DNA Crookedness Regulates DNA Mechanical Properties at Short Length Scales

Alberto Marin-Gonzalez,[1,*] J. G. Vilhena,[2,3,*] Fernando Moreno-Herrero,[1,†] and Ruben Perez[2,4,‡]
[1]Department of Macromolecular Structures, Centro Nacional de Biotecnología, Consejo Superior de Investigaciones Científicas, 28049 Cantoblanco, Madrid, Spain
[2]Departamento de Física Teórica de la Materia Condensada, Universidad Autónoma de Madrid, E-28049 Madrid, Spain
[3]Department of Physics, University of Basel, Klingelbergstrasse 82, CH 4056 Basel, Switzerland
[4]Condensed Matter Physics Center (IFIMAC), Universidad Autónoma de Madrid, E-28049 Madrid, Spain

(Received 28 May 2018; published 1 February 2019)

Sequence-dependent DNA conformation and flexibility play a fundamental role in the specificity of DNA-protein interactions. Here we quantify the DNA crookedness: a sequence-dependent deformation of DNA that consists of periodic bends of the base pair centers chain. Using extensive 100 μs-long, all-atom molecular dynamics simulations, we found that DNA crookedness and its associated flexibility are bijective, which unveils a one-to-one relation between DNA structure and dynamics. This allowed us to build a predictive model to compute the stretch moduli of different DNA sequences from solely their structure. Sequences with very little crookedness show extremely high stretching stiffness and have been previously shown to form unstable nucleosomes and promote gene expression. Interestingly, the crookedness can be tailored by epigenetic modifications, known to affect gene expression. Our results rationalize the idea that the DNA sequence is not only a chemical code, but also a physical one that allows finely regulating its mechanical properties and, possibly, its 3D arrangement inside the cell.

DOI: 10.1103/PhysRevLett.122.048102

The mechanism by which proteins interact with the genome with such extraordinary specificity is still an open question in biology. Since the discovery of the DNA double helix (dsDNA), it became clear that a sequence-dependent set of hydrogen bond donors and acceptors are exposed in the major groove and are specifically recognized by certain amino acids. However, there is increasing evidence that this mechanism is far from sufficient. In a number of DNA-protein complexes, DNA adopts a conformation that substantially deviates from the canonical B form [1–3], suggesting a structural deformation or an exceptional flexibility intrinsic to the DNA sequence. Among the most-studied cases are sequence-dependent DNA deformations (such as A-like structures, kinked base pair steps, and A tracts) that play an important role in transcription regulation [4–7]. In parallel, the high sequence-dependent flexibility of DNA is used by several proteins to achieve binding specificity [1,8].

However, many aspects of DNA flexibility have so far remained elusive. Indeed, it is not fully understood how a relatively stiff molecule, with a persistence length of $P \sim 50$ nm, is able to wrap around a histone octamer of $\sim$4 nm of radius. Even more intriguing is the fact that some sequences are hardly able to form stable nucleosomes, arguably as a consequence of a distinct conformation or mechanical properties [9,10]. The same question holds for other DNA-protein complexes, in particular, for many repressor systems where a highly bent loop is predicted in the DNA [11]. These considerations, supported by recent findings on high bendability of DNA at short-length scales [12,13], challenge the currently accepted worm-like chain model (WLC) and demand for an accurate description of sequence-dependent DNA flexibility at these scales.

Using constant-force molecular dynamics (MD) simulations [14], we observed that the extension of the DNA changed from one sequence to another for molecules with the same number of base pairs. We performed over 1 μs-long MD simulations of 18 base pair long DNA molecules with benchmark sequences of the form $CGCG(NN)_5CGCG$, where NN denotes AA, AC, AG, AT, CG, and GG, and computed its average structure at 1 pN force, see Fig. 1(a) and Fig. S1 [15]. The variability in the extension could not be attributed to a different separation between consecutive base pairs (Fig. S2 [15]) and, therefore, reflects an intrinsic curvature of the molecule. This curvature is apparent if we represent the centers of the base pairs [color beads in Fig. 1(a) and Fig. S1]. We will denote this curvature by crookedness, in analogy with a crooked road whose trajectory is not straight. Importantly, DNA crookedness is of the order of a few ($\sim$2) nanometers, a scale comparable with the histone octamer radius [39], the DNA curvature of the DNA-I-PpoI endonuclease complex [40], and several examples of sharply bent DNA found in regulatory regions [11]. This is illustrated in Fig. 1(b).

Additionally, it is useful to compare the crookedness with the curvature predicted by the WLC model. Notice that the WLC curvature is entropic (i.e., temperature dependent and absent as $T \to 0$), whereas the crookedness

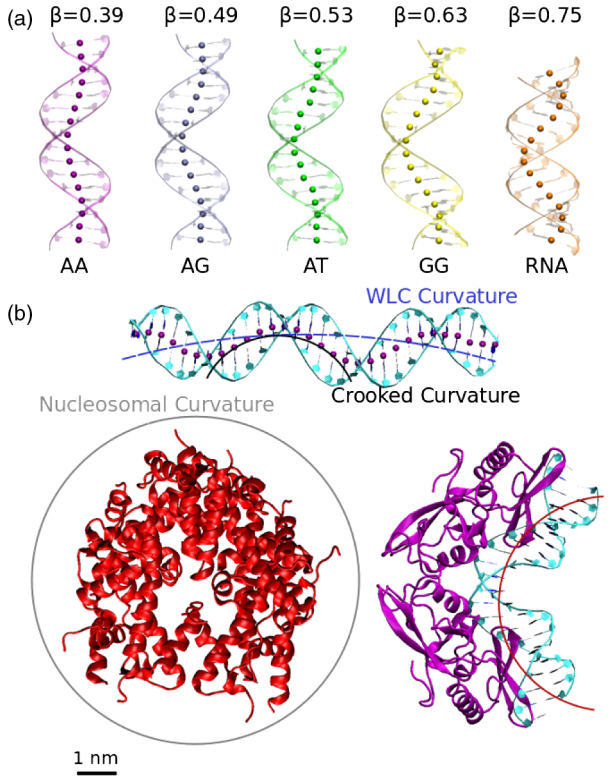

FIG. 1. Representation of DNA crookedness. (a) Average structures and computed β values (in rads) over 1 μs-long MD simulation at 1 pN force of the sequences $CGCG(NN)_5CGCG$ with NN = AA, AC, AT, and GG. The beads represent the centers of the base pairs. The terminal base pairs have been omitted. (b) Top: average structure over 250 ns MD of a 30 bps poly-G DNA molecule. The solid black line represents the crookedness curvature and the dashed blue line an estimation of the curvature predicted by the WLC. Bottom: examples of highly curved DNA when bound to proteins. (left) Histone octamer crystallized in Ref. [39] (PDB ID: 1AOI), where the histone tails have been removed for clarity. A grey circle of radius 41.8 Å represents the trajectory of nucleosomal DNA. (right) Crystal structure of the homing endonuclease I-PpoI DNA complex taken from Ref. [40] (PDB ID: 1A73), with an estimation of the DNA curvature represented by the solid red line.

is enthalpic and, thus, characteristic of the structure of the molecule and present at zero temperature. Figure 1 illustrates that, at length scales of the order of a few nm, the curvature predicted by the WLC (~16 nm of radius) is much smaller than the crookedness curvature intrinsic to the DNA molecule. This contrasts with the behavior expected for DNA molecules longer than the persistence length, for which entropic effects are no longer negligible and will eventually dominate over the crookedness.

Crookedness can be quantified via a parameter β defined as $\cos\beta \equiv x/\sum l_i$, with x being the extension (end-to-end distance) of the molecule and $\sum l_i$ the sum of distances between consecutive base pair centers. When the molecule is completely straight, the line that runs through the base pairs is perfectly aligned, $x = \sum l_i$, and the crookedness, β, is zero. As this line deviates more from the helical axis, the ratio $x/\sum l_i$ becomes smaller and therefore DNA crookedness increases. Double stranded RNA (dsRNA), which structure is A form, is the most crooked of the simulated molecules, whereas B-DNA molecules exhibit lower β values centered on 0.5 [Fig. 1(a)]. Indeed, as shown in Figs. S3 and S4, the crookedness is a reasonable parameter to distinguish DNA conformations along the A $\leftrightarrow$ B form spectrum and is closely related with major groove dimensions. Moreover, the β for a poly-G sequence of 30 bps ($\beta = 0.620$ rad) was very similar to the one found the 18 bps molecule ($\beta = 0.626$ rad), suggesting that β values are likely to persist for longer DNA molecules. Crookedness was inspired by the idea of *springiness*, proposed by J. Lipfert and co-workers in Refs. [41,42] to account for differences in mechanical properties of dsDNA and dsRNA.

Additionally, to sequence-dependent conformations, proteins often exploit DNA flexibility [1,8]. Therefore, a complete comprehension of the biological relevance of DNA crookedness requires understanding its effect on DNA mechanical properties. We propose a model to rationalize the relation between β and DNA stretch modulus, S, where DNA can elongate by reducing its β or by separating consecutive base pairs [Figs. 2(a) and 2(b) and

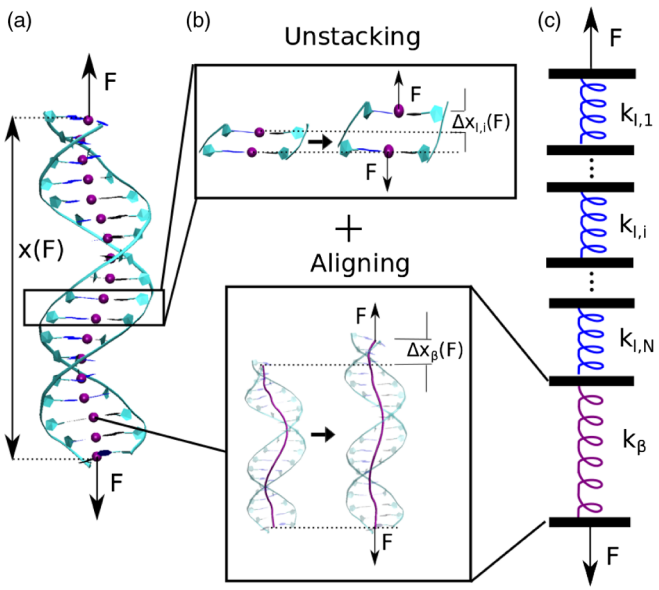

FIG. 2. A model to link DNA crookedness with DNA stretch modulus, S. (a) DNA molecule with the base pair centers represented by purple beads. An external force, F, induces a change in DNA extension, $x(F)$. (b) DNA can elongate by separating consecutive base pairs (top; $\Delta x_{l,i}$) or by aligning the base pair centers with the helical axis (bottom; Δx_β). (c) DNA is modelled as a set of N springs in series. The first N-1 springs account for the stiffness of elongating individual base pair steps, $k_{l,i}$, and the Nth spring is the crookedness stiffness, k_β.

Supplemental Material [15]]. Thus S can be written for an N-bps molecule as a set of N springs in series:

$$S^{-1} = \sum_{i=1}^{N-1} k_{l,i}^{-1} + k_\beta^{-1}. \quad (1)$$

The first N-1 springs correspond to the stiffness of separating consecutive base pairs, denoted by $k_{l,i}$, and the Nth spring to the crookedness stiffness, denoted by k_β [Fig. 2(c) and Fig. S4 [15]].

We ran five constant-force MD simulations at $F = 1$, 5, 10, 15, and 20 pN using the benchmark sequences described above to determine the sequence dependence of $k_{l,i}$ and k_β. We found that consecutive base pair elongation [$\sum k_{l,i}^{-1}$, Fig. 2(c), blue springs] has a minor contribution to S (Figs. S5, S6 [15]) and the dominant contribution comes from k_β, which accounts for a global deformation of the molecule as a whole [Fig. 2(c), purple spring]. We computed k_β by measuring the force induced change in DNA crookedness, $\Delta \cos\beta(F)$, according to this (see Supplemental Material [15]):

$$k_\beta \equiv \frac{F}{\Delta \cos\beta(F)/\cos\beta(0)}. \quad (2)$$

Figure 3(a) shows the calculated values (blue points) and the extraordinary fit provided by the function $k_\beta(\beta) = Ae^{-k\beta} + B$ with parameters $A = (2.24 \pm 1.24) \times 10^6$ pN, $B = 700 \pm 120$ pN, and $k = 16.2 \pm 1.5$. Remarkably, this phenomenological function reproduces the results for 11 additional sequences exhibiting a broad range of β values (Table S2 [15]) [43]. Importantly, this finding is consistent when defining β at 10 pN, see Fig. S5. This reveals an exceptional property of DNA crookedness: the relation with its associated stiffness is bijective. In other words, an *equilibrium* structural parameter, β, univocally determines a *dynamical* response, k_β, and vice versa.

The one-to-one correspondence between β and k_β provides predictive power to our model. Indeed, if the equilibrium structure of a DNA sequence is known (by NMR, crystallography, MD, etc.) one can measure β and the base pair separation distances. Then, the k_β obtained from Fig. 3(a) and the $k_{l,i}$, taken from Table S1 [15], can be used to compute the value of S with Eq. (1). Figure 3(b) confirms the good agreement with the measured values (our calculations and Refs. [50,51]) for a wide range of S values (~800–3000 pN).

In addition to predicting the stretch modulus, our model provides valuable information about the nature of its sequence dependence. Our results show that S is not dependent on the GC content, but, rather, on how guanines and cytosines are distributed along the strands. Molecules with alternating GCs are much less prone to stretching than those where several guanines are placed sequentially on the same strand. This adds to other examples where mechanical stability is independent of thermal stability [52]. Moreover, our model identified three molecules with anomalously enhanced stretching flexibility [Fig. 3(a), void green triangles]. These molecules include two strings of three

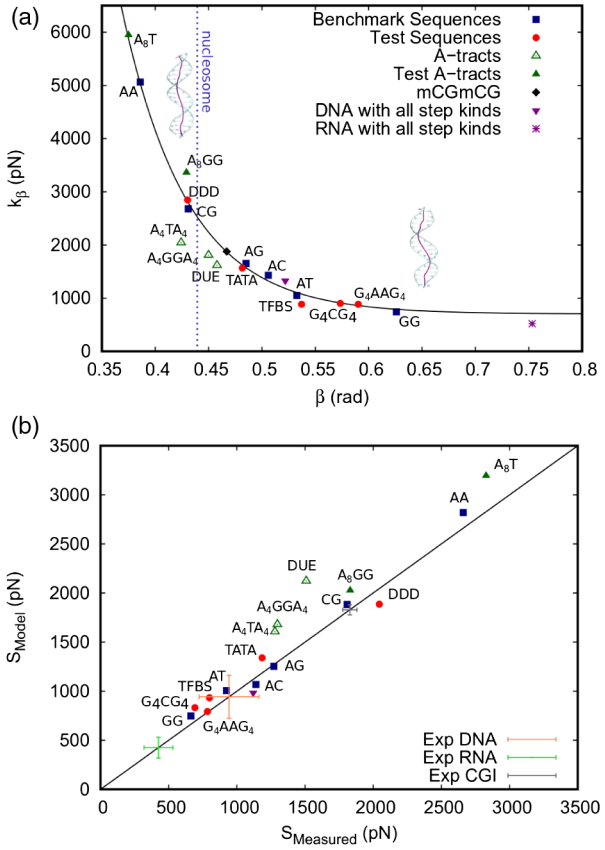

FIG. 3. Implications of DNA crookedness on DNA stretching flexibility. (a) k_β values as a function of the crookedness, β. k_β was computed from the MD simulations data using Eq. (1) and taking the $F = 1$ pN simulation as reference. β was computed from the $F = 1$ pN simulation as $\beta = \cos^{-1}(x/\sum l_i)$. The blue squares correspond to the benchmark sequences and are the points used for the fit $k_\beta(\beta) = Ae^{-k\beta} + B$ (fitting parameters in main text). The red circles represent the sequences simulated to test the model. The void green triangles are the A tracts containing sequences and the filled green triangles are the sequences used to test the effect of the A tracts. See Table S3 [15] for a list of the simulated sequences. The dashed line represents an estimation of nucleosome curvature according to [44]. (b) The computed value of S is plotted as a function of the value directly measured from the force-extension curve as done in Ref. [14]. The stretch modulus was computed from our model for the simulated sequences described above using Eq. (1), the value of the fit of $k_\beta(\beta)$, Table S1 [15], and the values of β and $\sum l_i$ obtained from the $F = 1$ pN simulation. The experimental value of the stretch modulus for random DNA and dsRNA sequences are the average value of the measurements from Refs. [42,45–48] and [42,45], respectively; and the error bars are the standard deviation of these values. The experimental value of the CpG island is the one from Ref. [49] with the error reported there.

or four A tracts, which are known to introduce anomalies in DNA curvature and mechanical properties [53,54]. The possibility that this enhanced stretching flexibility stems from the presence of two A tracts in A_4TA_4 and A_4GGA_4 molecules was confirmed by running additional simulations using the sequences A_8 T and A_8GG [Fig. 3(a), filled green triangles]. This highlights that other sources of flexibility may coexist with the crookedness mechanism proposed here.

Our benchmark simulations allowed us to identify sequence patterns with unusually high stretching stiffness, namely poly-A and alternating CG. They appear repeatedly throughout the genome and are frequently involved in gene expression regulation [9,10]. Long poly (dA:dT) and their flanking DNA have been shown to be depleted of nucleosomes *in vitro*, suggesting a mechanism of gene activation [9]. In parallel, about 70% of annotated gene promoters are associated with so-called CpG islands (CGI), rich in CpG steps [10]. Notice that our CG-alternating sequence shows the largest proportion of CpG steps possible in any molecule: one in every two steps. As in the case of the poly-A, CGI have been attributed to nucleosome destabilization [55]. These findings together with Fig. 3(a) suggest a possible relation between an unusually high crookedness stiffness and nucleosome destabilization. Hypermethylation of CGI commonly induces gene silencing, and in some cases this has been attributed to nucleosome stabilization [10]. We found that complete hypermethylation of the poly-CG molecule significantly increases its crookedness and stretching flexibility, in quantitative agreement (Table S3 [15]) with recent optical tweezers experiments reporting both an unusually high S and a softening induced by hypermethylation in CGI [49]. This could increase nucleosome affinity for a hypermethylated CGI [56]. Notice, however, that, outside this context, methylation is known to reduce DNA flexibility and destabilize nucleosomes [57–59].

We have shown that DNA crookedness can be tuned by specific sequences, being the main responsible for a sequence-dependent stretching flexibility. In addition, the crookedness mechanism, which regulates DNA enthalpic bending at short scales (10 bps), might extend to longer length scales. This would be similar to the case of A-tract induced bending. Although a single A tract shows a small kink at the 10 bps scale, placing several A tracts in phase with the helical pitch results in highly bent DNA molecules longer than 100 bps [53]. Similarly, one could think that placing alternating high and low crookedness motifs, such as poly-Gs and poly-As, would result in highly bent molecules. Supporting this idea, G-rich and A-rich motifs, with a periodic repeat of 10 bps and in antiphase with each other, are frequently found in nucleosome positioning sequences [60–62]. This crookedness regulation of DNA curvature is illustrated in Fig. 4(a). This mechanism would

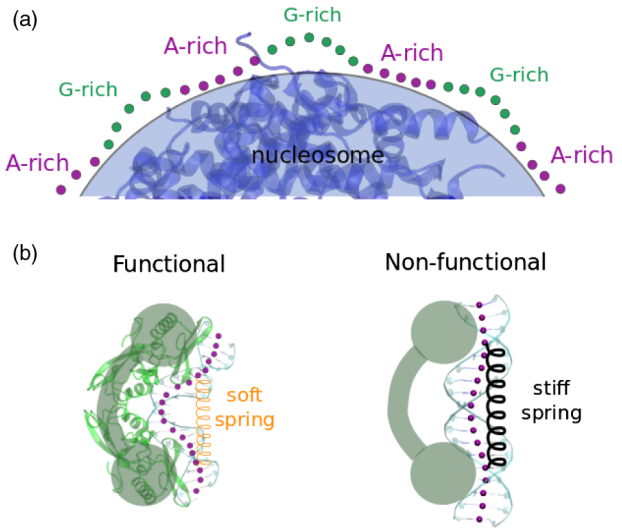

FIG. 4. Implications of DNA crookedness on DNA-protein interactions. (a) The base pair center chains of a DNA molecule are represented by color beads. We propose that net directional bending could be achieved by alternating high and low crookedness motifs, such as G-rich and A-rich sequences. (b) Additionally, crookedness flexibility could be exploited to induce a conformational change in DNA towards an A-like structure. On the left is the crystal structure of the I-PpoI DNA complex [40] (PDB ID: 1A73), the same as in Fig. 1(b). When bound to this protein a high distortion is found on the DNA, which would be more favorable for highly crooked and flexible sequences (left) than for sequences for which this flexibility is hindered (right).

add to other DNA features such as nicks [13,63], mismatches [13], and kinks [64,65] that are known to facilitate DNA bending at short scales.

Additionally, highly crooked molecules might be preferred for DNA-ligand binding where DNA flexibility is required. Indeed, interaction of DNA with proteins and drugs commonly modifies the B-DNA to a more A-like form and this occurs in a sequence-dependent manner [3]. According to Fig. 3 molecules with high β would be more prone to form an A-type helix when bound to a protein both because their structure is already closer to the A form and because of their enhanced flexibility [Fig. 4(b)]. Moreover, our finding that the stretching flexibility of DNA relating to sequence may have a relevant role in homologous recombination, where DNA must be stretched for homologous pairing and repair [66,67]. Finally, crookedness may modulate the charge distribution along the duplex, known to be a mechanism for protein-DNA recognition by electrostatic interactions [7].

In this Letter, we have introduced the crookedness: a "hidden code" imprinted in the DNA sequence that modulates its local curvature at short length scales. Our extensive all-atom simulations allowed us to unveil a striking one-to-one correspondence between this structural parameter and its associated flexibility. Such remarkable

correspondence was exploited to build a discrete model to compute the stretch modulus of any DNA sequence given solely by its equilibrium structure. Altogether, we have described how the DNA sequence finely tunes the stretch modulus of the molecule via the crookedness. Our results rationalize the idea that DNA sequence regulates the local curvature and the mechanical properties of the double helix at the length scale relevant for biological function.

We thank J. Lipfert, G. J. L. Wuite, C. L. Pastrana, and A. Gil for fruitful discussions. We thank the financial support from the Spanish MINECO [Projects No. MAT2014-54484-P, FIS2014-58328-P, MDM-2014-0377, MAT2017-83273-R (AEI/FEDER, UE), and BFU2017-83794-P (AEI/FEDER, UE)]. F. M.-H. acknowledges support from European Research Council (ERC) under the European Union Horizon 2020 research and innovation (Grant Agreement No. 681299). A. M.-G. acknowledges support from the International PhD Program of "La Caixa-Severo Ochoa" as a recipient of a PhD fellowship. J. G. V. acknowledges funding from a Marie Sklodowska-Curie Fellowship within the Horizons 2020 framework (DLV-795286). The authors acknowledge the computer resources, technical expertise, and assistance provided by the Red Española de Supercomputación at the Minotauro Supercomputer (BSC, Barcelona).

*A. M-G. and J. G. V. contributed equally to this work.
†Corresponding author.
fernando.moreno@cnb.csic.es
‡Corresponding author.
ruben.perez@uam.es

[1] R. Rohs, X. Jin, S. M. West, R. Joshi, B. Honig, and R. S. Mann, Annu. Rev. Biochem. **79**, 233 (2010).
[2] W. K. Olson, A. A. Gorin, X.-J. Lu, L. M. Hock, and V. B. Zhurkin, Proc. Natl. Acad. Sci. U.S.A. **95**, 11163 (1998).
[3] X.-J. Lu, Z. Shakked, and W. K. Olson, J. Mol. Biol. **300**, 819 (2000).
[4] M. H. Werner, A. M. Gronenborn, and G. M. Clore, Science **271**, 778 (1996).
[5] T. E. Haran and U. Mohanty, Q. Rev. Biophys. **42**, 41 (2009).
[6] M. McCall, T. Brown, W. N. Hunter, and O. Kennard, Nature (London) **322**, 661 (1986).
[7] R. Rohs, S. M. West, A. Sosinsky, P. Liu, R. S. Mann, and B. Honig, Nature (London) **461**, 1248 (2009).
[8] S. Ahmad, H. Kono, M. J. Arauzo-Bravo, and A. Sarai, Nucleic Acids Res. **34**, W124 (2006).
[9] E. Segal and J. Widom, Curr. Opin. Struct. Biol. **19**, 65 (2009).
[10] A. M. Deaton and A. Bird, Genes Dev. **25**, 1010 (2011).
[11] H. G. Garcia, P. Grayson, L. Han, M. Inamdar, J. Kondev, P. C. Nelson, R. Phillips, J. Widom, and P. A. Wiggins, Biopolymers **85**, 115 (2007).
[12] P. A. Wiggins, T. van der Heijden, F. Moreno-Herrero, A. Spakowitz, R. Phillips, J. Widom, C. Dekker, and P. C. Nelson, Nat. Nanotechnol. **1**, 137 (2006).
[13] R. Vafabakhsh and T. Ha, Science **337**, 1097 (2012).
[14] A. Marín-González, J. Vilhena, R. Pérez, and F. Moreno-Herrero, Proc. Natl. Acad. Sci. U.S.A. **114**, 7049 (2017).
[15] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevLett.122.048102 for more details on the derivation of the model, tests on the approximations made, and the benchmark sequences used to build and test the model, which includes Refs. [16–38].
[16] X. Lu and W. K. Olson, Nucleic Acids Res. **31**, 5108 (2003).
[17] F. Lankaš, J. Šponer, J. Langowski, and T. E. Cheatham, Biophys. J. **85**, 2872 (2003).
[18] S. Fujii, H. Kono, S. Takenaka, N. Go, and A. Sarai, Nucleic Acids Res. **35**, 6063 (2007).
[19] W. K. Olson, A. V. Colasanti, Y. Li, W. Ge, G. Zheng, and V. B. Zhurkin, *Computational Studies of RNA and DNA*, edited by J. Sponer and F. Lankas (Springer, New York, 2006), Chap. 9, pp. 235–257.
[20] F. Kilchherr, C. Wachauf, B. Pelz, M. Rief, M. Zacharias, and H. Dietz, Science **353**, aaf5508 (2016).
[21] D. A. Case *et al.*, AMBER 14, University of California, San Francisco.
[22] R. Salomon-Ferrer, A. W. Götz, D. Poole, S. Le Grand, and R. C. Walker, J. Chem. Theory Comput. **9**, 3878 (2013).
[23] A. W. Götz, M. J. Williamson, D. Xu, D. Poole, S. Le Grand, and R. C. Walker, J. Chem. Theory Comput. **8**, 1542 (2012).
[24] S. L. Grand, A. W. Gtz, and R. C. Walker, Comput. Phys. Commun. **184**, 374 (2013).
[25] A. Perez, I. Marchn, D. Svozil, J. Sponer, T. E. Cheatham, C. A. Laughton, and M. Orozco, Biophys. J. **92**, 3817 (2007).
[26] M. Zgarbová, M. Otyepka, J. Sponer, A. Mládek, P. Baná, T. E. Cheatham, and P. Jurecka, J. Chem. Theory Comput. **7**, 2886 (2011).
[27] W. D. Cornell, P. Cieplak, C. I. Bayly, I. R. Gould, K. M. Merz, D. M. Ferguson, D. C. Spellmeyer, T. Fox, J. W. Caldwell, and P. A. Kollman, J. Am. Chem. Soc. **117**, 5179 (1995).
[28] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, and M. L. Klein, J. Chem. Phys. **79**, 926 (1983).
[29] I. S. Joung and T. E. Cheatham, J. Phys. Chem. B **113**, 13279 (2009).
[30] P. Li, B. P. Roberts, D. K. Chakravorty, and K. M. Merz, J. Chem. Theory Comput. **9**, 2733 (2013).
[31] F. Lankas, T. E. Cheatham, N. Spacková, P. Hobza, J. Langowski, and J. Sponer, Biophys. J. **82**, 2592 (2002).
[32] V. Bloomfield, P. Killman, D. Crothers, I. Tinoco, J. Hearst, D. Wemmer, and D. Turner, *Nucleic Acids: Structure, Properties, and Functions* (University Science Books, Sausalito, CA, 2000).
[33] J. T. Waters, X.-J. Lu, R. Galindo-Murillo, J. C. Gumbart, H. D. Kim, T. E. Cheatham, and S. C. Harvey, J. Phys. Chem. B **120**, 8449 (2016).
[34] R. Lavery, M. Moakher, J. H. Maddocks, D. Petkeviciute, and K. Zakrzewska, Nucleic Acids Res. **37**, 5917 (2009).
[35] R. Wing, H. Drew, T. Takano, C. Broka, S. Tanaka, K. Itakura, and R. E. Dickerson, Nature (London) **287**, 755 (1980).
[36] J. L. Kim, D. B. Nikolov, and S. K. Burley, Nature (London) **365**, 520 (1993).

[37] D. Bramhill and A. Kornberg, Cell **52**, 743 (1988).
[38] K. Liebl, T. Drsata, F. Lankas, J. Lipfert, and M. Zacharias, Nucleic Acids Res. **43**, 10143 (2015).
[39] K. Luger, A. W. Mader, R. K. Richmond, D. F. Sargent, and T. J. Richmond, Nature (London) **389**, 251 (1997).
[40] K. E. Flick, M. S. Jurica, R. J. Monnat, Jr., and B. L. Stoddard, Nature (London) **394**, 96 (1998).
[41] F. C. Chou, J. Lipfert, and R. Das, PLoS Comput. Biol. **10**, e1003756 (2014).
[42] J. Lipfert, G. M. Skinner, J. M. Keegstra, T. Hensgens, T. Jager, D. Dulin, M. Köber, Z. Yu, S. P. Donkers, F.-c. Chou, R. Das, and N. H. Dekker, Proc. Natl. Acad. Sci. U.S.A. **111**, 15408 (2014).
[43] Our choice for the fitting function is purely phenomenological. Other functions, such as $A\beta^{-n} + B$ with $n = 6, 7$, also fit to the data. We chose the function $Ae^{-k\beta} + B$ because it is smooth and reasonably easy to converge in conventional fitting programs.
[44] T. J. Richmond and C. A. Davey, Nature (London) **423**, 145 (2003).
[45] E. Herrero-Galn, M. E. Fuentes-Perez, C. Carrasco, J. M. Valpuesta, J. L. Carrascosa, F. Moreno-Herrero, and J. R. Arias-Gonzalez, J. Am. Chem. Soc. **135**, 122 (2013).
[46] C. G. Baumann, S. B. Smith, V. A. Bloomfield, and C. Bustamante, Proc. Natl. Acad. Sci. U.S.A. **94**, 6185 (1997).
[47] S. B. Smith, Y. Cui, and C. Bustamante, Science **271**, 795 (1996).
[48] J. R. Wenner, M. C. Williams, I. Rouzina, and V. A. Bloomfield, Biophys. J. **82**, 3160 (2002).
[49] C. I. Pongor, P. Bianco, G. Ferenczy, R. Kellermayer, and M. Kellermayer, Biophys. J. **112**, 512 (2017).
[50] F. Lankas, J. Sponer, P. Hobza, and J. Langowski, J. Mol. Biol. **299**, 695 (2000).
[51] F. Lankas, Biopolymers **73**, 327 (2004).
[52] R. Vlijm, v. d. Torre J, and C. Dekker, PLoS One **10**, e0141576 (2015).
[53] F. Moreno-Herrero, R. Seidel, S. M. Johnson, A. Fire, and N. H. Dekker, Nucleic Acids Res. **34**, 3057 (2006).
[54] S. H. Kim, M. Ganji, E. Kim, J. van der Torre, E. Abbondanzieri, and C. Dekker, eLife **7**, e36557 (2018).
[55] V. R. Ramirez-Carrozzi, D. Braas, D. M. Bhatt, C. S. Cheng, C. Hong, K. R. Doty, J. C. Black, A. Hoffmann, M. Carey, and S. T. Smale, Cell **138**, 114 (2009).
[56] S. B. Baylin, Nat. Clin. Pract. Oncol. **2**, S4 (2005).
[57] A. Perez, C. L. Castellazzi, F. Battistini, K. Collinet, O. Flores, O. Deniz, M. Ruiz, D. Torrents, R. Eritja, M. Soler-López, and M. Orozco, Biophys. J. **102**, 2140 (2012).
[58] G. Portella, F. Battistini, and M. Orozco, PLoS Comput. Biol. **9**, e1003354 (2013).
[59] T. T. M. Ngo, J. Yoo, Q. Dai, Q. Zhang, C. He, A. Aksimentiev, and T. Ha, Nat. Commun. **7**, 10813 (2016).
[60] K. Brogaard, L. Xi, J.-P. Wang, and J. Widom, Nature (London) **486**, 496 (2012).
[61] N. Kaplan, I. K. Moore, Y. Fondufe-Mittendorf, A. J. Gossett, D. Tillo, Y. Field, E. M. LeProust, T. R. Hughes, J. D. Lieb, J. Widom, and E. Segal, Nature (London) **458**, 362 (2009).
[62] E. Segal, Y. Fondufe-Mittendorf, L. Chen, A. Thåström, Y. Field, I. K. Moore, J.-P. Z. Wang, and J. Widom, Nature (London) **442**, 772 (2006).
[63] H. Qu, C.-Y. Tseng, Y. Wang, A. J. Levine, and G. Zocchi, Europhys. Lett. **90**, 18003 (2010).
[64] F. Lanka, R. Lavery, and J. H. Maddocks, Structure **14**, 1527 (2006).
[65] J. Curuksu, M. Zacharias, R. Lavery, and K. Zakrzewska, Nucleic Acids Res. **37**, 3766 (2009).
[66] J. C. Bell and S. C. Kowalczykowski, Trends Biochem. Sci. **41**, 491 (2016).
[67] I. De Vlaminck, M. J. van Loenhout, L. Zweifel, J. den Blanken, K. Hooning, S. Hage, J. Kerssemakers, and C. Dekker, Mol. Cell **46**, 616 (2012).