J Comput Aided Mol Des (2012) 26:289–299
DOI 10.1007/s10822-012-9548-z

Stability and free energy calculation of LNA modified quadruplex: a molecular dynamics study

Amit Kumar Chaubey · Kshatresh Dutta Dubey · Rajendra Prasad Ojha

Received: 5 July 2011 / Accepted: 23 January 2012 / Published online: 29 March 2012

Abstract Telomeric ends of chromosomes, which comprise noncoding repeat sequences of guanine-rich DNA, which are the fundamental in protecting the cell from recombination and degradation. Telomeric DNA sequences can form four stranded quadruplex structures, which are involved in the structure of telomere ends. The formation and stabilization of telomeric quadruplexes has been shown to inhibit the activity of telomerase, thus establishing telomeric DNA quadrulex as an attractive target for cancer therapeutic intervention. Molecular dynamic simulation offers the prospects of detailed description of the dynamical structure with ion and water at molecular level. In this work we have taken a oligomeric part of human telomeric DNA, d(TAGGGT) to form different monomeric quadruplex structures d(TAGGGT)$_4$. Here we report the relative stabilities of these structures under K+ ion conditions and binding interaction between the strands, as determined by molecular dynamic simulations followed by energy calculation. We have taken locked nucleic acid (LNA) in this study. The free energy molecular mechanics Poission Boltzman surface area calculations are performed for the determination of most stable complex structure between all modified structures. We calculated binding free energy for the combination of different strands as the ligand and receptor for all structures. The energetic study shows that, a mixed hybrid type quadruplex conformation in which two parallel strands are bind with other two antiparallel strands, are more stable than other conformations. The possible mechanism for the inhibition of the cancerous growth has been discussed. Such studies may be helpful for the rational drug designing.

Keywords G-quadruplex · LNA · MD simulation · Free energy calculation

Electronic supplementary material The online version of this article (doi:10.1007/s10822-012-9548-z) contains supplementary material, which is available to authorized users.

A. K. Chaubey · K. D. Dubey · R. P. Ojha (✉)
Biophysics Unit, Department of Physics, DDU Gorakhpur University, Gorakhpur 273009, Uttar Pradesh, India
e-mail: rp_ojha@yahoo.com

Introduction

Telomeric DNA occurs at the ends of eukaryotic chromosomes. It protects genome from degradation and participates in chromosomal fusion and recombination [1]. Human telomeric DNA consists of guanine-rich tandem repeats d(TTAGGG]$_n$, and in human somatic cells is ~5–8 kb in length [2]. It is in duplex form for most of this length, with the exception of the 3' terminal 150–200 nucleotides, that comprise a single stranded overhang [3]. Due to the inability of DNA polymerase to fully copy the 3' ends, telomeric DNA progressive shortens during replication [4]. The end replication problem is overcome by the expression of the enzyme telomerase, which adds hexanucleotide repeats to the 3' end of the single stranded DNA, thereby preventing shortening of telomeres. Telomeres have been shown to be activated in 80–85% of human cancer cells [5], have suggested playing a key role in maintaining the malignant phenotype by stabilizing telomer length and integrity [6].

The structure and stability of telomeres are of great research interest as they are closely related with cancer [7–10], aging [11, 12] and genetic stability [13–15]. Single stranded G-rich telomeric sequences at the end of chromosomes can fold into intramolecular quadruplex structures, a DNA secondary structure consisting of stacked

G-tetrad planes connected by a network of Hoogsteen hydrogen bonds and stabilized by monovalent cations, such as Na+ and K+. Human telomeric DNA repeats are highly conserved, which has been suggested to be related to their ability to form DNA G-quadruplexes [10–16]. The formation and stabilization of the DNA G-quadruplex in the human telomeric sequence have been shown to inhibit the activity of telomerase. Thus the telomeric DNA quadruplex has been considered to be an attractive target for cancer therapeutic intervention [7, 9, 10, 17–19].

The folding and formation of short telomeric G-quadruplex structures have been studied extensively by a variety of biophysical, structural and chemical probe methods. Structural information on the intact human telomeric DNA G-quadruplex formed under physiologically relevant conditions is necessary for structure based rational drug design. The human single stranded telomer end which can potentially fold into a number of four repeat quadruplexes, is a longer and more stable than in many other vertebrates. The potential role of quadruplexes in vivo has been highlighted with the recent development of therapeutic strategies designed to stabilize telomeric ends as G-quadruplex structures using specific small molecules, which can destabilize telomer maintenance in tumor cells [20–23].

The crystal structure of a four-repeat human telomeric sequence $d[AG_3(T_2AG_3)_3]$ crystallized from a potassium ion environment, shows that it forms an intramolecular G-quadruplex structure in which all four phosphate backbone strands are parallel to each other and the tetrads stabilized by the presence of K+ ions in the central cavity [24]. Several NMR solution studies of the folded structure of the intramolecular human telomeric G-quadruplex formed in K+ solution, demonstrate a novel hybrid type folding topology with mixed parallel-antiparallel G-strands [25]. This hybrid type G-quadruplex structure appears to be the physiologically relevant conformation of the human telomeric DNA. This folding of structure has also been independently reported by two other groups [26, 27].

The intramolecular human telomeric quadruplex can attain antiparallel basket-type, antiparallel chair-type, parallel propeller-type and mixed hybrid-type conformations, depending on the nucleotide length and ionic conditions [24, 25, 28, 29]. Literature based on NMR data shows that the Tal-22 sequence forms a well defined basket type intramolecular G-quadruplex in the presence of Na+ ions [28]. From the platinum cross-linking method, it showed that the antiparallel basket-type conformation forms in the presence of both Na+ and K+ ions [29]. Another report based on the I-radioprobing technique supported the formation of the antiparallel chair-type conformation in the presence of K+ ions and the antiparallel basket-type conformation forms in the presence of Na+ ions [30]. Shafer et al. also found the antiparallel chair-type conformation to be present under K+ ions conditions with the chemical ligation method [31]. An X-ray crystal study of Tel-22 found the parallel propeller-type conformation in the presence of K+ ions [24]. Another report based on simulations showed that Tel-22 adopts the chair-type conformation in the presence of K+ ions [32]. Ambrus et al. [25] using 1D NMR analysis showed that, in the presence of K+ ions, the Tel-22 sequence adopts two conformations, which were found to be the antiparallel and mixed hybrid-type G-quadruplexes according to CD data.

Nucleic acid is central to transmission, expression and conservation of genetic information. Consequently, high affinity binding of complimentary nucleic acids has applications in biotechnology and medicine. In this context there are the development of nucleic acid with chemical modification rendering them high affinity and stability, since unmodified DNA or RNA oligonucleotides have moderate affinities for complementary targets and low stability in biological fluids [33]. Result of such modification is Locked Nucleic Acid (LNA) molecule, where the furanose conformation is chemically locked in an RNA like (C3′-endo) conformation by introduction of a 2′-O, 4′-C methylene linkage (Fig. 1). LNA nucleotides substantially increase the thermal stability when incorporated into both DNA duplexes and triplexes and LNA has the potential for applications in drug design. With such immense potential, it is of fundamental importance to understand the structural nature of complexes formed by LNA with DNA or RNA. There have been some studies in this direction, mainly with NMR spectroscopy and X-ray crystallography, where in most cases only selected nucleotides has been replaced by LNA nucleotide in a regular RNA or DNA based duplex [34–36]. These studies have provided useful information on the nature of duplex structure with the introduction of LNA. It was demonstrated that LNA modifications stabilize the tetrameric quadruplexes of TG_3T when it is fully modified [37]. Also fully modified RNA and 2′-O-methyl-RNA quadruplexes are more stable than their unmodified $d(TG_4T)$ counterpart [38]. However, there are no structures showing the effect of LNA on 24-mer monomeric human telomeric quadruplex structures. Thus, we decided to investigate the stability and existence of LNA modified human telomeric quadruplex structures with K+ ions.

We have used molecular modeling methods to build models of G-quadruplex arrays of monomeric structures formed from extended human telomeric DNA. Since different reports have suggested variable conformation of human telomeric quadruplex in the presence of K+ ions, we used seven different types of human telomeric $d(TAGGGT)_4$ modified quadruplex conformations (Fig. 2), viz., parallel propeller type (all strands are parallel), basket-type, chair-type and mixed hybrid types, for simulation studies to understand their relative stabilities and

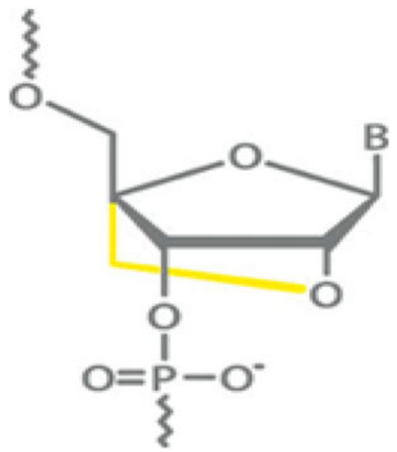

Fig. 1 Chemical structure of locked nucleic acid, in which a methylene bridge connecting the O2′ atom with C4′ atom locks the ribose ring

interaction between strands. Furthermore, to characterize the most energetically favorable quadruplex structures, we estimated the binding free energy and the entropy. As only a few earlier reports based on MD simulation of human telomeric quadruplex and the role of LNA with DNA in douplex [39, 40], triplex and quadruplex, this is the first exhaustive study providing detailed insight into the fully LNA modified different type of monomeric quadruplex structures.

Methodology

Model generation

The crystal structure [24] of the 22-mer human telomeric DNA d [AG3(T2AG3)3] (PDB code 1KF1) was used as the primary unit for the construction of the higher-order models. It contains three stacks of guanine tetrads. The extreme 5′ and 3′ thymine base were first added at start and end of the sequence respectively, to generate 24-mer d $(TAGGGT)_4$ structures. We have generated seven different types of modified structures (Fig. 2), where (LNA) was considered for modification. For modification, all the nucleotides were changed into LNA. In seven generated modified structures, four are mixed hybrid-type (MOD1, MOD2, MOD3 and MOD4), one basket-type (MOD5), one chair-type (MOD6) and one parallel propeller-type (MOD7). In MOD1 and MOD2, strand 1-2-4 are parallel and strand 3 are antiparallel (Fig. 2), in MOD3 and MOD4

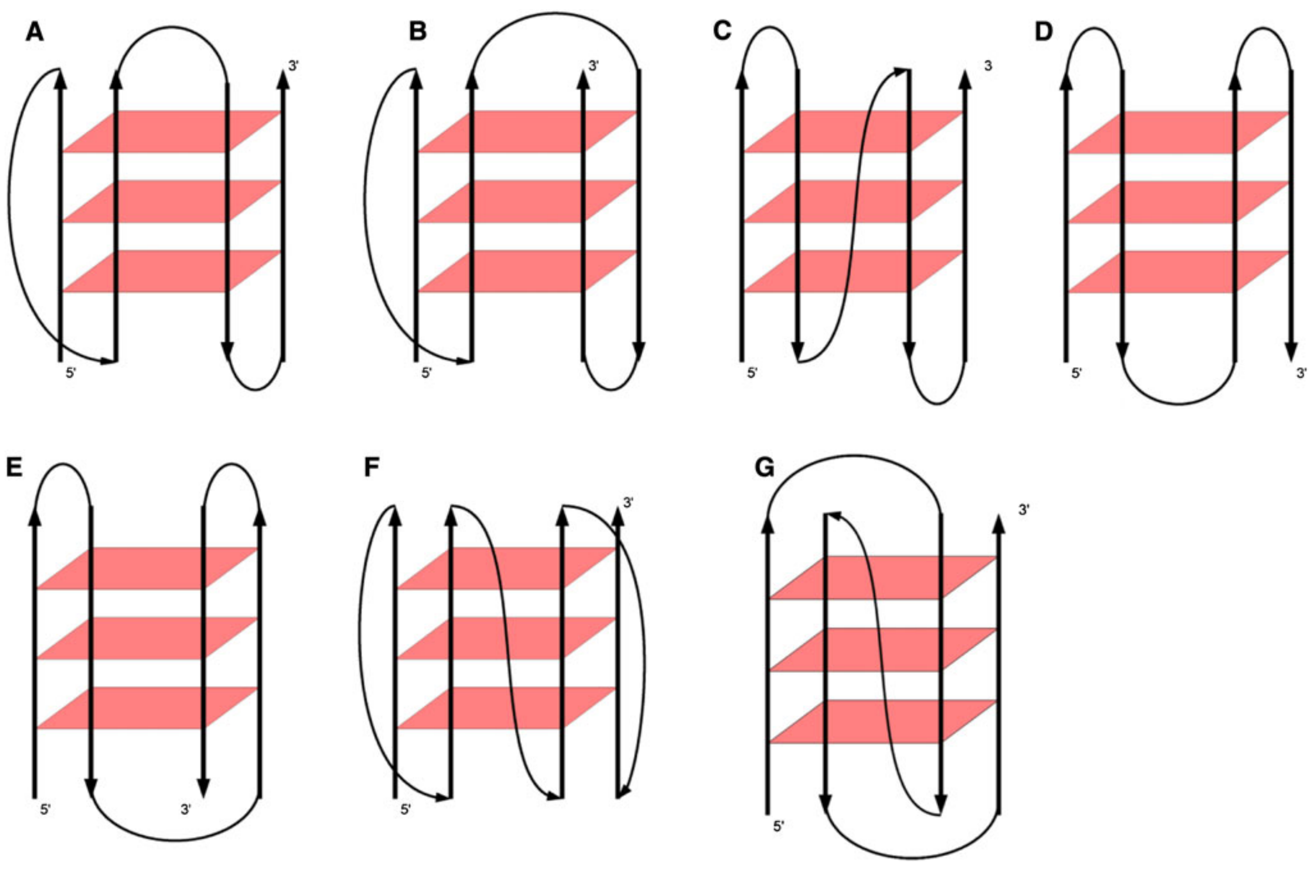

Fig. 2 Schematic representations of human telomeric quadruplex conformations: **a–d** mixed hybrid type, **e** basket type, **f** chair type, and **g** parallel propeller type

Table 1 Summary of the model names, nature of structures and production time of LNA modified 24-mer human telomeric quadruplex structures

Model	Nature and structure of modified quadruplex	Production time (ns)
MOD1	Parallel-1-2-4 antiparallel-3a (mixed hybrid-type)	20
MOD2	Parallel-1-2-4 antiparallel-3b (mixed hybrid-type)	20
MOD3	Parallel-1-4 antiparallel-2-3a (mixed hybrid-type)	20
MOD4	Parallel-1-4 antiparallel-2-3b (mixed hybrid-type)	20
MOD5	Parallel-1-3 antiparallel-2-4a (basket-type)	20
MOD6	Parallel-1-3 antiparallel-2-4b (chair-type)	20
MOD7	Parallell propeller-type	20

which is also a mixed hybrid-type structure, strand 1–4 are in parallel and strand 2–3 are in antiparallel (Fig. 2), in MOD5 and MOD6, strand 1-3 are in parallel and strand 2–4 are in antiparallel (Fig. 2) and in seventh modified structure (MOD7), all strands are in parallel (Fig. 2). Summary of the model names, their types and production time are listed in Table 1.

Force field parameter generation

Since the force field parameters for the modified nucleotides (LNA) are not available in literature, therefore, their parameters are generated using the GAUSSIAN 03 and RESP [41] program of AMBER10. All the ab-initio calculations are done at the DFT (B3LYP) level of theory with G-31G (d) basis set, using GAUSSIAN suite. Three letter codes for all the fitted nucleosides were developed to standardize the naming of the modified nucleosides in pdb files. The parameters for the modified nucleotides are available as supplementary material (Table S1).

The nomenclature of the modified nucleotides used in the text is as follows.

Modified adenine: LCA
Modified guanine: LCG
Modified thymine: TL5 and TL3

Where thymine is the ending nucleotide in all quadruplexes.

Simulation and equilibration

The X-ray structure shows a vertical alignment of consecutive K+ ions along the axis within the central core of the structure, in the middle between the G-tetrad. Thus, the ions were retained in the positions as observed in the crystal structure [24]. We placed 5 K+ ions to the central channel of all the models manually. The models were solvated in a periodic TIP3P water box [42] at least 10 Å from any solute atom. Additional positively charged K+ counter ions were included in the system to neutralize the charge on the DNA backbone. We have used recent amber force field ff03 [43–45] for parameter generation.

The simulation protocols were consistent for all of the systems. The systems were annealed from 0 to 300 K for 600 ps with continuous decrease of restrain (force constant was decreased from 100 to 5 kcal/mol $Å^2$ in six steps). The final stage of equilibration involved upto 1,200 ps runs using a force constant of 2 kcal/mol $Å^2$ on the solute and inner ions to fix them during equilibration. The final production run was carried out without any restrain on the system for 20 ns and coordinates were saved after every 10 ps for analysis of their trajectories.

All calculations were carried out with the SANDER module of AMBER10 [46]. Periodic boundary condition has been applied using the particle mesh Ewald (PME) method to treat long-range electrostatics. Hydrogen bonds were constrained using SHAKE [47]. A time step of 2 fs and a direct space non-bonded cut-off of 10 Å were used. The Langevin coupling with a collision frequency of 2.0 was used for temperature regulation. A constant pressure of 1 atm. has been attained with isotropic molecule based scaling with a relaxation time of 2 ps. The trajectories were analyzed using the PTRAJ module available in the AMBER10 and visualized by means of the CHIMERA program [48].

MM-PBSA calculation

We performed standard MM-PBSA (Molecular Mechanics Poission-Boltzman, surface area) [49] method for free energy calculation. Here, the total free energy of binding is expressed as the sum of the contribution from the gas phase and solvation energy and an additional term of solute entropy. This can be expressed by following equation

$$\Delta G_{TOT} = \Delta E_{GAS} + \Delta E_{SOLV} - T\Delta S$$

where ΔE_{GAS} is the total gas phase energy given by

$$\Delta E_{GAS} = \Delta E_{INT} + \Delta E_{VDW} + \Delta E_{ELEC}.$$

Here ΔE_{INT} corresponds to bond, angle and torsion terms in the molecular mechanical force field. ΔE_{SOLV} is the total solvation energy (polar and non polar), and $T\Delta S$ corresponds to solute entropy effect. The detail of these terms can be found in our recent publication [50]. Analysis is done for the last 6 ns (14–20 ns) trajectory of all the complexes. The snapshots for these quadruplex are extracted at intervals of 20 ps. Prior to the analysis, all water molecule and K+ ions were stripped from the trajectory. Solvation free energy is computed as the sum of polar and nonpolar contributions using a continuum solvent representation.

The polar contribution is calculated by Molsurf, implemented in AMBER10. The non-polar solvent contribution is estimated from a SASA dependent term

$\Delta E_{SNP} = \gamma.\mathrm{SASA} + \beta.$

Here γ is set to 0.0072 kcal/A^2 and β to 0. The calculation for solute entropy contribution is performed with the NMODE module in AMBER10. The snapshots were minimized in the gas phase using the conjugate gradient method for 1,000 steps, using a distance dependent dielectric of 4r (r is inter atomic distance) and with a convergence criterion of 0.1 kcal/mol Å for the energy gradient.

Results

Dynamic structure

From MD simulation of all seven modified quadruplex structures, there are series of snapshots extorted at periodic intervals, out of which only the initial structures are shown in Fig. 3. During dynamics, different snapshots at interval of 10, 15 and 20 ns of MOD2 are represented in Fig. 4, while other snapshots of MOD1 and MOD3-MOD7 for interval 10, 15 and 20 ns are shown in Figure S1. During MD simulation, hydrogen bonding pattern is convincingly well maintained. The central core of stacked tetraplex has a line of potassium ions that is coincident with the helical axis. Here the metal is bipyramidally coordinated by eight equidistant carbonyl oxygen atoms. However, potassium ions which were inside quadruplex, deviate from the symmetrical geometry. From the Fig. 4, we see that deviated ions form bipyramidal binding with O6 atom of guanine bases.

The parallel three-tetrad quadruplex remains stable with two cations in the channel. This confirms capability to achieve smooth equilibrium by exchange of cations with the bulk solvent. This phenomenon was experimentally observed on the hundreds of microsecond to millisecond time scale for the central ions [51]. The stability of the quadruplex is immediately lost when ions are absent in the channel [52–54]. In the initial stage of simulation, the ions are slightly displaced towards the outside in all structures. However, after a long simulation, two ions are bound between the plane of quartet in MOD2 (Fig. 4), which shows the stability and occurrence of this model. In all other structures, except MOD2, the ions are moved outwards during a long simulation (Figure S1).

Structural stability of the models

The root mean square deviation (rmsd) during simulation can be used as a standard of the conformational stability of a structure or a model. In this work, we intend to investigate the conformational stability of different modified G-quadruplex structures from telomeric DNA and in studying the dynamic effects of the structures. Figure 5 for MOD2 and Figure S2 for rest models compare the rmsd over the course of simulation. Each plot shows the rmsd of all-atom model (black), backbone atom only (red), base atom only (green) and middle base atom (blue). A jump in rmsd is observed during the first nanosecond which is a consequence of relaxation of the starting model. All trajectories are stabilized during simulation with relatively small fluctuation in the models. As evident from the plots of MOD1,2,4,5,6, the rmsd of all atom models are higher than that of the backbone atoms, where as in the plots of MOD3 and MOD7 the rmsd value of base atoms are higher than that of all atoms and backbone atoms. The backbone atoms are stabilized earlier during the simulation and the higher rmsd of all atoms model is a result of the wobbling effect of the nucleotide bases. All G-tetrad are held together in a coplanar array with eight hydrogen bonds which play a major role in ranking of these tetrads as most stable moieties in the models, which is shown by lowest rmsd for middle base atoms in all plots.

All models are more or less stable during the simulation except MOD7. The rmsd of MOD7 are extremely high which suggesting distortion of the structure from its starting conformation. The rmsd of backbone atoms are ~11.0 Å which also gives the rmsd value of all atoms, base atoms and middle base atoms. During the 4.7–13 ns simulation, its fluctuation remains constant with ~5.8–7.4 Å. The other more deviated structures are MOD1 and MOD3. The rmsd of MOD1 is ~8.5 Å for backbone atoms and ~9.1 Å for all atoms, which gives constant fluctuations during 14–20 ns simulation. In MOD3 the rmsd is ~8.0 Å for base atoms and ~8.3 Å for all atoms, which gives constant fluctuation during 11–19 ns simulation. In MOD4, the rmsd of all the atoms and base atoms coincide during last 6 ns run and give the rmsd of ~7.2 Å. For MOD5 and MOD6, the rmsd value is ~5.0 and ~5.4 Å for backbone atoms and for all atoms it gives ~5.5 and ~ 6.0 Å, which is constant during simulation. The rms deviation over the course of simulation for MOD1 and MOD3-7 are shown in Figure S2. Figure 5 shows the rmsd of different subsegments of MOD2. The rmsd value for all atoms is in the range of ~2.9 Å–3.3 Å, for backbone atoms it is ~2.6 Å and for base atoms it gives the deviation of ~1.8 Å. The model gives constant rmsd value of middle base atoms (~0.75 Å), during the simulation. This model gives lower rmsd value for all subsegments in

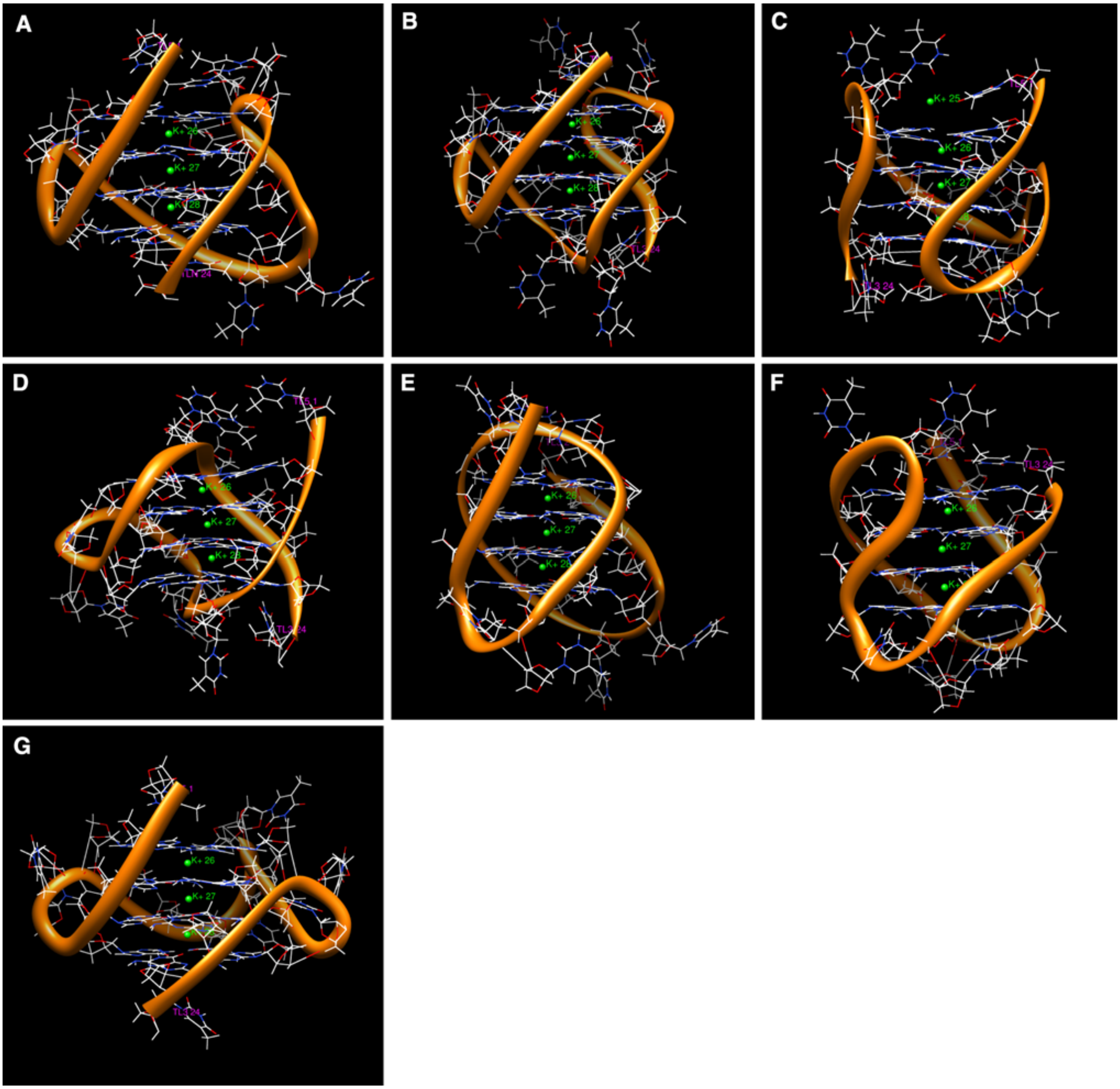

Fig. 3 Models of modified human telomeric (**a–d**) mixed hybrid type quadruplex, **e** basket type quadruplex, **f** chair type quadruplex, and **g** parallel propeller type quadruplex structure. Three K+ ion present in the central cavity are shown in *green*

comparison with all other models, which shows that MOD2 gives the highest stability, during simulation.

To identify the regions of high flexibility, per atom averaged mean square fluctuation (MSF) is reported for MOD2. Since, MOD2 exhibits a stable conformation; hence we performed β factor calculation of this model. The β factor regarding to MOD2 is represented in Fig. 6. Figure 6 reveals that atoms 575–685 show higher thermal fluctuation which corresponds to residue 17–20 of complex.

MM-PBSA calculations

The free energy calculation using MM-PBSA method can be used to study thermo chemical properties of quadruplex models and to provide a semi quantitative estimate of their stability [55, 56]. The binding energy components for all the models of modified human telomeric quadruplex structures are summarized in Table 2. Models with lower binding energies are expected to be more stable than those with higher value. The total energy of the solute (E_{GAS})

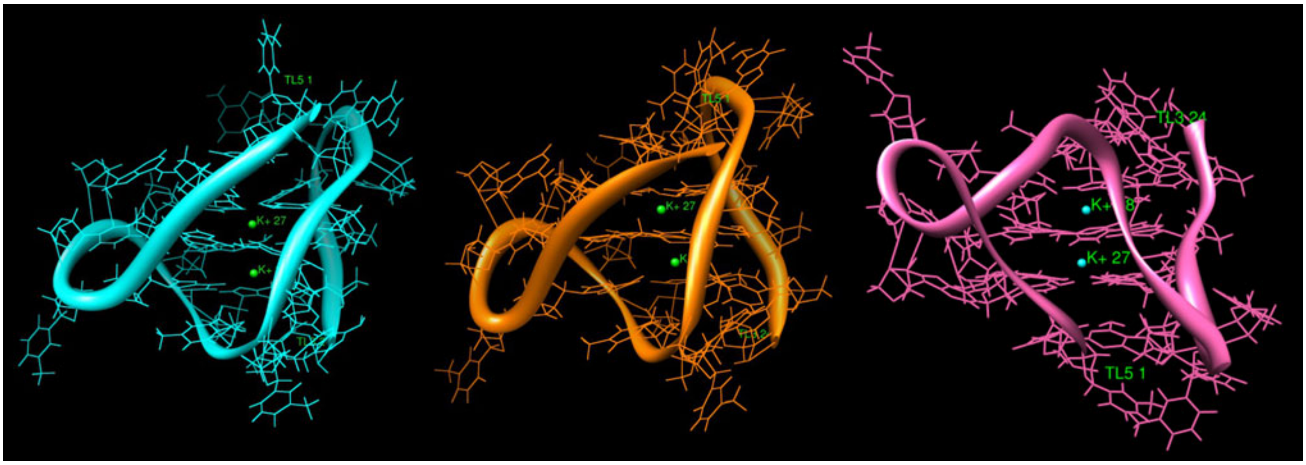

Fig. 4 Snapshots of MOD2 structure during production dynamics at 10, 15 and 20 ns (snapshot for 10 ns is in *cyan*, 15 ns is in *orange* and 20 ns is in *magenta colour*)

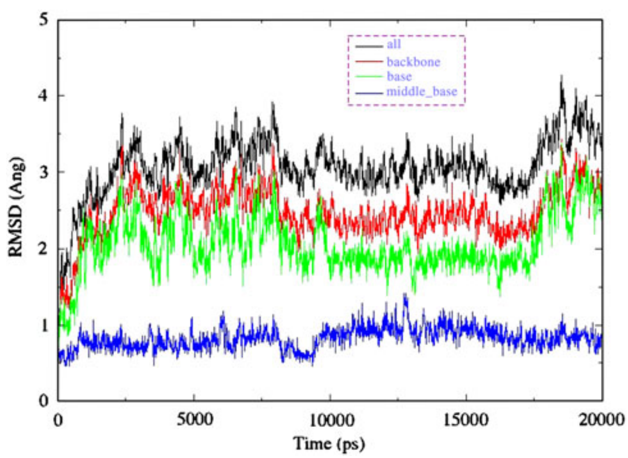

Fig. 5 RMSD plot showing the stability of the model during the MD run. RMSD values calculated for all atoms (*black*), backbone-only atoms (*red*), base only atoms (*green*) and middle base atoms (*blue*) are plotted for MOD2 structure

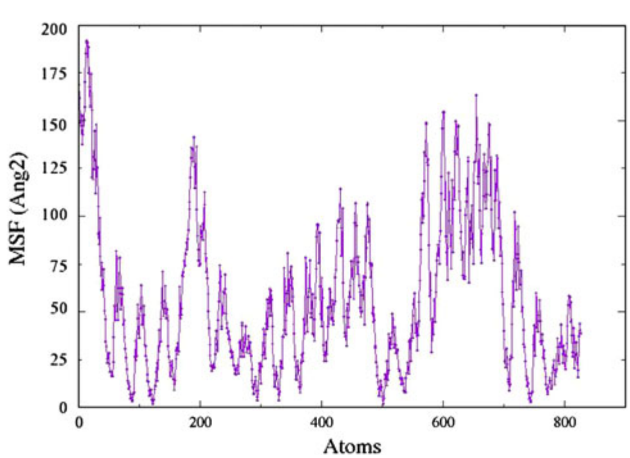

Fig. 6 All-atom mean square fluctuation versus atom number for MOD2

includes the electrostatic energy, van der Waals energy derived from a Lennard-Jones potential and the internal energy. Finite difference Poission–Bolttzman equations were used to calculate the electrostatic potential field. The binding energy calculation for the quadruplex suggested a favorable contribution for van der Waals energy (E_{VDW}) for all quadruplex models. The total binding energy (E_{TOT}) ranged between −1982.94 to −3482.13 kcal/mol for all models. These values show that MOD6, chair type structure, where Ist and IIInd strand are parallel and second and fourth are antiparallel, has the least favorable binding energy (−1,982.94 kcal/mol). Furthermore, MOD7, a parallel propeller type structure, gives the most favorable binding energy (−3,482.13 kcal/mol) relative to other models. Since all the models have sufficient negative binding energy, they show stability in their structures.

The free energy components, for all the seven quadruplex structures (where first two strands are bound with last two strands) are summarized in Table 3. Here, electrostatic energy (ΔE_{ELE}) makes an unfavorable contribution while van der Waals energy (ΔE_{VDW}) provides favorable contributions. Moreover, the nonpolar solvation energy (ΔE_{SNP}) also has favorable contributions to the total binding free energy. This implies that van der Waals and nonpolar solvation energy play an important role in binding of qudruplexes. The solvent electrostatic energy (ΔE_{SP}) also makes favorable contribution. Moreover, we observed unfavorable entropy contributions ($T\Delta S$) for all models of quadruplex. The solute entropy term of binding for all seven models ranged between −33.93 to −48.83 kcal/mol. The binding free energy (ΔG_{TOT}) of all seven structures varies from 68.09 to −4.87 kcal/mol. These values show that, MOD3, which is a mixed hybrid type structure is stable among other proposed model because it has lowest binding energy (−46.74 kcal/mol) and it has lowest

Table 2 Total Energies of all the quadruplex structures

	MOD1	MOD2	MOD3	MOD4	MOD5	MOD6	MOD7
E_{ELE}	1153.13 ± 65.15	1822.76 ± 66.66	1320.46 ± 76.38	1421.87 ± 118.48	1760.89 ± 78.17	1640.02 ± 70.83	1258.92 ± 132.32
E_{VDW}	−149.77 ± 9.13	−207.52 ± 12.33	−159.64 ± 12.31	−170.77 ± 15.76	−214.64 ± 11.01	273.14 ± 19.42	−160.32 ± 16.83
E_{INT}	2170.03 ± 22.10	2150.02 ± 27.30	2142.05 ± 23.83	2138.01 ± 15.54	2168.94 ± 19.87	3157.41 ± 27.65	2136.52 ± 19.17
E_{GAS}	3173.39 ± 63.66	3765.26 ± 60.06	3302.88 ± 73.19	3389.11 ± 115.33	3715.19 ± 75.83	5070.57 ± 70.19	3235.11 ± 111.59
E_{SNP}	42.41 ± 1.13	34.85 ± 0.82	40.30 ± 0.94	39.25 ± 1.04	34.13 ± 0.64	36.69 ± 0.80	40.24 ± 1.59
E_{SP}	−6686.28 ± 62.41	−7204.11 ± 55.19	−6818.80 ± 73.79	−6897.86 ± 110.85	−7173.57 ± 52.69	−7090.21 ± 65.71	−6757.50 ± 99.08
E_{SOLV}	−6643.87 ± 62.64	−7169.27 ± 55.66	−6778.50 ± 74.52	−6858.61 ± 117.70	−7139.44 ± 53.12	−7053.52 ± 66.06	−6717.28 ± 99.31
E_{TOT_ELE}	−5533.15 ± 14.15	−5381.36 ± 28.18	−5498.34 ± 11.4	−5475.99 ± 13.22	−5412.68 ± 32.71	−5450.19 ± 18.13	−5498.58 ± 27.57
E_{TOT}	−3470.48 ± 19.67	−3404.00 ± 28.18	−3475.63 ± 27.03	−3469.50 ± 15.32	−3424.24 ± 32.48	−1982.94 ± 19.70	−3482.15 ± 24.49

Table 3 Comparison of binding free energies of quadruplex structures

	MOD1	MOD2	MOD3	MOD4	MOD5	MOD6	MOD7
ΔE_{ELE}	2093.33 ± 37.69	2474.14 ± 48.67	2214.20 ± 45.68	2331.91 ± 78.71	2430.89 ± 35.11	2368.14 ± 49.25	2142.26 ± 73.57
ΔE_{VDW}	−67.41 ± 5.64	−65.06 ± 7.24	−68.63 ± 6.55	−75.86 ± 5.90	−92.58 ± 6.68	−56.27 ± 4.67	−43.36 ± 9.87
ΔE_{INT}	6.59 ± 1.71	4.73 ± 0.81	3.62 ± 1.28	6.28 ± 1.40	5.53 ± 1.38	20.52 ± 2.21	5.21 ± 1.55
ΔE_{GAS}	2032.50 ± 34.87	2413.80 ± 43.75	2149.18 ± 46.30	2262.33 ± 74.29	2343.84 ± 35.33	2332.39 ± 47.92	2104.11 ± 67.79
ΔE_{SNP}	−8.85 ± 0.86	−7.79 ± 0.79	−8.55 ± 0.64	−9.28 ± 0.74	−10.78 ± 0.48	−8.51 ± 0.45	−5.91 ± 1.08
ΔE_{SP}	−2067.99 ± 35.99	−2371.85 ± 43.55	−2187.37 ± 44.35	−2297.71 ± 73.91	−2365.22 ± 26.16	−2332.25 ± 48.26	−2125.45 ± 68.21
ΔE_{SOLV}	−2076.85 ± 36.10	−2379.65 ± 44.02	−2195.92 ± 44.40	−2306.99 ± 74.46	−2376.00 ± 26.25	−2340.77 ± 48.26	−2131.36 ± 68.77
Δ_{ETOT_ELE}	25.34 ± 6.36	102.28 ± 21.42	26.82 ± 8.38	34.21 ± 6.83	65.66 ± 22.60	35.89 ± 7.27	16.81 ± 14.82
$\Delta E_{GAS+SOLV}$	−44.33 ± 6.62	34.16 ± 18.75	−46.74 ± 9.21	−44.66 ± 5.56	−32.16 ± 25.27	−8.37 ± 5.96	−27.25 ± 14.05
$T\Delta S$	−48.83 ± 11.31	−33.93 ± 11.43	−41.87 ± 8.02	−47.41 ± 8.99	−46.54 ± 11.50	−45.35 ± 9.30	−38.06 ± 10.68
ΔG_{TOT}	4.50	68.09	−4.87	2.75	14.38	36.98	10.81

ΔE_{ELE} = coulombic energy, ΔE_{VDW} = vander walls energy, ΔE_{INT} = internal energy, $\Delta E_{GAS} = \Delta E_{ELE} + \Delta E_{VDW} + \Delta E_{INT}$, ΔE_{SNP} = nonpolar solvation energy, ΔE_{SP} = polar solvation energy, $\Delta E_{SOLV} = \Delta E_{SNP} + \Delta E_{SP}$ = total solvation energy, $\Delta E_{TOT\_ELE} = \Delta E_{ELE} + \Delta E_{SP}$ = total electrostatic energy, $\Delta E_{GAS+SOLV} = \Delta E_{GAS} + \Delta E_{SOLV}$ = enthalpy, $T\Delta S$ = solute entropy, $\Delta G_{TOT} = \Delta E_{GAS+SOLV} - T\Delta S$ = absolute free energy

All energy values are in kcal/mol

Table 4 Hydrogen bonding distances between the atoms of G-tetrads of MOD2

S. no.	Res. no.	Atom1	Res. no.	Atom2	Distance(Å)
1	LCG-9	O6	LCG-3	N1	2.942
2	LCG-3	N2	LCG-9	O6	3.511
3	LCG-3	N2	LCG-9	N7	3.083
4	LCG-9	N1	LCG-21	O6	2.989
5	LCG-21	O6	LCG-9	N2	3.662
6	LCG-9	N2	LCG-21	N7	3.055
7	LCG-16	O6	LCG-21	H1	2.174
8	LCG-16	N7	LCG-21	N2	2.989
9	LCG-16	N1	LCG-3	O6	2.999
10	LCG-17	O3′	LCG-17	O6	2.896
11	LCG-22	N1	LCG-15	N7	3.352
12	LCG-15	N7	LCG-22	N2	2.872
13	LCG-22	N2	LCG-23	N7	3.114
14	LCG-22	N7	LCG-10	N7	3.211
15	LCG-22	N7	LCG-10	N1	3.463
16	LCG-10	N1	LCG-22	O6	3.531
17	LCG-22	N1	LCG-15	O6	2.996
18	LCG-15	N1	LCG-4	O6	2.997
19	LCG-10	O6	LCG-4	N1	2.892
20	LCG-10	N7	LCG-4	N2	2.875
21	LCG-5	N1	LCG-11	N7	3.093
22	LCG-11	O2P	LCG-5	N2	3.188
23	LCG-11	N2	LCG-23	O2P	3.420
24	LCG-4	N7	LCG-15	N2	3.028

binding free energy (−4.87 kcal/mol). MOD4 is the second most stable structure with respect to other models, which is also a mixed hybrid type structure, where, first and fourth strands are parallel and second and third strand are anti-parallel. The binding energy for this model is −44.66 kcal/mol and binding free energy is 2.75 kcal/mol, which gives second lowest value. Among all structures, MOD2, which is also a mixed hybrid type structure where first, second and fourth strand are in parallel while third strand is in antiparallel, has lowest stability because its binding energy is 34.16 kcal/mol and also binding free energy for this model is most unfavorable (positive) relative to other models.

Consequently, the binding free energy calculation suggests that MOD3 exhibits higher stability than the other modified structures.

Hydrogen bonding patterns

The starting fiber model [57], as well as high resolution crystal structure [58] of parallel quadruplex with coordinated cations, contains N1–H1–O6 and N2–H2–N7 (Hoogsteen type) hydrogen bonds between adjacent guanines, leading to eight hydrogen bonds per G-tetrad. Ab initio study on G-tetrad [56] also suggest that, in the absence of a coordinated cation, G-tetrad is stabilized by bifurcated hydrogen bonds between N1–H1–N7 and N2–H2–O6 atoms, whereas MD simulation of the 4-mer parallel quadruplex structure without any coordinated cations [52]

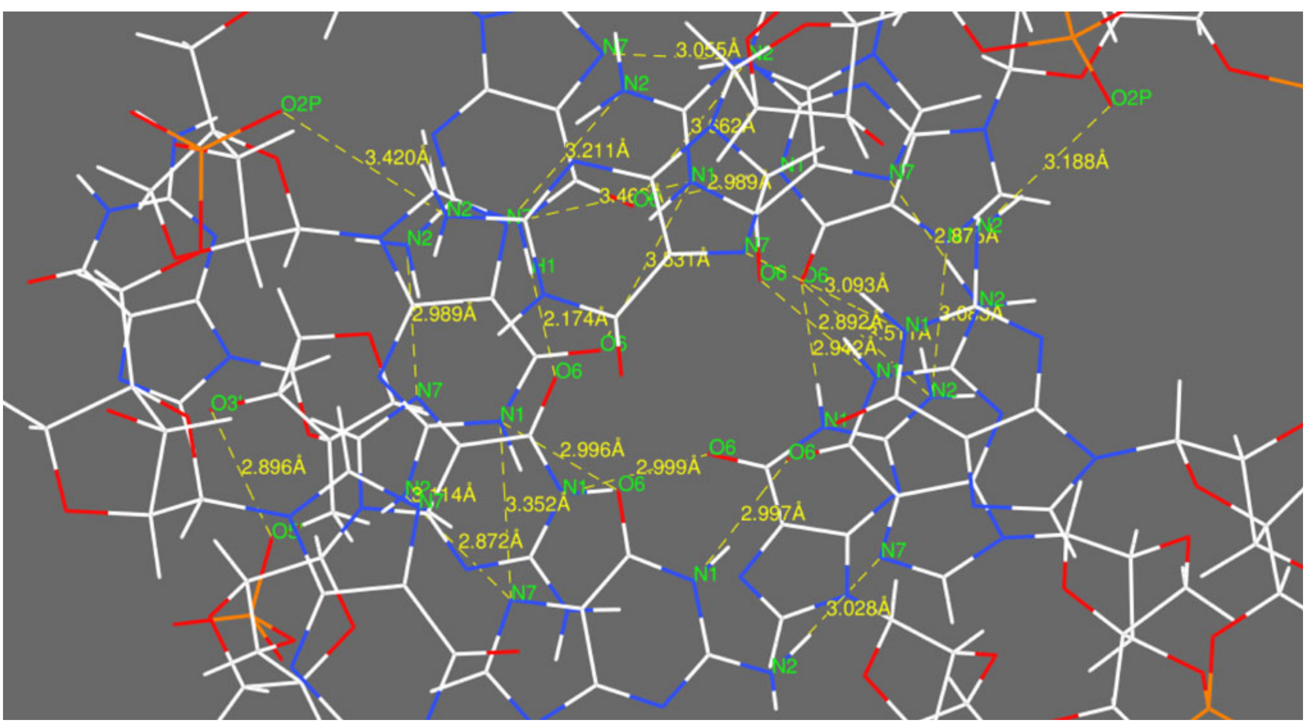

Fig. 7 Hydrogen bonds and corresponding distances between the atoms of G-tetrads in MOD2

reported disruption of the G-tetrad geometry due to strand slippage. It is observed that during the molecular dynamics simulation, hydrogen bonding scheme within G-tetrad depends on the electrostatic interaction between the polar atoms of guanine bases and the coordinating molecules. The inter atomic distances between potential hydrogen bond forming groups within a G-tetrad and distances between different O6 atoms in the MD structures obtained during 20 ns dynamic run, which are listed in Table 4 and Table S2. During the simulation, hydrogen bonds in all G-tetrads (Fig. 7 and S3) are retained for all seven structures. However, due to the strong attractive force between coordinated K+ ion and O6 atoms, guanine bases in some G-tetrads undergo in plane rotational motion. Because of this rotational motion, N_1–H_1–O_6 and in some cases N_2–H_2–N_7 hydrogen bonds are elongated. Whereas N_1 atom and N_2 atom of guanine base comes closer to N_7 and O_6 atom of neighboring guanine base and forms a N_1–H_1–N_7 and N_2–H_2–O_6 hydrogen bond. Thus, there centered hydrogen bond play additional role to stabilized these tetrads. As shown in Table 4, G-tetrad of MOD2 forms 24 hydrogen bonds, which is the maximum number of bond with association to all other models. The bonding distance and number of hydrogen bonds for MOD1 and MOD3-7 are listed in Table S2. Due to maximum number of H-bonding between tetrads, MOD2 structure shows more stability.

Conclusion

It is well known that LNA modifications stabilize duplex and triplex DNA. In previous report a fully LNA modified quadruplex d(TG3T) was stabilized relative to the DNA quadruplex (37). This study explains the relative stability of modified d(TAGGGT)$_4$ G-quadruplex configurations. MD simulations are performed for seven different modified quadruplex structures. Among all modified structures, one of the mixed hybrid type structures, with two parallel and two antiparallel strands, shows lowest free energy. The energy calculations suggest that this mixed hybrid type conformation is the most stable among other conformations. RMS deviation of MOD2 complex, which is also a mixed hybrid type structure, is smaller. RMS deviation values of other complexes are very high and vary during the simulation. This study suggest that the modified monomeric DNA sequence d(TAGGGT)$_4$ can form only mixed hybrid type of structure. These modified sequences may help to inhibit the growth of the tumor and cancerous cell. The tendency of human telomeric G-quadruplex structure to adopt different conformations under varying environmental conditions which posses significant challenges in the design of specific drugs that could efficiently bind and stabilize the structure. Our study provides in-depth insight into the conformational aspects of the quadruplex for subsequent evaluation of its respective interactions with drugs.

Acknowledgments The authors are thankful to DST, New Delhi for computational facility in the form of FIST scheme. We thankfully acknowledge the partial computational work at BRAF of C-DAC, Pune, India. KDD acknowledges CSIR for senior research fellowship.

References

1. Feldser DM, Hackett JA, Greider CW (2003) Telomere dysfunction and the initiation of genome instability. Nat Rev Cancer 3:623–627
2. Cech TR (2000) Life at the end of the chromosome: telomeres and telomerase. Angew Chem Int Ed 39:34–43
3. Wright WE, Tesmer VM, Huffman KE, Levene SD, Shay JW (1997) Normal human chromosomes have long G-rich telomeric overhangs at one end. Genes Dev 11:2801–2809
4. Organisian L, Bryan TM (2007) Physiological relevance of telomeric G-quadruplex formation: a potential drug target. Bioessays 29:155–165
5. Kim NW, Piatyszek MA, Prowse KR, Harley CB, West MD, Ho PL, Coviello GM, Wright WE, Weinrich SL et al (1994) Specific association of human telomerase activity with immortal cells and cancer. Science 266:2011–2015
6. Hanahan D, Weinberg RA (2000) The hallmarks of cancer. Cell 100:57–70
7. Mergny JL, Helene C (1998) G-quadruplex DNA: a target for drug design. Nat Med 4:1366–1367
8. Sun DY, Hurley LH (2001) Methods in enzymology, drug-nucleic acid interactions. Academic Press, Inc 340:573–592
9. Hurley LH (2002) DNA and its associated processes as targets for cancer therapy. Nat Rev Cancer 2:188–200
10. Neidle S, Parkinson G (2002) Telomere maintenance as a target for anticancer drug discovery. Nat Rev Drug Discov 1:383–393
11. Bodnar AG, Ouellette M, Frolkis M, Holt SE, Chiu CP, Morin GB, Harley CB, Shay JW, Lichtsteiner S et al (1998) Extension of life-span by introduction of telomerase into normal human cells. Science 279:349–352
12. Harley CB (1991) Telomere loss: mitotic clock or genetic time bomb? Mutat Res 256:271–282
13. Sun H, Karow JK, Hickson ID, Maizels N (1998) The Bloom's syndrome helicase unwinds G4 DNA. J Biol Chem 273:27587–27592
14. Sun H, Bennett RJ, Maizels N (1999) The saccharomyces cerevisiae Sgs1 helicase esciently unwinds G-G paired DNAs. Nucleic Acids Res 27:1978–1984
15. Hackett JA, Feldser DM, Greider CW (2001) Telomere dysfunction increases mutation rate and genomic instability. Cell 106:275–286
16. Salazar M, Thompson BD, Kerwin SM, Hurley LH (1996) Thermally induced DNA:RNA hybrid to G-quadruplex transitions: possible implications for telomere synthesis by telomerase. Biochemistry 35:16110–16115
17. Hurley LH (2001) Secondary DNA structures as molecular targets for cancer therapeutics. Biochem Soc Trans 29:692–696
18. Hurley LH, Wheelhouse RT, Sun D, Kerwin SM, Salazar M, Fedoro OY, Han FX, Han HY, Izbicka E et al (2000) G-quadruplexes as targets for drug design. Pharmacol Ther 85:141–158
19. Neidle S, Read MA (2000) G-quadruplexes as therapeutic targets. Biopolymers 56:195–208

20. Mergny J-L, Helene C (1998) G-quadruplex DNA: a target for drug design. Nat Genet 4:1366–1367
21. Bearss DJ, Hurley LH, Von Hoff DD (2000) Telomere maintenance mechanisms as a target for drug development. Oncogene 19:6632–6641
22. Gowan SM et al (2002) A G-Quadruplex-interactive potent small-molecule inhibitor of telomerase exhibiting in vitro and in vivo antitumor activity. Mol Pharmacol 61:1154–1162
23. Read MA et al (2001) Structure-based design of selective and potent G quadruplex-mediated telomerase inhibitors. Proc Natl Acad Sci USA 98(9):4844–4849
24. Parkinson GN, Lee MPH, Neidle S (2002) Crystal structure of parallel quadruplexes from human telomeric DNA. Nature 417:876–880
25. Ambrus A, Chen D, Dai JX, Bialis T, Jones RA, Yang DZ (2006) Human telomeric sequence forms a hybrid-type intra- molecular G-quadruplex structure with mixed parallel/antiparallel strands in potassium solution. Nucleic Acids Res 34:2723–2735
26. Xu Y, Noguchi Y, Sugiyama H (2006) The new models of the human telomere d[AGGG(TTAGGG)(3)] in K+ solution. Bioorg Med Chem 14:5584–5591
27. Luu KN, Phan AT, Kuryavyi V, Lacroix L, Patel DJ (2006) Structure of the human telomere in K+ solution: an intramolecular (3 + 1) G-quadruplex scafold. J Am Chem Soc 128:9963–9970
28. Wang Y, Patel DJ (1993) Solution structure of the human telomeric repeat d[$AG_3(T_2AG_3)_3$] G-tetraplex. Structure 1:263–282
29. He Y, Neumann RD, Panyutin IG (2004) Intramolecular quadruplex conformation of human telomeric DNA assessed with ^{125}I-radioprobing. Nucleic Acid Res 32:5359–5367
30. Redon S, Bombard S, Elizondo-Riojas MA, Chottard JC (2003) Platinum cross-linking of adenines and guanines on the quadruplex structures of the $AG_3(T_2AG_3)_3$ and $(T_2AG_3)_4$ human telomere sequences in Na+ and K+ solutions. Nucleic Acid Res 31:1605–1613
31. Qi J, Shafer RH (2005) Covalent ligation studies on the human telomere quadruplex. Nucleic Acid Res 33:3185–3192
32. Li J, Correia 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 Acid Res 33:4649–4659
33. Crooke ST (2004) Progress in antisence technology. Annu Rev Med 55:61–95
34. Bondensgaard K, Petersen M, Singh SK, Rajwanshi VK, Kumar R, Wengel J, Jacobsen JP (2000) Structural studies of LNA: RNA duplexes by NMR: conformations and implications for RNase H activity. Chemistry 6:2687–2695
35. Nielsen KE, Singh SK, Wengel J, Jacobsen JP (2000) Solution structure of an LNA hybridized to DNA: NMR study of the d(CT(L)GCT(L)T(L)CT(L)GC): d(GCAGAAGCAG) duplex containing four locked nucleotides. Bioconjug Chem 11:228–238
36. Petersen M, Bondensgaard K, Wengel J, Jacobsen JP (2002) Locked nucleic acid (LNA) recognition of RNA: NMR solution structures of LNA: RNA hybrids. J Am Chem Soc 124:5974–5982
37. Randazzo A, Esposito V, Ohlenschlager O, Ramachandran R, Mayol L (2004) NMR solution structure of a parallel LNA quadruplex. Nucleic Acids Res 32:3083–3092
38. Sacca B, Lacroix L, Mergny J-L (2005) The effect of chemical modifications on the thermal stability of different G-quadruplex-forming oligonucleotides. Nucleic Acids Res 33:1182–1192
39. Ivanova A, Rosch N (2007) The structure of LNA: DNA hybrids from molecular dynamics simulations: the effect of locked nucleotides. J Phys Chem A 111(38):9307–9319
40. Pande V, Nilsson L (2008) Insights into structure, dynamics and hydration of locked nucleic acid (LNA) strand-based duplexes from molecular dynamics simulation. Nucleic Acids Res 36(5): 1508–1516
41. Cornell WD, Cieplak P, Baylyl CI, Gould IR, Merz JKM et al (1995) A second generation force field for the simulation of proteins, nucleic acids and organic molecules. J Am Chem Soc 117:5179–5197
42. Price DJ, Brooks CL 3rd (2004) A modified TIP3P water potential for simulation with Ewald summation. J Chem Phys 121:10096–10103
43. Duan Y, Wu C, Chowdhury S, Lee MC, Xiong G, Zhang W, Yang R, Cieplak P, Luo R, Lee T (2003) A point-charge force field for molecular mechanics simulations of proteins based on condensed-phase quantum mechanical calculations. J Comput Chem 24:1999–2012
44. Lee MC, Duan Y (2004) Distinguish protein decoys by using a scoring function based on a new Amber force field, short molecular dynamics simulations, and the generalized born solvent model. Proteins 55:620–634
45. Yang L, Tan C, Hsieh M-J, Wang J, Duan Y, Cieplak P, Caldwell J, Kollman PA, Luo R (2006) New-generation Amber united-atom force field. J Phys Chem B 110:13166–13176
46. Case DA, Cheathem TE, Daren T, Gohlke H, Luo R, Merz KM, Onufriev A et al (2005) The AMBER biomolecular simulation programs. J Comput Chem 26:1668–1688
47. Ryckaert JP, Ciccotti G, Berendsen HJC (1977) Numerical integration of the Cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. J Comput Phys 23:327–341
48. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE (2004) UCSF Chimera—a visualization system for exploratory research and analysis. J Comput Chem 25:1605–1612
49. Gohlke H, Case DA (2004) Converging free energy estimates: MM-PB(GB)SA studies on the protein-protein complex Ras-Raf. J Comput Chem 25:238–250
50. Dubey KD, Ojha RP (2011) Binding free energy calculations with hybrid QM/MM methods for Abl kinase inhibitors. J Biol Phys 37:69–78
51. Hud NV, Schultze P, Sklenar V, Feigon J (1999) Binding sites and dynamics of ammonium ions in a telomere repeat DNA quadruplex. J Mol Biol 285:233–243
52. Spackova N, Berger I, Sponer J (1999) Nanosecond molecular dynamics simulation of parallel and antiparallel guanine quadruplex DNA molcules. J Am Chem Soc 121:5519–5534
53. Chowdhury S, Bansal M (2001) G-quadruplex structure can be stable with only some coordination sites being occupied by cations: a six-nanosecond molecular dynamics study. J Phys Chem B 105:7572–7578
54. Cavallari N, Calzolarii A, Garbesi A, Di Felice R (2006) Stability and migration of metal ions in G4-wires by molecular dynamics simulations. J Phys Chem B 110:26337–26348
55. Agrawal S, Ojha RP, Maiti S (2008) Energetics of the human Tel-22 quadruplex—telomestatin interaction: a molecular dynamics study. J Phys Chem B 112(22):6828–6836
56. Gu J, Leszczynski J, Bansal M (1999) A new insight into the structure and stability of Hoogsteen hydrogen-bonded G-tetrad: an ab initio SCF study. Chem Phys Lett 311:209
57. Arnott S, Chandrasekaran R, Marttila CM (1974) Structures for polyinosinic acid and polyguanylic acid. Biochem J 141:537–543
58. Phillips K, Dauter Z, Murchie AIH, Lilley DMJ, Luisi B (1997) The crystal structure of a parallel-stranded guanine Tetraplex at 0.95 Å resolution. J Mol Biol 273:171–182