Structural Dynamics and Cation Interactions of DNA Quadruplex Molecules Containing Mixed Guanine/Cytosine Quartets Revealed by Large-Scale MD Simulations

Nad'a Špačková,[†,‡,§] Imre Berger,[||] and Jiří Šponer*[,†,‡]

Contribution from the Institute of Biophysics, Academy of Sciences of the Czech Republic, and National Centre for Biomolecular Research, Královopolská 135, 612 65 Brno, Czech Republic, J. Heyrovský Institute of Physical Chemistry, Academy of Sciences of the Czech Republic, Dolejškova 3, 182 23 Prague, Czech Republic, Department of Physical Electronics, Faculty of Science, Masaryk University, Kotlářská 2, 611 37 Brno, Czech Republic, and Institute for Molecular Biology and Biophysics, ETH-Hönggerberg, CH-8093 Zürich, Switzerland

Received July 19, 2000. Revised Manuscript Received January 29, 2001

Abstract: Large-scale molecular dynamics (MD) simulations have been utilized to study G-DNA quadruplex molecules containing mixed GCGC and all-guanine GGGG quartet layers. Incorporation of mixed GCGC quartets into G-DNA stems substantially enhances their sequence variability. The mixed quadruplexes form rigid assemblies that require integral monovalent cations for their stabilization. The interaction of cations with the all-guanine quartets is the leading contribution for the stability of the four-stranded assemblies, while the mixed quartets are rather tolerated within the structure. The simulations predict that two cations are preferred to stabilize a four-layer quadruplex stem composed of two GCGC and two all-guanine quartets. The distribution of cations in the structure is influenced by the position of the GCGC quartets within the quadruplex, the presence and arrangement of thymidine loops connecting the guanine/cytosine stretches forming the stems, and the cation type present (Na^+ or K^+). The simulations identify multiple nanosecond-scale stable arrangements of the thymidine loops present in the molecules investigated. In these thymidine loops, several structured pockets are identified capable of temporarily coordinating cations. However, no stable association of cations to a loop has been observed. The simulations reveal several paths through the thymidine loop regions that can be followed by the cations when exchanging between the central ion channel in the quadruplex stem and the surrounding solvent. We have carried out 20 independent simulations while the length of simulations reaches a total of 90 ns, rendering this study one of the most extensive MD investigations carried out on nucleic acids so far. The trajectories provide a largely converged characterization of the structural dynamics of these four-stranded G-DNA molecules.

Introduction

DNA is a highly polymorphous molecule that can adopt an astounding variety of molecular conformations. These include the established A,B and Z duplex forms, and also multistranded conformations, for example four-stranded assemblies such as intercalated cytosine-rich i-DNA, and the G-DNA quadruplex formed by guanine-rich DNA.[1,2] In the genome of cells, DNA sequences containing guanine-rich stretches are present in gene regulatory regions, and also in centromeric and telomeric DNAs, among others. A four-stranded DNA structural assembly containing a guanine quadruplex stem was first proposed for guanine-rich DNA oligonucleotides comprising sequences that are present as unpaired overhangs at the ends of eukaryotic chromosomes, the telomeres.[1] Structural studies on guanine-rich telomeric sequences revealed many details of this quadruplex assembly at high resolution, also implicating the involvement of cations as essential stabilizing factors of this novel DNA motif.[2] An interesting feature of the G-DNA quadruplex molecule is the high degree of conformational variability of this structural motif in itself. Thus, oligonucleotides with multiple guanine-rich stretches can form inter- and intramolecular,

* To whom correspondence should be addressed. Institute of Physical Chemistry, Prague, Czech Republic. E-mail: sponer@indy.jh-inst.cas.cz. Fax: (4202) 858 2307.

‡ J. Heyrovský Institute of Physical Chemistry, Academy of Sciences of the Czech Republic.

§ Masaryk University.

|| Institute for Molecular Biology and Biophysics.

† Institute of Biophysics, Academy of Sciences of the Czech Republic.

(1) (a) Williamson, J. R.; Raghuram, M. K.; Cech, T. R. *Cell* **1989**, *59*, 871−880. (b) Sundquist, W. I.; Klug, A. *Nature* **1989**, *342*, 825−829. (c) Sen, D.; Gilbert, W. *Nature* **1988,** *334*, 364−366. (d) Zakian, V. A. *Science* **1995**, *279*, 1601.

(2) Guanine quadruplex molecules: (a) Laughlan, G.; Murchie, A. I. H.; Norman, D. G.; Moore, M. H.; Moody, P. C. E.; Lilley, D. M. J.; Luisi, B. *Science* **1994**, *265*, 520−524. (b) Smith, F. W.; Feigon, J. *Nature* **1992**, *356*, 164−168. (c) Phillips, K.; Dauter, Z.; Murchie, A. I. H.; Lilley, D. M. J.; Luisi, B. *J. Mol. Biol.* **1997**, *273*, 171−189. (d) Phillips, K.; Luisi, B. *Science* **1998**, *281*, 883. (e) Kang, C. H.; Zhang, X.; Ratliff, R.; Moyzis, R.; Rich, A. *Nature* **1992**, *356*, 126−131. (f) Smith, F. W.; Schultze, P.; Feigon, J. *Structure* **1995**, *3*, 997−1008. (g) Aboud-ela, F.; Murchie, A. I. H.; Norman, D. G.; Lilley, D. M. J. **1994**, *243*, 458−471. (h) Marathias, V. M.; Sawicky, M. J.; Bolton, P. H. *Nucleic Acids Res.* **1999**, 27, 2860−2867. (i) Schultze, P.; Smith, F. W.; Feigon, J. *Curr. Biol.* **1994**, *2*, 221−223. (j) Formation of a five-stranded cation-stabilized DNA assembly has recently been suggested, see: Chaput, J. C.; Switzer, C. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 10614−10619. i-DNA asemblies, see: (k) Gehring, K.; Leroy, J.-J.; Gueron, M. *Nature* **1993**, *363*, 561−565. (l) Nonin, S.; Leroy, J.-L. *J. Mol. Biol.* **1996**, *261*, 399−414. (m) Kang, C.; Berger, I.; Lockshin, C.; Ratliff, R.; Moyzis, R.; Rich, A. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 11636−11640. (n) Gueron, M.; Leroy, J.-L. *Curr. Opin. Struct. Biol.* **2000**, 10, 326−331.

10.1021/ja002656y CCC: $20.00
Published on Web 03/15/2001

parallel or antiparallel quadruplex assemblies, depending on the sequence connecting the guanine stretches and the choice of stabilizing cations.[2] This structural variety could conceivably contribute to regulatory mechanisms in biological processes. Thus, toward a more complete understanding of these DNA quadruplex conformations and the mechanisms involved in their interconversion, it is necessary to obtain detailed information about the factors governing the stability of G-DNA quadruplexes and the role of specific cations therein.

The evident dynamic properties of the G-DNA quadruplex render this motif particularly attractive to computational studies to complement the experimental approaches.[3] The recent years have seen a considerable qualitative improvement of the molecular dynamics simulation technique,[4−6] which now allow for carrying out reliable and extensive MD studies of hydrated nucleic acid conformations.[7] Utilizing these methodological advances, we have recently reported a large-scale MD study of parallel and antiparallel G-DNA quadruplex molecules.[8] The focus of that analysis was the guanine stem of the molecules built by four adjacent stacked guanine quartets, the role of monovalent cations in stabilizing four-stranded G-DNA, and the effect of selected nucleobase substitutions in the stem. Our studies showed that the all-guanine quadruplex stem is a rigid molecular assembly with the cations being essential for its stabilization. The MD results were in excellent agreement with high-resolution X-ray and NMR studies, and the simulations provided additional insight into the structure and dynamics of the molecules studied.[8]

In the present contribution, we extend our previous studies of four-stranded DNA assemblies[6g,8] by characterizing mixed G-DNA quadruplexes containing guanine/cytosine GCGC quartets in addition to conventional all-guanine G-quartets. Mixed quadruplexes are considerably less well characterized by atomic resolution experiments than all-guanine G-DNA. In particular, an atomic resolution X-ray structure does not exist for a mixed guanine quadruplex molecule incorporating GCGC quartets, which renders the use of molecular dynamics techniques here particularly attractive. Notwithstanding, structures of G-DNA molecules containing mixed GCGC quartets have been characterized in a series of high-resolution NMR analyses (Figure 1).[9] The NMR geometries are important to initiate MD studies, as the current status of molecular modeling is not yet sufficiently advanced to propose the folding of the structures entirely de novo. The predominant focus of our present investigation is, besides the basic dynamical properties of the quadruplexes containing mixed GCGC quartets, the role of monovalent Na^+ and K^+ cations stabilizing the quartet stems, features which could not be directly revealed by the experimental approaches. Here, we report 20 independent simulations of a variety of mixed guanine quadruplex molecules incorporating GCGC quartets, on a time scale 2.5−12 ns. The simulations provide a systematic analysis of a wide range of properties of these mixed quadruplex molecules, including the distribution and dynamics of monovalent cations in both stem and loop regions. The length of our simulations reaches a total of 90 ns, rendering this study the most extensive MD investigations carried out on nucleic acids so far. We believe that these simulations provide a largely converged characterization of the properties of quadruplex G-DNA molecules containing mixed GCGC quartets.

Methods

All calculations were carried out using the AMBER5.0 program[10] and the Cornell et al. force field.[5a] Where indicated, a more recent version of the Cornell et al. force field (PARM98) was employed which uses a different parametrization of the sugar pucker parameters.[5c] The nucleic acid molecules were surrounded by a periodic box of water molecules described by the TIP3P potential[11] extended to a distance of 10 Å from any solute atom. The number of explicit water molecules included in the simulations varied from 1800 to 4500, depending on the solute molecule. The nucleic acid molecules were neutralized by Na^+ or K^+ cations using standard parameters for the Cornell et al. force field (van der Waals radii of 1.868 Å for Na^+ and 2.658 Å for K^+, the corresponding well depths were 0.00277 and 0.000328 kcal/mol).[3a,5a]

(3) (a) Hardin, C. C.; Ross, W. S. *J. Am. Chem. Soc.* **1994**, *116*, 6070−6080. (b) Mohanty, D.; Bansal, M. *Biophys. J.* **1995**, *69*, 1046−1067. (c) Tohl, J.; Eimer, W. *Biophys. Chem.* **1997**, *67*, 177−186. (d) Strahan, G. D.; Keniry, M. A.; Shafer, R. H. *Biophys. J.* **1998**, *75*, 968−974. (e) Mohanty, D.; Bansal, M. *Biopolymers* **1994**, *34*, 1187−1211. (f) Strahan, G. D.; Shafer, R. H.; Keniry, M. A. *Nucleic Acids Res.* **1994**, 22, 5447−5455. (g) Balagurumoorthy, P.; Brahmachari, S. K.; Mohanty, D.; Bansal, M.; Sasisekharan, V. *Nucleic Acids Res.* **1992**, *20*, 4061−4067. (h) Gu, J.; Leszczynski, J.; Bansal, M. *Chem. Phys. Lett.* **1999,** *311*, 209−214. (i) Chowdhury, S., Bansal, M. *J. Biomol. Struct. Dyn.* **2000**, *18*, 11−28. (j) Gu, J.; Leszczynski, J. *J. Phys. Chem. A* **2000**, *104*, 6308−6313. (k) Read, M. S.; Neidle, S. *Biochemistry* **2000,** *39*, 13422−13432. (l) Gresh, N.; Pullman, B. *Int. J. Quantum Chem., Quantum. Biol. Symp.* **1986**, *12*, 49−56. (m) Meyer, M.; Steinke, T.; Brandl, M.; Suhnel, J. *J. Comput. Chem.* **2001**, *22*, 109−124.

(4) (a) York, D. M.; Darden, T.; Pedersen, L. G. *J. Chem. Phys*. **1993**, *99*, 8345−8348 (b) Lee, H.; Darden, T.; Pedersen, L. *Chem. Phys. Lett*. **1995**, *243*, 229−234. (c) Cheatham, T. E., III; Miller, J. L.; Fox, T.; Darden, T. A.; Kollman, P. A. *J. Am. Chem. Soc*. **1995**, *117*, 4193−4194.

(5) (a) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M., Jr.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. *J. Am. Chem. Soc*. **1995**, *117*, 5179−519. (b) Langley, D. R. *J. Biomol. Struct. Dyn.* **1998**, *16*, 487−509. (c) Cheatham, T. E., III; Cieplak, P.; Kollman, P. A. *J. Biomol. Struct. Dyn.* **1999**, *16*, 849−866. (d) Follope, N.; MacKerrel, A. D., Jr. *J. Comput. Chem.* **2000**, *21*, 86−104. Hobza, P.; Kabeláč, M.; Šponer, J.; Mejzlík, P.; Vondrášek, J. *J. Comput. Chem.* **1997**, *18*, 1136−1150. (e) Cieplak, P. *Encyclopedia Comput. Chem*. Schleyer, P. v. R., Allinger, N. L., Clark, T., Gasteiger, J., Kollman, P. A., Schaefer, H. F., III, Schreiner P. R., Eds.; John Willey & Sons: Chichester, UK, 1998; pp 1922−1930.

(6) (a) Young, M. A.; Jayram, B.; Beveridge, D. L. *J. Am. Chem. Soc*. **1997**, *119*, 59−69 (b) Young, M. A.; Beveridge, D. L. *J. Mol. Biol*. **1998**, *281*. (c) Srinivasan, J.; Cheatham, T. E., III; Cieplak, P.; Kollman, P. A.; Case, D. A. *J. Am. Chem. Soc*. **1998,** *120*, 9401−9409. (d) Feig, M.; Pettitt, B. M. *Biophys. J*. **1999**, *77*, 1769−1781. (e) Simmerling, C.; Fox, T.; Kollman, P. A. *J. Am. Chem. Soc*. **1998**, *120*, 5771−5782. (f) Sprous, D. M.; Young, A.; Beveridge, D. L. *J. Mol. Biol.* **1999**, *285*, 1623−1632. (g) Špaèková, N.; Berger, I.; Egli, M.; Šponer, J. *J. Am. Chem. Soc.* **1998**, *120*, 6147−6151. (h) Shields, G. C.; Laughton, C. A.; Orozco, M. *J. Am. Chem. Soc*. **1997**, *119*, 1463−1469. (i) Cubero, E.; Sherer, E. C.; Luque, F. J.; Orozco, M.; Laughton, C. A. *J. Am. Chem. Soc*. **1999**, *121*, 8653−8654. (j) Cheatham, T. E., III; Kollman, P. A. *J. Am. Chem. Soc.* **1997**, *119*, 4805−4825. (k) Lankaš, F.; Šponer, J.; Hobza, P.; Langowski, J. *J. Mol. Biol.* **2000**, *299*, 907−922. (l) Štefl, R.; Koča, J. *J. Am. Chem. Soc.* **2000**, *122*, 5025−5033. (m) Trantírek, L.; Štefl, R.; Vorlíčková, M.; Koča, J.; Sklenář, V. J. Kypr, J. *J. Mol. Biol.* **2000**, *297*, 907−922. (n) Špačková, N.; Berger, I.; Šponer, J. *J. Am. Chem. Soc.* **2000**, *122*, 7564−7572. (o) Konerding, D. E.; Cheatham, T. E., III; Kollman, P. A.; James, T. L. *J. Biomol. NMR* **1999**, *13*, 119−131. (p) Florian J.; Goodman, M. F.; Warshel, A. *J. Phys. Chem B* **2000**, *104*, 10092−10099. (q) Šponer, J.; Florián, J.; Ng, H.-L.; Šponer, J. E.; Špačková, N. *Nucleic Acids Res.* **2000**, *28*, 4893−4902. (r) Nagan, Kerimo, S. S.; Musier-Forsyth, K.; Cramer, C. J. *J. Am. Chem. Soc*. **2000**, *121*, 7310−7317. (s) Nagan, M. C.; Beuning, P.; Musier-Forsyth, K.; Cramer, C. J. *Nucleic Acids Res.* **2000**, *28*, 2527−2534.

(7) (a) Louise-May S.; Auffinger, P.; Westhof, E. *Curr. Opin. Struct. Biol*. **1996**, *6*, 289−298. (b) Cheatham, T. E., III; Kollman, P. A. *Annu. Rev. Phys. Chem.* **2000**, *51*, 435−471. (c) Cheatham, T. E., III; Brooks, B. R. *Theor. Chem. Acc.* **1998**, *99*, 279−288. (d) Beveridge, D. L.; McConnell K. J. *Curr. Opin. Struct. Biol*. **2000**, *10*, 182−196.

(8) (a) Špačková, N.; Berger, I.; Šponer, J. *J. Am. Chem. Soc*. **1999**, *121*, 5519−5534. (b) Štefl, R.; Špačková, N.; Berger, I.; Koča, J.; Šponer, J. *Biophys. J.* **2001**, *80*, 455−468.

(9) (a) Kettani, A.; Kumar, R. A.; Patel, D. J. *J. Mol. Biol.* **1995**, *254*, 638−656. (b) Kettani, A.; Bouaziz, S.; Zhao, H.; Jones, R. S. *J. Mol. Biol*. **1998**, *282*, 619−636. (c) Bouaziz, S.; Kettani, A.; Patel, D. J. *J. Mol. Biol*. **1998**, *282*, 637−652.

(10) Pearlman, D. A.; Case, D. A.; Caldwell, J. W.; Roos, W. S.; Cheatham, T. E., III; Ferguson, D. M.; Seibel, G. L.; Singh, P.; Weiner, P.; Kollman, P. A. *AMBER Program Package*; UCSF: San Francisco, CA.

(11) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926−937.

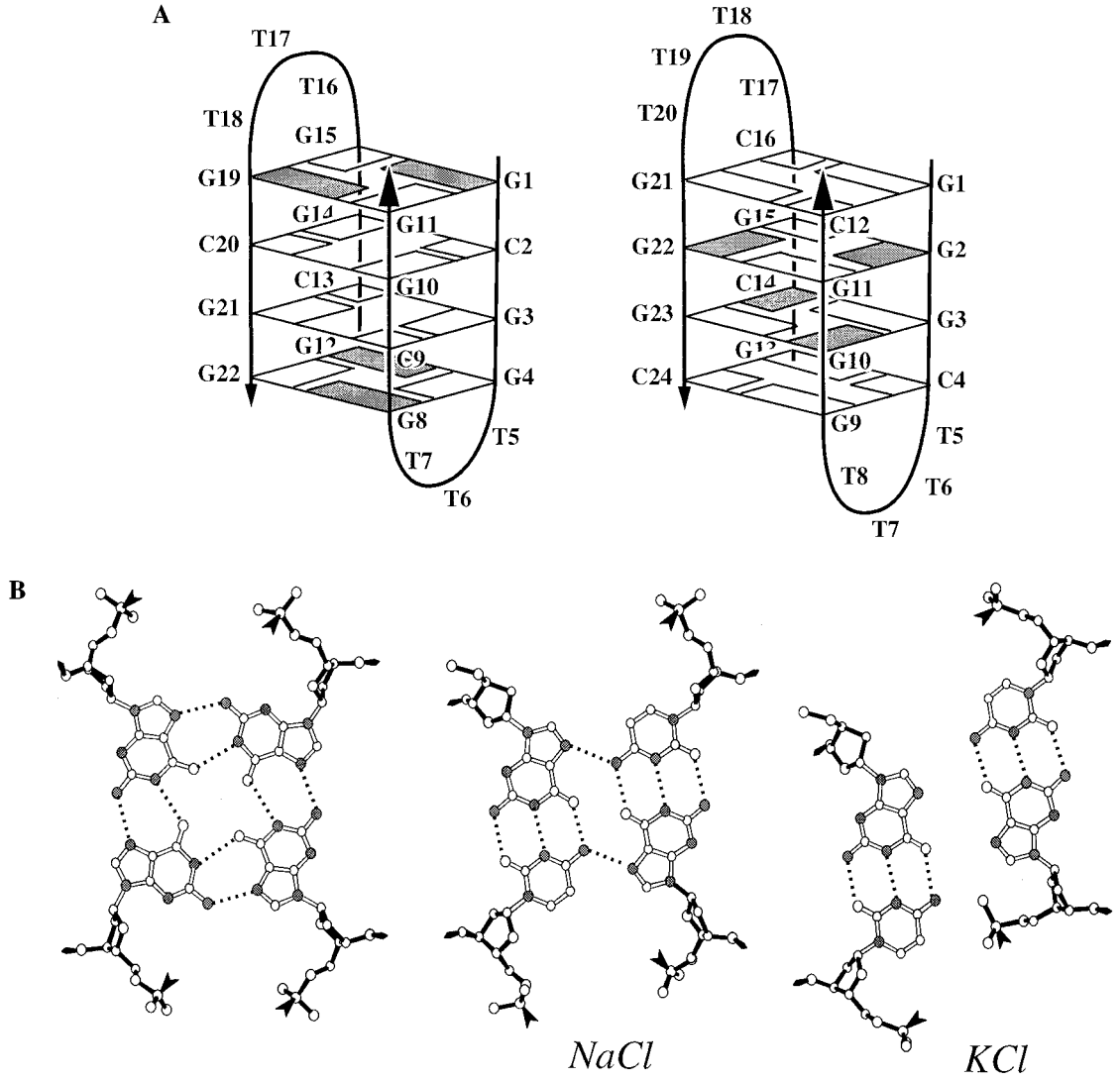

Figure 1. (A) Topology and residue numbering of four-stranded molecular assemblies characterized in NMR studies. On the left: d(GCG-GTTTGCGG)$_2$ (adapted from ref 9a) on the right: d(GGGCTTTTGGGG)$_2$ (adapted from refs 9b,c). Bases in *syn* are shaded, bases with *anti* glycosidic bond conformation are drawn as open boxes. (B) Base interactions in the quadruplex stems:[9] all-guanine GGGG quartet (left), "closed" structure of the mixed GCGC quartet (center) as observed in the presence of NaCl, "sheared" structure of the mixed GCGC quartet (right) in the presence of KCl. Strand polarity is indicated by arrows, H-bond interactions are drawn with dashed lines, nitrogens are shaded.

The concentrations of cations in our simulations were in the range of 0.22−0.25 M for systems without loops and 0.28−0.30 M for systems with stems only. Three of the initial simulations were carried out with a reduced radius of the Na^+ cation (1.618 Å) with the aim to improve the mobility of the ions in the channel.[8a] This modified Na^+ parameters are derived from our earlier study on G-DNA quadruplex molecules where we had observed that Na^+ ion in its original parametrization appears to be oversized when being placed into the quadruplex channel.[8a] Nevertheless, all simulations proved that even the larger K^+ cation is capable of passing through the plane of both guanine and mixed quartets on the time scale of the simulation. Internal cations were added into the channel manually into the centers of the cavities between consecutive guanine or mixed GCGC quartets. When placing cations into the loops, we located them within a cavity formed by ribose O4′ and O2 keto oxygen atoms of the thymidine residues. The remaining cations were initially positioned into the most negative locations using Coulombic potential terms with the LEAP[12a] module. One simulation has been carried out in the presence of NaCl, by adding 50 additional Na^+ and 50 additional Cl^- ions using LEAP, leading to Na^+ concentration of 0.98 1 M. (Parameters of Cl^-: radius of 2.47 Å, well depth of 0.1 kcal/mol). Calculations were carried out using the Sander[12b] module with SHAKE[12c] on the hydrogen atoms with a tolerance of 0.0005 Å and a 2 fs time step. A 9 Å cutoff was applied to Lennard-Jones interactions. Simulations were performed using the Berendsen temperature coupling[12d] algorithm with a time constant of 0.2 ps. The nonbonded pair list was updated every 10 steps. Equilibration started by 1000 steps of minimization, with the atomic positions of the nucleic acid fixed. All subsequent simulations were performed using the particle mesh Ewald method (PME).[4,7] The PME charge grid spacing was approximately 1.0 Å, and the charge grid was interpolated using a cubic B-spline with the direct sum tolerance of 10^{-6} at the 9 Å direct space cutoff. The size of the PME charge grid was chosen to be a product of powers 2, 3, and 5. For dynamics runs after minimizations initial velocities were assigned from a Maxwellian distribution. Equilibration was continued by 50 ps of PME dynamics (except for the high-salt

(12) (a) LEAP is a preparatory program which allows the preparation of input for the molecular modeling programs using graphical interfaces. See http://www.amber.ucsf.edu/amber/. (b) SANDER is the basic energy minimizer and molecular dynamics program. Molecular dynamics calculations are carried out by integrating Newtonian equations of motion. For a detailed description, see: http://www.amber.ucsf.edu/amber/. (c) Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. C. *J. Comput. Phys.* **1997**, *23*, 327−341. (d) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; DiNola, A.; Haak, J. R. *J. Chem. Phys.* **1984**, *81*, 3684−3690. (e) Harvey, S. C.; Tan, R.-K., Z.; Cheatham, T. E., III. *J. Comput. Chem.* **1998**, *19*, 726−740. (f) Carnal is a molecular dynamics analysis program. It is used for geometrical measurements, RMS coordinate fitting, trajectory averaging, and other structural analyses of MD trajectories. See http://www.amber.ucsf.edu/amber/.

Table 1. Simulations of DNA Quadruplex Molecules Containing Mixed GCGC Quartets[a]

simulation	quadruplex	ions	N_i/M_i	N_e/M_e	duration [ns]
A1[b]	$d(GCGG)_4$	Na^+	2	2	3.1
A2[b]	$d(GCGG)_4$	small Na^+	2[j]	3	5.0
A3[b]	$d(GCGG)_4$	small Na^+	2	2	2.5
A4[b]	$d(GCGG)_4$	K^+	2	2	5.0
A5[b]	$d(GCGG)_4$	K^+	2[j]	1	4.5
A6[b]	$d(GCGG)_4$	Na^+	0	1	2.5
A7L[b]	$d(GCGGTTTGCGG)_2$	Na^+	2/0	2/0	2.5
B1[c]	$d(GGGC)_4$	small Na^+	3	2	2.5
B2[c]	$d(GGGC)_4$	Na^+	0	0	2.5
B3L[d]	$d(GGGCTTTTGGGC)_2$, *KCl*	K^+	3/2	2/0	5.0
B4L[d]	$d(GGGCTTTTGGGC)_2$, *KCl*	Na^+	3/2	2/0	5.0
B5L[c]	$d(GGGCTTTTGGGC)_2$, *NaCl*	Na^+	3/0	2/1[k]	5.0
B6L[c]	$d(GGGCTTTTGGGC)_2$, *NaCl*	K^+	3/0	2/0	5.0
B7L[d,e]	$d(GGGCTTTTGGGC)_2$, *KCl*	K^+	3/2	3/0	12.0
B8L[d,e]	$d(GGGCTTTTGGGC)_2$, *KCl*	Na^+	3/2	2/0[l]	10.0
B9L[d,e,f]	$d(GGGCTTTTGGGC)_2$, *KCl*	K^+	3/2	3/0	5.0
B10L[e,g]	$d(GGGCTTTTGGGC)_2$, *KCl*	K^+	2/1	2/1	3.0
B11L[d,h]	$d(GGGCTTTTGGGC)_2$, *KCl*	NaCl	3/2	2/0	5.0
C1[i]	$d(GCGC)_4$	Na^+	0	0	2.5
C2[i]	$d(GCGC)_4$	Na^+	3	1	2.5

[a] Designation of the simulation: A = mixed GCGC quartets inside the stem, B = guanine quartets inside the stem, C = four GCGC quartets, L = the simulation includes the thymine loops. N_i/M_i = number of ions in the channel/loops in the initial structure. N_e/M_e = number of ions in the channel/loops at the end of the simulation. [b] Initial structures based on the NMR structure of $d(GCGGTTTGCGG)_2$.[9a] [c] Initial structures based on NMR structure of $d(GGGCTTTTGGGC)_2$ obtained in the presence of NaCl.[9b] [d] Initial structures based on NMR structure of $d(GGGCTTTTGGGC)_2$ obtained in the presence of KCl.[9c] [e] Modified Cornell et al.force field (parm98) has been utilized.[5c] [f] Repetition of simulation B7L with a modified equilibration protocol. [g] Starting structure: a snapshot taken from B7L at 10.0 ns. [h] High-salt simulation. [i] Hypothetical structure with four consecutive GCGC quartets. [j] One additional cation placed directly above the channel. [k] Cation oscillating between the loop and the outer quartet plane. [l] 2/1 after 5 ns.

simulation where the equlibration has been extended to 200 ps), with the position of the nucleic acid fixed. Subsequently, 1000 steps of minimization were carried out with 25 kcal/(mol·Å^2) restraints placed on all solvent atoms, continued by 3 ps MD simulation using the same restraint. This equilibration was followed by five rounds of a 1000-step minimization with solute restraints reduced by 5 kcal/(mol·Å^2) in the course of each round. Then 20 ps of MD followed, with the system heated from 100 to 300 K over 2 ps (except for the high salt simulation where the heating-up period reached 100 ps). Equilibration was continued by several nanoseconds of production simulation. Coordinates were written to trajectory files after each picosecond. The center of mass velocity was removed during the production dynamics periodically in intervals of 10 ps.[12e] The results were analyzed using the Carnal[12f] module of AMBER5.0 with no extra processing of the averaged structures.

Results

Twenty independent simulations have been performed in the course of this study, and these are listed in Table 1. The table shows the nucleotide sequence of the simulated molecules, which either consisted of only quadruplex stems or included thymine loops (marked by L). Further, the length of each simulation is given in nanoseconds. In addition to the cation type (Na^+ or K^+), the number of cations occurring in the central channel (N) and within the loops (M) is listed; subscript i and e mean the starting and final structure, respectively. In the following, the simulations are commented on in detail. We commence by analyzing the role of the monovalent cations in stabilizing the individual four-stranded assemblies that are investigated. Then, we will comment on other properties of these molecules such as geometries of individual quartet arrangements, loop geometries, and other structural features. Final structures of all simulated systems with thymine loops can be found in the Supporting Information, represented via stereofigures as well as using the PDB coordinate files.

Molecular Dynamics of Quadruplexes with Two Adjacent Mixed GCGC Quartet Layers in the Center of the Quadruplex Stem. The first simulation series utilized starting coordinates that are based on a high-resolution NMR study of a quadruplex formed by the sequence $d(GCGGTTTGCGG)_2$.[9a] In this structure, two oligonucleotides, each arranged as a fold-back hairpin, assemble by head-to-tail dimerization to form a quadruplex (Figure 1a). The antiparallel stem is built by two central mixed guanine/cytosine quartet layers sandwiched between two all-guanine quartets. This quadruplex stem is capped both at top and bottom by a three-nucleotide loop with the sequence -TTT-. The simulations are denoted A1−A7L. In the simulations A1−A6, only the stem of the molecule was considered. Figure 2 schematically depicts all cation rearrangements observed during the simulations.

In the first simulation (A1), we investigated the $d(GCGG)_4$ stem with two Na^+ cations initially placed into the cavities between the mixed GCGC quartet layers and the all-guanine quartets, while the central cavity between the two mixed guanine/cytosine quartets remained vacant. This disposition of ions was based on our assumption that mixed guanine/cytosine quartets should interact less favorably with cations as all-guanine quartets (see also below) and disposition of cations in the outer cavities reduces the electrostatic intercation repulsion. The two cations remained in their initial positions during the entire simulation. The structure is very stable and rigid, similar to that in previously reported simulations of parallel and antiparallel G-DNA molecules (Figure 3).

In the course of simulation A2, we again placed two Na^+ cations into the channel; however, one of them was placed between the mixed quartets (Figure 2), and one of the outer cavities was vacant in the starting structure.[13] We placed a third cation below the terminal GGGG layer, facing the empty cavity. This initially vacant cavity was immediately filled by the cation. The cations then redistributed in such a way that the inner cation remained sandwiched between the two mixed quartets while the outer two cations occupied positions in planes of the G-quartets.

Next, we repeated simulation A2 with the difference now that no external cation was placed initially close to the vacant cavity (A3).[13] During this simulation, one cation remained again sandwiched between the mixed GCGC quartets in the center of the quadruplex stem, and actually did not move into the adjacent—presumably more electronegative—cavity present between one GCGC and one GGGG quartet layer. Contrary to our expectation, the vacant cavity was not invaded by a water molecule. Only the channel entrance was hydrated during the entire simulation, with residency times of individual water molecules up to 300 ps. The cation at the other end of the stem (upper cavity in Figure 2) adjusted its position again into the plane of the outer G-quartet layer.

We have repeated the simulations A1 and A2 with larger K^+ cations instead of Na^+ (A4 and A5 in Table 1). Simulation A4 shows essentially the same trajectory as was observed in simulation A1 (Figure 2), with the exception of a temporary deformation (2.5−3.5 ns) of one of the outer all-guanine quartets, with one residue moving slightly out of the quartet plane. However, this perturbation was subsequently fully

(13) This simulation was carried out with slightly reduced radii of the Na^+ ions, with the aim to increase the mobility of the cations and to allow their faster movement throughout the channel (see Methods section).

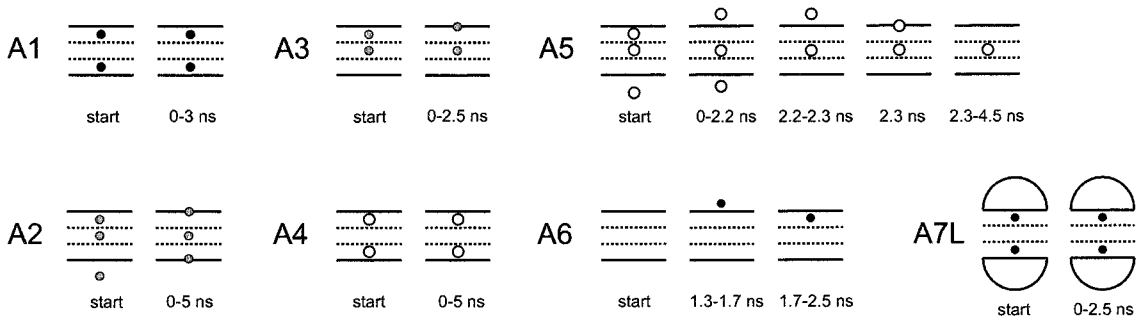

Figure 2. Schematic illustration of major rearrangements of cations in simulations A1−A7L (cf. Table 1). GCGC quartet layers are drawn as dashed lines, all-guanine quartets are depicted with black lines. Thymidine loops on top and bottom of the quartet stack are shown as half-circles. Na^+ or K^+ ions are drawn as black or white spheres, respectively. Na^+ with reduced ion radius (see Methods and Table 1) are drawn as gray spheres.

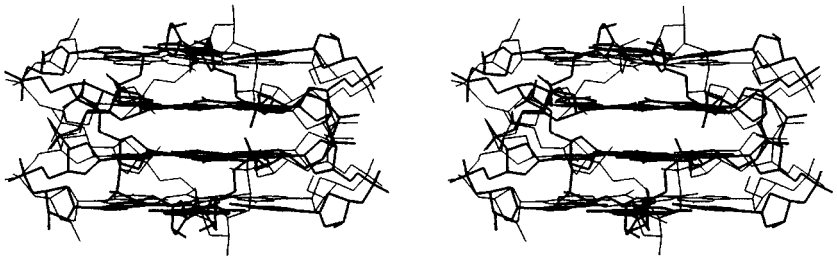

Figure 3. Stereo overlay of experimental and averaged (1−2.5 ns) theoretical structures of the $(GCGG)_4$ stem, simulation A1. Theoretical structure is drawn with thick lines.

repaired. Simulation A5 in contrast showed a different trajectory from that of the corresponding simulation with Na^+ (A2). The simulation was initiated with one K^+ cation in the central cavity, another in one of the outer cavities, and the third K^+ positioned outside the channel facing the entrance to the empty cavity. The first cation remained in the central cavity throughout the entire simulation. At the beginning of the simulation, the second K^+ relocated to the outside of the central channel, slightly away from the outer quartet plane, while the third cation at the opposite end of the molecule adopted a symmetrical position. However, within the period of 2.2−2.3 ns, both outer cations were released to the solvent. The difference between simulations A2 (Na^+) and A5 (K^+) can obviously be attributed to cation size. Nevertheless, we assume (based on simulations of other structures in the presence of K^+, see below) that simulation A5 does probably not show the optimal cation arrangement of K^+ in the stem.

In simulation A6 (Na^+), we placed no cations into the central channel. During the first part of the simulation we observed marked geometrical fluctuations of the vacant quadruplex stem structure. The two outer cavities were swiftly hydrated, although the water molecules apparently preferred to stay somewhat closer to the outer GGGG quartets. Residency times of water molecules in the cavities were in the range of 100−1300 ps. No water molecule entered the space between the two mixed quartets. However, *the central channel of the d(GCGG)$_4$ stem did attract a cation, which penetrated at 1.7 ns through the channel entry into the cavity between this outer G-quartet and the adjacent mixed GCGC quartet (Figure 2).* This was a two-step process, with the Na^+ first hovering nearby the channel entry for about 0.4 ns. Then, the cation and the water molecules that were at this time located in the cavity exchanged their positions. This was accompanied by an instant rigidification of the whole quadruplex stem. Let us emphasize that the Na^+ cation was initially localized near a phosphate group in one of the grooves of the quadruplex stem, at a comparatively large distance from the channel entry site.

Finally, we carried out a 2.5 ns simulation for the complete $d(GCGGTTTGCGG)_2$ hairpin dimer quadruplex (simulation A7L), in the presence of Na^+. The cations were initially distributed identically as in the simulation A1. Inclusion of the loops influenced neither the behavior of the stem nor the cation distribution. This simulation resulted in RMSd values along the trajectory between experimental and theoretical coordinates of around 2 Å (Figure 4).

Molecular Dynamics of Mixed Quadruplexes with GCGC Quartet Layers at the Ends of the Four-Stranded Stem. Our second extended set of simulations is based on the high-resolution NMR structure of the quadruplex formed by the sequence d(GGGCTTTTGGGC), which was determined in the presence of either KCl or NaCl.[9b,c] The studies revealed a formation of a mixed quadruplex (again an antiparallel hairpin dimer) formed by foldback hairpins of two d(GGGCTTTTGGGC) molecules arranged in a head-to-tail orientation (Figure 1). Here, however, two outer-mixed GCGC quartets envelop two GGGG quartets in the center of the stem. This stem is capped by two four-thymidine loops. The NMR studies indicated that the cations affect the geometry of the mixed GCGC layers (Figure 1), and different loop geometries were predicted in the presence of either KCl or NaCl (Figure 5). It was further suggested that in the presence of KCl (this molecule will be referred to as *KCl* structure in the following) the molecule accommodates a cation within each of the two loops. This would result in a string of five consecutive cations lining up along the entire *KCl* structure, with three K^+ cations filling the three cavities present in the quadruplex stem, and two K^+ ions both on top and bottom of the quadruplex stem tightly bound to the loop regions. On the other hand, it was suggested that in the presence of NaCl (*NaCl* structure), no cations could be accommodated within the loop (Figure 5). We investigated the distribution of cations in the $d(GGGCTTTTGGGC)_2$ quadruplex assembly, utilizing both the *KCl* and *NaCl* NMR geometries as starting structures, and varying the number and type of cations embedded within the initial structure. The individual simulations are listed in Table

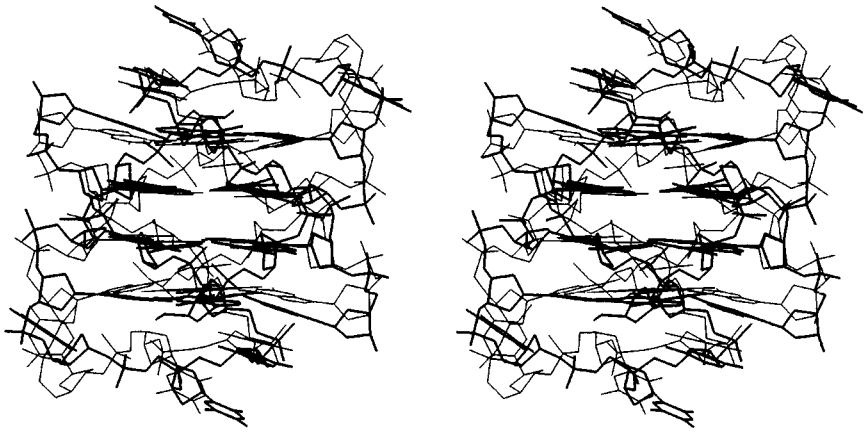

Figure 4. Stereo overlay of experimental and averaged (1−2.5 ns) theoretical structures of the d(GCGGTTTGCGG)$_2$. The theoretical structure is drawn with thick lines.

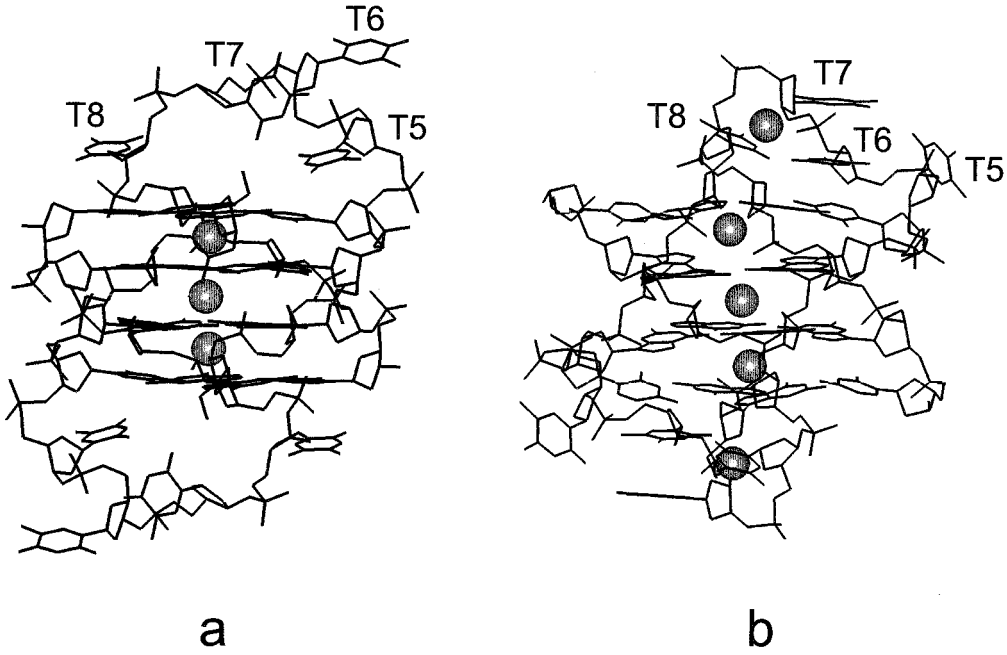

Figure 5. NMR structure of d(GGGCTTTTGGGC)$_2$ obtained in the presence of (a) NaCl and (b) KCl. These structures are designated as *NaCl* and *KCl* throughout the paper. Note the marked differences in the loop regions. The internal cations are positioned according to the original literature suggestion.[9b,c]

1 and denoted B1 to B11L, with L indicating the presence of the four-thymidine loop. Figure 6 summarizes all major cation relocations in these simulations.

We carried out two 2.5 ns simulations (B1[13] and B2) of the antiparallel quadruplex stem with the four-thymidine loops omitted from the molecule. First, we placed three Na^+ cations into the three cavities. This distribution of cations proved to be unstable along the trajectory, and at 0.7 ns one outer cation had actually left the channel. The two remaining cations adopted positions in the planes of the central two all-guanine quartets. Then we carried out a simulation while removing all cations from the channel (B2 in Table 1). The trajectory showed large fluctuations, as all G-DNA simulations carried out to date with vacant channel,[8] and we did not observe any cation entry into the channel.

Molecular Dynamics of d(GGGCTTTTGGGC)$_2$ Antiparallel Hairpin Quadruplex Using Either NaCl or KCl NMR Structures as Starting Geometries. We have then carried out six simulations considering the whole d(GGGCTTTTGGGC)$_2$ structure, with initial simulation time of 5 ns (B3L−B8L). We started with the *KCl* NMR structure and positioned five K^+ cations into the channel and loops (simulation B3L, see Figure 6). At the end of the simulation, however, only two cations remained in the channel, one of them occupying the central cavity. In the next simulation (B4L), we used the same *KCl* starting arrangement, but the simulation was carried out in the presence of Na^+ cations. Again, only two cations remained in the channel (cf. Figure 6).

Then, we used the alternative experimental molecule (*NaCl* structure) as the starting arrangement, with three cations located in the channel and without cations inside the loops. During the B5L simulation in the presence of Na^+, one of the outer cations shifted away from the center of the cavity at 0.6 ns and positioned itself exactly in the center of that mixed GCGC quartet layer. Then, until the end of the simulation, this cation oscillated between this position and a position within the neighboring four-thymidine loop, located somewhat above the quartet layer. As will be shown below, this distribution of cations is accompanied with a formation of a well defined, specific loop geometry. It is likely that in the absence of the four-thymidine loop, the cation would have left the quadruplex stem entirely. In the case of the B6L simulation with K^+, one cation left the

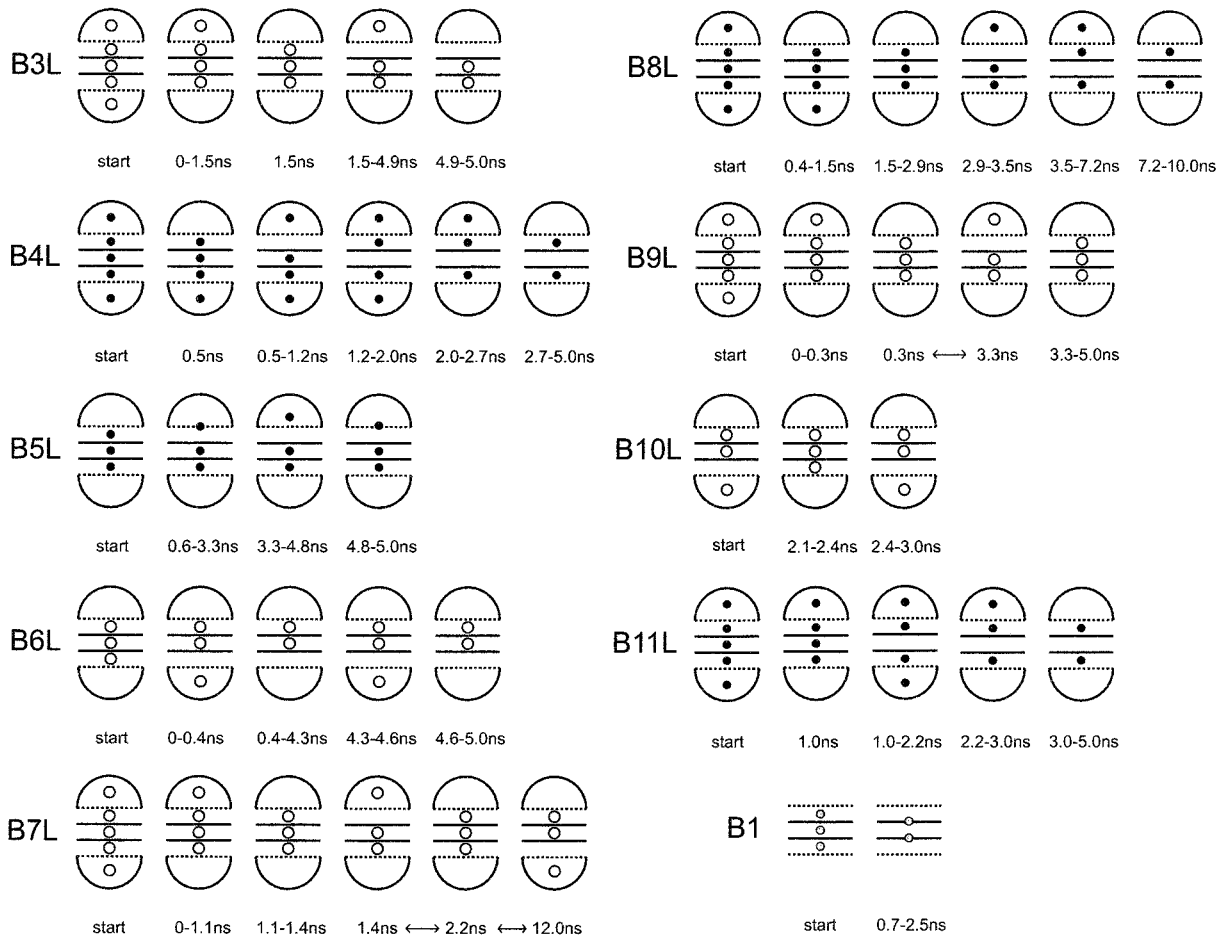

Figure 6. Schematic illustration of major rearrangements of cations in simulations B1−B11L (cf. Table 1). Patterns analogous to Figure 2 (see above). The arrows (↔) indicate multiple oscillations between the two states.

channel during the early stages of the simulation. This K^+ cation was then locked inside the loop for <0.5 ns, but then it left the pocket formed between the four-thymidine loop and the quadruplex stem and was located in the bulk solvent around the loop. Nevertheless, the same cation returned back into the loop at 4.3 ns, and stayed there for additional 0.3 ns. After this, the cation again left the loop area. This behavior of a cation distribution is associated with a specific development of the loop geometry (see below).

We then decided to carry out additional simulations to test the reproducibility of our predictions. We repeated simulations B3L and B4L, assuming the *KCl* structure starting geometry with five internal cations. However, we now used a more recent version of the Cornell et al. force field with modified internal parameters for sugar puckers (PARM98, see Methods section).[5c] Change of sugar puckers can obviously have an effect on loop geometries. These two simulations (B7L and B8L) were initially 5 ns long. During the simulation with K^+ cations (B7L), one cation left the four-thymidine loop cage within the first 0.1 ns. The other loop cation left its position at 1.1 ns: however, the loop region was then filled by a cation coming from the channel. At 2.2 ns, this K^+ cation returned back into the channel. At 4.8 ns, the cation originally placed within the loop and then expelled into the solvent temporarily tested its original position inside the loop again. The simulation B8L with Na^+ cations revealed again two cations leaving the loop region during the first 2 ns. At 2.9 ns, one cation left the channel and occupied a position within the adjacent loop close to the outer mixed GCGC quartet layer. At 3.5 ns, this cation relocated to a position buried deep inside the loop, while the cations inside of the channel occupied the outer cavities, being considerably closer to the guanine quartet planes.

Both of these simulations with the modified Cornell et al. force field (B7L and B8L) show after 5 ns still three cations inside the solute molecule, in contrast to two cations observed at the end of the corresponding simulations utilizing the older variant of the force field (B3L and B4L). Thus, we extended BL7 to 12 and BL8 to 10 ns. During the prolongation of B8L the Na^+ cation that had been localized after 5ns in the loop finally left this position. The simulation conclusively led to a structure equivalent to the outcome of the corresponding simulation B4L. On the other hand, when we extended simulation B7L (Table 1), three K^+ cations remained in the channel of the structure all through the end of the extended trajectory. The outer cations, however, three times tested positions outside the ion channel, being temporarily localized within the four-thymidine loops close to the channel entry site (cf. Figure 6). The cations thus appear to be entirely poised to leave the channel in accordance with all other simulations, but the loops obstruct this process. We have then designed two further simulations: Simulation B9L (5 ns) was a repetition of the B7L simulation, but we modified the equilibration protocol to ensure an entirely independent trajectory. The development of simulation B9L was very similar to that of B7L. Two cations were eliminated from the loops soon; however, the remaining three cations remained located inside the solute molecule, one of them testing a position within the loop just above the channel entrance. As the last attempt to expel the third cation from this structure, we have used a snapshot (taken at 10.0 ns) of simulation B7L as a starting point and carried out an entirely new simulation with the whole

equilibration protocol repeated (B10L, 3 ns). We expected that the new equilibration could shake up the structure and free the cation. However, the behavior of simulation B10L remained similar to that of B7L.

High-Salt Simulation of d(GGGCTTTTGGGC)$_2$. One of the reviewers of this contribution pointed out that the instability of the cations inside the solute molecules in our simulations could be due to a low cation concentration in the simulations. However, it should be noted that in fact our simulations with loops were done (considering the number of water molecules and cations included) at cation concentrations around 0.22−0.25 M, that is, higher than in the corresponding NMR experiments by Patel's group (0.1−0.15 M). Nevertheless, we have found the idea of carrying out a high-salt simulation quite interesting and executed an additional simulation of d(GGGCTTTTGGGC)$_2$, assuming the starting geometry with five internal cations (simulation B11L). The simulation has been carried out in the presence of NaCl with a Na^+ concentration of 1M (concentration of Cl^- of ∼0.7 M). As indicated in Figure 6, this simulation again shows the three excessive cations leaving the loops and channel, and is entirely consistent with the preceding simulations.

Molecular Dynamics of a Hypothetical Antiparallel d(GCGC)$_4$ Quadruplex Stem Consisting of Mixed GCGC Quartets Only. To assess the intrinsic stability of GCGC quartet layers, we have carried out a simulation with a strictly hypothetical stem consisting of four consecutive mixed GCGC quartets: antiparallel d(GCGC)$_4$. In the first simulation (C1 in Table 1) no cations were placed into the channel. This simulation showed large fluctuations, and the structure is evidently much less stable than regular cation-stabilized quadruplexes. A second simulation (C2) was initiated with three cations in the cavities; however, only one cation remained in the molecule center after 2.5 ns. The structure swiftly collapsed around the cation trapped inside.

Structural Properties of Quadruplexes Containing Mixed GCGC Quartet Layers. Above, we provided the basic description of the nanosecond-scale dynamics of quadruplex molecules containing mixed GCGC quartet layers. In the following, we will summarize additional properties of the simulated structures.

Movement of Cations Initially Located Outside the Solute Molecule. As shown above, many cations initially located inside the solute left the structure in the course of the simulation. A genuine question is whether some of these vacated sites were not then occupied by cations initially localized outside the channel and the loops. We have analyzed trajectories of all cations, and we have not seen any cation entering the channel from the outside with the exception of simulations A2 and A6 as described above. We have observed a number of times an external cation approaching the loop in such a way that it could eventually be coordinated inside the loop. However, no major event leading to a longer cation coordination has been observed, even in the high-salt simulation B11L.

Geometry of Mixed Guanine/Cytosine GCGC Quartet Layers. The NMR structures show two distinct configurations of the mixed guanine/cytosine GCGC quartets (Figure 1). In the *KCl* structure of d(GGGCTTTTGGGC)$_2$, the mixed quartet is composed of two separate GC Watson−Crick base pairs which are not mutually interconnected by H-bonds. This is explained as a consequence of the presence of large K^+ cations pulling the GC base pairs apart.[9b] This quartet geometry can be called a "sheared" GCGC quartet. The second type of a GCGC quartet ("closed" quartet) was found in the *NaCl* structures of d(GGGCTTTTGGGC)$_2$ and d(GCGGTTTGCGG)$_2$. In this arrangement, the two GC Watson−Crick base pairs are interconnected by bifurcated H-bonds, and the channel becomes very narrow (cf. Figure 1).[9]

Table 2. Interatomic Distances in Å in Two Types of Mixed GCGC Quartets (sheared, closed) Revealed by the NMR Studies[9]

	sheared quartet[a]	closed quartet[a]
O6−O6	3.2 (around 3)	4.6 (4−5)
N4−N4	4.6 (5−6)	8.8 (8−9)
O6−N7	5.4 (around 5)	3.6 (3−4)

[a] Values in parentheses show the distances observed during the MD simulations.

Our simulations show exactly the same two types of mixed quartets. Their geometry is determined by the disposition of cations around the quartet plane. The closed GCGC quartet layer is present all the time when no cation is in contact with the mixed quartet. Large K^+ cations induce a switch to the sheared arrangement when present in the adjacent cavity. The Na^+ cations induce a switch toward the sheared arrangement of the mixed quartet only when closely approaching—or penetrating into—the quartet plane. The mixed GCGC quartet remains in the closed geometry when the Na^+ cation is positioned between an all-guanine quartet and the mixed layer in the usual way, that is being closer to the plane of the all-guanine quartet. The range of the typical theoretical values for the intermolecular distances between the GC base pairs is summarized in Table 2. Transitions between sheared and closed quartets induced by cation re-locations are usually very fast, without intermediates. For a few times, however, we could see a nonsymmetrical intermediate state between closed and sheared quartets, with a lifetime on a scale of hundreds of ps. This geometry is characterized by one short N4(C)···N7(G) hydrogen bond as in the closed quartet, while the other N4(C)···N7(G) distance is considerably larger. The intermediate geometry was seen in simulation B5L and was associated with the oscillation of a Na^+ cation between the thymidine loop and the mixed GCGC quartet layer, accompanied by an unusual loop geometry (see below). While being within the loop, the cation is very close to the channel, and it prevents a relaxing of the quartet into the "closed" geometry, leading to the irregular intermediate. The quartet adopts the fully sheared arrangement when the cation resides exactly within the quartet plane.

In simulation B6L, one of the GC Watson−Crick base pairs is broken at 1.6 ns, and its cytosine is shifted ∼3 Å above the plane of the remaining bases originally involved in the quartet layer. Also this event is associated with a specific loop geometry (see below).

Geometry of the All-Guanine GGGG Quartets. The simulations based on the d(GCGGTTTGCGG)$_2$ structure (A1-A7L) show two basically symmetrical outer all-guanine G-quartets with typical N1···O6 (3.0 Å) and N2···N7 (2.9 Å) H-bonds and diagonal O6···O6 distances across the channel of around 4.3 Å. This is the same geometry as noticed for outer quartets in simulations of all-guanine quadruplex molecules.[8] The NMR structure of d(GCGGTTTGCGG)$_2$, in contrast, shows nonsymmetrical GGGG quartets with diagonal O6···O6 distances of 3.6 and 4.2 Å. *NMR studies often show nonsymmetrical guanine quartets, while MD simulations consistently predict symmetrical G-quartets.*[3d,8] The reason for this discrepancy still needs to be resolved.

The behavior of the inner all-guanine quartets in the simulations is somewhat more complicated. The NMR derived all-guanine quartets are slightly asymmetrical, with diagonal O6···O6 distances of 4.6 and 5.2 Å in the case of the *KCl* structure, and 4.0 and 4.5 Å for the *NaCl* structure. The computed

Table 3. Selected H−Bonding and Stacking Interaction Energies (kcal/mol) in Mixed Quadruplexes

hydrogen bonding energies of quartets			
quartet	total H-bond energy	Coulombic term	van der Waals term
GGGG	−51	−46	−5
GCGC closed[a]	−67(−13)	−66(−10)	−1(−3)
GCGC sheared[a]	−47(+9)	−44(+12)	−3(−3)
stacking between quartets			
stacked system	total stacking energy	Coulombic term	van der Waals term
GGGG···GGGG	−24	+21	−45
GCGC···GCGC	−28	+9	−37
GGGG···GCGC closed	−21	+22	−43
GGGG···GCGC sheared	−20	+23	−43
cation−quartet interactions			
interacting system	total interaction energy	Coulombic term	van der Waals term
GGGG···Na^+	−97	−105	+8
GCGC closed ··· Na^+	−55	−57	+2
GCGC sheared ··· Na^+	−69	−70	+1

[a] Values in parentheses shows the interaction between the two GC Watson−Crick base pairs, i.e., GC···GC interaction.

geometries are affected by the disposition and size of the surrounding cations. When there are two K^+ cations present above and below a particular guanine quartet, the diagonal O6···O6 distance is within the range of 4.1−4.6 Å, and the quartets are basically symmetrical. When only one K^+ cation is present, the O6···O6 distance increases to ∼4.6−4.8 Å, and the quartet becomes slightly asymmetrical. The presence of two adjacent Na^+ cations promotes a formation of bifurcated hydrogen-bonding[8] in the guanine quartet layer. This means that the N1−H1 groups involved in pairing are pointing to between the N2 and N7 atoms of the neighboring guanine base. In most cases, however, only two of the four N1−H groups were oriented in this way, resulting in a slight asymmetry of the quartet. When only one Na^+ cation is contacting a particular guanine quartet layer, standard bonding pattern is observed. Let us note that the bifurcated hydrogen-bonding of all-guanine quartets surrounded by Na^+ at each side was reported in our previous MD study on $d(GGGG)_4$[8] but is not in agreement with the atomic (0.95 Å) resolution X-ray crystal structure of $d(TGGGGT)_4$.[2a,c] The bifurcated bonding is considered to be a subtle artifact of the force field (for more details see ref 8).

Interaction Energies of Bases in the Mixed Quadruplex Stem. We have calculated selected interaction energies in the quadruplex stems (Table 3). The interaction energies were evaluated using the Cornell et al. force field[5a] assuming a dielectric constant of 1, utilizing averaged geometries of appropriate simulated structures having two cations in the channel. Thus, for "closed" mixed quartet layers and all-guanine quartets, the Na^+ cations were residing closer to the guanine quartets, while for a "sheared" GCGC quartet we utilized a geometry with the Na^+ cation close enough to the mixed quartet plane to introduce the sheared pairing.

Among the H-bonded quartets, the closed mixed quartets show the most favorable H-bonding interaction. The stabilization is dominated by the interaction energy of the two GC Watson−Crick base pairs. Interaction between these two pairs is considerably weaker (values in parentheses). Formation of the sheared quartet considerably reduces the H-bonding energy of the GCGC quartet since the GC···GC interaction becomes repulsive. Among stacking arrangements, stacking between two mixed quartets is more favorable compared to the other cases, but the base stacking does not show any pronounced variability in quadruplexes.[3f,8] Finally, concerning the quartet−cation interactions, the interaction between the all-guanine quartet and the cation is clearly the most stable arrangement. The interaction energy between a closed GCGC quartet and Na^+ is less favorable by 40 kcal/mol, although still attractive, ∼−55 kcal/mol. However, as evidenced by the simulations, the cations prefer to be outside the channel fully exposed to the solvent rather than to be coordinated with the mixed quartets. Close approach of a cation toward the plane of the mixed GCGC quartet improves the cation−quartet interaction energy by ∼14 kcal/mol. However, it does not fully compensate for the accompanying loss of H-bonding in a sheared GCGC quartet, which is about 20 kcal/mol.

Geometry of Four-Thymidine Loops in Simulations Assuming the KCl Starting Structure. We have carried out four independent simulations of the $d(GGGCTTTTGGGC)_2$ quadruplex based on the *KCl* NMR structure (Figure 5b) with 5 Na^+ or K^+ cations initially inside and utilizing both variants of the force field (B3L, B4L, B7L, and B8L). As each quadruplex contains two four-thymidine loops (T5−T8 and T17−T20), we are in possession of eight independent loop trajectories. In the NMR structure, T6 is stacked on top of one GC base pair of the adjacent mixed quartet, while T8 is stacked on top of the other GC base pair of this layer. T7 is stacked on T6, while T5 is close to T6 and T7, being, however, positioned perpendicular to these two bases. The methyl group of T7 is tightly packed against the O2 oxygen of T5. The loop T17−T20 has an identical arrangement in the NMR structure (see Supporting Information).

The *KCl* NMR structure shows a single N3(T6/T18)···O2(T8/T20) H-bond. During the equilibration protocol, the hydrogen bond was broken in all loops except for one. During the simulations, formation of H-bonded T−T base pairs was observed in four out of the eight simulated loops within 5 ns, correlated with expulsion of the ions from the loops (see above). Simulations B3L and B8L ended after 5 ns with the second loop showing T18···T20 base pair with: N3(T18)···O2(T20) and O4(T18)···N3(T20) H-bonds. These stable base pairs were established via a two-step process, forming a single O2···N3 bond first and completing the asymmetric base pair after an additional 0.9 and 3.0 ns. In the case of the simulation B4L, a formation of asymmetric base pairs was observed in both loops within the first 2 ns, albeit the H-bonding scheme was reversed to O2(T6/T18)···N3(T8/T20) and N3(T6/T18)···O4(T8/T20). The remaining loops did not form H-bonded base pairs during the first 5 ns, although during simulation B7L a base pair was temporarily formed around 1 ns and broken around 1.2 ns. In general, *all eight loops remained geometrically close to the starting KCl loop geometry and can be considered as one conformational substate, despite the H-bonding differences* (see Supporting Information).

Simulations B7L and B8L had have been extended to 12 and 10 ns, respectively (see above). In the simulation B8L a stable H-bonded T6−T8 base pair was formed at ∼7 ns, immediately after a Na^+ cation had finally been expelled from this loop region. The simulation finished with both loops having H-bonded pairs. It is interesting to note that the adjacent mixed quartet rearranged to a "closed" geometry at 6.8 ns, that is still during the presence of a cation within the loop region. It is due to a subtle shift of T8 above the channel entry site, which hinders the access of the cation to the channel. This was the final micro-rearrangement of the loop prior to eliminating the cation.

The outcome of the extended simulation B7L with 5 K^+ cations is more complicated. Not only did three cations remain here in the channel (see above), but we also have seen the only case of a significant rearrangement of an apical four-thymidine loop in a simulation based on the *KCl* NMR starting geometry. After a cation returned from the loop region back to the channel (see above) at 7.2 ns, the thymine residue T8 also approached the channel entry site. T7 was stacked on top of T6, but T6 became unstacked out of the adjacent quartet layer, leading to these thymines becoming inclined by ∼45° toward the channel entry site. At around 7.8 ns, thymine T7 shifted toward thymine T8 in this loop, and at 8.0 ns T7 and T8 established a full stacking. T8 then basically blocked the channel entry site by its O2 keto oxygen (probably attracted by the cation residing in the central ion channel). The observed behavior of the four-thymidine loop has implications for explaining why the three K^+ cations could not escape from the channel of this structure. Both channel entrances appear to be obstructed by the thymines (see Supporting Information).

The two additional simulations B9L and B10L (alternative simulations to B7L, see above) provided a loop picture essentially consistent with the outcome of the B7L simulation.

A rather interesting development has been seen during the very last simulation B11L, carried out in the presence of high salt. The T6−T8 loop adopted a similar geometry as in the other simulations, and two O2(T8)···N3(T6) and O4(T6)···N3(T8) H-bonds have been formed at ∼3 ns, after cations have been expelled from the loop. However, the other loop adopted (with a transition occurring at ∼2.4 ns) an entirely new geometry, where T20 rotates in such a way that its methyl groups points toward the channel entry and the two carbonyl groups are oriented away from the loop (see Supporting Information).

Exchange of Cations between Solute and Solvent. The simulations revealed three distinct positions in the loops, where the cation can be temporarily coordinated. (i) just above the outer quartet layer and below thymines T6 and T8, (ii) coordinated simultaneously to O2 atoms of T6 and T8, and (iii) interacting with the O2 oxygen of the outermost thymine T7 (see Supporting Information). Thus, the cations located within the loop are mostly coordinated to thymine O2 oxygens, while additional coordination sites may involve, for example, sugar oxygens. When a cation leaves the channel and enters the loop region, it first occupies the position just above the mixed GCGC quartet below thymines T6 and T8, preventing a "closed" arrangement of the adjacent GCGC quartet. Then the cation either returns to the channel or moves vertically to coordinate to T6 and T8 and eventually to T7. When leaving the loop regions, the cations utilize two distinct routes. In most cases, cations leave the loop cage vertically, initially coordinated to T6 and T8, then moving up to T7 and being released to the solvent through the apex of the loop. However, in three cases a cation left the loop region in a more horizontal direction, penetrating into a groove. We have seen a reversed event during the B7L simulation (see above) where one cation returned for ∼0.1 ns back to the four-thymidine loop from a position already outside the solute molecule. The cation was first intercepted by the O2 keto oxygen of T7 at the loop apex and then relocated to the T6/T8 coordination site. One can assume that both routes used by the cations to leave the loop region in our simulations can be reversed. On a longer time scale, these two routes can be used for exchange of the cations between the solute and solvent.

Geometry of Four-Thymidine Loops in Simulations Assuming the NaCl Starting Structure. The *NaCl* NMR structure of $d(GGGCTTTTGGGC)_2$ shows a different loop geometry compared to that of the *KCl* structure (Figure 5a). Thymines T5 and T8 are stacked on the cytosines of the adjacent "closed" mixed GCGC quartet. These bases are not involved in any H-bonding interactions. Thymine T7 is located perpendicular to the plane of T5, with its O2 oxygen pointing toward the channel entry. T6 is approximately in the plane of T7 and protrudes away from the loop. There is a contact between the O2 oxygen of T6 and a methyl group of T7. With T6 projecting away from the molecule this loop is less compact compared to the *KCl* loop arrangement. The T16-T120 NMR loop geometry is obviously identical. The two simulations initiated with the *NaCl* geometry (B5L and B6L) diverged into four considerably different loop arrangements (Figure 7). All of these still show the T6/T18 flipped away from the structure; thus, none of the four rearrangements leads to a structure similar to that of the *KCl* arrangement (see Supporting Information).

The T5-T8 loop in the B5L simulation developed in the following way: At ∼2.4 ns, a phosphate oxygen atom of T8 approached the N3 nitrogen of T5, with direct N3(T5)···O3′-(T7) and N3(T5)···O1P(T8) contacts. The O1P oxygen of T8 pointed toward the channel entry with a distance around 3 Å, interacting with the cation oscillating between a position just above the plane of the outer mixed GCGC quartet and a position in its plane (se above). We assume that this particular loop geometry is essential to lock the cation in this position. In the other loop the thymine T19 was flipped at 0.1 ns in a way that thymines T18, T19, and T20 formed a stacked trimer. Further, T19 formed a base pair with T17 via N3(T19)···O2(T17) and N3(T17)···O4(T19) H-bonds. This is the only inter-thymidine H-bonded base pair formed when we commenced the simulation from the *NaCl* NMR structure. Thymine T17 remained stacked on top of a cytosine residue from the adjacent mixed GCGC quartet. This also brought the bases of T20 and T17 close to each other. The channel entry is blocked by the methyl group of thymine T20, preventing cations from leaving the channel through this entrance. This loop geometry appears to be very compact and stabilized by base−base interactions (see Supporting Information).

During simulation B6L, thymine T8 started to rotate at 1.1 ns about its glycosidic bond. At 1.6 ns the proximal cytosine C24 is displaced from the plane of the mixed GCGC quartet and shifted around 3 Å up to the level of the T8 loop base, forming a base-on-deoxyribose stack with the sugar moiety of thymine T8. T8 is further engaged in a H-bond with the O2P phosphate oxygen of cytosine C24 (distance 2.9 Å). In the other loop of the molecule, at ∼4.0 ns, thymine T20 became unstacked from cytosine C12 of the adjacent outer mixed quartet layer, and moved toward the solvent. This effectively opened up the loop and exposed the region above the channel entry to the solvent. T17 remained stacked on top of cytosine C16 of the mixed quartet layer, while thymine T19 stacked on top of T17, with oxygens O2(T19) and O4(T17) lining up above the channel entry. This open loop configuration and the fact that the adjacent stem cavity is vacant should attract cations into the region above the channel entry, which is in fact exactly what happens during the simulation in the period 4.2−4.6 ns (see Supporting Information).

Discussion and Conclusions

Nanosecond-scale molecular dynamics simulations represent a state-of-the-art tool to study the structure and dynamics of nucleic acids. In the present paper, we utilized this technique for evaluating of G-DNA molecules containing mixed GCGC base quartets.

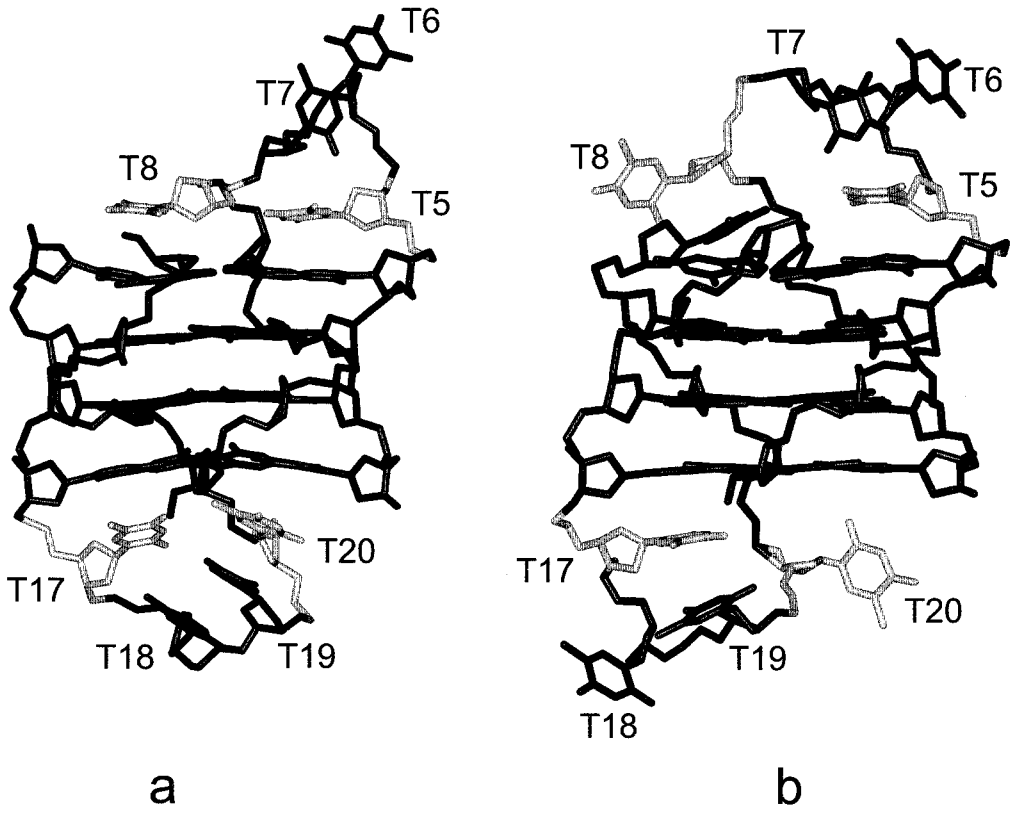

Figure 7. (a) Averaged geometry (4−5 ns) of the d(GGGCTTTTGGGG)$_2$ assuming the *NaCl* starting geometry and utilizing Na^+ cations in the simulation. (b) The same, but K^+ cations have been utilized in the simulation. Note that all four loops adopt distinct geometries.

Structure and Dynamics of Quadruplex Stems Containing Mixed GCGC Quartets. The simulations suggest that quadruplex stems containing mixed GCGC quartet layers in addition to all-guanine quartets are stabilized by monovalent cations residing in their central channel. The mixed GCGC quartets in our studies adopt two distinct conformations, termed "closed" and "sheared", in excellent agreement with high-resolution NMR studies by Patel and co-workers.[9] The "closed" quartet is stabilized by H-bonds connecting the two G·C Watson−Crick base pairs forming the quartet. The "sheared" quartet has no H-bonds between the G·C base pairs and is formed upon a close contact with a cation or its penetration into the quartet plane. The conformational switch between "sheared" and "closed" GCGC quartet geometries is usually swift without any intermediates. When a large K^+ cation is present in the cavity, the adjacent mixed quartet adopts the "sheared" arrangement while in the presence of Na^+ the mixed quartet remains in the "closed" architecture. The only exception occurs when the Na^+ closely contacts the mixed quartet (for example, due to its interaction with the loops).

The simulations suggest that mixed quadruplex stems with two GCGC and two GGGG quartets are preferably stabilized by two cations. In almost all simulations, a third cation was swiftly expelled. The cations prefer to stay in cavities between mixed and all-G quartets, being situated closer (but not in) the planes of the all-G quartets. Mixed GCGC quartets expel the cations from their planes in order to adopt the closed geometry with a fully developed network of H-bonds between the two G·C base pairs. The cations may be temporarily locked in the quartet planes by specific loop geometry, but this is not a prevalent event. Both Na^+ and K^+ cations are capable of passing through the plane of both types of quartets without causing any structural destabilization.

Interactions (Table 2) between monovalent cations and mixed GCGC quartets are quite favorable. They are nevertheless much less attractive than interactions between ions and all-guanine quartets. We suggest that the monovalent cations have the following preference for coordination: all-guanine quartet > full hydration outside the solute > mixed quartet. This view is consistent with our observation that hypothetical four-stranded assemblies consisting of four consecutive mixed GCGC quartets are unstable in the presence as well as in absence of cations. The simulations suggest that the stability of mixed stems originates in the presence of cation-stabilized guanine quartets, while mixed GCGC quartets are rather tolerated within quadruplexes.

A previous study on all-guanine quadruplex DNA revealed that its channel is fully hydrated when left vacant by cations.[8a] The mixed GCGC quartets do not exhibit an obvious tendency to be hydrated to such an extent. Close contact with water molecules would enforce a sheared mixed quartet arrangement with a disruption of several H-bonds. A water molecule apparently does not provide a sufficient stabilization energy for that to occur.

A mixed stem initially without any cations in the channel shows pronounced instabilities and fluctuations. However, we have seen that a vacant quadruplex stem spontaneously intercepted and incorporated an initially remote cation. So far, we have carried out less than 10 ns of simulations with vacant G-DNA stems.[8] During these simulations, we have seen one event of a remote cation spontaneously entering the channel and stabilizing the molecule. A spontaneous relocation of an external cation into an initially vacant channel has also been recently reported for an all-guanine parallel stranded G-DNA stem.[3i] Therefore, we estimate that a lifetime of a four-stranded G-DNA stem with a vacant channel could be on a scale of 10

ns. This observation further supports the genuine view that G-DNA stems are never left vacant by cations. Note that the experimentally observed time scale for a full equilibrium exchange of cations between the channel and surroundings is hundreds of microseconds for sodium cations.[14] It is genuine to assume that incorporation of the very first cation into an initially vacant stem must be orders of magnitude faster compared to the complete equilibrium exchange of all cations.

Interactions of the Thymidine Loops with Cations. It has been suggested that the $d(GGGCTTTTGGGC)_2$ molecule in the presence of KCl incorporates five consecutive closely spaced K^+ cations (including one in each loop) along the molecule axis, while only three internal cations are present in the structure in the presence of NaCl.[9b,c]

The simulations reveal that the arrangement with five cations inside the quadruplex is highly unstable. The majority of simulations lead to two cations inside the channel and no cation within the loop region. Some simulations, however, ended with three internal cations. Then, there are either three cations in the channel repeatedly testing a position outside the channel, or one cation is held at the stem−loop junction. Both situations are aided by specific loop geometries.

We did not find any stable coordination site for a cation within the loops. However, three distinct positions for cations to be temporarily coordinated inside the loops have been identified: above the channel entry, in the loop center, and at the apex of the loop. Coordination of ions in the loop is a process competing with a formation of inter-thymine H-bonds in the loops. We have identified two routes that appear to be utilized by the cations for an exchange between channel and solvent.

Conformational plasticity of the Loops. The simulations revealed multiple nanosecond-scale stable conformations of the loops of $d(GGGCTTTTGGGC)_2$. When starting the simulation assuming the *KCl* NMR structure at the beginning, the loops remained localized essentially within one sub-state close to the starting geometry. When the *NaCl* NMR structure was assumed at the beginning, the loops exhibited large transitions resulting in a set of mutually different loop geometries. None of them approached the *KCl* loop geometry, as it would require a complete inversion of the position of the T6/T18 thymine. It is evident, that the time scale of the simulations was not sufficient to unambiguously characterize the conformational space of the loops and to identify a global minimum of the loop region. We suggest a possibility that a real dynamics of the four-thymidine loops involves different structures populated with significant probabilities.[15]

Comparison of Theoretical and Experimental Data. The simulations are based on NMR starting structures,[9] and these structures show excellent stability during the simulations. There is full agreement between theory and experiment concerning the presence of two distinct mixed GCGC quartet geometries and a cation-induced switch between them. The suggested (based on calculated electrostatic potential terms) arrangement of five K^+ internal cations in $d(GGGCTTTTGGGC)_2$ in the presence of KCl is not stable in the simulations. The experimentally determined difference between the loop regions of the *NaCl* and *KCl* structures of $d(GGGCTTTTGGGC)_2$ has been related to extensive interactions of the loops with K^+ cations. This seems to contradict the simulations suggesting that ions do not interact with the loops and mostly only two cations are in the channel. However, closer analysis of the data reveals that there is no disagreement. The large K^+ cations induce a formation of "sheared" mixed quartets even when only two ions are present in the central ion channel. Indeed, it is entirely possible that the substantial difference in the outer quartet geometries for *NaCl* and *KCl* structures could induce substantial differences in the loop geometries even in the absence of extensive cation−loop interactions. Further, the NMR data for the *KCl* structure indicates a formation of inter-thymidine H-bonds, while the simulations suggest that inter-thymidine base pairs are disrupted by a direct coordination of ions to the loop. Thus, we propose that the differences in loop geometries in *KCl* and *NaCl* structures of $d(GGGCTTTTGGGC)_2$ can be rationalized when considering just two cations in the channel.

Limitations of the Computational Approach. During the review process, the reviewers raised the issue of statistical significance of our results and suggested that single observation of an event during MD simulation does not prove that the event is significant. Let us comment on this in more detail. In this study, we have carried out altogether 20 independent simulations. Rather than repeating one simulation many times, we designed carefully the individual simulations to cover as wide range of conditions/situations as possible. Thus, we cannot provide a simple statistical analysis of our data. Nevertheless, a vast majority of our conclusions is based on repetitive observations of identical and related events, while the amount of information gathered in this way is an order of magnitude larger than repeating few simulations many times. We further argue that simulations only rarely result in geometries that can be considered as truly nonrepresentative or misleading, and repeating of simulations is quite infrequently encountered in the literature. Our simulations show (with a clear statistical significance) that two cations inside the mixed stem represent the most likely arrangement of the simulated structures, because a vast majority of our simulations finished with just two cations inside, after eliminating excessive cations. We did not see any movement in the opposite direction, while structures with a further reduction of the number of internal cations show instabilities and tend to spontaneously incorporate cations. The analysis of mixed quartet geometries including the switch between closed and sheared arrangement is statistically significant too, being based on numerous consistent events. In numerous simulations, we did not find a single case of a cation with a stable coordination to a loop. Thus, this (negative) result is also statistically entirely significant. The same is true also for the observed mutual competition between inter-thymidine H-bond formation and cation coordination to loops. On the other hand, our loop geometry analysis is based on evaluation of individual loop simulations often resulting into quite distinct geometries. Here we obviously cannot evaluate their statistical significance. Nevertheless, our major conclusion is that the loops adopt multiple nanosecond-scale stable geometries. This qualitative conclusion is statistically significant, as it is based on multiple independent simulations each providing two independent loop trajectories.

Additional limitations of the simulations stem from the approximations inherent to the force field. We have utilized two versions of the Cornell et al. force field, differing in the sugar pucker parameters.[5a,c] These two force fields provide identical results for the stem part of the simulated structures. However, they appear to affect the behavior of the loop region to some extent. We are not in a position to decide which of the force fields is more realistic for our purpose.

(14) (a) Xu, Q.; Deng, H.; Braunlin, W. H. *Biochemistry* **1993**, *32*, 13130−13137. (b) Hud, N. V.; Schultze, P.; Sklenář, V.; Feigon, J. *J. Mol. Biol.* **1999**, *8*, 233−243.

(15) (a) Sun, S.; Bernstein, E. R. *J. Phys. Chem.* **1996**, *100*, 13348−13366. (b) Kratochvíl, M.; Šponer, J.; Hobza, P. *J. Am. Chem Soc.* **2000**, *122*, 3495−3499.

Another approximation is the pair-additivity of the force field, that is, lack of an explicit polarization term. Neglect of polarization limits the accuracy of a description of the strength and balance of cation−solute and cation−solvent interactions.[8] As noted in our preceding study, the currently used nonpolarizable force field renders the Na^+ cation slightly over-sized. This likely penalizes the in-plane position of the Na^+ cation with respect to all-G quartet planes. Thus, in reality, the Na^+ cations might spend longer time periods in the planes of the G-quartets compared with simulations showing rare ps-scale events only. No such bias is expected for K^+, as this cation unambiguously is too large to be placed in the plane of the quartets for longer time periods. However, the global distribution of the cations is captured well, since it is primarily determined by the long-range electrostatic interactions properly included in contemporary simulations.[6−8]

Concluding Remarks

MD simulations provide encouraging agreements with high-resolution experimental data for a wide range of unusual DNA structures.[8,6g,6o] The mixed quadruplex architectures examined here show many similarities as well as interesting differences when compared with the all-guanine G-DNA molecules. DNA quadruplex arrangements containing guanine quartets have been implicated in a variety of biological processes including, for example, telomere elongation by the enzyme telomerase, and thus they have attracted recently considerable pharmaceutical interest.[16] The ability of G-DNA to incorporate the mixed quartets expands the possible sequence and structural variability of G-DNA-based DNA assemblies, possibly offering variations in the four-stranded epitope that may be specifically recognized by ligands such as proteins. We anticipate that computational approaches, complementing atomic resolution experimental data, will be instrumental for designing DNA-based pharmaceuticals[3h,17] and may also provide further insight into the roles of these unusual DNAs in physiological systems.

Acknowledgment. This study was supported by the Grants LN00A016 (National Centre for Biomolecular Research), Ministry of Education, ČR, and partly by A4004002 by IGA AS ČR, and VW-Stiftung I/74657. I.B. acknowledges support by a Liebig fellowship from the *Fonds der Chemischen Industrie* (FCI, Germany). We especially thank the Supercomputer Center, Brno, where all calculations were carried out. We thank the reviewers for stimulating comments. All trajectories can be obtained from the authors upon request.

Supporting Information Available: Coordinate files for all starting structures utilized in simulations with thymidine loops (PDB); coordinate files for all final structures with thymidine loops (PDB); color stereo figures of all starting and final MD structures for simulations including the thymidine loops ; color stereo figures showing typical positions of cations inside the thymidine loops (PDF). This material is available free of charge via Internet at http://pubs.acs.org.

JA002656Y

(16) (a) Mergny, J. L.; Mailliet, P.; Lavelle, F.; Riou, J.-F.; Laoui, A.; Helene, C. *Anticancer Drug Des.* **1999**, *14*, 327−339. (b) Han, F. X.; Wheelhouse, R. T.; Hurley, L. H. *J. Am. Chem. Soc.* **1999**, *121*, 3561−3570. (c) Han, H.; Cliff, C. L; Hurley, L. H. *Biochemistry* **1999**, *38*, 6981−6986. (d) Read, M. A.; Wood, A. A.; Harrison, J. R.; Gowan, S. M.; Kelland, L. R.; Dosanjh, H. S.; Neidle, S. *J. Med. Chem.* **1999**, *42*, 4538−4546. (e) Haq, I.; Trent, J. O.; Chowdhry, B. Z.; Jenkins, T. C. *J. Am. Chem. Soc.* **1999**, *121*, 1768−1779.

(17) Smirnov, I.; Shafer, R. H. *Biochemistry* **2000**, *39*, 1462−1468.