Structural Dynamics of Human Telomeric G-Quadruplex Loops Studied by Molecular Dynamics Simulations

Hong Zhu¹, Shiyan Xiao¹, Haojun Liang¹,*

¹ CAS Key Laboratory of Soft Matter Chemistry, Department of Polymer Science and Engineering, University of Science and Technology of China, Hefei, Anhui, People’s Republic of China, 2 Hefei National Laboratory for Physical Sciences at Microscale, University of Science and Technology of China, Hefei, Anhui, People’s Republic of China

Abstract

Loops which are linkers connecting G-strands and supporting the G-tetrad core in G-quadruplex are important for biological roles of G-quadruplexes. TTA loop is a common sequence which mainly resides in human telomeric DNA (hTel) G-quadruplex. A series of molecular dynamics (MD) simulations were carried out to investigate the structural dynamics of TTA loops. We found that (1) the TA base pair formed in TTA loops are very stable, the occupied of all hydrogen bonds are more than 0.95. (2) The TA base pair makes the adjacent G-quartet more stable than others. (3) For the edgewise loop and the diagonal loop, most loop bases are stacking with others, only few bases have considerable freedom. (4) The stabilities of these stacking structures are distinct. Part of the loops, especially TA base pairs, and bases stacking with the G-quartet, maintain certain stable conformations in the simulation, but other parts, like TT and TA stacking structures, are not stable enough. For the first time, spontaneous conformational switches of TTA edgewise loops were observed in our long time MD simulations. (5) For double chain reversal loop, it is really hard to maintain a stable conformation in the long time simulation under present force fields (parm99 and parmbsc0), as it has multiple conformations with similar free energies.

Introduction

G-quartet is a layer formed by four guanines held together with eight hydrogen bonds. In guanine-rich nucleic acid sequences, the formation of several consecutive G-quartets can form inter-molecular or intra-molecular four-stranded structures, termed G-quadruplexes. Telomeres, at the end of most eukaryotic chromosomes, usually comprise simple tandem repeats of guanine rich sequences which can form non-canonical DNA conformations like quadruplexes. Maintenance of the quadruplex structure in telomere has been regarded as a target for anticancer drug discovery, and has been investigated widely [1,2].

Loops which are linkers connecting G-strands and supporting the G-tetrad core exist in all of the intra-molecular G-quadruplexes and most of the inter-molecular G-quadruplexes. Loops can be divided into three typical families: (1) edgewise loops connecting two adjacent guanines; (2) diagonal loops connecting two opposite guanines in the G-quartet; and (3) double chain reversal loops run across the G-quadruplex grooves, from one side to another side of G-quadruplex stems.

Loops are important for biological roles of G-DNA molecules. For instance, thrombin binding aptamer (TBA) is a 13-mer DNA oligonucleotide which can fold into a G-quadruplex with two TT loops and a TGT loop. Recent researches demonstrated that TT loops participate in direct binding to thrombin, and the function of the TGT loop is to stabilize the aptamer structure [3,4]. The sequence of loops has influence to the stability of G-quadruplexes. If TTA loops of hTel are replaced by AAA, the G-quadruplex stability decreases largely [5]. Similar results also found in TBA’s TT loops, when thymine residues are substituted by adenine, the G-quadruplex is more stable or more unstable than that of wild-type base on the position of the thymine which is been replaced [6]. Besides, the loop length and loop sequence affect the folding topologies and stability of G-quadruplexes, which have been investigated by experimental and computational methods [7,8]. In some cases, the behavior of loops also affects the process of ligands binding to the G-quadruplex. Hexaoxazole-containing macrocyclic derivative ligands bind to hTel G-quadruplex inducing an abnormal increase of the system entropy [9,10]. The possible reason is that the loop base stacking structure is destroyed by the ligands binding process which inducing the increase of conformational entropy.

MD simulations have been applied in investigating G-quadruplex for a long time [11,12]. Some significant works have done in recent years. For instance, the formation and stability of G-triplex DNA [13,14], and folding pathways of hybrid type hTel [15]. In another work, the relative stabilities of varied glycosidic conformations of G-tracts were compared with free energy analysis [16]. This work is particularly important as it unveiled some folding rules of the G-quadruplex.

Structural dynamics of the loops, which connect to its biological roles, also investigated by MD simulations. MD simulations combined with simulated annealing (SA) and locally enhanced sampling methods (LES) have been used to predict the favorable topologies and conformations of TT (T2) and TTT (T3) loops [17]. It concluded that even though the favored type of loops
accords with the experimental results, but the predicted loop conformations are distinct with experimental structures. The predicted optimal structure of TTTT (T4) loop also differs from the experiments [10]. The failure is due to the mismatch of the AMBER parm99 force field [19] which has been used in these works. The AMBER parmbsc0 force field [20], which recovers the π/γ torsion from parm99 force field, has been confirmed to be valid for DNA duplex in hundreds of nanoseconds state-of-the-art molecular dynamics simulations in aqueous solution [21]. A recent research also demonstrated that the parmbsc0 provides G-quadruplexes loops conformations that are clearly closest to the experiments [22].

To our knowledge, another type of loop sequence, which contains adenine base, like TTA, has not been widely researched yet. One distinction between this type and loops which contain only thymine base (like T3 and T4 loops mentioned above) is that the former one has a chance to form TA Watson-Crick base pair, while the latter one has a chance to form TA Hoogsteen base pair. In the simulation, the TTA loop can fold into different types of topologies, including anti-parallel basket structure and some higher-order structures [28]. In the present work, the AMBER parmbsc0 force field was used to conduct NMR experiments [22].

Results

In the present work, we chose anti-parallel structure (PDB code, 143D) and [3+1] hybrid structure (PDB code, 2GKU) for simulations (Figure 1). The first one has the sequence d[T2G3(T1AG2)], and another one has the sequence d[T12G5(T13AG13)A1]. These two structures contain all the three typical types of loop conformations. Simulations and abbreviations are listed in Table 1.

In order to identify the importance of time scale on the structural dynamics of these two G-quadruplexes, the pairwise root mean square deviation (RMSD) was assessed (Figure 2). One principal result is that in both structures, the conformational fluctuation has the time scale of hundreds nanoseconds. Besides, the two structures have different forms of fluctuations. In the anti-paralle<ref>l structure, the initial conformation maintained in the first 100 ns, and then two other conformations each maintained 200 ns, another conformation maintained in the last 400 ns. In contrast, the hybrid structure settled down to a conformation after 200 ns. These primary results indicate that in order to characterize the structural dynamics of G-quadruplex loops, the simulation time should be as long as several hundreds nanoseconds. It coincides with a recent work [23]. In addition, the RMSD of all G-stems were in the range of 1 Å with tiny fluctuations (Figure 8), which means that the G-stem of all models were stable and maintained their initial structure in the simulations. It accords with other works on similar G-quadruplexes in dozens of nanoseconds under the same force field [22,29]. So the conformation change mainly happened in loops, and more detailed discussion about loops was given in the next subsections.

Behavior of Loops in Anti-parallel Structure

### Description of the Structure

There are two edgewise loops (T5T6A7 and T17T18A19) and a diagonal loop (T11T12A13) in the anti-parallel hTel G-quadruplex. In the crystal structure, A7 and A13 planes are approximately parallel to the adjacent G-quartet plane with A13 stacking over the adjacent G2 and A17 stacking over the cross strand G22. A10 is parallel to the adjacent quartet and cross-strand stacks over G16, while T3 is parallel to and stacks over non-adjacent G4. The A7 base exhibits a pronounced tilt relative to the adjacent quartet plane and is the only adenine that does not stack with a guanine in this G-quadruplex [24].

### Behavior of the Upper Diagonal Loop T11T12A13

There is a diagonal loop T11T12A13 and a single terminal base A11 on the upper side of anti-parallel structure. In the simulation, two adenine bases A1 and A13 were stacking with adjacent G-quartet with considerable stability (distance and angle between adenine base

---

### Table 1. List of simulations involved in this work.

| PDB ID | ions type | force field | time(ns) | abbreviation |
|--------|-----------|-------------|----------|--------------|
| 143D   | Na⁺       | parm99      | 1000     | anti_99     |
| 2GKU   | K⁺        | parm99      | 1000     | hybrid_99   |
| 143D   | Na⁺       | parmbsc0    | 1000     | anti_bsc0   |
| 2GKU   | K⁺        | parmbsc0    | 1000     | hybrid_bsc0|
| 143D   | Na⁺       | parmbsc0    | 5        | anti_stem⁸  |
| 2GKU   | K⁺        | parmbsc0    | 5        | hybrid_stem⁸|

*In these two structures, loop bases are deleted and just G-quartets left.

---

*Figure 1. Scheme of two hTel G-quadruplex structures involved in this work. Syn and anti glycosidic bond orientations are drawn in white and gray. From layer 1 to layer 3 are three G-quartets named from the origin of the first strand of G-quadruplexes.*

doi:10.1371/journal.pone.0071380.g001
and adjacent G-quartet are listed in Table 2. By monitoring the distance and angle between two thymine bases T11 and T12, we found that they stacked with each other in the trajectory (Figure 3a), and also found that this stacking conformation only maintained in anti_bsc0, but did not stable enough to maintain in anti_99 (Figure 9a). Furthermore, these two residues also had a considerable fluctuation as a whole, and this fluctuation led to the deviation of the adjacent quartet bigger than other two quartets (this loop conformation is named ANTI_U, Figure 4).

Behavior of the Lower Edgewise Loops T5-T6A7 and T17-T18A19. There are two edgewise loops T5-T6A7 and T17-T18A19 on the lower side of anti-parallel structure. Among them, A canonical Watson-Crick DNA base pair formed with two hydrogen bonds between T6 and A19. This base pair was very stable, and the occupied of both hydrogen bonds were more than 0.95 (Table 3). In addition to this, T6 and A19 were stacking with the lower layer of G-quartet. The small fluctuation of distance and angle between T6/A19 and adjacent G-quartet manifests that this stacking structure is very stable (Table 2).

In the NMR structure, A7 is the only adenine does not stack with adjacent G-quartet. However, it was stacking with the adjacent G-quartet after the energy minimization step of the

| Table 2. Distance and angle between loop bases and their reference base groups. |

| Model  | Group 1     | Group 2     | Distance(Å) | Angle(degree) |
|--------|-------------|-------------|-------------|---------------|
| anti_bsc0 | A1          | layer 1     | 3.2±0.3     | 21.9±9.3     |
|         | A13         | layer 1     | 2.8±0.7     | 22.9±11.4    |
|         | T6          | layer 3     | 2.8±0.5     | 14.8±7.2     |
|         | A19         | layer 3     | 3.1±0.4     | 14.6±6.3     |
| hybrid_bsc0 | T1          | layer 1     | 2.6±0.7     | 13.7±5.8     |
|          | A20         | layer 1     | 3.0±0.3     | 9.0±5.0      |
|          | T13         | layer 3     | 2.5±0.5     | 20.3±9.0     |
|          | A24         | layer 3     | 2.7±0.4     | 16.1±7.5     |
|          | A14         | T13/A24     | 4.8±0.7     | 27.4±10.7    |

Results under parm99 force field are given in Table S1.

*a*The adjacent G-quartet or adjacent base pair are chosen as the reference base group.

doi:10.1371/journal.pone.0071380.t002

Figure 3. Structural dynamics of anti_bsc0 loops. (a), distance and angle between two thymine bases T11 and T12 of the upper diagonal loop. The small fluctuation reflects that these two thymine bases stacked with each other in the trajectory. (b), distance and angle between A7 and upper G-quartet, in the first 250 ns, A7 stacked with the G-quartet, but then this stacking structure was destroyed. (c), angle between T17 and T18 and angle between T18 and A19.

doi:10.1371/journal.pone.0071380.g003
simulations (the conformation is named ANTI_L2, Figure 4). This stacking conformation maintained hundreds of nanoseconds (100 ns in anti_99 and 250 ns in anti_bsc0, Figure 3b and Figure 9b), and then it was destroyed and formed a conformation (named ANTI_L1, Figure 4) similar to the NMR structure. The angle between A7 and adjacent G-quartet was 40±17 degrees in anti_bsc0 and 42±18 degrees in anti_99. A hydrogen bond formed between O4 of A7 and N2 of G8 to help stabilizing this conformation (occupied of the hydrogen bond is about 0.91, Table 3). It reflects that the loop conformation is the competition of diverse force, like stacking interaction and hydrogen bond. MM/PBSA results indicated that ANTI_L1 is the more stable conformation with free energy 18 kcal/mol lower in anti_basc0 (Table 4).

In T_17T_18A_19 loop, A_19 was stacking with the adjacent G-quartet, T_17 was vertical to the adjacent G-quartet and stabilized by a hydrogen bond between O4 of T_17 and N2 of G_16, and T_18 had two conformations, one is stacking with T_17 (named ANTI_L3, Figure 4) and another is stacking with A_19 (named ANTI_L4, Figure 4). In most of the time, T_18 was stacking with

### Table 3. Details of hydrogen bonds between loop bases and between loop base and quartet base.

| Model    | Donors | Acceptors | Occupied(%) | lifetime(ps) |
|----------|--------|-----------|-------------|--------------|
| anti_bsc0| N6@A19 | O4@T6     | 98.3        | 322.5        |
|          | N3@T6  | N1@A19    | 98.4        | 441.5        |
|          | N2@G8  | O4@A5     | 91.4*       | 101.0        |
| hybrid_bsc0| N3@T13 | N1@A24    | 99.2        | 1229.5       |
|          | N6@A24 | O2@T13    | 99.0        | 659.8        |
|          | N3@T1  | N1@A20    | 99.9        | 6989.5       |
|          | N6@A20 | O4@T1     | 99.4        | 931.0        |
|          | N2@G4  | O4@T6     | 79.2        | 96.0         |

Results under parm99 force field are given in Table S2. *This hydrogen bond formed after 250 ns in the simulation, so just last 700 ns of the trajectory was used to count the hydrogen bond.

doi:10.1371/journal.pone.0071380.t003
A19, and T18 stacking with T17 was found mainly from 100 ns to 200 ns and from 500 ns to 600 ns (Figure 3c). MM/PBSA analysis manifested that two conformations have similar free energies, the free energy difference is just 6 kcal/mol in anti_bsc0 (Table 4). Furthermore, it indicates that ANTI_L4 is more stable than ANTI_L3, which means that T18 inclines to stack with A20. It accords with the result of our observation (T18/A19 stacking conformation was found in most of the snapshots, 91% in anti_bsc0 and 80% in anti_99). Analyzing dihedral angles of the backbone indicated that the conformational switch of T18 connects to its $x_2/g_2$ torsion. The $x_2/g_2$ switched from $g_2/t$ to $g_2/g_2$ when T18 turn to stacking with T17 in anti_99, but changed from $t/g_2$ to $g_2/g_2$ in anti_bsc0 as $g_2/t$ is penalized in parm99.

Among these loop bases, only T2 had the biggest fluctuation in the simulation, even though T5 formed some weak hydrogen bonds with neighbor backbone atoms.

Behavior of Loops in [3+1] Hybrid Structure

Description of the Structure.

The [3+1] hybrid hTel G-quadruplex contains one double chain reversal loop (T2-T3-A8) and two edgewise loops (T12-T13-A14 on the top and T14-T18-A20 on the bottom of the G-tetrad core). The first thymine and the adenine of the double chain reversal loop have the similar positions to the bottom of the G-tetrad core. The first thymine and the adenine of

The second thymine has the different the double chain reversal loop have the similar positions to the bottom of the G-tetrad core). The first thymine and the adenine of

Furthermore, it indicates that ANTI_L4 is more stable than ANTI_L3, which means that T18 inclines to stack with A20. It accords with the result of our observation (T18/A19 stacking conformation was found in most of the snapshots, 91% in anti_bsc0 and 80% in anti_99). Analyzing dihedral angles of the backbone indicated that the conformational switch of T18 connects to its $x_2/g_2$ torsion. The $x_2/g_2$ switched from $g_2/t$ to $g_2/g_2$ when T18 turn to stacking with T17 in anti_99, but changed from $t/g_2$ to $g_2/g_2$ in anti_bsc0 as $g_2/t$ is penalized in parm99.

Among these loop bases, only T2 had the biggest fluctuation in the simulation, even though T5 formed some weak hydrogen bonds with neighbor backbone atoms.

Behavior of Loops in [3+1] Hybrid Structure

Description of the Structure.

The [3+1] hybrid hTel G-quadruplex contains one double chain reversal loop (T2-T3-A8) and two edgewise loops (T12-T13-A14 on the top and T14-T18-A20 on the bottom of the G-tetrad core). The first thymine and the adenine of the double chain reversal loop have the similar positions to the crystal parallel structure, but the second thymine has the different conformation. In addition, there are Watson-Crick base pair T1/A20 and reversed Watson-Crick base pair T13/A24 stacks on two sides of the G-tetrad [26].

Behavior of the Double Chain Reversal Loop T2-T3-A8.

The double chain reversal loop was flexible and did not maintain a single stable structure in the simulation. In detail, there was a hydrogen bond formed between O4 of the first thymine and middle quartet of G-stem (G4). It is the same with which found in previous simulations for hTel parallel G-quadruplex structure [22]. This hydrogen bond maintained in hybrid_bsc0 in most of the time as its occupied is 79%, but it just existed in 43% of the time in hybrid_99 (Table 3). In addition to this, a particular base stacking conformation was found in both models, in which the second thymine was stacking with the first thymine (named HYBRID_P, Figure 4). This conformation was not found in the relevant work of hTel parallel topology [22]. We noted that this TT stacking conformation was not stable and destroyed several times (Figure 5a and Figure 10a). The adenine, which had no special interaction with other bases, did not have a certain conformation in the simulation. The RMSD of this loop was calculated by choosing the experimental structure as the reference (Figure 5b). The value is about 1 Å with tiny fluctuation for hybrid_bsc0, but large fluctuation was found in hybrid_99. It means that the parm99 force field does not have the ability to keep a relatively stable double chain reversal TTAA loop.

Behavior of Edgewise Loops T12-T13-A14 and T14-T18-A20.

There is a edgewise loop T12-T13-A14 on the lower side of [3+1] hybrid structure, its stacking type was really similar to the lower side (T17-T18-A19 loop) of anti-parallel structure. Firstly, the canonical Watson-Crick base pair T1/A20 which stacking with the adjacent G-quartet maintained in trajectories (distance and angle between this base pair and G-quartet are listed in Table 2). This TA base pair is stable like the one in anti-parallel structure (the occupied of hydrogen bonds is listed in Table 3). Then, a hydrogen bond between O4 of T18 and N2 of G7 formed and stabilized T18, and T19 was vertical to the adjacent G-quartet. And then, the conformations of T18-T19 loop were found, in one, T19 stacked with T18 (named HYBRID_U1, Figure 4), and in another, T19 stacked with A20 (named HYBRID_U2, Figure 4). In hybrid_bsc0, the T18/T19 stacking structure existed in 98% of the time, and no long time T18/A20 stacking conformation found in the trajectory (Figure 5c). But in hybrid_99, the T18/T19 stacking structure was found in just 25% of the snapshots. Both conformations last dozens of nanoseconds, and conformational switch happened several times in hybrid_99 (Figure S3b). MM/PBSA results indicated that HYBRID_U1 is more stable with free energy 36 kcal/mol lower in hybrid_bsc0 (Table 4). It means that T19 inclines to stack with T18 but not A20, which is different from the anti-parallel T17-T18-A19 loop.

There is a edgewise loop T12-T13-A14 on the lower side of [3+1] hybrid structure. In this loop, the reversed Watson-Crick base pair T13-A24 stably stacked with adjacent G-quartet (Table 2 and Table 3), and A14 stacked to the adjacent T11/A24 base pair (the conformation named HYBRID_L, Figure 4). The distance and angle and their fluctuations between A14 and T13/A24 base pair are much bigger than the bases which stacked to the adjacent G-quartet. It means that, this stacking conformation is looser. The last loop base, T12 was very flexible, whose behavior likes the T5 in anti-parallel structure.

Behavior of G-stems

Generally, G-stems are rigid part of G-quadruplexes, and G-quartets which four guanine bases dwell in is a plane [30,31]. But the G-quartet in G-quadruplexes is usually distorted due to the influence of loops and also the dynamical fluctuation of themselves. So, we draw more attention on the distortion and fluctuation of G-quartets here. The parameter RMSD in the Z-axis (RMSDz) was used to measure the G-quartet distort away from the plane conformation. RMSDz is given by.

$$RMSD_z(t) = \sqrt{\frac{1}{N_i} \sum_{i} (z_i(t) - z_{zi}(t))^2}$$

where $i$ is the base number, $j$ is the atom number, $z_{ij}$ is the z component of the coordinate of atom $ij$, and $N_i$ is the number of atoms for base $i$.

### Table 4. Comparison of the absolute free energies* of the different conformations of the loops.

| Model        | Conformations | $\Delta E_{\text{int}}$ | $\Delta E_{\text{elec}}$ | $\Delta E_{\text{vdw}}$ | $\Delta E_{\text{sp}}$ | $\Delta E_{\text{opp}}$ | $\Delta PB$ | $\Delta TS$ | $\Delta G$ |
|--------------|---------------|------------------------|------------------------|------------------------|------------------------|------------------------|-----------|-----------|---------|
| anti_bsc0    | ANTI_L1-ANTI_L2 | -9                     | 85                     | -91                    | 0                      | -19                    | -1        | -18       |
| anti_99      | ANTI_L4-ANTI_L3 | -7                     | -14                    | 2                      | 17                     | 0                      | -2        | 4         | -6      |
| hybrid_bsc0  | HYBRID_L1-HYBRID_L2 | 13                   | -90                    | -12                    | 57                     | 1                      | -32       | 4         | -36     |

Results under parm99 force field are given in Table S3.

* $\Delta E_{\text{int}}$, internal energy; $\Delta E_{\text{elec}}$, Coulombic energy; $\Delta E_{\text{vdw}}$, van der Waals energy; $\Delta E_{\text{sp}}$, polar solvation energy; $\Delta E_{\text{opp}}$, nonpolar solvation energy; $\Delta PB$, enthalpy; $\Delta TS$, solute entropy; $\Delta G$, absolute free energy. All are in kcal/mol.

doi:10.1371/journal.pone.0071380.t004

PLOS ONE | www.plosone.org 5 August 2013 | Volume 8 | Issue 8 | e71380
The probability distribution of RMSD for G-quartets in anti_bsc0 and hybrid_bsc0 are given in Figure 6. As the RMSD results under parm99 force field are exactly similar to ones in parmbsc0 force field, they are not shown here. What we care about mostly are the peak point and the width of the distribution. Because the peak point represents the distortion of G-quartets, and the width represents the fluctuation of G-quartets. In this part, we care about whether the flexibility of loops contributes to the distortion of G-quartets. In order to prove this, we ran another two simulations in which loop bases are deleted (named anti_stem and hybrid_stem, Table 1). These two simulations just last 5 ns as it’s enough for G-stems. We drew the probability distribution of RMSD for G-quartets of these two models with dash line in Figure 6. In this figure, we found that the relative distortions and fluctuations are layer1 > layer2 > layer3 for anti_bsc0, and layer1 < layer2 < layer3 for hybrid_bsc0. By comparing it with the results of anti_stem and hybrid_stem, we can give some explanation. In anti-parallel structure, the flexible diagonal T12T13 loop led to the deviation of layer1 bigger than the one in anti_stem. On the other side, the stable T6/A19 base pair help stabilizing layer3, that it has smaller distortion than the one in anti_stem. In hybrid structure, the stable T1/A20 base pair help stabilizing layer1, as it has smaller distortion than in hybrid_stem. On the other side, bigger distortion of layer3 in hybrid_bsc0 comes from the flexibility of edgewise loop T12T13A14, even though there is a T13/A24 base pair.

One thing make us amazing is that the center layer is not the most stable one, whether in anti_bsc0 or in hybrid_bsc0. It’s contrary to the general understanding that central G-quartets which are away from the influence of loops and solvent are more stable than ones in the terminal. In order to explain it, we checked the G-quartets in anti_stem and hybrid_stem. In these two structures, the central one is the most stable one as it is away from the influence of the solvent. So, a brief explanation is that this abnormal phenomenon is the result of the dynamical behavior of loops. In detail, there are two reasons, one is layer3 of anti-parallel structure and layer1 of hybrid structure become more stable by the help of TA base pair, another is layer1 of anti-parallel structure and layer3 of hybrid structure have bigger distortion which comes from the flexibility of adjacent loops, and the central layer is influenced by them. Among this, we can get a conclusion that loops can make the adjacent G-quartet more stable or less stable, base on the behavior of adjacent loops.

Discussion

Human telomeric G-quadruplex was widely researched by MD simulations, but few people paid attention to the dynamical behavior of loop conformations as it is complicated and changeful. In this work, by the help of the improving of the force field by other groups [20], and also by the help of the new method created to monitor the base stacking conformation in this work, after extended the trajectories to 1 ms, we now have a new insight into the stability and dynamical behaviors of hTel G-quadruplexes TTA loops.

There are distinct stable part and flexible part for edgewise loops and diagonal loops. The stable part, mainly the TA Watson-Crick base pair and most of the bases stacking with G-quartets, is stable enough to maintain in the simulation. Other parts, especially TT base stacking conformations and TA base stacking conformations, are not stable enough to maintain in the simulation, and spontaneous conformational switches happened in our 1 ms trajectories. It indicates that there are multiple loop
conformations with similar energies for these structures. In order to prove this idea, we checked the NMR structure of anti-parallel structure and hybrid structure, and found that in some snapshots of anti-parallel structure, T18 stacked with A19, but in other snapshots T18 flipped into the bulk solvent. Even though no T17/T18 stacking conformation found in these snapshots, it is very likely to form by the help of hydrophobic interaction and π-π stacking interaction of these two thymine bases. Furthermore, similar free energies also indicate that these two stacking conformations maybe coexist in the solution. There are more details in MM/PSA results (Table 4), first is that the electrostatic energy (∆E_{elec}) and the polar solvation energy (∆E_{polar}) have big deviations with opposite direction. Even though the switch of middle thymine (T18 in anti-parallel structure and T19 in hybrid structure) also accompanies with the torsion of backbone, which shows in the variation of internal energy (∆E_{int}), this change is relatively small than ∆E_{elec} and ∆E_{polar}. It means that even though the conformation of loops is the result of the balance of these competing forces, but the base stacking conformation is mainly determined by the competition of electrostatic interaction and polar solvation interaction.

The TA Watson-Crick base pair is a unique feature of TTA edgewise loops. All three TA base pairs involved in this work (T6A19 in anti-parallel structure and T1A20, T13A14 in [3+1] hybrid structure) are very stable. These stable base pairs make a great contribution to steady loop conformations. By the help of these hydrogen bonds interactions, the TTA edgewise loops maintain more stable and simple conformations than T3 and T4. As the TTA edgewise loops mainly maintained their experimental conformation and only single thymine base was flapping in the hundreds of nanoseconds simulation, while the experimental conformation of T3 and T4 were almost always lost and diverse conformations were found in previous simulations [18,22]. And also, it’s understandable that the TTA loops have more stable conformation thanAAA loops, and TTA replaced with AAA led to the melting temperature of the G-quadruplex decreased [5].

Double chain reversal loops did not maintain a single stable conformation in the simulation. It is similar to the relevant work of the parallel structure of hTel [22]. In our opinion, there are two reasons for this unstability. On one hand, loop conformation is the balance of multiple interactions, and this balance is extremely hard to reach under the ability of present force fields for double chain reversal loops [22]. On the other hand, indeed the double chain reversal loop has multiple conformations with similar free energies. So, this loop will not maintain a certain conformation in the simulation. This view can be proofed by that the initial structure of the double chain reversal loop, which is given by the NMR experiment, also don’t has a certain conformation [26]. All these results reflect the unstability of double chain reversal loops.

Besides, we should note that the stability of loop conformations discussed above does not directly relate to the relative stabilities of these two G-quadruplexes topologies. Actually, due to the different sequences and sequence lengths of hybrid structure (dT2G1T3AG3A1) and anti-parallel structure (dG1A2T3AG3A1), we think it is not possible to compare the relative stabilities of these two structures base on the results of MM-PBSA and MD simulations directly.

Methods

Molecular Dynamics Simulations

The AMBER force field parm99c0 was used in this work as it works well for simulating nucleic acids [29]. Besides, the parm99 force field was chosen to make a comparison. As we realized the limitation of these force fields [22], we did simulations carefully and the results were compared to the experimental structures to make sure that they are valid.

The models were built by lea program in AMBERTOOLS 1.5 software package [32]. The acpype program [33] was used to convert the AMBER file formats to GROMACS file formats, and single precision GROMACS 4.5 software package [34] was used to run the simulations.

Explicit solvent simulations were performed at 300 K under the control of Berendsen thermostat with time constant 1 ps [35]. Isotropic constant-pressure boundary condition under the control of the Berendsen algorithm of pressure coupling with time constant 1 ps was used for NPT ensemble simulations. The Particle mesh Ewald (PME) method was used for calculating electrostatic interactions [36]. A 10 Å cut-off was applied to Lennard-Jones interactions. A cubic box of TIP3P water molecules were added around the DNA to a depth of 12 Å on each side of the solute [37]. The system was neutralized with the addition of potassium ions or sodium ions, and two of them were placed between the three G-quartets layers. Standard AMBER parm94 ion parameters were used here. Covalent bonds contain hydrogen atoms were constrained using the SIEHEAK algorithm [38]. The time step for integration in all simulations was 2 fs. Coordinates were written to trajectory files every 5 ps. Periodic boundary conditions were applied to avoid surface effects.

An energy minimization of 1000 steps using the steepest descent algorithm was followed by a 20 ps position-constrained MD simulation in order to equilibrate water and ions, and then, 200 ps additional simulation was add under NPT ensemble. At last, production simulations were run under NVT ensemble.

Free Energy Calculation

The Molecular Mechanic/Poisson-Boltzmann Surface Area (MM/PBSA) method was chosen to calculate the free energy [39]. The solvation free energy was computed as sums of polar and nonpolar contributions using APBS software [40]. The MM energy parts were calculated in GROMACS. The solute entropic part was calculated by normal mode analysis, and mdrun and g_mems in GROMACS were used to calculate eigenvalues. All water molecules and cations (except the two locate in the center of the G-quadruplex) were removed from the trajectories. Standard AMBER parm94 ions radius were used here. The trajectories were examined in 50 ps intervals. Values used for the dielectric constant were 1 for solute and 80 for the surrounding solvent. The radius of the solvent was set to 1.4 Å.

Conformational Analysis

A G-quartet consisting of four guanines shown in Figure 7a usually does not stay in the same geometrical plane during the dynamic processes due to deformation of the G-quadruplex. In order to calculate the distance and angle between G-quartet and loop base, a center point Oq and a Z-axis Zq are used to represent the G-quartet in our model, and the algorithms are established based on the strategy applied to the description of the double stranded DNA base and base pairs. Details are given below.

The coordinates of the base are described by the base-triad base on El Hassan’s work [41]. Then by using a least squares fitting procedure, the real base is fit to a standard reference base, the center of bases, O1, O2, O3, and O4, and Z-axes, Z1, Z2, Z3 and Z4 which perpendicular to bases (for a standard base, all the non-hydrogen atoms are in the XY plane) can be calculated. Figure 7a is an illustration of the base-triad for G-quartet, it was found that four base-triads have similar Z-axis but different X-axis and Y-axis. In the following, the four Z-axes, Z1, Z2, Z3 and Z4, are associated to create a new Z-axis, Zq.
First of all, we give an example to show the procedure for creating a new Z-axis, \( Z_{mst} \), via association of two unparallel Z-axes, say \( Z_1 \) and \( Z_2 \). The angle \( \Gamma \) between two axes \( Z_1 \) and \( Z_2 \) can be calculated using the formula below:

\[
\Gamma = \cos^{-1}(Z_1 \cdot Z_2)
\]  

Two bases with similar Z-axis can be associated to create a \( Z_{mst} \) vector by rotate the base 1 by \( \Gamma/2 \) about the \( rt \) axis:

\[
Z_{mst} = R_{rt}(\Gamma/2) \cdot Z_1
\]  

\( rt \) is the rotate axis determined by cross product two Z-axes:

\[
rt = Z_1 \times Z_2
\]  

This algorithm refers to Lu’s work [42] and the \( Z_{mst} \) represents the Z-axis of the two bases (Figure 7b).

The \( Z_g \) for G-quartet is created in the same principle. The four guanine bases in quartet are divided into two parts, each \( Z_{mst} \) is created by rotating the first base by \( \Gamma/2 \) about the \( rt \) axis in each part. In this way it will produce two values of \( Z_{mst} \). Finally, two generated \( Z_{mst} \) are associated together to create a new Z-axis, \( Z_g \), by rotating the base corresponding to one of two \( Z_{mst} \).

Centers of four guanine bases are associated to create a new center \( O_g \) which represents the center of the G-quartet as well.

\[
O_g = \frac{1}{4} \sum_{i=1}^{4} O_i
\]  

Now, the angle \( \theta \) and distance in Z-axis \( d_z \) between two G-quartets or between G-quartet and loop base can be calculated directly after we get the Z-axis and center of two base groups \( Z_{g1} \), \( Z_{g2} \), and \( O_{g1} \), \( O_{g2} \).

Figure 7. Illustration of the G-quartet base-triads. (a), G-quartet base-triads. All four bases have similar Z-axis but distinct X-axis and Y-axis. (b), Calculating for \( Z_{mst} \) by rotating the \( Z_1 \) by \( \Gamma/2 \) about the \( rt \) axis, and \( \Gamma \) is the angle between two Z-axes \( Z_1 \) and \( Z_2 \).

doi:10.1371/journal.pone.0071380.g007
$$Z_2, \ O_{g1}, \ O_{g2}.$$ 

$$\theta = \cos^{-1}(Z_{g1} \cdot Z_{g2}) \tag{7}$$

$\theta$ is the angle between two plane base groups. The two Z-axes are combined to create a $Z_{gmst}$ axis:

$$Z_{gmst} = R_{\theta}(\frac{h}{2}) \cdot Z_{g1} \tag{8}$$

Then we can get the distance of $O_{g1}, \ O_{g2}$ in $Z_{gmst}$ axis $d_z$, it’s the distance in Z-axis for two plane base groups.

$$d_z = (O_{g2} - O_{g1}) \cdot Z_{gmst} \tag{9}$$

In this work, we used the distance and angle (and also their fluctuation) between two bases, or between base and adjacent G-quartet to monitor the stability of stacking structures. We assumed that if the distance maintains in the range of 3 Å to 4 Å, and the angle maintains in the range of 0 to 30 degrees, then it forms a stacking structure.

MDAnalysis software [43] and some in-house programs were used for data analysis. Standard reference data of nucleic bases came from Olson’s work [44]. The base fitting algorithms in 3DNA [45] have been used in our methods. $\phi$,$\psi$ in AmberTools 1.5 was used to calculate the lifetime and occupied of hydrogen bonds, and the cut-off of distance and angle are 3.5 Å and 30 degrees.

Supporting Information

Figure S1 RMSD of G-stems for all models. (TIF)

References

1. Neidle S, Parkinson G (2002) Telomere maintenance as a target for anticancer drug discovery. Nature Reviews Drug Discovery 1: 383–393.

2. Huppert JL (2008) Four-stranded nucleic acids: structure, function and targeting of G-quadruplexes. Chemical Society Reviews 37: 1375–1384.

3. Red’chenkov R, Golovin A, Spurio-Vanossa V, Kopylov A, Sponer J (2010) Structural dynamics of thymin-binding DNA aptamer digctgogctgtgctggtgctg DNA quadruplex studied by large scale explicit solvent simulations. Journal of Chemical Theory and Computation 6: 3005–3014.

4. Olsen CM, Lee HT, Marky LA (2009) Unfolding thermodynamics of intramolecular DNA quadruplexes: Base sequence contributions of the loops. The Journal of Physical Chemistry B 113: 2587–2595.

5. Fox KR, Risitano A (2003) Stability of intramolecular DNA quadruplexes: Comparison with dhduplexes. Biochemistry 42: 6307–6313.

6. Nagatani S, Isono N, Tsunoto K, Sugimoto N (2011) Loop residues of thymin-binding DNA aptamer impact g-quadruplex stability and thymin binding Biochemistry 93: 1231–1238.

7. Hazel P, Huppert J, Balaban P, Sniegowski S, Neidle S (2004) Loop-length-dependent folding of G-quadruplexes. Journal of the American Chemical Society 126: 16405–16415.

8. Cang X, Sponer J, Cheatham TE III (2011) Insight into gDNA structural polymorphism and folding from sequence and loop connectivity through free energy analysis. Journal of the American Chemical Society 133: 14270–14279.

9. Barbieri CM, Semovian AR, Rzuczek SG, Rice JE, LaVoie EJ, et al. (2007) Defining the mode, energetics and specificity with which a macrocyclic hexaoxazole binds to human telomeric g-quadruplex DNA. Nucleic Acids Research 35: 3272–3280.

10. Pilch DS, Barbieri CM, Rzuczek SG, LaVoie EJ, Rice JE (2008) Targeting human telomeric g-quadruplex DNA with oxazole-containing macrocyclic compounds. Biochemistry 90: 1233–1249.

11. Sponer J, Spáčková N (2007) Molecular dynamics simulations and their application to four-stranded DNA. Methods 43: 278–290.

12. Sponer J, Cang X, Cheatham TE (2012) Molecular dynamics simulations of g-DNA and perspectives on the simulation of nucleic acid structures. Methods 57: 25–39.

13. Koivula D, Moshino T, Sannohe Y, Yu Z, Mao H, et al. (2012) Intramolecular folding in three tandem guanine repeats of human telomeric DNA. Chemical Communications 48: 2006–2008.

14. Limongelli V, DeTito S, Cerolfini L, Fraggi M, Pagano B, et al. (2013) The g-quadruplex DNA. Angewandte Chemie International Edition 52: 2269–2273.

15. Mashino T, Yagi H, Sannohe Y, Rajendra A, Sugiyama H (2010) Folding pathways of human telomeric type-1 and type-2 g-quadruplex structures. Journal of the American Chemical Society 132: 14910–14918.

16. Cang X, Sponer J, Cheatham TE (2011) Explaining the varied glycosidic conformational, extract length and sequence preferences for anti-parallel g-quadruplexes. Nucleic Acids Research 39: 4499–4512.

17. Neidle S, Hazel P, Parkinson GN (2006) Predictive modelling of topology and loop variations in dimeric DNA quadruplex structures. Nucleic Acids Research 34: 2117–2127.

18. Fadma E, Spáčková N, Stell R, Koca J, Cheatham TE, et al. (2004) Molecular dynamics simulations of guanine quadruplex loops: Enhances and force field limitations. Biophysical Journal 87: 227–242.

19. Wang J, Gieplak F, Kollman PA (2000) How well does a restrained electrostatic potential (RESP) model perform in calculating conformational energies of organic and biological molecules? Journal of Computational Chemistry 21: 1049–1074.

20. Orozco M, Perez A, Marchan I, Smouli D, Sponer J, et al. (2007) Refinement of the amber force field for nucleic acids: Improving the description of alpha/gamma conformers. Biophysical Journal 92: 3817–3829.

21. Perez A, Luque FJ, Orozco M (2007) Dynamics of b-DNA on the microsecond time scale. Journal of the American Chemical Society 129: 14739–14745.

22. Fadma E, Spáčková N, Szydna J, Koca J, Orozco M, et al. (2009) Single stranded loops of quadruplex DNA as key benchmark for testing nucleic acid force fields. Journal of Chemical Theory and Computation 5: 2514–2530.

23. Islam B, Sgobba M, Laughton C, Orozco M, Sponer J, et al. (2013) Conformational dynamics of the human propeller telomeric DNA quadruplex on a microsecond time scale. Nucleic Acids Research 41: 2723–2735.

24. Wang Y, Patel DJ (1993) Solution structure of the human telomeric repeat d[ag(3)(t(2)ag(3))3] g-tetraplex. Structure 1: 263–282.
25. Neidle S, Parkinson GN, Lee MPH (2002) Crystalstructure of parallel quadruplexes from human telomeric DNA. Nature 417: 876–880.

26. Phan AT, Lau KN, Kuryayyi V, Lacroix L, Patel DJ (2006) Structure of the human telomere in K+ solution: An intramolecular (3+1) g-quadruplex scaffold. Journal of the American Chemical Society 128: 9663–9670.

27. Li J, Czernia JJ, Wang L, Trent JO, Chaires JB (2005) Not so crystal clear: the structure of the human telomere g-quadruplex in solution differs from that present in a crystal. Nucleic Acids Research 33: 4649–4659.

28. Petraccone L, Spink C, Trent JO, Garberth NC, Mckinayyy CS, et al. (2011) Structure and stability of higher-order human telomeric quadruplexes. Journal of the American Chemical Society 133: 20951–20961.

29. Krepl M, Zgarbová M, Stadlbauer P, Otyepka M, Banaś P, et al. (2012) Reference simulations of noncanonical nucleic acids with different x variants of the amber force field: Quadruplex DNA, quadruplex RNA, and Z-DNA. Journal of Chemical Theory and Computation 8: 2506–2520.

30. Leszczynski J, Gu JD (2000) A remarkable alteration in the bonding pattern: An hF and dlF study of the interactions between the metal cations and the hoogsteen hydrogen-bonded g-tetrad. The Journal of Physical Chemistry A 104: 6308–6313.

31. Meyer M, Steinke T, Brandl M, Suhnel J (2001) Density functional study of guanine and uracil quartets and of guanine quartet/metal ion complexes. Journal of Computational Chemistry 22: 109–124.

32. Pearlman DA, Case DA, Caldwell JW, Ross WS, Cheatham TE, et al. (1995) Amber, a package of computer-program for applying molecular mechanics, normal-mode analysis, molecular-dynamics and free-energy calculations to simulate the structural and energetic properties of molecules. Computer Physics Communications 91: 1–41.

33. Sousa da Silva AW, Vranken WF (2012) Acpype - antechamber python parser interface. BMC research notes 5: 367.

34. Hess B, Kutzner C, van der Spoel D, Lindahl E (2008) Gromacs 4. Algorithms for highly efficient, load-balanced, and scalable molecular simulation. Journal of Chemical Theory and Computation 4: 435–447.

35. Berendsen HJC, Postma JPM, Van Gunsteren WF, Dinola A, Haak JR (1984) Molecular-dynamics with coupling to an external bath. Journal of Computational Physics 81: 54–57.

36. Darden T, York D, Pedersen L (1993) Particle mesh ewald - an n-log(n) method for ewald sums in large systems. Journal of Computational Physics 98: 100–122.

37. Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML (1983) Comparison of simple potential functions for simulating liquid water. Journal of Computational Physics 49: 120–148.

38. Ryckaert JP, Ciccotti G, Berendsen HJC (1977) Numerical-integration of cartesian equations of motion of a system with constraints - molecular-dynamics of n-alkanes. Journal of Computational Physics 23: 327–341.

39. Srinivasan J, Cheatham TE, Cieplak P, Kollman PA, Case DA (1998) Continuum solvent studies of the stability of DNA, RNA, and phosphoramide DNA helices. Journal of the American Chemical Society 120: 9401–9409.

40. Baker NA, Sept D, Joseph S, Holst MJ, McCammon JA (2001) Electrostatics of nanosystems: Application to microtubules and the ribosome. Proceedings of the National Academy of Sciences of the United States of America 98: 10037–10041.

41. El hassan MA, Calladine CR, (1995) The assessment of the geometry of dinucleotide steps in double-helical DNA - a new local calculation scheme. Journal of Molecular Biology 251: 648–664.

42. Lu XJ, ElHassan MA, Hunter CA (1997) Structure and conformation of helical nucleic acids: Analysis program (schnaap). Journal of Molecular Biology 273: 685–688.

43. Woolf TB, Michael-Agrawal N, Denning EJ, Beckstein O (2011) Software news and updates md-analysis: A toolkit for the analysis of molecular dynamics simulations. Journal of Computational Chemistry 32: 2319–2327.

44. Olson WK, Bansal M, Burley SK, Dickerson RE, Gerstein M, et al. (2001) A standard reference frame for the description of nucleic acid base-pair geometry. Journal of Molecular Biology 313: 229–237.

45. Olson WK, Lu XJ (2003) 3dna: a software package for the analysis, rebuilding and visualization of three-dimensional nucleic acid structures. Nucleic Acids Research 31: 5108–5121.