Structural dynamics of lateral and diagonal loops of Human telomeric G-quadruplexes in extended MD simulations

Barira Islam,[1] Petr Stadlbauer,[1] Marek Havrila,[1] Shozeb Haider[2] and Jiri Sponer[1]*

[1] Institute of Biophysics, Academy of Sciences of the Czech Republic, Královopolská 135, 612 65 Brno, Czech Republic

[2] UCL School of Pharmacy, 29-39 Brunswick Square, London WC1N 1AX, UK

* To whom correspondence should be addressed. Tel: 420 541 517 133; Fax: +420 541 212 179; Email: sponer@ncbr.muni.cz

KEYWORDS: G-quadruplex, lateral loop, diagonal loop, extended MD simulations, Human telomeric quadruplex

ABSTRACT

The NMR solution structures of human telomeric (Htel) G-quadruplexes (GQs) are characterized by the presence of two lateral loops complemented by ~~one~~ either diagonal or propeller loop. Bases of the loops can interact within the loops, between loops as well as with flanking bases to form base alignments above and below the GQ stems. These base alignments are known to affect the loop structures and relative stabilities of different Htel GQ folds. We have carried out a total of 217 μs of classical (unbiased) molecular dynamics (MD) simulations starting from the available solution structures of Htel GQs to characterize structural dynamics of the lateral and diagonal loops, using several recent AMBER DNA force-field variants. As the loops are involved in diverse stacking and H-bonding interactions, their dynamics is slow and extended sampling is required to capture different conformations. Nevertheless, although the simulations were far from being quantitatively converged, the data suggest that multiple 10 μs-scale simulations can provide a quite good assessment of the loop conformational space as described by the force field. The simulations suggest that the lateral loops may sample multiple co-existing conformations, which should be considered when comparing simulations with the NMR models as the latter include ensemble averaging. The adenine-thymine Watson Crick base pairing was the most stable base pairing in the simulations. Adenine-adenine and thymine-thymine base pairs were also sampled but were less stable. The data suggest that description of lateral and diagonal GQ loops in contemporary MD simulations is considerably more realistic than the description of propeller loops.

INTRODUCTION

Telomeres form the linear ends of eukaryotic chromosomes and protect them from end-to-end fusion and degradation.[1, 2] They comprise of short DNA repeats and specialized associated proteins.[3, 4] In normal human cells, telomeric DNA shortens at every cycle of replication, ultimately reaching the Hayflick limit and thus leading to replicative senescence.[4-7] The ribonucleoprotein telomerase provides a compensatory mechanism as it can preserve telomere length shortening by using its intrinsic RNA as the template.[3, 6, 8-11] Telomerase is inactive in normal cells but in 85% of all primary tumors, germ line cells, stem cells, hematopoietic cells and other rapidly dividing cells, high telomerase activity is observed.[8, 12-17] Therefore, telomerase inhibition is considered as an attractive selective therapy for cancer.[18-20] The 5' to 3' strand of human telomeric DNA (Htel) consists of tandemly repeated d(TTAGGG) sequence which extends beyond the duplex DNA producing a 3' single-stranded overhang of 100-280 nucleotides in length.[21, 22]

It has been shown by diverse experiments that guanine-rich sequences have unusual aggregation properties and can form four-stranded structures containing G-quartets, in a cyclic co-planar arrangement of four guanines stabilized by a Hoogsteen type hydrogen bonding.[23, 24] Formation of such secondary structures by G-rich sequences found at the chromosomal telomeres can prevent the telomerase binding and arrest cell division.[25] Therefore, G-rich telomeric sequence is considered an important target for small molecule drug design in cancer treatment.[26, 27] The guanine quartet is the basic structural unit of a G-quadruplex (GQ) and is formed when guanines utilize both Watson Crick (WC) and Hoogsteen edges to form square planar alignment.[28] Several G-quartets stack to form GQ stems (GSs) and cations between the quartets shield the electrostatic repulsion between the centrally facing carbonyl oxygen atoms of the guanines.[28, 29] The intervening sequences between the guanines (TTA in Htel GQs) form the loops of the GQ.[28]

Many DNA GQs are polymorphic in nature as their sequences can lead to different GQ topologies in different experimental conditions or even several folds can co-exist. The guanines in the quartet can adopt either -*syn* or -*anti* glycosidic orientation. The *syn-anti* distributions in the G-strands are interrelated with the overall topological variability of the GQs, because base-paired guanines belonging to parallel and antiparallel G-strands must have the same and opposite orientations, respectively.

The Htel GQ sequence possesses extreme topological variability (Table S1 in the Supporting Information). Till date, Htel GQs in crystallographic conditions have been observed as parallel-stranded GQ which is characterized by the presence of only propeller loops (PLs).[23] PLs link bases in the GSs that are not in the same quartet but share a groove (Figure S1a). They are also known as chain reversal loops as they orient the following guanine strand parallel to the preceding connected strand. In the solution conditions, in addition to parallel stranded, two- and three-quartet antiparallel, (2+2) antiparallel and (3+1) hybrid-1 and hybrid-2 topologies have been reported at the atomistic resolution (Figure 1).[30-36] For clarification, three-quartet antiparallel and (2+2) antiparallel GQ refers to antiparallel GQ with diagonal and propeller as the middle loop, respectively. Many more GQ folds may exist on the Htel folding landscape and those are yet to be visualized by experimental techniques.[37]

All the presently-known solution structures of Htel GQs are characterized by the presence of two lateral loops (LLs) and one either propeller (PL) or diagonal loop (DL). The PLs span medium grooves of GQ while LLs can span both wide and narrow grooves. The LLs link bases of the same quartet that share a hydrogen bond while the DLs link bases of the same quartet that do not share hydrogen bonds (Figure S1b and c in the Supporting Information).

Along with experimental techniques, molecular dynamics (MD) simulations have also been used to study the structural dynamics, ligand binding and folding/formation of GQs.[37-74] Obviously, the reliability of simulation studies of GQs depends on the completeness of sampling and specific accuracy of the force-field description of various aspects of the GQ structures. Several studies indicated that quality of the MD description of the quartet cores of the GQs and of the single-stranded loop regions differ. The accuracy of the force-field description of the single-stranded loops is in general inferior to the description of the GSs. The GSs and loops also represent a distinct sampling problem. In the MD simulations, cation-stabilized GSs are (with properly chosen solute force field) largely stable and show only some backbone substates and ion fluctuations. In contrast, the loop regions may sample numerous conformations with diverse lifetimes, depending on the specific interactions within the loops and with the other parts of the structures.[62-65, 71] It has also been shown that alignments formed by the loops along with the flanking bases may be important.[65, 71] With the increasing computational power as one can carry out multiple μs-scale simulations, the force field emerges as the major limitation in the GQ simulations.[64, 65, 71] For example, the following GQ-specific force-field limitations have been discussed in the literature: local imperfectness of the description of the ion-quartet interactions,[62, 75, 76] excessive ion-ion repulsion in the GS within the approximation of the pair additive force field,[77] lack of sampling of gamma *trans* conformations in PLs[62, 64, 71] and underestimated stability of the GG base pairing.[37, 76] In our previous analysis of Htel and *c-kit* promoter GQ PLs,[64, 65, 71] we noticed that the crystal-structure-like conformation of these loops could not be sampled in the simulations.[65, 71] Benchmark unbiased 1-10 μs MD simulations supplemented by some enhanced-sampling simulations revealed that it is presently possible to capture by MD the most significant structures on the free-energy landscapes of the PLs, as defined by the force

field.[71] However, the sampled structures indicated serious force-field imbalances in the description of the PLs.

In the present paper, we complement the preceding benchmark simulations on TTA PLs by analysis of LLs and DLs of the Htel DNA GQs. The sampling issue of these loops is complicated by the fact that both loop types can form alignments with bases of the other loops and with flanking bases, above and below the GS. As already noticed earlier,[65] the alignments formed by the bases may stabilize specific loop conformations, which are different in diverse folds. We show that some alignments native to the structures remained stable in simulations while few new were also formed during the simulations. In general, the results suggest that contemporary MD simulations can provide a rather satisfactory modelling of the TTA LLs, which appear to be considerably better described by the force field than the TTA PLs.

METHODS

Starting structures

We have simulated representatives of all LL- and DL-forming Htel GQ topologies (Figure 1 & Table 1), using the NMR structure coordinates for the three-quartet antiparallel (PDB id 143D),[36] two-quartet antiparallel (PDB id 2KF8),[32] antiparallel (2+2) (PDB id 2MBJ),[33] hybrid-1 (PDB id 2HY9),[31] and hybrid-2 (PDB id 2JPZ)[30] topologies. We also carried out simulations for hybrid-1 GQ topology using the NMR structure coordinates from PDB ids 2GKU and 2JSM.[34, 35] For more sampling of hybrid-2 topology, we carried out simulations using 5MVB without the ligand as the starting structure.[78] Note that 5MVB resembles 2JPZ in topology, sequence and loop succession. It has ligand stacked on the first quartet and due to this remodelling of second LL (LL2) bases have been observed. For all NMR structures, coordinates were taken from model 1.

Additional simulations were also carried out using model 4 of 143D as the LL2 of models 1 and 4 has different conformational arrangement of bases.[36] The results of 5MVB and model 4 of 143D simulations are discussed in separate section for clarity. The brominated base was replaced by corresponding unmodified bases in the starting structure derived from PDB id 2MBJ.[33]

Water and ionic conditions in standard simulations

Most simulations were carried out in 0.15 M excess NaCl. Two cations were manually placed between the quartets in the GSs. The solvent molecules and additional ions for simulations were added using *xleap* module of AMBER12. The system was first neutralized by Na^+ and then excess NaCl of 0.15 M concentration was further added to the system. SPC/E-adapted Joung and Cheatham parameters were used for Na^+ (radius 1.212 Å and well depth of 0.3526418 kcal mol^{-1}) and Cl^- ions (radius 2.711 Å and well depth 0.0127850 kcal mol^{-1}).[79] The systems were solvated in SPC/E water model and placed in truncated octahedral box with a minimal distance of 10 Å of solute from the box border.

As demonstrated in our earlier studies, these ionic conditions should be sufficient for the purpose of the present paper. Within the force-field approximation, the properties of the Na^+ ions inside the GSs should be somewhere in between real Na^+ and K^+ ions. Further, simulations of fully folded cation-stabilized GQs are known to be fairly insensitive to the chosen ion types, concentrations and water models; for a thorough discussion see the following review papers.[37, 75] In other words, uncertainties caused by the choice of a specific ion type and ion parametrization are expected to be smaller compared to the genuine sampling limitations of our investigations. To validate this further and for additional sampling, we carried out some simulations in excess

0.15 M KCl as well. Both SPC/E and TIP3P water models were tested with KCl to assess the simulation behaviour of GQ in different ions and water models.

DNA Force fields

We used three different variants of the AMBER DNA force field, starting from the parmbsc0 (bsc0)[80] modification of the Cornell *et al*.[81] force field supplemented by the χ_{OL4},[82] marked as bsc0χ_{OL4} throughout the paper. χ_{OL4} improves the behaviour of *syn* nucleotides of DNA GQs compared to simulations carried out with the bsc0 refinement alone and hence is essential in simulations of antiparallel and hybrid GQs. The χ_{OL4} refinement modulates the χ *anti*-region and facilitates transition through the 120º χ region by decreasing the energy barrier through this region and increasing for transitions through the 350º χ region.[82] Simulations were also carried out with $\varepsilon\zeta_{OL1}$[83] as a further refinement added to the bsc0χ_{OL4} force field, abbreviated as bsc0$\chi_{OL4}\varepsilon\zeta_{OL1}$. The $\varepsilon\zeta_{OL1}$ modifies the equilibrium of BI and BII B-DNA conformations thereby improving the BI/BII populations and helical twists.[83] Most simulation trajectories (cf. Table 1) were carried out with the latest OL15 version, i.e., upon adding the β_{OL1} modification[84] and completing the dihedral reparametrization of the Cornell *et al.* force field. For an overview of DNA force fields see ref.[85] Similar spectrum of AMBER force-field variants has been used in our preceding study of PLs,[71] which are structurally much more homogenous than the presently investigated loops and for which a more convincing sampling could have been achieved with similar amount of simulations as used in the present work. The results were essentially insensitive to the details of the dihedral parametrization. Therefore, we assume that the results presented in this work should also not be dramatically affected by the chosen force-field variants, as all of them are modern versions of the AMBER DNA force field and differ only in the tuning of dihedral potentials. If differences of the same loop simulated in different force field versions

are seen, they appear to primarily reflect the effect of random sampling rather than any systematic force-field dependence. This suggestion is consistent with the observed differences in simulations. In other words, we consider simulations of a given loop executed with different force-field variants as being comparable and of equal validity.

MD simulations

The starting structures were equilibrated using standard protocols described in the Supporting Information. The final MD simulations were performed with PMEMD CUDA version of AMBER12 or AMBER14.[86] The electrostatic interactions were calculated using Particle mesh Ewald method and periodic boundary conditions were used.[87] The cut-off distance for non-bonded Lennard-Jones interactions was set to 9 Å. Covalent bonds involving hydrogen atoms were constrained using SHAKE algorithm with a tolerance of 0.0001 Å, allowing us to use an integration time step of 2 fs. For simulations in AMBER14, hydrogen mass repartitioning of solute atoms was utilized and an integration time step of 4 fs was used. Berendsen weak coupling thermostat and barostat were used to maintain constant temperature and pressure of 300 K and 1 atm, respectively. The final production runs without restraints were carried out for continuous 5 or 10 μs and the frames were written at a time step of 10 ps. Analyses of trajectories performed using *cpptraj* module of AMBER and VMD and pymol programs were used for visualization.

Selected abbreviations

Human telomeric, G-quadruplex and G-stem are referred to as Htel, GQ and GS, respectively. The propeller, lateral and diagonal loops of the GQ are referred to as PL, LL and DL. The LLs are numbered such that the first and second LL starting from the 5'-end is called LL1 and LL2, respectively. The base pairing and stacking interactions are marked by symbols ':' and '|' in

between the bases. Majority of the simulations were carried out using NaCl and SPC/E water model (see above). However, when conditions other than these were used, they are mentioned specifically. List of all simulations is given in Table 1.

RESULTS

Classification of LLs in experimental structures of Htel GQs

All the solution structures of Htel GQs have two LLs which show a wide range of specific geometries. We have categorized them into four main types based on the arrangement of bases and their stacking interactions (Figure 2).

Type-1 LL is the most common TTA LL conformation class in Htel GQ experimental structures. The first thymine is exposed to the solvent and the second thymine and adenine are stacked above or below the GS, depending on the loop position, and interact with the other loops or flanking bases (Figure 2a and b). The LL2 of 143D (model 1), both the LLs of 2KF8, LL1 of 2MBJ, LL2 of hybrid-1 (PDB ids: 2HY9, 2GKU and 2JSM) and LL1 of hybrid-2 (PDB ids: 2JSL and 2JPZ) GQs belong to the type-1 LL class (Supporting Information Figure S2). In the 2KF8 structure, LL1 is longer with GTTA sequence but in this loop as well the first thymine is exposed to the solvent while the second thymine and adenine are stacked above the GQ. Therefore, we have categorized this LL also as type-1. In summary, type-1 structure has been identified for LLs spanning narrow grooves of all hybrid-1, three-quartet antiparallel, (2+2) antiparallel and two-quartet antiparallel GQs. It is also formed by LL spanning wide groove of two-quartet antiparallel and hybrid-2 GQs.

In type-2 LL, only the second thymine is stacked with the GS while the first thymine and adenine are exposed to the solvent (Figure 2c and d). The type-2 arrangement is formed by the LLs spanning the wide grooves of three-quartet antiparallel (LL1), (2+2) antiparallel (LL2) and LL1 of hybrid-1 GQ structures 2HY9 and 2GKU (Supporting Information Figure S3).

In the type-3 LL, the first thymine and adenine are stacked with the GS bases while the second thymine is exposed to the solvent. The LL1 (wide) of 2JSM and LL2 (narrow) of 2JSL are type-3 LL (Figure 2e and f). Note that the 2JSL structure has not been simulated.

In the type-4 LL, all the three loop bases are stacked with GS or the flanking bases aligned above or below the GS. The LL2 (wide) of 2JPZ is a type-4 LL (Figure 2g).

The backbone dihedral angles of some Htel solution NMR structures have significant variability within the models of the same GQ structure. Therefore, we did not attempt to identify the LL types based on the backbone dihedral angles. The NMR data might not allow a confident assignment of the details of the loop backbone conformations. NMR structures are ensembles of multiple molecular models and the position of flexible bases can be different in the models. We compared primary NMR data and PDB structures along with their respective behaviour in simulations and some noted ambiguities are presented in the Supporting Information.

Overall behaviour of LLs in simulations of the Htel GQs

As shown in Tables 1 and 2, we have accumulated series of forty eight 5-10 μs trajectories of the individual LLs in twenty four separate MD simulations initialized from altogether thirteen different starting LL geometries (the 5MVB and 143D model 4 trajectories will be analysed separately). As discussed below, hydrogen bonding and stacking interactions of LL bases

appeared to be the major determinants of the structural stability of the LLs in the simulations. Although RMSD should not be used for a quantitative assessment of the structural developments, let us first present overview of the RMSD, as it gives a useful qualitative insight into the sampling of the LL conformational space.

The overall backbone atom RMSDs of the whole GQs were broadly similar in the present simulations (Figures S4 and S5 in the Supporting Information), reflecting maintenance of stable GQ folds in all simulations, as expected.[37, 75]

Figure 3 and Figures S6-S9 in the Supporting Information summarize RMSD developments of all the individual LL trajectories. Altogether, the data shows that although some simulations keep very low RMSD with respect to the start, indicating that the starting structure is frozen during the course of the simulations, majority of the trajectories show significant dynamics. In some cases, the starting structures were lost irreversibly. In other cases, we observe reformations of the starting structures on variable time scales after a perturbation or loss of the starting structures.

For the LL type-1, some loops showed significant deviations from the starting structure while others remained close to the starting structures (Figure 3a and Figures S6, S8a and S9a in the Supporting Information). For the LL type-2, many simulations departed from the starting structures suggesting that the initial positions of loop bases were not stable in the simulations (Figure 3b and Figures S7, S8b and S9b in the Supporting Information). Note that in this loop type only the second thymine is stacked with the GS while the first thymine and adenine are exposed to the solvent. For the type-3 and type-4 LLs we have only two (Figure 3c and Figure S9c in the Supporting Information) and three trajectories (Figure 3d and Figures S8c and S9d in the Supporting Information), respectively.

We suggest that the data can be interpreted in the following way. First, we need to consider that the simulations correspond to a diverse set of thirteen LL starting structures embedded within diverse overall GQ folds and within different flanking base contexts, resulting in diverse possibilities to form additional interactions with other loops and flanking bases. Thus, the data for the individual LLs are certainly not converged and achieving a convergence for all the specific LLs would be extremely time-consuming. Nevertheless, it is evident that the LLs are capable to undergo structural rearrangements typically on μs time scale. Based on the simulations, and described in more detail below, we suggest that, when one needs to qualitatively characterize conformational landscape of a single specific LL, a series of ~10 μs standard simulations should be the minimal requirement. Consideration of several starting structures would be highly advisable to avoid initial trapping in one conformation.

LL behaviour in specific simulations

In majority of simulations, the loops maintained the initial loop type throughout the simulation (Table 2). It may reflect true stability of the initial arrangements as well bias due to the starting structure. To unambiguously differentiate between these two scenarios is not straightforward. In the following text, we will discuss the dynamics of those LLs that changed the LL type during the course of the simulation. In the Supporting Information, we provide a brief description of all the individual LL trajectories; note that in many simulations there has been visible structural dynamics without a change of the loop type.

LL type-1 is the most commonly observed loop arrangement in the Htel GQs. The first thymine of the loop is exposed to the solvent and in many cases stabilized by a hydrogen bond interaction between its O4 and hydrogen atoms of the GS guanine just preceding or following the loop. The

second thymine and adenine of the loop are stacked with the GS. Out of the twenty seven (fourteen 10 μs and thirteen 5 μs long) LL type-1 simulations, with all the force fields combined, change in loop type was observed in only three simulations (Table 2). In the simulation of 2KF8 in bsc0$\chi_{OL4}\varepsilon\zeta_{OL1}$ force field, the T4 of LL1 at ~2.8 μs moved to stack below the G3 of the same loop. T5 and A6 stacked below the T4 and GS, respectively (Figure 4a). The loop arrangement was thus similar to type-4 LL and was stable till the end of the simulation. Similarly, in the bsc0$\chi_{OL4}\varepsilon\zeta_{OL1}$ simulation of 2HY9, at ~ 0.9 μs T20 of LL2 oriented towards the solvent and stacked with T19 (Figure 4b). The T19|T20 stack was exposed to the solvent while A21 stacked below the GS. In the bsc0χ_{OL4} simulation of 2JPZ, LL1 changed to type-4 as the T7, T8 and A9 sequentially stacked below the G10 of third quartet. For the sake of completeness, in one of the simulations of 2MBJ in OL15 force field, LL1 fluctuated from type-1 to a new loop type when T8 from T8|A9 moved to stack with T7 in the solvent. Such conformation of LL has not been observed in the experimental structures of Htel GQs. Nevertheless, the loop re-attained type-1 conformation after 600 ns.

The type-2 LLs were very flexible in the simulations. This could be because in type-2 LL only second thymine is stacked with the GS bases and both the first thymine and adenine are exposed to the solvent which allowed flexibility to them. In the bsc0χ_{OL4} simulation of 2MBJ, A21 of LL2 at ~1.5 μs aligned in the same plane as T20, stacked over GS and formed a T20(O2)-A21(H8) bond. T19 remain exposed to the solvent and the LL thus changed from type-2 to type-1 and was then stable till the end of the simulation (Figure 4c). In the OL15 simulation of 2HY9, T8 moved to stack above A9 while T7 remained exposed to the solvent and attained type-1 like conformation.

The conformation of LL1 of 2GKU was changed in all the simulations (Table 2). In the bsc0$\chi_{OL4}\varepsilon\zeta_{OL1}$ simulation of 2GKU, at ~65 ns both the T12 and A14 of LL1 eventually moved to stack below the GS and formed a T12(O2)-A14(H62) hydrogen bond. T13 stacked below T12 and the loop thus attained stable type-4 structure. In the simulation of 2GKU with K^+ and SPC/E water model in OL15, early in the simulation A14 and T13 stacked below the GS and the loop attained type-1 conformation. At ~2.3 μs, T12 also stacked below the GS and the loop finally attained type-4 conformation which was stable till the end (5 μs) of the simulation (Figure 4d). However, in the equivalent simulation of 2GKU in TIP3P water model, only A14 moved to stack with GS along with T13. T12 remained exposed to the solvent and LL1 remained in the newly adopted type-1 conformation. In general, type-2 LLs have a visible tendency to convert to either type-1 or type-4 conformations.

The LL1 of 2JSM represents type-3 conformation. In the bsc0$\chi_{OL4}\varepsilon\zeta_{OL1}$ simulation, LL1 of 2JSM changed from type-3 to type-4 in the equilibration stage as the T13 stacked along GS axis below the T12. T12 and A14 stacked with guanines of the third quartet, G11 and G15, respectively. In the OL15 simulation as well the LL1 of 2JSM changed to type-4 conformation as T12 and A14 stacked below GS and T13 stacked below T12 (Figure 4e).

The type-4 structure was represented by LL2 of 2JPZ and it was stable in two out of the three simulations. In the starting structure, T13 and A15 of the loop stack on the GS while the T14 of the loop is stacked over T2 which is stacked on GS. This structure was maintained in the bsc0χ_{OL4} and bsc0$\chi_{OL4}\varepsilon\zeta_{OL1}$ simulations. In the OL15 simulation, T13 moved to the solvent in first 10 ns and loop attained type-1 confirmation which lasted till the end of the simulation (Figure 4f).

In summary, acknowledging the limits of sampling, simulations tentatively suggest that the force-field description supports clearly the most common experimental type-1 LL arrangement as well as the rare type-4 arrangement. The common type-2 and rare type-3 arrangements appear to be less stable since they have been often lost in simulations.

Behaviour of LLs in simulations is biased by the starting structures and multiple simulations are required to study its dynamics

To further test to what extent is the LL behaviour in simulations biased by the starting conformation, we carried out series of simulations with two different starting structures. In the first one, model 4 of PDB 143D was used as simulation start rather than model 1. LL2 in model 1 is in the type-1 conformation but in model 4 both the thymines of LL2 are exposed to the solvent and only its adenine stacks below the GS. This conformation does not correspond to any established loop type. In the 3-5 μs long simulations of 143D model 4, LL2 could attain 143D model 1 like conformation (type-1) in four out of five simulations (Figure S10). In three simulations, type-1 conformation was attained within 200 ns and in one simulation it was attained at ~2.3 μs. In the simulation when type-1 conformation was unattained, both the thymines of the LL2 remained exposed to the solvent and the adenine also moved towards the solvent. LL2 did not sample type-1 conformation in this 5 μs long simulation and the final conformation still did not correspond to any of the established loop types (Figure S10f). Nevertheless, the simulations show a visible tendency to move towards the type-1 LL arrangement of LL2.

In the second series of simulations, hybrid-2 structure 5MVB without the ligand was used as the starting structure. In 2JPZ, all the three bases of LL2 are stacked on the GS but in 5MVB the

LL2 bases could not stack on the GS due to the presence of the ligand. We carried out six 3-5 μs long simulations in this series. The unstructured LL2 of 5MVB transitioned to different classified LL types in independent simulations. In one simulation LL2 first attained type-4 conformation similar to 2JPZ at 50 ns and then shifted to type-1 at 200 ns and further to type-2 at 700 ns which lasted till the end of the 3 μs long simulation (Figure S11b). In three simulations, type-1 was the main conformation sampled during the simulation (Figure S11c, e and f). In one simulation, LL2 was trapped in the type-3 conformation throughout the simulation (Figure S11d). In one 5 μs long simulation, T13 and T14 stacked on the GS but A15 remained aligned in the groove and was rather trapped by A15(H61)-G11(O3') hydrogen bond (Figure S11g). The simulation thus moved in the direction of the type-4 conformation seen in the 2JPZ structure, but full transition appeared to be obstructed by the above-noted hydrogen bond. LL1 of 5MVB, which was not remodelled by the ligand behaved similarly to the LL1 of 2JPZ in all the simulations. Further details of the simulations are presented in the Supporting Information.

Structural dynamics of DLs of Htel GQs

The GQs 143D and 2KF8 have DL as the middle loop. In 143D, T11 is exposed to the solvent while T12 and A13 are arranged on the GS (Figure 5a). The 5'-flanking base A1 is aligned over G2 of the first quartet, the A13 is stacked over G22 of the first quartet and A1:A13 form *cis* Hoogsteen-Hoogsteen base pair. In all the simulations, T11|T12 stack was formed early and was sampled for majority of the simulation time (Figure 5b). The alignment or stacking of T11|T12 over A1 or A13 dominated the subsequent dynamics of the DL.

In the bsc0χ_{OL4} simulation of 143D, A1 and A13 rearranged early in the simulation to form *trans* WC-Hoogsteen base pair that was stable till 4.2 μs (Figure S12a). Also, T11|T12|A1 stack was

observed till ~4.2 μs and then T11|T12 moved to stack over A13. The T11|T12|A13 stack was stable till the end of the simulation.

The DL bases of 143D were more flexible in the simulation in bsc0$\chi_{OL4}\varepsilon\zeta_{OL1}$ force field where the A1:A13 base pair was not sampled. The T11|T12 stack arranged perpendicularly to the GS and their backbone hydrogen atoms formed CH/π interaction and hydrogen bonds with A13 (Figure S13). This affected the alignment of A13 and it could not stack on G22 but rather moved along with T11|T12 (Figure S13). The backbone interactions locked the loop in one conformation till the end of the simulation (Figure S14). The dynamics of 143D DL in bsc0χ_{OL4} and bsc0$\chi_{OL4}\varepsilon\zeta_{OL1}$ simulations is shown in Figure S15a and b, respectively.

In the 5 μs long simulation of 143D in OL15, T12 formed an A1(H62)-T12(OP2) hydrogen bond and CH/π interaction with A1 through its CH_3 group up to 2.2 μs of the simulation (Figure S16). This fixed the position of T11|T12 perpendicularly to the GS. A1 and A13 stacked on G2 and G22 during this time but no hydrogen bond interaction was observed between them. T11|T12 then showed some fluctuations as A1 moved to stack over G14. A1:A13 sampled *cis* WC/sugar edge pair from 2.6 to 4.6 μs (Figure S12b). T11|T12 could stack on either A13 or A1 such that the -CH_3 of T12 was oriented towards the GS. The dynamics of DL in this simulation is shown in Figure 5b.

The DL of 143D sampled similar dynamics in the simulations with K^+ as the stabilizing ions. In the simulation with SPC/E and K^+, T11|T12 stacked on A13 early in the simulation to form stable T11|T12|A13 structure resembling an untwisted ladder formed by partial stacking of bases. A1:A13 formed *cis* WC/sugar edge interactions in the simulation (Figure S12c). The T11|T12|A13 structure was also established in the simulation with TIP3P and K^+. A1:A13

formed *cis* Hoogsteen/Hoogsteen base pair in this simulation similar to the starting structure (Figure S12d). Thus, the dominant interactions in all the simulations were broadly similar but not identical.

The DL of 2KF8 is a five-nucleotide loop with GTTAG sequence. Both the guanines of the loop, G9 and G13 are aligned over the first quartet to form a base triple with G21 (Figure 5c). T10 is exposed to the solvent while A12|T11 stack over guanine (G9) of the loop. T11 also forms a symmetric base pair with 3'-thymine, T22 (Figure 5c).

In the bsc0χ_{OL4} force field, all the diagonal loop alignments, as in the starting structure, were maintained throughout the simulation. T11 and T22 base pair in the starting structure is formed by mutual hydrogen bond interactions between O4 and H3 atoms of both the thymines. However, in the bsc0χ_{OL4} simulation, different interactions, T11(O4)-T22(H1') and T11(H3)-T22(O2) formed the T11:T22 pair. The positions of G9, T11, A12 and G13 were stabilized by various interactions. Only the solvent exposed first thymine (T10) showed some flexibility (Figure S15c).

In the bsc0$\chi_{OL4}\varepsilon\zeta_{OL1}$ simulation, T11 and T22 interacted by one or more hydrogen bonds till 5.4 μs after which A12 moved to stack between the two thymines such that T11|A12| T22 stack was formed. This stack lasted till 6.9 μs and then again a hydrogen bond between T11 and T22 was formed which lasted till 9.7 μs. The DL of 2KF8 was slightly more flexible in bsc0$\chi_{OL4}\varepsilon\zeta_{OL1}$ simulation (Figure S15d). In the simulation of 2KF8 with OL15, T10 remained exposed to the solvent while T11 and A12 stacked on G9 throughout the simulation (Figure 5d). The T11:T22 base pair was maintained in the simulation.

Thymine:Adenine (TA) base pairs of the loops were typically stable in the simulations

In the solution structures of Htel GQs, many base alignments are present above and below the GSs (Table 3). TA is the most common base pair formed by the interaction of loops and flanking bases, usually by WC base pairing (Figure 6a). Such base pair is observed between T6 of LL1 and A19 of LL2 in the 143D and was stable in all the force fields (Figure S17). In the 2KF8, no stable TA base pairing was observed in any of the force fields. In the 2MBJ, A3:T20 WC base pair and A21 of the LL2 align over the first quartet. A21 formed a hydrogen bond with T20 and A3:T20:A21 triple was sampled in the bsc0χ_{OL4} simulation while A3:T20 WC base pair was sampled in all the other simulations (Figure S18). In the starting structure of 2MBJ, A15:A9:T25 triple is present below the GS. This triple was also sampled in the bsc0χ_{OL4} simulation (Figure S19a). Also, T8 of LL1 formed *trans* WC base pair with the flanking 3'-terminal adenine (A27) below the triple in bsc0χ_{OL4} simulation (Figure 6b, Figure S19b). In all the other simulations, A15 did not interact with either A9 or T25 and only A9:T25 WC base pair was sampled (Figure S19c-f). A27 interacted with sugar hydrogen of T25 in OL15 simulation in K^+ and SPC/E water model (Figure S19e).

In the 2HY9, A3:A9:A21 triple is present on the top of the GS. In all the simulations A9 moved away and the native A3:A21 base pair between 5'-flanking base A3 and adenine of LL2 was sampled (Figure S20). In 2HY9, along with the WC and *trans* WC base pairing, Hoogsteen base pairing was also observed in the bsc0χ_{OL4} and OL15 simulations as new TA interaction between T20 of LL2 and the flanking base A2 was formed above the GS (Figure 6c and Figure S21a and b in the Supporting information). T14 of LL1 and 3'-flanking base A25 form *trans* WC base pair below the GS which was mostly stable in bsc0χ_{OL4} and bsc0$\chi_{OL4}\varepsilon\zeta_{OL1}$ but not in the OL15 simulation (Figure S21).

Similarly, 5'-terminal thymine T1 forms WC base pair with A20 of LL2 above the first quartet in the 2GKU. This base pair was also stable in all the simulations (Figure S22a, c, e in the Supporting Information). In the 2GKU, T13 of LL1 forms *trans* WC base pair with 3'-flanking base A24 but in the bsc0$\chi_{OL4}\varepsilon\zeta_{OL1}$ and one of the OL15 simulations instead of T13:A24, T12:A24 base pair was sampled (Figure S22b, d). In the OL15 simulation with K^+ and TIP3P water model, native T13:A24 base pair was sampled (Figure S22f).

The native T1:A20 WC and T12:A14 Hoogsteen base pairings of 2JSM were stable in the simulations (Figure S23). The WC base pairing between A20 of LL2 and 5' base T1 was mostly stable in both the OL15 and bsc0$\chi_{OL4}\varepsilon\zeta_{OL1}$ simulations (Figure S23a and c). In the bsc0$\chi_{OL4}\varepsilon\zeta_{OL1}$ simulation of 2JSM, T12 and A14 Hoogsteen base pair in LL1 was stable till 3.6 μs of the simulation. The nucleotides showed some fluctuations between 3.6-4.5 μs and then T12(O2)-A14(H61) and A14(N1)-T12(H1') hydrogen bonds were formed which lasted till the end of the simulations (Figure S23b). In the OL15 simulation as well, native T12:A14 lasted till 3.8 μs and then interactions similar as in bsc0$\chi_{OL4}\varepsilon\zeta_{OL1}$ simulations were formed (Figure S23d).

In the 2JPZ, base triples are present both above and below the GS. At the top of the GS, the triple is formed by 5'-flanking base A3 and T13 and A15 of LL2. This triple was mostly stable in bsc0χ_{OL4} (Figure S23 a and b) and bsc0$\chi_{OL4}\varepsilon\zeta_{OL1}$ simulation (Figure S23d) but in the OL15 simulation T13 moved away and only A3:A15 base pair was sampled (Figure S24f). Below the GS, T8 and A9 of LL1 and flanking base T25 form a TAT triple. In the bsc0χ_{OL4} simulation of 2JPZ, this triple was not stable and A9 and T26 formed WC base pair at 2.5 μs of the simulation (Figure S24c). Once formed, this base pair was stable throughout the simulation. The triple was sampled in the bsc0$\chi_{OL4}\varepsilon\zeta_{OL1}$ and OL15 simulations (Figure S24e, g).

TT interactions were not very stable in the simulations

TT base pairing between the thymines of two loops or flanking bases are less common than the TA base pairs in the GQ experimental structures. Consistent behaviour was observed in MD simulations. TT base pairing is present above the GS in the 2KF8 between second thymine of the diagonal loop (T11) and the 3'-thymine (T22) (Figure 7a). In the 2JPZ, LL1 base T8 and 3'-flanking base T25 formed a base pair as a part of triple below the GS (Figure 7b). In the 2GKU and 2JSL, single hydrogen bond was observed in between LL thymine and flanking thymine bases (Figure 7c and d).

In the simulations of 2KF8, T11 and T22 base pairing was maintained in the bsc0χ_{OL4} force field but was intermittent in the bsc0$\chi_{OL4}\varepsilon\zeta_{OL1}$ and OL15 simulations (Figure S25). In the simulations of this GQ, below the GS, thymines of LL1 and LL2 also formed base pair. In the bsc0χ_{OL4} and OL15 simulations, T5 and T17 formed base pair while in bsc0$\chi_{OL4}\varepsilon\zeta_{OL1}$ it was formed by T4 and T17 (Figure S26). In 2MBJ, T8 and T26 interact by a single intermittent hydrogen bond (Figure S27). In 2GKU, T2(O4) and T19(H3) form a single hydrogen bond in the starting structure which was sampled in a similar way in all three simulations (Figure S28). In the bsc0$\chi_{OL4}\varepsilon\zeta_{OL1}$ simulation of 2JPZ, T8 and T25 formed hydrogen bonds, T8(H3):T25(O2) and T25(O2):T25(H3) but these were intermittent and not very stable. Additional T2:T14 interaction was sampled in the OL15 simulation (Figure S29). No TT interactions were observed in the simulations of 143D, 2HY9 and 2JSM.

DISCUSSION AND CONCLUSIONS

Lateral and diagonal loops represent common structural features of GQs and it is thus important to understand their structural dynamics in GQs as well as the reliability of their modelling in MD

simulation studies. We have analysed behaviour of LLs and DLs of the Htel GQs in a series of 5-10 μs explicit solvent MD simulations, complementing our earlier analysis of the propeller GQ loops.[71] Before discussing the main outcome of our benchmark simulations, let us mention several limitations that unavoidably affect our study. These limitations in many cases preclude us to reach unambiguous conclusions.

In the studies of LLs, the only experimental data that we have is a set of solution NMR studies. This contrasts the numerous X-ray structures available for the TTA PLs. Although, in principle, solution data should be more suitable than X-ray data to study the loop conformations due to absence of crystal packing, we also need to take into consideration that some of the loop conformations may be underdetermined due to paucity of the primary NMR data. Thus, it is fair to expect that the available experimental structures can suffer from some uncertainties. Further, the observed LLs belong to different GQ folds possessing also different 3' and 5' flanking bases, creating structure-specific contexts for the individual studied cases of LLs. In other words, the LLs are involved in omnipresent and variable molecular interactions, and it is not straightforward to suggest any reference LL conformation that could be used as a clear benchmark. As indicated by the MD simulations, at least some of the loops may in reality sample coexisting conformations which would be difficult to capture by the NOE-based ensemble-averaged NMR experiments unambiguously. For the DLs, we have an even more limited amount of experimental information to draw any firm conclusions about their conformational properties.

On the other side, also the simulation technique has its genuine limitations. Although we suggest (see below) that the description of the LLs by the force field may be quite realistic, it certainly is not perfect. The second limitation is the simulation time scale. Although we have seen a number of structural transitions in our simulations, they are certainly not converged. Due to the above-

noted limits of the experimental data and of the force-field description, we decided to not push the simulations to longer time scale. Likewise, we did not use any enhanced sampling technique. It is worth mentioning that in our previous study on PL[71] we employed the RECT[88] enhanced sampling method but its efficiency was not fully convincing. Sampling of the LLs is complicated by the numerous molecular interactions described throughout this work.

Despite all the limitations, our study allows several conclusions. It is evident, that the force-field description of the LLs is considerably more realistic than in case of the TTA PLs, for which the experimentally observed geometries were essentially entirely un-sampled.[71] For the LLs, many simulations neatly keep key aspects of the starting structures and even when departing from the starting structures, they sample geometries that are broadly consistent with the experimentally suggested LL conformations. We do not suggest that the MD description of the LLs is flawless, but certainly MD simulations can be used to provide qualified predictions of potential LL interactions for example in the GQ – ligand complexes.

The simulation behaviour of LLs was visibly dependent on their starting structures and interactions with other loops, with flanking bases, with the GSs and within the loops themselves. The AT WC interactions (or other AT base pairs) likely play a dominant role in shaping up specific LL conformations. These interactions have been observed in nearly all of the solution structures of Htel GQs and were mostly stable in the simulations (Table 3). The adenine-adenine and thymine-thymine base pairings were also sampled in the simulations but they were not as stable and common as AT base pairs.

The available structural experimental data for DLs are quite scarce. This also limited our capability to characterize the DL by simulations. Nevertheless, we hypothesise that the

suggestions and conclusions presented above for simulating the LLs should be well transferable to the Htel DLs, as their simulations typically sampled geometries that in many aspects resembled those suggested in the experimental solution structures.

All the topologies of Htel GQs other than the parallel-stranded GQ have two LLs. These loops are important for the process of folding of GQs as antiparallel hairpins compatible with LLs are one of the most easily folded structural elements in the Htel GQ folding landscape.[69, 89] These antiparallel hairpins can act as a scaffold as they can readily participate in the further folding of the GQ. It has been noted in our previous studies of Htel hairpins that interactions within the loop, such as stacking of the second thymine with adenine of the loop, stabilize the conformation of LLs and antiparallel hairpin as a whole.[69] Such stacking interactions have been widely observed in our current simulations and it has been noted that LLs tend to maximise such interactions. Specific molecular interactions involving the LLs may contribute to the rules determining relative stabilities of different GQ topologies. Thus using extensive sampling, our study provides an overview or a catalogue of interactions that could stabilize the LL conformations of the GQs, some of which have not yet been observed in the experimental structures.

In conclusion, our study shows that force-field description of Htel LLs and likely also of DLs is considerably more realistic than description of the PLs. It does not mean that the force field is quantitatively accurate regarding its ability to correctly predict relative stabilities of different competing substates and interactions on the LL free-energy landscapes which appear to be quite rich. Nevertheless, structures closely resembling those suggested by experiments are readily sampled. The simulations clearly support type-1 and type-4 LL arrangements. On the other hand, it appears that representative sampling of the free-energy landscape of LLs (as described by the

force field) is significantly hampered by their interactions with other elements of the GQ structures. Based on the simulations we suggest that to achieve a qualitative characterization of conformational landscape of a single specific LL, i.e., to identify the main conformations on the landscape, an extended series of ~10+ μs standard simulations should be the minimal requirement. It should be noted that in the present work we did not attempt to reach that level of sampling for any specific Htel LL, as we intentionally studied their full spectrum. When investigating a specific LL, consideration of several starting structures would be highly advisable to avoid initial trapping in one conformation. This should be taken into account when assessing the literature data and when designing further studies of GQ loops, since conclusions are often derived from much shorter simulation time scales; to the best of our knowledge, none of the available studies is anywhere close in sampling to this work. We intentionally did not apply any enhanced sampling techniques, since their application may be complicated by multidimensionality of the LL intrinsic conformational space and variability of the context-dependent interactions that affect the LL conformations. This could make utilization of collective-variable-based methods challenging.

SUPPORTING INFORMATION

The method of equilibration, comparison of MD simulations with primary structures and description of individual MD trajectories is presented in the Supporting Information. Supporting results are also presented in Figures S1-S29 and Table S1. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author

*Prof. Jiri Sponer. Tel: +420541517133; Fax: +420541212179; Email: sponer@ncbr.muni.cz

ACKNOWLEDGMENTS

This work was supported by grant 16-13721S from the Czech Science Foundation and by the project SYMBIT reg. number: CZ.02.1.01/0.0/0.0/15_003/0000477 financed by the ERDF. JS acknowledges support by Praemium Academiae. B.I. and P.S. greatly appreciate access to storage facilities owned by parties and projects contributing to the National Grid Infrastructure MetaCentrum, provided under the programme "Projects of Large Research, Development, and Innovations Infrastructures" (CESNET LM2015042).

REFERENCES

1. O'Sullivan, R. J.; Karlseder, J., Telomeres: Protecting Chromosomes Against Genome Instability. *Nat. Rev. Mol. Cell Biol.* **2010**, *11*, 171-181.

2. Bailey, S. M.; Murnane, J. P., Telomeres, Chromosome Instability and Cancer. *Nucleic Acids Res.* **2006**, *34*, 2408-2417.

3. Lewis, Karen A.; Wuttke, Deborah S., Telomerase and Telomere-Associated Proteins: Structural Insights into Mechanism and Evolution. *Structure* **2012**, *20*, 28-39.

4. Greider, C. W., Telomeres and Senescence: The History, the Experiment, the Future. *Curr. Biol.* **1998**, *8*, R178-R181.

5. Zhao, Y.; Hoshiyama, H.; Shay, J. W.; Wright, W. E., Quantitative Telomeric Overhang Determination Using a Double-Strand Specific Nuclease. *Nucleic Acids Res.* **2008**, *36*, e14.

6. Shay, J. W.; Wright, W. E., Role of Telomeres and Telomerase in Cancer. *Semin Cancer Biol.* **2011**, *21*, 349-353.

7. Kiyono, T.; Foster, S. A.; Koop, J. I.; McDougall, J. K.; Galloway, D. A.; Klingelhutz, A. J., Both Rb/p16INK4a Inactivation and Telomerase Activity are Required to Immortalize Human Epithelial Cells. *Nature* **1998**, *396*, 84-88.

8. Blackburn, E. H., Telomeres and Telomerase: Their Mechanisms of Action and the Effects of Altering Their Functions. *FEBS Letters* **2005**, *579*, 859-862.

9. Greider, C. W.; Blackburn, E. H., The Telomere Terminal Transferase of Tetrahymena is a Ribonucleoprotein Enzyme with Two Kinds of Primer Specificity. *Cell* **1987**, *51*, 887-898.

10. Morin, G. B., The Human Telomere Terminal Transferase Enzyme is a Ribonucleoprotein That Synthesizes TTAGGG Repeats. *Cell* **1989**, *59*, 521-529.

11. Schmidt, J. C.; Cech, T. R., Human Telomerase: Biogenesis, Trafficking, Recruitment, and Activation. *Genes Dev.* **2015**, *29*, 1095-1105.

12. Pech, M. F.; Garbuzov, A.; Hasegawa, K.; Sukhwani, M.; Zhang, R. J.; Benayoun, B. A.; Brockman, S. A.; Lin, S.; Brunet, A.; Orwig, K. E.; Artandi, S. E., High Telomerase is a Hallmark of Undifferentiated Spermatogonia and is Required for Maintenance of Male Germline Stem Cells. *Genes Dev.* **2015**, *29*, 2420-2434.

13. Armstrong, L.; Saretzki, G.; Peters, H.; Wappler, I.; Evans, J.; Hole, N.; von Zglinicki, T.; Lako, M., Overexpression of Telomerase Confers Growth Advantage, Stress Resistance, and Enhanced Differentiation of ESCs Toward the Hematopoietic Lineage. *Stem cells* **2005**, *23*, 516-529.

14. Marion, R. M.; Strati, K.; Li, H.; Tejera, A.; Schoeftner, S.; Ortega, S.; Serrano, M.; Blasco, M. A., Telomeres Acquire Embryonic Stem Cell Characteristics in Induced Pluripotent Stem Cells. *Cell stem cell* **2009**, *4*, 141-154.

15. Artandi, S. E.; DePinho, R. A., Telomeres and Telomerase in Cancer. *Carcinogenesis* **2010**, *31*, 9-18.

16. Bernardes de Jesus, B.; Blasco, M. A., Telomerase at the Intersection of Cancer and Aging. *Trends Genet.* **2013**, *29*, 513-520.

17. Shay, J. W., Role of Telomeres and Telomerase in Aging and Cancer. *Cancer Disc* **2016**, *6*, 584-593.

18. Chen, H.; Li, Y.; Tollefsbol, T. O., Strategies Targeting Telomerase Inhibition. *Mol. Biotechnol.* **2009**, *41*, 194-199.

19. Damm, K.; Hemmann, U.; Garin-Chesa, P.; Hauel, N.; Kauffmann, I.; Priepke, H.; Niestroj, C.; Daiber, C.; Enenkel, B.; Guilliard, B.; Lauritsch, I.; Muller, E.; Pascolo, E.; Sauter, G.; Pantic, M.; Martens, U. M.; Wenz, C.; Lingner, J.; Kraut, N.; Rettig, W. J.; Schnapp, A., A Highly Selective Telomerase Inhibitor Limiting Human Cancer Cell Proliferation. *EMBO J.* **2001**, *20*, 6958-6968.

20. Qi, Z.; Mi, R., Inhibition of Human Telomerase Reverse Transcriptase in Vivo and in Vitro for Retroviral Vector-Based Antisense Oligonucleotide Therapy in Ovarian Cancer. *Cancer Gene Ther.* **2016**, *23*, 36-42.

21. Wright, W. E.; Tesmer, V. M.; Huffman, K. E.; Levene, S. D.; Shay, J. W., Normal Human Chromosomes Have Long G-Rich Telomeric Overhangs at One End. *Genes Dev.* **1997**, *11*, 2801-2809.

22. Makarov, V. L.; Hirose, Y.; Langmore, J. P., Long G Tails at Both Ends of Human Chromosomes Suggest a C Strand Degradation Mechanism for Telomere Shortening. *Cell* **1997**, *88*, 657-666.

23. Parkinson, G. N.; Lee, M. P. H.; Neidle, S., Crystal Structure of Parallel Quadruplexes from Human Telomeric DNA. *Nature* **2002**, *417*, 876-880.

24. Gellert, M.; Lipsett, M. N.; Davies, D. R., Helix Formation by Guanylic Acid. *Proc. Natl. Acad. Sci. USA* **1962**, *48*, 2013-2018.

25. Moye, A. L.; Porter, K. C.; Cohen, S. B.; Phan, T.; Zyner, K. G.; Sasaki, N.; Lovrecz, G. O.; Beck, J. L.; Bryan, T. M., Telomeric G-quadruplexes are a Substrate and Site of Localization for Human Telomerase. *Nat. Commun.* **2015**, *6*, 7643.

26. Riou, J. F.; Guittat, L.; Mailliet, P.; Laoui, A.; Renou, E.; Petitgenet, O.; Megnin-Chanet, F.; Helene, C.; Mergny, J. L., Cell Senescence and Telomere Shortening Induced by a New Series of Specific G-quadruplex DNA Ligands. *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 2672-2677.

27. Han, H.; Hurley, L. H., G-quadruplex DNA: A Potential Target for Anti-cancer Drug Design. *Trends Pharmacol. Sci. 21*, 136-142.

28. Burge, S.; Parkinson, G. N.; Hazel, P.; Todd, A. K.; Neidle, S., Quadruplex DNA: Sequence, Topology and Structure. *Nucleic Acids Res.* **2006**, *34*, 5402-5415.

29. Neidle, S.; Balasubramanian, S. Fundamentals of Quadruplex Structures, In *Quadruplex Nucleic Acids*; Neidle, S.; Balasubramanian, S., Eds.; RSC:London, 2006, 1-30.

30. Dai, J.; Carver, M.; Punchihewa, C.; Jones, R. A.; Yang, D., Structure of the Hybrid-2 Type Intramolecular Human Telomeric G-quadruplex in K^+ Solution: Insights into Structure Polymorphism of the Human Telomeric Sequence. *Nucleic Acids Res.* **2007**, *35*, 4927-4940.

31. Dai, J.; Punchihewa, C.; Ambrus, A.; Chen, D.; Jones, R. A.; Yang, D., Structure of the Intramolecular Human Telomeric G-quadruplex in Potassium Solution: A Novel Adenine Triple Formation. *Nucleic Acids Res.* **2007**, *35*, 2440-2450.

32. Lim, K. W.; Amrane, S.; Bouaziz, S.; Xu, W.; Mu, Y.; Patel, D. J.; Luu, K. N.; Phan, A. T., Structure of the Human Telomere in K^+ Solution: A Stable Basket-Type G-quadruplex with Only Two G-tetrad Layers. *J. Am. Chem. Soc.* **2009**, *131*, 4301-4309.

33. Lim, K. W.; Ng, V. C.; Martin-Pintado, N.; Heddi, B.; Phan, A. T., Structure of the Human Telomere in Na^+ Solution: An Antiparallel (2+2) G-quadruplex Scaffold Reveals Additional Diversity. *Nucleic Acids Res.* **2013**, *41*, 10556-10562.

34. Luu, K. N.; Phan, A. T.; Kuryavyi, V.; Lacroix, L.; Patel, D. J., Structure of the Human Telomere in K^+ Solution: An Intramolecular (3+1) G-quadruplex Scaffold. *J. Am. Chem. Soc.* **2006**, *128*, 9963-9970.

35. Phan, A. T.; Kuryavyi, V.; Luu, K. N.; Patel, D. J., Structure of Two Intramolecular G-quadruplexes Formed by Natural Human Telomere Sequences in K^+ Solution. *Nucleic Acids Res.* **2007**, *35*, 6517-6525.

36. Wang, Y.; Patel, D. J., Solution Structure of the Human Telomeric Repeat $d[AG_3(T_2AG_3)_3]$ G-tetraplex. *Structure* **1993**, *1*, 263-282.

37. Šponer, J.; Bussi, G.; Stadlbauer, P.; Kührová, P.; Banáš, P.; Islam, B.; Haider, S.; Neidle, S.; Otyepka, M., Folding of Guanine Quadruplex Molecules–Funnel-like Mechanism or Kinetic Partitioning? An Overview from MD Simulation Studies. *Biochim. Biophys. Acta* **2017**, *1861*, 1246-1263.

38. Lech, C. J.; Heddi, B.; Phan, A. T., Guanine Base Stacking in G-quadruplex Nucleic Acids. *Nucleic Acids Res.* **2013**, *41*, 2034-2046.

39. Zhu, H.; Xiao, S.; Wang, L.; Liang, H., Communication: Asymmetrical Cation Movements Through G-quadruplex DNA. *J. Chem. Phys.* **2014**, *141*, 041103.

40. Wang, Z.; Liu, J.-P., Effects of the Central Potassium Ions on the G-quadruplex and Stabilizer Binding. *J. Mol. Graph. Model.* **2017**, *72*, 168-177.

41. Jarmila, H.; K., T. A.; A., P. J.; Stephen, N., Small-Molecule G-quadruplex Interactions: Systematic Exploration of Conformational Space Using Multiple Molecular Dynamics. *Biopolymers* **2013**, *99*, 989-1005.

42. Siebenmorgen, T.; Zacharias, M., Origin of Ion Specificity of Telomeric DNA G-Quadruplexes Investigated by Free-Energy Simulations. *Biophys. J.* **2017**, *112*, 2280-2290.

43. Deng, N.; Wickstrom, L.; Cieplak, P.; Lin, C.; Yang, D., Resolving the Ligand-Binding Specificity in c-MYC G-quadruplex DNA: Absolute Binding Free Energy Calculations and SPR Experiment. *J. Phys. Chem. B* **2017**, *121*, 10484-10497.

44. Mulholland, K.; Siddiquei, F.; Wu, C., Binding Modes and Pathway of Rhps4 to Human Telomeric G-quadruplex and Duplex DNA Probed by All-Atom Molecular Dynamics Simulations with Explicit Solvent. *Phys. Chem. Chem. Phys.* **2017**, *19*, 18685-18694.

45. Machireddy, B.; Kalra, G.; Jonnalagadda, S.; Ramanujachary, K.; Wu, C., Probing the Binding Pathway of Braco19 to a Parallel-Stranded Human Telomeric G-quadruplex Using Molecular Dynamics Binding Simulation with Amber DNA OL15 and Ligand Gaff2 Force Fields. *J. Chem. Inf. Model.* **2017**, *57*, 2846-2864.

46. Bian, Y.; Tan, C.; Wang, J.; Sheng, Y.; Zhang, J.; Wang, W., Atomistic Picture for the Folding Pathway of a Hybrid-1 Type Human Telomeric DNA G-quadruplex. *PLoS Computat. Biol.* **2014**, *10*, e1003562.

47. Ghosh, S.; Jana, J.; Kar, R. K.; Chatterjee, S.; Dasgupta, D., Plant Alkaloid Chelerythrine Induced Aggregation of Human Telomere Sequence—A Unique Mode of Association between a Small Molecule and a Quadruplex. *Biochemistry* **2015**, *54*, 974-986.

48. Bergues-Pupo, A. E.; Arias-Gonzalez, J. R.; Morón, M. C.; Fiasconaro, A.; Falo, F., Role of the Central Cations in the Mechanical Unfolding of DNA and RNA G-quadruplexes. *Nucleic Acids Res.* **2015**, *43*, 7638-7647.

49. Heddi, B.; Martín-Pintado, N.; Serimbetov, Z.; Kari, Teuku Mahfuzh A.; Phan, A. T., G-quadruplexes with (4n - 1) Guanines in the G-tetrad core: Formation of a G-triad·water Complex and Implication for Small-molecule Binding. *Nucleic Acids Res.* **2016**, *44*, 910-916.

50. Zeng, X.; Zhang, L.; Xiao, X.; Jiang, Y.; Guo, Y.; Yu, X.; Pu, X.; Li, M., Unfolding Mechanism of Thrombin-binding Aptamer Revealed by Molecular Dynamics Simulation and Markov State Model. *Sci. Rep.* **2016**, *6*, 24065.

51. Yang, C.; Kulkarni, M.; Lim, M.; Pak, Y., In-silico Direct Folding of Thrombin-Binding Aptamer G-quadruplex at All-atom Level. *Nucleic Acids Res.* **2017**, *45*, 12648-12656.

52. Moraca, F.; Amato, J.; Ortuso, F.; Artese, A.; Pagano, B.; Novellino, E.; Alcaro, S.; Parrinello, M.; Limongelli, V., Ligand Binding to Telomeric G-quadruplex DNA Investigated by Funnel-Metadynamics Simulations. *Proc. Natl. Acad. Sci. USA* **2017**, *114*, E2136-E2145.

53. Rebic, M.; Mocci, F.; Laaksonen, A.; Ulicny, J., Multiscale Simulations of Human Telomeric G-quadruplex DNA. *J. Phys. Chem. B* **2015**, 119, 105-113.

54. Luo, D.; Mu, Y., All-atomic Simulations on Human Telomeric G-quadruplex DNA Binding with Thioflavin T. *J. Phys. Chem. B* **2015**, *119*, 4955-4967.

55. Ray, A.; Panigrahi, S.; Bhattacharyya, D.; Case, D., A Comparison of Four Different Conformations Adopted by Human Telomeric G-quadruplex Using Computer Simulations. *Biopolymers* **2015,** *105,* 83-99.

56. Zhu, H.; Xiao, S.; Liang, H., Structural Dynamics of Human Telomeric G-quadruplex Loops Studied by Molecular Dynamics Simulations. *PloS One* **2013**, *8*, e71380.

57. Akhshi, P.; Acton, G.; Wu, G., Molecular Dynamics Simulations to Provide New Insights into the Asymmetrical Ammonium Ion Movement Inside of the $[d(G_3T_4G_4)]_2$ G-quadruplex DNA Structure. *J. Phys. Chem. B* **2012**, *116*, 9363-9370.

58. Luo, D.; Mu, Y., Computational Insights into the Stability and Folding Pathways of Human Telomeric DNA G-Quadruplexes. *J. Phys. Chem. B* **2016**, *120*, 4912-4926.

59. Petraccone, L.; Garbett, N. C.; Chaires, J. B.; Trent, J. O., An Integrated Molecular Dynamics (MD) and Experimental Study of Higher Order Human Telomeric Quadruplexes. *Biopolymers* **2010**, *93*, 533-48.

60. Rueda, M.; Luque, F. J.; Orozco, M., G-quadruplexes can Maintain Their Structure in the Gas Phase. *J. Am. Chem. Soc.* **2006**, *128*, 3608-3619.

61. Akhshi, P.; Wu, G., Umbrella Sampling Molecular Dynamics Simulations Reveal Concerted Ion Movement Through G-quadruplex DNA Channels. *Phys. Chem. Chem.Phys.* **2017**, *19*, 11017-11025.

62. Fadrna, E.; Spackova, N.; Sarzynska, J.; Koca, J.; Orozco, M.; Cheatham, T. E.; Kulinski, T.; Sponer, J., Single Stranded Loops of Quadruplex DNA As Key Benchmark for Testing Nucleic Acids Force Fields. *J. Chem. Theory Comput.* **2009**, *5*, 2514-2530.

63. Fadrna, E.; Spackova, N.; Stefl, R.; Koca, J.; Cheatham, T. E., 3rd; Sponer, J., Molecular Dynamics Simulations of Guanine Quadruplex Loops: Advances and Force Field Limitations. *Biophys. J.* **2004**, *87*, 227-242.

64. Islam, B.; Sgobba, M.; Laughton, C.; Orozco, M.; Sponer, J.; Neidle, S.; Haider, S., Conformational Dynamics of the Human Propeller Telomeric DNA Quadruplex on a Microsecond Time Scale. *Nucleic Acids Res.* **2013**, *41*, 2723-2735.

65. Islam, B.; Stadlbauer, P.; Krepl, M.; Koca, J.; Neidle, S.; Haider, S.; Sponer, J., Extended Molecular Dynamics of a c-Kit Promoter Quadruplex. *Nucleic Acids Res.* **2015**, *43*, 8673-8693.

66. Islam, B.; Stadlbauer, P.; Neidle, S.; Haider, S.; Sponer, J., Can We Execute Reliable MM-PBSA Free Energy Computations of Relative Stabilities of Different Guanine Quadruplex Folds? *J. Phys. Chem. B* **2016**, *120*, 2899-2912.

67. Sponer, J.; Banas, P.; Jurecka, P.; Zgarbova, M.; Kuhrova, P.; Havrila, M.; Krepl, M.; Stadlbauer, P.; Otyepka, M., Molecular Dynamics Simulations of Nucleic Acids. From Tetranucleotides to the Ribosome. *J. Phys. Chem. Lett.* **2014**, *5*, 1771-1782.

68. Stadlbauer, P.; Krepl, M.; Cheatham, T. E.; Koča, J.; Šponer, J., Structural Dynamics of Possible Late-Stage Intermediates in Folding of Quadruplex DNA Studied by Molecular Simulations. *Nucleic Acids Res.* **2013**, *41*, 7128-7143.

69. Stadlbauer, P.; Kuhrova, P.; Banas, P.; Koca, J.; Bussi, G.; Trantirek, L.; Otyepka, M.; Sponer, J., Hairpins Participating in Folding of Human Telomeric Sequence Quadruplexes Studied by Standard and T-REMD Simulations. *Nucleic Acids Res.* **2015**, *43*, 9626 -9644.

70. Stefl, R.; Cheatham, T. E., 3rd; Spackova, N.; Fadrna, E.; Berger, I.; Koca, J.; Sponer, J., Formation Pathways of a Guanine-Quadruplex DNA Revealed by Molecular Dynamics and Thermodynamic Analysis of the Substates. *Biophys. J.* **2003**, *85*, 1787-804.

71. Islam, B.; Stadlbauer, P.; Gil-Ley, A.; Perez-Hernandez, G.; Haider, S.; Neidle, S.; Bussi, G.; Banas, P.; Otyepka, M.; Sponer, J., Exploring the Dynamics of Propeller Loops in Human Telomeric DNA Quadruplexes Using Atomistic Simulations. *J. Chem. Theory Comput.* **2017**, *13*, 2458-2480.

72. Bian, Y.; Song, F.; Cao, Z.; Zhao, L.; Yu, J.; Guo, X.; Wang, J., Fast-Folding Pathways of the Thrombin-Binding Aptamer G-Quadruplex Revealed by a Markov State Model. *Biophys J* **2018**, *114*, 1529-1538.

73. Shen, Z.; Mulholland, K. A.; Zheng, Y.; Wu, C., Binding of Anticancer Drug Daunomycin to a TGGGT G-quadruplex DNA Probed by All-atom Molecular Dynamics Simulations: Additional Pure Groove Binding Mode and Implications on Designing More Selective G-quadruplex Ligands. *J. Mol. Model.* **2017**, *23*, 256.

74. Oscar, M.; Massimiliano, P.; F., S. G.; Valérie, G.; Jean-Louis, M., Orienting Tetramolecular G-Quadruplex Formation: The Quest for the Elusive RNA Antiparallel Quadruplex. *Chem. Eur. J.* **2015**, *21*, 6732-6739.

75. Sponer, J.; Cang, X.; Cheatham, T. E., 3rd, Molecular Dynamics Simulations of G-DNA and Perspectives on the Simulation of Nucleic Acid Structures. *Methods* **2012**, *57*, 25-39.

76. Havrila, M.; Stadlbauer, P.; Islam, B.; Otyepka, M.; Sponer, J., Effect of Monovalent Ion Parameters on Molecular Dynamics Simulations of G-quadruplexes. *J. Am. Chem. Soc.* **2017**, *13*, 3911-3926.

77. Gkionis, K.; Kruse, H.; Platts, J. A.; Mládek, A.; Koča, J.; Šponer, J., Ion Binding to Quadruplex DNA Stems. Comparison of MM and QM Descriptions Reveals Sizable Polarization Effects Not Included in Contemporary Simulations. *J. Am. Chem. Soc.* **2014**, *10*, 1326-1340.

78. Wirmer-Bartoschek, J.; Bendel, L. E.; Jonker, H. R. A.; Grun, J. T.; Papi, F.; Bazzicalupi, C.; Messori, L.; Gratteri, P.; Schwalbe, H., Solution NMR Structure of a Ligand/Hybrid-2-G-Quadruplex Complex Reveals Rearrangements that Affect Ligand Binding. *Angew Chem Int Ed Engl.* **2017**, *56*, 7102-7106.

79. Joung, I. S.; Cheatham, T. E., Determination of Alkali and Halide Monovalent Ion Parameters for Use in Explicitly Solvated Biomolecular Simulations. *J. Phys. Chem. B* **2008**, *112*, 9020-9041.

80. Perez, A.; Marchan, I.; Svozil, D.; Sponer, J.; Cheatham, T. E., 3rd; Laughton, C. A.; Orozco, M., Refinement of the AMBER Force Field for Nucleic Acids: Improving the Description of Alpha/Gamma Conformers. *Biophy. J.* **2007**, *92*, 3817-3829.

81. 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., A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules. *J. Am. Chem. Soc.* **1995**, *117*, 5179-5197.

82. Krepl, M.; Zgarbova, M.; Stadlbauer, P.; Otyepka, M.; Banas, P.; Koca, J.; Cheatham, T. E., 3rd; Jurecka, P.; Sponer, J., Reference Simulations of Noncanonical Nucleic Acids with Different Chi Variants of the Amber Force Field: Quadruplex DNA, Quadruplex RNA and Z-DNA. *J. Am. Chem. Soc.* **2012**, *8*, 2506-2520.

83. Zgarbova, M.; Luque, F. J.; Sponer, J.; Cheatham, T. E., 3rd; Otyepka, M.; Jurecka, P., Toward Improved Description of DNA Backbone: Revisiting Epsilon and Zeta Torsion Force Field Parameters. *J. Am. Chem. Soc.* **2013**, *9*, 2339-2354.

84. Zgarbova, M.; Sponer, J.; Otyepka, M.; Cheatham, T. E., 3rd; Galindo-Murillo, R.; Jurecka, P., Refinement of the Sugar-Phosphate Backbone Torsion Beta for AMBER Force Fields Improves the Description of Z- and B-DNA. *J. Am. Chem. Soc.* **2015**, *11*, 5723-5736.

85. Galindo-Murillo, R.; Robertson, J. C.; Zgarbova, M.; Sponer, J.; Otyepka, M.; Jurecka, P.; Cheatham, T. E., 3rd, Assessing the Current State of Amber Force Field Modifications for DNA. *J. Am. Chem. Soc.* **2016**, *12*, 4114-4127.

86. Case, D.; Darden, T.; Cheatham III, T.; Simmerling, C.; Wang, J.; Duke, R.; Luo, R.; Walker, R.; Zhang, W.; Merz, K., AMBER 12. *University of California, San Francisco* **2012**.

87. Darden, T.; York, D.; Pedersen, L., Particle Mesh Ewald: An N·log(N) Method for Ewald Sums in Large Systems. *J. Chem. Phys.* **1993**, *98*, 10089-10092.

88. Gil-Ley, A.; Bussi, G., Enhanced Conformational Sampling Using Replica Exchange with Collective-Variable Tempering. *J. Chem. Theory Comput.* **2015**, *11*, 1077-1085.

89. Hou, X.-M.; Fu, Y.-B.; Wu, W.-Q.; Wang, L.; Teng, F.-Y.; Xie, P.; Wang, P.-Y.; Xi, X.-G., Involvement of G-triplex and G-hairpin in the Multi-pathway Folding of Human Telomeric G-quadruplex. *Nucleic Acids Res.* **2017**, *45*, 11401-11412.

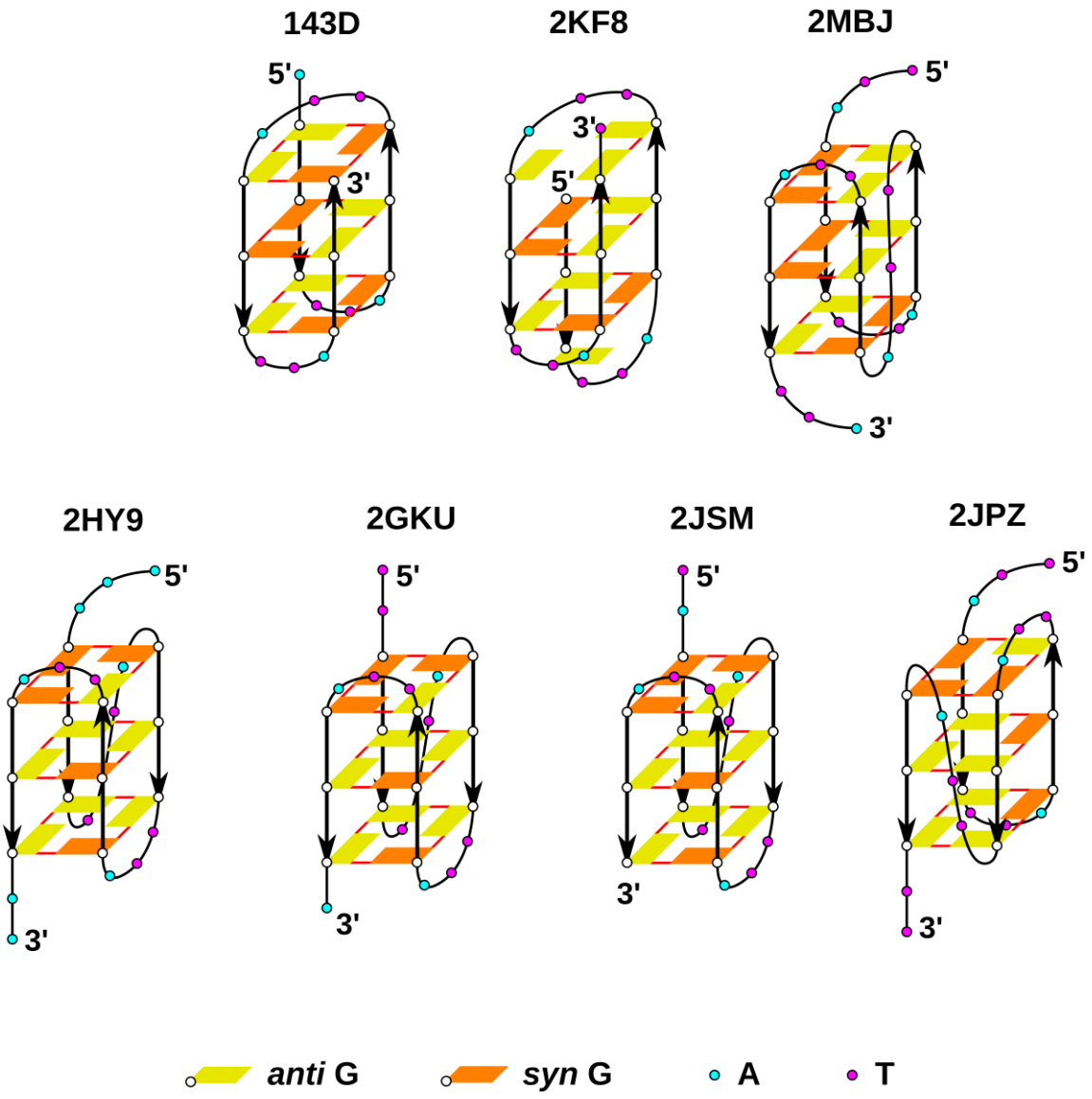

Figure 1: Schematic representation of Htel GQs with diagonal and lateral loops used in the present study. The backbone is shown as black lines. White, magenta and cyan circles mark the guanine, thymine and adenine nucleotides, respectively. Yellow and orange rectangles show guanines in *anti* and *syn* orientation. The PDB ids are shown above the respective structures.

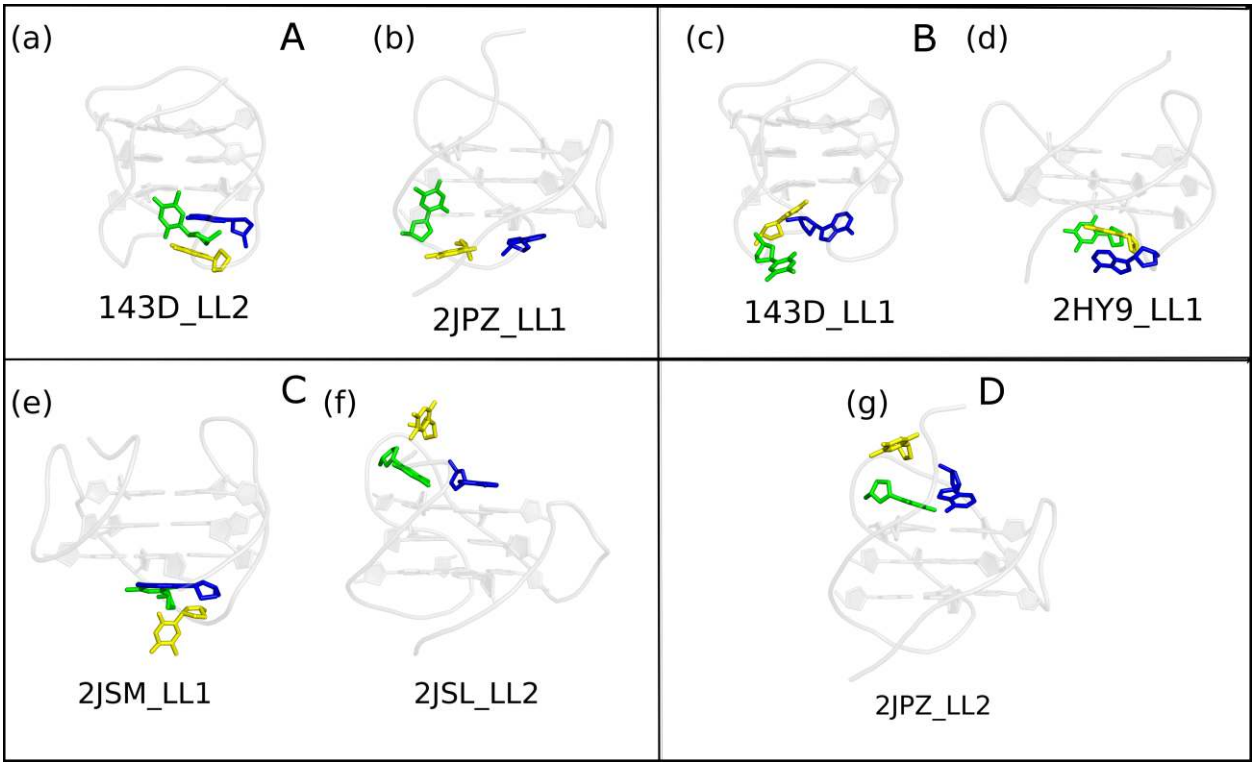

Figure 2: Type of LLs observed in experimental structures of Htel GQs. Each LL type is shown in separate panel. Panels A, B, C and D show type-1, 2, 3 and 4 LLs respectively. The PDB code and LL number is shown below each structure. Only the GS, backbone and loops of interest are shown. The GS and backbone are shown in grey transparent cartoon. The first thymine, second thymine and adenine of loops of interest are shown in green, yellow and blue sticks, respectively.

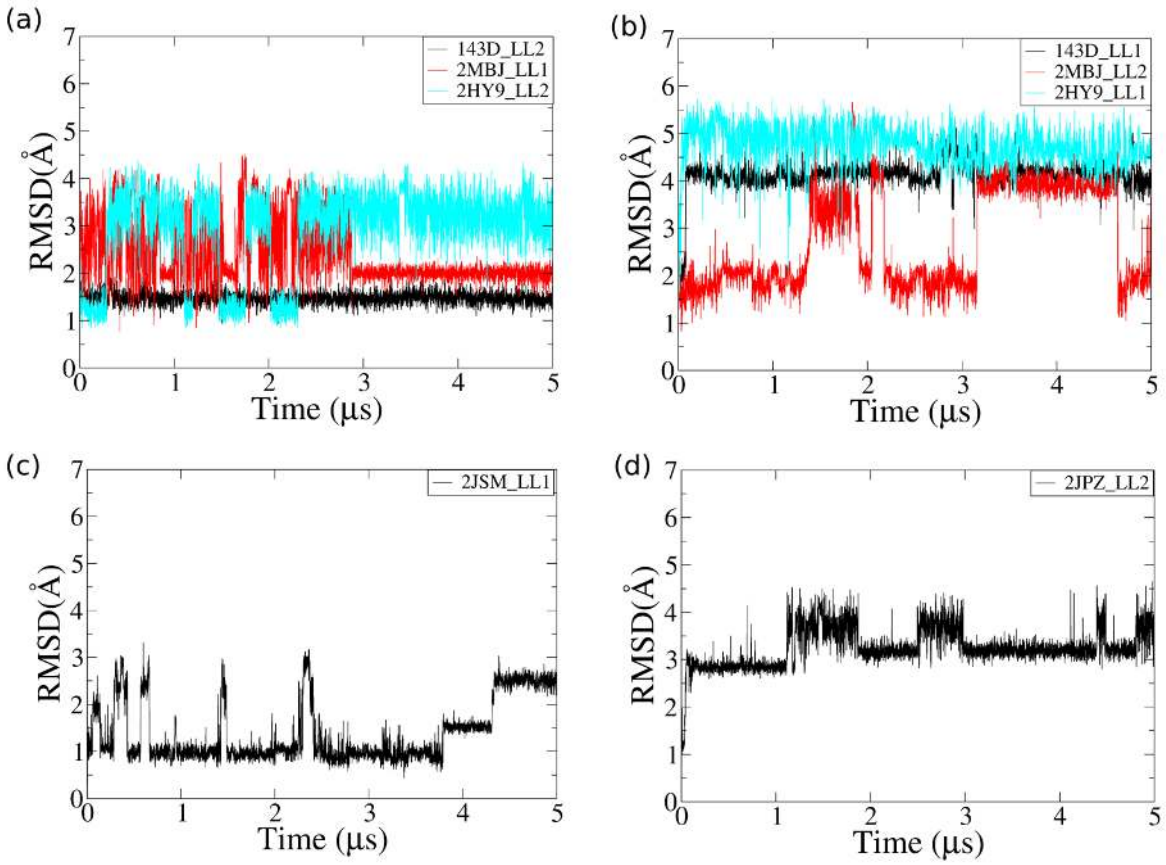

Figure 3: All-atom RMSDs of (a) type-1 (b) type-2 (c) type-3 and (d) type-4 LLs of Htel GQs in selected OL15 simulations carried out in Na^+ (i.e., NaCl) and with SPC/E water model. Type-1 is represented by LL2 of 143D and LL1 of 2MBJ and 2HY9. Type-2 is represented by LL1 of 143D and LL2 of 2MBJ and 2HY9. Type-3 and type-4 are represented by LL1 of 2JSM and LL2 of 2JPZ, respectively. For complete data see Supporting Information.

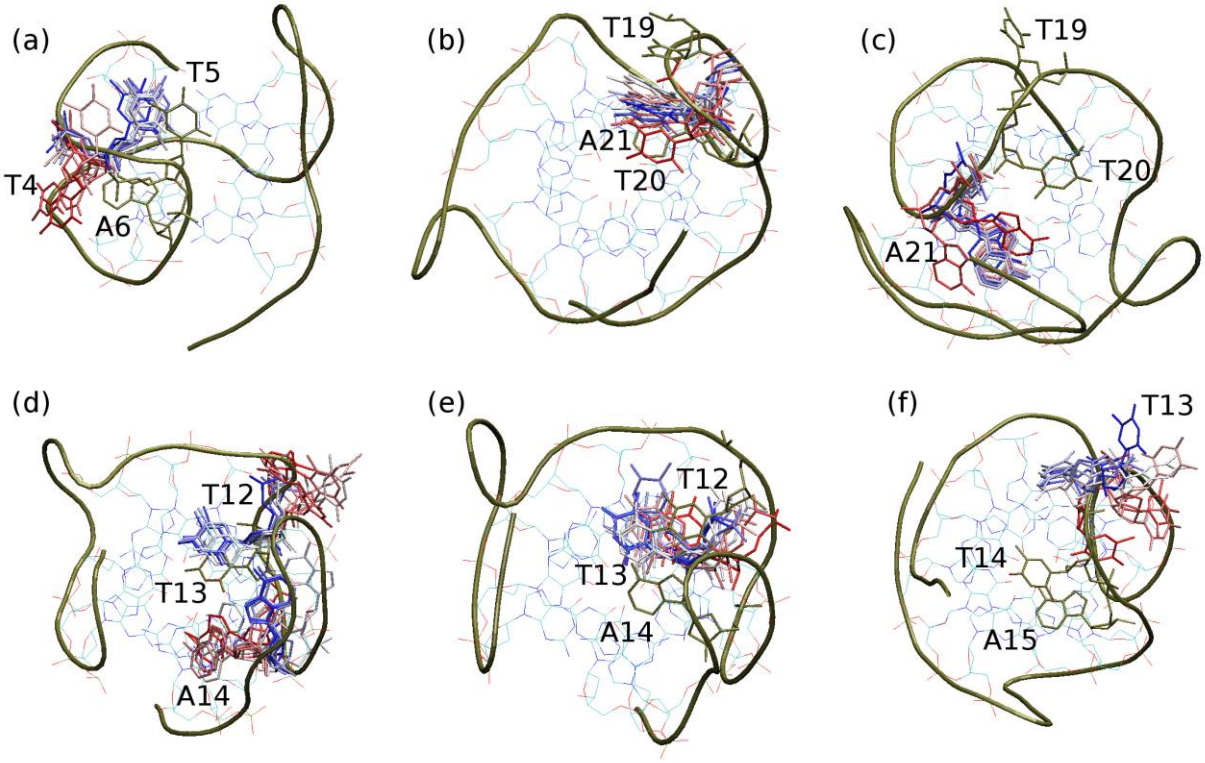

Figure 4: Dynamics of some LLs that changed loop types during the simulations. Top view of GQ is shown to highlight the loop dynamics. The backbone is shown in tan cartoon and GS is shown in lines. The LLs of interest are shown in sticks and the bases are labelled in the figures. Within the loop, base/bases showing significant dynamics are colored according to the trajectory time progression; red is the starting conformation and blue is the end conformation. The stable bases of the same loop are shown in tan sticks. (a) LL1 of 2KF8 in bsc0$\chi_{OL4}\varepsilon\zeta_{OL1}$ simulation with dynamics of T4, (b) LL2 of 2HY9 in bsc0$\chi_{OL4}\varepsilon\zeta_{OL1}$ simulation with dynamics of T20, (c) LL2 of 2MBJ in bsc0χ_{OL4} simulation with dynamics of A21, (d) LL1 of 2GKU in OL15 simulation in K^+ with dynamics of T12 and A14, (e) LL1 in 2JSM in OL15 with dynamics of T13 and (f) LL2 of 2JPZ in OL15 with dynamics of T13.

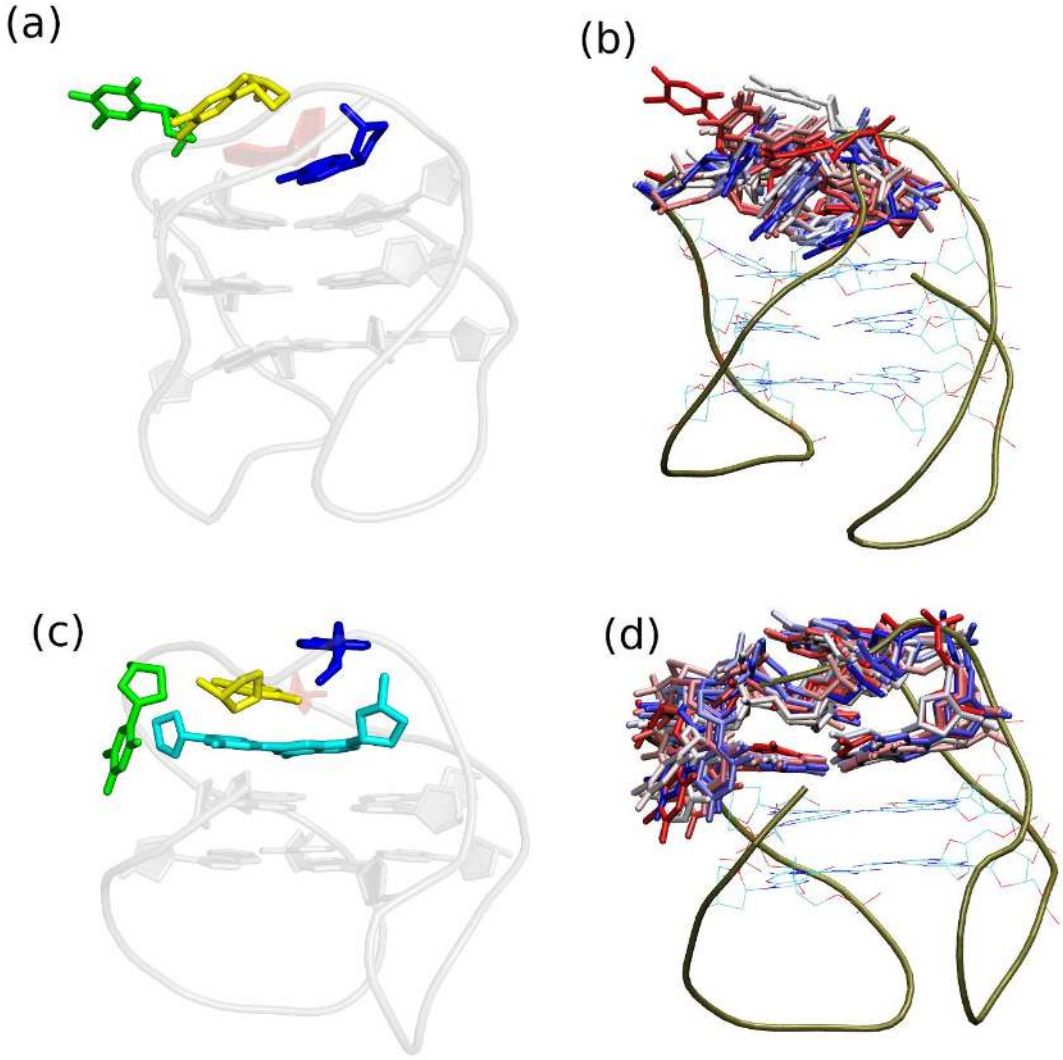

Figure 5: Diagonal loops of Htel GQs. Panels (a) and (c) show the experimental structures 143D and 2KF8, respectively while (b) and (d) show their dynamics in 5 μs long OL15 simulation in Na^+ and SPC/E water model at a time step of 500 ns. In panels (a) and (c), the backbone and GS are shown in grey cartoon. The diagonal loop nucleosides are shown in sticks, the first thymine, second thymine and adenine are shown in green, yellow and blue sticks, respectively. The terminal adenine of 143D interacts with adenine of diagonal loop and is shown in red in panel a. The diagonal loop guanines of 2KF8 are shown in cyan sticks while T22 that interacts with second thymine of diagonal loop is shown as red cartoon in panel c. In panels (b) and (d) the diagonal loop bases are colored according to the trajectory time progression; red is the starting conformation and blue is the end conformation.

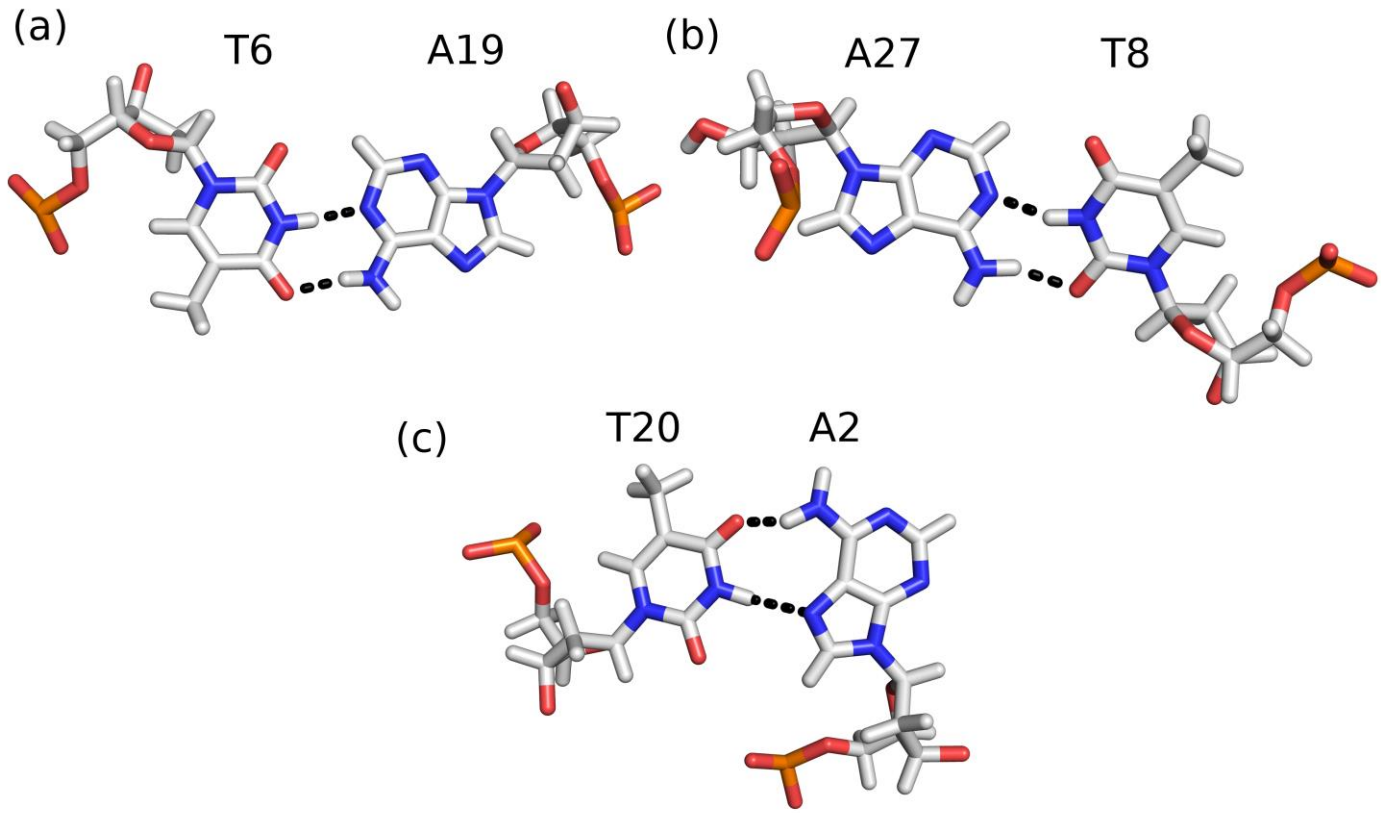

Figure 6: Example of TA base pairs in simulations of LLs. (a) Stable WC base pairing in simulations of 143D, (b) *trans* WC base pairing formed in 2MBJ in bsc0$\chi_{OL4}\varepsilon\zeta_{OL1}$ simulation and (c) Hoogsteen base pair formed in 2HY9 in bsc0χ_{OL4} simulation.

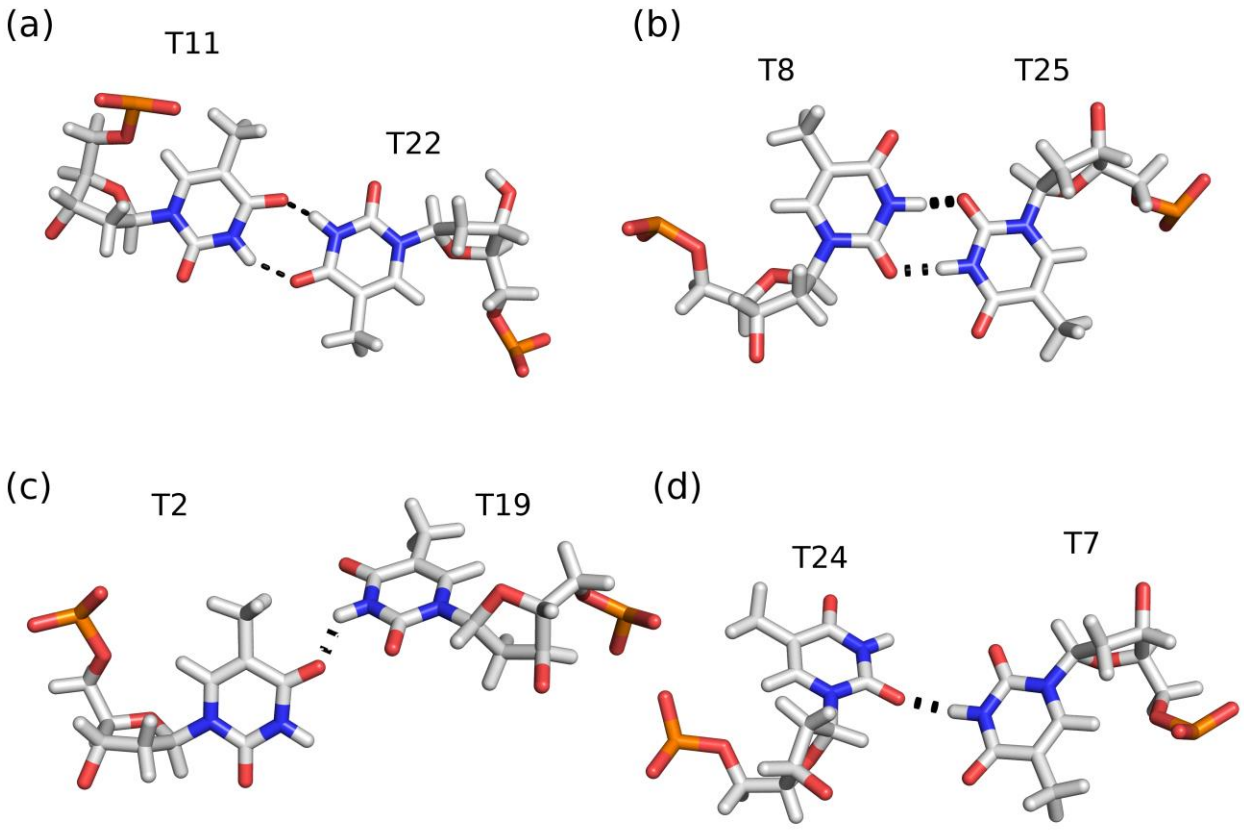

Figure 7: TT base pairs in LL of Htel GQ experimental structures. (a) T11:T22 in 2KF8, (b) T8:T25 in 2JPZ, (c) T2-T19 in 2GKU and (d) T24-T7 in 2JSL

Table 1: List of simulations carried out on in the present study

PDB id of the starting structure	Force-field variant	Ion type	Water model	Length of the simulation (μs)
143D	bsc0χ_{OL4}	Na^+	SPC/E	10
	bsc0$\chi_{OL4}\varepsilon\zeta_{OL1}$	Na^+	SPC/E	10
	OL15	Na^+	SPC/E	5
	OL15	K^+	SPC/E	5
	OL15	K^+	TIP3P	5
143D model 4	OL15	Na^+	SPC/E	3
	OL15	Na^+	SPC/E	3
	OL15	Na^+	SPC/E	3
	OL15	Na^+	SPC/E	3
	OL15	Na^+	SPC/E	5
2KF8	bsc0χ_{OL4}	Na^+	SPC/E	10
	bsc0$\chi_{OL4}\varepsilon\zeta_{OL1}$	Na^+	SPC/E	10
	OL15	K^+	SPC/E	5
2MBJ	bsc0χ_{OL4}	Na^+	SPC/E	10
	bsc0$\chi_{OL4}\varepsilon\zeta_{OL1}$	Na^+	SPC/E	10
	OL15	Na^+	SPC/E	5
	OL15	K^+	SPC/E	5
	OL15	K^+	TIP3P	5
2HY9	bsc0χ_{OL4}	Na^+	SPC/E	10
	bsc0$\chi_{OL4}\varepsilon\zeta_{OL1}$	Na^+	SPC/E	10
	OL15	Na^+	SPC/E	5
2GKU	bsc0$\chi_{OL4}\varepsilon\zeta_{OL1}$	Na^+	SPC/E	10
	OL15	K^+	SPC/E	5
	OL15	K^+	TIP3P	5
2JSM	bsc0$\chi_{OL4}\varepsilon\zeta_{OL1}$	Na^+	SPC/E	10
	OL15	Na^+	SPC/E	5
2JPZ	bsc0χ_{OL4}	Na^+	SPC/E	10
	bsc0$\chi_{OL4}\varepsilon\zeta_{OL1}$	Na^+	SPC/E	10
	OL15	Na^+	SPC/E	5
5MVB,ligand removed	OL15	Na^+	SPC/E	3
	OL15	Na^+	SPC/E	3
	OL15	Na^+	SPC/E	3
	OL15	Na^+	SPC/E	3
	OL15	Na^+	SPC/E	3
	OL15	Na^+	SPC/E	5

Table 2: Summary of LL behaviour in the simulations carried out in the present study

Loop type	Loop	groove	force field	behaviour in the simulations
Type-1	143D_LL2	narrow	bsc0χ_{OL4}	no change in loop type
			bsc0$\chi_{OL4}\varepsilon\zeta_{OL1}$	no change in loop type
			OL15	no change in loop type
			OL15 _K	no change in loop type
			OL15_TIP3P_K	no change in loop type
	2KF8_LL1	wide	bsc0χ_{OL4}	no change in loop type
			bsc0$\chi_{OL4}\varepsilon\zeta_{OL1}$	changed to type-4
			OL15	no change in loop type
	2KF8_LL2	narrow	bsc0χ_{OL4}	no change in loop type
			bsc0$\chi_{OL4}\varepsilon\zeta_{OL1}$	no change in loop type; first thymine of the loop was flexible but remain exposed to the solvent
			OL15	no change in loop type
	2MBJ_LL1	narrow	bsc0χ_{OL4}	no change in loop type
			bsc0$\chi_{OL4}\varepsilon\zeta_{OL1}$	no change in loop type
			OL15	no change in loop type
			OL15 _K	no change in loop type
			OL15_TIP3P_K	second thymine also moved towards the solvent for 600 ns and then re-attained type-1
	2HY9_LL2	narrow	bsc0χ_{OL4}	no change in loop type
			bsc0$\chi_{OL4}\varepsilon\zeta_{OL1}$	changed to a new loop type in which the second thymine also moved towards the solvent and stacked with first thymine of the LL
			OL15	no change in loop type
	2GKU_LL2	narrow	bsc0$\chi_{OL4}\varepsilon\zeta_{OL1}$	no change in loop type
			OL15 _K	no change in loop type
			OL15_TIP3P_K	no change in loop type
	2JSM_LL2	narrow	bsc0$\chi_{OL4}\varepsilon\zeta_{OL1}$	no change in loop type
			OL15	no change in loop type
	2JPZ_LL1	wide	bsc0χ_{OL4}	changed to type-4
			bsc0$\chi_{OL4}\varepsilon\zeta_{OL1}$	no change in loop type
			OL15	no change in loop type
Type-2	143D_LL1	wide	bsc0χ_{OL4}	no change in loop type
			bsc0$\chi_{OL4}\varepsilon\zeta_{OL1}$	no change in loop type
			OL15	no change in loop type
			OL15 _K	no change in loop type
			OL15_TIP3P_K	no change in loop type
	2MBJ_LL2	wide	bsc0χ_{OL4}	changed to type-1
			bsc0$\chi_{OL4}\varepsilon\zeta_{OL1}$	no change in loop type
			OL15	no change in loop type
			OL15 _K	no change in loop type
			OL15_TIP3P_K	no change in loop type

	2HY9_LL1	wide	bsc0χ_{OL4}	no change in loop type
			bsc0$\chi_{OL4}\varepsilon\zeta_{OL1}$	no change in loop type
			OL15	changed to type-1
Type-2 (contd)	2GKU_LL1	wide	bsc0$\chi_{OL4}\varepsilon\zeta_{OL1}$	changed to type-4
			OL15_ K	changed to type-4
			OL15_TIP3P_K	changed to type-1
Type-3	2JSM LL1	wide	bsc0$\chi_{OL4}\varepsilon\zeta_{OL1}$	changed to type-4
			OL15	changed to type-4
Type-4	2JPZ-LL2	narrow	bsc0χ_{OL4}	no change in loop type
			bsc0$\chi_{OL4}\varepsilon\zeta_{OL1}$	no change in loop type
			OL15	changed to type-1

Table 3: Comparison of base pairs[a] formed by the LL and flanking bases in the Htel GQ experimental structures and MD simulations

PDB structure	**Experimental**	**bsc0χ_{OL4}**	**bsc0$\chi_{OL4}\varepsilon\zeta_{OL1}$**	**OL15**		
				SPC/E	**SPC/E_K**	**TIP3P_K**
143D	T6:A19 A1:A13	T6:A19 A1:A13	T6:A19	T6:A19 A1:A13	T6:A19 A1:A13	T6:A19 A1:A13
2KF8	T11:T22	T11:T22 T5:T17	T11:T22 intermittent bonds T5:T17 (till 3.1 μs) T4:T17 (3.1-10 μs)	T11:T22 intermittent T5:T17 intermittent	no simulation	no simulation
2MBJ	A3:T20:A21 A15:A9:T25	A3:T20:A21 A15:A9:T25 T8:A27	A3:T20 A9:T25	A3:T20 A9:T25	A3:T20 A9:T25:A27 T8:T26 intermittent	A3:T20 A9:T25 T8:T26 intermittent
2HY9	A3:A21:A9 T14:A25	A3:A21 A2:T20 T14:A25	A3:A21 T14:A25	A3:A21 A2:T20	no simulation	no simulation
2GKU	T1:A20 T2:T19 (single HB) T13:A24	no simulation	T1:A20 T2:T19 (Single HB) intermittent T12:A24	no simulation	T1:A20 T2:T19 (single HB) T13:T24 (46%) T12:A24 (54%)	T1:A20 T2:T19 (single HB) T13:A24
2JSM	T1:A20:A2 T12:A14	no simulation	T1:A20 A2:T19 T12:A14	T1:A20 T12:A14	no simulation	no simulation
2JPZ	A3:T13:A15 A9:T8:T25	A3:T13:A15 A9:T26	A3:T13:A15 A9:T8:T25	A3:A15 A9:T8:T25 T2:T14	no simulation	no simulation
2JSL	T12:A14 A8:T25	no simulation	no simulation	no simulation	no simulation	no simulation

[a] Gaps indicate that the base pair is not observed in the experiment or simulation. This applies to the base pairs which are sampled in some and not all of the simulations.

For Table of Contents Use Only

Title: Structural dynamics of lateral and diagonal loops of Human telomeric G-quadruplexes in extended MD simulations

Authors: Barira Islam, Petr Stadlbauer, Marek Havrila, Shozeb Haider and Jiri Sponer