THE JOURNAL OF PHYSICAL CHEMISTRY Letters

Letter

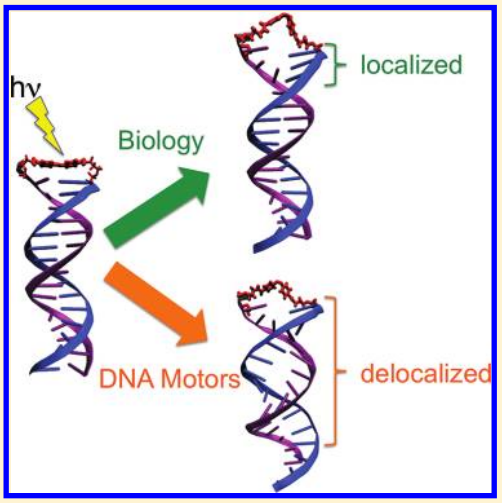
pubs.acs.org/JPCL

Defects in DNA: Lessons from Molecular Motor Design

Martin McCullagh,[‡] Ignacio Franco,[¶] Mark A. Ratner, and George C. Schatz*

Department of Chemistry, Northwestern University, Evanston, Illinois 60208-3113, United States

ABSTRACT: The degree of localization of structural damage in DNA is computationally investigated in the framework of molecular motors. Damage is induced on DNA hairpins with two and three guanine-cytosine (GC) base pairs by the photoinduced isomerization of their azobenzene containing cap. Light-induced changes in elasticity of such hairpins can be used to transduce photon energy into mechanical work in a single-molecule pulling setup through optomechanical cycles. The maximum extractable work per cycle is, in fact, a good measure of the degree of disruption of the hairpin structure upon isomerization. The extractable work was quantified by means of free-energy reconstruction techniques and several microseconds of molecular dynamics simulations. The maximum work that can be extracted from the 2GC and 3GC systems starting from their native B-DNA conformation ($d_{O3'-O5'} \approx 16$ Å) is 2.70 kcal/mol in both cases. The fact that the extractable work does not increase when transitioning from the dimer to the trimer implies that the DNA damage induced by azobenzene isomerization is localized to the two base pairs adjacent to the photoswitchable unit. From the perspective of DNA-based molecular motors, these findings indicate that a dense azobenzene arrangement would be required for effective actuation. From a biological perspective, the results highlight the remarkable ability of the DNA design to mitigate the propagation of damage, thus limiting detrimental effects that this may have on healthy cell function.



SECTION: Biophysical Chemistry

DNA damage can alter or eliminate the cell's ability to transcribe the gene that the affected DNA encodes.[1,2] In response, the cell has a sophisticated collection of processes by which DNA damage is identified and corrected, including nucleotide excision repair and direct repair by proteins such as DNA photolyase.[2] Here, we are concerned with one of the most basic of the protective mechanisms, that of the response of the DNA structure by itself to the presence of a defect that breaks a few base pairs. The question that we ask is simple but important: (in the absence of repair enzymes) how far along a double-stranded DNA sequence does a structural defect propagate?

The proposed problem is also of central interest in the design of DNA-based actuators.[3−10] In these systems, externally switchable chemical units are incorporated along the DNA architecture to be able to control the elastic, melting, or even gene expression properties of the DNA through application of external stimuli. Determining the density and distribution of the switchable molecules for optimal actuation requires knowledge of the number of base pairs over which the switchable defect will have an impact.

There are two possible hypothetical limiting behaviors for how the DNA architecture responds to the presence of defects. One possibility is that once a defect is introduced along the DNA architecture, this defect in the base pairing propagates along the structure in a manner analogous to crack propagation in solids.[11] This class of behavior would be ideal for molecular motor design because it would amplify the structural change induced by the external stimulus. A second possibility is that the DNA architecture exploits its inherent flexibility[12,13] to localize the effect of the defect to a given number of base pairs, thus mitigating any potential damage to the DNA strand, a behavior with evident biological advantages.

In this Letter, we quantitatively approach this problem from the perspective of molecular motors. As a specific system, we consider an azobenzene-capped DNA hairpin with a variable number of complementary guanine (G)-cytosine (C) base pairs (see Figure 1 for a scheme of the structures). Previously, we investigated the structure with two GC pairs as a prototype of an optomechanical molecular motor.[5] In these motors,[5,14,15] light is used to induce trans-to-cis isomerization of the photoactive azobenzene unit,[16] a process that disrupts the π-stacking and hydrogen-bonding network of the hairpin. Specifically, when the azobenzene is in its trans configuration, the hairpin cap is flat and the base pairing structure intact. By contrast, a defect is created in the hairpin structure when the azobenzene is in the cis conformation, making it thermodynamically less stable. Such disruption can be harnessed to extract net work by constructing analogues of thermodynamic cycles in a single-molecule pulling setup.[5,14]

Received: December 14, 2011
Accepted: February 21, 2012
Published: February 21, 2012

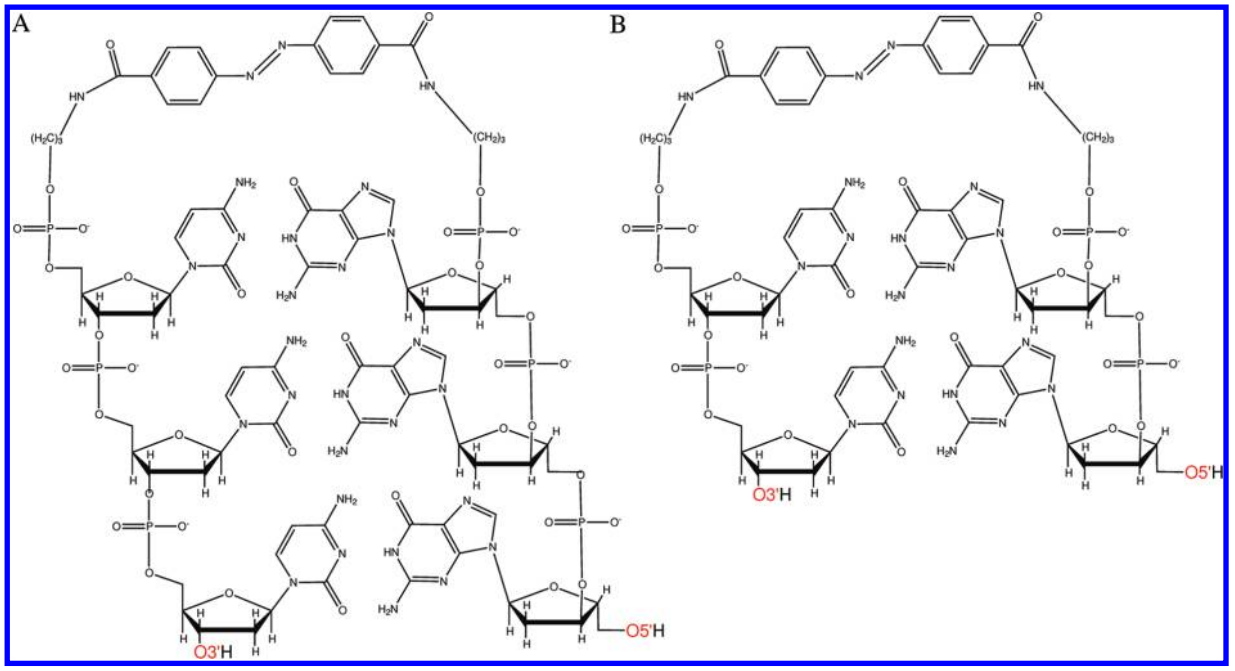


Figure 1. Schematics of the three GC (A) and two GC (B) azobenzene-capped DNA hairpins. The O3′ and O5′ atoms restrained during simulation to mimic AFM pulling experiments are highlighted in red.

Here, the degree of disruption of the hairpin structure induced by the defect is quantified by determining the changes in the potential of mean force (PMF) along the end-to-end distance coordinate ξ due to photoisomerization of the azobenzene moeity. The PMF is the Helmholtz free-energy profile. When determined along ξ, it summarizes thermodynamic changes during folding and captures the elastic properties observed in single-molecule pulling experiments.[17] The influence of the photoinduced kink on the DNA structure can be determined by investigating how the PMF changes upon isomerization for hairpins with varying numbers of base pairs. Previously,[5] we reconstructed the PMF of both the cis and trans forms for a 2GC structure. Here, we build upon those efforts and reconstruct the PMF of both isomers for a hairpin with three GC base pairs by means of molecular dynamics (MD) simulations and the weighted histogram analysis method (WHAM).[18,19] As detailed below, these four PMFs are sufficient to quantitatively address the question proposed above.

All MD simulations were performed using NAMD[20] at 300 K in the *NVT* ensemble with a Langevin thermostat with a damping coefficient of 5.0 ps^{-1}. The 3GC system was composed of 241 DNA hairpin atoms, 6 sodium ions, and 2543 TIP3P water molecules in a 32.64 Å × 37.13 Å × 64.16 Å box (yielding a solvent density of 1.011 g/mL). An integration time step of 2 fs was used in combination with the SHAKE algorithm. A real space cutoff of 12 Å and the particle mesh Ewald method were used to compute nonbonded interactions. The DNA hairpin was described using the CHARMM27 force field[21,22] and supplemented with additional parameters (that have been presented before[5]) required to describe the central nitrogens in the *cis*- and *trans*-azobenzene.

In order to reconstruct the PMF using WHAM, umbrella sampling was performed along the terminal O5′−O3′ (see Figure 1) molecular end-to-end coordinate ξ. For this, the terminal O5′ atom was restrained by a stiff isotropic harmonic potential, and simultaneously, the terminal O3′ atom was subjected to a harmonic bias with stiffness of k_0 = 1.10 N/m (1.58 kcal mol^{-1} $Å^{-2}$) along ξ and stiff in the perpendicular direction. Such harmonic bias was centered at a distance L away from the terminal O5′ atom along the ξ coordinate. After an initial energy minimization in vacuum, the system was placed in a pre-equilibrated solvent box. The energy of the solvated system was minimized and the system allowed to equilibrate for 4 ns. Subsequently, the dynamics was followed for 8 ns and the end-to-end distance ξ recorded every 1 ps. In total, 267 extensions for the cis and 271 extensions for the trans were simulated. Values of L ranged from 6.75 to 51.06 Å in both cases. This set of biased simulations was combined using the WHAM to extract the molecular PMF $\phi(\xi)$ along the end-to-end distance coordinate, as described in detail elsewhere.[17] In the WHAM procedure, we employed bins of 1 Å along ξ and used as a convergence criterion an average difference of 10^{-7} kcal mol^{-1} among consecutive estimates of the free energy in the self-consistent procedure. The total analyzed simulation time was 4.384 μs for the trans and 4.272 μs for the cis.

The PMF $\phi(\xi)$ of the dimer (top panel) and trimer (bottom panel) DNA hairpins are shown in Figure 2. Regions of convexity along the PMF signify mechanically stable conformations, while regions of concavity signify molecular unfolding events. Substantial changes in the structure of the PMF are induced by photoisomerization. Importantly, note that the effect of the photoisomerization in the dimer and the trimer is qualitatively different. While for the dimer the isomerization affects the free-energy profile for all sampled extensions, for the trimer, the native end-to-end distance and the curvature around the minimum ($\xi \approx 16$ Å) are approximately the same for both cis and trans structures. Major free-energy differences in the trimer isomers are only observed for $\xi \gtrsim 22$ Å, a region where the terminal base pair is already unfolded. This simple observation has important consequences; in the presence of defects, the DNA hairpin responds in such a way that it effectively localizes the structural damage to just two base pairs.

dx.doi.org/10.1021/jz201649k | *J. Phys. Chem. Lett.* 2012, 3, 689−693

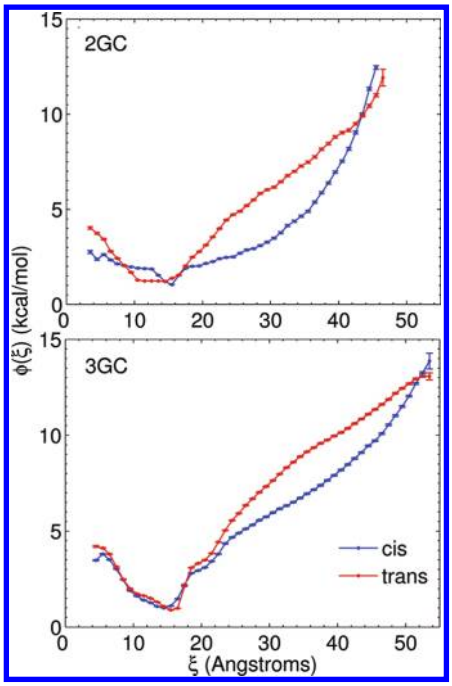


Figure 2. PMF along the end-to-end distance coordinate ξ for the 2GC and 3GC DNA hairpins in both cis and trans forms. The error bars correspond to twice the standard deviation obtained from a bootstrapping analysis.

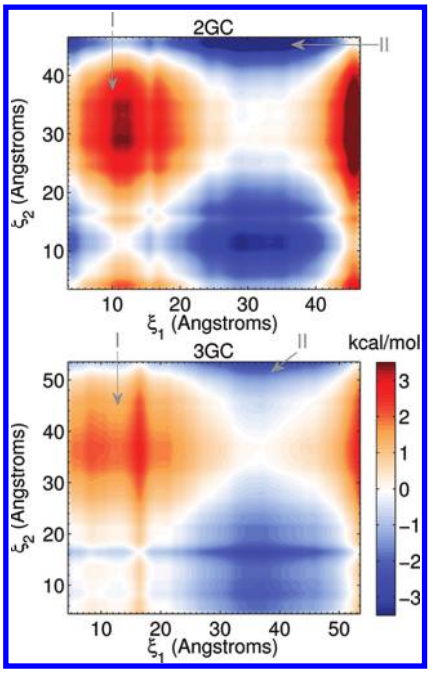


Figure 3. Extractable work W_{cycle} (eq 1) during an optomechanical cycle for the 2GC and 3GC DNA hairpins when the cycle is operated between extensions ξ_1 and ξ_2.

A quantitative way to assess the effect of the defect is by considering an optomechanical cycle[5] involving the cis and trans isomers. During this cycle, the molecule is first pulled from extension ξ_1 to extension ξ_2 in its cis form. At ξ_2, light is used to induce the cis-to-trans isomerization. Then, the molecule in the trans form is contracted from ξ_2 back to ξ_1. The cycle is closed by using light to induce trans-to-cis isomerization. Such a cycle can be performed in a single-molecule pulling setup using a stiff cantilever. Maximum work is obtained for reversible pulling, where the extractable work

$$W_{cycle} = \phi_{trans}(\xi_2) - \phi_{cis}(\xi_2) - [\phi_{trans}(\xi_1) - \phi_{cis}(\xi_1)] \quad (1)$$

is determined by the free-energy changes during the cycle. Importantly, W_{cycle} quantifies the difference in elasticity between the cis and trans structures. Figure 3 shows W_{cycle} for different choices of ξ_1 and ξ_2. There are two possible modes of work extraction[5] for these molecules, labeled I and II in the figure. Mode I exploits the photoinduced disruption of the hairpin structure for extracting work, and it is relevant for extensions where the interactions between the two strands of the hairpin are important. In turn, mode II exploits the difference in net length between the cis and trans isomers and operates for long extensions where this difference in length is significant. Here, we focus solely on mode I because it is the mode that quantifies the disruptions in the hairpin structure induced by the isomerization.

For the dimer, the maximum extractable work from mode I is 3.45 kcal/mol, and it occurs when ξ_1 = 10.9 Å and ξ_2 = 29.0 Å. By contrast, the maximum extractable work for the trimer is 2.70 kcal/mol, with ξ_1 = 16.4 Å and ξ_2 = 37.0 Å. Because the maximum extractable work does not increase when transitioning from the dimer to the trimer, we conclude that the DNA structure does in fact localize the disruptions introduced by the isomerization to two base pairs. The slight decrease of 0.75 kcal/mol in the extractable work of the trimer with respect to the dimer arises because the work cycle for the dimer exploits elasticity differences seen during contraction (ξ < 16 Å) of the dimer that are not present in the trimer because the terminal base pair of the trimer is not affected by the isomerization. In fact, if the cycle is started at ξ_1 = 16.4 Å for both molecules, the maximum extractable work coincides and is equal to 2.70 kcal/mol. This maximum extractable work is not expected to change when considering even longer DNA hairpins.

The localization of the defect to the two adjacent base pairs can also be seen in images of the unfolding process of the 3GC system (Figure 4). Starting in the native state (ξ = 16 Å), the trans system has all three base pairs neatly stacked and hydrogen-bonded. The cis isomer, on the other hand, has noticeable stacking of the azobenzene adjacent base pair instead of the usual Watson−Crick hydrogen bonding. The initial unfolding behavior of both isomers is similar due to the dissociation of the terminal base pair at around ξ = 18 Å for both systems. At this point, the unfolding pathways of the two isomers diverge. The cis system undergoes a concerted unfolding of the azobenzene adjacent base pairs for 23 < ξ < 26 Å. The trans system, however, undergoes a stepwise unfolding of these two base pairs. The middle GC base pair starts to unpair at ξ = 26 Å, and finally, the azobenzene adjacent base pair starts to unfold at ξ = 32 Å. Thus, the disruption due to the trans-to-cis isomerization of azobenzene only affects the unfolding behavior of the two inner GC base pairs in the 3GC system.

Such observed localization of structural defects in DNA is consistent with previous experimental and theoretical findings. In ref 23, an increase in the melting temperature of DNA hairpins was observed when a short linker (triethylene glycol) was replaced by a longer one (hexaethylene glycol). This melting temperature difference is comparable to the melting temperature difference between hairpins that differ in length by

dx.doi.org/10.1021/jz201649k | J. Phys. Chem. Lett. 2012, 3, 689−693

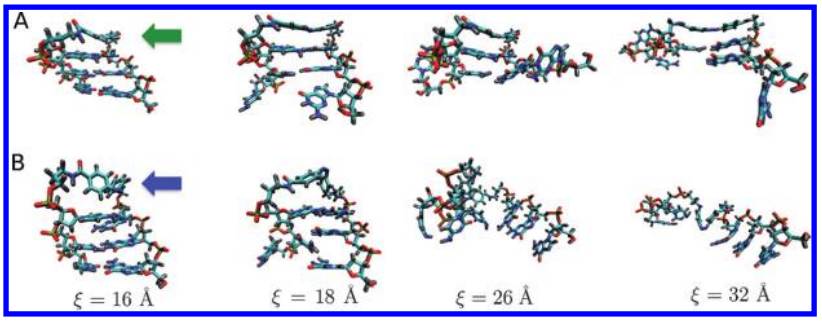


Figure 4. Pictures of the unfolding process for the 3GC system in its (A) trans and (B) cis isomers. The arrows signal the position of the azobenzene moiety. These images represent average structures encountered during the unfolding for different end-to-end distances. The averaging was performed by first collecting all recorded MD snapshots in 1 Å bins along ξ. Within each bin, the structures where then aligned and averaged. The energy of the averaged structure was then minimized to remove unphysical structural attributes inherent to this qualitative averaging procedure.

one base pair. The localization of the deformation was further evidenced by MD simulations that showed significant base pair restructuring only in the linker-adjacent base pair for triethylene glycol. Similarly, in ref 7, an increase in the melting temperature of a seven base pair DNA was observed upon cis-to-trans isomerization of an incorporated azobenzene moiety attached to only one of the strands. Again, the melting temperature increase (8.9 °C) is consistent with the melting temperature increase expected for a DNA sequence one or two base pairs longer. Both studies provide additional examples of defects in DNA affecting only one to two base pairs. This inherent flexibility of the DNA backbone that allows for defect localization is an important benchmark that coarse-grained models of DNA[24−29] need to capture to accurately model denaturation dynamics.

In conclusion, we have quantified the effect of structural defects induced by azobenzene isomerization in DNA hairpins of variable lengths using molecular dynamics simulations and free-energy reconstruction techniques. The results indicate that the structural flexibility of the DNA localizes the effect of the defect to two base pairs. Effective actuation using azobenzene in similarly designed DNA motors will require either a dense arrangement of azobenzenes along the DNA structure or use of modified DNA structures with more rigid backbones. While the results evidence a basic challenge in DNA motor design, they could not be more comforting from a biological perspective; despite the order of Watson−Crick base pairing, the chemical structure of DNA is flexible enough to efficiently prevent propagation of defects along its structure.

AUTHOR INFORMATION

Corresponding Author

*E-mail: schatz@chem.northwestern.edu.

Present Addresses

[‡]Chemistry Department, University of Chicago, Chicago, IL [¶]60637.

[¶]Theory Department, Fritz Haber Institute of the Max Planck Society, Faradayweg 4-6, 14195 Berlin, Germany.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This work was supported by the Nonequilibrium Energy Research Center (NERC), which is an Energy Frontier Research Center funded by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences under Award Number DE-SC0000989.

REFERENCES

(1) Lindahl, T. Instability and Decay of the Primary Structure of DNA. *Nature* **1993**, *362*, 709−715.

(2) Sancar, A.; Lindsey-Boltz, L. Molecular Mechanisms of Mammalian DNA Repair and The DNA Damage Checkpoints. *Annu. Rev. Biochem.* **2004**, *73*, 39−85.

(3) Balzani, V.; Credi, A.; Venturi, M. *Molecular Devices and Machines*, 2nd ed.; Wiley: New York, 2008.

(4) Tang, X.; Dmochowski, I. J. Regulating Gene Expression With Light-Activated Oligonucleotides. *Mol. BioSyst.* **2007**, *3*, 100−110.

(5) McCullagh, M.; Franco, I.; Ratner, M. A.; Schatz, G. C. DNA-Based Optomechanical Molecular Motor. *J. Am. Chem. Soc.* **2011**, *133*, 3452−3459.

(6) Kang, H.; Liu, H.; Phillips, J. A.; Cao, Z.; Kim, Y.; Chen, Y.; Yang, Z.; Li, J.; Tan, W. Single-DNA Molecule Nanomotor Regulated by Photons. *Nano Lett.* **2009**, *9*, 2690−2696.

(7) Asanuma, H.; Liang, X.; Yoshida, T.; Komiyama, M. Photocontrol of DNA Duplex Formation by using Azobenzene-Bearing Oligonucleotides. *Chem. Biol. Chem.* **2001**, *2*, 39.

(8) Liang, X.; Asanuma, H.; Kashida, H.; Takasu, A.; Sakamoto, T.; Kawai, G.; Komiyama, M. NMR Study on the Photoresponsive DNA Tethering an Azobenzene. Assignment of the Absolute Configuration of Two Diastereomers and Structure Determination of Their Duplexes in the *trans*-Form. *J. Am. Chem. Soc.* **2003**, *125*, 16408−16415.

(9) Liu, M.; Asanuma, H.; Komiyama, M. Azobenzene-Tethered T7 Promoter for Efficient Photoregulation of Transcription. *J. Am. Chem. Soc.* **2005**, *128*, 1009−1015.

(10) Krishnan, Y.; Simmel, F. C. Nucleic Acid Based Molecular Devices. *Angew. Chem. Intl. Ed.* **2011**, *50*, 3124−3156.

(11) Lawn, B. *Fracture of Brittle Solids*, 2nd ed.; Cambridge University Press: New York, 1993.

(12) Liphardt, J.; Onoa, B.; Smith, S. B.; Tinoco, I.; Bustamante, C. Reversible Unfolding of Single RNA Molecules by Mechanical Force. *Science* **2001**, *292*, 733−737.

(13) Huguet, J. M.; Bizarro, C. V.; Forns, N.; Smith, S. B.; Bustamante, C.; Ritort, F. Single-Molecule Derivation of Salt Dependent Base-Pair Free Energies in DNA. *Proc. Natl. Acad. Sci. U.S.A.* **2010**, *107*, 15431−15436.

(14) Hugel, T.; Holland, N. B.; Cattani, A.; Moroder, L.; Seitz, M.; Gaub, H. E. Single-Molecule Optomechanical Cycle. *Science* **2002**, *296*, 1103−1106.

(15) Shi, W.; Giannotti, M. I.; Zhang, X.; Hempenius, M. A.; Schönherr, H.; Vancso, G. J. Closed Mechanoelectrochemical Cycles of Individual Single-Chain Macromolecular Motors by AFM. *Angew. Chem.* **2007**, *119*, 8552−8556.

(16) Ciminelli, C.; Granucci, G.; Persico, M. The Photoisomerization Mechanism of Azobenzene: A Semiclassical Simulation of Non-adiabatic Dynamics. *Chem.—Eur. J.* **2004**, *10*, 2327−2341.

(17) Franco, I.; Schatz, G. C.; Ratner, M. A. Single-Molecule Pulling and the Folding of Donor−Acceptor Oligorotaxanes: Phenomenology and Interpretation. *J. Chem. Phys.* **2009**, *131*, 124902.
(18) Ferrenberg, A. M.; Swendsen, R. H. Optimized Monte Carlo Data Analysis. *Phys. Rev. Lett.* **1989**, *63*, 1195−1198.
(19) Kumar, S.; Bouzida, D.; Swendsen, R. H.; Kollman, P. A.; Rosenberg, J. M. The Weighted Histogram Analysis Method for Free-Energy Calculations on Biomolecules: I. The Method. *J. Comput. Chem.* **1992**, *13*, 1011−1021.
(20) Kale, L.; Skeel, R.; Bhandarkar, M.; Brunner, R.; Gursoy, A.; Krawetz, N.; Phillips, J.; Shinozaki, A.; Varadarajan, K.; Schulten, K. NAMD2: Greater Scalability for Parallel Molecular Dynamics. *J. Comput. Phys.* **1999**, *151*, 283−312.
(21) Foloppe, N.; MacKerell, A. D. All-Atom Empirical Force Field for Nucleic Acids: I. Parameter Optimization Based on Small Molecule and Condensed Phase Macromolecular Target Data. *J. Comput. Chem.* **2000**, *21*, 86−104.
(22) MacKerell, A. D.; Banavali, N. K. All-Atom Empirical Force Field for Nucleic Acids: II. Application to Molecular Dynamics Simulations of DNA and RNA in Solution. *J. Comput. Chem.* **2000**, *21*, 105−120.
(23) McCullagh, M.; Zhang, L.; Karaba, A. H.; Zhu, H.; Schatz, G. C.; Lewis, F. D. Effect of Loop Distortion on the Stability and Structural Dynamics of DNA Hairpin and Dumbbell Conjugates. *J. Phys. Chem. B* **2008**, *112*, 11415−11421.
(24) Drukker, K.; Schatz, G. C. A Model for Simulating Dynamics of DNA Denaturation. *J. Phys. Chem. B* **2000**, *104*, 6108−6111.
(25) Knotts, T. A.; Rathore, N.; Schwartz, D. C.; de Pablo, J. J. A Coarse-Grained Model for DNA. *J. Chem. Phys.* **2007**, *126*, 084901.
(26) Prytkova, T. R.; Eryazici, I.; Stepp, B.; Nguyen, S.-B.; Schatz, G. C. DNA Melting in Small-Molecule−DNA−Hybrid Dimer Structures: Experimental Characterization and Coarse-Grained Molecular Dynamics Simulations. *J. Phys. Chem. B* **2010**, *114*, 2627−2634.
(27) Dans, P. D.; Zeida, A.; Machado, M. R.; Pantano, S. A Coarse Grained Model for Atomic-Detailed DNA Simulations with Explicit Electrostatics. *J. Chem. Theory Comput.* **2010**, *6*, 1711−1725.
(28) Ouldridge, T. E.; Louis, A. A.; Doye, J. P. K DNA Nanotweezers Studied with a Coarse-Grained Model of DNA. *Phys. Rev. Lett.* **2010**, *104*, 178101.
(29) Savelyev, A.; Papoian, G. A. Chemically Accurate Coarse Graining of Double-Stranded DNA. *Proc. Natl. Acad. Sci. U.S.A* **2010**, *107*, 20340−20345.