JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY
Communication
pubs.acs.org/JACS

Direct Measurement of Sequential Folding Pathway and Energy Landscape of Human Telomeric G-quadruplex Structures

Wei Li,*[,†,∥] Xi-Miao Hou,[‡,∥] Peng-Ye Wang,[†] Xu-Guang Xi,[‡,§] and Ming Li*[,†]

[†]Beijing National Laboratory for Condensed Matter Physics and Key Laboratory of Soft Matter Physics, Institute of Physics, Chinese Academy of Sciences, Beijing 100190, China

[‡]School of Life Sciences, Northwest A&F University, Yangling 712100, China

[§]Laboratoire de Biologie et Pharmacologie Appliquée, Ecole Normale Supérieure de Cachan, Centre National de la Recherche Scientifique, 94235 Cachan, France

Supporting Information

ABSTRACT: Single-stranded guanine-rich sequences fold into compact G-quadruplexes. Although G-triplexes have been proposed and demonstrated as intermediates in the folding of G-quadruplexes, there is still a debate on their folding pathways. In this work, we employed magnetic tweezers to investigate the folding kinetics of single human telomeric G-quadruplexes in 100 mM Na^+ buffer. The results are consistent with a model in which the G-triplex is an in-pathway intermediate in the folding of the G-quadruplex. By finely tuning the force exerted on the G-quadruplex, we observed reversible transitions from the G-quadruplex to the G-triplex as well as from the G-triplex to the unfolded coil when the force was increased from 26 to 39 pN. The energy landscape derived from the probability distribution shows clearly that the G-quadruplex goes through an intermediate when it is unfolded, and vice versa.

The guanine-rich region with tandem TTAGGG repeats in human telomeres folds into compact G-quadruplex structures in which four guanines are held in a plane by Hoogsteen bonds.[1] The guanine planes are further stabilized by cations such as K^+ and Na^+.[2] G-quadruplexes have attracted great attention because of their physiological significance in genome stability maintenance,[3,4] chromosome end protection,[5] and anticancer drug targeting.[6] Knowledge of the folding pathway of the G-quadruplex, which is directly related to the dynamics of enzyme-catalyzed unwinding,[7] is essential for understanding the functions of telomeres. Many experiments, including circular dichroism (CD), differential scanning calorimetry, and isothermal titration calorimetry analyses,[8] fluorescence emission and fluorescence resonance energy transfer studies,[9] and optical tweezers experiments,[10] have established the existence of the G-triplex. Recently, NMR studies have even revealed the atomic structure of the G-triplex of thrombin binding aptamer (TBA).[11] However, the transition kinetics of the G-quadruplex, namely, whether the G-triplex serves as an in-pathway intermediate or exists independently of the G-quadruplex, is still the subject of debate. The folding kinetics revealed by ensemble-average assays indicated the involvement of two kinetic steps in the unfolding pathway.[8,12] However, single-molecule techniques seemed not to support the multistep sequential folding mechanism.[10,13−18]

The controversial results in the literature indicate that the folding kinetics of G-quadruplexes is far from being understood. The present research aimed to answer three questions: (1) Does a G-triplex exists in a force-induced unfolding process? (2) If it does exist, does it serve as an in-pathway intermediate in the unfolding process? (3) What is the corresponding energy landscape? To this end, we employed magnetic tweezers (MT) to study the folding kinetics of human telomeric G-quadruplex structures with the sequence GGGTTAGGGTTAGGGTTAGGG. We took advantage of the MT to tune the force very finely, which is essential for folding/unfolding studies under equilibrium conditions.

As the ionic conditions have an important impact on the topology of the G-quadruplexes,[19,20] we measured their CD spectra in a Na^+ buffer (100 mM NaCl, 20 mM Tris-HCl, pH 8.0) and in a K^+ buffer (100 mM KCl, 20 mM Tris-HCl, pH 8.0) before the MT assay. As shown in Figure 1a, the CD spectra indicate that the G-quadruplexes have a characteristic antiparallel structure in the Na^+ buffer[21] and a mixture of hybrids in the K^+ buffer, consistent with previous reports.[19,20] To avoid the uncertainty and reduce the complexity, we performed the MT experiments in the 100 mM Na^+ buffer at room temperature (24 °C).

The construction of the DNA is illustrated in Figure 1b. One end of the DNA was anchored to the antidigoxigenin-coated glass coverslip and the other to a 2.8 μm streptavidin-coated superparamagnetic Dynabead (Invitrogen Norway), as in our previous studies.[22,23] The tension exerted on the DNA was finely tuned by the position of magnets above the flow cell, and the end-to-end extension was recorded. The unfolding trajectories of the G-quadruplex under different tensions exhibited a multistate unfolding behavior (Figure 1c). In the low-tension region (below ~26 pN), the extension of the DNA increased smoothly with tension following the wormlike chain (WLC) model (data not shown).[24] Between ~26 and ~29 pN, the extension hopped between two states (Figure 1c, curves *i* and *ii*). When the tension exerted on the same DNA was increased to ~32 pN, the trajectory clearly showed the

Received: February 21, 2013

   dx.doi.org/10.1021/ja4019176 | J. Am. Chem. Soc. XXXX, XXX, XXX−XXX

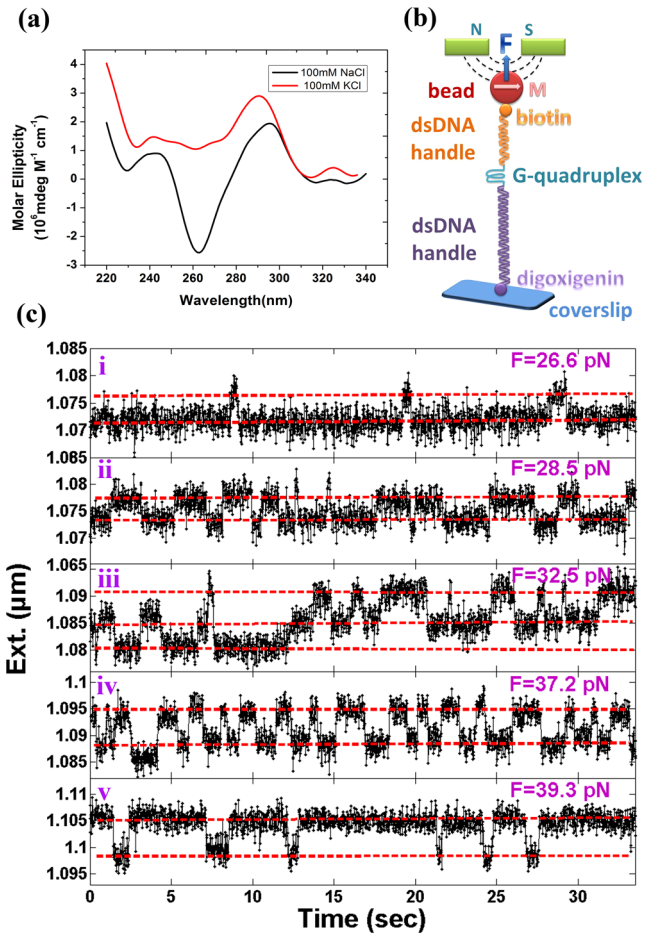

Figure 1. (a) CD spectra of the G-quadruplexes in 100 mM Na^+ buffer (black line) and 100 mM K^+ buffer (red line). (b) Schematic view of the magnetic tweezers assay (not to scale). The G-quadruplex is connected to a lower 699 bp dsDNA handle and an upper 2271 bp dsDNA handle. (c) Sequential unfolding trajectories of the G-quadruplex under different constant tensions. Here the data are from a single G-quadruplex, but the starting time of each measurement is shifted. The dashed lines correspond to the peaks in the probability distributions in Figure 2.

existence of three states (Figure 1c, curve *iii*). The extension of the third state had the largest value. As the tension was further increased, the first state became invisible and the extension hopped between the second and third states (Figure 1c, curves *iv* and *v*). Neither double-stranded DNA (dsDNA) nor single-stranded DNA (ssDNA) exhibits structural transitions at tensions below ~60 pN.[25,26] The hops in the extension in our single-molecule assays are therefore attributed to structural transitions of the G-quadruplex.

The corresponding tension-dependent probability distributions of the extension of the DNA, $P(x)$, are shown in Figure 2. The distributions were obtained by first removing the thermal motions of the beads at the end of the dsDNA handles (see Figure 1b) using the Harr wavelet transform tools in MATLAB. The details of the evaluations are described in the Supporting Information (SI).

Two peaks separated by 3.8 ± 0.4 nm were observed in the probability distributions when the tension on the DNA was tuned to between ~26 and ~29 pN (Figure 2, curves *i* and *ii*). The first peak corresponds to the extension of the DNA with a folded G-quadruplex (F). The second peak corresponds to the

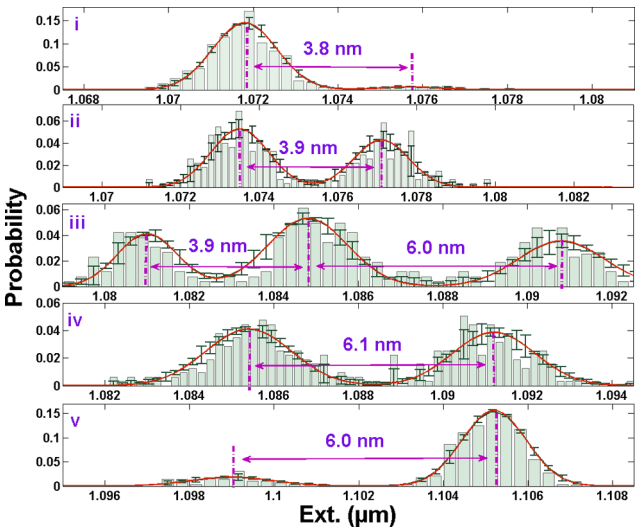

Figure 2. Probability distributions of the extension under different tensions corresponding to the trajectories in Figure 1c. The red lines are the Gaussian fits to the peaks, and the error bars are the residual errors of the corresponding fits. Two different step sizes of 3.9 ± 0.4 and 6.0 ± 0.4 nm indicate two different sequential unfolding processes.

extension of the DNA with a partially unfolded intermediate state (I). Details about I will be discussed later. Three states were observed at ~32 pN (Figure 2, curve *iii*). The extension of the DNA hopped among the three states with a small step of 3.9 ± 0.4 nm and a large step of 6.0 ± 0.4 nm. The first transition is similar to the one under lower tension. The second one can be attributed to the transition between I and the totally unfolded state (U). Beyond ~37 pN, the first transition became almost invisible but the second transition continued. When the tension was increased beyond 40 pN, the intermediate was destroyed and the structure remained in the unfolded state. The results clearly show that there is an in-pathway intermediate during the unfolding of the G-quadruplex. First, the three states appeared in order as the force was increased. Second, two reversible transitions, from F to I and from I to U, were observed. Finally, many transitions with two sequential jumps were observed in a single experiment (see curve *iii* in Figure 1c and more in Figure S3 in the SI).

If the intermediate state is the G-triplex as proposed elsewhere,[8−12,27,28] the first transition can be attributed to the stripping of a six-nucleotide arm (GGGTTA or TTAGGG) from the G-quadruplex. The reason for this is as follows. The length per nucleotide is ~0.55 nm under a tension near 30 pN.[25] Hence, the extension of the arm is 3.3 nm. There are two types of antiparallel G-quadruplexes: the basket type and the chair type.[29,30] As shown in Figure S2, the change in end-to-end distance is 0.2 nm for the transition of the basket-type G-quadruplex and 0.8 nm for the chair-type one.[31] The total increases in extension upon transition from these G-quadruplexes to the corresponding G-triplexes are therefore 3.5 and 4.1 nm, respectively. These values are very close to our experimental result of 3.9 ± 0.4 nm. A similar analysis shows that disruption of the G-triplex results in changes in extension of 6.1 and 6.0 nm for the basket-type and chair-type G-quadruplexes, respectively. These values are equal to the measured value of 6.0 ± 0.4 nm within experimental error. The estimated overall changes in extension are 9.6 and 10.1 nm for the complete disruption of the basket-type and chair-type G-quadruplexes, respectively. The measured steps fit the chair-

dx.doi.org/10.1021/ja4019176 | J. Am. Chem. Soc. XXXX, XXX, XXX−XXX

type G-quadruplex better than the basket-type one. The above analysis, together with evidence collected from the literature,[8−12,27,28] indicates that the G-triplex is the most plausible candidate for the intermediate structure, although some other unknown structures cannot be excluded at present.

Energy landscape formalisms provide a fundamental conceptual framework for describing the formation of three-dimensional structures of biomolecules.[32,33] In this work, the folding kinetics was measured under equilibrium conditions. The free energy landscape $G(x)$ can hence be directly derived from the probability distribution via the relation $G(x) = -k_BT \ln[P(x)]$. We repeatedly measured the folding trajectories and used those that exhibited multiple states to construct the free energy landscapes. As shown in Figure 3a, three minima

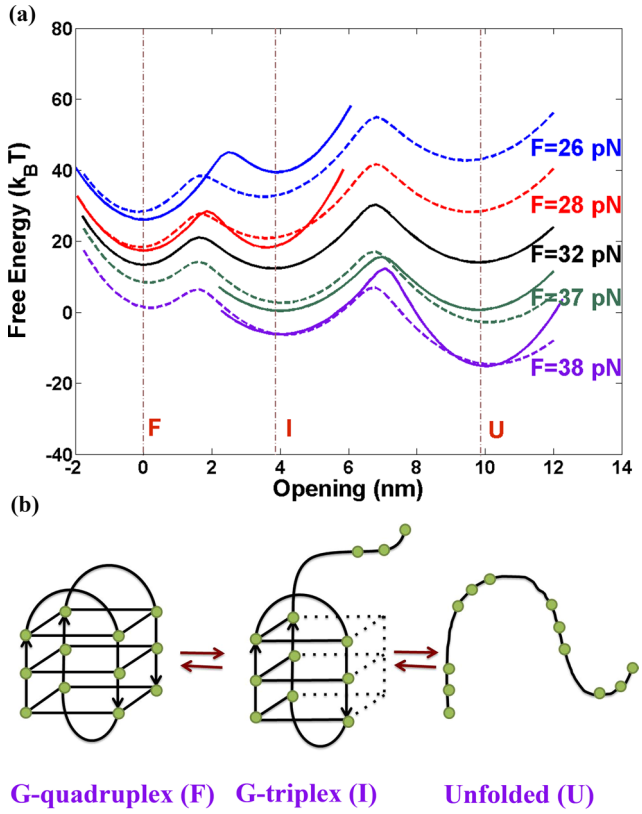

Figure 3. (a) Free energy landscapes of the G-quadruplex under different tensions (solid lines). These are related to each other by the expression $G = G_3 + x\Delta F$, where G_3 is the landscape at 32 pN; landscapes calculated using this formula are shown as dashed lines. The curves have been shifted vertically for clarity. (b) Model proposed to explain the sequential unfolding pathway of the G-quadruplex.

separated by two energy barriers can be identified in the energy landscape. The three minima correspond to the F, I, and U states. Under a finely tuned tension of $F_3 \approx 32$ pN, the free energy landscape [black curve in Figure 3a, denoted as $G_3(x)$ in the following] clearly shows that the three states have almost the same free energy. Neglecting the change in extension with force around F_3, the free energy landscapes under other tensions F are related to $G_3(x)$ by $G(x) = G_3(x) + x\Delta F$, where $\Delta F = F - F_3$. The landscapes calculated using this equation (dashed lines in Figure 3a) and the ones constructed from the experiments (solid lines) agree well with each other. However, this equation cannot be applied to zero tension because the reduced entropy of the unfolded segment due to stretching is not considered. According to Liphardt et al.,[34] the free energy change under zero tension, ΔG^0, is related to the equilibrium constant for the transition under nonzero tension, K_{eq}, by $\Delta G^0 = F\Delta x - \Delta G_{stre} - k_BT \ln(K_{eq})$, where ΔG_{stre} accounts for stretching of the unfolded state. The equilibrium constant K_{eq} can be derived from the lifetimes of the states (see the SI). The resulting values of ΔG^0 were found to be 6.3 kcal/mol for the transition from F to I and 13.3 kcal/mol for the transition from I to U (see the SI for details). The values are slightly higher than those reported previously,[8,28] which is understandable because the step sizes in the present work (Figure 2) are slightly larger than those in ref 28.

We performed measurements on more than 200 DNA molecules, with each being used several times, and analyzed 1641 extension steps. The steps were identified automatically by using a χ^2 step-fitting algorithm.[35] The intermediate state was robust in our experiments. Between 25 and 28 pN, more than 60% exhibited the transition from F to I as well as that from I to F with a step size of 3.7 ± 0.4 nm. Between 30 and 35 pN, more than 20% exhibited at least one event with two sequential jumps. The probability distribution of 863 steps in the force range between 30 and 35 pN is illustrated in Figure 4.

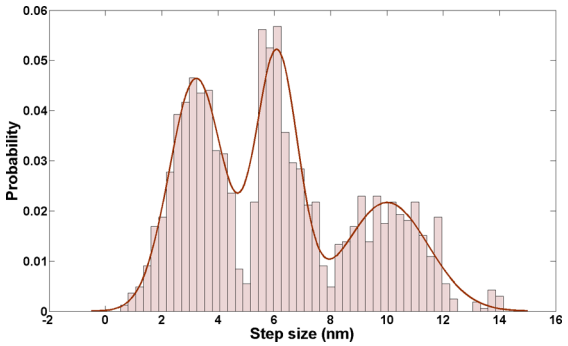

Figure 4. Probability distribution of the 863 steps identified in the data sets measured between 30 and 35 pN. The peaks are located at 3.7 ± 0.4, 6.1 ± 0.4, and 10.1 ± 0.5 nm.

Besides the two transitions with step sizes of 3.7 ± 0.4 and 6.1 ± 0.4 nm, many direct transitions from F to U and from U to F with a step size of 10.1 ± 0.5 nm were also observed. The magnetic tweezers enabled us to perform the folding experiments in the force-clamp mode easily. One can catch almost all of the events so long as one waits long enough. On the other hand, care must be taken to perform the unfolding experiments in the velocity-clamp mode.

It is noteworthy that the stability of the G-quadruplex depends on many conditions, such as the buffer used in the measurements and the detailed structure of the DNA construct (Figure 1b). As mentioned above, we used CD spectroscopy to check the structure of the G-quadruplexes and chose a buffer containing 100 mM Na^+. The intermediate state was robust in our experiments and began to collapse when the tension was higher than 32 pN. This value is similar to those for the unfolding of antiparallel G-quadruplexes measured using optical tweezers.[10] It has been reported that ssDNA protrusions destabilize G-quadruplexes,[4,36] so a DNA construct containing a G-quadruplex connected to the dsDNA handles through ssDNA segments may result in a low disruption force in the stretching experiments.[18]

In summary, we have clearly shown that the human telomeric G-quadruplex goes through an intermediate state before it is

dx.doi.org/10.1021/ja4019176 | J. Am. Chem. Soc. XXXX, XXX, XXX−XXX

forced to unfold totally, and vice versa. We have also constructed the tension-dependent free energy landscape, which possesses three minima separated by two barriers. The distances between the folded, intermediate, and unfolded states imply that the intermediate is a G-triplex. The extension of the DNA hops among the three states near a certain finely tuned tension. Away from this tension, the structure hops back and forth between the G-quadruplex and the G-triplex as well as between the G-triplex and the fully extended state (see Figure 3b for the proposed model). We believe that the present study reveals a clear sequential unfolding pathway of the G-quadruplex, which should be helpful for further studies on helicase-catalyzed unwinding of G-quadruplexes.

■ ASSOCIATED CONTENT

Supporting Information

Experimental details and analysis of data. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

weili007@iphy.ac.cn; mingli@iphy.ac.cn

Author Contributions

‖W.L. and X.-M.H. contributed equally.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work was supported by the National Natural Science Foundation of China (Grants 11104341, 61275192, and 11274374), the 973 Program of China (Grants 2013CB837200 and 2009CB930704), the 863 Program of China (Grant 2012AA02A104), and Northwest A&F University startup funding for X.-G.X. (Z101021102) and X.-M.H. (Z111021205). We thank Nanv Liu for the help with CD measurements.

■ REFERENCES

(1) Hoogsteen, K. *Acta Crystallogr.* **1963**, *16*, 907.
(2) Gilbert, D. E.; Feigon, J. *Curr. Opin. Struct. Biol.* **1999**, *9*, 305.
(3) Kamath-Loeb, A.; Loeb, L. A.; Fry, M. *PLoS One* **2012**, *7*, No. e30189.
(4) Lane, A. N.; Chaires, J. B.; Gray, R. D.; Trent, J. O. *Nucleic Acids Res.* **2008**, *36*, 5482.
(5) Blackburn, E. H. *Cell* **2001**, *106*, 661.
(6) Balasubramanian, S.; Hurley, L. H.; Neidle, S. *Nat. Rev. Drug Discovery* **2011**, *10*, 261.
(7) Liu, J. Q.; Chen, C. Y.; Xue, Y.; Hao, Y. H.; Tan, Z. *J. Am. Chem. Soc.* **2010**, *132*, 10521.
(8) Boncina, M.; Lah, J.; Prislan, I.; Vesnaver, G. *J. Am. Chem. Soc.* **2012**, *134*, 9657.
(9) Gray, R. D.; Buscaglia, R.; Chaires, J. B. *J. Am. Chem. Soc.* **2012**, *134*, 16834.
(10) Koirala, D.; Ghimire, C.; Bohrer, C.; Sannohe, Y.; Sugiyama, H.; Mao, H. *J. Am. Chem. Soc.* **2013**, *135*, 2235.
(11) Limongelli, V.; De Tito, S.; Cerofolini, L.; Fragai, M.; Pagano, B.; Trotta, R.; Cosconati, S.; Marinelli, L.; Novellino, E.; Bertini, I.; Randazzo, A.; Luchinat, C.; Parrinello, M. *Angew. Chem., Int. Ed.* **2013**, *52*, 2269.
(12) Zhang, A. Y.; Balasubramanian, S. *J. Am. Chem. Soc.* **2012**, *134*, 19297.
(13) Ying, L.; Green, J. J.; Li, H.; Klenerman, D.; Balasubramanian, S. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 14629.
(14) Lee, J. Y.; Okumus, B.; Kim, D. S.; Ha, T. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 18938.
(15) Lynch, S.; Baker, H.; Byker, S. G.; Zhou, D.; Sinniah, K. *Chemistry* **2009**, *15*, 8113.
(16) de Messieres, M.; Chang, J.-C.; Brawn-Cinani, B.; La Porta, A. *Phys. Rev. Lett.* **2012**, *109*, No. 058101.
(17) Qureshi, M. H.; Ray, S.; Sewell, A. L.; Basu, S.; Balci, H. *J. Phys. Chem. B* **2012**, *116*, 5588.
(18) Long, X.; Parks, J. W.; Bagshaw, C. R.; Stone, M. D. *Nucleic Acids Res* **2013**, *41*, 2746.
(19) Yu, H.; Gu, X.; Nakano, S. I.; Miyoshi, D.; Sugimoto, N. *J. Am. Chem. Soc.* **2012**, *134*, 20060.
(20) Xue, Y.; Kan, Z. Y.; Wang, Q.; Yao, Y.; Liu, J.; Hao, Y. H.; Tan, Z. *J. Am. Chem. Soc.* **2007**, *129*, 11185.
(21) Wang, Y.; Patel, D. J. *Structure* **1993**, *1*, 263.
(22) Li, W.; Wang, P.-Y.; Yan, J.; Li, M. *Phys. Rev. Lett.* **2012**, *109*, No. 218102.
(23) Fu, W. B.; Wang, X. L.; Zhang, X. H.; Ran, S. Y.; Yan, J.; Li, M. *J. Am. Chem. Soc.* **2006**, *128*, 15040.
(24) Marko, J. F.; Siggia, E. D. *Macromolecules* **1995**, *28*, 8759.
(25) Chen, W. S.; Chen, W. H.; Chen, Z.; Gooding, A. A.; Lin, K. J.; Kiang, C. H. *Phys. Rev. Lett.* **2010**, *105*, No. 218104.
(26) Zhang, X.; Chen, H.; Fu, H.; Doyle, P. S.; Yan, J. *Proc. Natl. Acad. Sci. U.S.A.* **2012**, *109*, 8103.
(27) Mashimo, T.; Yagi, H.; Sannohe, Y.; Rajendran, A.; Sugiyama, H. *J. Am. Chem. Soc.* **2010**, *132*, 14910.
(28) Koirala, D.; Mashimo, T.; Sannohe, Y.; Yu, Z.; Mao, H.; Sugiyama, H. *Chem. Commun.* **2012**, *48*, 2006.
(29) Dai, J.; Carver, M.; Punchihewa, C.; Hones, R. A.; Yang, D. *Nucleic Acids Res.* **2007**, *35*, 4927.
(30) Amrane, S.; Ang, R. W.; Tan, Z. M.; Li, C.; Lim, J. K.; Lim, J. M.; Lim, K. W.; Phan, A. T. *Nucleic Acids Res.* **2009**, *37*, 931.
(31) Zhang, N.; Phan, A. T.; Patel, D. J. *J. Am. Chem, Soc.* **2005**, *127*, 17277.
(32) Woodside, M. T.; Anthony, P. C.; Behnke-Parks, W. M.; Larizadeh, K.; Herschlag, D.; Block, S. M. *Science* **2006**, *314*, 1001.
(33) Hyeon, C.; Morrison, G.; Thirumalai, D. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 9604.
(34) Liphardt, J.; Onoa, B.; Smith, S. B.; Tinoco, I., Jr.; Bustamante, C. *Science* **2001**, *292*, 733.
(35) Kerssemakers, J. W.; Munteanu, E. L.; Laan, L.; Noetzel, T. L.; Janson, M. E.; Dogteron, M. *Nature* **2006**, *442*, 709.
(36) Viglasky, V.; Bauer, L.; Tluckova, K.; Javorsky, P. *J. Nucleic Acids* **2010**, No. 820356.