Indian Journal of Chemistry
Vol. 56A, September 2017, pp. 907-912

Understanding the unfolding mechanism of human telomeric G-quadruplex using steered molecular dynamics simulation

Pralok K Samanta[a] & Swapan K Pati[a, b, *]

[a]Theoretical Sciences Unit, [b]New Chemistry Unit, Jawaharlal Nehru Center for Advanced Scientific Research, Jakkur PO, Bangalore 560 064, India

Email: pati@jncasr.ac.in

Received 4 July 2017; revised and accepted 18 August 2017

The unfolding pathway of human telomeric G-quadruplex with three G-tetrads in presence of K^+ and Na^+ ions, separately using steered molecular dynamics (SMD) simulation is reported. The isothermal-isobaric all-atoms classical molecular dynamics simulation results show that three K^+ and three Na^+ ions are required within the central channel of the G-quadruplex (PDB ID: 143D and 2HY9, respectively) to stabilize the respective overall structure. To obtain the unfolded G-quadruplex which is ~5-6 times of its initial contour length, SMD simulation has been carried out by fixing one end of the G-quadruplex and constraining the other end to move only along the long axis (z-axis). The SMD results suggest that the unfolding of G-quadruplex occurs via G-triplex intermediates independent of the presence of cations (K^+, Na^+).

Keywords: Molecular dynamics, Steered molecular dynamics, Unfolding kinetics, H-bonding, G-Quadruplex, DNA, Triplex DNA, Telomeres, Guanine

G-quadruplex DNA structures formed by the self-assembly of guanine-rich poly-nucleotides have a major role in maintaining the ends of chromosomes, the telomere.[1] Human telomeric DNA consists of tandem repeats ($d[5'\text{-}T_2AG_3]_n$) for 5-8 kb in length towards the end of chromosome and terminate in a single-stranded 3'-overhang of 100-200 nucleobases in length.[2,3] In G-quadruplex, four guanines are held in a plane by Hoogsteen hydrogen-bonds (H-bonds) to each other that form a cyclic structure (G-tetrad) with a central cavity. Each guanine comes from a strand and hence, the G-quadruplex forms a four stranded structure. Also, the G-tetrad planes are stabilized by π-stacking interaction between the stacked-tetrads and the electrostatic interaction involving monovalent cations (e.g., K^+, Na^+) sandwiched between two G-tetrads which reduce the repulsion between the carbonyl oxygen atoms in neighbouring guanines.[4-7] The G-quadruplex protects telomere ends and regulates telomere length.[8] It can regulate gene expression and has an important role in several aspects of metabolism, such as transcription termination, splicing and translations.[9-13] It also has recently received great interest because of possible targets for cancer therapy.[14,15] In addition, synthetic G-quadruplexes are used as drugs for treating cancer,[16-19] prevention of thrombosis[20] and inhibition of HIV replication.[21] Also, the drugs that bind to specific quadruplex DNAs can be used for the treatment of a variety of pathological conditions including cancer.[22] The dynamics of enzyme-catalyzed unwinding depends on the unfolding pathway of G-quadruplex.[23-25] Therefore, it is important to study the unfolding pathway of the G-quadruplex to understand the functions of telomeres.

The kinetics of folding and unfolding of G-quadruplexes has been studied with various experimental techniques, including circular dichroism (CD), differential scanning calorimetry, isothermal titration calorimetry, optical tweezers experiments[26-28] and theoretically with molecular dynamics (MD) simulations.[29-32] Though, there are a few experimental studies which reported the existence of G-triplex (three guanine strands) intermediate[28,33] during the unfolding of G-quadruplex, the unfolding pathway of the G-quadruplex has not been clearly understood yet.

Recently, Li *et al.*[34] have employed magnetic tweezers experiments to investigate the unfolding kinetics of single human telomeric G-quadruplex with the sequence, $d[G_3(T_2AG_3)_3]$ in 100 m*M* Na^+ buffer. Their results probe the existence of triplex intermediate in the unfolding pathway of the G-quadruplex. Motivated by the experimental work by Li *et al.*[34] we have performed steered molecular

dynamics (SMD) simulation for the detail study of the unfolding kinetics of G-quadruplexes. We have considered the ends of the molecule to be linked to Hookean spring serving as force sensors whose ends move in opposite direction with constant velocity as reported in the experimental work by Li *et al*.[34] The stereochemistry of the glycosidic bonds (sys or anti), directions of strands (parallel or anti-parallel), molecularity, the sequence and topology of the loops connecting the G-tetrads and the sequences flanking the G-tetrads lead to different three-dimensional (3D) orientation of a G-quadruplex. Also, the stability and configurations of the quadruplex depends upon the presence of specific metal cations.[35] In this article, we have considered two different metal ions (K^+ and Na^+) with two different monomolecular G-quadruplex configurations and by calculating free energy profile of the unfolding pathways show that in both cases, unfolding kinetics involves G-triplex intermediates.

Computational Details

For our study, we have considered two different monomolecular quadruplexes, (i) human telomeric repeat (seq1=d[$AG_3(T_2AG_3)_3$]) G-quadruplex in Na^+ solution (PDB ID: 143D)[36], and, (ii) human telomeric G-quadruplex (seq2=d[$A_3G_3(T_2AG_3)_3A_2$]) in K^+ solution (PDB ID: 2HY9)[37]. The quadruplexes were solvated with water in a box of 5.0×5.0×25.0 nm^3. We considered transferable intermolecular potential three point (TIP3P) model for water solvent.[38,39] Appropriate numbers of Na^+ or K^+ ions were added to electroneutralize the systems. Simulations were run in the NPT ensemble using leap-frog algorithm for integrating Newton's equation of motion at constant temperature (300 K) and pressure (1 bar). The temperature was kept constant at 300 K by velocity rescaling with a stochastic term[40] and the pressure was kept at 1 bar using isotropic coupling to a Berendsen barostat[41]. An integrations time step of 1.5 fs was used and van der Waals interactions were calculated using a cutoff of 1.4 nm. At a distance smaller than 1.4 nm, electrostatic interactions were calculated explicitly, whereas long-range electrostatic interactions were calculated by Particle-Mesh Ewald (PME) summation.[42] All covalent hydrogen bonds were constrained using the LINCS algorithm.[43] All the simulations were carried out using GROMACS-4.0.7[44] software package with Amber99 force field[45], which is commonly used to study all major kinds of nucleic acid systems.[46-50]

Initially, the high energy contacts between the atoms in the initial conformations were removed by minimizing the energy using the steepest decent method following which, NPT simulations were carried out for 200 ps to equilibrate the entire system. Finally, production phase of 3 ns long simulation was performed. In the present system, there were around ~800 atoms of each G-quadruplex (seq1 and seq2) and ~20,000 solvent (water) molecules in a box of 5.0×5.0×25.0 nm^3. After 3 ns of NPT simulations, we considered the final geometry and fixed one end (5'-endmost nucleotide) of the monomolecular G-quadruplex of both seq1 and seq2 and allowed the other end to move along only the z-axis. A guiding potential, $h(r; \lambda) = (k/2)[\xi(r)-\lambda]^2$, was added to control the end-to-end distances, ξ. The parameter, λ, was varied between 0–10 nm. A force constant of k = 1000 kJ/(mol nm^2) with two different constant velocities (1 nm/ns and 10 nm/ns) were used for the SMD simulations for 10 ns and 1 ns, respectively.

The MD trajectory was visualized by Visual Molecular Dynamics (VMD) software.[51] Free energy changes during the stretching were calculated based on umbrella sampling method.[52] The free energies were calculated using weighted histogram analysis method (WHAM) as ensemble average.[52-54]

Results and Discussion

Equilibrium simulations

Seq1 in Na^+ solution forms a basket structure characterized by anti-parallel strands, two lateral loops and one diagonal loop, whereas, seq2 (in K^+ solutions) forms propeller structure characterized by mixed parallel and anti-parallel strands with lateral loops only. After 3 ns of NPT simulation, for both the seq1 and seq2, the central units of G-quadruplexes are H-bonded arrays of guanine bases (G-tetrads) and there are three such G-tetrads in each sequence. Each guanine base in a G-tetrad is involved in the formation of four Hoogsteen H-bonds with two neighbouring guanines. Thus, there are eight such H-bonds in each G-tetrads.[7] We observe that three Na^+ and three K^+ ions are sandwiched between two G-tetrads in the central channel of the G-quadruplex for seq1 and seq2, respectively (see Fig. 1) and the orientation of each is similar to the respective PDB structure.[36,37]

Force induced conformational changes

After 3 ns of NTP simulations, we considered the final geometry and fixed the endmost 5'-end nucleotide of each monomolecular G-quadruplex and allowed the other end to move along only the z-axis to study the unfolding mechanism. The detailed

structural changes along the pulling trajectory with pulling rate 1 nm/ns are described in Fig. 2. The initial quadruplex structure (~20 Å) changed to the stretched (unfolded) one which is ~5-6 times of its initial contour length. With increase in the force (with time) the DNA extension increased and finally became completely stretched. We found that in the beginning one strand separated and then complete unfolding occurred. In fact, the unfolding occurred via G-triplex intermediates. This is because the G-triads and G-tetrad are more stable than the GG base pair.[33]

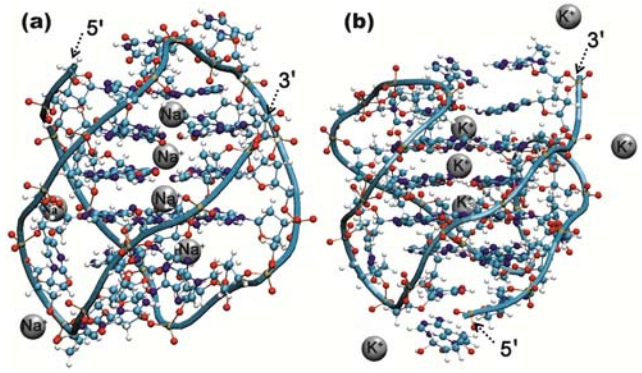

Fig. 1 — Equilibrium structure (after 3 ns of NPT simulation) of (a) seq1 (d[$AG_3(T_2AG_3)_3$]) in Na^+ solution, and, (b) seq2 (d[$A_3G_3(T_2AG_3)_3A_2$]) in K^+ solution.

The triplex structures remain intact around ~4 ns (from ~3 ns to 7 ns) and the alkali metal ions (K^+, Na^+) get sandwiched between the G-triplexes to stabilized it.[57] Subsequently, the metal ions are released into the solvent and complete unfolding occurs.

Interaction with metal ions (Na^+ and K^+)

As mentioned above, three Na^+ and three K^+ ions are positioned within the central channel of the G-quadruplex. We note that, the metal ions are not positioned in the center of G-tetrads because of their large ionic radius as compared to cavity size.[7] The G-tetrads have a region of negative electrostatic potential due to charge localization on the carbonyl oxygen atoms in the central cavity. Such electrostatic repulsion in the cavity is reduced by the presence of the metal ions (Na^+, K^+).[55,56] In Fig. 3, we plot the electrostatic and Lennard-Jones (LJ) interactions of metal ions for both the sequences (seq1 and seq2) throughout the unfolding processes. In the beginning (i.e., G-quadruplex conformer) the interaction between the sequences and metal ions is mainly electrostatic (~300–400 kcal/mol).[55,56] The strength of both electrostatic and LJ interactions decreases with unfolding. The electrostatic interaction dominates the LJ interaction (~30 kcal/mol) which is

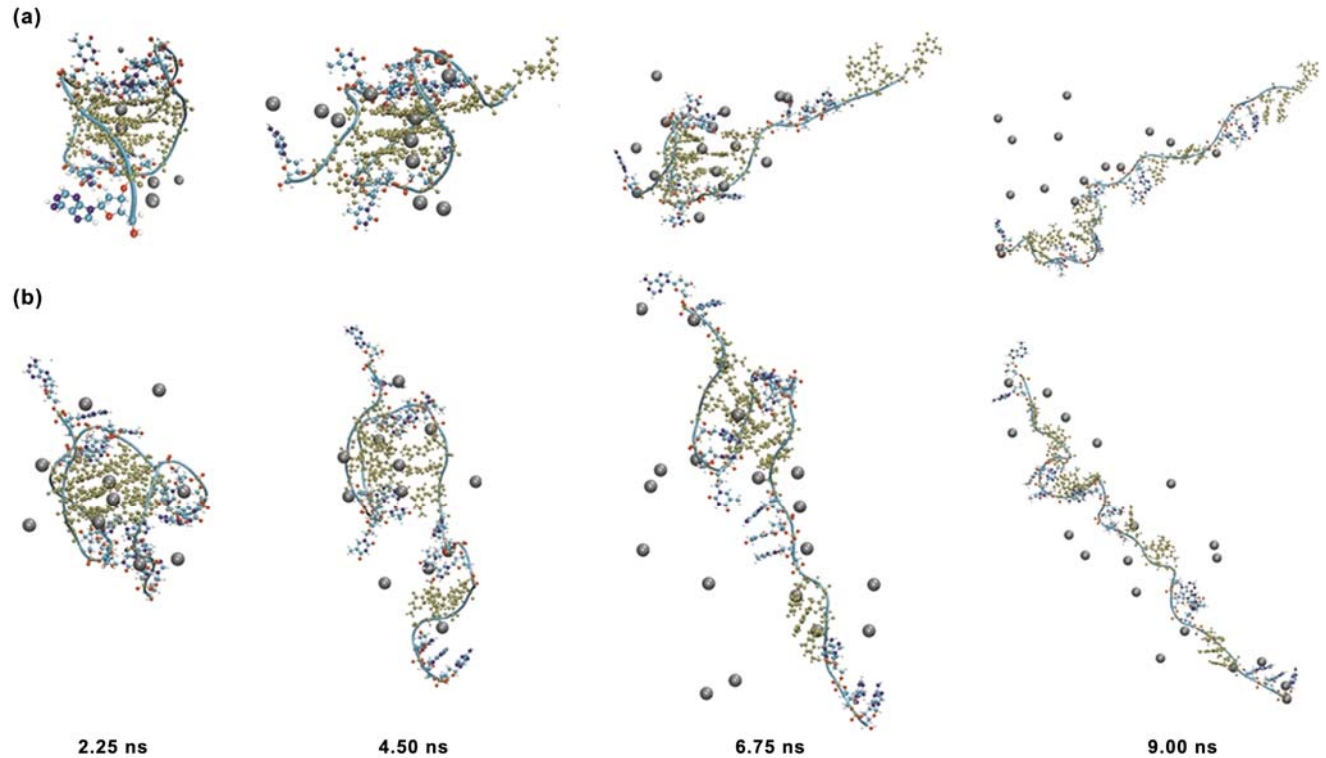

Fig. 2 — Evolution of the structure of the G-quadruplex during pulling of (a) seq1 in Na^+ solution, and, (b) seq2 in K^+ solution. [Pulling rate: 1 nm/ns].

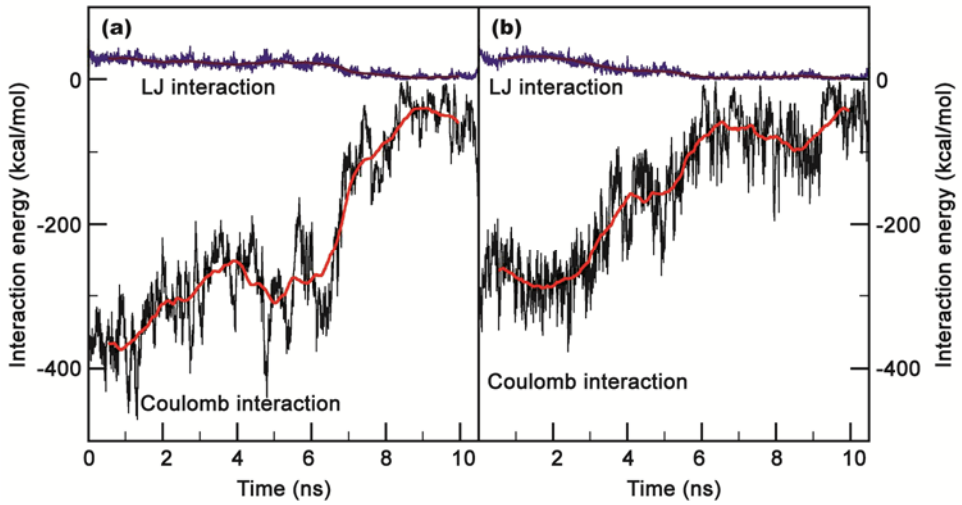

Fig. 3 — Columbic and Lennard-Jones (LJ) interaction along the way of unfolding of (a) seq1 and Na^+ ion, and, (b) seq2 and K^+ ion. [Pulling rate: 1 nm/ns].

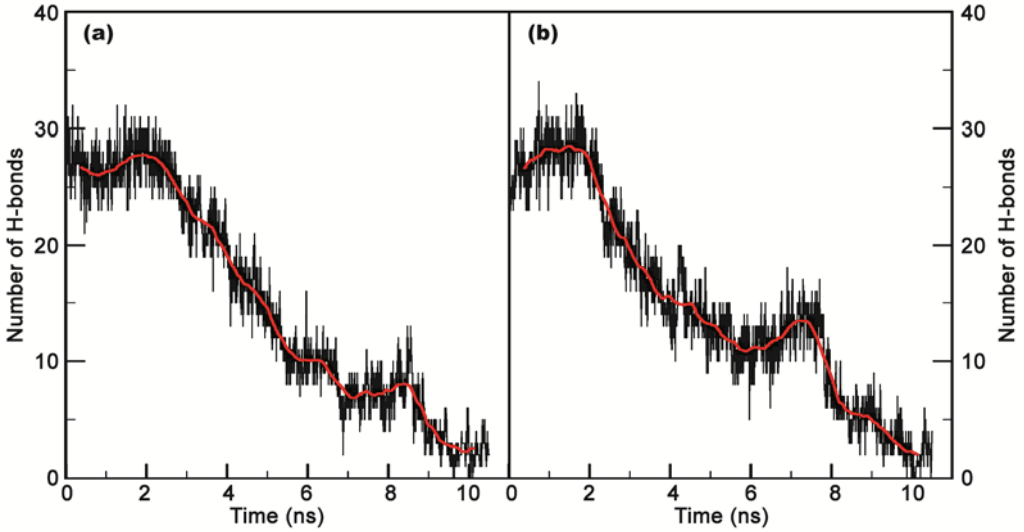

Fig. 4 — Number of H-bonds with time for (a) seq1 in Na^+ solution, and, (b) seq2 in K^+ solution [Pulling rate: 1 nm/ns].

unfavorable in this case. The unfavorable LJ interaction indicates large ionic radius of metal ions (Na^+, K^+) compared to the central cavity size in G-quadruplex.[7] In the case of G-triplex, the G-triads are stacked together. Unlike G-tetrads, there is no central cavity in G-triads and hence the interaction with metal ions decreases. However, like G-quadruplex, the G-triplex is also stabilized by metal ions.[57] In the unfolded species, the metal ions were released into the solvent and hence, both the electrostatic and LJ interactions tend to vanish.

Hydrogen bonding profile

The H-bond is represented as D–H...A where D is the donor atom and A is the acceptor atom. In the case of DNA, D is nitrogen (N) atom and A is N or oxygen (O) atom. We have used the g_hbond GROMACS-4.0.7 utility to calculate the number of H-bonds with a cutoff D–H…A distance of 0.35 nm and cutoff ∠H-D-A angle of 30°. These parameters of H-bonding have been consider based on H-bonding in proteins as reported previously.[58,59] With increase in the force, the DNA extension increases sharply and hence the number of H-bonds decreases (Fig. 4 and Fig. S1, Supplementary Data). The number of H-bonds becomes zero for the complete stretched structure. For both seq1 in Na^+ solution (Fig. 4a) and seq2 in K^+ solution (Fig. 4b), we found that initially there was a decrease in the number of H-bonds, which

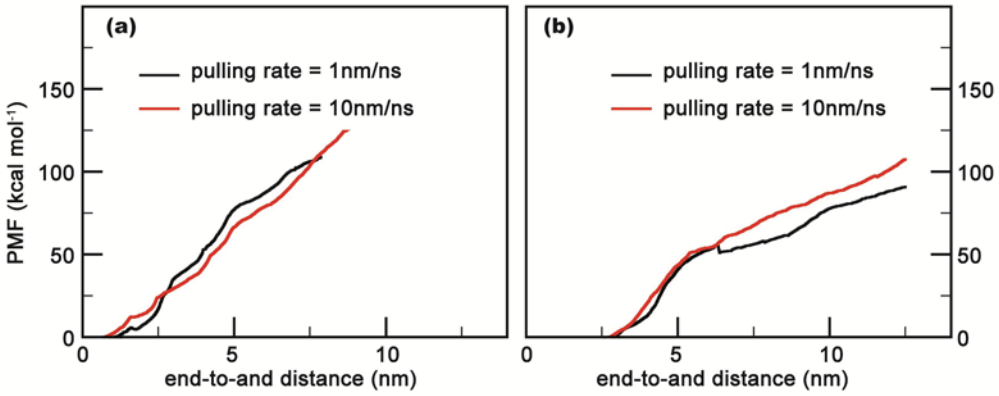

Fig. 5 — Potential of mean force (PMF) as a function of end-to-end distance (in nm) for (a) seq1 in Na^+ solution, and, (b) seq2 in K^+ solution.

implies that the G-quadruplex structure begins to break up. The number of H-bonds becomes stable (~12 H-bonds) around 7 ns of the SMD simulation, which reflects the existence of some intermediate conformation. Each G-terad forms 8 H-bonds. Since there three G-tetrads are present in the initial G-quadruplex, therefore the initial number of H-bonds is about 24 and some other H-bonds due to loop part of the DNA also contribute to the total number of H-bonds. Thus, we observe a total number of H-bonds ~25-30 at the very beginning of the SMD simulations. In a G-triad, the central guanine is involved in formation of total four H-bonds with other two guanines. Therefore, the total number of H-bonds of a G-triplex with three stacks is ~12. This result clearly indicates the presence of G-triplex intermediate during the unfolding of G-quadruplex.

Free energy profile

Details of unfolding kinetics of the G-quadruplexes are described from the calculation of potential of mean force (PMF). The PMF as a function of the end-to-end distance is shown in Fig. 5. Different humps in the PMF vs. end-to-end distance plot represent existence of an intermediate. Therefore, many different humps in the case of G-quadruplex in Na^+ solution (Fig. 5a) point towards existence different intermediate conformers and hence, a complicated kinetics of unfolding pathway is observed as reported experimentally.[28] The kinetics of unfolding pathway for G-quadruplex in K^+ solution is, however, simple with a single hump ~7 nm.

Conclusions

To summarize, the unfolding pathway of human telomeric G-quadruplex in presence of two different alkali metal ions (K^+ and Na^+) is studied by atomistic steered molecular dynamics (SMD) simulation. Seq1 in Na^+ solution forms a basket structure characterized by anti-parallel strands and seq2 in K^+ solutions forms propeller structure characterized by mixed parallel and anti-parallel strands. At equilibration, three K^+ and three Na^+ ions were positioned within the central channel of the G-quadruplex to stabilize the overall structure. When one end of the G-quadruplex was fixed and constrained the other end to move only along z-axis to make the quadruplex unfolded, the unfolding occurs via G-triplex intermediates for both the ions solutions. Thus the unfolding of G-quadruplex occurs via G-triplex intermediate, independent of the presence of cations (K^+, Na^+) and hence the special configuration of the G-quadrplexes. We believe that these new unfolding pathways involving triplex intermediate could be helpful for the development of G-quadruplex binding ligands and anti-cancer drugs. Furthermore, these pathways could be applied to understand the folding/unfolding mechanism of other poly-nucleic acids (DNA/RNA).

Supplementary Data

Supplementary data associated with this article is available in the electronic form at http://www.niscair.res.in/jinfo/jinfo/ijca/IJCA_56A(09)907-912_SupplData.pdf.

Acknowledgement

We acknowledge the Centre for Computational Materials Science (CCMS), Jawaharlal Nehru Centre for Advanced Scientific Research (JNCASR), Bangalore, India, for computational and storage resources. SKP acknowledges Department of Science & Technology, Govt. of India, New Delhi, India, for JC Bose fellowship.

References

1 Shafer R H & Smirnov I, *Biopolymers*, 56 (2000) 209.
2 Blackburn E H & Gall J G, *J Mol Biol*, 120 (1978) 33.

3 Moyzis R K, Buckingham J M, Cram L S, Dani M, Deaven L L, Jones M D, Meyne J, Ratliff R L & Wu J R, *Proc Natl Acad Sci USA*, 85 (1988) 6622.
4 Ida R & Wu G, *J Am Chem Soc*, 130 (2008) 3590.
5 Gilbert D E & Feigon J, *Curr Opin Struct Biol*, 9 (1999) 305.
6 Bugaut A, Murat P & Balasubramanian S, *J Am Chem Soc*, 134 (2012) 19953.
7 Jissy A K, Ashik U P M & Datta A, *J Phys Chem C*, 115 (2011) 12530.
8 Rice C & Skordalakes E, *Comput Struct Biotech J*, 14 (2016) 161.
9 Han H Y & Hurley L H, *Trends Pharmacol Sci*, 21 (2000) 136.
10 Kerwin S M, *Curr Pharm Des*, 6 (2000) 441.
11 Mergny J L & Helene C, *Nat Med*, 4 (1998) 1366.
12 Bugaut A & Balasubramanian S, *Nucleic Acids Res*, 40 (2012) 4727.
13 Collie G W & Parkinson G N, *Chem Soc Rev*, 40 (2011) 5867.
14 Islam M K, Jackson P J M, Rahman K M & Thurston D E, *Future Med Chem*, 8 (2016) 1259.
15 Neidle S & Read M A, *Biopolymers*, 56 (2000) 195.
16 Patel D J, Phan A T & Kuryavyi V, *Nucleic Acids Res*, 35 (2007) 7429.
17 Ireson C R & Kelland L R, *Mol Cancer Ther*, 5 (2006) 2957.
18 Neidle S & Parkinson G, *Nat Rev Drug Discov*, 1 (2002) 383.
19 Siddiqui-Jain A, Grand C L, Bearss D J & Hurley L H, *Proc Natl Acad Sci USA*, 99 (2002) 11593.
20 Krauss I R, Merlino A, Giancola C, Randazzo A, Mazzarella L & Sica F, *Nucleic Acids Res*, 39 (2011) 7858.
21 Mukundan V T, Ngoc Quang D & Anh Tuan P, *Nucleic Acids Res*, 39 (2011) 8984.
22 Olaussen K A, Dubrana K, Dornont J, Spano J P, Sabatier L & Soria J C, *Crit Rev Oncol Hematol*, 57 (2006) 191.
23 Gray R D, Trent J O & Chaires J B, *J Mol Biol*, 426 (2014) 1629.
24 Liu J-q, Chen C-y, Xue Y, Hao Y-h & Tan Z, *J Am Chem Soc*, 132 (2010) 10521.
25 Hu X L, Li C, Feng C, Mao X X, Xiang Y & Li G X, *Chem Commun*, 53 (2017) 4692.
26 Kankia B I & Marky L A, *J Am Chem Soc*, 123 (2001) 10799.
27 Olsen C M, Gmeiner W H & Marky L A, *J Phys Chem B*, 110 (2006) 6962.
28 Gray R D & Chaires J B, *Nucleic Acids Res*, 36 (2008) 4191.
29 Gavathiotis E & Searle M S, *Org Biomol Chem*, 1 (2003) 1650.
30 Cavallari M, Calzolari A, Garbesi A & Di Felice R, *J Phys Chem B*, 110 (2006) 26337.
31 Stefl R, Cheatham T E, Spackova N, Fadrna E, Berger I, Koca J & Sponer J, *Biophys J*, 85 (2003) 1787.
32 Gray R D, Li J & Chaires J B, *J Phys Chem B*, 113 (2009) 2676.
33 Mashimo T, Yagi H, Sannohe Y, Rajendran A & Sugiyama H, *J Am Chem Soc*, 132 (2010) 14910.
34 Li W, Hou X-M, Wang P-Y, Xi X-G & Li M, *J Am Chem Soc*, 135 (2013) 6423.
35 Engelhart A E, Plavec J, Persil O & Hud N V, *Metal Ion Interactions With G-quadruplex Structures in Nucleic Acid-Metal Ion Interactions,* edited by N V Hud, (Royal Society of Chemistry:Cambridge, UK) 2008, p. 114.
36 Wang Y & Patel D J, *Structure*, 1 (1993) 263.
37 Dai J, Punchihewa C, Ambrus A, Chen D, Jones R A & Yang D, *Nucleic Acids Res*, 35 (2007) 2440.
38 Mahoney M W & Jorgensen W L, *J Chem Phys*, 112 (2000) 8910.
39 Jorgensen W L, *J Am Chem Soc*, 103 (1983) 335.
40 Bussi G, Donadio D & Parrinello M, *J Chem Phys*, 126 (2007) 014101.
41 Berendsen H J C, Postma J P M, van Gunsteren W F, DiNola A & Haak J R, *J Chem Phys*, 81 (1984) 3684.
42 Darden T, York D & Pedersen L, *J Chem Phys*, 98 (1993) 10089.
43 Hess B, Bekker H, Berendsen H J C & Fraaije J, *J Comput Chem*, 18 (1997) 1463.
44 Hess B, Kutzner C, van der Spoel D & Lindahl E, *J Chem Theory Comput*, 4 (2008) 435.
45 Cornell W D, Cieplak P, Bayly C I, Gould I R, Merz K M, Ferguson D M, Spellmeyer D C, Fox T, Caldwell J W & Kollman P A, *J Am Chem Soc*, 117 (1995) 5179.
46 Haider S M, Parkinson G N & Neidle S, *J Mol Biol*, 326 (2003) 117.
47 Haider S, Parkinson G N & Neidle S, *Biophys J*, 95 (2008) 296.
48 Chandran A, Ghoshdastidar D & Senapati S, *J Am Chem Soc*, 134 (2012) 20330.
49 Samanta P K, Periyasamy G, Manna A K & Pati S K, *J Mater Chem*, 22 (2012) 6774.
50 Chakraborty S, Sharma S, Maiti P K & Krishnan Y, *Nucleic Acids Res*, 37 (2009) 2810.
51 Humphrey W, Dalke A & Schulten K, *J Molec Graphics*, 14 (1996) 33.
52 Kaestner J, *WIREs Comput Molec Sci*, 1 (2011) 932.
53 Kumar S, Bouzida D, Swendsen R H, Kollman P A & Rosenberg J M, *J Comput Chem*, 13 (1992) 1011.
54 Souaille M & Roux B, *Comput Phys Commun*, 135 (2001) 40.
55 Gu J, Leszczynski J & Bansal M, *Chem Phys Lett*, 311 (1999) 209.
56 Meyer M, Steinke T, Brandl M & Sühnel J, *J Comput Chem*, 22 (2001) 109.
57 Cerofolini L, Amato J, Giachetti A, Limongelli V, Novellino E, Parrinello M, Fragai M, Randazzo A & Luchinat C, *Nucleic Acids Res*, 42 (2014) 13393.
58 Jeffrey G A, *An Introduction to Hydrogen Bonding*, (Oxford University Press, New York) 1997.
59 Lipsitz R S, Sharma Y, Brooks B R & Tjandra N, *J Am Chem Soc*, 124 (2002) 10621.