J|A|C|S
ARTICLES
Published on Web 03/01/2006

G-Quadruplexes Can Maintain Their Structure in the Gas Phase

Manuel Rueda,[†,‡] F. Javier Luque,*[,§] and Modesto Orozco*[,†,‡,⊥]

Contribution from the Institut de Recerca Biomédica, Parc Científic de Barcelona, Josep Samitier 1-5, Barcelona 08028, Spain, and Departament de Bioquímica i Biologia Molecular, Facultat de Química, Universitat de Barcelona, Martí i Franqués 1, Barcelona 08028, Spain, Departament de Fisicoquímica, Facultat de Farmacia, Universitat de Barcelona, Avgda Diagonal 643, Barcelona 08028, Spain, and Computational Biology Program, Barcelona Supercomputing Center, Jordi Girona 31, Edifici Torre Girona, Barcelona 08028, Spain

Received September 5, 2005; E-mail: modesto@mmb.pcb.ub.es; fjluque@ub.edu

Abstract: Several very extended (0.5−1 μs) molecular dynamics (MD) simulations of parallel and antiparallel G-quadruplex DNA strongly suggest that in the presence of suitable cations the quadruplex not only remains stable in the gas phase, but also displays a structure that closely resembles that found in extended (25-ns long) trajectories in aqueous solution. In the absence of the crucial cations, the trajectories become unstable and in general the quadruplex structure is lost. To our knowledge, this is the first physiologically relevant structure of DNA for which very large MD simulations suggest that the structure in water and in the gas phase are indistinguishable.

Introduction

More than 50 years after the discovery of the double helix,[1] we know that DNA is very polymorphic and can adopt a variety of helical structures, including duplexes, triplexes, and G-quadruplexes,[2] whose stability is largely related to the solvent screening of the phosphate−phosphate repulsion.[3−5] Removal of solvent should, therefore, promote the unfolding of the DNA structures. However, several research groups[6−19,84] have shown that multistranded DNA complexes and even noncovalent drug−DNA complexes can be detected in the gas phase.

Recent extended molecular dynamic (MD) studies performed in our group support the experimental claims that the DNA structures retain a significant fraction of their structural integrity under the conditions used in electrospray ionization mass spectrometry (ESI-MS) experiment.[20,21] Thus, though vaporization of a DNA duplex induces notable structural alterations, the two strands remain bound by nucleobase−nucleobase interactions, and many structural features typical of the duplex in solution are preserved.[20] These results have been confirmed by combined MD and ESI-MS experiments by other groups,[15] showing that MD simulations can help to understand the conformational changes induced by vaporization. Recently,[21] we analyzed the structure of three drug−DNA complexes in the gas phase by means of very extended MD simulations. We found that while the helical structure of the duplex is severely distorted, the two strands remain bound and the drug is tightly anchored to them. In fact, the pattern of contacts between the drug and the host DNA is quite well preserved. Once again, MD simulations are able to provide a structural explanation to ESI-MS experiments.[6−10,12−14]

[†] Parc Científic de Barcelona.
[‡] Departament de Bioquímica i Biologia Molecular, Universitat de Barcelona.
[§] Departament de Fisicoquímica, Universitat de Barcelona.
[⊥] Barcelona Supercomputing Center.

(1) Watson, J. D.; Crick, F. H. C. *Nature* **1953**, *171*, 1303−1324.
(2) Saenger, W. *Principles of Nucleic Acid Structure*; Springer-Verlag: New York, 1984; pp 2−457.
(3) Bloomfield, V. A.; Crothers, D. M.; Tinoco, I. *Nucleic Acids: Structures, Properties, and Functions*; University Science Books: Sausalito, CA, 2000; pp 165−217.
(4) Levine, L.; Gordon, J. A.; Jencks, W. P. *Biochemistry* **1963**, *2*, 168−175.
(5) Turner, D. H. *Nucleic Acids: Structures, Properties, and Functions*; Bloomfield, V. A., Crothers, D. M., Tinoco, I., Eds.; University Science Books: Sausalito, CA, 2000; pp 308−310.
(6) Gale, D. C.; Smith, R. D. *J. Am. Soc. Mass Spectrom.* **1995**, *6*, 1154−1164.
(7) Hofstadler, S. A.; Griffey, R. H. *Chem. Rev.* **2001**, *101*, 377−390.
(8) Gabelica, V.; De Pauw, E. *Int. J. Mass Spectrom.* **2002**, *219*, 151−159.
(9) Gabelica, V.; De Pauw, E.; Rosu, F. *J. Mass Spectrom.* **1999**, *34*, 1328−1337.
(10) Wan, K. X.; Shibue, T.; Gro, M. L. *J. Am. Chem. Soc.* **2000**, *122*, 300−307.
(11) Reyzer, M. L.; Brodbelt, J. S.; Kerwin, S. M.; Kuman, D. *Nucleic Acids Res.* **2001**, *29*, 103−116.
(12) Gabelica, V.; Rosu, R.; Houssier, C.; De Pauw, E. *Rapid Commun. Mass Spectrom.* **2000**, *14*, 464−467.
(13) Rosu, F.; Valerica, G.; Houssier, C.; De Pauw, E. *Nucleic Acids Res.* **2002**, *30*, e82.
(14) Gabelica, V.; De Pauw, E. *J. Mass Spectrom.* **2001**, *36*, 397−402.
(15) Gidden, J.; Ferzoco, A.; Baker, E. S.; Bowers, M. T. *J. Am. Chem. Soc.* **2004**, *126*, 15132−15140.
(16) Gidden, J.; Baker, E. S.; Ferzoco, A.; Bowers, M. T. *Int. J. Mass. Spectrom.* **2005**, *240*, 183−193.
(17) Gabelica, V.; Rosu, F.; Witt, M.; Baykut, G.; De Pauw, E. *Rapid Commun. Mass Spectrom.* **2005**, *19*, 201−208.
(18) Rosu, F.; Gabelica, V.; Houssier, C.; Colson, P.; De Pauw, E. *Rapid Commun. Mass Spectrom.* **2002**, *16*, 1729−1736.
(19) Vairamani, M.; Gross, M. L. *J. Am. Chem. Soc.* **2003**, *125*, 42−43.
(20) Rueda, M.; Kalko, S. G.; Luque, F. J.; Orozco, M. *J. Am. Chem. Soc.* **2003**, *125*, 8007−8014.
(21) Rueda, M.; Luque, F. J.; Orozco, M. *J. Am. Chem. Soc.* **2005**, *127*, 11690−11698.

10.1021/ja055936s CCC: $33.50

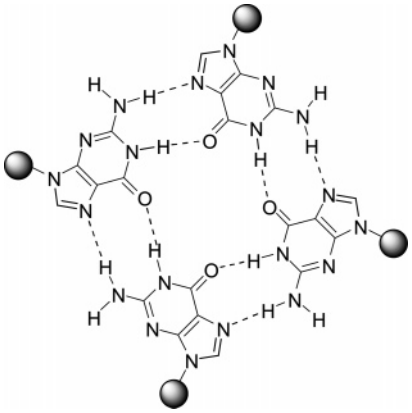

Figure 1. Schematic representation of a tetrad of G4-DNA.

In this article, we focus our attention in the G4-DNA motif, a right-handed G-quadruplex of DNA formed by four guanines linked through hydrogen bonds (see Figure 1). This structure is key in the definition of telomeres,[22,23] an important chromosomic region, which when altered can lead to processes such as cancer or aging.[24,26−28] Furthermore, G-quadruplex offers promising biotechnological properties as biosensors,[29] as liquid crystals,[30] or as template for ion channels.[31] In normal physiological conditions, the G4-DNA is a very stable structure[26,28,32,33] for both parallel and antiparallel arrangements of the four strands, as shown in a large variety of NMR and X-ray crystallography experiments.[22,23,34−47] State-of-the-art MD simulations[48−54] provided trajectory sampling conformations very close to the experimental ones and supported the large stability and rigidity of the tetrads, as well as the important role played by ions in the central channel.

The large stability of the G-quadruplex can be justified[48,54] by a variety of factors, including stacking and hydrogen bonding between bases and electrostatic interactions between the carbonyl group in guanines and cations along the central channel. These interactions, in conjunction with the solvent screening effects, have to overcome the electrostatic repulsion between the very densely packed phosphates. Clearly, one should expect a drastic unbalance of these forces upon vaporization, which would alter and probably destroy the quadruplex. However, Vairamani and Gross[55] studied the intramolecular quadruplex formed by the thrombin-binding aptamer sequence and found signals of the molecule bonded to ions such as K^+ or Sr^{2+}, which suggested that some quadruplex structure should exist in the gas phase to coordinate these ions, as also indicated by preliminary H/D-exchange experiments.[55] Similarly, Gabelica and co-workers[18] studied different sequences able to form intra- and intermolecular quadruplexes in the presence of cations and in some cases of mesoporphyrin IX (i.e., a drug that binds specifically G4-DNA in solution) and found signals in the ESI-MS spectra corresponding to quadruplexes, as well as to the tertiary complex between the G4-DNA and the drug. These results suggest that G4-DNA retains some of its solution structure in the gas phase, a result also supported by preliminary MD simulations by the same group.[56] Finally, Bowers and co-workers[16] performed cross section measurements in intermolecular quadruplexes and found that the cross section detected in gas-phase experiments is not far from that expected for the solvated structure. All these results were interpreted as additional evidence that some G-quadruplex structures can exist in the gas phase. Unfortunately, none of these groups was able to analyze at atomic detail the structure of the G4-DNA in the gas phase, and additional experiments such as those based on gas-phase H/D exchange experiments in an FTICRMS failed to shed light into the fine structural details.[56]

Here we present an extended study of the structure and dynamics of parallel and antiparallel G4-DNAs. Trajectories were obtained in the presence of different cations and represent more than 19 μs of simulation in the gas phase, including eight very long single trajectories covering 1 μs of simulation time. To our knowledge, these are the largest simulations ever published for nucleic acids. The analysis was complemented by quite extended reference MD simulations in aqueous solution, covering a total of 275 ns (11×25 ns) of simulation, which represent the more complete set of trajectories available for G4-DNA in conditions mimicking the physiological ones.

Methods

The behavior of G4-DNA was examined using two model compounds: (i) the thrombin-binding aptamer sequence (dGGTTGGTGTGGTTGG), which forms an intramolecular antiparallel G-quadruplex

(22) Wang, Y.; Patel, D. J. *Structure* **1993**, *1*, 263−282.
(23) Parkinson, G. N.; Lee, M. P. H.; Neidle, S. *Nature* **2002**, *417*, 876−880.
(24) Gowan, S. M.; Harrison, J. R.; Patterson, L.; Valenti, M.; Read, M. A.; Neidle, S.; Kelland, S. *Mol. Pharmacol.* **2002**, *61*, 1154−1162.
(25) Schnier, P. D.; Klassen, J. S.; Strittmatter, E. F.; Williams, E. R. *J. Am. Chem. Soc.* **1998**, *120*, 9605−9613.
(26) Williamson, J. R.; Raghuraman, M. K.; Cech, T. R. *Cell* **1989**, *59*, 871−880.
(27) Neidle, S.; Read, M. A. *Biopolymers* **2000**, *56*, 195−208.
(28) Sen, D.; Gilbert, W. *Nature* **1988**, *334*, 364−366.
(29) Alberti, P.; Mergny, J. L. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 1569−1573.
(30) Bonazzi, S.; Capobianco, M.; DeMorais, M. M.; Garbesi, A.; Gottarelli, G.; Mariani, P.; Ponzi, B.; Maria, G.; Spada, G. P.; Tondelli, L. *J. Am. Chem. Soc.* **1991**, *113*, 5809−5816.
(31) Forman, S. L.; Fettinger, J. C.; Pieraccini, S.; Gottareli, G.; Davis, J. T. *J. Am. Chem. Soc.* **2000**, *122*, 4060−4067.
(32) Mergny, J.-L.; DeCian, A.; Ghelab, A.; Saccà, B.; Lacroix, L. *Nucleic Acids Res.* **2005**, *33*, 81−94.
(33) Kumar, N.; Maiti, S. *Biochem. Biophys. Res. Commun.* **2004**, *319*, 759−767.
(34) Phillips, K.; Dauter, Z.; Murchie, A. I. H.; Lilley, D. M. J.; Luisi, B. *J. Mol. Biol.* **1997**, *273*, 171−182.
(35) 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.
(36) Aboul-ela, F.; Murchie, A. I. H.; Lilley, D. M. J. *Nature* **1992**, *360*, 280−282.
(37) Aboul-ela, F.; Murchie, A. I. H.; Norman, D. G.; Lilley, D. M. J. *J. Mol. Biol.* **1994**, *243*, 458−471.
(38) Bouaziz, S.; Kettani, A.; Patel, D. J. *J. Mol. Biol.* **1998**, *282*, 637−652.
(39) Horvath, M. P.; Schultz, S. C. *J. Mol. Biol.* **2001**, *310*, 367−377.
(40) Hud, N. V.; Smith, F. W.; Anet, F. A. L.; Feigon, J. *J. Mol. Biol.* **1999**, *285*, 233−243.
(41) Haider, S.; Parkinson, G. N.; Neidle, S. *J. Mol. Biol.* **2002**, *320*, 189−200.
(42) Marathias, V. M.; Wang, K. Y.; Kumar, S.; Pham, T. Q.; Swaminathan, S.; Bolton, P. H. *J. Mol. Biol.* **1996**, *260*, 378−394.
(43) Marathias, V. M.; Bolton, P. H. *Nucleic Acids Res.* **2000**, *28*, 1969−1977.
(44) Schultze, P.; Smith, F. W.; Feigon, J. *Structure* **1994**, *2*, 221−223.
(45) Smith, F. W.; Feigon, J. *Nature* **1992**, *356*, 164−168.
(46) Strahan, G. D.; Keniry, M. A.; Shafer, R. H. *Biophys. J.* **1998**, *75*, 968−981.
(47) Smirnov, I.; Shafer, R. H. *Biochemistry* **2000**, *39*, 1462−1468.
(48) Hazel, P.; Huppert, J.; Balasubramanian, S.; Neidle, S. *J. Am. Chem. Soc.* **2004**, *126*, 16405−16415.
(49) Fadrna, E.; Spackova, N.; Stefl, R.; Koca, J.; Cheatham, T. E.; Sponer, J. *Biophys. J.* **2004**, *87*, 227−242.
(50) Spackova, N.; Berger, I.; Sponer, J. *J. Am. Chem. Soc.* **1999**, *121*, 5519−5334.
(51) Spackova, N.; Berger, I.; Sponer, J. *J. Am. Chem. Soc.* **2001**, *123*, 3295−3307.
(52) Stefl, R. Cheatham, T. E.; Spackova, N.; Fadrna, E.; Berger, I.; Koca, J.; Sponer, J. *Biophys. J.* **2003**, *85*, 1787−1804.
(53) Stefl, R.; Spackova, N.; Berger, I.; Koca, J.; Sponer, J. *Biophys. J.* **2001**, *80*, 455−468.
(54) Spackova, N.; Cubero, E.; Sponer, J.; Orozco, M. *J. Am. Chem. Soc.* **2004**, *126*, 14642−14650.
(55) Vairamani, M.; Gross, M. L. *J. Am. Chem. Soc.* **2003**, *125*, 42−43.
(56) Gabelica, V.; Rosu, F.; Witt, M.; Baykut, G.; DePauw, E. *Rapid Commun. Mass Spectrosc.* **2005**, *19*, 201−208.

Table 1. List of Main Simulations Performed in This Article[a]

148D	Time (ns)
8-W-NoIon	25
8-W-Na+	25
8-W-K+	25
8-W-Li+	25
8-G[L/H]T(-4)-NoIon	1000/unstable
8-G[L/H]T(-3)-Na+	1000/1000
8-G[L/H]T(-3)-K+	1000/1000
8-G[L/H]T(-3)-Li+	1000/1000
352D	**Time (ns)**
16-W-NoIon	25
16-W-Na+	25
16-W-K+	25
16-W-Li+	25
16-GLT(-12)-NoIon	unstable
16-GLD(-10)-NoIon	unstable
16-GLD(-9)-NoIon	500
16-G[L/H]D(-8)-NoIon	500/unstable
16-G[L/H]T(-9)-Na+	500/500
16-G[L/H]D(-5)-Na+	500/500
16-G[L/H]T(-9)-K+	500/500
16-G[L/H]D(-5)-K+	500/500
16-G[L/H]T(-9)-Li+	500/500
16-G[L/H]D(-5)-Li+	500/500

[a] Simulations starting with 8 are those corresponding to the 8-nucleotide antiparallel G4-DNA, while those starting with 16 refer to the 16-nucleotides parallel G4-DNA. W stands for simulations in aqueous solution (always at 298 K and 1 atm), G stands for gas phase, L and H refer to low (298 K) or high (448 K) temperatures, and T stands for the total localized charge scheme, while the distributed model (where total charge is spread over all the phosphates) is denoted as D. The total charge in the system and the nature of the ion placed originally inside the cavity/channel are also indicated. The total length of the trajectories is noted in all cases, except in those leading to unstable trajectories, where the DNA is very severely distorted prior to corruption in the trajectory.

with two G-tetrads, and (ii) an intermolecular parallel G-quadruplex formed by four d(TG_4T) strands. The thrombin-binding aptamer is the standard model to perform structural analysis of antiparallel G4-DNA and has been solved by different authors by both NMR and X-ray crystallography.[23,42,57,58] Here we used as starting geometry in solution the NMR structure solved by Feigon's group[57] (PDB entry 148D). The intermolecular parallel G-quadruplex d$(TG_4T)_4$ has been studied by different authors using X-ray crystallography.[34,35,59] For our purposes, we examined the high-resolution structure[35] determined by Luisi and co-workers (PDB entry 352D). The original structures were manipulated to remove nontetrad fragments and to build up ideal G-quadruplexes. Following the advice of one reviewer, we performed additional test calculations to check the importance of loops on the integrity of G-quadruplexes in the gas phase and the role of size in determining antiparallel G-quadruplex structure. These test calculations included the entire aptamer 148D, including loops and a longer antiparallel G-quadruplex (d$(GGGGTTTTGGGG)_2$) (PDB entry 156D)[44] with and without loops.

Control MD Simulations in Water. The cation coordination positions along the central channel known experimentally were kept empty or occupied by Na^+, K^+, or Li^+, yielding four starting systems for each G4-DNA (see Table 1). In general, the loops were removed so the final structures contain only the central guanines of the original structure. Therefore, 8-mer refers to the structure built from the 15-nucleotide (148D), and 16-mer refers to the structure built from the 4*6-nucleotide (352D). Then the systems were immersed in truncated octahedral boxes of water containing around 2200 (148D) and 2800 (352D) water molecules. Systems were neutralized by adding a number of Na^+ cations equal to the number of charged phosphates and a number of Cl^- anions equal to the number of cations (Na^+, K^+, or Li^+) in the central G-channel. Such a process was performed using Poisson−Boltzmann potentials as implemented in the iterative cMIP titration program.[60] As in previous simulations,[20,21] these systems were then optimized and equilibrated using our standard multistep protocol,[61,62] which lasts for more than 100 ps of restricted and unrestricted MD simulation, plus additional 200 ps of unrestrained MD. The systems were then subject to 25 ns of unrestrained MD simulation in the isothermic−isobaric ensemble (T = 298 K, P = 1 atm). Periodic boundary conditions and the PME method[63] were used to account for long-range interactions. SHAKE[64] was used to maintain all the chemical bonds at their equilibrium distances, which allowed us to use a 2-fs time scale for integration of Newton equations. Global center-of-mass movements were removed every 0.1 ns to avoid artifactual cooling of the system. PARM-98[65,66] and TIP3P[67] force fields were used to represent molecular interactions. A few words of caution seems necessary regarding the use of standard PARM-98 parameters for cations, since the neglect of polarization interaction might introduce some bias in the results.

Test calculations on the full 148D and 156D structures with loops (and 156D structure without loop) were prepared and equilibrated as described above, with trajectories extending to 25 ns. Only Na^+ was considered as central cation in these test calculations.

MD Simulations in the Gas Phase. A crucial decision in the simulation of gas-phase DNAs is the determination of the charge state of the molecule. As in our previous works,[20,21] we tried to follow as close as possible the charge state found in ESI-MS experiments. However, a delicate question is how to distribute such a net charge between the different nucleotides in the G-quadruplex. CID studies on DNA duplexes indicate that the overall charge can be distributed along all titrable groups,[12,14,25] while ion mobility experiments in single-stranded oligonucleotides suggest that a localized charge distribution is more realistic.[84]Thus, we have considered here both charge models. In the distributed model, the net charge is spread equally on all the phosphates by appropriate scaling of their charges. In contrast, in the localized model, charges are assigned to those phosphates that minimize the Coulombic repulsion in the DNA structure. These two charge models, therefore, represent two limit neutralization scenarios. For our purposes here, a wide set of simulations that exploit different charge neutralization protocols has been examined to obtain a complete picture of the behavior of the G-quadruplex in the gas phase.

For the intramolecular antiparallel G-quadruplexes, experiments[55] detected signals in the spectrum corresponding to structures with charges from −3 to −8 (in units of electron), the largest signals being found for charge states of −6 and −7. Considering the presence of the loops, a total charge state of −3 (including 1 cation) appears to be a reasonable choice for the isolated G-quadruplex. For the intermolecular parallel G-quadruplexes, different charge states have been found experimentally for d$(TG_4T)_4$, the predominant one being −5.[16,18] We performed simulations with a total charge of −5 (including three cations) distributed equally on the phosphates and also a total charge of −9

(57) Schultze, P.; Macaya, R. F.; Feigon, J. *J. Mol. Biol.* **1994**, *235*, 1532−1547.
(58) Lin, C. H.; Patei, D. J. *Chem. Biol.* **1997**, *4*, 817−832.
(59) Caceres, C.; Wright, G.; Gouyette, C.; Parkinson, G.; Subirana, J. A. *Nucleic Acids Res.* **2004**, *32*, 1097−1102.
(60) Gelpí, J. L.; Kalko, S. G.; Barril, X.; Cirera, J.; de la Cruz, X.; Luque, F. J.; Orozco, M. *Proteins: Struct., Funct., Genet.* **2001**, *45*, 428−437.
(61) Shields, G. C.; Laughton, C. A.; Orozco, M. *J. Am. Chem. Soc.* **1997**, *119*, 7563−7469.
(62) (a) Soliva, R.; Laughton, C. A.; Luque, F. J.; Orozco, M. *J. Am. Chem. Soc.* **1998**, *120*, 11226−11233. (b) Shields, G. C.; Laughton, C. A.; Orozco, M. *J. Am. Chem. Soc.* **1998**, *120*, 5895−5904.
(63) Darden, T A.; York, D.; Pedersen, L. *J. Chem. Phys.* **1993**, *98*, 10089−10092.
(64) Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. C. *J. Comput. Phys.* **1977**, *23*, 327−341.
(65) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. *J. Am. Chem. Soc.* **1995**, *117*, 11946−11975.
(66) Cheatham, T. E.; Cieplak, P.; Kollman, P. A. *J. Biomol. Struct. Dyn.* **1999**, *16*, 845−862.
(67) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926−935.

Table 2. Geometrical Parameters of Antiparallel G4-DNA (8-mer) in Solution (W) and Gas Phase (G) for Different Trajectories Differing in the Presence or Absence of Ions in the Central Cavity[a]

trajectory (148D)	rmsd (exptl)	rmsd (ave)	collision cross section	SAS	*D* (ion COM)
8-W-NoIon	1.1 ± 0.1	0.5 ± 0.1	416 ± 12	1780 ± 15	0.4 ± 0.2
8-W-Na+	1.1 ± 0.1	0.6 ± 0.1	419 ± 13/*422*	1778 ± 15	0.4 ± 0.2
8-W-K+	1.1 ± 0.1	0.5 ± 0.1	420 ± 13	1800 ± 20	0.3 ± 0.2
8-W-Li+	1.2 ± 0.1	0.6 ± 0.1	417 ± 12	1794 ± 19	1.5 ± 0.4
8-GLT(-4)-NoIon	4.4 ± 0.3	1.5 ± 0.6	419 ± 13	1795 ± 30	–
8-GLT(-3)-Na+	2.0 ± 0.4	1.0 ± 0.4	403 ± 13/*411*	1722 ± 19	0.5 ± 0.3
8-GHT(-3)-Na+	5.8 ± 0.2	1.5 ± 0.6	334 ± 7/*352*	1496 ± 35	2.2 ± 0.3
8-GLT(-3)-K+	1.3 ± 0.2	1.0 ± 0.2	415 ± 12	1753 ± 25	0.4 ± 0.2
8-GHT(-3)-K+	6.3 ± 1.2	2.4 ± 1.3	329 ± 6	1475 ± 34	5.0 ± 0.7
8-GLT(-3)-Li+	5.3 ± 0.5	0.6 ± 0.7	346 ± 6	1457 ± 20	5.4 ± 0.2
8-GHT(-3)-Li+	5.9 ± 0.5	2.2 ± 1.8	344 ± 13	1533 ± 46	2.0 ± 0.4

[a] L and H refer to low (298 K) or high (448 K) temperatures, and T denotes total charge. The rmsd with respect to the experimental (exptl) NMR structure and the average (ave) conformation obtained in the trajectory are in Å. SAS and CCS are in Å^2 (numbers in italics correspond to values obtained with the EHSS method); as reference the NMR structure shows cross section of 418 Å^2 and SAS of 1810 Å^2. The distance of the internal ion to the center of mass of the G4-DNA is in Å; as reference in the NMR structure a K+ is located at 1.2 Å from the center of mass of the G4-DNA. For nomenclature, see also Table 1.

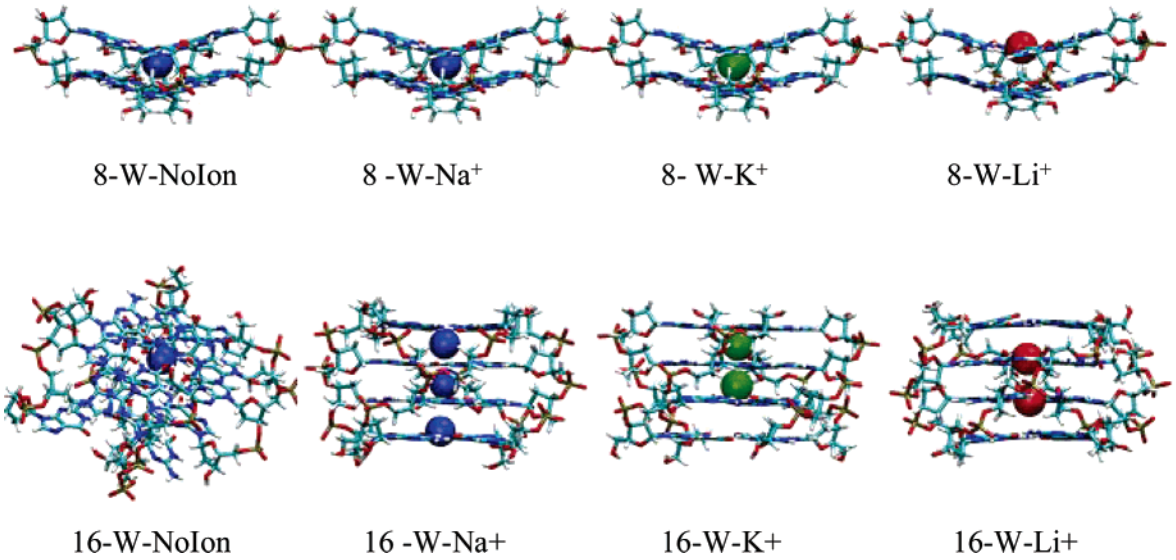

Figure 2. Average structures obtained at the final portion of the water (W) MD trajectories of the antiparallel G4-DNA from PDB entry 148D (8-mer) and the parallel G4-DNA from PDB entry 352D (16-mer). The central ions are represented by a colored sphere. Note that some ions can be found inside the structures although these positions were empty at the beginning of the simulation (see NoIon). Loops were not considered in these simulations. For nomenclature, see Table 1.

(including three cations) placed following the localized charge model,[20,21] which helped us to determine the resistance of the structure of highly charged states. Furthermore, we performed simulations to analyze the resistance of the structure to high charge states in the absence of ions inside the tetrads. We considered total charges of −8, −9, and −10 placed using the distributed model and also a total charge of −12 placed using the localized protocol for parallel quadruplex and a total charge of −4 placed using the localized model for the 8-mer. Simulations were carried out at both room (298 K) and high (448 K) temperatures (see Table 1). The first simulations will serve to study ideal vaporization processes, while the second are closer to the conditions used in ESI experiments. PARM-98 force field[65,66] was used in all cases with standard charges (on the basis of our previous studies,[20,21] we did not perform polarity reduction).

For test calculations on the full 148D structure (with loops), a total charge of −6 spread over all the phosphates (distributed charge model) was used.[19] Test simulations for the full 156D structure (with loops) were performed also with a total charge of −6[16,18] distributed equally along all the phosphates, while for the no-loops 156D G-quadruplex two simulations were done: one with a charge of −5 (distributed charge model) and another with charge −9 (localized charge model).

The structures corresponding to the first nanosecond of the aqueous MD simulations were used as starting conformations for the simulations in vacuo. After we removed waters and external counterions, we modified the structures to fit the conditions of the simulation in the gas phase (see above) and equilibrated at 298 K for 100 ps starting from the velocities of the MD simulation in water. Calculations in related systems[21] showed that no dramatic differences are expected when a softer dehydration protocol is adopted.

After equilibration, MD simulations were performed for 0.5−1 μs at constant temperature (298 K). The structures obtained after 1 ns were used as starting points for simulations at 448 K. Heating from 298 to 448 K was performed in 100 ps, and the new MD simulations were run for 0.5−1 μs. Translational and rotational displacements (small in any case) were removed periodically every 1 ns. As for the simulations in solution, SHAKE[64] was used to constrain all the bonds, but to avoid instabilities in the trajectories the integration time step was 1 fs. All trajectories were collected using the AMBER8.0 suite of programs.[68] Test calculations were also performed for 0.5−1 μs in the case of 148D (with loops) and 156D (with and without loops) but only at low temperature. MD simulations presented here cover more than 19 μs of MD simulations of G4-DNA in the gas phase and 275 ns in aqueous solution (see Table 1 and tables in Supporting Information).

(68) Case, D. A.; et al. *AMBER 8*; University of California: San Francisco, CA, 2004.

Table 3. Parameters Describing the Flexibility of Antiparallel G4-DNA (8-mer) in Water and in the Gas Phase[a]

trajectory (148D)	log V	dimensionality	entropy	force constants
8-W-NoIon	21.3	19	0.8	34
				82
				116
8-W-Na+	23.6	21	0.8	34
				81
				96
8-W-K+	22.7	20	0.8	30
				77
				80
8-W-Li+	24.4	20	0.8	24
				73
				85
8-GLT(-4)-NoIon	38.6	30	1.1	3
				6
				17
8-GLT(-3)-Na+	30.9	28	0.9	22
				36
				39
8-GHT(-3)-Na+	35.1	32	1.2	25
				54
				89
8-GLT(-3)-K+	28.2	24	0.8	21
				38
				51
8-GHT(-3)-K+	41.8	38	1.4	3
				28
				39
8-GLT(-3)-Li+	19.2	16	0.7	27
				60
				91
8-GHT(-3)-Li+	42.1	35	1.3	5
				18
				28

[a] L and H refer to low (298 K) or high (448 K) temperatures, and T denotes total charge. The configurational volume (in $Å^{3n-6}$) is shown in logarithmic (natural) form. Entropies (in kcal/mol K; see Methods section) are determined using Schlitter's method (see also Andricioaei−Karplus estimates in Supporting Information, Table S5). The force constants associated with the first three essential modes are shown in cal/mol $Å^2$. See also Table 1.

Analysis of Trajectories. Unless otherwise stated, the analysis was performed for 100 structures collected in the last 5 ns in water simulations and in the last 25 ns of the gas-phase simulations. Following our previous studies,[20,21] intramolecular contact maps were computed assuming that two guanines are in contact when the distance between them is smaller than 3.5 Å. Furthermore, the interaction was considered to be hydrogen-bonded if there is at least one interaction such that the heteroatom−heteroatom distance is less than 3.5 Å and the donor-hydrogen/donor−acceptor angle deviates less than 30° from linearity; otherwise the interaction is labeled as "stacking". The position of the central ion was followed by the distance to the center-of-mass of the nearest tetrad. Solvent accessible surfaces (SAS) were determined using NACCESS[69] and considering only DNA atoms.

Principal component analysis (PCA[70−73]) was performed to examine the essential dynamics of DNA. Accordingly, Cartesian covariance matrices were diagonalized to obtain a set of eigenvectors (describing the essential deformation modes) and a set of eigenvalues, which indicates the amount of variance explained by each mode, and can be transformed into harmonic force constants[72−75] (see eq 1 in Supporting Information), from which we can estimate the size (volume) and dimensionality (number of deformation modes with a significant contribution (more than 1 $Å^2$) to the variance) of conformational space accessible for a molecule (see Supporting Information and ref 72). The methods of Schlitter and Andricioaei-Karplus[76,77] were used to determine the intramolecular entropy associated with the different structures (Harris's procedure was used to derive entropies at infinite time[78]). Due to the agreement between the Schlitter and Andricioaei−Karplus (A−K) methods, only the former is presented here, and the A−K estimates are presented as Supporting Information. Finally, the similarity in the deformation movements was evaluated by computing similarity indexes κ and δ[20,71,72] (see eqs 2−5 in Supporting Information). The κ index measures similarity between two sets of important eigenvectors (those explaining most of the variance in the trajectory), whereas the similarity δ index is obtained by weighting differently the eigenvectors according to the amount of variance explained (see Supporting Information). Due to the qualitative similarity between κ and δ measures, the Results and Discussion section displays only the former, and δ measures are shown in the Supporting Information.

MD simulations were used to predict the cross sections expected for the different complexes. The cross section provides a low-resolution picture of the general structural shape of the molecule.[15,79−83] Calculations were performed using the SIGMA program developed by Bowers's group[80,81] using a Lennard-Jones collision model on 100 snapshots collected during the last part of the trajectories. Average cross sections for each snapshot were obtained using a Monte Carlo procedure[15,80,81] until a convergence better than 1% was achieved. For selected cases, the quality of SIGMA estimates was tested by comparison with the much more CPU-demanding methodology developed by Jarrold and co-workers[82,83] as implemented in the MOBCAL software ("Exact hard sphere scattering" method; EHSS). As described in the Results and Discussion section, for our purposes both methods agree, giving confidence on the less expensive SIGMA calculations.

Results and Discussion

MD Simulations in Water. The structure of G4-DNA in aqueous solution has been already explored by other authors at both experimental and theoretical levels.[22,23,34−54] Particularly, Sponer and co-workers have published several nanosecond-long MD simulations analyzing different structural aspects of G4-DNA.[48−54] We will then summarize here only the most relevant features of our control solution simulations (the reader is addressed to previous studies[48−54] for a more extended description of G4-DNA in solution).

Antiparallel G4-DNA. Simulations of the thrombin aptamer G-quadruplex in water are extremely stable in all the cases, with root-mean-square deviations (rmsd; Table 2) ∼0.5 Å from the MD-averaged structure and ∼1 Å from the experimental NMR structure taken as starting conformation (see Methods section). No backbone transitions, base opening, or puckering changes

(69) Hubbard, S. J.; Thornton, J. M. *NACCESS Computer Program*; University College London: London, UK, 1993.
(70) Sherer, E.; Harris, S. A.; Soliva, R.; Orozco, M.; Laughton, C. A. *J. Am. Chem. Soc.* **1999**, *121*, 5981−5991.
(71) Orozco, M.; Perez; A.; Noy, A.; Luque, F. J. *Chem. Soc. Rev.* **2003**, *32*, 350−364.
(72) Pérez, A.; Blas, J. R.; López-Bes, J. M.; de la Cruz, X.; Orozco, M. *J. Chem. Theor. Comput.* **2005**, *1*, 790−800.
(73) Hess, B. *Phys. Rev. E* **2000**, *62*, 8438−8448.
(74) Noy, A.; Pérez, A.; Lankas; F.; Luque, F. J.; Orozco. M. *J. Mol. Biol.* **2004**, *343*, 627−638.
(75) Pérez, A.; Noy, A.; Lankas, F.; Luque, F. J.; Orozco, M. *Nucleic Acids Res.* **2004**, *32*, 6144−6151.
(76) Schlitter, J. *Chem. Phys. Lett.* **1993**, *215*, 617−621.
(77) Andricioaei, I.; Karplus, M. *J. Chem. Phys.* **2001**, *115*, 6289−6292.
(78) Harris, S.; Gavathiotis, E.; Searle, M. S.; Orozco, M.; Laughton, C. A. *J. Am. Chem. Soc.* **2001**, *123*, 12658−12663.
(79) Von Helden, G.; Hsu, M.; Gotts, N.; Bowers, M. T. *J. Phys. Chem.* **1993**, *97*, 8182−8192.
(80) Wyttenbach, T.; von Helden, G.; Batka, J. J.; Carlat, D.; Bowers, M. T. *J. Am. Soc. Mass Spectrom.* **1997**, *8*, 275−282.
(81) Wyttenbach, T.; Witt, M.; Bowers, M. T. J. *Am. Chem. Soc.* **2000**, *122*, 3458−3464.
(82) Shvartsburg, A. A.; Jarrold, F. *Chem. Phys. Lett.* **1996**, *261*, 86−91.
(83) Mesleh, M. F.; Hunter, J. M.; Shvartsburg, A. A.; Schwartz, G. C.; Jarrold, M. F. *J. Phys. Chem.* **1996**, *100*, 16082−16086.
(84) Hoaglund, C. S.; Liu, Y.; Ellington, A. D.; Pagel, M.; Clemmer, D. E. *J. Am. Chem. Soc.* **1997**, *119*, 9051−9052.

Table 4. Geometrical Parameters of Parallel G4-DNA (16-mer) in Four Trajectories Obtained by Placing None or Three Ion (Indicated in the Label) in the Central Channel[a]

trajectory (352D)	rmsd (exptl)	rmsd (ave)	collision cross section	SAS	*D* (ion COM)
16-W-NoIon	2.9 ± 1.7	2.8 ± 1.6	640 ± 10	3031 ± 39	1.4 ± 0.4
					>10
					>10
16-W-Na+	1.1 ± 0.2	0.9 ± 0.1	628 ± 12/*672*	3073 ± 35	1.4 ± 1.2
					0.4 ± 0.2
					0.5 ± 0.3
16-W-K+	1.2 ± 0.2	0.9 ± 0.1	632 ± 14	3104 ± 43	>10
					0.3 ± 0.2
					0.4 ± 0.2
16-W-Li+	1.2 ± 0.1	1.0 ± 0.1	619 ± 13	3041 ± 38	1.4 ± 1.2
					2.0 ± 0.3
					>10
16-GLD(-9)-NoIon	14.0 ± 0.7	3.0 ± 0.9	817 ± 34	3614 ± 49	–
16-GLD(-8)-NoIon	3.9 ± 0.5	1.1 ± 0.6	631 ± 9	2859 ± 24	–
16-GLT(-9)-Na+	1.8 ± 0.1	0.6 ± 0.2	604 ± 8/*652*	2730 ± 16	0.3 ± 0.1
					0.3 ± 0.1
					0.6 ± 0.2
16-GHT(-9)-Na+	2.7 ± 0.1	0.6 ± 0.1	580 ± 9/*658*	2715 ± 32	0.4 ± 0.2
					0.3 ± 0.2
					0.8 ± 0.3
16-GLD(-5)-Na+	1.6 ± 0.1	0.9 ± 0.4	569 ± 15/*593*	2585 ± 76	0.7 ± 0.3
					0.4 ± 0.2
					1.1 ± 0.2
16-GHD(-5)-Na+	4.4 ± 1.4	3.0 ± 1.7	565 ± 10/*605*	2609 ± 33	1.7 ± 0.3
					>10
					6.1 ± 0.2
16-GLT(-9)-K+	1.9 ± 0.1	0.4 ± 0.1	606 ± 10	2741 ± 18	0.4 ± 0.2
					0.3 ± 0.1
					0.6 ± 0.2
16-GHT(-9)-K+	1.9 ± 0.1	0.6 ± 0.1	588 ± 9	2773 ± 34	0.4 ± 0.2
					0.3 ± 0.1
					0.6 ± 0.2
16-GLD(-5)-K+	1.6 ± 0.1	0.8 ± 0.3	557 ± 13	2430 ± 28	0.7 ± 0.2
					0.3 ± 0.2
					1.2 ± 0.2
16-GHD(-5)-K+	1.6 ± 0.1	0.9 ± 0.2	555 ± 16	2587 ± 88	0.7 ± 0.3
					0.4 ± 0.2
					1.1 ± 0.3
16-GLT(-9)-Li+	2.8 ± 0.1	0.7 ± 0.1	598 ± 9	2673 ± 19	0.8 ± 0.3
					0.9 ± 0.2
					1.8 ± 0.3
16-GHT(-9)-Li+	6.1 ± 1.8	7.6 ± 1.9	741 ± 30	3498 ± 68	5.1 ± 0.4
					9.7 ± 0.5
					8.6 ± 1.0
16-GLD(-5)-Li+	1.7 ± 0.1	0.6 ± 0.1	570 ± 14	2587 ± 55	1.3 ± 0.3
					0.7 ± 0.2
					0.8 ± 0.3
16-GHD(-5)-Li+	5.2 ± 0.4	1.6 ± 0.8	577 ± 11	2712 ± 53	0.8 ± 0.2
					0.8 ± 0.2
					4.1 ± 0.2

[a] The rmsd with respect to the experimental (exptl) X-ray structure and the average (ave) conformation obtained in the trajectory are in Å. SAS and CCS are in Å^2 (numbers in italics correspond to values obtained with the EHSS method); as reference the X-ray structure shows a cross section of 592 Å^2 and SAS of 2900 Å^2. The distance of the internal ion to the center of mass of the G4-DNA is in Å; as reference in the X-ray structure a K+ is located at 1.0, 0.2, and 0.7 Å from the center of mass of the surrounding nucleotides of the G4-DNA. L and H refer to low (298 K) or high (448 K) temperatures, T stands for the total localized charge scheme, while the distributed model is denoted as D. For nomenclature, see also Table 1.

are detected in any of the four MD simulations. SAS and collision cross sections (CCS) are almost identical to those found in the NMR structure (see Table 2), as is also found for any other general geometrical parameter of the G-quadruplex. The stability of trajectories collected for structures with cations in the central cavity is not surprising,[48−54] but that of the trajectories started without cations is unexpected. Analysis of the simulation revealed, however, that during the first nanosecond a Na^+ ion was captured, leading to a trajectory that resembled the 8-W-Na^+ system (Figure 2; structures and rmsd plots available at Molecular Modelling & Bioinformatics Unit Web site, http://mmb.pcb.ub.es/~manu/G4DNA/). The largest cations (Na^+, K^+) fluctuate around the center-of-mass (COM) of the G4-DNA, thus mimicking the positions found in the NMR structure (see Table 2 and Figure 2). However, Li^+ is displaced 1.5 Å apart from the COM and becomes trapped in the center of one of the tetrads (see Table 2 and Figure 2).

At least for simulations with Na^+ inside the cavity, loops do not introduce any structural change in the central G-quadruplex and the global rmsd from NMR structure increases to 2.5 Å (2.0 with respect to the MD-averaged conformation), due to some movements in the loops. However, the central G-quadruplex is only 1 Å from the NMR structure (0.6 Å from the MD average), indicating that the G-quadruplex structure is

Table 5. Parameters Describing the Flexibility of Parallel G4-DNA (16-mer) in Water and in the Gas Phase[a]

trajectory (352D)	log V	dimensionality	entropy	force constants
16-W-NoIon	104	83	3.2	0.2
				2
				4
16-W-Na+	63	53	1.9	6
				20
				24
16-W-K+	64	53	1.9	6
				13
				20
16-W-Li+	66	53	2.0	6
				8
				20
16-GLD(-9)-NoIon	76	58	2.0	1
				2
				3
16-GLD(-8)-NoIon	47	38	1.7	13
				20
				24
16-GLT(-9)-Na+	29	21	1.4	66
				69
				73
16-GHT(-9)-Na+	49	40	2.3	52
				55
				60
16-GLD(-5)-Na+	40	29	1.6	10
				14
				25
16-GHD(-5)-Na+	76	66	2.7	5
				16
				27
16-GLT(-9)-K+	28	20	1.4	62
				95
				104
16-GHT(-9)-K+	48	38	2.2	50
				78
				83
16-GLD(-5)-K+	28	21	1.4	68
				89
				117
16-GHD(-5)-K+	67	53	2.7	10
				17
				28
16-GLT(-9)-Li+	36	26	1.5	11
				12
				15
16-GHT(-9)-Li+	98	79	3.2	1
				2
				5
16-GLD(-5)-Li+	38	29	1.6	14
				34
				39
16-GHD(-5)-Li+	76	64	2.7	8
				12
				19

[a] L and H refer to low (298 K) or high (448 K) temperatures, T stands for the total charge, while the distributed model is denoted as D. The configurational volume (in $Å^{3n-6}$) is shown in logarithmic (natural) form. Entropies (in kcal/mol K; see Methods section) are determined using Schlitter's method (see also Andricioaei−Karplus estimates in Supporting Information, Table S6). The force constants associated with the first three essential modes are shown in cal/mol $Å^2$. For nomenclature, see also Table 1.

the same in simulations with or without loops (see Table S1 in Supporting Information).

Even in the absence of connecting loops, the antiparallel G4-DNA is very rigid, as noted in the force constants for the first three modes, whose values (24−116 cal/mol $Å^2$; Table 3) are larger than those obtained for duplexes with the same number of bases (10−20 cal/mol $Å^2$),[72] and in the small oscillations in the rmsd (see Table 2). The pattern of deformability is quite

Table 6. Similarity Indexes (κ; See Text) between the Different Trajectories of the Antiparallel G-Quadruplexes[a]

8-W-NoIon	8-W-Na+	8-W-K+	8-W-Li+	8-GLT(-4)-NoIon	8-GLT(-3)-Na+	8-GHT(-3)-Na+	8-GLT(-3)-K+	8-GHT(-3)-K+	8-GLT(-3)-Li+	8-GHT(-3)-Li+	
1.0	1.0	1.0	0.9	0.4	0.6	0.3	0.8	0.3	0.4	0.3	8-W-NoIon
	1.0	1.0	1.0	0.4	0.6	0.3	0.8	0.3	0.4	0.3	8-W-Na+
		1.0	1.0	0.4	0.5	0.3	0.8	0.3	0.4	0.3	8-W-K+
			1.0	0.4	0.5	0.3	0.7	0.3	0.4	0.3	8-W-Li+
				1.0	0.5	0.3	0.3	0.4	0.4	0.3	8-GLT(-4)-NoIon
					1.0	0.3	0.7	0.3	0.3	0.3	8-GLT(-3)-Na+
						1.0	0.3	0.4	0.3	0.4	8-GHT(-3)-Na+
							1.0	0.3	0.3	0.3	8-GLT(-3)-K+
								1.0	0.4	0.4	8-GHT(-3)-K+
									1.0	0.3	8-GLT(-3)-Li+
										1.0	8-GHT(-3)-Li+

[a] G stands for gas phase, L and H refer to low (298 K) or high (448 K) temperatures, and T stands for the total charge. See also simulation codes in Table 1.

simple, as noted in the small dimensionality of the configurational space (19−21 modes) compared to that expected for a duplex with the same number of bases (around 28).[72] Test calculations for the same aptamer with loops and Na^+ inside the channel (see Methods section) show that the connecting loops are very mobile, increasing then the global flexibility of the structure (see Table S2 in Supporting Information). However, the flexibility of the central G-quadruplex is not much altered by the presence of loop (just a small increase in the stiffness of the quadruplex can be attributed to the presence of connecting loops; see Table S2 in Supporting Information and http://mmb.pcb.ub.es/∼manu/G4DNA/).

Concern might exist as to whether the thrombin aptamer is too short to be a good model of antiparallel G-quadruplex. Thus, as described in the Methods, we performed test calculations with a larger G-quadruplex (156D) in the presence or absence of connecting loops. The rmsd from experimental data always reaches very small values for the central quadruplex (1.0−1.4 Å). Collision cross sections, solvent accessible surfaces, and other descriptors fit very well those computed for the NMR structure, confirming the suitability of the simulation protocol to describe these structures (see Tables S3 and S4 in Supporting Information and http://mmb.pcb.ub.es/∼manu/G4DNA/). All flexibility indexes computed indicate that once the size effect is corrected the flexibility of the central G-quadruplex is the same for 148D and 156D. In summary, neither the small size of the antiparallel G4-DNA model nor the neglect of loops is expected to introduce detectable bias in our results.

Parallel G4-DNA. The three model systems built up with monovalent ions in the central channel are stable along the 25-ns trajectory, as noted in the small rmsd values (see Table 4), which indicate the lack of any remarkable conformational transition (Figure 2; structures and rmsd plots available at http://mmb.pcb.ub.es/∼manu/G4DNA/). The only relevant change during the trajectory is a small enlargement in the SAS, which reflects a certain expansion of the G4-DNA probably due to the release of lattice packing forces in the X-ray structure. Two of the channel ions remain at positions identical to those found in the crystal (see Table 4 and Figure 2), but the third ion is quite mobile. Thus, in 16-W-Na^+, one of the central cations is displaced from the inner tetrad position to a coordination point between carbonyls in one of the terminal tedrads (see Figure 2). For 16-W-K^+ and 16-W-Li^+, the mobile cation is released

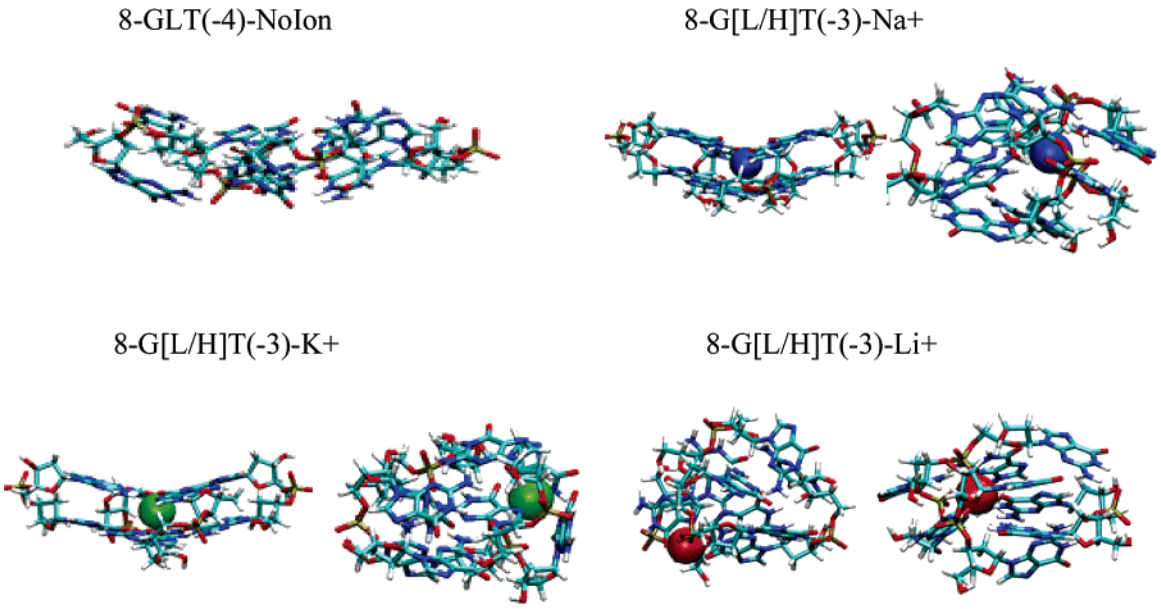

Figure 3. Average structures obtained at the final portion of the MD trajectories in the gas phase of the antiparallel G4-DNA from PDB entry 148D (8-mer). In each pair of structures, G stands for gas phase, L and H refer to low (298 K) or high (448 K) temperatures, and T stands for the total charge. Loops were not considered in these simulations. For simulation codes, see Table 1.

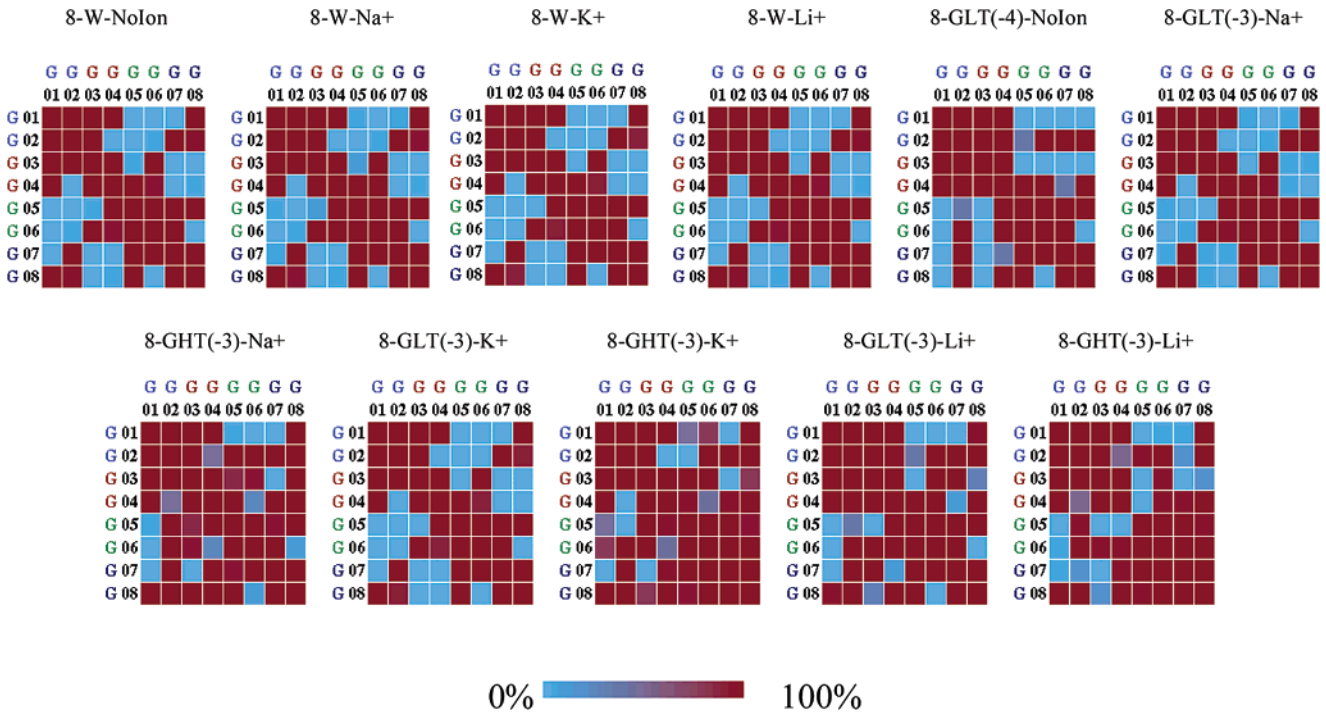

Figure 4. Contact maps for the final portion of the MD trajectories in the gas phase of the antiparallel G4-DNA from PDB entry 148D (8-mer). W stand for water, G stands for gas phase, L and H refer to low (298 K) or high (448 K) temperatures, and T stands for the total charge. See also simulation codes in Table 1.

to the ionic solution, though the structure of the G-quadruplex remains unaltered (see Figure 2 and Table 4).

As pointed out by others,[48−54] the parallel G4-DNA is unstable in the absence of cations in the central channel. However, simulation 16-W-NoIon captures very rapidly a Na^+ cation from the environment, which explains the small rmsd value with respect to the crystal structure for near 15 ns (data not shown). However, a single ion does not suffice to maintain the integrity of the structure, which collapses before a second Na^+ enters the channel (Figure 2), leading at the end of the trajectory to structures with rmsd values above 6 Å. Though the general G4-DNA conformation is severely altered in these structures, both SAS and CCS are close to those expected for a canonical G4-DNA.

In the absence of ions, the G-quadruplex unfolds and jumps in the nanosecond time scale between different states (see Table 5: 16-W-NoIon simulation). Even when the central channel is occupied with cations, the parallel G4-DNA (352D) has a more complex deformability pattern than the short antiparallel structure (148D) described above. Such a difference cannot be attributed to the different size of both structures, since for example the dimensionality (which scales linearly with the

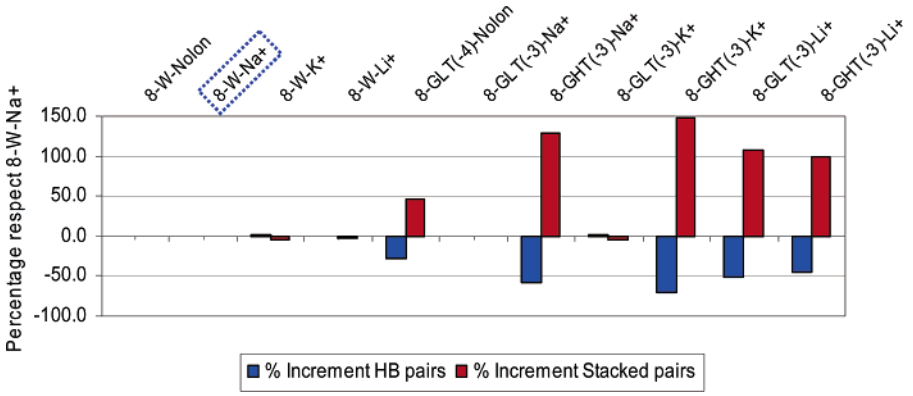

Figure 5. Percentage of gained/lost interactions (hydrogen bond and stacking) in the different simulations for antiparallel G4-DNA from PDB entry 148D (8-mer) with respect to the control simulation in water in the presence of Na^+ in the central cavity. G stands for gas phase, L and H refer to low (298 K) or high (448 K) temperatures, and T stands for the total charge. See also simulation codes in Table 1.

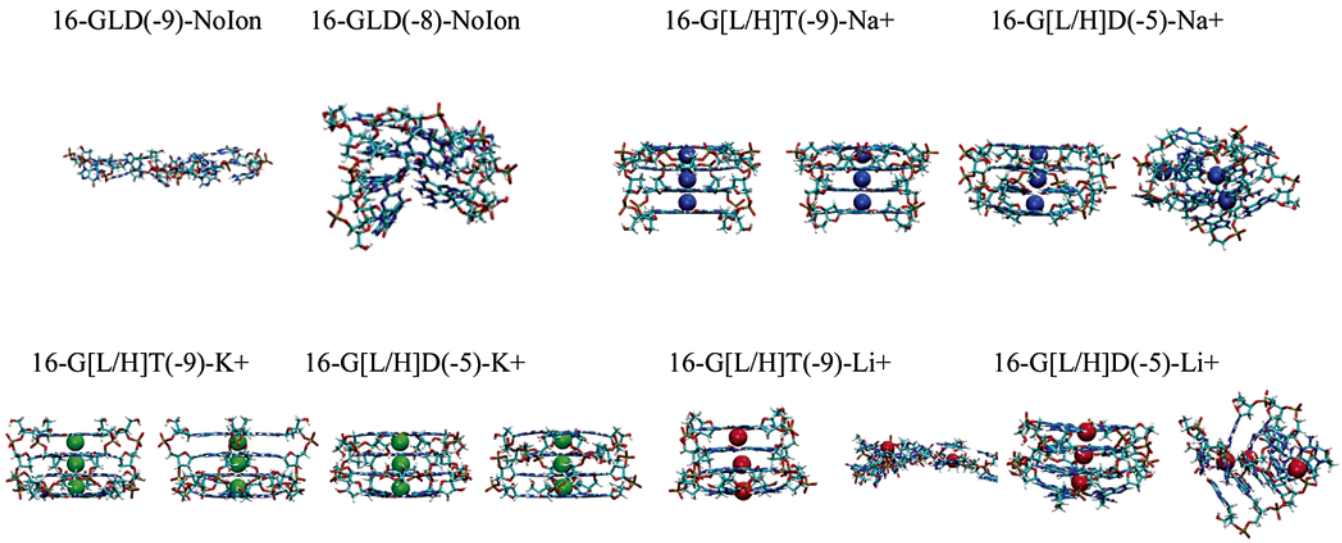

Figure 6. Average structures obtained at the final portion of the MD trajectories in the gas phase of the parallel G4-DNA from PDB entry 352D (16-mer). In each pair of structures, G stands for gas phase, L and H refer to low (298 K) or high (448 K) temperatures, and T stands for the total charge, while the distributed model is denoted as D. The less compacted structures have been reduced (16GLD(-9)-NoIon and 16GHT(-9)-Li+) to keep the sizes of the pictures. See also simulation codes in Table 1.

number of nucleotides[72]) is 3-fold larger in 352D than in 148D, when the number of nucleotides only increases by a factor of 2. Furthermore, a test simulation for the antiparallel 156D quadruplex (see above) confirms the larger intrinsic flexibility of parallel quadruplexes, since despite their identical size 156D (no-loop simulation) is clearly more rigid than 352D (see Table S4 in Supporting Information and http://mmb.pcb.ub.es/~manu/G4DNA/.).

MD Simulations in the Gas Phase. Antiparallel G4-DNA. Simulations of 148D in the absence of ions lead to extremely distorted geometries after a few nanoseconds (see Figure 3 and Table 2) due to the large magnitude of the phosphate−phosphate electrostatic repulsion. Interestingly, both SAS and CCS are still close to the values expected for G4-DNA, suggesting that caution is required in the interpretation of these structural parameters, since similar CCS can be associated with extremely different conformations. When an ion is placed in the cavity, the G4-DNA−ion complex is retained during the entire microsecond trajectory. However, different structures are obtained depending on the temperature and the nature of the central ion. Thus, in the presence of Li^+, simulations yield very distorted G-quadruplexes, which appear to be collapsed (note the small values of SAS and CCS; see Table 2 and Figure 3) and where many native contacts are lost. A similar situation is found for simulations with Na^+ and K^+ at high temperature (Table 2; Figures 3 and 4). However, the G-quadruplex is completely maintained during the entire microsecond in the trajectories run at room temperature with Na^+ or K^+ in the central cavity (Table 2; Figures 3 and 4). The only noticeable difference between the structures sampled in those trajectories and the control simulations in water are a certain increase in the *North* (phase angle centered around 0 ± 45°; typically C3′endo conformations) puckering (negligible in aqueous simulations and around 30−50% in the gas phase), though this does not have a major impact in the global structure (structures and rmsd plots available at http://mmb.pcb.ub.es/~manu/G4DNA). The cation is found in the central cavity in all simulations where the G-quadruplex structure is preserved (i.e, with Na^+ or K^+ at low temperature). However, when the structure collapses, the guanine−cation interactions are replaced with a large variety of contacts involving the DNA backbone. Even in very distorted G-DNA structures there is a large number of guanine−guanine interactions (see Figure 5). In general when the G-quadruplex structure is lost, guanines increase their stacking interactions, decreasing

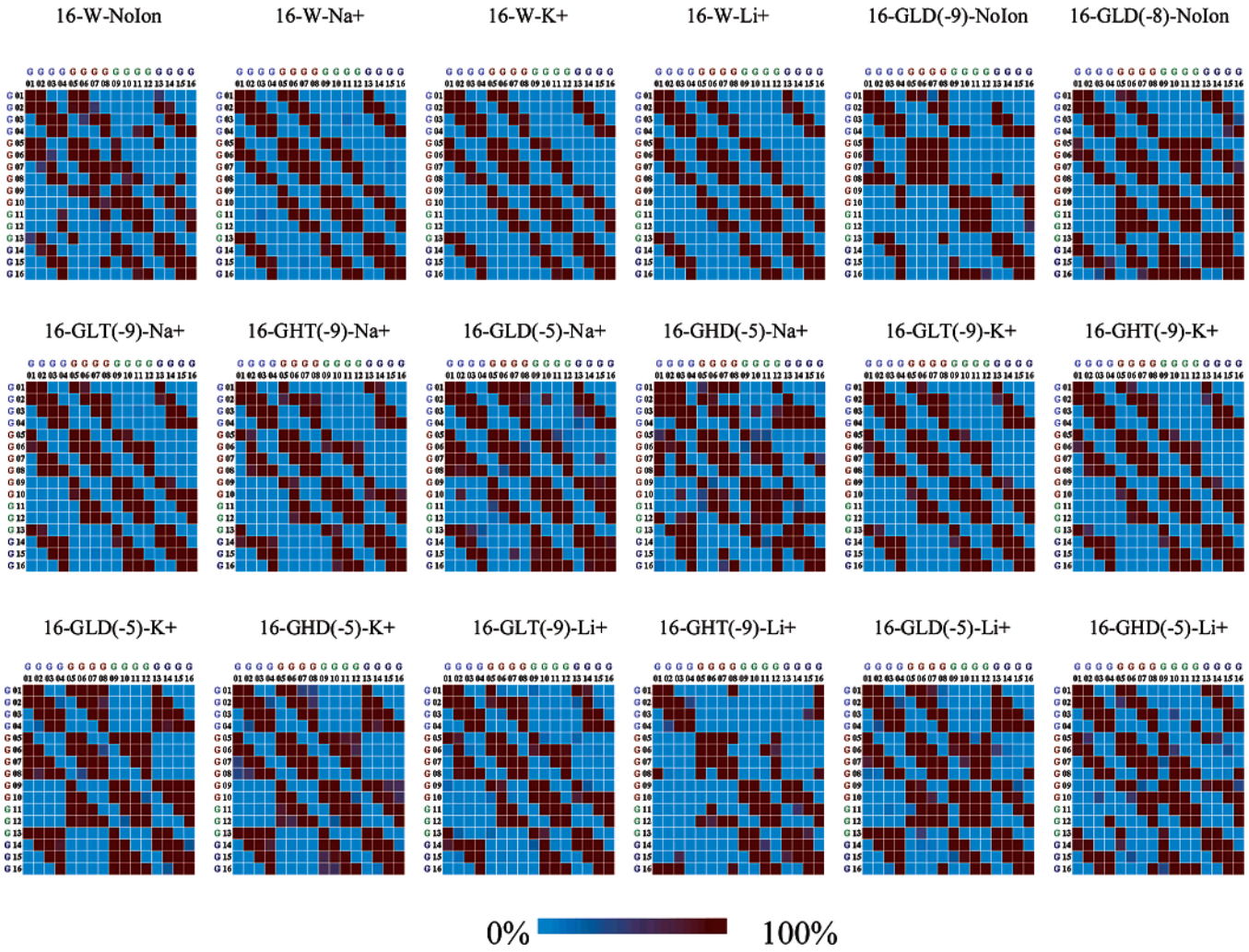

Figure 7. Contact maps for the final portion of the MD trajectories in the gas phase of the parallel G4-DNA from PDB entry 352D (16-mer). W stands for water, G stands for gas phase, L and H refer to low (298 K) or high (448 K) temperatures, and T stands for the total charge, while the distributed model is denoted as D. See also simulation codes in Table 1.

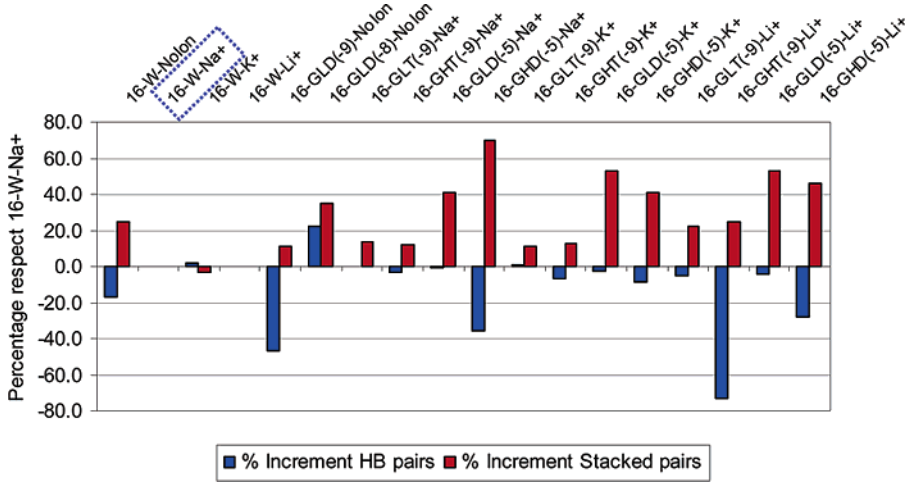

Figure 8. Percentage of gained/lost interactions (hydrogen bond and stacking) in the different simulations for parallel G4-DNA from PDB entry 352D (16-mer) with respect to the control simulation in water (W) in the presence of Na^+ in the central channel. G stands for gas phase, L and H refer to low (298 K) or high (448 K) temperatures, and T stands for the total charge, while the distributed model is denoted as D. See also simulation codes in Table 1.

their hydrogen bonding. Overall, the pattern of interactions becomes more complex in the gas phase than in solution.

The changes in the stiffness of cation-bound G4-DNA associated with its isothermal transfer (298 K) from solution to the gas phase are small and lead to a small gain in flexibility in cases where the structure is preserved (Na^+ and K^+) and a small loss of flexibility in the Li^+ simulation (see Table 3 and Supporting Information, Table S5). As expected, when the temperature is higher or when the ion is removed from the central cavity, the system gains flexibility (see Table 3 and

Table 7. Similarity Indexes (κ; See Text) between the Different Trajectories of the Parallel G-Quadruplexes (16-mer)[a]

16-W-NoIon	16-W-Na+	16-W-K+	16-W-Li+	16-GLD(-9)-NoIon	16-GLD(-8)-NoIon	16-GLT(-9)-Na+	16-GHT(-9)-Na+	16-GLD(-5)-Na+	16-GHD(-5)-Na+	16-GLT(-9)-K+	16-GHT(-9)-K+	16-GLD(-5)-K+	16-GHD(-5)-K+	16-GLT(-9)-Li+	16-GHT(-9)-Li+	16-GLD(-5)-Li+	16-GHD(-5)-Li+	
1.0	0.9	0.9	0.9	0.5	0.7	0.7	0.7	0.8	0.6	0.7	0.8	0.7	0.8	0.7	0.6	0.8	0.7	16-W-NoIon
	1.0	1.0	0.9	0.4	0.5	0.6	0.7	0.8	0.5	0.6	0.7	0.7	0.8	0.6	0.4	0.7	0.6	16-W-Na+
		1.0	0.9	0.4	0.5	0.6	0.7	0.7	0.5	0.6	0.7	0.6	0.8	0.6	0.4	0.7	0.6	16-W-K+
			1.0	0.4	0.5	0.7	0.7	0.7	0.5	0.6	0.7	0.7	0.8	0.6	0.4	0.7	0.5	16-W-Li+
				1.0	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	16-GLD(-9)-NoIon
					1.0	0.5	0.5	0.5	0.4	0.5	0.5	0.5	0.5	0.5	0.4	0.5	0.5	16-GLD(-8)-NoIon
						1.0	0.8	0.7	0.5	1.0	1.0	0.7	0.7	0.8	0.4	0.7	0.5	16-GLT(-9)-Na+
							1.0	0.7	0.5	0.8	0.9	0.6	0.7	1.0	0.4	0.6	0.5	16-GHT(-9)-Na+
								1.0	0.4	0.7	0.7	0.8	0.8	0.6	0.4	0.7	0.5	16-GLD(-5)-Na+
									1.0	0.5	0.5	0.4	0.5	0.5	0.4	0.4	0.6	16-GHD(-5)-Na+
										1.0	1.0	0.7	0.7	0.8	0.4	0.7	0.5	16-GLT(-9)-K+
											1.0	0.7	0.8	0.8	0.4	0.7	0.6	16-GHT(-9)-K+
												1.0	0.8	0.6	0.4	0.8	0.4	16-GLD(-5)-K+
													1.0	0.6	0.4	0.8	0.6	16-GHD(-5)-K+
														1.0	0.4	0.6	0.5	16-GLT(-9)-Li+
															1.0	0.4	0.4	16-GHT(-9)-Li+
																1.0	0.5	16-GLD(-5)-Li+
																	1.0	16-GHD(-5)-Li+

[a] W stands for water and G stands for gas phase, L and H refer to low (298 K) or high (448 K) temperatures, T stands for the total charge, while the distributed model is denoted as D. See also simulation codes in Table 1.

Supporting Information, Table S5). Finally, it is worth noting that the nature of the essential movements of the antiparallel G4-DNA in water is largely retained (around 60−70%) in the gas-phase simulations 8-GLT(-3)-Na^+ and 8-GLT(-3)-K^+, and that nonnegligible 20−30% similarity is found in simulations leading to distorted quadruplexes (see Table 6 and Supporting Information, Table S7).

According to our test simulations, the presence of connecting loops in 148D does not introduce major changes (see http://mmb.pcb.ub.es/∼manu/G4DNA) in the structure of the G-quadruplex in the gas phase (see Table S1 in Supporting Information). The total rmsd from NMR structure increases to more than 3 Å, but for the quadruplex portion the rmsd is only 1.9 Å, very close to the results found in the equivalent simulation without loops (see Table 2). The only remarkable structural difference induced by vaporization in the entire 148D structure is a small compression of the loops, which leads to a reduction of 5% in the CCS, but this small change does not affect the structure of the central G-quadruplex. Interestingly, connecting loops, which were very flexible in solution, becomes clearly more rigid in the gas phase, leading to a small increase in the stiffness of the central G-quadruplex upon vaporization. Test calculations with a longer antiparallel quadruplex (156D) do not provide additional insights, since most results obtained mirror those derived from the smaller quadruplex (see Tables S3 and S4 in Supporting Information and http://mmb.pcb.ub.es/∼manu/G4DNA). The only detectable difference is that for this largest quadruplex vaporization always leads to a reduction of flexibility (irrespective of whether loops are considered or not). In summary, no-loops thrombin aptamer seems to be a good model for antiparallel quadruplexes also in the gas phase.

Parallel G4-DNA. Three of the five simulations of parallel G4-DNA performed in the absence of ions lead to unstable trajectories after less than 20 ns, which avoid extension to 0.5 μs. The only two stable simulations were those performed at low temperature with the smallest net charges (−8 and −9) spread over all the phosphate groups. However, even in these cases the structures largely deviate from the canonical G4-DNA motif (see Figures 6 and 7; structures and rmsd plots available at http://mmb.pcb.ub.es/∼manu/G4DNA). In agreement with ESI-MS experiments, all the ion−G-quadruplex complexes are maintained during the entire trajectories (see Figure 6), even in those where the tetrad structure is lost (see Figure 6 and http://mmb.pcb.ub.es/∼manu/G4DNA). Most simulations performed in the presence of ions lead to compact structures, with CCS close to those of quadruplexes in solution.

However, these compact structures do not always correspond to real G-quadruplexes (see Figure 6 and Table 5), which suggests that conservation of CCS alone is not enough to demonstrate structure conservation. The degree of conservation of the G-quadruplex structure in the gas phase depends on the nature of the bound cation. Thus, four trajectories performed with K^+ and three done with Na^+ yielded structures very close to the canonical G-quadruplex after 0.5 μs of MD simulation in the gas phase (see Table 5, Figures 6 and 7, and http://mmb.pcb.ub.es/∼manu/G4DNA). As expected, Li^+ showed a lower ability to stabilize the parallel G-quadruplex in the gas phase than Na^+ or K^+, though in some cases G-quadruplexes resembling the canonical ones are found at low temperature (see Figures 6 and 7 and Table 5).

Even in trajectories leading to distorted G-quadruplexes, a large amount of intra- and interstrand contacts is found (see

Figure 7), with a general tendency to decrease the number of guanine−guanine hydrogen bonds and to increase the number of stacking contacts (Figure 8). It is not clear that all the distorted structures will have their residues much more accessible when the G-quadruplex structure is lost, which again casts doubts on the use of low resolution measures as evidence that the G-quadruplex structure is not altered in the gas phase.

Cation−DNA interactions are fully maintained in simulations leading to canonical G-quadruplexes, whereas when the G-quadruplex structure is lost the canonical guanine−cation interactions are replaced by a complex pattern of interactions with bases, phosphates, and sugars. As happens in the antiparallel structure, there is a small increase in the *North* puckering (2−5% in aqueous simulations and 1−25% in the gas phase), but this does not have a major impact in the global structure.

The parallel G4-DNA is in general stiffer in the gas phase than in solution, as noted in the entropies, configurational volume, and dimensionality determined at the same temperature (see Table 5 and Supporting Information, Table S6). The larger rigidity of the system can be explained not only by the maintenance of the three cation−DNA contacts (in most cases only two cation−DNA contacts are found in solution), but also by the slight compression of the G4-DNA structure upon transfer to the gas phase (see Table 5 and Supporting Information, Table S6). The only expected exceptions to this behavior correspond to distorted G4-DNAs, where a moderate increase in flexibility from aqueous values is found (see Table 5 and Supporting Information, Table S6). The nature of the essential movements of the parallel G4-DNA is conserved around 60−70% for cases where the G-quadruplex structure is maintained (see Table 7 and Supporting Information, Table S8), while such a conservation decreases, but not dramatically (typically to around 40−50%), in cases where the G-quadruplex structure is lost.

Conclusions

Both antiparallel and parallel G4-DNA are very stable in aqueous solution, provided that ions are placed in the central channel of the G-quadruplex. If cations are removed, the structure tries to capture them from the environment, but if the G-quadruplex is large the time scale required to capture several cations is greater than that for collapse of the G-quadruplex, which leads to disruption in the G4-DNA structure. In agreement with the experimental evidence, all simulations show that Na^+ and K^+ are more efficient that Li^+ to stabilize the G-quadruplex.

As expected from ESI-MS experiments, we have not detected any strand separation or ion-diffusion process when the G4-DNA is transferred from solution to the gas phase. Despite the large magnitude of phosphate−phosphate repulsions, the G-quadruplex structure is resistant to structural alterations if suitable cations are present. For short antiparallel G-quadruplexes, the structure persists after 1-μs MD simulation at room temperature. Indeed, longer parallel G4-DNA can maintain their structure intact in the microsecond time scale in the presence of Na^+ or K^+ even at very high temperatures. Very interestingly, even in cases where the G-quadruplex structure is lost, the molecule still exhibits a large amount of interactions, stacking then being enriched at the expense of some lost in canonical hydrogen bonding.

To the best of our knowledge, the surprising stability of the canonical structure of G4-DNA in the gas phase does not have a similar precedent among nucleic acids or any other biologically important macromolecule. The symmetry of the molecule is expected to be partially responsible for the rigidity of the system in hostile environments, but certainly the strength of cation−DNA interactions and the magnitude of hydrogen bond and stacking contacts should be responsible for the structural stability in the gas phase. Our findings strongly support the use of ESI-MS experiments to describe G4-DNA structure. Furthermore, combining the present results with those obtained in the analysis of duplex DNA−drug complexes in the gas phase, it can be expected that ESI-MS experiments can be well suited to analyze complexes of G4-DNA with small G-quadruplex binders of potential use as telomerase inhibitors.

Acknowledgment. We are indebted to Dr. Wyttenbach for a copy of SIGMA. This work has been supported by the Spanish DGCYT (BIO2003-06848 and SAF2002-04282), the Fundación BBVA, and the Fundació La Caixa. M.R. is a predoctoral fellow of the Spanish Ministry of Education and Science. Calculations were partially carried out on the *Mare Nostrum* computer at the Barcelona Supercomputing Center.

Supporting Information Available: Complete ref 68, extended Methods, results of test calculations with structures containing loops and larger antiparallel quadruplexes (Tables S1−S4) and complete Tables 3, 5, 6, and 7 (Tables S5−S8). This material is available free of charge via the Internet at http://pubs.acs.org.

JA055936S