Published in final edited form as:

J Phys Chem A. 2011 September 1; 115(34): 9377–9391. doi:10.1021/jp110871g.

B-DNA to Zip-DNA: Simulating a DNA Transition to a Novel Structure with Enhanced Charge-Transport Characteristics

Alexander Balaeff^{*,†}, Stephen L. Craig[†], and David N. Beratan^{†,‡,§}

[†]Department of Chemistry, Duke University, Durham, North Carolina 27708, United States

[‡]Department of Biochemistry, Duke University, Durham, North Carolina 27708, United States

[§]Department of Physics, Duke University, Durham, North Carolina 27708, United States

Abstract

The forced extension of a DNA segment is studied in a series of steered molecular dynamics simulations, employing a broad range of pulling forces. Throughout the entire force range, the formation of a zipper-like (zip-) DNA structure is observed. In that structure, first predicted by Lohikoski et al., the bases of the DNA strands interdigitate with each other and form a single-base aromatic stack. Similar motifs, albeit only a few base pairs in extent, have been observed in experimental crystal structures. Analysis of the dynamics of structural changes in pulled DNA shows that S-form DNA, thought to be adopted by DNA under applied force, serves as an intermediate between B-DNA and zip-DNA. Therefore, the phase transition plateau observed in force–extension curves of DNA is suggested to reflect the B-DNA to zip-DNA structural transition. Electronic structure analysis of purine bases in zip-DNA indicates a several-fold to order of magnitude increase in the π – π electronic coupling among nearest-neighbor nucleobases, compared to B-DNA. We further observe that zip-DNA does not require base pair complementarity between DNA strands, and we predict that the increased electronic coupling in zip-DNA will result in a much higher rate of charge transfer through an all-purine zip-DNA compared to B-DNA of equal length.

INTRODUCTION

Single-molecule force spectroscopy measurements on DNA have drawn significant interest since the early 1990s.^{1–12} Experimental studies reveal changes in DNA structure and dynamics under applied tension; fitting the experimental data to models yields direct estimates of such basic DNA properties as the entropy and enthalpy of melting per base

© 2011 American Chemical Society

*Corresponding Author: abalaeff@duke.edu.

Supporting Information. Detailed SMD simulation protocols, Voronoi partition of the MD unit cell for the electrostatics analysis, additional energy and structural data from the SMD simulations, distribution plots for electronic coupling between nucleobases in zip-DNA, and derivation of the critical tilt for S-zip DNA transition. This material is available free of charge via the Internet at <http://pubs.acs.org>.

pair,^{13,14} the energy of base stacking,¹⁵ the elastic modulus,^{5,6} and the amount of bend-to-twist coupling.¹¹ In addition, micromanipulation experiments directly probe DNA interactions with drugs^{16,17} and proteins, including RNA polymerase⁹ and helicase.¹⁰ Understanding fundamental DNA properties, and the structural response of DNA to its environment, is essential for establishing a quantitative description of such key biological processes as transcription, replication, and DNA packing, as well as for the design of DNA-based nanodevices.^{18–25}

Importantly, structural changes during the forced extension of DNA affect the DNA conductance.²⁶ Charge flow through DNA and between DNA and repair proteins plays an important role in genome maintenance and has been suggested as the basis for nanotechnology applications.^{18,19,23,27} For such applications, it is crucial to learn how to manipulate DNA conductance via its sequence, structure, and thermal fluctuations, all of which are known to have profound effects on DNA charge transfer.^{24,25,27–36} To this end, forced extension is one possible way of manipulating DNA structure and dynamics, creating a direct feedback to the applied force through changes in the electric current.

In a typical micromanipulation experiment, DNA is stretched between two silicon beads that are trapped by optical or magnetic “tweezers” or are secured by a micropipette.^{1,5,6,8,17,37} Alternatively, the DNA can be stretched between a gold or glass surface and the cantilever of an atomic force microscope (AFM).^{15,38–40} As the ends of the captured DNA are pulled apart, the pulling force F is recorded vs DNA extension. A typical force–extension curve (Figure 1) begins with an “entropic elasticity” plateau where F of several piconewtons (pN) straightens the entangled DNA. As the DNA end-to-end length approaches its full contour length (L_0), the force begins to grow steeply, reflecting the elastic response of the stretched DNA. At $F \sim 70$ pN, the DNA undergoes a significant extension to $\sim 1.7\text{--}1.9 L_0$, resulting in a wide plateau in the curve. This plateau is often attributed to the DNA transition from the classical B-form to the so-called S-form.^{5,41,42} Beyond the 70 pN plateau, the steep force growth resumes, presumably reflecting the elastic response of S-DNA. Eventually, the growing force causes the DNA to melt, which is reflected by another plateau or even a drop in F .^{38,39} The final steep portion of the curve corresponds to pulling on the single strand(s) of melted DNA. The number, width, and height of the plateaus in the force–extension curve may vary depending on the temperature,^{2,14,22,37,43} the DNA sequence,^{22,38,40} the number of nicks in the DNA strands,^{37,44,45} solvent pH,^{2,37,46} salt concentration,^{2,22,37,40,43} torsional constraints,³⁸ pulling speed,^{22,38} and the stiffness of the AFM cantilever.³⁹ Under the right conditions, the S-DNA state may conceivably be skipped altogether, resulting in a force–extension curve with a single plateau corresponding to direct force-induced melting of B-DNA.^{2,14,22,37,45}

The structure of S-DNA was predicted in modeling studies.^{5,41,42,47,48} Compared to B-DNA, the S-form is not only extended but is also almost completely unwound, and its Watson–Crick (WC) base pairs are tilted with respect to the helix axis.^{41,42,47} Modeled stretching of S-DNA beyond $1.7L_0$ produced two possible DNA forms, depending on whether the DNA strands were pulled on their 3′ or 5′ ends.⁴¹ Pulling on the two 3′ ends produced a “flat ribbon” where the unwound DNA stack was split into clusters of two base pairs, each with significant gaps between the clusters.⁴¹ Pulling on the two 5′ ends produced a “narrow fiber” where the backbone strands approached each other closely, the strongly tilted nucleobases lost most of their WC contacts, and the bases formed a twisted sequence of 4-base stacks.⁴¹ Several recent molecular dynamics (MD) studies supported the picture of a B-to-S DNA transition and demonstrated that the forced DNA extension may produce complex structural changes, including local DNA melting.^{49–51}

S-DNA or melted DNA are not the only possible forms of extended DNA. A simulation by Lohikoski et al.⁵² revealed that a 22 bp long DNA segment may convert into a novel zipper-like structure, if pulled simultaneously on both strands by a force gradually increasing from 0 to 600 pN. In the zipper structure, the WC base pairs are broken, and the nucleobases from the opposite DNA strands interdigitate like the teeth of a zipper, forming a continuous aromatic stack. Short DNA structures of this kind have been observed in X-ray structures of DNA containing mismatches or bulges,^{53,54} as well as in X-ray structures of the four-stranded i-DNA.⁵⁵

Here, we simulate the extension of a 16 bp DNA segment using constant-force steered molecular dynamics (SMD),^{56,57} and we investigate how the resulting structural changes affect the DNA electronic properties. The structural transition from B-DNA to the zipper form (zip-DNA) is consistently produced by a broad range of pulling forces (100 pN to 11.2 nN). In large-force simulations, zip-DNA self-assembles from force-melted DNA, whereas in small-force simulations B-DNA evolves into zip-DNA through an S-DNA intermediate. Consistency of the zip-conformation with the $1.7L_0$ upper end of the 70 pN transition plateau of the force–extension curve, as well as the surprisingly low energy of zip-DNA (~ 2 (kcal/mol)/base), indicate that the 70 pN plateau could in fact reflect the B–S–zip structural transition. Experiments designed to detect zip-DNA are suggested here. Analysis of the zip-DNA electronic structure shows that the electronic coupling between the highest filled orbitals of the nearest-neighbor bases is several-fold larger in zip-DNA than in B-DNA. Therefore, a much higher conductance is predicted to be found in zip-DNA than in B-DNA of equal length, suggesting potential applications for zip-DNA in the rapidly developing field of DNA-based nanotechnology.

COMPUTATIONAL MODEL AND METHODS

Model Building

The 16 bp long DNA segment studied has the sequence GGTATACCGCTTAAGC (Scheme 1), which consists of two palindromic halves. This DNA segment was shown to be able to self-assemble into long chains of nicked DNA.⁵⁸ The initial all-atom structure of the DNA was built using the online tool *model.it*.⁵⁹ The structure was solvated in a $65 \text{ \AA} \times 65 \text{ \AA} \times 95 \text{ \AA}$ box of TIP3 water with 30 Na^+ counterions and equilibrated in a 5 ns MD simulation. The simulation was performed using NAMD⁶⁰ and employed the CHARMM 27 force field,⁶¹ periodic boundary conditions, an NPT ensemble ($T = 300 \text{ K}$, $P = 1 \text{ atm}$), full electrostatics computed with the particle mesh Ewald method, and a multiple time-stepping integration scheme. Snapshots of the DNA structure were saved every 1 ps during the simulation. A representative DNA snapshot with a small average root mean square deviation (rmsd) from the other snapshots was selected from the preliminary MD trajectory (see Supporting Information for details) and was used to initiate the subsequent SMD simulations.^{56,57} Test runs using different initial structures yielded qualitatively similar results.

SMD Simulations

The selected DNA structure (Figure 2A) was resolvated in a $120 \text{ \AA} \times 65 \text{ \AA} \times 65 \text{ \AA}$ box of TIP3 water with 30 Na^+ ions. The end-to-end axis of the DNA was aligned with the long axis of the box. Prior to commencing the SMD runs, the water and ions were equilibrated for 0.5 ns around the harmonically restrained DNA. The average values of the MD energy components during the last 0.4 ns of water equilibration were used as baseline values for the energy components in the subsequent SMD runs (cf. Tables 2 and 3).

The SMD simulations were performed in the constant-force regime: on one end of the DNA segment, a constant force of $F/2$ was applied to each DNA strand, on the other end, one atom of each strand was harmonically restrained to its initial position (see Supporting

Information for details). Importantly, both DNA strands were pulled simultaneously in the same direction, which is different from most previous SMD simulations of DNA,^{41,42,47–50,62} as well as many optical tweezers experiments.^{1,5,6,8} The harmonic restraints imposed here had the additional effect of maintaining the distance between the DNA strands on the restrained end.

Table 1 summarizes all of the simulations performed, including the pulling forces and the simulation times. The force field and the simulation conditions during the SMD runs were the same as those used during the preliminary equilibration run. The atomic coordinates of the full system were saved every 1 ps.

SMD Simulation Analysis

The parameters characterizing the DNA structure during the SMD simulations, including the DNA end-to-end length, the dihedral angles, the root-mean-square deviation, etc., were computed using VMD.⁶³ The end-to-end DNA length was calculated for the 14 bp (28 bases) DNA segment between the midpoint of the restrained atoms and the midpoint of the pulled atoms (see Supporting Information). The terminal base pairs were excluded from the analysis because of (i) the boundary effect disorder and (ii) the effect of the SMD restraints and pulling forces on the conformation of the base pairs (see Supporting Information for details). The baseline DNA length L_0 was set to 47.6 Å, which is the length of a 14 bp B-DNA segment.

The structure of the simulated DNA was compared to reference structures of B-form and zip-form DNA via the root mean square deviation (rmsd) of each SMD DNA snapshot from those structures. The reference B-DNA structure was represented by the DNA snapshot used to initiate all of the SMD trajectories (vide supra). The reference zip-DNA structures for the large-force SMD trajectories were represented by the final DNA snapshots in those trajectories. Each such snapshot corresponded to a well-equilibrated zip-DNA (cf. Table 1). The reference zip-DNA structures for the small-force SMD trajectories were represented by the final DNA snapshot from the SMD1 trajectory, because the small-force simulations did not converge to a stable zip-DNA structure during the allotted simulation time (cf. Table 1, Figures 3 and 4). The latter reference structure has to be used with a caveat that an equilibrated zip-DNA structure, presumably achievable in the long run, may be force-dependent and therefore different among the small-force simulations (vide infra).

The rmsd was computed following the best-fit alignment of a DNA structure in an SMD snapshot with the reference DNA structure. Only the coordinates of non-hydrogen DNA atoms of nucleotides A₃ through A₁₄ of each strand (cf. Scheme 1) were used for the alignment and the rmsd analysis. Nucleotides 1, 2, 15, and 16 were excluded from the rmsd analysis as noted above.

The SMD trajectory energy analysis was performed with NAMD.⁶⁰ The various components of the MD energy function (cf. Table 2, Figure 4C,F) were recorded every 20 fs during each simulation. In addition, the saved SMD snapshots were used to compute the energies of electrostatic interaction between different subsystems of the simulated system, e.g., the repulsion between the two DNA strands (Table 3). Because of the charged nature of the subsystems analyzed, it was impossible to compute their electrostatic energies using Ewald summation on an infinite grid, as was done for the whole system during the SMD simulation. The interactions were therefore computed for a single unit cell using an infinite electrostatic cutoff. In order to better approximate the solvent screening observed in the simulated periodic system, the solvent molecules of the unit cell were relocated to their images closest to the DNA, effectively resulting in a Voronoi partitioning of the solvent with respect to the DNA (cf. Supporting Information). The absolute values of the resulting

electrostatic energies differ from those obtained with periodic boundary conditions (cf. Tables 2 and 3). Yet, comparing the subsystem energies between different SMD simulations still provides useful information about the physical effects driving the DNA structural change during the B-zip transition.

Electronic Coupling

Purine–purine electronic couplings in zip-DNA were calculated for G and A bases stacked against each other in the final zip-DNA structure: A_{12B}/A_{6A}, A_{5B}/A_{13A}, and G_{9B}/G_{9A} (Scheme 1). The bases A_{14B}/A_{4A} were not included in the QM computations because of a larger degree of disorder that reduced the A/A coupling compared to the more ordered A/A bases (Figure S6, Supporting Information). Only the purine–purine couplings were computed because the positive charge carriers in DNA reside primarily on purines.^{24,27,32–34} Structural ensembles for the selected 2-purine stacks were obtained from the last 500 ps (one snapshot per picosecond) of a large-force SMD trajectory SMD-H1 that converged sufficiently well to zip-DNA (Figures 2 and 4, Table 1). Only the nucleobase atoms were extracted from the MD snapshots for each 2-base stack. The rest of the DNA, as well as water and counterions, were excluded from the QM computations. The dangling N₉–C₁' bonds resulting from the base extraction were capped with hydrogens. See ref 30 for further details.

Molecular orbitals were computed for each extracted 2-base stack using the semiempirical INDO/s method,⁶⁴ as implemented in the program CNDO.⁶⁵ The orbitals were computed both for the whole 2-base stack and for each nucleobase separately. Then, the coupling V_{12} between the highest occupied molecular orbitals (HOMOs) of the bases was computed as

$$V_{12} = \langle \phi_2' | H | \phi_1' \rangle \quad (1)$$

where H is the Fock matrix of the 2-base system, and $|\phi_1'\rangle$, $|\phi_2'\rangle$ are the isolated-nucleobase HOMOs $|\phi_1\rangle$, $|\phi_2\rangle$ extended to the 2-base system in the following way

$$\begin{aligned} |\phi_1\rangle &= \sum_i a_i |\lambda_i^1\rangle \rightarrow |\phi_1'\rangle = \sum_i a_i |\lambda_i^1\rangle + \sum_j 0 \cdot |\lambda_j^2\rangle \\ |\phi_2\rangle &= \sum_j b_j |\lambda_j^2\rangle \rightarrow |\phi_2'\rangle = \sum_i 0 \cdot |\lambda_i^1\rangle + \sum_j b_j |\lambda_j^2\rangle \end{aligned} \quad (2)$$

Here, $|\lambda_i^1\rangle$, $|\lambda_j^2\rangle$ are the atomic orbital (AO) basis functions for nucleobases 1 and 2, respectively, and a_i and b_j are the linear coefficients of $|\phi_1\rangle$ and $|\phi_2\rangle$ in their respective AO basis sets. Such an extension is possible because CNDO simply merges the AO basis sets of the two nucleobases to form the AO basis set of the 2-nucleobase system.

Reference values of GG and AA electronic couplings in B-DNA were calculated using the ensemble of 5000 DNA structures generated in the preliminary 5 ns MD simulation of the modeled DNA segment (vide supra). 2-Base systems G_{9B}/G_{10B}, A_{13A}/A_{14A}, and A_{5B}/A_{6B} were used to compute the intrastrand GG and AA couplings, and 2-base systems G_{9A}/G_{9B}, G_{9A}/G_{7B}, A_{4A}/A_{14B}, A_{4A}/A_{12B}, A_{6A}/A_{12B}, and A_{13A}/A_{5B} were used to compute the interstrand GG and AA couplings (cf. Scheme 1). Due to the limited number of 2-base systems, and for the sake of simplicity, we do not distinguish between the 3'-3' and 5'-5' interstrand coupling in this study.³²

RESULTS

SMD Simulations Produce a B–Zip DNA Transition

The SMD simulations of the DNA segment GGTATACCGCTTAAGC⁵⁸ employed either a small pulling force (F) in the range of 100–240 pN or a large pulling force of several nanonewtons. A complete or partial transition from B-DNA to zip-DNA was observed in all of the simulations. The small forces are typical for SMD simulations^{56,57} and correspond to the upper limit of the force range in DNA pulling experiments (cf. Figure 1). In the simulations here, the small forces produced a partial B–zip transition during the accessible simulation time (cf. Figures 3 and 4). The large (nanonewton range) forces produced a complete transition from B-DNA to a stable zip-DNA structure (cf. Figures 2 and 4), even if the transition pathway to the structure was likely unrealistic, as described below.

Six large-force simulations were run with $F = 5.6$ nN or $F = 11.2$ nN for 10 ns or 5 ns (Table 1). Each large-force simulation produced a B–zip DNA transition that proceeded via an almost completely melted DNA state, reached in ~ 0.1 ns after the beginning of the simulation (Figure 2B). The backbone of the melted DNA was extended to its final length of $2.3\text{--}2.5L_0$ (Figures 2B and 4A), while the nucleobases flipped out at random angles, losing most of the interstrand WC hydrogen bonds (except at the restrained end) and most of the intrastrand stacking contacts (Figure 2B). The melted state was less disordered in $F = 5.6$ nN simulations than in $F = 11.2$ nN simulations; in the former case, nucleobases from the opposite strands frequently formed stacking contacts in lieu of the lost Watson–Crick bonds. However, the melted state of the DNA did not persist and the nucleobase rotation around the extended backbone eventually lead to formation of the zip-DNA: an aromatic stack of interlocked nucleobases (Figure 2D). The structural reorganization took 1–2 ns (5 ns in one case), as illustrated by the rmsd dynamics in Figure 4B. Apparently, the reorganization time was affected by the degree of structural disorder resulting from the fast initial DNA extension. Once assembled, the zip-stack persisted until the end of the simulation. Thermal fluctuations around the equilibrium structure produced an extremely small rmsd of ~ 1 Å, due to an occasional nucleobase flipping out of the stack (Figures 2D and 4B and Table 1).

The small-force SMD simulations produced a more gradual and, presumably, more realistic evolution of the DNA structure from B- to zip-form. Eleven small-force SMD simulations employed F from 100 to 240 pN and ran between 20 and 50 ns (Table 1). A small initial extension of B-DNA in these simulations produced one or several S-DNA nucleation sites where the DNA unwound and the base pairs tilted. Further DNA extension caused either (i) spreading of the S-DNA structure from the nucleation sites, as seen in other simulations,^{41,42,49–51} or (ii) the conversion of the already formed S-DNA sites to zip-DNA form, consistent with the observations of Lohikoski et al.⁵² The S-to-zip conversion occurs in the simulations because continued DNA extension increases both the tilt and the stretch of the S-form base pairs. The increased tilt produces increased interstrand base overlap (Figure S7, Supporting Information), while the increased stretch eventually ruptures the WC hydrogen bonds. When the WC bonds rupture, the overlapping bases from the opposite strands slide past each other, forming a zip-DNA nucleation site.

The mixed DNA structures resulting from the small-force SMD simulations are shown in Figure 3 and are described in Table 1. In 20–40 ns long simulations with $F = 160\text{--}240$ pN, the DNA segment adopted a mixed S/zip-conformation. In 40–50 ns long simulations with $F = 100\text{--}130$ pN, the DNA segment adopted mixed B/S/zip-conformations; the mix of all three structures was accessed because some of the S-DNA sites adopted a zip-conformation before the entire B-DNA converted to the S-form. The final number of bases adopting a zipconformation varied from 4 to 20 bases, out of 32 total bases (Table 1). The number of zip-conformation bases correlated not only with the pulling force, but also with the

simulation time. None of the small-force simulations apparently reached equilibrium (Figure 4D,E and Table 1), so it appears likely that the small-force simulations would have converted the entire DNA segment to zip-form too, had they been run for a longer time. It is, however, unclear how much longer that simulation time would need to be, or whether a mixed B-/S-/zip-DNA structure (rather than a pure zip-DNA) represents the equilibrium point at the low end of the force spectrum.

Molecular Structure of Zip-DNA

The structural properties of zip-DNA may be deduced from the equilibrium DNA structures reached in the large-force simulations (Figure 2D). The structures closely resemble the one obtained by Lohikoski et al.⁵² The backbone of each DNA strand adopts a straight untwisted conformation; a similar conformation is observed in extended S-DNA.^{41,42,47} The separation distance between the two strands is about 10 Å, i.e., half of the canonical B-DNA diameter. Both the chemical bond lengths and the bond angles in the backbone are strained by the pulling force (cf. the energy plots in Figures 4C,F and S2). However, the main structural changes that result in a B-zip DNA transition are the torsional rotations in the DNA backbone that convert the backbone to an all-trans state (Figure 2C). The largest rotations are those by 120° around each P–O_{5'} bond, by 100° around each C_{4'}–C_{5'} bond, by 55° around each C_{3'}–O_{3'} bond, and by 40° around each P–O_{3'} bond (the rotation angles are measured as ensemble-average changes in corresponding dihedral angles with respect to the initial equilibrated DNA structure). In the large-force simulations, the torsional rotations occur almost simultaneously during the early stages of the DNA pulling, whereas in the small-force simulations, the rotations occur gradually, one by one, during the late stages of the DNA pulling. The torsional rotations are in fact sufficient for the backbone to convert to the all-trans state, while chemical bond lengths and angles need not change significantly. For example, the DNA pulled by a 200 pN force in simulation SMD7 (cf. Table 1) converted almost completely to zip-form (Figure 3) with a small increase in conformational energy: 2.1 (kcal/mol)/base, of which 1.1 (kcal/mol)/base was accounted for by the bond and angle energy and 1.0 (kcal/mol)/base by the torsional energy (Figure 4F, Table 2). This conformational energy is 20–30 times less than the conformational energy of DNA attained in the large-force simulations (Figures 4C and S2 and Table 2). The pucker angles of the backbone sugars do not change much during the B-zip transition either. With the exception of the terminal residues, all of the pucker angles fluctuate in the 130–170° range characteristic of the C_{2'}-*endo* sugar conformation found in B-DNA (Figure S3, Supporting Information).

The nucleobases from the opposite strands of zip-DNA interdigitate, forming a single-base aromatic stack (Figure 2D). The polar moieties of the bases form hydrogen bonds with the oxygen atoms of the opposite backbone strand. The bases are tilted by 20–30° with respect to the helix axis, which is much less than in the S-structure.^{41,42,47} It should be noted that the parameters of a zip-DNA structure, such as the tilt, the base-to-base distance along the stack, and the resulting end-to-end length of the DNA depend on the stretching force F . By increasing the force, one increasingly stretches the backbone, thus increasing the base–base distance in the zip-stack. For example, the end-to-end length of zip-DNA produced in 5.6 nN simulations is $2.19L_0$, and that produced in 11.2 nN simulations is $2.33L_0$ (Figure 4A and Table 1). In the small-force simulations, the end-to-end DNA length at the end of the simulations varied from $1.21L_0$ to $1.70L_0$ (Figure 4D and Table 1), but the DNA structure in those simulations continued to evolve. At present, it is unclear what the equilibrium zip-DNA length would be under small forces, if, indeed, the small forces are sufficient to achieve a B-zip transition of the entire DNA segment. It can be assumed that the smallest possible length of a pure zip-DNA would be equal to the length of the DNA backbone converted to the all-trans state, as described above, but without changing the DNA bonds

and angles. The geometry of the converted DNA steps in our small-force simulations indicates that the length of the all-trans DNA backbone is $\sim 2.1L_0$ with respect to L_0 of an ideal B-DNA. With respect to L_0 of a specific DNA sequence, that ratio would be slightly different due to sequence-specific variations in L_0 .

Electronic Couplings in Zip-DNA

The hole-type conductance of DNA is determined primarily by the electronic couplings between purine bases that act as hole traps (G) or bridge states (A).^{24,27,32–34,36} The geometric overlap between neighboring bases is larger in zip-DNA than in B-DNA;⁵² therefore, larger couplings between the hole-localizing orbitals are expected in zip-DNA than in B-DNA. We approximated the hole orbitals by the HOMOs of the purine bases^{24,30,32} and computed the HOMO–HOMO couplings between neighboring purine bases in B-DNA and zip-DNA (Figure 5 and Table 4).

The computed ensemble of V_{XX} values ($X = A, G$) exhibits a log–log–normal distribution (Figure 5). The width of the distribution is smaller for zip-DNA couplings than for B-DNA couplings, which may reflect a smaller degree of structural disorder in zip-DNA compared to B-DNA, but may also be a computational artifact due to a smaller ensemble size for zip-DNA couplings (see Methods). There are several possible ways to compute the ensemble-averaged coupling; Table 4 shows the average absolute coupling $\langle |V_{XX}| \rangle$ and the root-mean-

square coupling $\left(\langle |V_{XX}|^2 \rangle \right)^{1/2}$. The former reflects the average coupling strength, whereas the latter reflects the difference in base–base charge transfer rate, as determined by Fermi's golden rule.^{25,27,34} By either averaging method, the average GG coupling in zip-DNA is ~ 3 times larger than the average intrastrand GG coupling in B-DNA and 14–19 times larger than the average interstrand GG coupling in B-DNA. The zip-DNA AA coupling is ~ 2 times larger than the intrastrand B-DNA AA coupling and 4–5 times larger than the interstrand B-DNA AA coupling. Multiplied over several purine–purine steps, the larger nearest-neighbor coupling may lead to a significantly larger charge transfer rate through zip-DNA than through B-DNA of the same length (vide infra).

DISCUSSION

Earlier Simulations of Zip-DNA

Structures of zip-DNA, similar to the ones obtained in our large-force simulations, were seen twice before. Lohikoski et al.⁵² produced a zip-conformation for a 22 bp DNA sequence with lower GC content (27%) than in the sequence studied here (50%). Santosh and Maiti⁶⁶ produced a zip-conformation for a 12 bp DNA structure with a high GC content (67%), but did not distinguish the zip-form from an S-form conformation. Structurally, the zip-DNA in ref 52 is very similar to the family of structures produced in the present simulations. The final zip-DNA length in that study is $2.3L_0$ (similar to the length achieved here in the large-force simulations), the nucleobases form a zipper-like stack, with base tilt and base-to-backbone hydrogen bonds that are very similar to those seen here. Visual comparison indicates that the zip-DNA fragments in ref 66 are also structurally similar to those found here.

In both refs 52 and 66, the DNA was subjected to a pulling force that changed dynamically from 0 to 600 pN over 2 ns⁵² or 6 ns.⁶⁶ Lohikoski et al. simulated their DNA using the CHARMM force field; Santosh and Maiti used AMBER. In our study, the simulations used the CHARMM force field and the zip-DNA structure was produced by a constant rather than changing force. The constant-force simulations with different pulling forces may be considered to probe the DNA structure at specific points of the experimental force–extension curve. Indeed, the experimental DNA pulling is slow (less than 1 $\mu\text{m/s}$) and the

pulling force is essentially constant on the time scale of tens of nanoseconds probed by the MD. We demonstrate that even a small constant force of 100–200 pN may produce a B-zip transition in the SMD simulation, which is not unexpected considering that the experimental force corresponding to the DNA structural transition plateau is ~70 pN (Figure 1).

The B-zip transition in the simulations of Lohikoski et al. occurred gradually, with a few base pairs briefly adopting an S-like intermediate structure. The simulations of Santosh and Maiti proceeded via a partially melted DNA state. The B-zip transition in the small-force simulations here proceeded through a variety of mixed B-/S-/zip-DNA structures. Indeed, we find that all three DNA states may coexist in the transition region. In the large-force simulations, we found that explosively melted DNA strands consistently self-assembled into a zip-DNA stack. The speed of this self-assembly on a nanosecond time scale is remarkable. Apparently, the self-assembly is driven by a large free energy gradient from a force-melted DNA state toward zip-DNA. That provides a significant argument in favor of zip-DNA being a viable DNA form adapted under stress, rather than being a structure that arises as a simulation artifact.

Zip-DNA Energetics

How realistic is the proposed single-base aromatic stack of zip-DNA? In general, aromatic molecules are known to assemble into a great variety of stacks.⁶⁷ For example, two aedamer chains may fold into a heterogeneous stack of interdigitating aromatic units.⁶⁸ Therefore, the ability of nucleobases to form similar stacks is not surprising. Indeed, a zip-stack of four adenines was observed in the X-ray structure of a double-helical homodimer (GCGAAAGCT)₂^{53,54} and was found to be stable in a 1 ns unrestrained MD simulation.⁶⁹ In the four-stranded i-DNA, nucleobases of one pair of strands partially intercalate between the nucleobases of the other pair of strands.⁵⁵ Therefore, the only heretofore unobserved aspect of the zip-DNA structure found here and in ref 52 is the scale of tens of nucleobase steps on which the interstrand intercalation occurs.

Energetically, zip-DNA appears to be a viable structure that exists under applied tension, compared to the extended S-DNA structures.⁴¹ In fact, it is remarkable that such a large conformational transition as the B-to-zip transition can be achieved for an enthalpic cost of less than 5 (kcal/mol)/base (Figure 4F, Table 2). The main enthalpic penalty during the B-zip transition comes from the conformational energy of the stretched sugar–phosphate backbone (Figure 4C,F and Table 2). Neither the van der Waals nor the electrostatic energies change much, except for a short initial spike in the electrostatic energy during the large-force simulations (Figures 4C, S2).

The absence of a larger electrostatic penalty for the B-zip transition seems surprising, considering that two negatively charged DNA strands are brought near each other in zip-DNA. However, the backbone extension increases the distance between the phosphate groups where the DNA charge is concentrated. As a result, the distance between all the phosphate groups increases except for those located opposite to each other on the two strands of the double helix. For example, the average phosphate–phosphate distance $\langle R_{pp} \rangle$ in the baseline DNA structure is 23.1 Å ($\langle 1/R_{pp} \rangle = 0.05 \text{ Å}^{-1}$), whereas in zip-DNA seen in the SMD-H1 simulation, $\langle R_{pp} \rangle = 42.3 \text{ Å}$ ($\langle 1/R_{pp} \rangle = 0.04 \text{ Å}^{-1}$). Therefore, both the intrastrand and the interstrand electrostatic repulsion of DNA decrease upon extension (Table 3). The net electrostatic penalty for the B-zip transition results solely from the increase in the energy of DNA–solvent interaction (Table 3). The origin of the latter increase is clear: the more compact a system of charges the less is the energy of the system interaction with the solvent medium, especially counterions. An immediate conclusion is that the force-induced B-zip transition should occur more easily at low salt concentrations, which are known to reduce the DNA overstretching force in general.^{43,45}

Finally, the hydrophobic penalty incurred by forming zip-DNA arises only from opening the DNA grooves, which increases the exposure of the base edges to solvent. The amount of exposure is, naturally, larger than that in a relaxed B-DNA but similar to that in S-DNA at moderate extension. At large extension ($L > 1.7L_0$), the hydrophobic and π -stacking interactions appear to favor zip-DNA over either of the two over-extended S-DNA forms.⁴¹ In the “flat ribbon” S-DNA, there are gaps in the base stack that are large enough to be filled by water, thus incurring a significant hydrophobic penalty. In the “narrow fiber” S-DNA, the continuous base pair stack is broken into short four-base stacks that are also significantly exposed to the solvent. In contrast, zip-DNA maintains a continuous stack of aromatic bases with no gaps up to at least $2.33L_0$. The base-to-base distance in the zip-stack increases with DNA extension but remains sufficiently small to prevent water penetration between the bases.

Zip-DNA and the Force–Extension Curve

If a B-zip DNA transition under applied tension is plausible, which part of the DNA force–extension curve corresponds to the transition? One possibility is that the B-zip transition occurs on the 70 pN plateau. In that case, the steep slope following the plateau (labeled “4” in Figure 1A) reflects the elastic resistance of zip-DNA, and the third plateau or a sharp drop in F (labeled “5” in Figure 1A) would mark the beginning of the DNA melting phase.^{38,39} This sequence of structural changes in DNA is supported by our observation that zip contacts appear early in B-DNA extension by small force. However, the experimental plateau ends at $L \approx 1.7\text{--}1.9L_0$, while the minimal length of a fully converted zip-DNA strand is estimated to be around $2.1L_0$. There are several ways in which this apparent discrepancy may be resolved. First, the end-to-end length of zip-DNA may be reduced compared to its contour length of $\sim 2.1L_0$ by some residual helical twist, as seen in our small-force simulations (Figure 3). Second, different DNA sequences may have different propensities to form zip-DNA, and parts of a long DNA molecule used in experiment (typically, several thousand base pairs) could adopt other conformations at $L \approx 1.7L_0$. Finally, precession of the rigid zip-DNA, observed in the large-force simulations here (Figure S4, Supporting Information), may cause the measured distance between the optical tweezer bead and the micropipette (or the AFM tip and the gold surface) to appear to be less than the actual DNA end-to-end distance.

It is also possible that the 70 pN plateau in the force–extension curve could indeed correspond to the B–S transition, and the upward part of the curve beyond the plateau ($L > 1.7L_0$) could correspond to the S-zip transition. This point of view is supported by the estimate that the critical tilt of S-DNA, where the nucleobases from the opposite strands effectively form zip-contacts, equals $\sim 56^\circ$ (for derivation, see Supporting Information). The S-DNA length at the critical tilt equals $L_{\text{cr}} = L_0/\cos(56^\circ) \approx 1.78L_0$, which is precisely the DNA length at the upper end of the 70 pN plateau. However, the pulling force growth during the S-zip transition would imply a nonlinear dependence of zip-DNA energy on the number of nucleobases converted to the zip-conformation. The physical basis of such nonlinear dependence is not apparent.

Many studies suggest that the 70 pN plateau is indicative of force-induced DNA melting rather than of a structural transition.^{2,12,14,37,43,45,46,70} Our large-force simulations indicate that even a force-melted double-stranded DNA has a strong propensity to reassemble into a zip-form, thereby supporting the structural transition view of forced DNA extension. Indeed, a recent study of DNA extension kinetics²² found that the DNA extension begins with a fast transition to an elongated state (presumed by the authors of ref 22 to be S-DNA), followed by a slow unpeeling of the strands off of each other. Other experiments show that a transition plateau in the force–extension curve is seen not only for double-stranded but also for *single-stranded* DNA¹⁵ (albeit at a smaller force of ~ 23 pN rather than at 70 pN). There

is no duplex to melt in the single-stranded DNA; however, the backbone transition to the elongated all-trans state should occur under stress in both the single- and double-stranded DNA.

The forced melting and the B–S–zip transition views are not necessarily incompatible.²² Many factors may affect the structural transitions of DNA during forced extension, including the temperature,^{2,14,22,37,43} the pulling speed,^{22,38} the stiffness of the AFM cantilever or the optical trap,^{38,39} and the ionic strength of the solvent.^{2,22,37–39,71} It might be possible for DNA to skip the S/zip phase and undergo melting while passing through the 70 pN plateau, if the experimental conditions favor melting.^{22,37,45} Under conditions less favorable for melting, the DNA could first undergo the B–zip structural transition, and melt afterward. That would be compatible with *two* transition plateaus in the DNA melting curves, observed in many experiments (Figure 1 and refs 3 and 38–40).

In fact, it is even possible that forced melting may take place *prior* to the DNA reassembly into the zip-stack. Stabilization of double-stranded DNA near the melting point by “moderate” forces was an interesting prediction by Rouzina and Bloomfield.⁴³ It is conceivable that a similar stabilization mechanism could apply to a force-melted DNA. The stretching force would limit the conformational space available to each of the two separated DNA strands and hold the tightly extended strands next to each other, thereby promoting the reassembly of double-stranded DNA. Considering the extended conformations of each DNA strand, the double-stranded DNA would adapt a zip-conformation, rather than a B-conformation.

The height of the plateau in the force–extension curve that corresponds to the critical force of DNA overextension is known to decrease with temperature.^{2,14,37,43} That observation is consistent with the B–zip transition. The critical force is determined by the free energy of the transition state between B- and zip-DNA phases, which corresponds to either S-DNA or melted DNA. Both S-DNA and melted DNA exhibit a significantly larger degree of fluctuations than B-DNA, as seen in the simulations here and elsewhere.^{47,49,50,66} Thus, the difference in entropy $\Delta S_{B-S/melt}$ is positive and both the energy of the transition state $\Delta G_{B-S/melt} = \Delta H_{B-S/melt} - T\Delta S_{B-S/melt}$, and the critical pulling force should decrease with temperature. (We assume that the enthalpy $\Delta H_{B-S/melt}$ is less affected by the temperature than the entropy term $T\Delta S_{B-S/melt}$.^{43,45}) Interestingly, the entropy difference $\Delta S_{zip-S/melt}$ between S-DNA and zip-DNA should also be positive because of a smaller number of degrees of freedom accessible to zip-DNA compared to S-DNA. Therefore, higher temperatures should reduce the energy difference between zip-DNA and S-DNA, destabilizing zip-DNA. Such zip-DNA destabilization by temperature could be tested in the proposed experiments seeking to detect zip-DNA (*vide infra*).

Other Experimental Findings Explained by Zip-DNA

Our calculations show that the nearest-neighbor nucleobase electronic coupling in zip-DNA exceeds that in B-DNA (Table 4). The strong coupling should result in a measurable conductivity of zip-DNA that could be observed in conductive AFM or other experiments. Such experiments monitor changes in the current through a polymer as it is stretched between an AFM tip and an electrode.⁷² Experiments of Cohen et al.²⁶ showed that the current through a 26-base DNA strand drops sharply upon stretching, apparently reflecting disrupted base stacking and increased base-to-base distances. However, the downward course of the current–distance curve is interrupted by several peaks. The first peak at $\sim 1.2L_0$ was suggested to result from an increase in base-to-base coupling caused by the initial unwinding of B-DNA.⁷³ The actual origin of that peak could be more subtle, as a significant base tilting accompanies unwinding during the B–S transition.^{41,42,47} The second, lower peak at $\sim 1.9L_0$ may indicate the formation of zip-DNA with other DNA strands within the

monolayer, where the ordering of the bases may cause a temporary increase in conductivity after a more disordered S-DNA phase.

Another recent experiment⁷⁰ demonstrates that DNA stretching (i) suppresses DNA staining by the fluorescent dye YOYO⁷⁴ and (ii) induces DNA binding by a single-stranded DNA marker, replication protein A (RPA).⁷⁵ The authors of ref 70 attribute both of these effects to force-induced DNA melting, yet both effects can be explained by a B-zip transition as well. It clearly may be difficult for the intercalator dye YOYO to bind to zip-DNA because the dye needs to compete for the intercalation space between nucleobases on the same DNA strand with the bases of the other DNA strand. Moreover, the DNA backbone in the zip form is already stretched, so the dye cannot push the nucleobases apart even further and intercalate into the zip stack. On the other hand, the stretched backbone of zip-DNA provides a good binding target for RPA. Indeed, the structure of the RPA–DNA complex shows that RPA binds to a stretched DNA backbone with interphosphate distances of 5.4–6.3 Å and has minimal interactions with nucleobases.⁷⁵

Computational Recipe for Producing zip-DNA

Except for refs 52 and 66 and the present study, no other modeling studies of DNA stretching have observed a B-zip or an S-zip transition under applied force. We attribute the absence of such transitions to differences among the computational protocols. On the basis of our study and those from refs 52 and 66, zip-DNA can be produced computationally if (i) the DNA strands are pulled for a sufficiently long time, (ii) the two DNA strands are pulled in the same direction rather than in the opposite directions, and (iii) the transition to zip-DNA is not constrained. Below, we elaborate on these points.

a. Simulation Time—Some of the earlier simulations might simply be insufficiently long to produce an S-zip transition. For example, the B–S transition in ref 42 occurred at $F \sim 100$ pN and zip-DNA nucleation sites appeared in the S-DNA stack. However, the small-force part of that simulation lasted for only several hundred picoseconds (albeit in implicit solvent), and the subsequent large forces separated the DNA strands without giving the zip-form a chance to spread over the entire DNA.

b. Pulling Mode—In most simulations,^{41,42,49,50} the pulling force was applied either to the 5' or to the 3' termini of each DNA strand, effectively pulling the DNA strands past each other. In contrast, the termini of the DNA strands in refs 52 and 66, and in this study, are constrained on one end of the DNA segment and are pulled in the same direction on the other end. Therefore, the DNA strands effectively remain aligned. Even a significant DNA disruption by a nanonewton-scale force does not lead to indefinite DNA strand separation, allowing the nucleic bases to relax into a zip-stack.

DNA pulling modes differ in experimental studies too. The early optical tweezers experiments often had the silicon beads attached to only one end of each DNA strand.^{1,5,6,8,76} Therefore, the DNA strands in those experiments were pulled past each other as in the simulations of refs 41, 42, 49, and 50. In more recent AFM experiments,^{15,38,39} the contacts between the DNA strands and the gold substrate/AFM tip are established randomly, resulting in different possible pulling modes. The DNA strands could be pulled past each other, as in refs 1, 5, 6, and 8 or stretched parallel to each other, as simulated here and in refs 52 and 66. A recent magnetic tweezers experiment probed different pulling modes directly;⁷⁶ another experiment employed a DNA hairpin in an optical trap pulling experiment, very closely reproducing the pulling conditions simulated here.⁴⁴

Should the simulation results be compared only to experiments with a similar pulling mode? We believe that the simulation-experiment comparison should not be limited by such

specifics. The time scale and system size of the present-day simulations are sufficiently remote from the experiment and therefore only comparisons at the most general level make sense. The simulations employ short DNA segments (tens of base pairs) and perform the DNA pulling on a nanosecond time scale. In contrast, the DNA in most experiments is long (thousands to tens of thousands of base pairs^{5,6,8,15,22,40}), and the experimental pulling time ranges from 10 μ s to 1 s.³ The DNA length and local structural variations allow the pulling stress to become evenly distributed between the DNA strands at a sufficient distance from the DNA ends, thereby mitigating the differences between the pulling modes. The pulling time scale allows the DNA to explore multiple conformations and to adopt energetically favorable conformations at any extension. Simulations, in contrast, tend to emphasize the effect of stress distribution for a specific pulling mode. For example, the short DNA strands pulled past each other would separate faster than the long experimental DNA strands. As a consequence, the energetically favorable DNA conformation can be missed. We therefore posit that the simulation setup should allow detecting the favorable DNA conformation *on the simulated time scale* rather than faithfully reproduce the experimental pulling mode. Following our analysis above, we propose that zip-DNA might be such a favorable conformation of extended DNA, at least beyond the critical length of $\sim 1.7L_0$.

It is also conceivable that the zip-DNA phase could follow the force-melted DNA phase, as discussed above. In that case, the pulling mode would become critical. The DNA strands pulled past each other should indeed separate following the forced melting and zip-DNA would form only transiently if the pulling is sufficiently slow. Maintaining a zip-DNA conformation for an extended period of time would demand stretching the DNA strands in parallel. Further studies are required to determine whether this scenario, or the one described in the previous paragraph, is valid.

c. Simulation Constraints—Finally, specific details of some of the earlier simulations could have precluded the formation of zip-DNA. For example, the DNA strands in ref 51 are effectively constrained on both ends by periodic boundary conditions that do not allow the DNA to unwind. A nick in the middle of the DNA segment does not alleviate this constraint as the periodic boundary conditions make the DNA on both sides of the nick essentially the same segment, which cannot rotate its two ends without imposing a prohibitively large twist. In refs 41 and 47, rupturing of the WC hydrogen bonds might have been impeded either by explicit hydrogen bond energy terms⁴¹ or by elastic energy terms that favor an “ideal” base pair structure.⁴⁷ In addition, the energy minimization procedures in refs ^{41,47} might have limited the scope of the conformational search compared to that achieved in MD simulations, thus making the free energy barrier crossing between S-DNA and zip-DNA unlikely.

In the present study, the DNA twist is not constrained and the strand-to-strand distances are allowed to change along most of the DNA, as the pulled termini of the two DNA strands are free to move in the plane perpendicular to F . The distance between the strands remains fixed only at the restrained end of the DNA. Consequently, the nucleobases at that end never adopt a zip-conformation.

The established protocol for obtaining zip-DNA structures *in silico* enables further modeling studies. Such studies could establish the free energy profile for the B–S–zip transition and could attempt to reproduce the dependence of zip-DNA length on the pulling force. Matching the computed force–extension profile with the experimental force–extension curve for DNA could provide a “smoking gun” for the proposed B–S–zip transition.

Experimental Tests for Zip-DNA Formation

An experimental probe for the existence of zip-DNA could be based on one of its most striking structural features, namely, the possible **noncomplementarity** of the two DNA strands. Indeed, the base-to-base interactions in zip-DNA consist solely of stacking interactions; thus no WC complementarity between the bases on the opposite strands is required. Demonstrating a stable complex between two noncomplementary DNA strands would be a persuasive argument in favor of the existence of a zip-form of DNA.

A possible experimental test for the existence of zip-DNA could employ Förster resonance energy transfer (FRET)⁷⁷ within the zip structure. The proposed experimental setup is illustrated in Figure 6A. A DNA strand carrying the donor fluorophore label near its 5' end could be captured by optical tweezers. Once that strand is captured, a noncomplementary DNA strand, carrying the acceptor fluorophore label near its 3' end, could be added to the solution. If a zip-complex between the two strands can indeed be formed, then stretching the captured DNA molecule to $2L_0$ should trigger the formation of such a complex. The formation of the complex should be facilitated by low ionic strength of the solvent due to the reduced electrostatic energy difference between B-DNA and zip-DNA (vide supra). The formation of the complex would bring the two fluorophore labels within a quenching distance resulting in a FRET signal.

Conducting AFM experiments²⁶ could provide another test for the existence of zip-DNA (Figure 6C). A single-stranded DNA captured between the tip of a conducting AFM and a gold surface can carry electric current of tens of nanoamperes if stretched to its normal length L_0 (see ref 26 for the details of experimental setup). Stretching the DNA strand to the length of $2L_0$ turns off the current due to destacking of the bases and an increase in the base-to-base distance. However, a formation of the zip-complex between the stretched DNA strand and another DNA strand should partially restore the current due to increased base–base couplings. In fact, if our suggestion that the formation of the zip-DNA complex within the DNA monolayer is responsible for the peaks on the current–distance curve of stretched DNA²⁶ (possibly, with the exception of the first peak, vide supra), then we predict that the peaks will disappear as the density of the monolayer decreases. The increased separation between the DNA strands in a low-density monolayer should preclude the possibility of formation of a zip-DNA complex.

We further suggest that in the case of a low-density monolayer, a DNA strand noncomplementary to the strands of the mono-layer could be added to the solution. If a zip-DNA structure is formed between the added and the stretched DNA strands, then adding the new DNA strand should partially restore the current through the stretched DNA. The restored current would not be as high as the initial current through the unstretched DNA because the zip-DNA would possess twice the number of base–base steps. Nonetheless, we expect a current in the nanoampere range, similar to the current in the second peak of the current distance curve.²⁶ A zip-complex between two all-purine strands should produce the largest current among the possible DNA sequences due to the lowest energy of the hole states on the purine bases.^{27,32–34}

Admittedly, the success of the two experiments described above will rely on the ability of the free DNA strand to spontaneously undergo a B–zip transition in the vicinity of the stretched DNA. Local transient openings of the DNA helix producing zip-type DNA steps were indeed observed in MD simulations of single-stranded DNA (Balaeff et al., unpublished data). Yet it is unclear how realistic or frequent such openings are.

If the free energy barrier between B- and zip-conformations is too large, then the proposed experiments can be modified by replacing a continuous free DNA strand with mixtures of

tri-, di-, or even mononucleotides. The intercalation of the bases of the short oligonucleotides between the bases of the stretched DNA strand should, in principle, stabilize the zip-conformation. However, entropic penalties and competition for binding sites on the stretched DNA could come to disfavor zip-DNA formation.

Another alternative would be to replace the combination of the bound and the free DNA strands with a DNA hairpin containing extended mismatched regions (Figure 6B,D). Pulling “handles” would be embedded in the middle and at the end of the hairpin as done in a recent study of Paik and Perkins.⁴⁴ Converting mismatched sections of the hairpin to a zip-stack should increase the hairpin conductivity in the conducting AFM experiment (Figure 6D). In the FRET experiment, the fluorescent labels could be embedded in the mismatched section of the hairpin (Figure 6B). Bringing the mismatched sections together in a zip-stack should turn on the FRET signal.

Possible Applications of Zip-DNA

If the existence of zip-DNA structures is confirmed, then its unusual structural and electronic properties could find applications in the design of DNA-based nanoscale circuits. A multitude of recently developed technologies direct DNA self-assembly into complex 2D and 3D constructs.²⁰ Using DNA conducting properties for converting those constructs into integrated circuits is a tantalizing goal.^{23,78} Such circuits would find applications in future electronics, biomedical diagnostics, materials design, and catalysis.^{20,21,23}

Natural DNA conductance might be insufficient for some electronics applications and may need to be increased, for example, by the integration of metal ions.²³ Zip-DNA, if confirmed, could become an alternative option for generating enhanced-conductivity “wire” sections of DNA nanocircuits arising from the increased electronic coupling between zip-DNA nucleobases (Table 4). Interestingly, manipulating the zip-DNA sequence could either significantly increase or significantly decrease the conductivity of zip-DNA compared to the conductivity of B-DNA (vide infra). Decreased-conductivity sections of zip-DNA could be used as low-conductance “isolator” elements holding the DNA constructs together mechanically without interfering with charge conductance along the designed paths.

A high-conductance zip-DNA could be engineered of two noncomplementary all-purine strands (e.g., G_nG_n). The nearest-neighbor G–G electronic coupling in such DNA is at least 2.6 times larger than in B-DNA (Table 4). On the basis of the McConnell model,^{27,35,79} the tunneling conductance of the G_{2n} zip-DNA would be $2.6^{8n} \approx 10^{3.3n}$ higher than the conductance of B-DNA of equivalent length (a $G_{2n}C_{2n}$ duplex).

A low-conductance zip-DNA could be engineered of two noncomplementary sequences with low purine content (e.g., GT_nT_nG). Note that a B-DNA duplex composed of two complementary strands has a purine base at every step. The energies of the hole states on pyrimidines are ~0.4 eV higher than those on purines.³³ Therefore, the G–G hole transfer rate through the $GT_{2n}G$ zip-DNA in the hopping regime^{24,27,30,31,33} would be $\exp[-0.4 \text{ eV}/kT] \approx 4.8 \times 10^6$ slower than through B-DNA of equivalent length/structure ($GT_{2n}G:CA_{2n}C$). In the McConnell superexchange regime, the rate of G–G hole transfer through the $GT_{2n}G$ zip-DNA would be $(0.6 \text{ eV}/0.2 \text{ eV})^{2n} = 3^{2n}$ times slower than through the $GT_{2n}G:CA_{2n}C$ B-DNA. Here, 0.6 and 0.2 eV are taken as the differences in the hole energies between G and T and G and A, respectively.³³

Another application of zip-DNA may be found in smart stress-sensing materials. DNA hairpins, like the ones shown in Figure 6, could be attached to nanoscale electrodes and incorporated into the materials of interest. An applied stress would convert the hairpin DNA

into larger-conductance zip-DNA (cf. Figure 6D), resulting in an electric signal in response to an applied stress.

CONCLUSIONS

In conclusion, we have used steered molecular dynamics simulations to demonstrate a B-DNA transition to a novel DNA form, zip-DNA, using a broad range of pulling forces. The zip-DNA form, first predicted by Lohikoski et al., is highly unusual: the nucleobases from the opposite DNA strands interdigitate with each other, forming a single-nucleobase aromatic stack. The analysis of the simulated B-zip transition indicates that the earlier predicted S-DNA structure may be an intermediate formed during the B-zip transition, and the well-known 70 pN plateau of the DNA force–extension curve, often attributed to the B–S transition, could correspond to the B-zip transition. The zip-DNA structure is also consistent with the alternative theory that ascribes the 70 pN plateau to forced melting of DNA: our simulations indicate that force-melted DNA strands may self-assemble into the zip-form once extended. The analysis of the electronic structure of zip-DNA reveals that the nearest-base purine–purine electronic coupling in zip-DNA is several times stronger than that in B-DNA. The strong purine–purine coupling, aided by the fact that the two zip-DNA strands do not have to be complementary, may lead to future zip-DNA applications in the rapidly developing field of DNA-based nanoelectronics. Several experimental tests for the existence of zip-DNA are suggested. The tests seek a FRET or conductance signal indicating the formation of the zip-complex between two noncomplementary DNA strands, or a zip conversion of a DNA hairpin.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We gratefully acknowledge financial support from the National Institutes of Health Grant GM-048043 (A.B. and D.N.B.), the National Science Foundation CRC Grant CHE-0628218 (A.B. and D.N.B.), and the National Science Foundation CDI Grant CBET-0835794 (S.L.C.). We thank Professor Jeffrey R. Reimers for the CNDO computer code. The computational facilities were provided in part by Duke Shared Cluster Resource. We thank Ravi Venkatramani, Shahar Keinan, Ron Naaman, Catalina Achim, and David Waldeck for stimulating discussions. We thank David W. Pratt for two decades of friendship, advice, and encouragement.

REFERENCES

1. Smith SB, Finzi L, Bustamante C. *Science*. 1992; 258:1122. [PubMed: 1439819]
2. Williams MC, Rouzina I. *Curr. Opin. Struct. Biol.* 2002; 12:330. [PubMed: 12127451]
3. Clausen-Schaumann H, Seitz M, Krautbauer R, Gaub HE. *Curr. Opin. Chem. Biol.* 2000; 4:524. [PubMed: 11006539]
4. Bustamante C, Chemla YR, Forde NR, Izhaky D. *Annu. Rev. Biochem.* 2004; 73:705. [PubMed: 15189157]
5. Cluzel P, Lebrun A, Heller C, Lavery R, Viovy JL, Chatenay D, Caron F. *Science*. 1996; 271:792. [PubMed: 8628993]
6. Smith SB, Cui YJ, Bustamante C. *Science*. 1996; 271:795. [PubMed: 8628994]
7. Strick TR, Allemand JF, Bensimon D, Croquette V. *Annu. Rev. Biophys. Biomol. Struct.* 2000; 29:523. [PubMed: 10940258]
8. Bustamante C, Marko JF, Siggia ED, Smith S. *Science*. 1994; 265:1599. [PubMed: 8079175]
9. Davenport RJ, Wuite GJL, Landick R, Bustamante C. *Science*. 2000; 287:2497. [PubMed: 10741971]

10. Cheng W, Dumont S, Tinoco I, Bustamante C. *Proc. Natl. Acad. Sci. U.S.A.* 2007; 104:13954. [PubMed: 17709749]
11. Lionnet T, Joubaud S, Lavery R, Bensimon D, Croquette V. *Phys. Rev. Lett.* 2006; 96:178102. [PubMed: 16712339]
12. Chaurasiya KR, Paramanathan T, McCauley MJ, Williams MC. *Phys. Life Rev.* 2010; 7:299. [PubMed: 20576476]
13. Bockelmann U, Thomen P, Essevaz-Roulet B, Viasnoff V, Heslot F. *Biophys. J.* 2002; 82:1537. [PubMed: 11867467]
14. Williams MC, Wenner JR, Rouzina I, Bloomfield VA. *Biophys. J.* 2001; 80:1932. [PubMed: 11259306]
15. Ke C, Humeniuk M, S-Gracz H, Marszalek PE. *Phys. Rev. Lett.* 2007; 99:018302. [PubMed: 17678193]
16. Paramanathan T, Westerlund F, McCauley MJ, Rouzina I, Lincoln P, Williams MC. *J. Am. Chem. Soc.* 2008; 130:3752. [PubMed: 18311981]
17. Vladescu ID, McCauley MJ, Rouzina I, Williams MC. *Phys. Rev. Lett.* 2005; 95:158102. [PubMed: 16241765]
18. Kwon YW, Lee CH, Choi DH, Jin JI. *J. Mater. Chem.* 2009; 19:1353.
19. Taniguchi M, Kawai T. *Phys. E (Amsterdam)*. 2006; 33:1.
20. Lo PK, Metera KL, Sleiman HF. *Curr. Opin. Chem. Biol.* 2011; 14:597. [PubMed: 20869905]
21. Yang DY, Campolongo MJ, Tran TNN, Ruiz RCH, Kahn JS, Luo D. *Wiley Interdiscip. Rev.: Nanomed. Nanobiotechnol.* 2010; 2:648. [PubMed: 20730873]
22. Fu H, Chen H, Marko JF, Yan J. *Nucleic Acids Res.* 2010; 38:5594. [PubMed: 20435680]
23. Clever GH, Shionoya M. *Coord. Chem. Rev.* 2010; 254:2391.
24. Venkatramani R, Keinan S, Balaeff A, Beratan DN. *Coord. Chem. Rev.* 2011; 255:635. [PubMed: 21528017]
25. Priyadarshy S, Risser SM, Beratan DN. *J. Phys. Chem.* 1996; 100:17678.
26. Cohen H, Nogues C, Naaman R, Porath D. *Proc. Natl. Acad. Sci. U.S.A.* 2005; 102:11589. [PubMed: 16087871]
27. Schuster, GB., editor. *Long Range Charge Transfer in DNA I and II*. Springer; Berlin and New York: 2004. Topics in Current Chemistry 236 and 237.
28. Wolfgang J, Risser SM, Priyadarshy S, Beratan DN. *J. Phys. Chem. B.* 1997; 101:2986.
29. Skourtis SS, Waldeck DH, Beratan DN. *Annu. Rev. Phys. Chem.* 2010; 61:461. [PubMed: 20192814]
30. Hatcher E, Balaeff A, Keinan S, Venkatramani R, Beratan DN. *J. Am. Chem. Soc.* 2008; 130:11752. [PubMed: 18693722]
31. Grozema FC, Tonzani S, Berlin YA, Schatz GC, Siebbeles LDA, Ratner MA. *J. Am. Chem. Soc.* 2008; 130:5157. [PubMed: 18324767]
32. Senthilkumar K, Grozema FC, Guerra CF, Bickelhaupt FM, Lewis FD, Berlin YA, Ratner MA, Siebbeles LDA. *J. Am. Chem. Soc.* 2005; 127:14894. [PubMed: 16231945]
33. Endres RG, Cox DL, Singh RRP. *Rev. Mod. Phys.* 2004; 76:195.
34. Beratan DN, Skourtis SS, Balabin IA, Balaeff A, Keinan S, Venkatramani R, Xiao DQ. *Acc. Chem. Res.* 2009; 42:1669. [PubMed: 19645446]
35. Paul A, Bezer S, Venkatramani R, Kocsis L, Wierzbinski E, Balaeff A, Keinan S, Beratan DN, Achim C, Waldeck DH. *J. Am. Chem. Soc.* 2009; 131:6498. [PubMed: 19382798]
36. Keinan S, Venkatramani R, Balaeff A, Beratan DN. *J. Phys. Chem. C.* 2010; 114:20496.
37. Williams MC, Rouzina I, Bloomfield VA. *Acc. Chem. Res.* 2002; 35:159. [PubMed: 11900519]
38. Rief M, Clausen-Schaumann H, Gaub HE. *Nat. Struct. Biol.* 1999; 6:346. [PubMed: 10201403]
39. Calderon CP, Chen WH, Lin KJ, Harris NC, Kiang CH. *J. Phys.: Condens. Matter.* 2009; 21:034114.
40. Clausen-Schaumann H, Rief M, Tolksdorf C, Gaub HE. *Biophys. J.* 2000; 78:1997. [PubMed: 10733978]
41. Lebrun A, Lavery R. *Nucleic Acid Res.* 1996; 24:2260. [PubMed: 8710494]

42. Konrad MW, Bolonick JI. *J. Am. Chem. Soc.* 1996; 118:10989.
43. Rouzina I, Bloomfield VA. *Biophys. J.* 2001; 80:894. [PubMed: 11159456]
44. Paik DH, Perkins TT. *J. Am. Chem. Soc.* 2011; 133:3219. [PubMed: 21207940]
45. Rouzina I, Bloomfield VA. *Biophys. J.* 2001; 80:882. [PubMed: 11159455]
46. Williams MC, Wenner JR, Rouzina L, Bloomfield VA. *Biophys. J.* 2001; 80:874. [PubMed: 11159454]
47. Kosikov KM, Gorin AA, Zhurkin VB, Olson WK. *J. Mol. Biol.* 1999; 289:1301. [PubMed: 10373369]
48. Lebrun A, Lavery R. *J. Biomol. Struct. Dyn.* 1998; 16:593. [PubMed: 10052616]
49. Harris SA, Sands ZA, Laughton CA. *Biophys. J.* 2005; 88:1684. [PubMed: 15626714]
50. Piana S. *Nucleic Acid Res.* 2005; 33:7029. [PubMed: 16356925]
51. Luan B, Aksimentiev A. *Phys. Rev. Lett.* 2008; 101:118101. [PubMed: 18851334]
52. Lohikoski R, Timonen J, Laaksonen A. *Chem. Phys. Lett.* 2005; 407:23.
53. Sunami T, Kondo J, Hirao I, Watanabe K, Miura K, Takenaka A. *Acta Crystallogr., Sect. D.* 2004; 60:422. [PubMed: 14993665]
54. Shepard W, Cruse WBT, Fourme R, de la Fortelle E, Prange T. *Structure.* 1998; 6:849. [PubMed: 9687367]
55. Chen LQ, Cai L, Zhang XH, Rich A. *Biochemistry.* 1994; 33:13540. [PubMed: 7947764]
56. Isralewitz B, Gao M, Schulten K. *Curr. Opin. Struct. Biol.* 2001; 11:224. [PubMed: 11297932]
57. Sotomayor M, Schulten K. *Science.* 2007; 316:1144. [PubMed: 17525328]
58. Kersey FR, Lee G, Marszalek P, Craig SL. *J. Am. Chem. Soc.* 2004; 126:3038. [PubMed: 15012119]
59. Munteanu MG, Vlahovicek K, Parthasarathy S, Simon I, Pongor S. *Trends Biochem. Sci.* 1998; 23:341. [PubMed: 9787640]
60. Phillips JC, Braun R, Wang W, Gumbart J, Tajkhorshid E, Villa E, Chipot C, Skeel RD, Kalé L, Schulten K. *J. Comput. Chem.* 2005; 26:1781. [PubMed: 16222654]
61. MacKerell AD, Banavali NK. *J. Comput. Chem.* 2000; 21:105.
62. Harris SA. *Philos. Trans. R. Soc., A.* 2006; 364:3319.
63. Humphrey W, Dalke A, Schulten K. *J. Mol. Graphics.* 1996; 14:33.
64. Ridley J, Zerner M. *Theor. Chim. Acta.* 1973; 32:111.
65. Tejerina B, Reimers J. *CNDO/INDO.* 2008
66. Santosh M, Maiti PK. *J. Phys.: Condens. Matter.* 2009; 21:034113. [PubMed: 21817258]
67. Klosterman JK, Yamauchi Y, Fujita M. *Chem. Soc. Rev.* 2009; 38:1714. [PubMed: 19587964]
68. Gabriel GJ, Iverson BL. *J. Am. Chem. Soc.* 2002; 124:15174. [PubMed: 12487583]
69. Spackova N, Berger I, Sponer J. *J. Am. Chem. Soc.* 2000; 122:7564.
70. van Mameren J, Gross P, Farge G, Hooijman P, Modesti M, Falkenberg M, Wuite GJL, Peterman EJG. *Proc. Natl. Acad. Sci. U.S.A.* 2009; 106:18231. [PubMed: 19841258]
71. Wenner JR, Williams MC, Rouzina I, Bloomfield VA. *Biophys. J.* 2002; 82:3160. [PubMed: 12023240]
72. Chen F, Hihath J, Huang ZF, Li XL, Tao NJ. *Annu. Rev. Phys. Chem.* 2007; 58:535. [PubMed: 17134372]
73. Song B, Elstner M, Cuniberti G. *Nano Lett.* 2008; 8:3217. [PubMed: 18759487]
74. Rye HS, Yue S, Wemmer DE, Quesada MA, Haugland RP, Mathies RA, Glazer AN. *Nucleic Acid Res.* 1992; 20:2803. [PubMed: 1614866]
75. Bochkarev A, Pfuetzner RA, Edwards AM, Frappier L. *Nature.* 1997; 385:176. [PubMed: 8990123]
76. Danilowicz C, Limouse C, Hatch K, Conover A, Coljee VW, Kleckner N, Prentiss M. *Proc. Natl. Acad. Sci. U.S.A.* 2009; 106:13196. [PubMed: 19666582]
77. Joo C, Balci H, Ishitsuka Y, Buranachai C, Ha T. *Annu. Rev. Biochem.* 2008; 77:51. [PubMed: 18412538]
78. Slinker JD, Muren NB, Renfrew SE, Barton JK. *Nat. Chem.* 2011; 3. [PubMed: 21160501]

79. McConnell HM. J. Chem. Phys. 1961; 35:508.

Figure 1.

Experimental DNA force–extension curves, adapted from (A) ref 38, (B) ref 5, and (C) ref 39. Labeled sections of the curve in (A) are (1) the entropic elasticity plateau, (2) elastic extension of B-DNA, (3) the B–S transition plateau, (4) elastic extension of S-DNA, (5) the DNA melting plateau, and (6) elastic extension of single-stranded (melted) DNA.

(A) Antiparallel strands in the B-DNA form: top, strand A, bottom, strand B.

3'- G_{1A} G_{2A} T_{3A} A_{4A} T_{5A} A_{6A} C_{7A} C_{8A} G_{9A} C_{10A} T_{11A} T_{12A} A_{13A} A_{14A} G_{15A} C_{16A} -5'

5'- C_{16B} C_{15B} A_{14B} T_{13B} A_{12B} T_{11B} G_{10B} G_{9B} C_{8B} G_{7B} A_{6B} A_{5B} T_{4B} T_{3B} C_{2B} G_{1B} -3'

(B) Interlocked strands in the zip-DNA form

G_{1A}C_{16B}G_{2A}C_{15B}T_{3A}A_{14B}A_{4A}T_{13B}T_{5A}A_{12B}A_{6A}T_{11B}C_{7A}G_{10B}C_{8A}G_{9B}G_{9A}C_{8B}C_{10A}G_{7B}T_{11A}-

-A_{6B}T_{12A}A_{5B}A_{13A}T_{4B}A_{14A}T_{3B}G_{15A}C_{2B}C_{16A}G_{1B}

Scheme 1.

Base Sequence of the Simulated DNA

Figure 2.

(A) The initial B-form adopted by DNA at the beginning of the SMD simulations. The restrained atoms are shown as dark colored spheres and the pulled atoms are shown as light colored spheres. The arrows indicate the direction of the pulling forces. (B) The intermediate melted DNA conformation adapted by the DNA during the transition from B- to zip-conformation in an SMD simulation with $F = 11.2$ nN (SMD-H6). The view is rotated sideways with respect to that in (A). The snapshot was taken 0.391 ns after the simulation began. Note the loss of base stacking and Watson–Crick base pairing by the nucleobases along the DNA, as well as the initial formation of zip-DNA contacts at the restrained end of the DNA. (C) The dihedral rotations in the phosphodiester backbone that are responsible for the B-zip DNA transformation. (D) Zip-DNA structures resulting from the large-force SMD simulations. Final snapshots from the simulations are shown. The large forces convert the entire DNA segment into the zip-form during the 5–10 ns simulation time.

Figure 3.

Small-force structures of extended DNA. Final snapshots from the small-force SMD runs are shown. The structures contain a mixture of DNA conformations because the simulation times of 20–50 ns were not sufficient to convert the entire DNA segment into the zip-form. The 100 pN/40 ns simulations SMD8, SMD9, and SMD10 end up in a mixed B/S/zip conformation. The other simulations ($F = 130\text{--}240$ pN) end up in a mixed S/zip conformation, with a high percentage of nucleobases in zip-conformation and most of the remaining S-conformation bases highly disordered.

Figure 4.

Dynamics of the simulated B-zip transition. (A) Extension L/L_0 of the 16 bp DNA segment in the large-force SMD simulations, where L is the instant length and $L_0 = 47.6$ Å is the B-DNA length between the anchored and the pulled atoms. Inset: most of the extension occurs during the early phase of the SMD simulation. (B) Rmsd of the DNA structure from zip-DNA in the large-force SMD simulations. The rmsd quickly drops during the initial rapid extension of the DNA backbone and then gradually converges to ~ 1 Å as the nucleobases assemble into a zip-stack. (C) Energy dynamics in a representative 5.6 nN SMD run (SMD-H1). The curves represent a 200 ps running average of each energy component. The conformational energy is the sum of the chemical bond, angle, dihedral, and improper energy terms. Base line energy E_0 for each energy component is the average value of that component during the 0.5 ns water equilibration that preceded the SMD runs (see Methods and Table 2). (D) DNA extension L/L_0 in the small-force SMD simulations. (E) Rmsd of the DNA structure from zip-DNA in the small-force SMD simulations. The rmsd values may illustrate not only the DNA structural dynamics under applied force but also the force-dependence of the zip-DNA structure, as discussed in the paper. (F) Energy dynamics in a representative 200 pN SMD run (SMD7).

Figure 5. Distribution of electronic coupling between neighboring purine bases in zip-DNA and B-DNA. The distributions are shown as normalized histograms on a logarithmic scale. See Methods for the list of nucleobases used to compute each ensemble of coupling values.

Figure 6.

Suggested experimental tests for the existence of zip-DNA. (A) A FRET-based experiment probes for binding between two noncomplementary DNA strands, one solvated and the other captured by optical tweezers. The green diamond indicates the donor fluorophore label; the white diamond indicates the acceptor fluorophore label; the flash indicates the FRET emission from the acceptor when the two labels are brought into proximity. (B) A modification of (A) based on a DNA hairpin. The hairpin is presumed to rotate freely around the bonds that bind it to the optical tweezers.⁴⁴ (C) A conducting AFM experiment looks for an increase in the conductance of a stretched DNA strand that should occur upon the binding of another DNA strand in zip-conformation. The DNA conductance should be measurable when the DNA is not stretched (left), should vanish when the DNA is stretched and disordered (middle), and should be partially restored in the zip-DNA state (right). (D) A modification of (C) based on a DNA hairpin. Stretching the hairpin should increase the electric current.

Table 1

The List of SMD Simulations Performed

code name	F , pN	simulation time, ns	N_{zip}	a	$\langle L_{DNA}/L_o \rangle^{b,d}$	$\langle \text{rmsd} \rangle^{c,d}$, Å
					from zip-DNA ^c	from B-DNA ^d
Large Force Simulations						
SMD-H1	558	10.0	26	2.185 (0.005)	0.93 (0.11)	17.60 (0.08)
SMD-H2	558	10.0	26	2.186 (0.005)	0.88 (0.11)	17.62 (0.07)
SMD-H7	558	4.9	28	2.184 (0.005)	0.87 (0.12)	17.62 (0.08)
SMD-H4	1116	5.0	28	2.330 (0.005)	0.90 (0.10)	19.27 (0.07)
SMD-H5	1116	5.5	26	2.330 (0.005)	1.29 (0.21)	19.26 (0.08)
SMD-H6	1116	5.0	24	2.334 (0.005)	1.73 (0.45)	19.40 (0.08)
Small Force Simulations						
SMD8	100	40	6	1.41 (0.02)	6.13 (0.21)	9.69 (0.20)
SMD9	100	40	7	1.29 (0.02)	8.04 (0.18)	7.09 (0.19)
SMD10	100	40	4	1.21 (0.02)	10.02 (0.22)	5.02 (0.30)
SMD1	130	50	20	1.61 (0.02)	2.19 (0.46)	12.55 (0.16)
SMD2	130	50	14	1.58 (0.02)	4.13 (0.28)	11.49 (0.24)
SMD12	160	30	11	1.62 (0.02)	3.92 (0.14)	12.40 (0.18)
SMD13	160	40	13	1.62 (0.01)	3.51 (0.17)	11.79 (0.21)
SMD7	200	40	19	1.69 (0.01)	3.70 (0.12)	13.14 (0.18)
SMD11	200	20	19	1.70 (0.01)	3.72 (0.11)	13.02 (0.19)
SMD14	240	20	19	1.70 (0.01)	3.64 (0.15)	13.18 (0.15)
SMD15	240	20	14	1.68 (0.02)	4.15 (0.14)	12.89 (0.16)

^aNumber of bases in zip-conformation at the end of the simulation (out of 32 total). A nucleobase was considered to have a zip conformation if (i) the base did not form Watson–Crick bonds with another base; (ii) the base was stacked against a base from the opposite DNA strand against which it would be also stacked in the fully extended zip-DNA conformation; and (iii) the base was positioned between the two DNA strands and the base plane was roughly perpendicular to the DNA axis (i.e., the zip-contacts formed by rotated-out bases outside the main DNA stack did not count). Note that even after a complete transition of the entire DNA segment to zip-form only 26 bases are likely to adopt a stable zip-conformation. The terminal bases (G1A, G2A, C15B, C16B, C16A, G1B) are unlikely to adopt a zip-conformation due to the harmonic restraints imposed and thermal disorder (see Methods).

^b L_{DNA} is the DNA end-to-end length (see Methods), averaged over the last 500 ps of the simulation. $L_o = 47.6$ Å. Standard deviations are shown in brackets.

^cThe rmsd from zip-DNA and B-DNA is computed as described in Methods and averaged over the ensemble of the last 500 snapshots (0.5 ns) of each simulation. Standard deviations are shown in brackets.

^dThe averages of L_{DNA} and $rmsd$ computed for the large-force simulations characterize well-converged structural ensembles of zip-DNA. In contrast, the averages computed for the small-force simulations characterize structural ensembles in dynamic transition to a stable zip-DNA structure, presumably achievable in the long run. Cf. Figures 2, 3, and 4.

Table 2

DNA Energetics in the SMD Simulations^a

code name	conformational energy					electrostatic	van der Waals	total
	bond	angle	dihedral	improper	total			
baseline	8.7 (0.4)	23.5 (0.6)	31.9 (0.2)	0.5 (0.1)	64.6 (0.7)	-5465 (10)	489 (7)	-4911 (12)
Water Equilibration (0.5 ns)								
Large-Force SMD Simulations (energies with respect to the baseline above)								
SMD-H1	3.8 (0.6)	7.9 (0.8)	4.0 (0.3)	0.1 (0.1)	15.9 (1.0)	2.2 (10.0)	-0.8 (6.8)	17.3 (6.1)
SMD-H2	3.9 (0.6)	8.0 (0.8)	3.8 (0.3)	0.1 (0.1)	15.8 (1.0)	1.5 (9.2)	-0.6 (6.7)	16.7 (5.2)
SMD-H7	3.8 (0.6)	7.9 (0.8)	3.9 (0.3)	0.1 (0.1)	15.7 (1.0)	1.6 (9.7)	-0.9 (6.7)	16.4 (5.7)
SMD-H4	14.5 (0.8)	24.3 (1.0)	3.7 (0.3)	0.1 (0.1)	42.7 (1.4)	-2.5 (9.8)	-0.5 (6.8)	39.7 (6.0)
SMD-H5	14.5 (0.8)	24.3 (1.0)	3.9 (0.3)	0.1 (0.1)	42.8 (1.4)	-2.8 (9.9)	-0.4 (6.8)	39.5 (6.2)
SMD-H6	14.9 (0.8)	24.9 (1.0)	3.7 (0.3)	0.1 (0.1)	43.5 (1.4)	-0.4 (9.2)	-0.1 (6.7)	43.0 (5.3)
Small Force SMD Simulations (energies with respect to the baseline above)								
SMD8	0.5 (0.4)	0.8 (0.6)	0.6 (0.3)	0.1 (0.1)	1.9 (0.8)	4.3 (10.0)	-0.3 (6.9)	5.9 (6.0)
SMD9	0.5 (0.5)	0.8 (0.6)	0.6 (0.3)	0.1 (0.1)	1.9 (0.8)	3.1 (9.6)	0.0 (6.8)	5.0 (5.5)
SMD10	0.4 (0.5)	0.9 (0.7)	0.6 (0.3)	0.1 (0.1)	2.0 (0.9)	3.8 (10.9)	-0.4 (6.8)	5.4 (7.1)
SMD1	0.5 (0.5)	0.7 (0.6)	0.9 (0.3)	0.1 (0.1)	2.2 (0.8)	2.8 (9.5)	-0.1 (6.7)	5.0 (5.4)
SMD2	0.5 (0.5)	1.0 (0.6)	0.6 (0.3)	0.1 (0.1)	2.2 (0.8)	0.2 (9.6)	0.8 (6.8)	3.2 (5.6)
SMD12	0.5 (0.5)	0.8 (0.6)	0.8 (0.3)	0.1 (0.1)	2.1 (0.8)	1.7 (9.2)	0.3 (6.8)	4.1 (5.2)
SMD13	0.5 (0.5)	1.0 (0.7)	0.8 (0.3)	0.1 (0.1)	2.3 (0.8)	2.6 (9.4)	0.1 (6.8)	4.9 (5.4)
SMD7	0.5 (0.5)	0.6 (0.6)	0.9 (0.3)	0.1 (0.1)	2.1 (0.8)	2.3 (9.1)	0.1 (6.8)	4.5 (5.1)
SMD11	0.5 (0.5)	1.0 (0.6)	0.8 (0.3)	0.1 (0.1)	2.4 (0.8)	1.5 (9.1)	0.4 (6.8)	4.3 (4.9)
SMD14	0.5 (0.5)	0.7 (0.6)	0.8 (0.3)	0.1 (0.1)	2.1 (0.8)	1.9 (8.8)	0.1 (6.8)	4.0 (4.8)
SMD15	0.5 (0.5)	0.9 (0.6)	1.0 (0.4)	0.1 (0.1)	2.5 (0.8)	2.4 (9.6)	0.2 (6.8)	5.1 (5.6)

^aThe individual energy components and the total energies are averaged over the last 0.5 ns of each SMD simulation. Standard deviations are shown in brackets. The energies are measured in (kcal/mol)/base. The table presents relative energies computed with respect to the baseline values, which are defined as the averages of the corresponding energies from the 0.5 ns water equilibration that preceded the SMD runs (see Methods). Due to rounding errors, the numbers in the rows may not add up to exactly the listed totals.

Table 3

Electrostatic Energy in the SMD Simulations^a

code name	DNA, intrastrand	DNA, interstrand	DNA-solvent	solvent-solvent	total
Water Equilibration (0.5 ns)					
baseline	0.4 (0.4)	111.0 (0.3)	-632.5 (3.3)	-4587.4 (12.0)	-5108.6 (11.7)
Large-Force SMD Simulations (energies with respect to the baseline above)					
SMD-H1	-23.7 (0.7)	-31.9 (0.6)	113.3 (3.9)	-50.0 (12.7)	7.7 (12.3)
SMD-H2	-24.5 (0.7)	-29.9 (0.7)	108.7 (4.6)	-47.8 (11.8)	6.6 (11.5)
SMD-H7	-23.8 (0.7)	-31.2 (0.6)	113.0 (3.8)	-50.2 (12.2)	7.8 (11.9)
SMD-H4	-30.1 (0.6)	-32.9 (0.7)	121.1 (3.9)	-49.0 (11.8)	9.1 (11.7)
SMD-H5	-30.3 (0.7)	-35.8 (0.9)	129.5 (4.2)	-54.0 (13.3)	9.4 (12.7)
SMD-H6	-30.4 (0.7)	-35.8 (0.8)	128.5 (4.3)	-50.8 (11.8)	11.5 (11.9)
Small-Force SMD Simulations (energies with respect to the baseline above)					
SMD8	-5.2 (1.0)	-20.4 (0.9)	52.3 (4.5)	-24.8 (12.2)	2.1 (12.0)
SMD9	-4.1 (1.0)	-13.4 (1.1)	36.6 (4.6)	-13.8 (11.1)	5.2 (10.9)
SMD10	-1.4 (1.2)	-10.4 (0.8)	25.7 (4.4)	-10.3 (12.8)	3.7 (12.4)
SMD1	-6.1 (1.2)	-25.4 (0.8)	63.8 (5.0)	-33.4 (11.9)	-1.1 (11.4)
SMD2	-5.1 (1.0)	-26.6 (0.8)	64.7 (4.5)	-36.3 (11.5)	-3.3 (10.8)
SMD12	-6.4 (1.0)	-26.1 (0.7)	67.2 (4.1)	-38.9 (11.6)	-4.2 (11.3)
SMD13	-6.8 (1.0)	-25.5 (0.8)	66.0 (4.1)	-36.0 (11.5)	-2.2 (11.1)
SMD7	-10.2 (0.8)	-23.3 (0.6)	68.5 (3.7)	-38.3 (10.9)	-3.3 (10.8)
SMD11	-9.8 (1.0)	-22.2 (0.9)	65.2 (3.9)	-35.5 (11.0)	-2.4 (11.1)
SMD14	-9.5 (0.9)	-25.4 (0.8)	73.2 (3.8)	-41.4 (11.2)	-3.1 (11.3)
SMD15	-9.4 (1.1)	-21.0 (0.8)	61.7 (4.0)	-34.8 (11.5)	-3.6 (11.4)

^aThe electrostatic energies are computed as explained in Methods and averaged for the last 500 snapshots of each SMD simulation (corresponding to the last 0.5 ns of each simulation). Standard deviations are shown in brackets. The energies are measured in (kcal/mol)/base and computed with respect to the baseline values, which are the averages of the corresponding energies from the 0.5 ns water equilibration run that preceded the SMD runs (see Methods). Due to rounding errors, the numbers in the rows may not add up to exactly the listed totals.

Table 4

Average Electronic Couplings between Nearest-Neighbor Purine HOMOs in Zip- and B-DNA

		zip-DNA ($F = 5.6$ nN)	B-DNA, intrastrand	B-DNA, interstrand
GG	$\left(\langle V_{GG}^2 \rangle\right)^{1/2}$, eV	0.225	0.087	0.016
	$\langle V_{GG} \rangle$, eV	0.213	0.067	0.011
AA	$\left(\langle V_{AA}^2 \rangle\right)^{1/2}$, eV	0.204	0.107	0.051
	$\langle V_{AA} \rangle$, eV	0.189	0.086	0.039