

Sequence-dependent mechanical properties of the RNA double-helix

Alberto Marin-Gonzalez, J.G. Vilhena, Fernando Moreno-Herrero, Rubén Pérez

Supplementary Information

1. Mechanical parameters of dsRNA obtained from the elastic rod model
2. Stiffness analysis of dsRNA dinucleotide flexibility
3. Supplementary Figures

1. Mechanical parameters of dsRNA obtained from the elastic rod model

According to the elastic rod model, if bending fluctuations are negligible, the energy of a dsDNA/dsRNA molecule under an external force, F , can be written as [1-4]

$$E = \frac{1}{2L}Sx^2 + \frac{1}{2L}C\theta^2 + \frac{g}{L}x\theta - xF \quad (1)$$

where L is the equilibrium extension in the absence of force; x and θ are the elongation and change in helical twist at a given force with respect to their value at zero force; and S , C and g are the stretch modulus, torsion modulus and twist-stretch coupling parameters, respectively. We have previously shown [5] that the elastic parameters can be obtained by measuring three observables: a) the elongation as a function of force, $x(F)$; b) the force-induced change in helical twist $\theta(F)$; and c) how the thermal fluctuations in extension and helical twist are correlated $\partial x/\partial \theta$. The elastic parameters can then be computed by means of the following set of equations:

$$S = \frac{1}{A_1 - A_2A_3} \quad (2)$$

$$C = \frac{\frac{A_1A_3}{A_2}}{A_1 - A_2A_3} \quad (3)$$

$$g = -\frac{A_3}{A_1 - A_2A_3} \quad (4)$$

where A_1 and A_2 are the slopes of a linear fit to $x(F)/L$ and $\theta(F)/L$ and A_3 is the helical rise-helical twist covariance. It is also convenient to compute the effective stretch modulus, which is simply given by

$$\tilde{S} = \frac{1}{A_1} \quad (5)$$

In order to compute the elastic parameters of the benchmark and random duplexes we measured the force-induced change in extension and helical twist using the 3DNA software [6] and taking the $F = 1\text{ pN}$ simulation as reference. These data are shown in Fig. 1, Fig. 4 of the main text and also in Fig. S2 a-d. We additionally measured the helical rise-helical twist covariance at different forces, see Fig. S2 e, f. The slopes from Fig. S2 a, b yield A_1 ; the slopes from Fig. S2 c, d are equal to A_2 ; and A_3 was taken to be the helical rise-helical twist covariance at 1 pN force (Fig. S2 e, f). The elastic rod parameters were then obtained using Eq. 2-5. The values of these elastic rod parameters for the benchmark and random sequences are shown in Table S1 and Table S2 respectively. We repeated this analysis using the alternative Curves+ software [7] and found very similar results, see Table S3, S4 and Fig. S3.

Parameter	Alternating R-Y			Poly-R		
	Poly-CG	Poly-AC	Poly-AU	Poly-A	Poly-AG	Poly-G
S (pN)	442 (17)	524 (13)	559 (25)	498 (16)	635 (11)	641 (14)
\tilde{S} (pN)	242 (4)	356 (4)	438 (14)	482 (15)	607 (9)	581 (11)
C (pN nm ²)	326 (19)	230 (14)	308 (23)	244 (24)	460 (40)	500 (40)
$g/k_B T$, no units	62 (4)	48 (3)	47 (4)	15 (1)	27 (2)	42 (3)

Table S1. Rod model elastic parameters obtained for the benchmark RNA duplexes (Table 1) using the 3DNA software [6]. These elastic parameters were computed from the slopes of Fig. S2 a, c and the helical rise-helical twist covariance at 1 pN (Fig. S2e). Errors were obtained from the uncertainties of the fit and using quadratic error propagation. The sequences studied were classified into two groups: alternating pyrimidine-purine and poly-purines.

Parameter	Seq-1	Seq-2	Seq-3	Seq-4	Computational Work [5]	Experimental Work [4, 8]
S (pN)	527 (14)	457 (25)	546 (12)	637 (23)	480 (11)	
\tilde{S} (pN)	427 (7)	352 (8)	411 (5)	543 (16)	416 (7)	350, 500
C (pN nm ²)	301 (23)	280 (50)	361 (18)	227 (12)	310 (24)	409
$g/k_B T$, no units	42 (3)	42 (7)	54 (3)	36 (2)	34 (1)	11.5

Table S2. Elastic parameters obtained for the random dsRNA sequences (Table 1) using the 3DNA software [6]. Similarly to the ones from Table S1, these parameters were obtained from the slopes of Fig. S2b, d and the helical rise-helical twist covariance at 1 pN of Fig. S2f. Errors were computed as described in Table S1. For comparison, we included the values of the elastic parameters reported in a previous computational work [5] for a 16 bp-RNA duplex containing all dinucleotides and the experimental measurements for kbp-long dsRNA molecules [4, 8].

Parameter	Alternating R-Y			Poly-R		
	Poly-CG	Poly-AC	Poly-AU	Poly-A	Poly-AG	Poly-G
S (pN)	413 (16)	450 (10)	496 (21)	445 (10)	562 (7)	593 (13)
\tilde{S} (pN)	224 (3)	316 (2)	389 (11)	438 (9)	541 (6)	534 (10)
C (pN nm ²)	358 (23)	248 (16)	332 (34)	109 (35)	310 (34)	496 (44)
$g/k_B T$, no units	63 (4)	44 (3)	46 (5)	6.8 (2.1)	20 (2)	42 (3)

Table S3. Elastic parameters of the benchmark RNA duplexes (Table 1) computed using the Curves+ software [7]. The elastic parameters were obtained as described in Table S1, using the data from Fig. S3, instead of Fig. S2.

Parameter	Seq-1	Seq-2	Seq-3	Seq-4	Computational Work [5]	Experimental Work [4, 8]
S (pN)	457 (10)	439 (27)	503 (10)	567 (29)	429 (11)	
\tilde{S} (pN)	376 (5)	335 (11)	375 (3)	487 (21)	382 (7)	350, 500
C (pN nm ²)	304 (26)	339 (49)	381 (19)	206 (14)	298 (24)	409
$g/k_B T$, no units	38 (3)	46 (6)	54 (3)	31 (2)	29 (1)	11.5

Table S4. Elastic rod parameters of the random dsRNA sequences (Table 1) computed using Curves+ [7]. The parameters were obtained as in Table S2, using the data from Fig. S3. Under “Computational Work” we included the values of the elastic parameters that were calculated in that work using Curves+. The values of “Experimental Work” are the same as in Table S2.

2. Stiffness analysis of dsRNA dinucleotides

Force constants were computed for each dinucleotide step in the context of their base pair step parameters shift, slide, rise, tilt, roll and twist. We assumed a mechanical model in which the energy required to drive a dinucleotide step away from its equilibrium conformation is harmonic [9, 10]. In this case, the deformation energy can be written as

$$E = \frac{1}{2} \sum_{i=1}^6 \sum_{j=1}^6 k_{ij} \Delta q_i \Delta q_j \quad (6)$$

Where the subindices i, j refer to each of the six base pair step parameters; $\Delta q_i = (q_i - q_i^0)$ is the deviation of the i^{th} parameter from its equilibrium value; and k_{ij} are the elements of the stiffness matrix K . That the elements of K can be computed by inversion of the covariance matrix $(C)_{ij} = \langle \Delta q_i \Delta q_j \rangle$ as [10-12]

$$K = k_B T C^{-1} \quad (7)$$

Where k_B is the Boltzmann constant and T the temperature of the system, which in our case is 300K. Using Eq. (7) we computed the elements of the stiffness matrix for each base pair step in all the molecules and then averaged over the same base pair step kinds, *e.g.* all the AU steps. We also computed the standard deviation of the k_{ii} 's which reflects the variation among base pair steps of the same kind located in different positions along the molecule. The diagonal matrix elements were separately analyzed for the benchmark and random sequences, see Fig. S4. The similar results obtained from the benchmark and random sequences support the idea presented in the main text that the dinucleotide flexibility of dsRNA is approximately independent of the global sequence context. The complete stiffness matrix computed for the benchmark sequences is presented in Table S5.

Finally, the covariance matrix allows obtaining the conformational volume accessible to each base pair step. This quantity provides an estimate of the general flexibility of the dinucleotide and can be computed according to [12]

$$V = \det(C)^{1/2} \quad (8)$$

We compared the dinucleotide conformational volumes from the benchmark and random dsRNA sequences and found similar results, see Fig S5. We additionally computed the dinucleotide conformational volume of the benchmark DNA sequences from a previous work [13] and found that sequence variations of this parameter follow very different trends for dsRNA and dsDNA, see Fig. S1.

	CG	CA	UA	AA	AG	GA	GG	AC	AU	GC
Shift-shift	1.0602	1.3257	2.0498	1.4093	1.4752	1.7265	2.7347	1.1560	1.0025	2.0159
Slide-Slide	3.9006	4.1788	4.5051	3.6316	4.7133	4.5612	5.9300	2.8563	3.4989	3.5433
Rise-Rise	3.7337	3.8541	4.4180	8.0142	8.3721	9.4769	10.071	11.470	11.132	12.904
Tilt-tilt	0.0323	0.0270	0.0216	0.0310	0.0358	0.0436	0.0532	0.0317	0.0261	0.0391
Roll-roll	0.0134	0.0135	0.0131	0.0171	0.0198	0.0200	0.0250	0.0223	0.0194	0.0290
Twist-twist	0.0546	0.0525	0.0555	0.0546	0.0588	0.0603	0.0680	0.0524	0.0591	0.0593
Shift-Slide	-0.0063	0.3048	-0.0152	-0.2234	-0.2321	-0.4754	-0.5012	0.0157	0.0111	-0.0041
Shift-Rise	-0.0077	-0.0018	0.0092	0.2504	0.1733	0.6909	0.5153	-0.1884	-0.0182	0.0042
Shift-Tilt	-0.0561	-0.0079	0.0371	0.0106	-0.0250	-0.0233	-0.0732	-0.0054	0.0085	-0.0176
Shift-Roll	0.0008	0.0219	-0.0011	0.0081	-0.0072	0.0189	0.0038	-0.0067	0.0006	0.0004
Shift-Twist	-0.0021	-0.0071	-0.0015	0.0176	0.0359	0.0203	0.0427	-0.0023	0.0000	-0.0009
Slide-Rise	1.4427	0.8043	0.0693	0.4518	0.0639	0.9683	0.9845	2.1302	2.8590	1.8321
Slide-Tilt	0.0019	-0.0322	0.0002	0.0511	0.0361	0.0309	0.0421	-0.0343	0.0022	0.0001
Slide-Roll	-0.0218	-0.0422	-0.0174	0.0129	0.0143	0.0405	0.0201	-0.0313	-0.0502	-0.0130
Slide-Twist	-0.2169	-0.1399	-0.0467	-0.1377	-0.1999	-0.2072	-0.2176	-0.1816	-0.1491	-0.2434
Rise-Tilt	-0.0008	-0.0178	0.0001	-0.2370	-0.2399	-0.3215	-0.3426	0.0156	-0.0030	-0.0020
Rise-Roll	0.0076	-0.0475	-0.0965	-0.1584	-0.0856	-0.0939	-0.0202	-0.0728	-0.1472	0.0041
Rise-Twist	-0.0767	-0.0721	-0.0931	-0.1484	-0.1564	-0.1418	-0.1352	-0.1771	-0.1507	-0.2021
Tilt-Roll	0.0000	0.0011	-0.0001	0.0009	-0.0005	-0.0020	-0.0032	0.0028	-0.0001	-0.0000
Tilt-Twist	-0.0001	0.0006	-0.0003	0.0015	0.0013	-0.0005	-0.0038	0.0013	0.0002	0.0000
Roll-Twist	-0.0071	-0.0018	0.0008	0.0028	-0.0010	0.0022	-0.0020	0.0018	0.0026	0.0017

Table S5. Complete dinucleotide stiffness matrix obtained for the benchmark sequences. The stiffness matrix was obtained by inversion of the covariance matrices and subsequent averaging over the steps of the same kind. Diagonal elements are the ones shown in Figure S5. Translation-translation force constants are in kcal/(Å² mol), rotation-rotation ones in kcal/(deg² mol) and translation-rotation in kcal/(deg Å mol).

3. Supplementary Figures

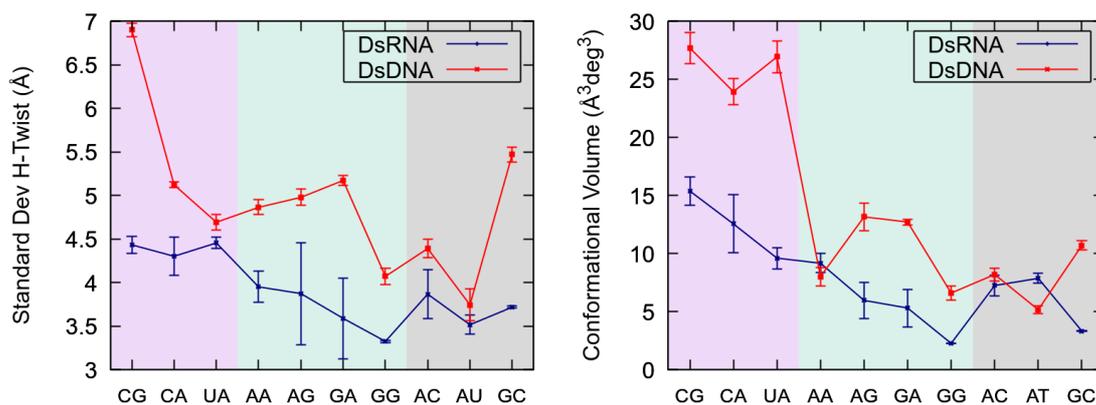


Figure S1. Helical twist fluctuations and conformational volumes of dsDNA and dsRNA dinucleotides. The helical twist standard deviation and the conformational volume (Eq 8) was computed for all the base pair steps of the benchmark dsDNA and dsRNA sequences. We then averaged over the base pair steps of the same kind, i.e. all the CG steps. Error bars are the standard error of this average. The shaded regions indicate the different dinucleotide families: pyrimidine-purine (pink), purine-purine (green) and purine-pyrimidine (gray). A line connecting the points was included to guide the eye.

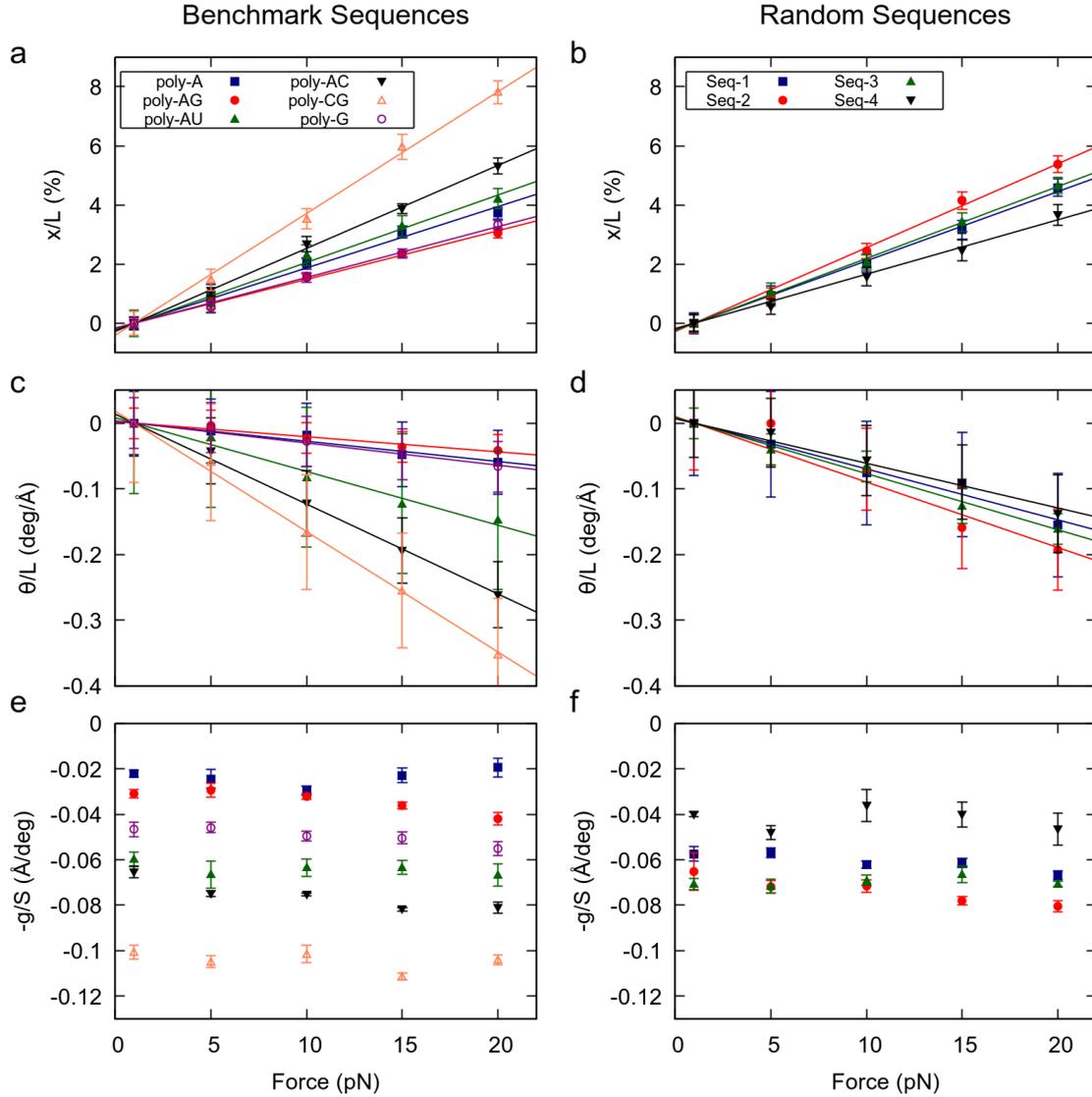


Figure S2. Measurements of the observables required for determining the elastic rod parameters of the RNA duplexes using the 3DNA software [7]. (a) The relative change in extension with respect to the value at 1 pN is plotted as a function of the applied force for the benchmark dsRNA sequences. The inverse of the slopes of the linear fits to the data yield \tilde{S} , see Table 1. (b) Force-induced elongation measured for the random sequences. Data analysis and representation was done as described in (a). (c) The change in helical twist with respect to the value at 1 pN force divided by the extension at this force is plotted as a function of the external force for the benchmark sequences. (d) Same as c, for the random sequences. (e) Extension-helical twist covariance computed at each constant force as done in [5, 14]. (f) Same as e, for the benchmark sequences. The fits of top and middle panels were constrained to go through the point (1,0) as done in [5]. Color legend is the same for both columns of panels, i.e. for a, c, e and b, d, f panels. All the errors were computed by splitting the data into five windows of 200ns and calculating the standard error of the mean of these five measurements.

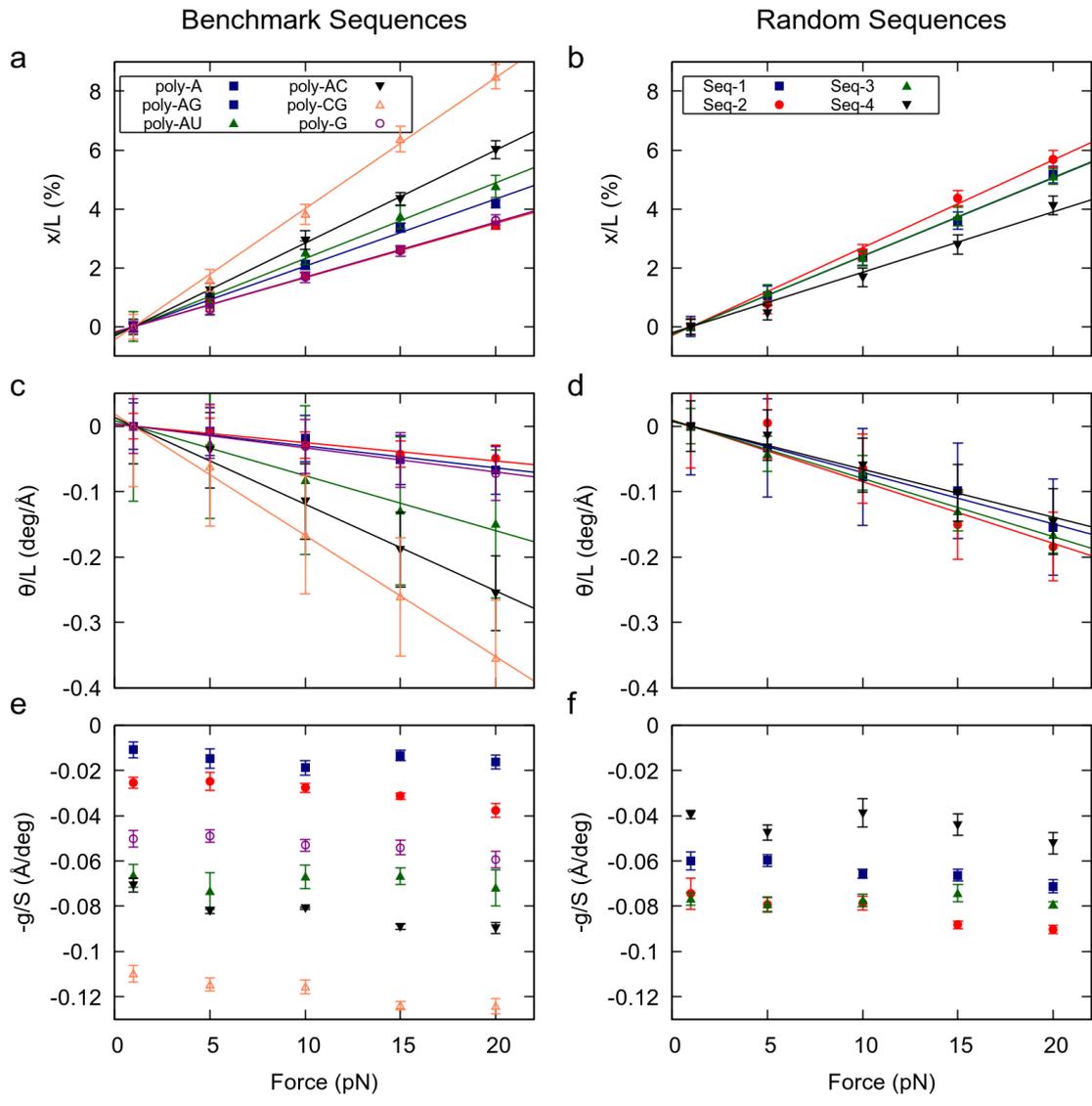


Figure S3. Mechanical properties of the simulated dsRNA sequences were measured using the software Curves+ [7]. All figure details regarding data analysis and representation are similar to Fig S2.

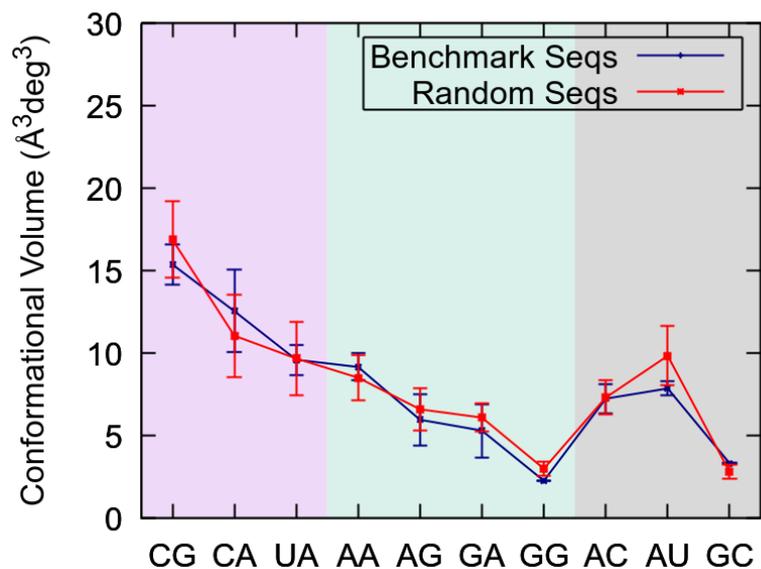


Figure S4. Conformational volume of the ten base pair step kinds. The conformational volumes were computed for all the base pair steps using Eq. (8). These values were then averaged for each base pair step kind. The error bars are the standard deviation of the conformational volumes of all the analyzed base pairs of the same kind, i.e. the standard deviation of all CG base pairs. For comparison, this analysis was performed for the benchmark and random sequences separately. The shaded regions and the line connecting the points were drawn as indicated in Fig S1.

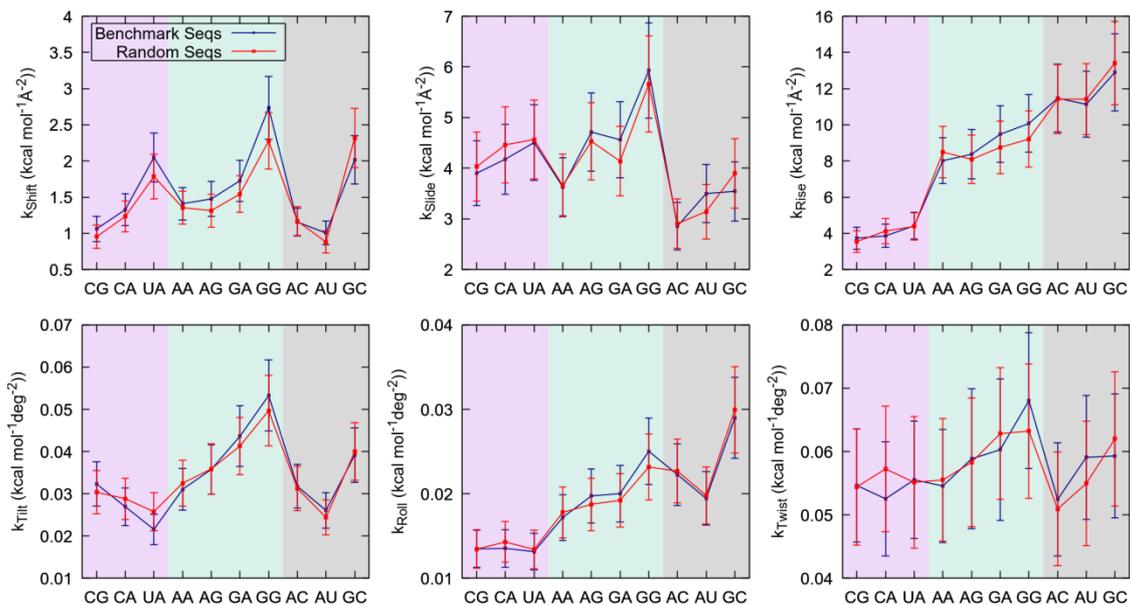


Figure S5. Diagonal elements of the stiffness matrix of all base pair step kinds. The stiffness matrix was computed for all base pair steps by inversion of the covariance matrix following the procedure from [10]. We then computed the mean force constants by averaging the matrices corresponding to all the steps of the same kind. The error bars are the standard deviation of the different steps of the same kind, i.e. all the CG dinucleotides. The analysis was done separately for the benchmark and random for comparison purposes. Here are shown the diagonal terms, the off-diagonal terms can be found in Table S5. The shaded regions delimit the different dinucleotide families and a connecting line was added in a similar fashion to FigS1.

References

1. Marko, J.F., *Stretching must twist DNA*. EPL (Europhysics Letters), 1997. **38**(3): p. 183.
2. Gross, P., et al., *Quantifying how DNA stretches, melts and changes twist under tension*. Nature Physics, 2011. **7**(9): p. 731-736.
3. Gore, J., et al., *DNA overwinds when stretched*. Nature, 2006. **442**(7104): p. 836-839.
4. Lipfert, J., et al., *Double-stranded RNA under force and torque: Similarities to and striking differences from double-stranded DNA*. Proceedings of the National Academy of Sciences, 2014. **111**(43): p. 15408-15413.
5. Marin-Gonzalez, A., et al., *Understanding the mechanical response of double-stranded DNA and RNA under constant stretching forces using all-atom molecular dynamics*. Proceedings of the National Academy of Sciences, 2017. **114**(27): p. 7049.
6. Lu, X.J. and W.K. Olson, *3DNA: a software package for the analysis, rebuilding and visualization of three-dimensional nucleic acid structures*. Nucleic Acids Research, 2003. **31**(17): p. 5108-5108.
7. Lavery, R., et al., *Conformational analysis of nucleic acids revisited: Curves+*. Nucleic Acids Res, 2009. **37**(17): p. 5917-5929.
8. Herrero-Galán, E., et al., *Mechanical Identities of RNA and DNA Double Helices Unveiled at the Single-Molecule Level*. Journal of the American Chemical Society, 2013. **135**(1): p. 122-131.
9. Olson, W.K., et al., *DNA sequence-dependent deformability deduced from protein-DNA crystal complexes*. Proceedings of the National Academy of Sciences, 1998. **95**(19): p. 11163-11168.
10. Lankaš, F., et al., *DNA Basepair Step Deformability Inferred from Molecular Dynamics Simulations*. Biophysical Journal, 2003. **85**(5): p. 2872-2883.
11. Noy, A., et al., *Relative Flexibility of DNA and RNA: a Molecular Dynamics Study*. Journal of Molecular Biology, 2004. **343**(3): p. 627-638.
12. Pérez, A., et al., *The relative flexibility of B-DNA and A-RNA duplexes: database analysis*. Nucleic Acids Research, 2004. **32**(20): p. 6144-6151.
13. Marin-Gonzalez, A., et al., *DNA Crookedness Regulates DNA Mechanical Properties at Short Length Scales*. Physical Review Letters, 2019. **122**(4): p. 048102.
14. Liebl, K., et al., *Explaining the striking difference in twist-stretch coupling between DNA and RNA: A comparative molecular dynamics analysis*. Nucleic Acids Research, 2015. **43**(21): p. 10143-10156.