Porphyrin Intercalation in G4-DNA Quadruplexes by Molecular Dynamics Simulations

Manuela Cavallari, Anna Garbesi, and Rosa Di Felice*

INFM-CNR National Center on nanoStructures and bioSystems at Surfaces (S3), I-41100 Modena, Italy

Received: April 28, 2009; Revised Manuscript Received: July 13, 2009

We investigated the intercalation of a porphyrin derivative (TMPyP) between guanine tetrads (G4-tetrads, G4t's) in 4-stranded G4-DNA oligomers by classical molecular dynamics simulations. Contrary to experimental evidence on very short oligomers that contain stabilizing cations, we find that TMPyP can stack with the G4-tetrads in the absence of interplane cations. A high TMPyP/G4t stoichiometric ratio of 1/2 induces strong deformations of the G4-quadruplexes. A lower ratio of 1/8 is better compatible with the helical conformation. When a TMPyP is accommodated between two tetrads, the stacking distance between the intercalated molecule and a G4-tetrad is ∼4.3−4.7 Å. We find that the possibility of regular TMPyP intercalation depends on the length of the quadruplex, on the stoichiometric ratio and on the edge termination motif.

1. Introduction

The interaction of porphyrins with G4-DNA quadruplexes has been long studied in the context of telomerase inhibition. It is known, in fact, that binding of small ligands in the G-rich telomeric regions of DNA can be effective in preventing cancer initiation and growth by promoting quadruplex folding.[1−3] Therefore, understanding the mechanisms that underlie such binding is a relevant issue for drug design. In a quite new scenario, it can also have relevant impact in the design/realization of molecular nanowires with electrical nonlinearities, for the implementation of nanoelectronics.[4−7]

Most experimental investigations of ligand binding to G4-DNA have concentrated on short strands whose folding is stabilized by positively charged alkali cations. For such samples, most results indicate that the TMPyP(4) porphyrin derivative[8] most favorably either stacks at the edges of a quadruplex or adsorbs laterally in a weak-coupling mode.[9−13] However, few studies highlight the possibility of intercalative coupling in small amounts.[14,15] In general, it can be inferred that the presence of inner cations is an obstacle to the insertion of positively charged intercalators because of steric hindrance and electrostatic repulsion. Recent synthesis and characterization experiments on long quadruplexes that can exist without inner cations[4] validated this hypothesis: Lubitz and co-workers[16] showed efficient intercalation of TMPyP in monomolecular G4-quadruplexes of about 1000 tetrads produced by enzymatic synthesis[4] and found a high TMPyP/G4t stoichiometric ratio of 1/2.

The intercalative mode for porphyrin/G4-quadruplex association was not probed so far by any theoretical approaches. Thus, we decided to undertake classical molecular dynamics simulations to inquire on whether and in which conditions G4-quadruplex structures with TMPyP intercalators resist unfolding, at least on the 10−20 ns time scale. In particular, we have simulated different TMPyP/G4t stoichiometries, different quadruplex lengths and different edge terminations. We find that the regularity of the quadruplex motif is affected by all such factors.

* Corresponding author. Mailing address: National Center on nanoStructures and bioSystems at Surfaces (S3) of INFM-CNR, c/o Dipartimento di Fisica, Università di Modena e Reggio Emilia, Via Campi 213/A, 41125 Modena, Italy. Phone: +39-059-2055320. Fax: +39-059-2055651. E-mail: rosa.difelice@unimore.it.

The paper is organized as follows: in section 2 we describe the adopted method and some aspects that influenced our choice of the simulated structures; in section 3 we report and discuss our results articulated in (i) average structures, (ii) H-bonding and stacking, (iii) dihedrals, (iv) root-mean-square deviations and (v) variance analysis; in section 4 we summarize our work and draw a few essential conclusions.

2. Method

Molecular Dynamics (MD) Simulations. The MD simulations were carried out using the NAMD[17] code with the all-atom CHARMM[18] force field. For validation, we repeated with CHARMM the 5-ns simulation of a non-intercalated 9-plane G4-quadruplex containing 8 inner K^+ ions, whose dynamics was recently performed by us with AMBER:[19] we found perfect agreement between the two methods for what concerns the average structure and statistical distribution of relevant quantities. Note that the natural choice for the nucleic acids in CHARMM is the ribonucleic backbone, and we retained this standard: hence, our test of CHARMM versus AMBER is also an indication that the different backbone does not significantly alter the main structural features in the dynamics of G4-quadruplexes. The porphyrin force field parameters were not developed in CHARMM. We parametrized a suitable force field. CHARMM-like parameters were built, tested and optimized using QUANTA (Accelrys)[20] and starting from the TMPyP4 crystallographic coordinates at a resolution of 2.09 Å.[20] Charges were computed using the PM3[21] semiempirical quantum-mechanical method.

Each simulated porphyrin/G4-quadruplex system was explicitly solvated using the TIP3P[22] model for water and the system charge was neutralized by an appropriate number of K^+ counterions. Periodic boundary conditions were applied to avoid surface effects. Water shells around the porphyrin/G4-quadruplex systems were chosen large enough to avoid van der Waals or electrostatic interactions between neighboring replicas. Precisely, we used water thicknesses ranging from 12.5 to 20 Å, depending on the *xyz* direction and the number of intercalated porphyrins. The latter are in fact responsible for quadruplex stretching along the main axis during the molecular dynamics (as described later), which must be taken into account to fix the water thickness in the direction of the quadruplex axis.

10.1021/jp9039226 CCC: $40.75
Published on Web 09/15/2009

Each simulation was carried out in two stages: an equilibration stage followed by a production MD run. In the equilibration phase we inserted porphyrins one by one, performing structural relaxation after each insertion. Each porphyrin was initially located in the middle of two adjacent G4-tetrads, at a tetrad−porphyrin stacking distance of 1.7 Å, maintaining the 3.4 Å stacking distance between the tetrads. Each relaxation comprised the following 5 steps: (i) a conjugate gradient[23] minimization of the whole system; (ii) a solvent equilibration; (iii) a further all-system minimization; (iv) a 60 ps dynamics run of the whole system in a heat bath where the temperature rose from 0 to 300 K at a rate of 50 K every 10 ps; (v) a final 100 ps equilibration during which the system temperature and pressure were kept constant at 300 K and 1.01325 bar, respectively, as in the subsequent MD production runs. In step (ii) the water molecules and counterions (solvent) were subjected to 20 ps dynamics: while keeping the nucleic acids and porphyrin atoms fixed with the SHAKE[24] algorithm with a tolerance of 10^{-8}, the temperature of the solvent was slowly raised to 100 K by coupling to a heat bath and the pressure was kept at 1.01325 bar (atmospheric pressure at sea level) using the Berendsen method.[25] In the remainder of the text, the structures resulted from these equilibration phases will be termed equilibrated structures.

During the production MD runs in the NPT scheme, the temperature was maintained constant through the addition of friction and random forces in a Langevin dynamics[26] scheme. A modified Nosé−Hoover[27] method, in the context of Langevin dynamics, controlled the fluctuations in the barostat. A 9.0 Å cutoff was applied to the Lennard-Jones interactions. To avoid the abrupt truncation of the electrostatic interactions, they were computed using the particle mesh Ewald method[28] (PME) with a charge grid spacing smaller than 1 Å and interpolated using a cubic B-spline. During the production runs distances between hydrogens and parent atoms were kept fixed at the equilibrated values, allowing a 2 fs time step. The system coordinates were collected every 1 ps for subsequent analysis. For each system we performed molecular dynamics simulations ranging from 10 to 20 ns after the equilibration phase, as detailed in the following, for a total dynamics simulation time of 80 ns.

Starting Structures. Starting atomic models are based on the X-ray crystal structure of the $d(TG_4T)_4$ quadruplex at 0.95 Å resolution[29] and the X-ray crystal structure of the TMPyP4 cationic porphyrin at 2.09 Å resolution.[13] We extracted one planar G4-tetrad from the experimental PDB file and constructed the stacked oligomers by subsequent rotations by 30° and translations by 3.37 Å of this primary tetrad. The terminal thymines were omitted, and the deoxyribonucleic backbone was replaced with a ribonucleic backbone for consistency with the adopted force field. The K^+ counterions in solution were initially placed in the most negative locations around the backbone. Details are given elsewhere.[19] No internal ions were present in the quadruplexes simulated in this work (except one structure mds1), consistently with the viability of empty "long" channels that we have found in a recent work.[19] The porphyrins were inserted one by one in the G4-quadruplexes during the equilibration phase, according to the protocol described above.

In Figure 1 we summarize all the simulated systems in a schematic way. Because previous theoretical and experimental investigations systematically indicated that porphyrin ligands are accommodated either at the sides or at the edges of short G4-quadruplexes, we started by inquiring on the possibility of inserting just one porphyrin in an intercalative mode in a 9-plane quadruplex with and without inner K^+ ions (mds1 and mds2, respectively), longer than ever explored in the context of intercalation studies. Then we took into account the TMPyP/G4t 1/2 stoichiometric ratio (mds3, mds4, mds5), as suggested by the results of recent biochemical synthesis experiments by the group of A. Kotlyar.[16] Such early experiments were done on monomolecular G4 wires. Later experiments conducted by the same group on four-stranded, parallel, G4 wires[30] revealed instead a smaller ratio of about 1/8−1/10.[31] Because the G4 wires that we simulated were all four-stranded, we finally also carried out a simulation for a TMPyP/G4-quadruplex system with the coarser TMPyP/G4t ratio of 1/8 (mds6).

Rationale for the Choice of the Structures. As we already mentioned, it is common wisdom that porphyrins do not bind to short G4-quadruplex oligomers by stacking between the tetrads. It is well-known that such short oligomers exist only in the presence of internal cations, which probably inhibit the "penetration" of one or more porphyrins by steric hindrance and electrostatic repulsion. On the other hand, Kotlyar and co-workers recently demonstrated that it is possible to produce by enzymatic synthesis long G4-quadruplexes that exist also in the absence of inner cations.[32] We also previously validated this finding by MD simulations.[19] Our choice of structures is guided by the questions of whether or not the viable absence of cations enables the penetrations of porphyrins and whether or not long empty fragments are important for edge stabilization of porphyrin-intercalated G4-quadruplexes.

Within this framework, mds1 is the obvious choice to start with: from a stable 9-plane quadruplex with inner K^+ ions,[19] we remove one ion to monitor whether the absence of an ion between two tetrad planes facilitates the entrance of a charged porphyrin at that location. Mds2 is expected to be less regular than mds1 because the 4-plane and 5-plane cation-empty edges are too short to maintain the regular quadruplex fold.[19,33] Mds3 is the structure that comes directly from the experimental findings for long monomolecular quadruplexes,[16] with the 1/2 TMPyP/G4t ratio throughout the 9-plane structure. Mds4 and mds5 are chosen to check whether and how much the edge structure is important for the overall stability: in fact, mds4 has the same 1/2 stoichiometric ratio as mds3, but with a more favorable edge termination, namely porphyrin capping,[13,34] while mds5 features a central segment with the same stoichiometry, flanked by long terminal empty[19,32] G4 tracts. Finally, mds6 has the lower 1/8 stoichiometric ratio and favorable edge cappings with one TMPyP at both 3′ and 5′ terminals.

The simulation of any possible structure that may be formed when porphyrins intercalate in G4-quadruplexes of different lengths is beyond the scope of this work and impossible within an atomistic approach. Coarse-grained models of polymer folding, in conjunction with molecular dynamics, would probably serve the scope and actually multiscale methods of this kind to solve the issue of intercalation mechanisms are not yet available and are still a very lively field of theoretical development. Thus, we limited ourselves to selected structures that allowed us to explore the viability of some experimentally suggested structural features and the details of the structures that may be obtained following such indications[16,31] (as explained in detailed above in this subsection), with the main question in mind being whether or not intercalation of porphyrins in G4-quadruplexes may be realized.

3. Results and Discussion

Average Structures. In Figure 2 we show the average structures relative to the whole production runs. Note that not all the simulations have the same duration. Most of them are

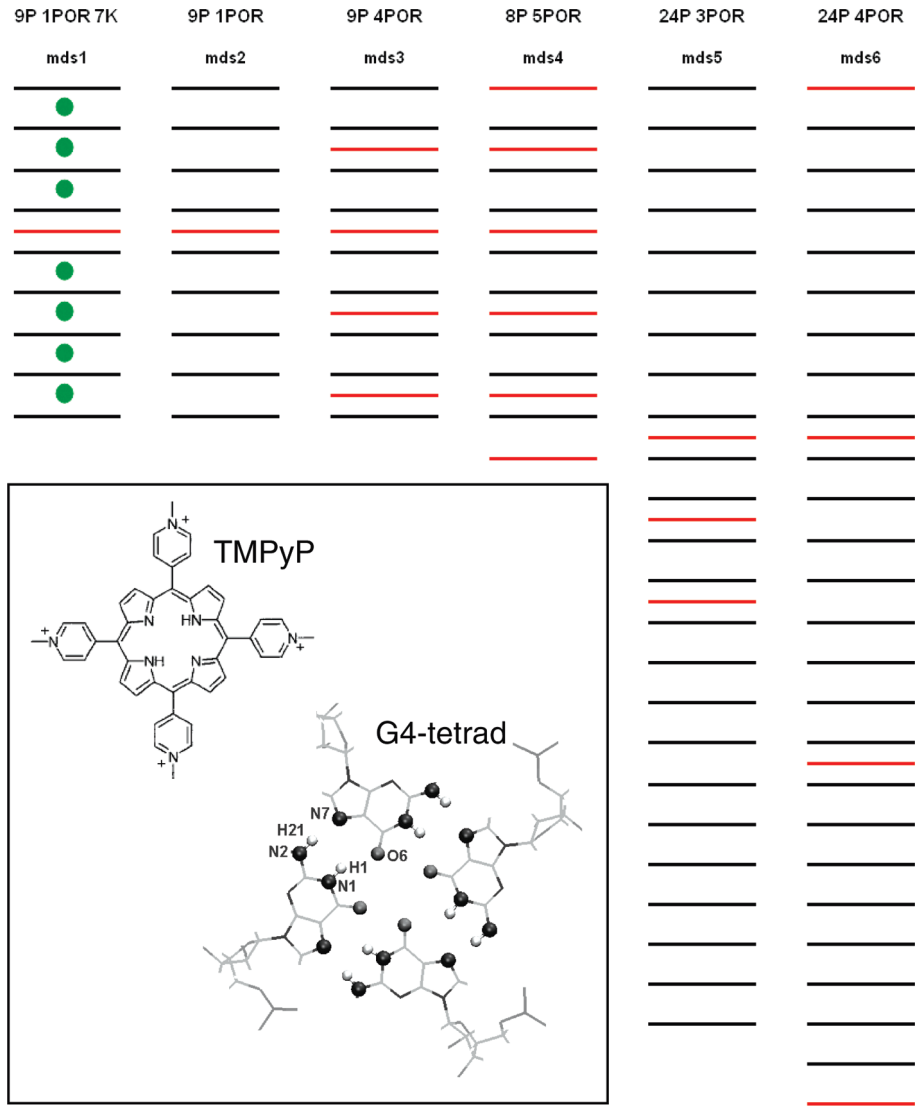

Figure 1. Schematic models of the simulated structures. The top labels indicate the number of planes (P), the number of porphyrins (POR) and the number of K^+ cations (K). The other labels (mds1−mds6) are used throughout the text and in the following Figures. Black segments represent G4-planes, referred to as G4-tetrads or simply tetrads in the text. Red segments represent porphyrins and green dots represent K^+ cations. Mds1 and mds2 mimic the intercalation of a single porphyrin. Mds3, mds4 and mds5 mimic the 1/2 TMPyP/G4t ratio with different edge terminations: random termination in mds3, edge-capped termination in mds4, termination by long G4 segments in mds5. Mds6 mimics the 1/8 ratio with edge capping. The edge-capping motif is the same in mds4 and mds6. Simulation times: 10 ns for mds1, mds2, mds3, mds5; 20 ns for the edge-capped mds4 and mds6. Inset: structure formula of the TMPyP derivative used in this study as the intercalator species and ball and stick representation of a G4-tetrad with labels for the atoms involved in H-bonds.

10 ns long (mds1, mds2, mds3 and mds5). Only for two systems, those capped by porphyrins at the edges (mds4 and mds6), we prolonged the molecular dynamics to 20 ns, in order to monitor the possible uncapping. Looking at the average structures of Figure 2, it is possible to outline some first qualitative considerations. A more quantitative statistical analysis is also presented later on.

The intercalation of a single porphyrin is viable in the presence of K^+ cations, as deduced from the regularity of the mds1 average structure: the edges are well stabilized by the cations (not shown), and the existence of an unoccupied interplanar location enables the accommodation of an intercalating porphyrin there. In the starting structure K^+ ions were located in all the intertetrad sites, with the exception of the one occupied by the porphyrin. During the dynamics three cations leave the structure. During the equilibration phase the outer cations lose their intertetrad coordination and move outside the quadruplex while staying close to the 5′ and 3′ tetrads, but during the production run they definitively abandon the quadruplex and enter the solution. This fact causes an overall rearrangement of the other cations, which shift toward the terminal sites of the quadruplex. Consequently, another ion leaves the structure from 3′ during the eighth nanosecond of the simulation and the intertetrad sites adjacent to the porphyrin remain empty, most likely because of its huge positive charge (+4*e*, being *e* the electron charge) that strongly repulses the cations in its proximity.

In the absence of coordinated K^+ ions, the 4-plane and 5-plane empty edges in mds2 are too short[19,33] and tend to unfold: in fact, the quadruplex mds2 suffers from major distortions, especially at the 3′ edge. This behavior is in line with well-known facts from theoretical and experimental studies, which indicate that in short G4 wires inner ions are essential for the regularity of quadruplex structure.[19,33,35−37]

While a single porphyrin intercalates quite easily, the accommodation of regularly spaced porphyrins is critical. Inter-

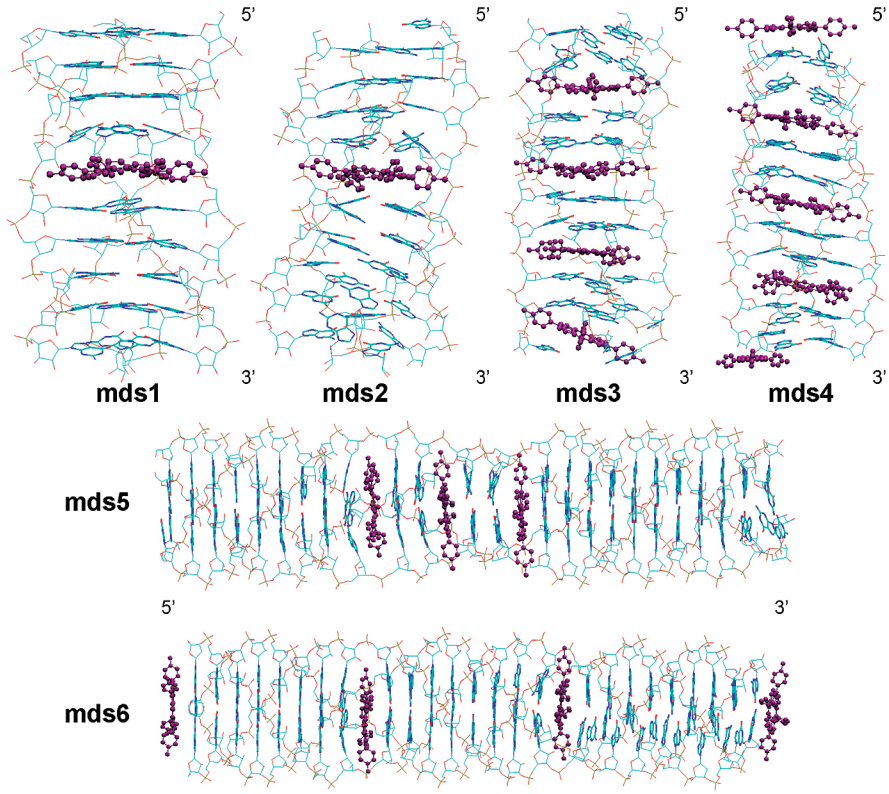

Figure 2. Average structures along the whole dynamics (production run). Mds1, mds2 (inner K^+ cations not shown), mds3 and mds5: 10 ns. Mds4 and mds6: 20 ns. The average structure of mds6 over 50 ns (Supporting Information) does not reveal any significant variation.

calation of porphyrins in the 1/2 TMPyP/G4t ratio inside short 9-plane G4-quadruplexes (mds3 and mds4) is also feasible without unfolding the G4-motif on the length scale of our simulations. However, notable structural distortions occur, especially in mds3, where the porphyrin close to the 3′ terminal loses the coplanarity with the G4-tetrads and destroys the 3′ tetrad. We remark that mds3 has very short 1-plane and 2-plane assemblies at the terminals. A tendency to disruption is also appreciable in the two 5′ tetrads of mds3. On the other hand, the terminal edges of the quadruplex are on average better preserved when porphyrins cap the G4-ends as in mds4, whose simulation was protracted to 20 ns, while the dynamics of mds3 was stopped at 10 ns. The prolongation of the dynamics was done because during the ninth nanosecond of the mds4 simulation we observed that the 3′ porphyrin started to lose its coordination with the quadruplex (also visible in the average structure as a lateral shift of the porphyrin). Hence, we continued the dynamics for other 10 ns and observed the detachment of that porphyrin from the tetrad-stack during the eleventh nanosecond. After this event the porphyrin did not leave the quadruplex, but it continued to be close to the G4-backbone.

The mds5 and mds6 simulations have been carried out in order to investigate the role of porphyrin intercalation in longer quadruplexes, taking into account different *local* stoichiometries and different terminal structures. In mds5 we maintained in the central portion the 1/2 ratio as for mds3 and mds4, but we terminated the quadruplex with long cation-empty stacks of G4-tetrads at both 3′ and 5′ edges. The average structure resulted from a 10 ns dynamics looks much more regular than the shorter systems with the same stoichiometry (mds3 and mds4), with a better preservation of both tetrad and porphyrin stacking. We ascribe this higher regularity to the fact that the empty terminal tracts are longer than for the previously described systems and long enough to maintain the quadruplex folding, as proved elsewhere.[19] Also for mds6, where we adopted a coarser TMPyP/G4t stoichiometry with the most external porphyrins capping the G4-quadruplex in 5′ and 3′, we observed a higher regularity of the structure relative to the shorter oligomers. We simulated mds6 for 20 ns, in order to probe a possible detachment of the capping porphyrins from the quadruplex, as seen in mds4, but it was not the case: these porphyrins remain well stacked with the G4-tetrads for the whole molecular dynamics run.

In summary, from the observation of the average structures and the dynamics trajectories, we infer that (i) TMPyP is a viable intercalator in G4-quadruplexes in the absence of cations in the channel; (ii) since the existence of an empty channel critically depends on the length of the empty tracts, the edge termination with either long empty tracts or capping porphyrins is an important stabilization factor; (iii) a high TMPyP/G4t stoichiometric ratio is quite difficult for these four-stranded quadruplexes, while a looser ratio is more doable.

Stacking and H-Bonding. The effects of porphyrin intercalation in G4-quadruplexes were also studied by analyzing the G4t−G4t and TMPyP−G4t stacking distances. The probability distributions[38] are shown in Figure 3, and the average values and standard deviations are collected in Table 1. We see that the average distance between two adjacent tetrads without an intercalated porphyrin (G4t−G4t) is 3.9 Å and the average distance between a tetrad and its adjacent porphyrin (TMPyP−G4t) is 4.5 Å. These average distances are obtained by averaging over all the directly stacked couples inside a single quadruplex and over all the simulated quadruplexes. Obviously, the average TMPyP−G4t distance is definitely larger than the value of 1.7 Å that was fixed in the starting structures: the insertion of a porphyrin induces a significant separation between the previously adjacent tetrads, which should also be reflected in an overall elongation of the quadruplex. Moreover, when a single porphyrin is intercalated (mds1 and mds2), its distance from

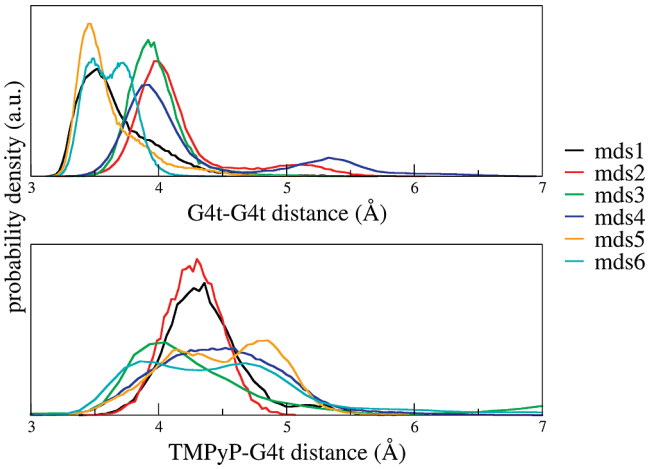

Figure 3. Probability density distributions of stacking distances between adjacent G4-tetrads without intercalating porphyrins (top) and between porphyrins and their adjacent G4-tetrads (bottom). Each distribution includes all the pairs of planes in the corresponding structure, except the 3′ terminal couple in mds4, because the edge-capping porphyrin abandons the edge.

TABLE 1: Average Stacking Distances with Standard Deviations

	distance (Å)	
structure	G4t−G4t[a,b]	TMPyP−G4t[b,c]
mds1	3.7 ± 0.3	4.3 ± 0.3
mds2	4.1 ± 0.4	4.3 ± 0.2
mds3	4.0 ± 0.2	4.7 ± 1.5
mds4	4.3 ± 0.7	4.5 ± 0.5
mds5	3.6 ± 0.3	4.5 ± 0.4
mds6	3.6 ± 0.2	4.5 ± 0.8

[a] Average values and standard deviations of distances between adjacent G4-tetrads not separated by a porphyrin. [b] Each average value includes all the pairs of planes in the corresponding structure, except the 3′ terminal couple in mds4 because the edge-capping porphyrin abandons the edge. [c] Average values and standard deviations of distances between a porphyrin and its adjacent G4-tetrads.

adjacent tetrads is well-defined with one net peak around 4.3 Å (Figure 3), while the distributions of this distance in the presence of more than one porphyrin in close sites are broad with an average value of 4.5 Å, and some of them exhibit double peaks. The G4t−G4t distance in mds1 has an average value of 3.7 Å, and the statistical distribution is well-defined around this value, but without coordinated cations (mds2) this distance becomes larger, that is 4.1 Å on average, making the stacking interaction weaker and the quadruplex conformation less regular. It may be expected that mds2 unfolds on time scales longer than the simulation time of 10 ns. In the systems with the high 1/2 TMPyP/G4t ratio we observe sharp distributions of this distance peaked around 4.0 and 4.3 Å for mds3 and mds4, respectively. On the other hand, in mds5, where the porphyrins are intercalated only in the central part of the quadruplex, with long empty edges and several nonintercalated dimers, this distance becomes smaller with the relative distribution well-defined around 3.6 Å. This is a further evidence of a good stacking behavior and overall regularity of the quadruplex motif when the systems have stable edges. This is also revealed in mds6. In fact, the distribution of the G4t−G4t distance in mds6, which has a smaller 1/8 TMPyP/G4t ratio than in mds5 and where the porphyrins are equally distributed along the G4 length, is different in shape from the mds5 distribution, but with the same average value.

To better evaluate the behavior of the G4-tetrads during the dynamics, we also computed the probability distributions of the donor−acceptor distances and angles in the Hoogsteen hydrogen bonds. The distribution curves[38] are shown in Figure 4, and the average values are collected in Table 2. A nominal hydrogen bond is considered to be formed if the distance between donor (N1 and N2) and acceptor (O6 and N7) atoms is below ∼3.5 Å and the donor−hydrogen−acceptor angle is within 30° from the flat angle. The smaller is the distance and the closer is the angle to 180°, the stronger is the H-bond. Consider, as a reference, that d_1(N1−O6) = 3.5 ± 0.5 Å, α(N1−H1−O6) = (140 ± 16)°, d_2(N2−N7) = 3.2 ± 0.4 Å and β(N2−H21−N7) = (152 ± 17)° for a 9-plane G4-quadruplex with inner K^+ ions.[19] Considering these H-bonding ranges, it appears that the strongest inner Hoogsteen hydrogen bonds (N1−H1−O6) are observed in mds5 and mds6, even stronger than those in mds1, with sharp distance and angle distributions markedly peaked around the average values. Moreover, the distance and angle distributions and average values in mds5 and mds6 are quite similar. In all the other quadruplexes these distributions are broader and the hydrogen bonds are on average weaker, with the weakest interaction found in mds3. A similar behavior is also observed for the outer Hoogsteen hydrogen bonds (N2−H21−N7): these are overall weaker than the inner hydrogen bonds (N1−H1−O6), as was already found in molecular dynamics simulations of G4-quadruplexes without intercalated porphyrins.[19]

The statistical data in Tables 1 and 2 serve to inquire on the stability of a G4-quadruplex in terms of intra- and inter-G4t interactions, namely in terms of Hoogsteen hydrogen bonds and distances, as well as of the coplanarity and stacking distance of adjacent tetrameric planes along the stack. As a matter of fact, these data reinforce our comments deduced more qualitatively above on the basis of Figure 2. Concisely, the most stable structures are the long (24 tetrads) quadruplexes with different *local* stoichiometric ratios and different edge motifs (mds5 and mds6), while the shorter quadruplexes with the 1/2 ratio are decidedly less regular and prone to unfolding.

Dihedral Angles. The behavior of the G4 backbone in the presence of intercalated porphyrins was evaluated with respect to two selected dihedrals, able to describe the backbone torsional mobility. The chosen dihedral angles are illustrated in Figure 5. One chosen dihedral is that formed by the C3′ and O3′ atoms of one tetrad and the P and O5′ atoms of the adjacent tetrad in the 3′ direction (A dihedral). The other chosen dihedral is that formed by the P, O5′, C5′ and C4′ atoms of the same tetrad (B dihedral). We did the dihedral analysis only for mds6, to have an additional proof of its good regularity. We examined three different couples of G4-tetrads, shown schematically in Figure 5: (i) far from intercalated porphyrins ($G4t_4$−$G4t_5$), (ii) in the proximity of an intercalated porphyrin ($G4t_7$−$G4t_8$) and (iii) at an intercalated site ($G4t_8$−$G4t_9$). These data for the mds6 TMPyP-intercalated G4-quadruplex were also compared to those for the central tetrad couple of the "empty" 24-plane quadruplex ($G4t_{12}$*−$G4t_{13}$*), whose dynamics (10 ns of production run) was described in our previous work.[19] Figure 6 shows the probability density distributions of these dihedrals. The distribution of the A dihedral in the empty quadruplex chosen as reference structure substantially consists of a main peak around 285° and a minor peak around 90°. Dihedral values between 90° and 285° have very small, nearly negligible, probabilities to occur. Considering the A dihedral formed by the $G4t_4$ and $G4t_5$ tetrads in mds6 (far from porphyrins) we still observe a 2-peak distribution: a main peak around 260° and a secondary

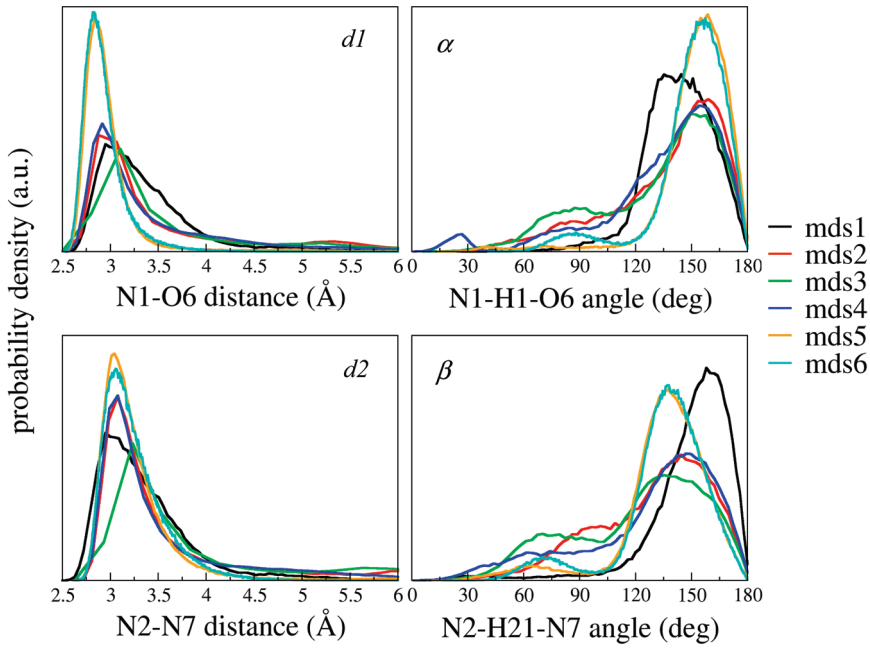

Figure 4. Probability density distributions of Hoogsteen hydrogen bond lengths (left) and angles (right) with respect the whole dynamics production runs. All the distance/angle values found in the production runs were considered to compute these distributions.

TABLE 2: Average Values and Standard Deviations of the Lengths (d_1, d_2) and Angles (α, β) of Hoogsteen Hydrogen Bonds

structure	d_1(N1−O6) (Å)	α(N1−H1−O6) (deg)	d_2(N2−N7) (Å)	β(N2−H21−N7) (deg)
mds1	3.3 ± 0.5	142.0 ± 18.2	3.3 ± 0.5	148.3 ± 23.7
mds2	3.5 ± 1.0	135.2 ± 30.9	3.6 ± 1.1	127.9 ± 30.1
mds3	4.5 ± 3.6	129.8 ± 33.2	4.7 ± 3.5	120.1 ± 34.9
mds4	3.6 ± 1.3	133.1 ± 33.5	3.7 ± 1.3	126.2 ± 35.9
mds5	3.0 ± 0.5	150.9 ± 20.7	3.3 ± 0.5	134.0 ± 24.9
mds6	2.9 ± 0.2	149.1 ± 21.1	3.2 ± 0.3	134.4 ± 24.5

broader peak around 180°. With respect to the former case, the range of allowed dihedral values is more restricted, but values between the two peaks have a more pronounced probability of occurrence. Also the distribution of the A dihedral for the $G4t_8$ and $G4t_9$ tetrads, which are intercalated by a porphyrin, exhibits two peaks, with the main peak around 280° and the other,

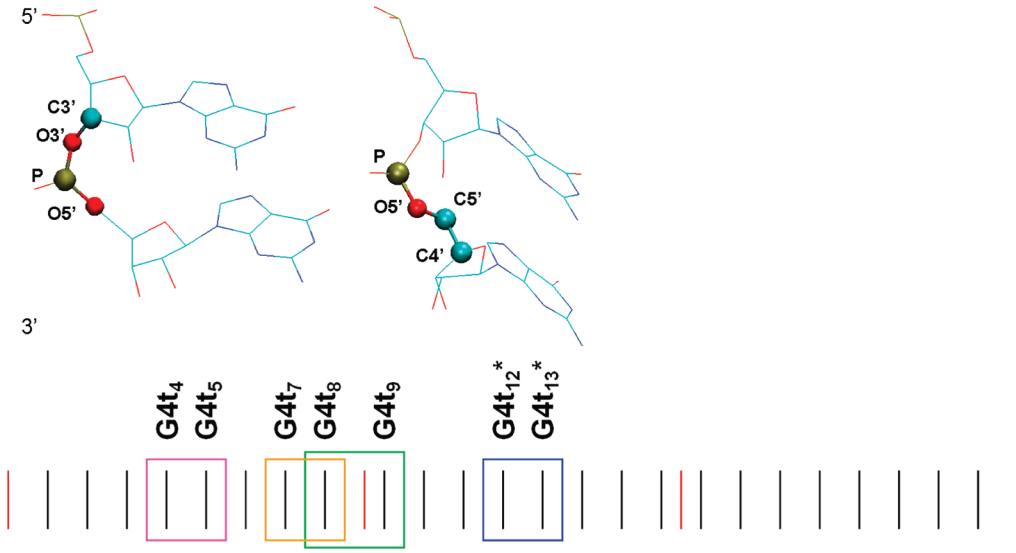

Figure 5. Top: ball and stick representation of the two selected backbone dihedrals discussed in this work and summarized in Table 2. Bottom: schematic picture of the md6 structure to define the couples of tetrads chosen for the analysis of the dihedrals. G4-tetrads (porphyrins) are in black (red). G4-tedrads involved in backbone dihedral evaluation are explicitly indicated with rectangles of different colors, with the same color code as used for the plots in Figure 6. The couples $G4t_4$−$G4t_5$, $G4t_7$−$G4t_8$ and $G4t_8$−$G4t_9$ belong to the mds6 quadruplex of this work. The couple $G4t_{12}$*−$G4t_{13}$* belongs to the 24-plane "empty" quadruplex simulated with AMBER in our previous work:[19] given the equivalence of AMBER and CHARMM in the representative test case of the 9-plane quadruplex with 8 K^+ ions (see text at the beginning of the Method section), we can reasonably assume that the two force fields perform equally also for this system and therefore take it as a reference to analyze what happens in the dihedrals upon intercalation with TMPyP.

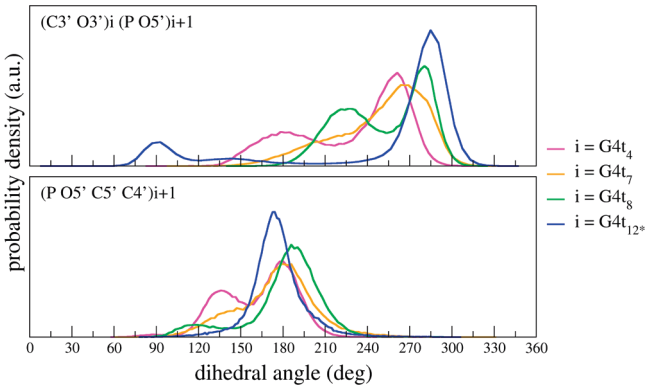

Figure 6. Probability density distributions of the backbone dihedrals for the couples of tetrads that are defined in Figure 5, calculated over the whole production runs. All the dihedral values found in the production runs (20 ns for mds6 and 10 ns for the "empty" G4 quadruplex) were used to obtain these distributions.

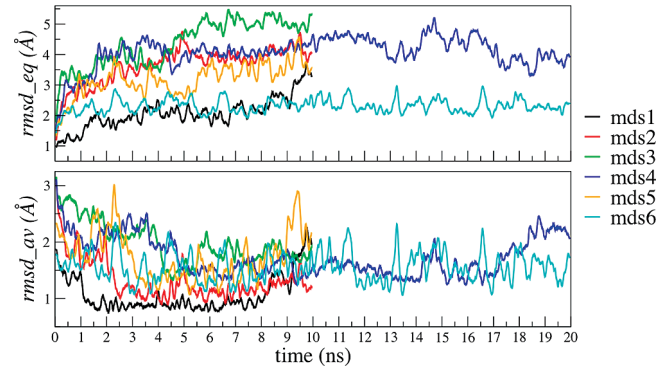

Figure 7. Root mean square deviations relative to the equilibrated structure (top) and the average structure (bottom), computed over all the G4 atoms except hydrogens, every 1 ps of the dynamics (production runs) according to the defining formula for rmsd in the text.

broader, around 225°. Both these peaks correspond to the probability minima of the $G4t_4-G4t_5$ distribution, pointing out a substantial change of the backbone dynamics in the immediate vicinity of a porphyrin. A somewhat intermediate trend characterizes the dihedral distribution of the intermediate couple of tetrads ($G4t_7-G4t_8$): it is characterized by just one wide peak that covers the main peaks of the other distributions, with a shoulder that comprises all the other possible values in those distributions. The distribution of the B dihedral in the empty reference quadruplex shows a single sharp peak around 170°, while those of the mds6 tetrads are broader, with main peaks in proximity of the peak for the empty structure and shoulders or secondary peaks at smaller values.

In short, we conclude that in the presence of intercalated porphyrins the flexibility of the G4 backbone, evaluated in terms of the selected dihedrals, increases, allowing a broader range of values than those assumed in the empty quadruplex, along with a higher tendency to bimodality in particular for the B dihedral.

Root Mean Square Deviations of Atomic Coordinates. The time-dependent relative motion of the quadruplexes during the dynamics can be described by the root-mean-square deviations (rmsd) of the atomic coordinates, relative to chosen reference structures. Rmsd's can be calculated for each saved step of the production runs with the formula

$$\mathrm{rmsd}(t) = \left[\frac{1}{N}\sum_{i=1}^{N}(\vec{r}_i(t) - \langle\vec{r}_i\rangle_T)^2\right]^{1/2}$$

where $\vec{r}_i(t)$ are the coordinates of the ith atom at time t, $\langle\vec{r}_i\rangle_T$ are the coordinates of the ith atom in the reference structure, which is most often either the equilibrated structure or the average structure. In the latter case, the average is performed over the time interval T. N is the number of atoms in the domain of interest.

In Figure 7, rmsd curves are calculated with reference to both the equilibrated structure (rmsd_eq) and the average structure (rmsd_av), performing the sum in the above equation over all the G4 atoms with the exception of the hydrogens. Table 3 collects the corresponding average values and standard deviations, together with those obtained by summing only on the guanine atoms and the backbone atoms separately, again with the exception of the hydrogens. Rmsd_eq's are on average larger than rmsd_av's for each quadruplex, meaning that the molecular dynamics induces a significant structural alteration of the quadruplex with respect to the starting geometry, but the resulting structure is generally stable around an average conformation. Moreover, rmsd_eq's exhibit a wider range of variability as a function of the simulated quadruplex than rmsd_av's. Namely, deviations from the equilibrated structure can change significantly as a function of the length of the quadruplex and the porphyrin/tetrad stoichiometry, whereas deviations from the average structure achieved with the dynamics are less affected by these parameters. Atoms forming the G4-tetrads experience smaller deviations than backbone atoms in all the simulated quadruplexes, considering both the equilibrated and the average structures as references.

Variance Analysis. In order to investigate more rigorously the internal flexibility/rigidity of the simulated G4-quadruplexes, we computed atomic mean square displacements (msd) using a technique[39] inspired by the analysis of the variance. In brief, this method consists of the following steps: (i) dividing the whole simulation time in n "blocks" and performing "within-blocks" averages of the linear and squared atomic positions, $\langle\vec{r}_i\rangle_j$ and $\langle\vec{r}_i^{\,2}\rangle_j$, over such finite time intervals j rather than over the whole trajectory; (ii) defining the within-blocks atomic mean square displacement, $\langle\mathrm{msd}_i\rangle_{\mathrm{wb}}$, for atom i; (iii) analyzing the dependence of the resulting $\langle\mathrm{msd}\rangle_{\mathrm{wb}}$ distribution of all the G4-quadruplex atoms on the extent of the blocks on which the averages are computed; (iv) selecting from the test (iii) the block extent for which there is a sort of plateau of the relevant index obtained through a χ^2 test; (v) performing the distribution analysis for the selected block width. The $\langle\mathrm{msd}\rangle_{\mathrm{wb}}$ in this analysis is defined for atom i as

$$\langle\mathrm{msd}_i\rangle_{\mathrm{wb}} = \frac{1}{3n}\sum_{j=1}^{n}(\langle x_i^2\rangle_j - \langle x_i\rangle_j^2 + \langle y_i^2\rangle_j - \langle y_i\rangle_j^2 + \langle z_i\rangle_j^2 - \langle z\rangle_j^2)$$

where the subscript wb stands for within-blocks and x_i, y_i, z_i are the Cartesian components of vector $\vec{r}_i$. The reader is referred to the original work for more details.[39]

We report for each simulated system the probability distribution of the $\langle\mathrm{msd}\rangle_{\mathrm{wb}}$ calculated for all the guanine atoms excluding hydrogens.

We performed the χ^2 test on the guanine atoms of mds4 and mds6 as representative of short and long quadruplexes, respectively, both with edge capping by TMPyP and representative of different stoichiometries. As suggested by Maragliano et al.,[39] the χ^2 test was carried out considering both the whole trajectory of 20 ns and the final 10 ns, calculating $\langle\mathrm{msd}\rangle_{\mathrm{wb}}$ for 30 different

TABLE 3: Average Root Mean Square Deviations with Their Standard Deviations[a]

structure	rmsd_eqG4 (Å)	rmsd_avG4 (Å)	rmsd_eqguanine (Å)	rmsd_avguanine (Å)	rmsd_eqbackbone (Å)	rmsd_avbackbone (Å)
mds1	2.1 ± 0.5	1.1 ± 0.3	1.4 ± 0.4	0.8 ± 0.3	2.6 ± 0.6	1.3 ± 0.4
mds2	3.6 ± 0.6	1.3 ± 0.4	2.8 ± 0.4	1.1 ± 0.3	4.3 ± 0.8	1.5 ± 0.5
mds3	4.4 ± 0.7	2.0 ± 0.4	3.7 ± 0.7	2.0 ± 0.4	4.8 ± 0.8	1.9 ± 0.5
mds4	4.1 ± 0.5	1.7 ± 0.4	3.3 ± 0.4	1.5 ± 0.3	4.6 ± 0.6	1.9 ± 0.4
mds5	3.2 ± 0.5	1.7 ± 0.4	2.4 ± 0.4	1.5 ± 0.4	3.8 ± 0.7	1.9 ± 0.5
mds6	2.3 ± 0.3	1.6 ± 0.3	1.9 ± 0.3	1.3 ± 0.3	2.6 ± 0.3	1.7 ± 0.3

[a] Rmsd_eq is computed relative to the equlibrated structure. Rmsd_av is computed relative to the average dynamical structure. The label "G4" indicates that the rmsd is computed by summing over all the G4 atoms, including guanine and backbone. The label "guanine" indicates that the rmsd is computed by summing over only the guanine atoms. The label "backbone" indicates that the rmsd is computed by summing over only the backbone atoms. The H atoms are always excluded from the rmsd evaluation.

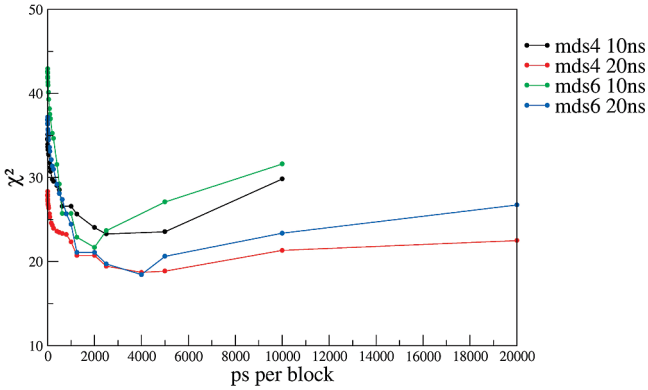

Figure 8. Values of the χ^2 variable for the distributions of within-blocks mean square displacements with different block lengths, calculated over the whole 20 ns trajectory and over the final 10 ns for the mds4 and mds6 G4 quadruplexes. The stability region can be roughly identified between 1250 and 4000 ps, and we chose 2000 ps as a proper block length for the calculation of atomic mean square displacements within blocks for all the simulated structures.

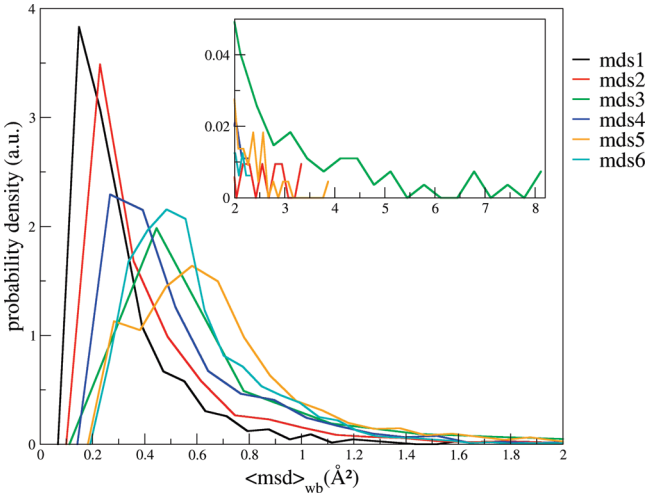

Figure 9. Probability density distributions of the mean square displacements of G4 atoms excluding hydrogens calculated after the procedure developed by Maragliano and co-workers.[39]

block lengths in the case of the 20 ns trajectory and 24 block lengths in the case of the 10 ns trajectory. From the results of the test (Figure 8), we chose a block length of 2 ns for the calculation of atomic mean square displacements within blocks for all the simulated structures.

Figure 9 reports the probability density distributions of the $\langle msd \rangle_{wb}$ for each G4-quadruplex computed in this way, and in Table 4 $\langle msd \rangle_{wb}$ average values and standard deviations are collected. The smallest $\langle msd \rangle_{wb}$'s are observed in the two quadruplexes with only one intercalated porphyrin (mds1 and mds2), especially in mds1, where potassium cations are present

TABLE 4: Average Values and Standard Deviations of the $\langle msd \rangle_{wb}$

structure	$\langle msd \rangle_{wb}$ G4[a] (Å^2)
mds1	0.27 ± 0.21
mds2	0.37 ± 0.34
mds3	0.60 ± 0.86
mds4	0.46 ± 0.31
mds5	0.64 ± 0.37
mds5 Gt_9-Gt_{14}	0.70 ± 0.21
mds6	0.56 ± 0.27

[a] The label "G4" indicates that the $\langle msd \rangle_{wb}$ is computed by summing over all the G4 atoms, including guanine and backbone, excluding hydrogen atoms.

between G4-tetrads in the interplanar spaces where porphyrins are not inserted. The corresponding probability distributions are sharp and narrow around the average values, revealing a high degree of rigidity of the whole quadruplex. Porphyrin intercalation makes the $\langle msd \rangle_{wb}$ larger, in terms of both average values and distributions. Higher $\langle msd \rangle_{wb}$ average values mean an overall increase of the quadruplex mobility, and broader distributions reveal an increased structural flexibility. The most flexible quadruplex is mds3, where $\langle msd \rangle_{wb}$ covers a range from nearly 0 to 0.8 Å^2. In mds4 the TMPyP/G4t ratio is the same as in mds3, but, differently from mds3, the terminal tetrads in 3′ and 5′ are both capped by a porphyrin, and this is enough to reduce the structural mobility and flexibility. This was already stated above from other data, and is revealed here by the fact that the peak of the $\langle msd \rangle_{wb}$ distribution is shifted to a lower value. As previously said, mds5 can be viewed as a short quadruplex with intercalated porphyrins with a 1/2 stoichiometric ratio like mds3 and mds4, but with a different edge capping with long G4 segments. Indeed, the overall $\langle msd \rangle_{wb}$ distribution of mds5 is the convolution of the distributions of these structural elements. We also inspect the $\langle msd \rangle_{wb}$ distribution obtained by considering only the short central quadruplex, comprising the tetrads from 9 to 14, with the indication of its average value and standard deviation in Table 4. This distribution is sharper than the mds4 (and consequently mds3) distribution and peaked around an average value larger than that for the mds3 (and consequently mds4) distribution. From these observations it results that different edge capping motifs differently influence the structural flexibility/mobility of short porphyrin-intercalated G4 quadruplexes. A short porphyrin-intercalated quadruplex capped by only one or two G4-tetrads at each edge (mds3) exhibits a high flexibility (broad $\langle msd \rangle_{wb}$ distribution), whereas capping by porphyrins (mds4) or long G4 segments (mds5) makes the quadruplex more rigid (sharp $\langle msd \rangle_{wb}$ distributions) and induces a different degree of mobility ($\langle msd \rangle_{wb}$ average value) dependent on the capping motif. Finally, the 24-tetrad mds6 is edge capped by porphyrins as mds4, but has a smaller TMPyP/G4t stoichiometric ratio. The $\langle msd \rangle_{wb}$ average value in mds6 is somewhat

larger than that in mds4, and the $\langle msd \rangle_{wb}$ probability distribution is somewhat broader for mds4 than for mds6. Apart from these slight differences, the variance analysis reveals that the degree of mobility and flexibility of these two quadruplexes is very similar, again reinforcing the observations deduced from the other kinds of analyses presented above.

4. Concluding Remarks

We presented the results of classical molecular dynamics simulations of porphyrin-intercalated parallel G4-quadruplexes of different lengths, with different TMPyP/G4t ratio and with different edge-capping motifs. The trajectories were analyzed in terms of average structure, observation of dynamical events and probability distributions for stacking distances, dihedral angles, lengths and angles of Hoogsteen H-bonds. In addition, we analyzed the time evolution and statistical distribution of total root-mean-square deviations. All these analyses reinforce one another in revealing that TMPyP may be a viable intercalator in G4-DNA oligomers of at least 8 G4-tetrads. The establishment of an optimal stoichiometry and of a convenient edge termination is crucial for maintaining the quadruplex folding upon TMPyP intercalation.[40]

Acknowledgment. We thank Francesca Fanelli for aid in the generation of the CHARMM-like force field parameters for TMPyP4 and Stefano Corni for help in visualization. This work was funded by the EC through project "DNA-based Nanodevices" (Contract FP6-029192). Access to the CINECA supercomputing facilities was granted by INFM-CNR.

Supporting Information Available: Test on the simulation length for mds6. This material is available free of charge via the Internet at http://pubs.acs.org.

References and Notes

(1) Izbicka, E.; Wheelhouse, R. T.; Raymond, E.; Davidson, K. K.; Lawrence, R. A.; Sun, D.; Windle, B. E.; Hurley, L. H.; Von Hoff, D. D. *Cancer Res.* **1999**, *59*, 639–644.

(2) Kim, M.-Y.; Gleason-Guzman, M.; Izbicka, E.; Nishioka, D.; Hurley, L. H. *Cancer Res.* **2003**, *63*, 3247–3256.

(3) Grand, C. L.; Han, H.; Muñoz, R. M.; Weitman, S.; Von Hoff, D. D.; Hurley, L. H.; Bearss, D. J. *Mol. Cancer Ther.* **2002**, *1*, 565–573.

(4) Kotlyar, A. B.; Borovok, N.; Molotsky, T.; Fadeev, L.; Gozin, M. *Nucleic Acids Res.* **2005**, *33*, 525–535.

(5) Kotlyar, A. B.; Borovok, N.; Molotsky, T.; Cohen, H.; Shapir, E.; Porath, D. *Adv. Mater.* **2005**, *17*, 1901–1905.

(6) Cohen, H.; Sapir, T.; Borovok, N.; Molotsky, T.; Di Felice, R.; Kotlyar, A. B.; Porath, D. *Nano Lett.* **2007**, *7*, 981.

(7) Porath, D.; Cuniberti, G.; Di Felice, R. *Top. Curr. Chem.* **2004**, *237*, 183–227.

(8) TMPyP4 = *meso*-tetrakis(4-*N*-methylpyridyl)porphyrin.

(9) Wheelhouse, R. T.; Sun, D.; Han, H.; Han, F. X.; Hurley, L. H. *J. Am. Chem. Soc.* **1998**, *120*, 3261–3262.

(10) Han, F. X.; Wheelhouse, R. T.; Hurley, L. H. *J. Am. Chem. Soc.* **1999**, *121*, 3561–3570.

(11) Han, H.; Langley, D. R.; Rangan, A.; Hurley, L. H. *J. Am. Chem. Soc.* **2001**, *123*, 8902–8913.

(12) Wei, C.; Jia, G.; Yuan, J.; Feng, Z.; Li, C. *Biochemistry* **2006**, *45*, 6681–6691.

(13) Parkinson, G. N.; Ghosh, R.; Neidle, S. *Biochemistry* **2007**, *46*, 2390–2397.

(14) Anantha, N. V.; Azam, M.; Sheardy, R. D. *Biochemistry* **1998**, *37*, 2709–2714.

(15) Haq, I.; Trent, J. O.; Chowdhry, B. Z.; Jenkins, T. C. *J. Am. Chem. Soc.* **1999**, *121*, 1768–1779.

(16) Lubitz, I.; Borovok, N.; Kotlyar, A. B. *Biochemistry* **2007**, *46*, 12925–12929.

(17) Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kale, L.; Schulten, K. *J. Comput. Chem.* **2005**, *26*, 1781–1802. NAMD was developed by the Theoretical and Computational Biophysics Group in the Beckman Institute for Advanced Science and Technology at the University of Illinois at Urbana-Champaign.

(18) (a) Foloppe, N.; MacKerell, A. D., Jr. *J. Comput. Chem.* **2000**, *21*, 86–104. (b) MacKerell, A. D., Jr.; Banavali, N. *J. Comput. Chem.* **2000**, *21*, 105–120.

(19) Cavallari, M.; Calzolari, A.; Garbesi, A.; Di Felice, R. *J. Phys. Chem. B* **2006**, *110*, 26337.

(20) http://accelrys.com/products/quanta/.

(21) Stewart, J. J. P. *J. Mol. Model.* **2004**, *10*, 155–164.

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

(23) Fletcher, R.; Reeves, C. M. *Comput. J.* **1964**, *7*, 149.

(24) Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. C. *J. Comput. Phys.* **1977**, *23*, 237.

(25) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; DiNola, A.; Haak, J. R. *J. Chem. Phys.* **1984**, *81*, 3684.

(26) Kolb, A.; Dünweg, B. *J. Chem. Phys.* **1999**, *111*, 4453.

(27) Nosé, S. *J. Chem. Phys.* **1984**, *81*, 511.

(28) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. *J. Chem. Phys.* **1995**, *103*, 8577.

(29) Phillips, K.; Dauter, Z.; Murchie, A. I. H.; Lilley, D. M. J.; Luisi, B. *J. Mol. Biol.* **1997**, *273*, 171.

(30) Borovok, N.; Iram, N.; Zikich, D.; Ghabboun, J.; Livshits, G. I.; Porath, D.; Kotlyar, A. B. *Nucleic Acids Res.* **2008**, *36*, 5050–5060.

(31) Kotlyar, A. B. Private communication.

(32) Borovok, N.; Molotsky, T.; Ghabboun, J.; Porath, D.; Kotlyar, A. B. *Anal. Biochem.* **2008**, *374*, 71–78.

(33) Fadrna, E.; Špackova, N.; Štefl, R.; Koèa, J.; Cheatham, T. E., III; Šponer, J. *Biophys. J.* **2004**, *87*, 227–242.

(34) Yamashita, T.; Uno, T.; Ishikawa, Y. *Bioorg. Med. Chem.* **2005**, *13*, 2423–2430.

(35) Hardin, C. C.; Watson, T.; Corregan, M.; Bailey, C. *Biochemistry* **1992**, *31*, 833.

(36) Xu, Q.; Deng, H.; Braunlin, W. *Biochemistry* **1993**, *32*, 13130.

(37) Guschlbauer, W.; Chantot, J. F.; Thiele, D. *J. Biomol. Struct. Dyn.* **1990**, *8*, 491.

(38) In this work the represented probability density distributions of the quantities of interest are obtained as linear convolutions of the computed normalized histograms collecting all the values of each quantity found in the production runs. We used Scott's choice (Scott, David W. (1979). "On optimal and data-based histograms". Biometrika 66 (3)) to compute the histogram bin with h, that is: $h = 3.5s/n^{1/3}$, where s is the sample standard deviation and n is the number of collected values.

(39) Maragliano, L.; Cottone, G.; Cordone, L.; Ciccotti, G. *Biophys. J.* **2004**, *86*, 2765–2772.

(40) Our results and our main conclusions essentially derive from the stacking interactions between the porphyrins and the bases, and are thus unlikely to be dependent on the ribonucleic backbone.

JP9039226