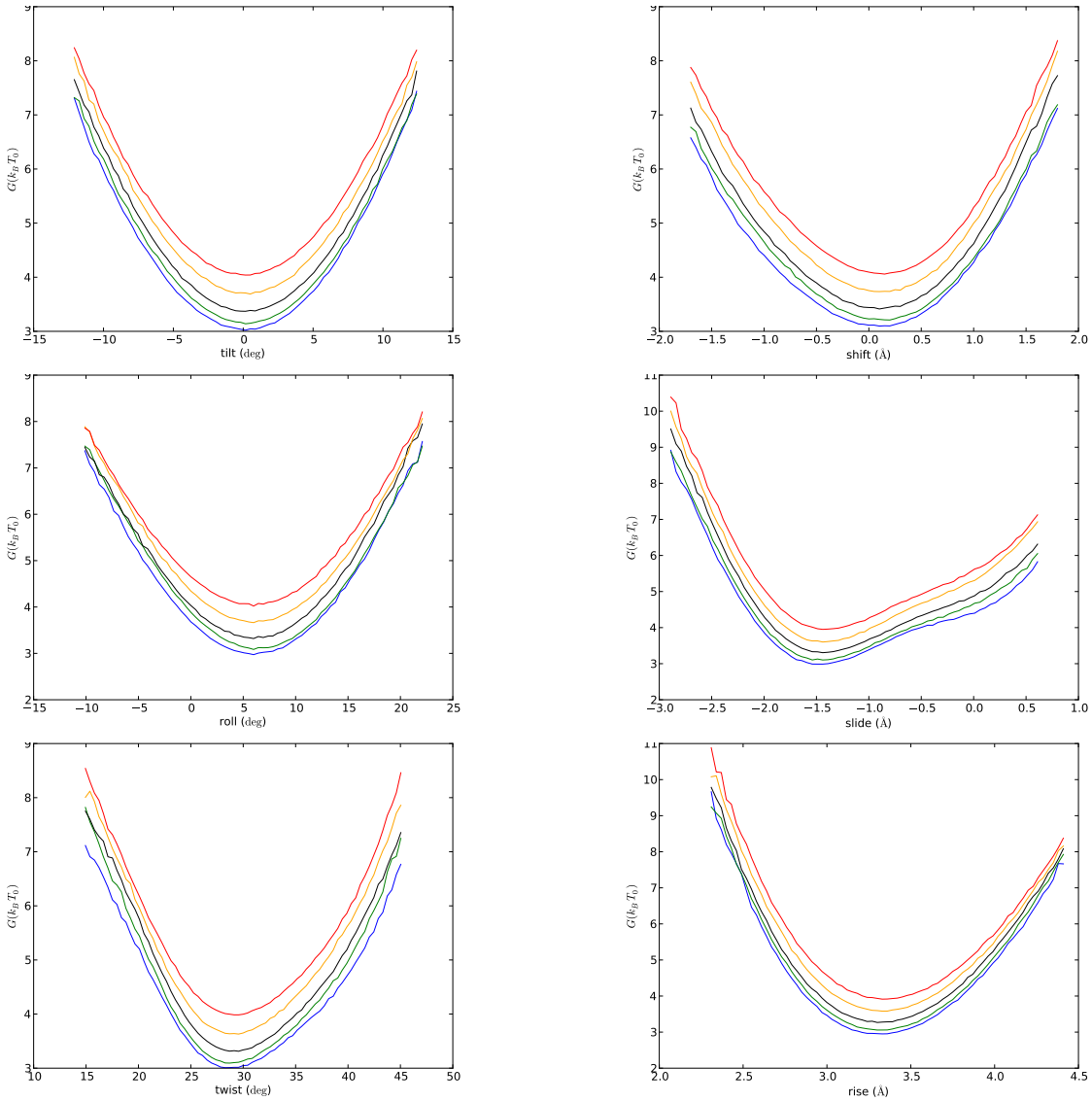

Figure S1: Free energy functions, obtained by Boltzmann inversion of the MD distributions projected along the 6 step degrees of freedom (left column: angular, right column: translational), for the GG sequence. Simulated temperatures: 273K (blue), 283K (green), 300K (black), 325K (orange), 350K (red). The distributions have the same shape at all temperatures (even 350 K), indicating the absence of melting in the analyzed base-pairs. In particular, they are all monomodal: we consider the harmonic approximation in the subsequent analysis. Bimodal distributions were observed in some very specific cases, already noticed in (6).

We separate the enthalpic and entropic contributions by finding the best linear fit for the temperature dependence of (i) each of the stiffness matrix elements (Eq. 6), and (ii) each degree of freedom of the equilibrium values (Eq. 7). The weighted least square fitting procedure from (8) gives a reduced weight to data points with poor precision, and provides not only the best fit parameters (the enthalpic and entropic stiffness contributions), but also error estimates for these parameters and correlations between them. The values and errors of the stiffness matrix at any temperature can be calculated from these numbers, as depicted in the plots by a shaded area rather than a single line: see for instance Fig. 1B.

We introduce a quantitative criterion, to discriminate between the cases where the effect of temperature on the data can be reliably estimated, from those where it is too small to be quantified. The "f-test" (8) compares the accuracy of a model with an enthalpic and entropic contribution (two parameters) and that of a purely enthalpic model, and determines if the gain in precision justifies the more detailed model. Our criterion is a threshold on a number obtained from a combination of this test with the relative uncertainty on the temperature t_s and Pearson's correlation coefficient between the data points. The two latter contributions were added to eliminate some specific cases where the f-test was positive, while the results of the fit lacked precision.

Validation of the analysis method The covariance and stiffness matrices are related by an inversion operation (Eq. 2), which mixes the errors of the different stiffness matrix elements. We made the simplistic hypothesis that the latter are

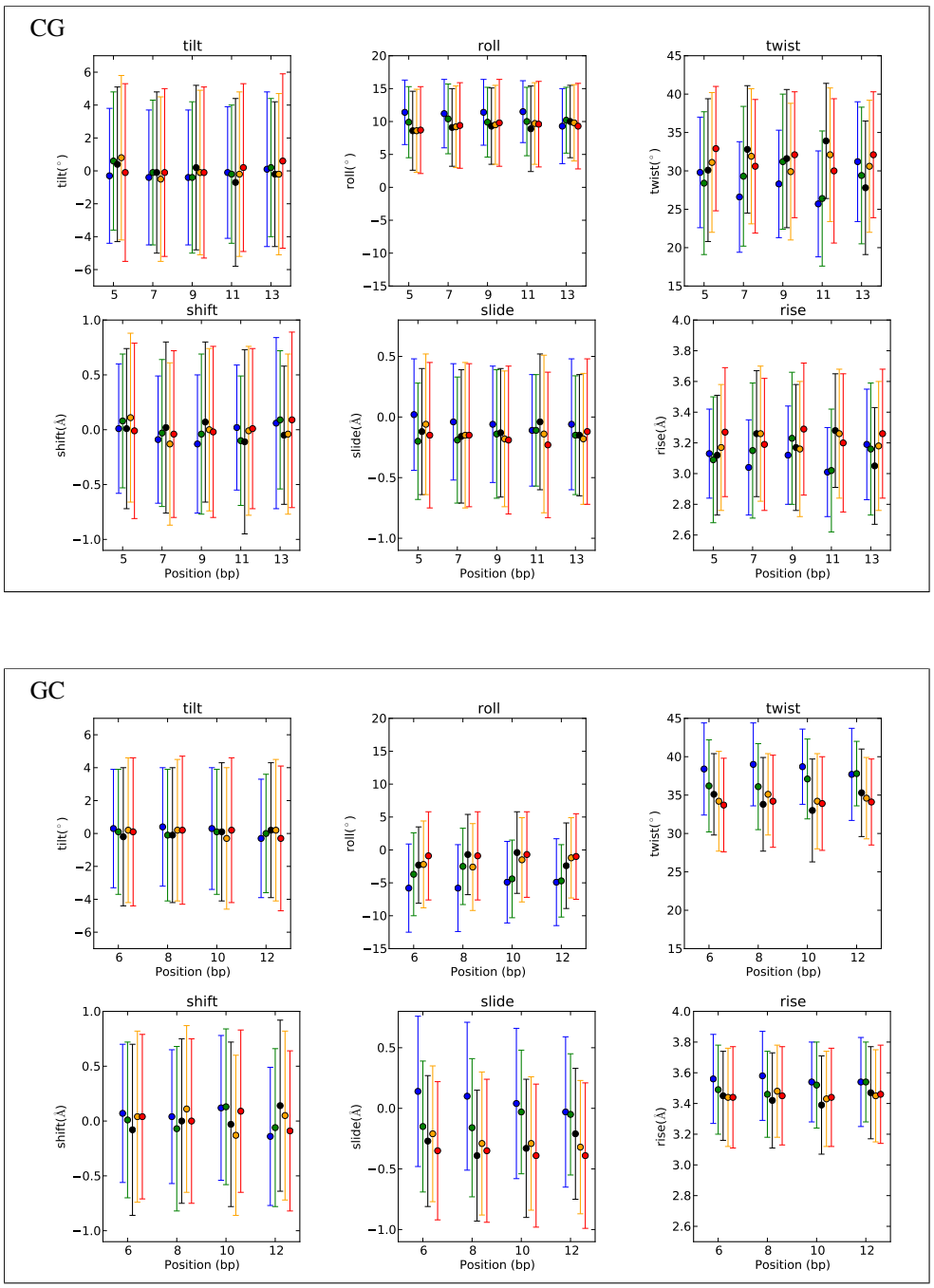

Figure S2: Mean values (dots) and standard deviations (bars) of the base-pair step parameters computed along the oligomer CGCG, at the simulated temperatures: from 273 K (blue) to 350 K (red) (same color code as Fig. S1). The outer 4 bps on either side of the oligomer have been excluded of the analysis to eliminate possible end effects. **(Upper panel)** Odd bps (CG steps); **(lower panel)** even bps (GC steps). The overall regularity of the values shows the absence of long-acting fraying or melting, even at $T = 350K$. Comparison of the values obtained at the different positions for the same dinucleotide indicates that for the lowest temperatures, the remaining statistical noise is sometimes important: see for instance the mean values of twist, for the GC steps (lower panel).

independent. From the fitted stiffness model, the computation of a model-derived covariance matrix (including the error estimates) then involves (i) generating an ensemble of stiffness matrices representative of the estimated uncertainties by Monte Carlo ; (ii) inverting each of these matrices and (iii) computing the average and standard deviation of the resulting ensemble of covariance matrices. The hypothesis of independence of the stiffness errors implies that the computed covariance errors are an upper bound. A comparison with the covariance error bars computed directly on the data shows that they are indeed larger, but are of same order of magnitude, and can therefore be considered acceptable given the possible equilibration problems (see for instance Fig. 1, lower panel).

To validate the entire procedure, we apply it to "artificial" trajectories, which were generated so as to mimic the properties of the real data, but where we know by construction the "true" parameters. The application of the procedure indeed allows the "input" properties to be recovered within the estimated error bars.

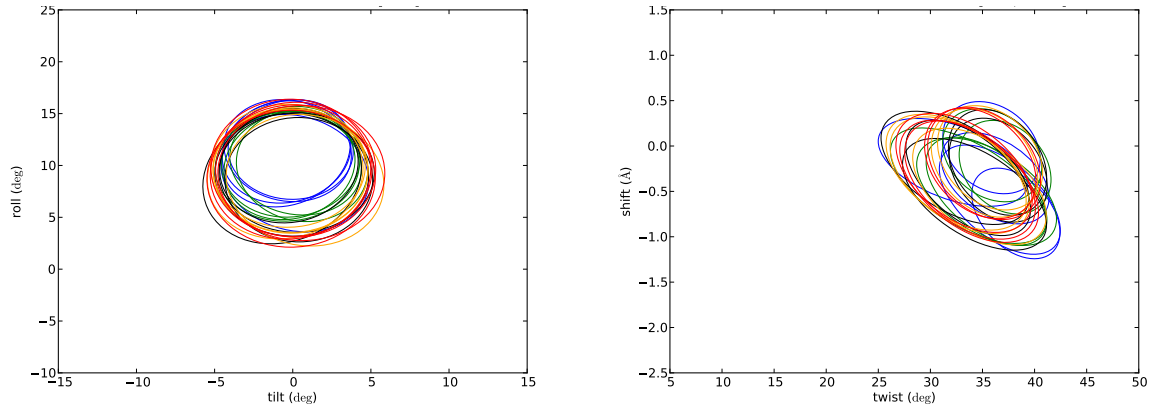

Figure S3: Covariance ellipses computed on the individual positions for the different temperatures (from blue-cold to red-hot, same colors as previous figures). **(Left)** Rise-slide coupling for the CG dinucleotides. Despite a remaining level of statistical noise, one can follow the evolution of the covariance with temperature : the Gaussian approximation is valid. **(Right)** Twist-shift coupling for AA dinucleotides located in the A**AA**C and C**AA**A tetranucleotides. They clearly divide into 2 groups (corresponding to the two different contexts), especially for larger temperatures where the system is better equilibrated, confirming the influence of the neighbor bps on the dinucleotide conformations.

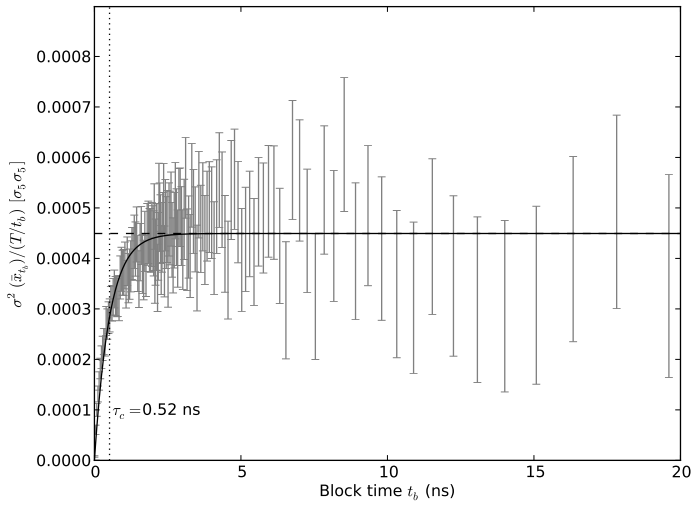

Figure S4: Computation of the statistical errors by the "block averaging" procedure, here for the rise variance of GC at 325K. For each "block size" t_b, we split the trajectory into parts, and the "block error" is then given by standard deviation of the resulting values, divided by the square root of the number of blocks $\sqrt{T/t_b}$. When the block size is small compared to the correlation time of the considered quantity (left part), the block values are not independent, and the computed error is therefore underestimated. For longer sizes, the samples become uncorrelated and the computed error becomes independent of the block size: the value of this plateau provides a reliable error estimate. In each case, we compute the confidence interval (vertical bar) of the computed error, and we fit the curve with an exponential function, yielding the correlation time τ_c. The details of the method are given in (7). In most cases, the correlation time is of the order of 1 ns; in some cases however, the data exhibited very long correlation times, which we attributed to large-scale rearrangements of the molecule: the computed error was then multiplied by a factor 2 for security.

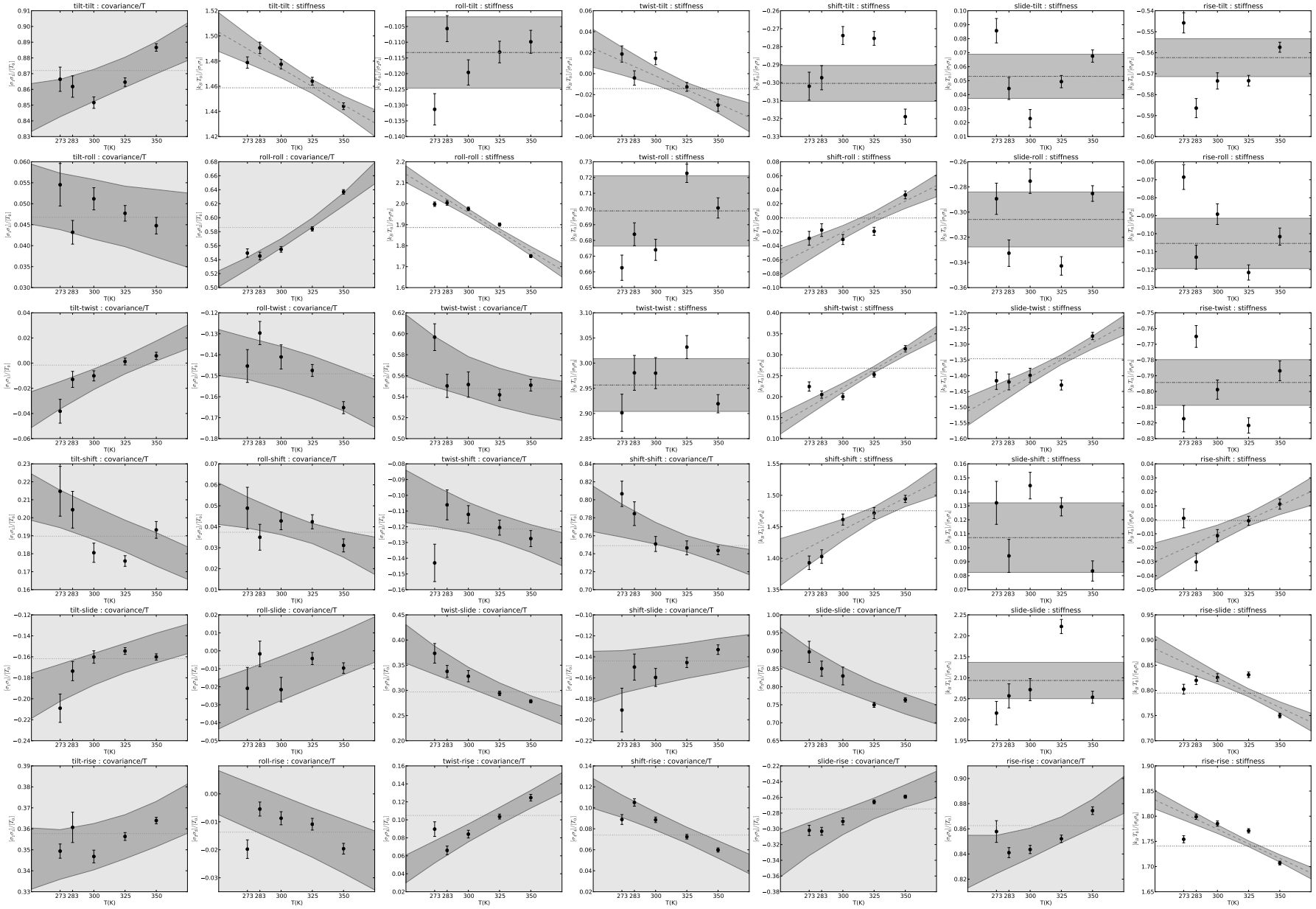

Figure S5: Synthetic illustration of the complete fitting procedure of the GG step parameters. All stiffness matrix elements (white background, upper right half) are fitted independently of each other. The covariance elements (grey background) are obtained by inverting the stiffness matrix. In this operation, the different degrees of freedom get mixed: a given covariance element may therefore exhibit a temperature dependence even though its stiffness counterpart (transposed) does not.

II. Internal bp elasticity

Entropic contribution to the stiffness and equilibrium values: Results for the different parameters and sequences

Seq	buckle	propel	opening	shear	stretch	stagger
AAA	∞	∞	$\mathbf{2.82 \pm 0.12}$	$\mathbf{5.21 \pm 0.34}$	$\mathbf{2.32 \pm 0.04}$	$\mathbf{3.74 \pm 0.25}$
AAC	∞	$\mathbf{2.60 \pm 0.19}$	$\mathbf{3.05 \pm 0.12}$	$\mathbf{6.50 \pm 0.46}$	$\mathbf{2.39 \pm 0.04}$	$\mathbf{2.99 \pm 0.14}$
CAA	∞	∞	$\mathbf{2.78 \pm 0.12}$	$\mathbf{5.62 \pm 0.31}$	$\mathbf{2.21 \pm 0.04}$	$\mathbf{4.80 \pm 0.48}$
GAT	∞	∞	$\mathbf{2.34 \pm 0.15}$	$\mathbf{4.23 \pm 0.33}$	$\mathbf{1.95 \pm 0.05}$	$\mathbf{2.92 \pm 0.14}$
TAG	∞	∞	$\mathbf{2.83 \pm 0.13}$	$\mathbf{6.76 \pm 0.55}$	$\mathbf{1.95 \pm 0.04}$	$\mathbf{5.50 \pm 0.44}$
TAT	∞	∞	$\mathbf{3.25 \pm 0.20}$	$\mathbf{4.34 \pm 0.21}$	$\mathbf{1.98 \pm 0.04}$	$\mathbf{2.93 \pm 0.13}$
ACA	∞	∞	$\mathbf{2.95 \pm 0.09}$	$\mathbf{3.22 \pm 0.13}$	$\mathbf{2.51 \pm 0.06}$	$\mathbf{4.59 \pm 0.38}$
CCC	∞	$\mathbf{4.44 \pm 0.25}$	$\mathbf{3.41 \pm 0.06}$	$\mathbf{4.15 \pm 0.13}$	$\mathbf{2.51 \pm 0.04}$	$\mathbf{3.94 \pm 0.09}$
GCG	∞	∞	$\mathbf{2.90 \pm 0.06}$	$\mathbf{3.91 \pm 0.14}$	$\mathbf{2.22 \pm 0.03}$	$\mathbf{3.58 \pm 0.16}$
TCT	∞	∞	$\mathbf{3.38 \pm 0.14}$	$\mathbf{2.91 \pm 0.09}$	$\mathbf{2.50 \pm 0.06}$	$\mathbf{3.40 \pm 0.16}$

Table S2: Spinodal temperature computed separately for the diagonal elements of the intra-bp stiffness matrix, in units of 300K: $t_s = k_h/(300K.k_s)$. This value is the inverse of the relative entropic weight shown in Fig. 3: an infinite spinodal temperatures signifies the absence of an entropic contribution (buckle, propeller), and a strong contributions yields a value close to 1 (stretch, opening).

Seq	buckle (°)	propel (°)	opening (°)	shear (Å)	stretch (Å)	stagger (Å)
AAA	8.21	-6.25	-1.52	0.0	-0.05	0.21
	8.79	-17.45	2.91	0.12	0.01	0.00
AAC	0.0	-5.95	0.0	0.0	-0.06	0.29
	5.65	-16.38	2.21	0.14	0.02	0.20
CAA	5.45	0.0	0.0	-0.06	-0.06	0.12
	7.34	-11.02	3.92	0.06	0.02	-0.16
GAT	0.0	0.0	1.54	-0.12	-0.08	0.37
	-3.40	-13.80	3.13	0.09	0.03	0.13
TAG	3.81	-3.66	0.0	-0.03	-0.06	0.27
	6.39	-12.84	5.12	0.05	0.03	0.17
TAT	0.0	0.0	-2.48	0.0	-0.07	0.18
	2.81	-11.46	3.52	-0.07	0.04	0.11
ACA	0.0	0.0	-1.12	0.0	-0.06	0.21
	4.41	-9.97	1.87	-0.03	0.05	0.08
CCC	1.62	1.90	-1.27	0.03	-0.03	0.0
	-1.83	-4.28	0.49	0.01	0.01	-0.25
GCG	5.73	0.0	-1.81	0.0	-0.05	0.18
	2.27	-9.27	0.22	0.03	0.02	0.20
TCT	0.0	-7.20	-0.84	0.0	-0.04	0.07
	2.71	-13.45	1.19	0.04	0.02	0.05

Table S3: Temperature dependence of the equilibrium values, for the intra-bp parameters. For each sequence, the first line indicates the entropic contribution at room temperature ($T_0.q_0'(s)$ in the notation of Eq. 7 in the main text), and the second line gives the value at room temperature ($q_0^0(s)$). From these two numbers, the equilibrium value at any temperature T is given by: $q_0(s,T) = q_0^0(s) - (T/T_0 - 1)(T_0.q_0'(s))$.

Spinodal decomposition of the base-pairs: temperature and direction

	t_s [300K]	buckle	propel	opening	shear	stretch	stagger
AAA	2.034	-0.287	0.337	-0.56	0.12	-0.606	-0.33
AAC	2.179	-0.019	0.149	-0.05	-0.014	-0.856	-0.492
CAA	2.102	0.067	-0.051	0.456	-0.057	0.873	0.139
GAT	1.743	0.109	0.156	-0.427	0.04	0.799	0.376
TAG	1.848	0.129	-0.072	-0.065	-0.032	-0.955	-0.246
TAT	1.802	0.075	0.094	-0.479	-0.072	-0.84	-0.213
ACA	1.753	0.096	0.082	-0.565	-0.36	-0.707	-0.188
CCC	1.877	-0.04	-0.14	0.554	-0.23	0.782	0.083
GCG	2.034	0.002	0.213	-0.521	0.039	-0.792	-0.235
TCT	1.802	-0.024	0.141	-0.569	0.361	-0.711	-0.143

Table S4: Spinodal decomposition of the base-pairs: temperature and direction of instability. The components of the eigenvector are expressed in reduced units of the different parameters, where they have comparable values at T_0. It therefore represents their respective weight in the direction of instability. The element of strongest weight is stretch, then opening.

III. Step elasticity

All following values are also given in separate ASCII files, with the associated error estimates

Equilibrium values

Param	CG	CA	TA	AG	GG	AA	GA	AT	AC	GC
Twist (°)	-7.92	0.	4.74	4.30	0.	0.	7.10	0.	2.27	9.943
Tilt (°)	0.	0.	0.	0.	0.	0.	0.	0.93	0.	0.
Roll (°)	2.86	0.	-9.60	-2.20	1.10	-3.95	-6.18	-1.59	-2.77	-11.96
Shift (Å)	-0.076	-0.192	0.565	-0.436	0.	-0.363	0.483	-0.400	0.	0.
Slide (Å)	0.189	0.	0.	0.339	-0.222	0.367	0.169	-0.199	0.	0.956
Rise (Å)	-0.49	0.	-0.348	-0.128	-0.163	-0.308	-0.161	-0.264	-0.102	0.

Table S5: Temperature dependence of the equilibrium conformation ($q'_0(s).T_0$ in the notation of Eq. 7). From these values and those of Table S6, the equilibrium value at any temperature T is given by: $q_0(s,T) = q^0_0(s) - (T/T_0 - 1)(T_0.q'_0(s))$.

Param	CG	CA	TA	AG	GG	AA	GA	AT	AC	GC
Twist (°)	**30.30**	**27.76**	**31.21**	**32.37**	**29.99**	**33.54**	**36.2**	**29.71**	**32.67**	**35.56**
	36.1	37.3	37.8	31.9	32.9	35.1	36.3	29.3	31.5	33.6
	27.93	25.69	29.74	28.84	29.93	31.04	33.32	29.02	31.46	33.13
Tilt (°)	**-0.02**	**1.80**	**0.11**	**-1.66**	**0.18**	**-2.21**	**-0.73**	**0.30**	**0.02**	**0.03**
	0.	0.5	0.	-1.7	-0.1	-1.4	-1.5	0.	-0.1	0.
	0.	-0.01	0.	0.99	0.84	-1.52	-0.28	0.	-0.36	0.
Roll (°)	**9.74**	**11.80**	**8.70**	**4.31**	**6.02**	**1.15**	**1.35**	**-0.75**	**-1.13**	**-2.86**
	5.4	4.7	3.3	4.5	3.6	0.7	1.9	1.1	0.7	0.3
	8.75	8.36	10.33	2.75	5.28	2.31	2.27	0.21	0.63	1.23
Shift (Å)	**-0.01**	**-0.15**	**0.20**	**-0.47**	**0.05**	**-0.41**	**-0.01**	**0.00**	**0.32**	**0.01**
	0.	0.09	0.	0.09	0.05	-0.03	-0.28	0.	0.13	0.
	0.	0.18	0.	0.05	0.07	-0.14	-0.01	0.	-0.12	0.
Slide (Å)	**-0.13**	**-0.40**	**-0.28**	**-0.54**	**-1.16**	**-0.17**	**-0.66**	**-0.80**	**-0.78**	**-0.21**
	0.41	0.53	0.05	-0.25	-0.22	-0.08	0.09	-0.59	-0.58	-0.38
	-0.76	-0.57	-1.02	-1.33	-1.64	-1.04	-0.92	-1.08	-1.16	-1.09
Rise (Å)	**3.16**	**3.30**	**3.17**	**3.38**	**3.36**	**3.27**	**3.42**	**3.20**	**3.39**	**3.45**
	3.39	3.33	3.42	3.34	3.42	3.27	3.37	3.31	3.36	3.4
	3.26	3.13	3.34	3.38	3.61	3.31	3.37	3.2	3.38	3.37

Table S6: Equilibrium conformations at room temperature, q^0_0 . We give the value estimated from our simulations (bold), from the analysis of DNA-protein crystallographic structures (9) (second line), and from a previous generation of MD simulations (10) (third line).

Stiffness values

Param	CG	CA	TA	AG	GG	AA	GA	AT	AC	GC
Twist-twist	0.0	0.0	0.027	0.0	0.0	-0.019	0.0	0.0	0.0	0.0
Twist-tilt	0.0	0.0	-0.003	0.0	0.003	0.0	-0.007	0.0	0.0	0.002
Twist-roll	0.005	0.0	0.0	0.0	-0.002	0.0	0.0	0.0	0.0	-0.006
Twist-shift	0.0	-0.041	0.063	0.0	-0.054	0.033	0.0	0.0	0.0	0.0
Twist-slide	0.0	-0.109	-0.193	0.0	-0.078	0.147	-0.245	0.065	0.0	-0.0
Twist-rise	-0.173	-0.154	-0.207	0.0	0.0	0.0	0.066	0.0	0.0	-0.121
Tilt-tilt	0.015	0.0	0.0	0.011	0.005	0.015	0.019	0.007	0.011	0.025
Tilt-roll	0.0	0.0	0.0	0.0	0.0	-0.005	0.0	-0.002	0.0	-0.0
Tilt-shift	-0.023	-0.031	0.0	-0.037	0.0	-0.045	-0.103	0.0	0.0	-0.096
Tilt-slide	0.037	0.0	0.0	0.075	0.0	0.124	0.0	0.0	0.0	-0.041
Tilt-rise	0.0	0.0	0.0	-0.095	0.0	-0.099	-0.226	-0.104	0.0	-0.0
Roll-roll	0.020	0.004	0.0	0.0	0.012	0.007	0.005	0.011	0.015	0.011
Roll-shift	0.0	0.0	0.0	0.0	-0.028	0.0	0.071	0.0	0.0	-0.0
Roll-slide	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.121	0.163
Roll-rise	-0.141	0.0	0.0	0.104	0.0	0.114	0.0	0.0	0.178	0.250
Shift-shift	0.0	0.0	-1.249	0.0	-0.427	0.676	0.0	0.0	0.0	0.0
Shift-slide	0.0	0.273	-0.309	0.431	0.0	0.343	0.0	0.0	-0.623	0.0
Shift-rise	0.0	0.0	0.0	0.675	-0.268	0.847	0.0	0.0	0.0	0.293
Slide-slide	0.866	0.0	0.836	0.0	0.0	0.0	2.002	2.712	1.019	1.225
Slide-rise	0.0	2.177	1.518	0.0	0.0	-0.833	0.0	3.406	1.560	0.0
Rise-rise	5.722	2.613	2.264	3.655	1.537	5.867	5.326	6.833	6.007	3.45

Table S7: Entropic stiffness constants, k_s, in units of (kcal/mol)/(deg^2.[300K]), (kcal/mol)/(Å^2.[300K]), or (kcal/mol)/(deg.Å.[300K]). From the values k_0 at T_0=300K (Table S8) and these numerical values, the stiffness at a given temperature T is given by (Eq. 6): $k(T) = k_0 - (T/T_0 - 1)k_s$.

Param	CG	CA	TA	AG	GG	AA	GA	AT	AC	GC
Twist-twist	0.0184	0.0169	0.0248	0.0305	0.0386	0.0318	0.0433	0.0422	0.0483	0.034
	(0.0227)	(0.021)	(0.0357)	(0.0441)	(0.0482)	(0.0461)	(0.0422)	(0.0463)	(0.0489)	(0.0421)
Twist-tilt*	0.0001	0.0026	0.0001	0.0087	0.0	0.0097	0.0003	-0.001	0.0047	0.0002
Twist-roll	0.008	0.0089	0.0117	0.0094	0.0085	0.0115	0.0107	0.0126	0.0108	0.0062
Twist-shift*	-0.0003	-0.0097	0.0125	0.0922	0.0278	0.0805	0.012	-0.0166	0.0192	0.0029
Twist-slide	-0.0688	-0.0429	-0.0687	-0.0725	-0.174	-0.1107	-0.216	-0.101	-0.175	-0.139
Twist-rise	-0.255	-0.213	-0.212	-0.207	-0.205	-0.218	-0.174	-0.160	-0.189	-0.256
Tilt-tilt	0.0348	0.0284	0.0241	0.04	0.0442	0.0453	0.0416	0.0404	0.0433	0.0459
	(0.0278)	(0.0275)	(0.0245)	(0.0371)	(0.0414)	(0.0389)	(0.0392)	(0.0404)	(0.0411)	(0.0396)
Tilt-roll*	-0.0001	0.001	0.0001	0.0033	-0.0022	-0.0006	-0.0028	0.0003	0.0041	-0.0003
Tilt-shift	-0.107	-0.0451	-0.0097	-0.0367	-0.0547	-0.0269	-0.0623	0.0239	-0.0297	-0.1154
Tilt-slide*	0.0017	-0.003	-0.0156	0.0138	0.0101	0.0073	0.0082	-0.0147	-0.0436	-0.0046
Tilt-rise	0.0051	-0.0096	-0.0104	-0.166	-0.219	-0.236	-0.275	-0.0213	0.0575	0.0034
Roll-roll	0.0224	0.0209	0.0178	0.0215	0.0231	0.0231	0.0211	0.0262	0.0261	0.0248
	(0.0153)	(0.0184)	(0.0136)	(0.0227)	(0.0241)	(0.0235)	(0.0211)	(0.0272)	(0.0267)	(0.0275)
Roll-shift	-0.0019	0.0258	0.0077	0.0248	-0.0025	-0.0036	0.0287	0.0215	0.0338	0.0028
Roll-slide	-0.0034	-0.0124	0.0087	-0.0567	-0.036	-0.0788	-0.0262	-0.0369	0.03	0.0904
Roll-rise	-0.1039	-0.1066	-0.1016	-0.0447	-0.0251	-0.0066	-0.0292	0.0624	0.104	0.1651
Shift-shift	1.480	1.524	0.781	1.595	1.690	2.257	1.700	1.425	1.526	1.420
	(1.346)	(1.600)	(1.529)	(1.657)	(1.984)	(1.975)	(1.430)	(1.193)	(1.341)	(1.761)
Shift-slide*	0.0098	-0.0735	-0.076	0.252	0.1347	0.4312	0.3307	-0.269	-0.3643	-0.0199
Shift-rise*	0.0058	0.0585	0.145	0.3118	-0.022	0.653	0.572	-0.168	0.133	-0.0122
Slide-slide	2.362	1.716	1.457	1.951	2.469	2.373	2.592	4.346	3.559	3.184
	(2.034)	(2.286)	(2.269)	(2.706)	(3.215)	(2.914)	(2.518)	(3.310)	(2.974)	(2.708)
Slide-rise	1.095	1.314	0.860	1.329	1.928	0.960	1.405	2.857	2.539	2.258
Rise-rise	7.358	6.083	6.967	7.906	8.879	9.468	9.098	10.56	10.35	10.54
	(4.390)	(6.290)	(5.055)	(6.388)	(7.335)	(7.621)	(8.330)	(10.50)	(9.88)	(10.28)

Table S8: Stiffness k_0 at room temperature, $T_0 = 300K$, in units of (kcal/mol)/deg^2, (kcal/mol)/Å^2, or (kcal/mol)/(deg.Å) for angular, translational, or mixed deformations, respectively. For the diagonal elements of the stiffness matrix, we indicate in parenthesis the previous MD values from (10), obtained with different force fields, water model, and oligomer sequences. The non-diagonal parameters indicated with a star should be zero for self-complementary sequences (CG, GC, AT, TA) for symmetry reasons.

IV. Torque-control of DNA stability

Param	CG	CA	TA	AG	GG	AA	GA	AT	AC	GC
w_0 (°)	30.3	27.8	31.2	32.4	30.0	33.5	36.2	29.7	32.7	35.6
w_0' (°)	-7.9	0.0	4.7	4.3	0.0	0.0	7.1	0.0	2.3	9.9

Table S9: Natural helical twist w for the 10 different base-pair steps. At any temperature T, $w = w_0 - (T/T_0 - 1)w_0'$ ($T_0 = 300$K) and represents the equilibrium value of "Twist" given in Tables S5 and S6

Param	CG	CA	TA	AG	GG	AA	GA	AT	AC	GC
K_0	0.81	0.78	1.15	1.47	2.33	1.38	2.09	3.16	3.24	1.84
K_0'	-0.44	-0.72	0.77	-0.35	-0.56	-1.44	-2.30	0.73	0.05	-0.59

Table S10: Torsional stiffness K for the 10 different base-pair steps. At any temperature T, $K = K_0 - (T/T_0 - 1)K_0'$ ($T_0 = 300$K) and represents the inverse of the twist-twist element of the matrix $\underline{k}^{-1}$ with $\underline{k}$ the stiffness matrix given in Tables S7 and S8. In units of 10^{-2} kcal/mol/degree^2.

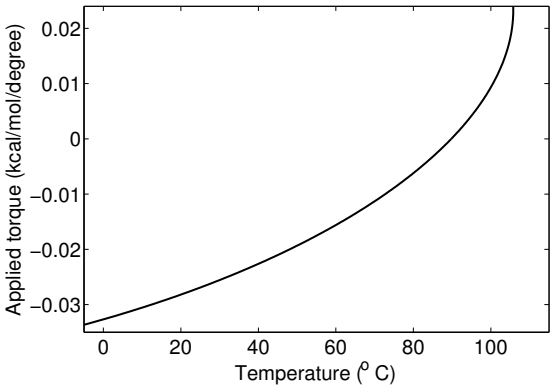

Figure S6: External torque applied on DNA *in vivo* for temperatures between 0°C and 110°C, maintaining a constant open fraction of 1% of the *T. thermophilus* genome (see text). Without external torque, the DNA melting occurs at $T_m \approx 91°$ C. A positive supercoiling maintains the stability of the double helix for an additional 15° C, up to a critical temperature $T_c \approx 106°$ C. The resulting twist excess is shown in Fig. 4 for the double-helical DNA, and on Fig. S7 for the small open fraction, which absorbs most of this excess.

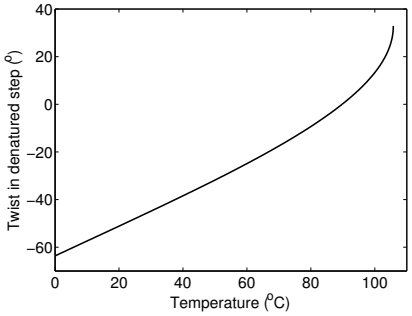

Figure S7: Average twist in the $\sim 1\%$ open DNA, as a result of the external torque (Fig. S6). Because of the much larger twist rigidity of double-helical DNA, the open fraction absorbs much more twist than the closed parts (compare with Fig. 4).

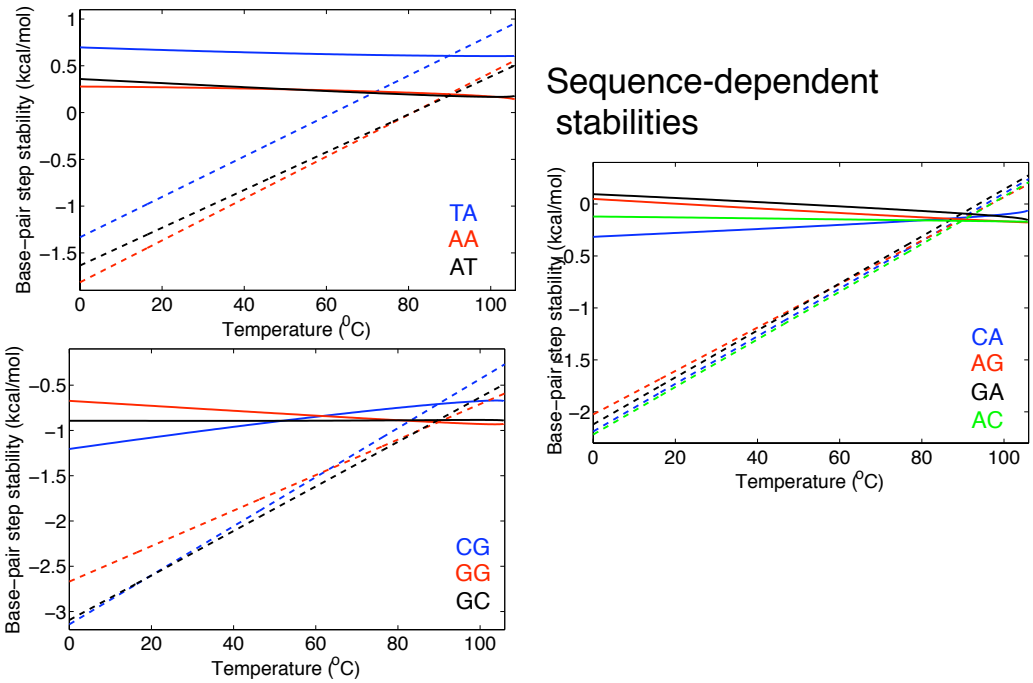

Figure S8: Double-helical stability of the individual step sequences, for unconstrained DNA (dashed) and under the biologically relevant torque (solid lines) as computed in Fig. S6. While the opening penalty becomes approximately constant at the sequence-neutral level, it is not rigorously true for the individual sequences. In particular, the order of stability is temperature-dependent, and is also different from that of unconstrained DNA (see for instance GC and GG). Note that the less stable sequences (upper left) are destabilized by the torque, but the stability of the double-helix is maintained by the cooperative terms between successive base-pairs.

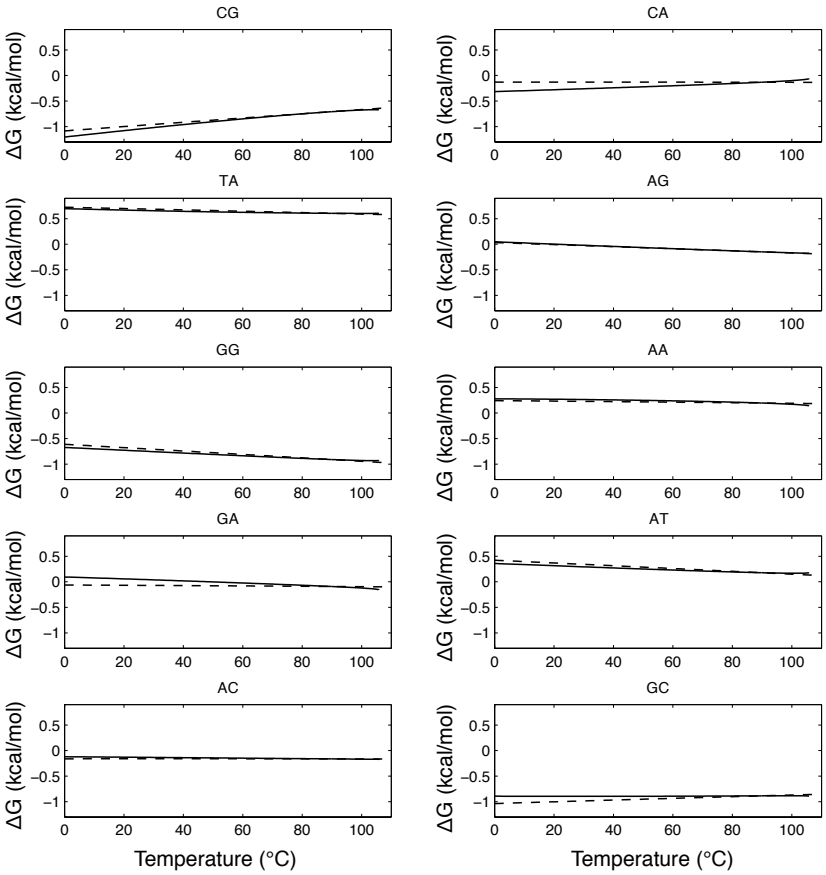

Figure S9: Base-pair step stabilities as predicted by the Benham model (11) (dashed line), and our model which includes the sequence- and temperature-dependent elasticity of the double-helix (solid line). The differences are only marginal, since the dominant contribution comes from the sequence-dependent melting properties.

Supporting references

(1) Lavery, R., M. Moakher, J. Maddocks, D. Petkeviciute, and K. Zakrzewska. 2009. Conformational analysis of nucleic acids revisited: Curves+. *Nucleic Acids Res.* 37(17):5917–5929.

(2) Olson, W. K., M. Bansal, S. K. Burley, R. E. Dickerson, M. Gerstein, S. C. Harvey, U. Heinemann, X.-J. Lu, S. Neidle, Z. Shakked, H. Sklenar, M. Suzuki, C.-S. Tung, E. Westhof, C. Wolberger, and H. M. Berman. 2001. A standard reference frame for the description of nucleic acid base-pair geometry. *J. Mol. Biol.* 313(1):229 – 237.

(3) Dickerson, R. 1989. Definitions and nomenclature of nucleic acid structure components. *Nucleic Acids Res.* 17(5):1797–1803.

(4) Jones, E., T. Oliphant, and P. Peterson. 2001. Scipy: Open source scientific tools for python.

(5) Hunter, J. 2007. Matplotlib: a 2d graphics environment. *Computing in Science & Engineering* pages 90–95.

(6) Lavery, R., K. Zakrzewska, D. Beveridge, T. C. Bishop, D. A. Case, T. r. Cheatham, S. Dixit, B. Jayaram, F. Lankas, C. Laughton, J. H. Maddocks, A. Michon, R. Osman, M. Orozco, A. Perez, T. Singh, N. Spackova, and J. Sponer. 2010. A systematic molecular dynamics study of nearest-neighbor effects on base pair and base pair step conformations and fluctuations in B-DNA. *Nucleic Acids Res.* 38(1):299–313.

(7) Flyvbjerg, H. and H. Petersen. 1989. Error estimates on averages of correlated data. *J. Chem. Phys.* 91(1):461–466.

(8) Press, W., S. Teukolsky, W. Vetterling, and B. Flannery. 2007. *Numerical Recipes 3rd Edition: The Art of Scientific Computing*. Cambridge University Press. ISBN 0521880688.

(9) Olson, W. K., A. A. Gorin, X. J. Lu, L. M. Hock, and V. B. Zhurkin. 1998. DNA sequence-dependent deformability deduced from protein-DNA crystal complexes. *Proc. Natl. Acad. Sci. USA* 95(19):11163–11168.

(10) Lankas, F., J. Sponer, J. Langowski, and T. E. r. Cheatham. 2003. DNA basepair step deformability inferred from molecular dynamics simulations. *Biophys. J.* 85(5):2872–2883.

(11) Benham, C. J. 1996. Theoretical analysis of the helix-coil transition in positively superhelical DNA at high temperatures. *Phys. Rev. E* 53(3):2984.