DNA Structures

Drastic Effect of a Single Base Difference between Human and *Tetrahymena* Telomere Sequences on Their Structures under Molecular Crowding Conditions**

*Daisuke Miyoshi, Hisae Karimata, and Naoki Sugimoto**

G-rich sequences, which are abundant throughout the genomes of most organisms, can fold into G-quadruplexes. There is little direct evidence for the formation of G-quadruplexes in vivo,[1] but there is growing interest in their potential roles in many biological systems.[2] In addition, various functional molecules can form G-quadruplex structures in vitro.[3] G-rich sequences have extraordinary structural polymorphism that depends on the sequence and the environmental conditions.[4] Importantly, the polymorphic nature of the G-quadruplexes makes them a promising nanomolecular material, because a regulated structural transition between different types of G-quadruplexes can provide the basis for switchable molecular devices.[5] Moreover, the G-rich sequences are attracting interest as functional elements in molecular electronics.[6] Therefore, regulation of the polymorphic nature of the G-quadruplex presents a novel methodology for both developing molecular devices in vitro and controlling biological phenomena in vivo.

Changes in the environmental conditions can result in various structural changes of nucleic acids,[7] especially G-quadruplexes. For example, we reported that structural transitions between antiparallel and parallel G-quadruplexes, and between duplex and quadruplex, are induced by molecular crowding,[8] which is a critical environmental factor affecting the structure of biomacromolecules.[9] The effect of molecular crowding on the structures and functions of biomacromolecules has been examined because living cells are inherently molecularly crowded environments that contain a wide variety of biomolecules. For example, the total biomolecule concentration in *Escherichia coli* is in the range of 300 to 400 $g L^{-1}$.[9] This value is different from the typical biomolecule concentrations ($< 1 g L^{-1}$) generally used for

[*] Prof. Dr. N. Sugimoto
Frontier Institute for Biomolecular Engineering Research (FIBER) and Department of Chemistry
Faculty of Science and Engineering, Konan University
8-9-1 Okamoto, Higashinada-ku, Kobe 658-8501 (Japan)
Fax: (+81) 78-435-2766
E-mail: sugimoto@konan-u.ac.jp
Dr. D. Miyoshi
FIBER, Konan University (Japan)
H. Karimata
Department of Chemistry, Faculty of Science and Engineering
Konan University (Japan)

[**] This work was supported in part by Grants-in-Aid from the Ministry of Education, Science, Sports, and Culture, Japan.

Supporting information for this article is available on the WWW under http://www.angewandte.org or from the author.

  DOI: 10.1002/ange.200462667

in vitro experiments. However, only a few systematic experimental studies of G-quadruplex structures under molecular crowding conditions have been reported.[8] Herein we present evidence that molecular crowding induces drastically different structures in the G-quadruplexes formed by *Tetrahymena* (Tet) and human (Hum) telomere sequences. These sequence motifs differ by only one base, which indicates that the single mutation in the telomere sequence is crucial for the regulation of its polymorphic nature and, therefore, its potential biological functions and material properties.

Figure 1 shows the circular dichroism (CD) spectra of samples (50 μmol L^{-1}) of four telomere DNA sequences (sites where Tet and Hum differ are underlined): intermolecular

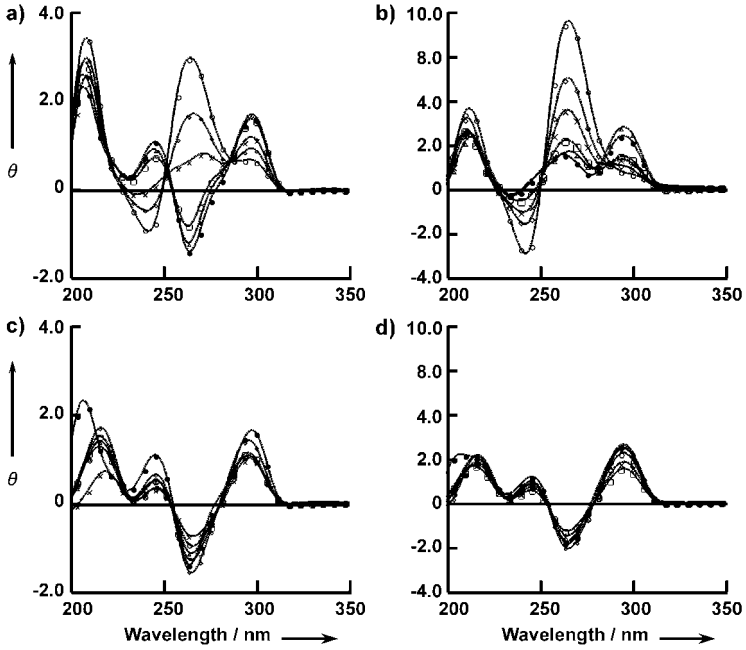

Figure 1. CD spectra (θ in 10^5 deg cm^2 $dmol^{-1}$) of samples of inter-Tet (a), intra-Tet (b), inter-Hum (c), and intra-Hum (d) at 4 °C without (●) and with 40 wt% cosolute. Cosolutes include ethylene glycol (△), diethylene glycol (□), triethylene glycol (×), PEG 200 (◇), and PEG 2000 (○).

Tetrahymena (inter-Tet; d[$TGG_3T_2\underline{G}G_3T$]; Figure 1a), intramolecular *Tetrahymena* (intra-Tet; d[$T_2G(G_3T_2\underline{G})_3G$]; Figure 1b), intermolecular human (inter-Hum; d[$TAG_3T_2\underline{A}G_3T$]; Figure 1c), and intramolecular human (intra-Hum; d[$(G_3T_2\underline{A})_3G_3$]; Figure 1d). CD measurements were carried out with and without a cosolute.[10] Polyethylene glycol (PEG) and its related small molecules are commonly used as cosolutes in aqueous solution to mimic molecular crowding conditions, because they do not react with nucleotides and a variety of molecular weights are available.[8–10] Except for intra-Tet, the CD spectra of the telomere sequences in the absence of cosolute have positive and negative peaks at 295 and 265 nm, respectively, which indicates that these telomere sequences fold into antiparallel G-quadruplexes.[8,11] Intermolecular *Oxytricha* (inter-Oxy; d[$G_4T_4G_4$]) and intramolecular *Oxytricha* (intra-Oxy; d[$(G_4T_4)_3G_4$)]) telomere sequences also show an antiparallel structure in the absence of the cosolute (see the Supporting Information). However, the CD spectrum of intra-Tet has positive peaks at 295 and 260 nm in the absence of cosolute, thus indicating a mixture of parallel and antiparallel G-quadruplexes.[8,11] These results are consistent with those of previous reports that describe the structure of these telomere sequences.[12]

Surprisingly, the CD spectra in the presence of cosolutes with various molecular weights reveal that, under crowding conditions, a single G-to-A replacement in the loops affects the entire telomere structure. Specifically, the CD spectra of inter-Tet and intra-Tet in the presence of higher-molecular-weight cosolutes have a positive peak at approximately 260 nm and a shoulder near 295 nm. In contrast, the CD spectra of inter-Hum and intra-Hum do not change under these conditions. This finding indicates that the *Tetrahymena* but not the human telomere DNA undergoes a structural transition from an antiparallel to a parallel G-quadruplex in the presence of higher-molecular-weight cosolutes such as PEG 2000. Furthermore, the CD spectra of inter-Oxy, intra-Oxy, d[$(T_3G_4)_4$], d[$(G_3T_3)_3G_3$], d[$(G_4T_2)_3G_4$], and d[$(T_2AG_3)_4$] in the presence and absence of the cosolute demonstrated that these telomere sequences, except d[$(T_2AG_3)_4$], undergo the structural transition by adding the cosolute (see the Supporting Information). These telomere sequences have different molecularities (intermolecular or intramolecular), sequence lengths, and numbers of G-quartets (guanine plane), nucleotides in the loop, and extra nucleotides at the termini, which indicates that these factors are not important participants in the transition. Therefore, the single base difference between human and *Tetrahymena* telomere sequences critically affects the entire structure of the telomere DNA. However, the CD spectra suggest that the cosolutes cause a more dramatic shift to a parallel G-quadruplex for inter-Tet (Figure 1a) than for intra-Tet (Figure 1b).

Phan et al.[12a,b] reported the solution structure of inter-Tet in the presence of Na^+ ions and of inter-Hum in the presence of K^+ ions. Inter-Tet forms two antiparallel dimeric G-quadruplexes with different positions of the loops. Inter-Hum also has two different structures: a dimeric parallel G-quadruplex with loops located in the grooves (propeller structure) and a dimeric antiparallel G-quadruplex with loops located on the G-tetrad planes. Wang and Patel[12c,d] reported that intra-Tet and intra-Hum form intramolecular antiparallel G-quadruplexes with Na^+ ions. Parkinson et al.[13] used X-ray crystallography to show that the inter-Hum and intra-Hum form dimeric and monomeric propeller-type parallel G-quadruplex structures in the presence of K^+ ions, although the human telomere sequences used in their study were slightly different from those used here. On the basis of these structural studies, one can conclude that inter-Tet and intra-Tet fold into antiparallel G-quadruplexes in the presence of Na^+ ions. Inter-Hum and intra-Hum can fold into dimeric and monomeric antiparallel G-quadruplexes, respectively, in the presence of Na^+ ions and

can potentially fold into propeller-type parallel G-quadruplexes as a monomer and dimer, respectively.

The excluded volume, one of the factors of molecular crowding,[9] can induce the association of monomers into multimeric complexes.[9,14] For example, monomeric and dimeric antiparallel G-quadruplexes can associate into multistranded G-wires.[8] Inter-Tet and intra-Tet likely underwent this structural transition because their parallel structure must be either four- or multistranded, such as in a G-wire structure (Figure 2). Inter-Hum and intra-Hum, however, could not

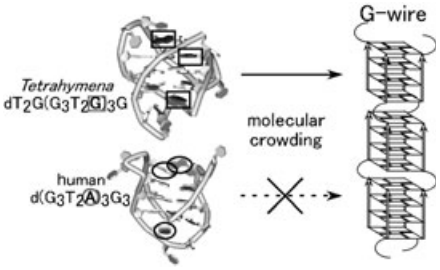

Figure 2. Schematic illustration of the structures of intra-Tet and intra-Hum in dilute and molecular crowding conditions. G (□) and A (○) correspond to the single-mutation sites.

undergo the antiparallel-to-parallel transition under conditions of molecular crowding because the parallel structures form propeller-type G-quadruplexes, which may have excluded volumes larger than that of the antiparallel G-quadruplex.[12,13] In addition, the thermodynamic properties of intra-Tet and intra-Hum in the presence and absence of 40 wt% PEG200 were examined by using UV melting curves monitored at 295 nm. The melting temperatures (T_m) of intra-Tet were 70.4 and 62.0 °C, respectively, whereas these T_m values for intra-Hum were 65.0 and 60.8 °C, respectively. Therefore, a larger stabilization of intra-Tet by the cosolutes may induce the structural transition.

To further verify the formation of the G-wire structure, we performed native gel electrophoresis on a 10% polyacrylamide gel.[15] Figure 3a shows the migration of intra-Tet and intra-Hum in the presence of 0 to 40 wt% PEG2000. The migration pattern of intra-Tet is typical for a G-wire (indicated by asterisks).[16] Furthermore, the ratio of intra-Tet migrating at the top of the gel increased as the concentration of PEG2000 was increased (indicated by arrows). In denaturing gel electrophoresis, the band at the top of the gel with intra-Tet samples disappeared, and another moderately migrating band appeared (arrow in Figure 3b). All of these results indicate that the band observed at the top of the native gel for intra-Tet samples corresponds to a highly ordered G-wire complex, whereas intra-Hum remains as a compact intramolecular G-quadruplex structure even in the presence of 40 wt% PEG2000. This concept is further confirmed by the CD spectra, which show that titration with PEG2000 causes a structural transition from an antiparallel G-quadruplex to a parallel-oriented G-quadruplex of intra-Tet but

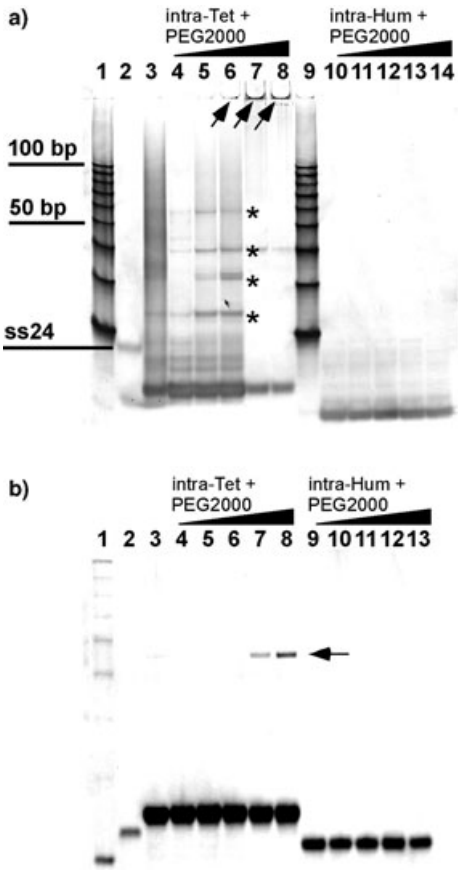

Figure 3. a) Native gel electrophoresis of intra-Tet and intra-Hum in the absence and presence of PEG2000. See the Experimental Section for details. Lanes 1 and 9: 10-bp DNA ladder; lane 2: single-stranded 24-mer DNA; lane 3: inter-Tet with Mg^{2+} and spermidine; lanes 4–8: intra-Tet in the presence of 0, 10, 20, 30, and 40 wt% PEG2000, respectively; lanes 10–14: intra-Hum in the presence of 0, 10, 20, 30, and 40 wt% PEG2000, respectively. b) Denaturing gel electrophoresis of intra-Tet and intra-Hum in the absence and presence of PEG2000. See the Experimental Section for details. Lane 1: 10-base DNA ladder; lane 2: single-stranded 24-mer DNA; lane 3: inter-Tet with Mg^{2+} and spermidine; lanes 4–8: intra-Tet in the presence of 0, 10, 20, 30, and 40 wt% PEG2000, respectively; lanes 9–13: intra-Hum in the presence of 0, 10, 20, 30, and 40 wt% PEG2000, respectively.

not intra-Hum (Figure 4). The CD spectrum of intra-Tet in the presence of 40 wt% PEG2000 (○) in Figure 4a) has a

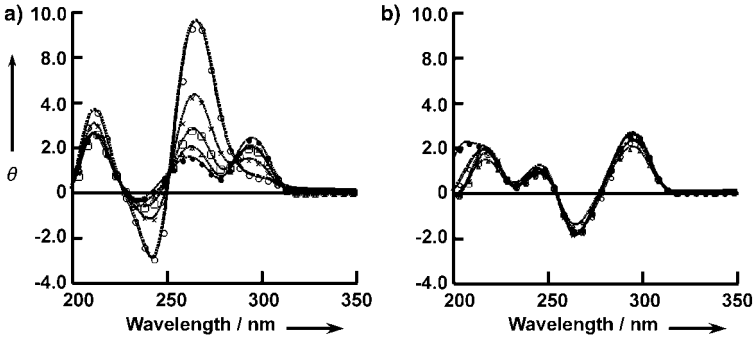

Figure 4. CD spectra of samples of intra-Tet (a) and intra-Hum (b) at 4 °C without (●) or with 10 (△), 20 (□), 30 (×), or 40 wt% (○) PEG2000.

small shoulder around 295 nm, which indicates that the antiparallel G-quadruplex still remains in this condition. This shoulder may correspond to the remaining band present in the antiparallel G-quadruplex of intra-Tet (Figure 3a). In addition, the migration of inter-Tet in the presence of Mg^{2+} ions and spermidine (lane 3 in Figure 3a), where there is a high potential for the formation of G-wires,[16] shows a very diffuse band, and lanes 4 to 8 provide some evidence of discrete retarded complexes. Therefore, the migration of intra-Tet in the presence of PEG 2000 shows a higher-ordered and, thus, a longer G-wire structure. Finally, although it is difficult to prepare homogeneous G-wires, they can be obtained by cutting out the band in the denaturing gel.

In summary, we have demonstrated that a single G-to-A substitution in the loops of telomere sequences leads to drastically different structures. Under molecular crowding conditions, *Tetrahymena* telomere sequences fold into very long, highly ordered G-wires, whereas human telomere sequences fold into antiparallel G-quadruplexes. Further studies are required, such as on the effect of monovalent cations on the structure of the telomere sequences under molecular crowding conditions. However, the results reported here indicate that a single mutation in the telomere sequence is a critical factor affecting the polymorphic nature of the G-quadruplex and, therefore, of telomere function in cell-like conditions.

Apart from the biological aspects, there is growing interest in G-rich sequences as functional elements in molecular electronics.[6,17] A theoretical study suggested that the G-wire structure is promising for nanoscale biomolecular electronics because of its highly order structure.[18] However, its electrical properties have not yet been described because of the difficulty in creating and controlling the G-wire structure. Herein, we have described the control of the G-wire structure by a single mutation and by adjusting the solution conditions. These findings make it possible to measure electrical transport through G-wires.

Experimental Section

All cosolutes were purchased from Wako Pure Chemical Co. Ltd. (Japan) and used without further purification. Oligonucleotides were synthesized, purified, and confirmed as described elsewhere.[19] The DNA samples were heated to 80 °C, gently cooled at a rate of 2.0 K min^{-1}, and incubated overnight at 4 °C.

CD spectra of DNA samples (50 μmol L^{-1} total strand concentration) were obtained by using a J-820 spectropolarimeter (JASCO Co. Ltd., Japan) with a 0.1-cm-path-length quartz cell. CD measurements were carried out in a buffer containing NaCl (100 mmol L^{-1}), Na_2HPO_4 (10 mmol L^{-1}), and ethylenediaminetetraacetic acid disodium salt (Na_2EDTA, 1 mmol L^{-1}, pH 7.0) with or without cosolutes at the desired temperature. It was possible to induce the formation of G-wires by inter-Tet and intra-Tet by adding a cosolute and without heating. In this study, however, to ensure that the structures reached equilibrium, CD spectra were recorded after heating and slow cooling to 4 °C followed by an overnight incubation.

UV melting curves for DNA samples (50 μmol L^{-1} total strand concentration) were recorded at 295 nm[20] by using a Shimadzu UV-1700 instrument (Shimadzu Co. Ltd., Japan) with a 0.1-cm-path-length quartz cell. The measurements were carried out at 4 °C in a buffer containing NaCl (100 mmol L^{-1}), Na_2HPO_4 (10 mmol L^{-1}), and Na_2EDTA (1 mmol L^{-1}, pH 7.0) with or without cosolutes. Before measurements were made, the DNA samples were heated, cooled, and incubated to allow them to equilibrate, as described above. The samples were heated at 0.5 K min^{-1} to obtain the melting curves.

Native gel electrophoresis was carried out at 4 °C and 5 V cm^{-1} on a 10 % nondenaturing polyacrylamide gel in a buffer containing NaCl (100 mmol L^{-1}), Na_2HPO_4 (10 mmol L^{-1}), and Na_2EDTA (1 mmol L^{-1}, pH 7.0). Ice-cold loading buffer (2 μL, 40 % glycerol and 1 % blue dextran) was mixed with the DNA sample (2 μL, 25 μmol L^{-1}). Before measurements were made, the DNA samples were heated, cooled, and incubated to allow them to equilibrate, as described above. Denaturing gel electrophoresis was carried out using the same procedure as that for native gel electrophoresis, except that the loading buffer contained 50 % formamide and electrophoresis was performed at room temperature. Gels were stained with GelStar nucleic acid gel stain (Cambrex, ME, USA) and imaged using FLS-5100 film (Fuji Photo Film Co. Ltd., Japan).

Received: November 19, 2004
Published online: April 29, 2005

Keywords: DNA structures · molecular devices · nanostructures · oligonucleotides · self-assembly

[1] a) C. Schaffitzel, I. Berger, J. Postberg, J. Hanes, H. J. Lipps, A. Plückthun, *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 8572 – 8577; b) A. Siddiqui-Jain, C. L. Grand, D. J. Bearss, L. H. Hurley, *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 11 593 – 11 598; c) M. L. Duquette, P. Handa, J. A. Vincent, A. F. Taylor, N. Maizels, *Genes Dev.* **2004**, *18*, 1618 – 1629.

[2] a) S. Neidle, M. A. Read, *Biopolymers* **2001**, *56*, 195 – 208; b) R. H. Shafer, I. Smirnov, *Biopolymers* **2001**, *56*, 209 – 227; c) J.-L. Mergny, C. Helene, *Nat. Med.* **1998**, *4*, 1366 – 1367; d) S. Neidle, G. Parkinson, *Nat. Rev. Drug Discovery* **2002**, *1*, 383 – 393.

[3] a) R. F. Macaya, P. Schultze, F. W. Smith, J. A. Roe, J. Feigon, *Proc. Natl. Acad. Sci. USA* **1993**, *90*, 3745 – 3749; b) N. Jing, M. E. Hogan, *J. Biol. Chem.* **1998**, *273*, 34 992 – 34 999; c) N. Sugimoto, T. Toda, T. Ohmichi, *Chem. Commun.* **1998**, 1533 – 1534; d) A. M. Whitney, S. Ladame, S. Balasubramanian, *Angew. Chem.* **2004**, *116*, 1163 – 1166; *Angew. Chem. Int. Ed.* **2004**, *43*, 1143 – 1146.

[4] a) D. Sen, W. Gilbert, *Nature* **1990**, *344*, 410 – 414; b) D. Sen, W. Gilbert, *Biochemistry* **1993**, *31*, 65 – 70; c) M. A. Keniry, *Biopolymers* **2001**, *56*, 123 – 146; d) F. M. Chen, *Biochemistry* **1992**, *31*, 3769 – 3776.

[5] a) C. Mao, W. Sun, Z. Shen, N. C. Seeman, *Nature* **1999**, *397*, 144 – 146; b) P. Alberti, J.-L. Mergny, *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 1569 – 1573; c) J. T. Davis, *Angew. Chem.* **2004**, *116*, 684 – 716; *Angew. Chem. Int. Ed.* **2004**, *43*, 668 – 698.

[6] V. A. Szalai, H. H. Thorp, *J. Am. Chem. Soc.* **2000**, *122*, 4524 – 4525.

[7] Y. Fang, T. S. Spisz, J. H. Hoh, *Nucleic Acids Res.* **1999**, *27*, 1943 – 1949.

[8] a) D. Miyoshi, A. Nakao, N. Sugimoto, *Biochemistry* **2002**, *41*, 15 017 – 15 024; b) D. Miyoshi, S. Matsumura, S. Nakano, N. Sugimoto, *J. Am. Chem. Soc.* **2004**, *126*, 165 – 169.

[9] a) S. B. Zimmerman, A. P. Minton, *Annu. Rev. Biophys. Biomol. Struct.* **1993**, *22*, 27 – 65; b) J. R. Wenner, V. A. Bloomfield, *Biophys. J.* **1999**, *77*, 3234 – 3241; c) P. R. Davis-Searles, A. J. Saunders, D. A. Erie, D. J. Winzor, G. J. Pielak, *Annu. Rev. Biophys. Biomol. Struct.* **2001**, *30*, 271 – 306; d) A. P. Minton, *J. Biol. Chem.* **2001**, *276*, 10 577 – 10 580.

[10] The molecular weights of the cosolutes used, ethylene glycol, diethylene glycol, triethylene glycol, PEG 200, and PEG 2000, are 62, 106, 150, 200 (average), and 2000 (average), respectively:

S. Nakano, H. Karimata, T. Ohmichi, J. Kawakami, N. Sugimoto, *J. Am. Chem. Soc.* **2004**, *126*, 14330 – 14331.
[11] a) M. Lu, Q. Guo, N. R. Kallenback, *Biochemistry* **1992**, *31*, 2455 – 2459; b) P. Balagurumoorthy, S. K. Brahmachari, *J. Biol. Chem.* **1994**, *269*, 21858 – 21869; c) A. Risitano, K. R. Fox, *Nucleic Acids Res.* **2004**, *32*, 2598 – 2606.
[12] a) A. T. Phan, Y. S. Modi, D. J. Patel, *J. Mol. Biol.* **2004**, *338*, 93 – 102; b) A. T. Phan, D. J. Patel, *J. Am. Chem. Soc.* **2003**, *125*, 15021 – 15027; c) Y. Wang, D. J. Patel, *Structure* **1994**, *2*, 1141 – 1156; d) Y. Wang, D. J. Patel, *Structure* **1993**, *1*, 263 – 282.
[13] G. N. Parkinson, M. P. H. Lee, S. Neidle, *Nature* **2002**, *417*, 876 – 880.
[14] A. P. Minton, *Biopolymers* **1981**, *20*, 2093 – 2120.
[15] V. Dapic, V. Abdomerovic, R. Marrington, J. Peberdy, A. Rodger, J. O. Trent, P. J. Bates, *Nucleic Acids Res.* **2003**, *31*, 2097 – 2107.
[16] a) D. Sen, W. Gilbert, *Biochemistry* **1992**, *31*, 65 – 70; b) T. C. Marsh, J. Vesenka, E. Henderson, *Nucleic Acids Res.* **1995**, *23*, 696 – 700.
[17] a) S. Delaney, J. K. Barton, *Biochemistry* **2003**, *42*, 14159 – 14165; b) D. Porath, A. Bezryadin, S. de Vries, C. Dekker, *Nature* **2000**, *403*, 635 – 637; c) G. B. Schuster, *Acc. Chem. Res.* **2000**, *33*, 253 – 260.
[18] A. Calzolari, R. D. Felice, E. Molinari, *Appl. Phys. Lett.* **2002**, *80*, 3331 – 3333.
[19] a) T. Ohmichi, S. Nakano, D. Miyoshi, N. Sugimoto, *J. Am. Chem. Soc.* **2002**, *124*, 10367 – 10372; b) S. Nakano, T. Kanzaki, N. Sugimoto, *J. Am. Chem. Soc.* **2004**, *126*, 1088 – 1095.
[20] J. L. Mergny, A. T. Phan, L. Lacroix, *FEBS Lett.* **1998**, *435*, 74 – 78.

