Supporting Information

Savin et al. 10.1073/pnas.1218677110

SI Text

Minimal Energy Conformation of the Polymer Chain Is Determined by Convex Properties of the Effective Deformation Energy Curve of Individual Monomer Site. Consider a linear polymer chain of $N \gg 1$ identical sites (monomeric units). Each site i is extended by Δl_i, with the corresponding deformation energy $E(\Delta l_i)$. When pulled by both ends, the equilibrium energy of the entire chain can be found by solving the constrained (conditional) minimization problem:

$$NE_{NU}(\Delta L) = \min_{\Delta l_1 + \ldots + \Delta l_n = N\Delta L} \{E(\Delta l_1) + \ldots + E(\Delta l_n)\},$$

where $N\Delta L$ is the total extension of the chain. Here, as in the text, ΔL denotes the mean per site extension of the chain. Using the method of Lagrange multipliers, the problem is converted to the unconditional minimization problem over $(N + 1)$ variables:

$$\min\{E(\Delta l_1) + \ldots + E(\Delta l_n) - \lambda(\Delta l_1 + \ldots + \Delta l_n - N\Delta L)\}.$$

Differentiating with respect to Δl_i yields $E'(\Delta l_i) - \lambda$ that should equal zero for each i. Derivative with respect to λ also equals zero; this condition yields the original constraint. Excluding λ from all of the equations, we obtain the following system of N equations:

$$E'(\Delta l_1) = E'(\Delta l_2) = \ldots = E'(\Delta l_N) \text{ and} \qquad \text{[S1]}$$

$$\Delta l_1 + \ldots + \Delta l_N = N\Delta L. \qquad \text{[S2]}$$

When the function $E(\Delta l)$ is convex, its derivative is a monotonically increasing function, and the condition 1 can be satisfied only when extensions of all of the sites are equal to each other: $\Delta l_a = \ldots = \Delta l_n = \Delta l = \Delta L$ (that is, when the chain is extended uniformly). In contrast, when the energy function $E(\Delta l)$ is nonconvex, it is possible for its derivative (tangent) to have the same value at two distinct points Δl_a and Δl_b (Fig. 1). In this case, some Δl_i (for simplicity, $i = 1, \ldots, pN$) are equal to Δl_b, whereas the rest of Δl_i ($i = pN + 1, \ldots, N$) are equal to Δl_a, $\Delta l_a < \Delta l_b$. If $N \gg 1$ (technically, if $N \to \infty$), one can always find such p, $0 < p < 1$, that the constraint $\Delta L = p\Delta l_b + (1-p)\Delta l_a$ is satisfied. The deformation energy in this case of nonuniform stretching is equal to $pNE(\Delta l_b) + (1-p)NE(\Delta l_a) = N(pE(\Delta l_b) + (1-p)E(\Delta l_a))$, which is a linear function of ΔL that connects points $(\Delta l_a, E(\Delta l_a)), (\Delta l_b, E(\Delta l_b))$. This linear function is the convex hull of $E(\Delta l)$ (Fig. 1). By definition of nonconvex function, $pE(\Delta l_b) + (1-p)E(\Delta l_a) < E(p\Delta l_b + (1-p)\Delta l_a) = E(\Delta L)$, which proves that the two-phase extension is energetically preferred relative to the uniform extension in this case. As the chain is stretched farther and ΔL increases, p increases accordingly so that $\Delta L = p\Delta l_b + (1-p)\Delta l_a$ is satisfied. The extension of the chain through the change in the fraction of sites extended by Δl_b can continue until $p = 1$ at $\Delta L = \Delta l_b$.

The above reasoning does not take into account phase boundary effects. However, these effects, as well as end effects, are negligible as long as the polymer chain is long, $N \gg 1$, which is always the case experimentally. Our numerical calculations presented in the text and below—in which all of the boundary effects are included—give the same results supporting our general conclusions.

In what follows, we consider three microscopic polymer models in detail: 2D zigzag, α-helix, and dsDNA. Because we are interested in the regime where polymer extension approaches its contour length, entropic contributions (key in weak stretching) are neglected.

2D Zigzag. Consider a dimensionless 2D model of zigzag chain shown in Fig. S1. Such a chain can be considered as a quasi-1D crystal with the elementary cell being two neighboring sites; that is, each cell can be obtained from the previous one by translation along the x axes. The potential energy of the zigzag chain is given by Eq. **1** in the text. The equilibrium bond length is $\rho_0 = 1$, and zigzag angle is $\phi_0 = \arccos(-1/3) = 109.47°$ (so that equilibrium longitudinal step is $l_0 = 0.8165$). The nonvalent interaction is given by $W(r) = \varepsilon_{hb}\left[\left(\frac{r_0-d}{r-d}\right)^6 - 1\right]^2$, where $\varepsilon_{hb} \geq 0$ is the interaction energy, $r_0 = 2h_x = 1.633$ is the equilibrium length, and $d = 0.5$ is the diameter of the inner hard core. For simplicity, but without loss of generality, we set the equilibrium length r_0 to correspond to the equilibrium value of the angle θ_0. The nonvalent interaction energy coefficient was fixed at $\varepsilon_{hb} = 0.0178$ [so that $W''(r_0) = 72\varepsilon_{hb}/(r_0-d)^2 = 1$].

Let the x axis be along the chain, and the y axis be in the perpendicular direction (Fig. S1). The longitudinal (x) step l and transverse (y) step h for site n are shown in Fig. S1. To find the uniform extension energy that is energy of a single zigzag site extended by $\Delta l > 0$ along the x axis, one has to solve the minimization problem over transverse step h:

$$E(l) = V(\rho) + \varepsilon_\phi U(\phi) + W(r) \to \min : h, \qquad \text{[S3]}$$

with fixed value of the longitudinal step $l_0 + \Delta l$. Note that the distance between next-nearest neighbors is $r = 2(l_0 + \Delta l)$, and the value of the angle ϕ is defined by the bond length ρ. Solving the minimization problem (1) yields the energy (per one site) of a uniformly stretched zigzag as a function of the relative longitudinal extension Δl. Conjugate gradient is used to find the minimum.

The angle ϕ, energy, and tension force as functions of the extension are shown in Figs. S2 and S3 for different values of the coefficient ε_ϕ. This coefficient controls the relative contribution of a valent interaction (angle deformation energy) to the elastic response of the chain.

To find the energy of the extended (generally nonuniform) chain for each value of the mean (per site) extension $\Delta L \geq 0$, one has to solve the minimization problem

$$\sum_{n=1}^{N-1} V(\rho_n) + \sum_{n=2}^{N-1} \varepsilon_\phi U(\phi_n) + \sum_{n=1}^{N-2} W(r_n) \to \min_{\{x_n y_n\}_{n=1}^N} \qquad \text{[S4]}$$

with the condition of fixed ends:

$$x_1 \equiv -\frac{1}{2}(N-1)\Delta L, \quad x_N = (N-1)l_0 + \frac{1}{2}(N-1)\Delta L.$$

Conjugate gradient is used to find the minimum. The initial condition is chosen to be that of a uniformly stretched chain. It should be mentioned that, in what follows, only the x coordinate is fixed, and all other coordinates are allowed to change freely. If $\{x_n, y_n\}_{n=1}^N$ is a solution of the minimization problem (2), then the distribution of the longitudinal extension in the chain is given by the function $\Delta l_n = x_{n+1} - x_n - l_0$, and the distribution of the transverse step is given by the function $h_n = |y_{n+1} - y_n|$.

One can see from Fig. S4 that, when $\varepsilon_\phi = 0$, the distribution of the longitudinal step extension Δl_n and the transverse step h_n

along the chain is in agreement with the prediction that can be made based on convexity of the function $E(\Delta l)$. That is, when $0 \leq \Delta l \leq 0.037$ and $\Delta l \geq 0.2056$, uniform stretching takes place, whereas when $0.037 < \Delta l < 0.2056$, nonuniform stretching is observed. Here, the end sites of the zigzag are in the strongly extended state with extension $\Delta l_n = 0.2056$, whereas the central part is in weakly stretched state with extension $\Delta l_n = 0.037$ (Fig. S4). Gradual transition from one state to the other is observed along the chain as the domain boundary. As the chain extends, the size of the weakly stretched part in the center monotonously decreases and vanishes at $\Delta l = 0.2056$. At even higher tension, the zigzag is stretched uniformly. The two-phase stretching scenario persists for as long as $0 \leq \varepsilon_\phi \leq 0.015$. The range of extensions where the scenario is realized gradually shrinks as the contribution of the valent bond potential grows. After $\varepsilon_\phi > 0.015$, all of the sites are stretched uniformly (Fig. S5). One can see from Fig. S2 that the strongly extended state corresponds to the completely stretched out zigzag (angle $\phi = 180^0$).

α-Helix. Consider a 3D molecular chain corresponding to an ideal (3) α-helix. The equilibrium atomic helix coordinates are

$$\mathbf{R}_n^0 = (R_0 \cos(n\varphi_0), R_0 \sin(n\varphi_0), nl_0), \quad \text{[S5]}$$

with $n = 0, \pm 1, \pm 2, \ldots$ being the atom number, R_0 being the helical radius, φ_0 and l_0 being the angular and longitudinal helix period. For the sake of simplicity, we consider dimensionless model of the helix (Fig. 4A), where the (dimensionless) helix radius is $R_0 = 0.4919$, angular step is $\varphi_0 = 100°$, and longitudinal step is $l_0 = 0.6572$ (4). Such a helix can be treated as a quasi-1D crystal; that is, each site can be obtained from the previous one by the appropriate translation along longitudinal axes and rotation around the same axis.

The chain potential energy is

$$H = \sum_n \{V(\rho_n) + \varepsilon_\phi U(\phi_n) + \varepsilon_\theta Z(\theta_n) + W(r_n)\}. \quad \text{[S6]}$$

The term $V(\rho_n)$ gives the energy of interaction between neighbor sites n and $(n + 1)$, where ρ_n is the distance between them. Bond rigidity is $K = 10$, and the equilibrium bond length is $\rho_0 = 1$. The angle deformation energy is described by $\varepsilon_\phi U(\phi_n)$, where ϕ_n is the angle between sites $(n - 1)$, n, and $(n + 1)$ (the vertex is on site n). Equilibrium angle is $\phi_0 = \arccos(-1/3) = 109.47°$, and coefficient is $\varepsilon_\phi = 1$. The specific form of these terms is described in *Methods*. The third term $\varepsilon_\theta Z(\theta_n)$ is the energy of the torsional deformation (rotation) around the nth bond:

$$\varepsilon_\theta Z(\theta) = \varepsilon_\theta (\cos\theta - \cos\theta_0)^2, \quad \text{[S7]}$$

where $\varepsilon_\theta \geq 0$ is the torsional rigidity; different values of ε_θ are considered while keeping other interactions fixed. The equilibrium torsion angle is $\theta_0 = \arccos(0.2395)$. The function $W(r_n)$ is the energy of the nth hydrogen bond connecting sites n and $(n + 3)$. It is given by the formula

$$W(r) = \varepsilon_{hb} \left[\left(\frac{r_0 - d}{r - d} \right)^6 - 1 \right]^2. \quad \text{[S8]}$$

The equilibrium hydrogen bond length is $r_0 = 2.0322$. Other parameters of the hydrogen bonding potential are $d = 0.7$ (inner core diameter) and $\varepsilon_{hb} = 0.0246$, so that rigidity of nonvalent interactions $W''(r_0) = 72\varepsilon_{hb}/(r_0 - d)^2 = 1$. Equilibrium values of angles and distances correspond to specified values of helix radius R_0, angle step φ_0, and longitudinal step l_0 of the helix in ground state.

To determine the dependence of the effective helix site energy E (per one step) on the relative uniform extension Δl, we solve the following conditional minimization problem

$$E(R, \varphi, l_0 + \Delta l) = V(\rho) + \varepsilon_\phi U(\phi) + \varepsilon_\theta Z(\theta) + W(r) \to \min_{R,\varphi} \quad \text{[S9]}$$

with the fixed value of longitudinal step $l_0 + \Delta l$. Here, R is the helix radius, and φ is its angular step (*Methods*). Conjugate gradient is used to find the minimum.

In general, to find variable extensions Δl_n of each helix site in the case of nonuniform stretching, we solve

$$\sum_{n=1}^{N-1} V(\rho_n) + \sum_{n=2}^{N-1} \varepsilon_\phi U(\phi_n) + \sum_{n=2}^{N-2} \varepsilon_\theta Z(\theta_n) + \sum_{n=1}^{N-3} W(r_n)$$
$$\to \min : \{(R_n, \varphi_n, l_0 + \Delta l_n)\}_{n=1}^N \quad \text{[S10]}$$

with the fixed ends condition:

$$l_1 \equiv -\frac{1}{2}(N-1)\Delta L, \quad l_N \equiv (N-1)l_0 + \frac{1}{2}(N-1)\Delta L.$$

Here, as before, $N\Delta L$ is the total extension of the chain of N sites. Conjugate gradient is used to find the minimum. The initial condition corresponds to the uniformly stretched helix.

dsDNA. *Model details.* The 12CG model of the DNA double helix used in this work is shown in Figs. S7 and S8. To provide additional information to what was presented in the text, we switch to atomic units.

The total potential energy of the system has the following form:

$$H = [E_v + E_b + E_a + E_t + E_{el} + E_{vdW}] + E_{hb}^* + E_{st}. \quad \text{[S11]}$$

The terms in the brackets describe the deformation energy of both strands. In the text, a shorthand notation was used to avoid unnecessary details. For example, the energy E_A of strand A, which appears in the text, equals the bracketed terms above in which only the untied atoms from strand A are retained. The terms E_v, E_a, and E_t correspond to valent bond, angle, and torsion deformation energy, respectively. These potentials have a common form; bond deformation energy is calculated as

$$U_{\alpha\beta}(r) = \frac{1}{2} K_{\alpha\beta} (r - R_{\alpha\beta})^2,$$

with the rigidity coefficients $K_{\alpha\beta}$ and equilibrium values $R_{\alpha\beta}$ being different for different grains. Angle deformation energy has the form

$$U_a(\phi) = \varepsilon_a (\cos\phi - \cos\phi_a)^2,$$

where values of the coefficient ε_a and the equilibrium angle differ for different angle types. Torsion deformation energy is described by the potential

$$U_t = \varepsilon_t (1 - \cos(\theta - \theta_0)),$$

where values of the coefficient ε_t and the equilibrium torsion angle θ_0 differ for different torsions. Rotational axis of the torsional potential is shown in Fig. S8. The second term E_b in the energy function (S11) describes deformation energy of a nitrogen base. The nitrogen base is a rather rigid chemical structure modeled here by rigid harmonic potentials, which keep four points near one plain: grain C1 and the three points on each nucleobase.

The next two terms E_{el} and E_{vdW} describe electrostatic and van der Waals interactions between backbone grains. Solvent is treated implicitly through the Generalized Born model (2, 6) (in the text). The methodology has been used to model free DNA in solution (7, 8), binding between proteins and nucleic acids (9–11), and conformational changes such as the $A \to B$ transition (5), as well as explore dynamics of long DNA fragments (12). The Generalized Born model approximates solvation energy of two interacting charges by the following formula originally proposed by Still et al. (2):

$$\Delta G_{solv} \approx -\frac{1}{2}\left(1 - \frac{1}{\varepsilon_{out}}\right)\sum_{ij}\frac{q_i q_j}{f\left(r_{ij}, R_i, R_j\right)}, \qquad \textbf{[S12]}$$

where ε_{out} is the dielectric constant of water, r_{ij} is the distance between atoms i and j, q_i is the partial charge of atom i, R_i is the so-called effective Born radius of atom i, and $f = [r_{ij}^2 + R_i R_j \exp(-r_{ij}^2/4R_iR_j)]^{\frac{1}{2}}$.

The empirical function f is designed to interpolate between the limits of large $r_{ij} \gg \sqrt{R_i R_j}$, where the Coulomb law applies, and the opposite limit, where the two atomic spheres fuse into one, restoring the famous Born formula for solvation energy of a single ion. The effective Born radius of an atom represents its degree of burial within the low dielectric interior of the molecule: the farther away the atom is from the solvent, the larger its effective radius is. In our model, we assume constant effective Born radii, which we calculate one time from the first principles (6) for the B-form DNA. The screening effects of monovalent salt are introduced approximately at the Debye–Huckel level by substitution:

$$1 - \varepsilon_{out}^{-1} \to 1 - \varepsilon_{out}^{-1}\ \exp(-0.73\kappa f).$$

The 0.73 prefactor was found empirically to give the best agreement with the numerical Poisson–Boltzmann treatment (13). Here, κ is the Debye–Huckel screening parameter $\kappa\left[\overset{\circ}{A}^{-1}\right] \approx 0.316\sqrt{[salt][mol/L]}$. Implementation details in the context of the 12CG DNA coarse-grained model can be found in ref. 1.

The last two terms E_{hb}^* and E_{st} in Eq. **S11** describe interactions between nitrogen bases (including stacking and hydrogen bonds). Because nitrogen base is a rather rigid structure, we can calculate coordinates of all of the original atoms (corresponding to the all-atom representation) from positions of the three united atoms. This useful property allows us to use, directly, the all-atom AMBER (14) Coulomb and van der Waals potentials used to mimic hydrogen bonds and stacking (1).

Assuming no sequence variability along the strand [e.g., poly (A)-poly(T)], such double helix can be considered as a quasi-1D crystal, with the elementary cell being one nucleotide pair of the double helix. In the ground (minimum energy) state, each successive nucleotide pair is obtained from its predecessor by translation along the z axis by step l followed by a rotation around the same axis through helical step $\Delta\phi$. (Here, we use l instead of $\Delta l = l - l_0$, because unlike in the case of a simple structure such as the 2D zigzag, the equilibrium value of the DNA longitudinal step l_0 in our model is not known a priori and is obtained by solving the corresponding minimization problem.) Thus, the energy of the ground state is a function of 38 variables: 36 Cartesian coordinates of 12 grains in the first nucleotide pair, $\Delta\phi$, and l.

To find the minimum energy (ground) state of the homogeneous (that is, no sequence variability along the strand) double helix under tension, we solve the following minimization problem over 37 variables,

$$H = (E_v + E_b + E_a + E_t + E_{el} + E_{vdW}) + E_{hb}^* + E_{st}$$
$$\to \min : \{\mathbf{x_j}\}_{j=1}^{12}, \Delta\phi, \qquad \textbf{[S13]}$$

under the fixed value of the longitudinal step l. The summation is taken over only 1 bp, and neighboring base pairs are obtained from it by rotation and translation. Conjugate gradient is used to find the minimum; the initial condition corresponds to all-atom B-form DNA. The minimization yields the effective site energy (per 1 bp) $E(l)$ as a function of the longitudinal step l (Fig. 6). The energy minimum is reached for the longitudinal step $l_0 = 3.352$ Å, which corresponds to the B-form of dsDNA.

Because the effective site energy is nonconvex, we predict two-phase stretching for dsDNA based on the general mechanism described in the text. Namely, a part of the double helix is in the weakly extended state with the longitudinal step l_a, while the rest of the base-pairs are in the strongly extended state with longitudinal step l_b (Fig. S9). The DNA structure and its model potentials are most complex among the three polymer models analyzed in this work. We have, therefore, chosen the DNA to further test our general predictions through molecular dynamics simulations at 300 K (for the 2D zigzag and the α-helix, only purely mechanical stretching without thermal fluctuations was considered; more information in the text).

Origins of the nonconvex shape of the effective site deformation for dsDNA. Variation of the base stacking and hydrogen bond components of the effective site deformation energy as a function of the chain extension for the double-helix DNA is shown in Fig. S11*A*. In this computation, thermal fluctuations are not considered. As the tension grows, the base stacking weakens, and the corresponding energy curve has a distinct nonconvex region. The hydrogen bonds also weaken but do not break; the nonconvex region on this curve is much less prominent.

Room temperature simulations of the DNA. To bring the temperature of the molecule to the desired value $T = 300$ K, we integrate over time the Langevin system of equations of motion:

$$\mathbf{M}_n\ddot{\mathbf{r}}_n = -\partial H/\partial r_n - \Gamma\mathbf{M}_n\dot{\mathbf{r}}_n + \Xi_n, \qquad \textbf{[S14]}$$

where the index $n = 1, 2, \ldots, N$ runs over all of the united atoms (grains) (Figs. S7 and S8). $\Gamma = 1/t_r$ is the Langevin collision frequency with $t_r = 1$ ps being the corresponding particle relaxation time, $\mathbf{M}_n$ is the mass of nth united atom, and $\Xi_n(t)$, $n = 1^N$ is a set of N 3D vectors of independent Gaussian distributed stochastic forces describing the interaction of nth united atom with the thermostat with correlation functions

$$\langle \Xi_n(t_1)\Xi_m(t_2)\rangle = 2M_n\Gamma k_B T\delta_{nm}\delta(t_2 - t_1).$$

The initial conditions correspond to the equilibrium state of the double helix.

After the system is thermalized, the temperature is maintained at $T = 300$ K, and the trajectory continues for the desired simulation time. We use Verlet integrator with the integration time step of 0.5 fs.

Dependence of the dsDNA site length and angular step (twist) is shown in Fig. S12. Twist was calculated using in-house software based on the algorithms described in ref. 15. The length was calculated as the distance between neighboring phosphorus atoms along the longitudinal axis (averaged over both strands).

Insensitivity of the plateau transition to the Watson–Crick bond strength. Within the framework of the 12CG coarse-grained model (1) used in the paper, the relative strength of the Watson–Crick (WC) hydrogen bonds is controlled by parameter c_0 in $E_{hb}^* = c_0 E_{hb}$ (Methods). In the main text, we used $c_0 = 0.4$, which gives an excellent agreement with the relevant single-molecule stretching experiment (16) (Fig. 9, main text).

Here, we vary c_0 to test the effect that the hydrogen bond strength may have on the overstretching plateau of dsDNA. The main conclusion is that doubling the strength of the WC bonds—to the point that they no longer break on stretching at 300 K—has little effect on the existence of the overstretching plateau in the force-extension diagram. Specifically, we have performed a 700-ps-long simulation of the same 500-bp poly(A)-poly(T) DNA fragment at 300 K but now, with an unphysically large value of $c_0 = 1.0$ intended to keep the WC bonds from breaking. Now, even in the stretched state (Fig. S13), the bonds do not break. However, the plateau in the force-extension diagram still exists (Fig. S14). Moreover, one can see from Fig. S14 that the value of the tension at the plateau and its range differ only slightly from the $c = 0.4$ case, where the bonds break as the chain is stretched, which is in perfect agreement

with the experiment (discussed in the text). Therefore, the existence and key characteristics of the plateau in the force-extension diagram for dsDNA are rather insensitive to hydrogen bond strength. The comparison of the DNA stretching behavior in these two parameter regimes—with regular and artificially strong WC bonds—is another confirmation that the DNA overstretching plateau does not arise from WC bond breaking: the plateau exists even when the hydrogen bonds remain unbroken.

Computational Resources. Most of computationally intense calculations presented here, such as the minimization and molecular dynamics simulations of 500-bp DNA, were performed at the Joint Supercomputer Center of the Russian Academy of Sciences.

1. Savin AV, Mazo MA, Kikot IP, Manevitch LI, Onufriev AV (2011) Heat conductivity of the DNA double helix. *Phys Rev B* 83(24):245406.
2. Still WC, Tempczyk A, Hawley RC, Hendrickson T (1990) Semianalytical treatment of solvation for molecular mechanics and dynamics. *J Am Chem Soc* 112(16):6127–6129.
3. Grosberg AY, Khokhlov A (1994) *Statistical Physics of Macromolecules* (American Institute of Physics, New York).
4. Christiansen PL, Zolotaryuk AV, Savin AV (1997) Solitons in an isolated helix chain. *Phys Rev E Stat Phys Plasmas Fluids Relat Interdiscip Topics* 56(1):877–889.
5. Tsui V, Case D (2000) Molecular dynamics simulations of nucleic acids using a generalized Born solvation model. *J Am Chem Soc* 122(11):2489–2498.
6. Onufriev A, Case DA, Bashford D (2002) Effective Born radii in the generalized Born approximation: The importance of being perfect. *J Comput Chem* 23(14):1297–1304.
7. Tsui V, Case DA (2000–2001) Theory and applications of the generalized Born solvation model in macromolecular simulations. *Biopolymers* 56(4):275–291.
8. Sorin EJ, Rhee YM, Nakatani BJ, Pande VS (2003) Insights into nucleic acid conformational dynamics from massively parallel stochastic simulations. *Biophys J* 85(2):790–803.
9. De Castro LF, Zacharias M (2002) DAPI binding to the DNA minor groove: A continuum solvent analysis. *J Mol Recognit* 15(4):209–220.
10. Allawi HT, et al. (2003) Modeling of flap endonuclease interactions with DNA substrate. *J Mol Biol* 328(3):537–554.
11. Chocholousová J, Feig M (2006) Implicit solvent simulations of DNA and DNA-protein complexes: Agreement with explicit solvent vs experiment. *J Phys Chem B* 110(34): 17240–17251.
12. Ruscio JZ, Onufriev A (2006) A computational study of nucleosomal DNA flexibility. *Biophys J* 91(11):4121–4132.
13. Srinivasan J, Trevathan M, Beroza P, Case D (1999) Application of a pairwise generalized Born model to proteins and nucleic acids: Inclusion of salt effects. *Theor Chem Acc* 101:426–434.
14. Case DA, et al. (2005) The Amber biomolecular simulation programs. *J Comput Chem* 26(16):1668–1688.
15. Lu XJ, Olson WK (2003) 3DNA: A software package for the analysis, rebuilding and visualization of three-dimensional nucleic acid structures. *Nucleic Acids Res* 31(17): 5108–5121.
16. van Mameren J, et al. (2009) Unraveling the structure of DNA during overstretching by using multicolor, single-molecule fluorescence imaging. *Proc Natl Acad Sci USA* 106 (43):18231–18236.
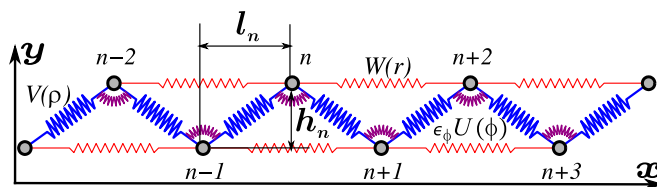
Fig. S1. Schematic of a 2D zigzag polymer chain. $\varepsilon_\phi U(\phi)$ and $V(\rho)$ are the valent angle bending and bond stretching potentials; $W(r)$ is the nonvalent interaction between next-nearest neighbors. The longitudinal step of site n is l_n, and the transverse step is h_n.
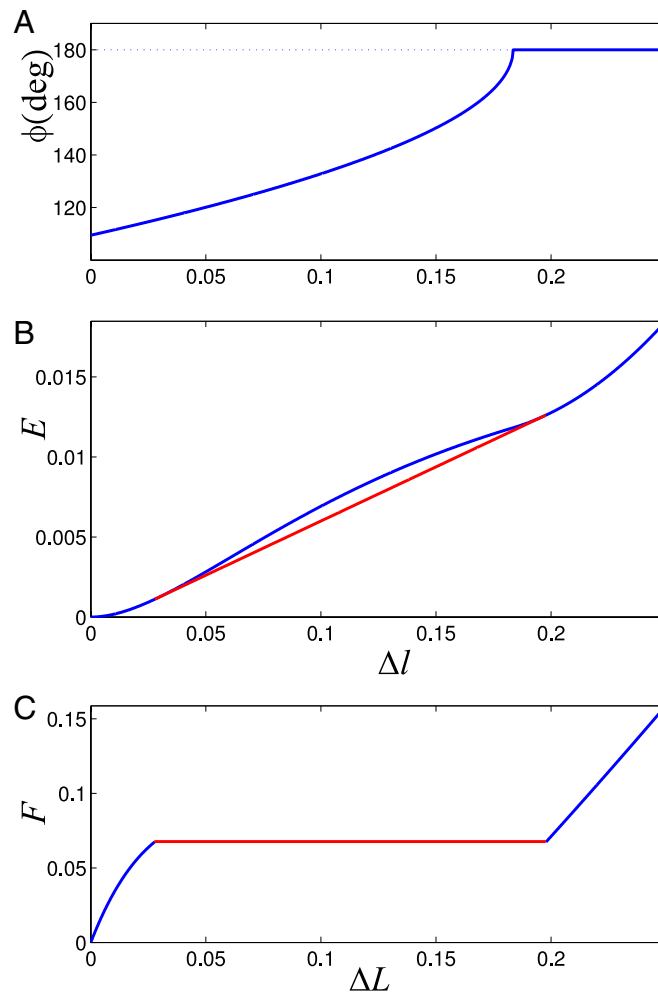
Fig. S2. 2D zigzag. The dependence of (*A*) angle ϕ, (*B*) effective site deformation energy *E* of the zigzag on extension Δl of its sites, and (*C*) tension as a function of the mean site extension. Angle deformation stiffness is $\varepsilon_\phi = 0$, which leads to the two-phase stretching regime. The red line in *B* is the convex hull of *E*(*l*). Dimensionless units.
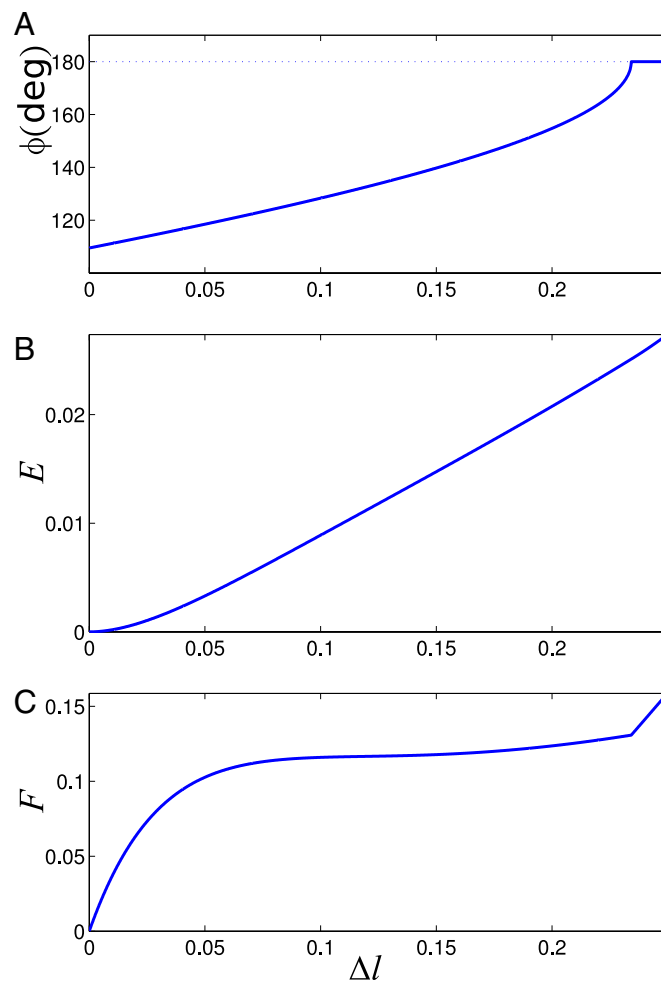
Fig. S3. 2D zigzag. The dependence of (*A*) angle ϕ, (*B*) effective site deformation energy *E*, and (*C*) tension as a function of the site extension. Angle deformation stiffness is $\varepsilon_\phi = 0.02$; nonvalent interactions dominate the elastic response, and therefore, only uniform stretching is possible. Dimensionless units.
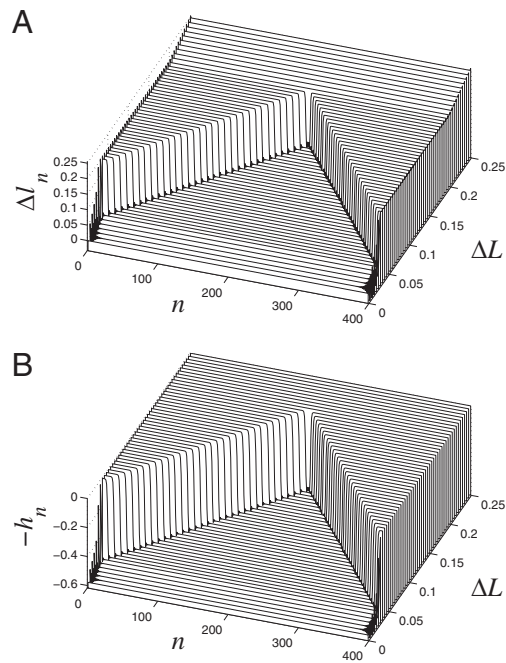
Fig. S4. The two-phase stretching in 2D zigzag chain. Shown is the dependence of site extension Δl_n and transverse step h_n on the site number n and the mean (per site) extension ΔL. A chain with $n = 400$ sites (atoms) with fixed ends. Angle deformation stiffness is $\varepsilon_\phi = 0$. Dimensionless units.

Fig. S5. The uniform stretching of the 2D zigzag chain. Shown is the dependence of site extension Δl_n and transverse step h_n on the site number n and the mean (per site) extension ΔL. A chain with $n = 400$ sites (atoms) with fixed ends. Angle deformation stiffness is $\varepsilon_\phi = 0.02$. Dimensionless units.
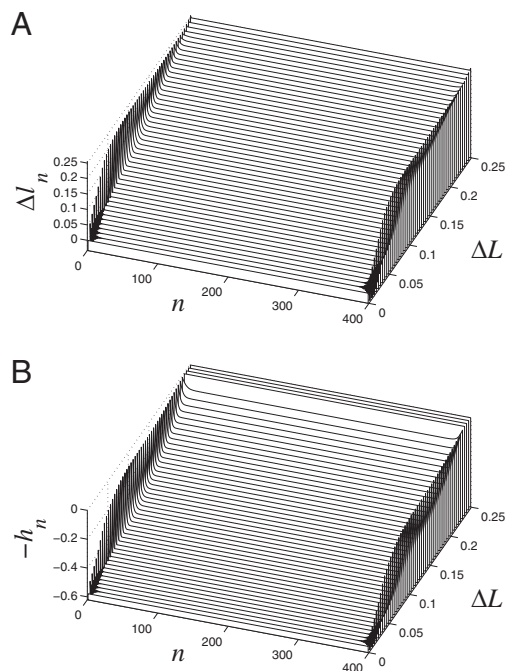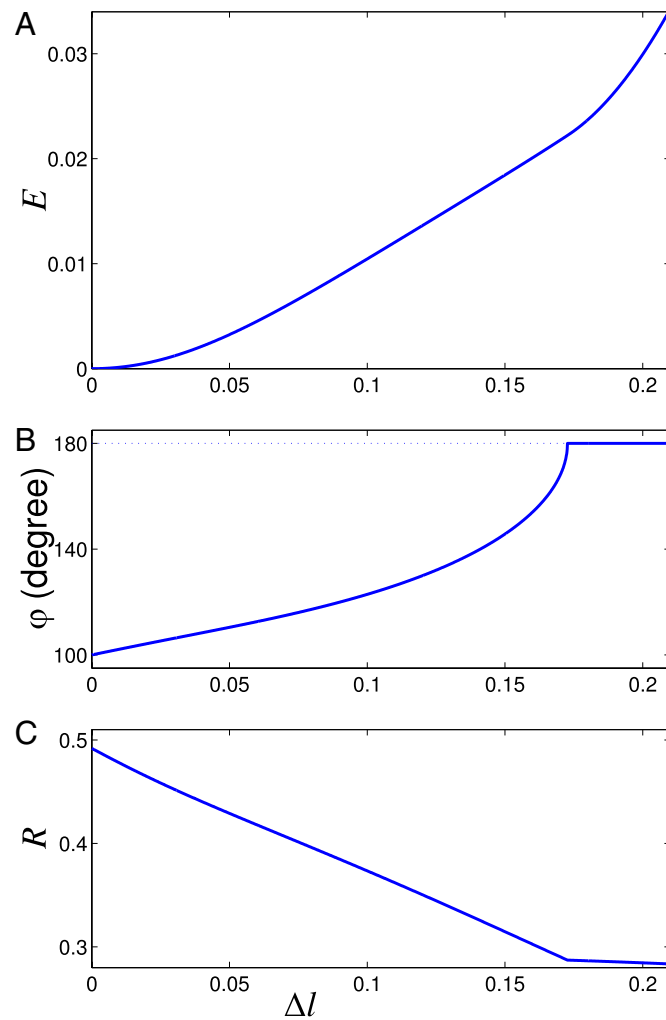
Fig. S6. (A) Dependence of the effective site energy E, (B) angular step φ, and (C) radius R of the helix from on its relative (uniform) longitudinal site extension Δl. Torsional rigidity is $\varepsilon_\theta = 0.002$. Dimensionless units.

Fig. S7. A DNA fragment in the coarse-grained representation (1) used here. Each base pair (site) is modeled by 12 united atom particles (grains).

1. Savin AV, Mazo MA, Kikot IP, Manevitch LI, Onufriev AV (2011) Heat conductivity of the DNA double helix. *Phys Rev B* 83(24):245406.
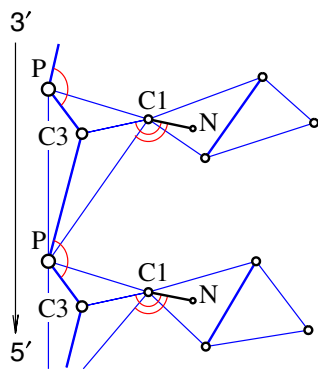
Fig. S8. United atom particles (grains) involved in the valent interactions in the 12CG coarse-grained DNA model. Blue lines denote valent (harmonic) bonds, red arcs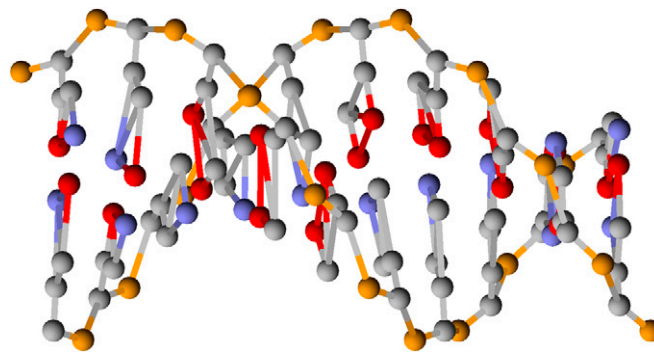 mark valent angles, and bold lines are axes of rotation in the torsional potentials. The circles marked as N stand for atoms N9 in A and G bases and N1 in T and C bases (no grain is situated on these atoms; their coordinates are calculated directly from positions of the base grains as detailed in ref. 1).

1. Savin AV, Mazo MA, Kikot IP, Manevitch LI, Onufriev AV (2011) Heat conductivity of the DNA double helix. *Phys Rev B* 83(24):245406.
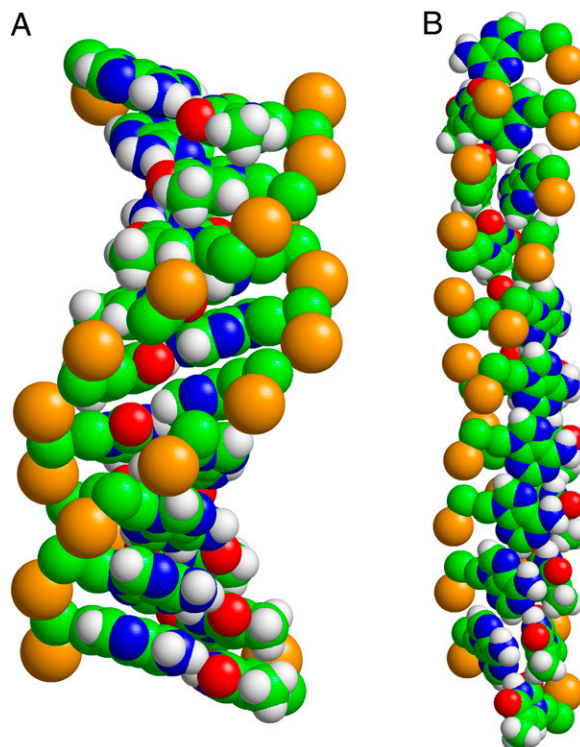
**Fig. S9.** Schematic view of the minimum energy (ground) state of (*A*) weakly stretched (longitudinal step is $l = l_a = 3.75$ Å) and (*B*) strongly stretched (longitudinal step is $l = l_b = 6.15$ Å) poly(A)-poly(T) DNA double helix.

Fig. S10. Schematic view of a fragment of (*A*) weakly stretched (averaged longitudinal step is $L = l_a = 3.75$ Å) and (*B*) strongly stretched (averaged longitudinal step is $L = l_b = 6.15$ Å) poly(A)-poly(T) DNA double helix at $T = 300$ K.

Fig. S11. (*A*) Effective energy $E(l)$ per base pair of extended poly(A)-poly(T) DNA in ground state, hydrogen bond energy E_{hb}^*, and neighbor base pair stacking interaction energy E_{st} as a function of longitudinal step *l*. (*B*) The tension as a function of the average longitudinal step *L*.

Fig. S12. The two-phase stretching in dsDNA. Shown is the dependence of site length l_n and angular step (twist) $\Delta\phi_n$ on the base pair number n along the DNA chain. A 500-bp poly(A)-poly(T) fragment was used in the simulation at 300 K; several values of the relative mean extension L/l_0 are tested. (*A* and *B*) $L/l_0 =$ 1.33; (*C* and *D*) $L/l_0 =$ 1.48; and (*E* and *F*) $L/l_0 =$ 1.63.

Fig. S13. Fraction of remaining hydrogen bonds as a function of relative chain extension for dsDNA. Blue line, artificially strong WC bonds ($c_0 = 1.0$). Red line, regular-strength WC bonds ($c_0 = 0.4$). Simulation at $T = 300$ K.

Fig. S14. Force-extension diagram for an ideal dsDNA helix. Blue line, artificially strong WC bonds ($c_0 = 1.0$). Red line, regular strength WC bonds ($c_0 = 0.4$). Simulation at $T = 300$ K.