A molecular dynamics study on the binding of an anti-cancer DNA G-quadruplex stabilizer, CX-5461, to human telomeric, cKIT-1, and c-Myc G-quadruplexes and a DNA duplex

Holli-Joi Sullivan[1], Brian Chen[1] and Chun Wu[1]*

[1] Rowan University, College of Science and Mathematics, Glassboro, NJ, 08028 USA

* To whom correspondence should be addressed. Email: wuc@rowan.edu

ABSTRACT

DNA G-quadruplex (G4) stabilizer, CX-5461 is in Phase I/II clinical trials for advanced cancers with BRCA1/2 deficiencies. A FRET-melting temperature increase assay measured the stabilizing effects of CX-5461 to a DNA duplex (~10 K) and three G4 forming sequences negatively implicated in the cancers upon its binding: human telomeric (~30 K), c-KIT1 (~27 K) and c-Myc (~25 K). Without experimentally solved structures of these CX-5461-G4 complexes, CX-5461's interactions remain elusive. In this study, we performed a total of 73.5 μs free ligand molecular dynamics binding simulations of CX-5461 to the DNA duplex and three G4s. Three binding modes (top, bottom and side) were identified for each system and their thermodynamic, kinetic, and structural nature were deciphered. The MM/PBSA binding energies of CX-5461 were calculated for the human telomeric (-28.6 kcal/mol), c-KIT1 (-23.9 kcal/mol), c-Myc (-22.0 kcal/mol) G4s, and DNA duplex (-15.0 kcal/mol) systems. These energetic differences coupled with structural differences at the 3' site explained the different melting temperatures between the G4s, while CX-5461's lack of intercalation to the duplex explained the difference between the G4s and duplex. Based on the interaction insight, CX-5461 derivatives were deigned and docked, showing higher selectivity to the G4s over the duplex.

INTRODUCTION

The use of DNA G-Quadruplexes (G4s) as novel therapeutic targets has been a rapidly developing field over the last decade because compounds targeting the G4s have demonstrated a high potential against a variety of cancer cell lines. DNA G-quadruplexes demonstrate very useful characteristics as drug targets including high diversity, stability and much slower dissociation when compared to DNA duplexes. With well over 300,000 sequence motifs identified within the human genome, the design of small molecules targeting G4s as anti-cancer agents has become a primary focus of many researchers.

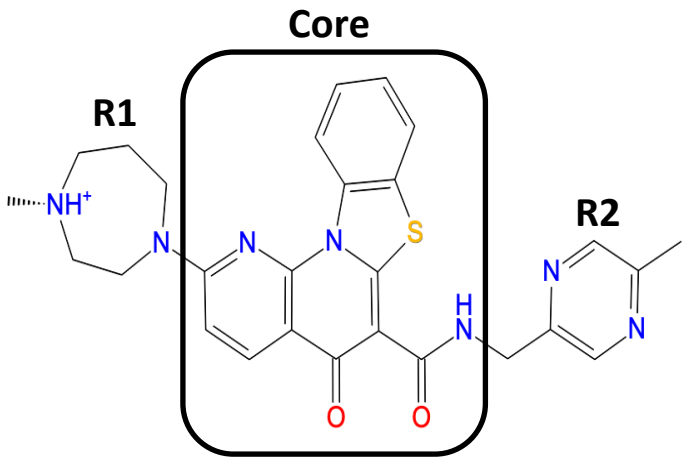


Figure 1. Chemical Structure of CX-5461. Two distinct regions are defined here where the region outlined in black represents the rigid core of CX-5461 and the two side chains (R1 and R2) represent flexible regions of CX-5461.

Small molecule CX-5461 (**Figure 1**) is a DNA G-quadruplex stabilizer whose structure contains a rigid benzothiazole-based core and two flexible side chains which are methyl diazepane based (R1) and methylpyrazine based (R2). CX-5461 was designed for superior *in vivo* stability and pharmacokinetics and is currently in Phase I/II clinical trials for advanced hematologic malignancies and cancers with BRCA1/2 deficiencies. CX-5461's first working mechanism was identified as a G4 binder directly inhibiting the binding of RNA Polymerase I, which has implications in cancer therapeutics.[1] More specifically, by binding to ribosomal DNA (rDNA) G4s formed within the rDNA promoter, CX-5461 prevents the binding of the transcription factor, SL1, and subsequently RNA Polymerase I, to the rDNA promoter which prevents the transcription of rDNA.[2, 3] Recently a second, unexpected, mechanism was identified for CX-5461 whereby it disrupts the cells replication fork by binding to and stabilizing chromosomal DNA G4 structures in cancer cells. Although experiments have yet to identify specific G4 targets for CX-5461 in the human genome, experimental evidence discovered CX-5461's specific roles at DNA G4s include the ability to selectively bind to and stabilize G4 structures of human cells lines *in vitro*, and increase the number of *in vivo* G4 structures. [4] These properties are extremely advantageous for cancer therapeutics, and evident from recent work, CX-5461 is a promising therapeutic agent for a variety of targets. In fact, as research expands, so do the number of potential targets for CX-5461 including solid tumors[5], acute myeloid leukemia[6, 7], multiple myeloma[8, 9],

neuroblastoma tumors[10], prostate cancer[11], osteosarcoma[12], acute lymphoblastic leukemia[13, 14], epithelial ovarian cancer[15-17], arterial injury-induced neointimal hyperplasia[18], and even non-cancerous diseases such as cytomegalovirus[19, 20], Herpes Simplex type I virus[19], and African trypanosomiasis[21]. However, without an experimentally solved structure of CX-5461 in complex with any G4 structure the specific interactions associated with the binding of CX-5461 and ultimate stabilization of the G4 remains to be fully understood.

It is critical for G4 stabilizers to have a high binding affinity to G4's and demonstrate high selectivity over DNA duplexes to reduce adverse side effects. Experimentally, this has been shown using a DNA duplex as a negative control when comparing the binding of several G4 targeting ligands which have effectively demonstrated a higher affinity and selectivity towards G4s over DNA duplexes.[22, 23] In one study, Xu et. al performed a FRET-melting temperature increase assay to test CX-5461's stabilizing effects to the canonical DNA duplex structure and three different G4 forming sequences which have been implicated in the cancerous complications resulting from BRCA1/2 mutations (human telomeric, c-KIT1, and c-Myc) [24-29]. Using the double stranded DNA duplex as a negative control to the G4 systems[4], the melting temperatures of each system was measured in the apo form. Then CX-5461 was added to each system to measure the increase in melting temperature upon CX-5461 binding to each DNA fragment, where a higher increase in melting temperature indicates a higher stabilizing effect and thus higher binding affinity. The results of the FRET melting temperature assay indicated that with 10 μM CX-5461 the highest melting temperature increase was demonstrated by the human telomeric system (~30 K) followed by the c-KIT1 (~27 K) and c-Myc (~25 K) G4s and the DNA duplex (~10 K). Thus, these results show that the stabilizing effect due to CX-5461 binding was highest in the human telomeric complex followed by the c-KIT1 and c-Myc G4s and then the DNA duplex. The difference in melting temperature increase between the G4s and duplex complex systems (15+ K) suggest that CX-5461 can selectively bind to and stabilize G4 structures over duplex DNA[4]. Along with the three G4 systems having significantly higher melting temperatures than the DNA duplex, they also varied from each other. Due to this, it is essential to compare the binding modes and mechanisms of CX-5461 in complex with the G4s versus the DNA duplex to identify specific differences that may help explain the higher binding affinity and selectivity demonstrated by CX-5641 to the G4s over the duplex and the variance among the three G4s.

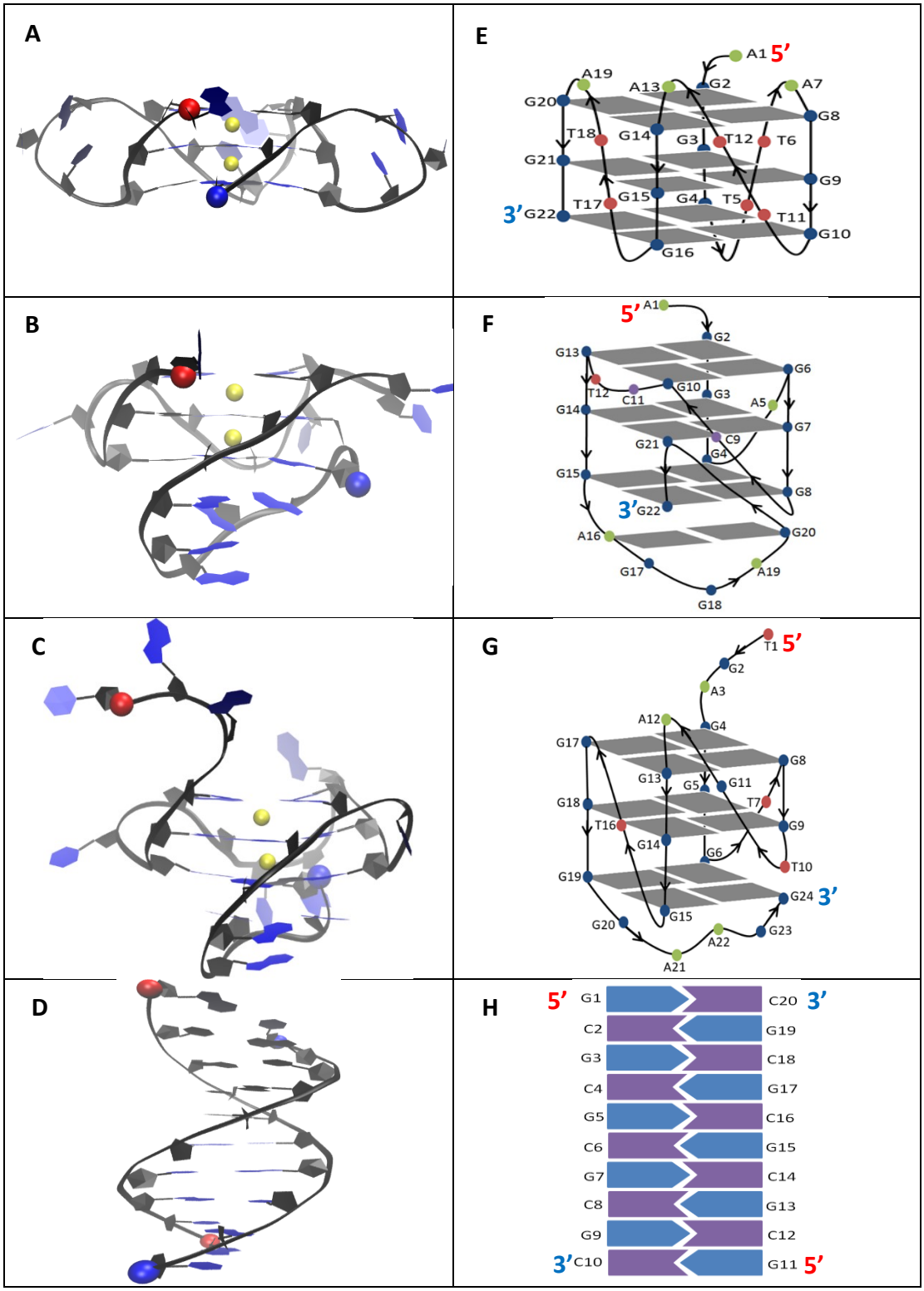


Figure 2. The initial configuration (A-D) and topology models of the simulation systems (E-H). **A&E:** Human telomeric quadruplex G4 (PDB ID: 1KF1), **B&F:** c-KIT1 quadruplex (PDB ID: 4WO3), **C&G:** c-Myc quadruplex (PDB ID: 2MGN) and **D&H:** Duplex DNA. 5’ and 3’ are indicated by red and blue spheres, respectively. K^+ ions are represented by yellow spheres.

Based on the three G4 forming sequences used in the FRET melting temperature assay performed by Xu and coworkers[4], the solved G4 scaffolds were obtained from the Protein Data Bank and used in this study. These include a human telomeric G4 (PDB ID: 1KF1), c-KIT1 G4 (PDB ID: 4WO3), and c-MYC promoter G4 (PDB ID: 2MGN)[30-32] (**Figure 2**). Due to the sequence difference these G4s also vary in structure. The human telomeric DNA G4 (**Figure 2A;E**) is made of four parallel DNA strands with three linking TTA trinucleotide loops which connect the top of one strand to the bottom of another forcing the strands into a parallel configuration which is highly similar on each face. This parallel scaffold was chosen based on the understanding gained in our previous work where we deciphered two lines of conflicting evidence on the major target form of BRACO19: (1) under solution conditions with cellular extracts as crowding agents, multiple conformations coexist including parallel, anti-parallel and the hybrid topologies. (2) however, there is no high-resolution complex structures of BRACO19 binding to antiparallel or the hybrid scaffold, except for parallel stranded. Our binding energy data suggested a hypothesis that reconciled the conflict: the relative population shift of three scaffolds upon BRACO19 binding (i.e., an increase of population of parallel scaffold, a decrease of populations of antiparallel and/or hybrid scaffold). We felt the hypothesis appeared to be consistent with the facts that BRACO19 was specifically designed based on the structural requirements of the parallel scaffold and has since proven effective against a variety of cancer cell lines.[33] Recent experimentation by the Kumar group has further supported our conformational selection mechanism hypothesis through a binding affinity assay and their circular dichroism spectra of BRACO19 binding to parallel G4 Structures of *Klebsiella pneumoniae*.[34] On the other hand, the c-KIT1 DNA G4 (**Figure 2B;F**) has an anti-parallel scaffold with double chain-reversal and a long lateral stem loop at the 3' region made of five nucleotides, two of which (A16 and G20) are capable of pairing. There is also one non-G-tract guanine that is part of the core of stacked G-quartets and the short single and dinucleotide loops of this c-KIT1 G4 are extremely flexible and show extensive base flipping. Whereas the c-Myc DNA G4 (**Figure 2C;G**) has a hybrid scaffold with a snapback motif that is adopted by the 3′-end GAAGG segment that forms a stable diagonal loop containing a G(A-G) triad and caps the 3' side of the G-tetrad. Finally, for our DNA duplex system we use a GC rich DNA duplex (**Figure 2D;G**), rather than using an oligonucleotide comprised of a polyethylene glycol linker able to fold into a hairpin as used in the FRET melting temperature assay by Xu and coworkers, which we feel is a more suitable comparison under physiological conditions.

To date, a variety of studies have demonstrated using molecular modeling and simulations as a powerful approach to identify structural details at the molecular level. [35] Hou et al. used this approach to probe the stability of six ligand-G-quadruplex DNA complexes which were structurally determined by experimental approaches [36]. Many MD simulations studies are ligand binding studies that effectively provide mechanistic insight into the binding of small molecules to G4 DNA. [37-42] Both AMBER forcefields such as BSC1 and OL15[35, 43-47] and polarizable force fields[48] are commonly used for these studies. Information such as DNA-ligand binding free energy calculations, identification of ligand/G4 binding sites, and ligand binding modes were successfully determined using a modeling system that utilized the standard parm99 Amber force field using parmbsc0 parameters and a K^+ cation in the center of the G-tetrads to neutralize the system[39, 40, 49]. Deng et. al resolved ligand-binding specificity using absolute

free binding energy calculations for both c-MYC [50] and human telomeric [51] G-quadruplex DNA . The work of Luo and Mu studied the binding of small molecules to human telomeric G-quadruplex using all-atomic molecular dynamics simulations[52]. Kumar and coworkers studied the binding of small molecules to G4 formed by the c-MYC promoter[32]. Some studies, like that performed by Chatterjee and coworkers have also had success performing an in silico screening on G4 structures formed by the c-MYC oncogene [53]. Further, the Lemkuls group has performed extensive work on the binding of small molecules to the c-KIT1 promoter G4 using molecular dynamics simulations with polarizable force field [54-56]. In addition to small molecules targeting single G-quadruplexes, Praadeepkumar and coworkers studied the binding of small molecules that stabilize multiple G-quadruplex forming sequences including the c-MYC and c-KIT1 promoter G4s[57]. Modeling studies have also produced insight into a variety of G-quadruplex forming sequences. Research done on telomeric G4s have successfully calculated realistic intermolecular and relative binding energies as well as determined binding modes and pathways. [42, 58, 59] Extensive research was performed by Liu and others which studied potassium binding with human telomeric intra-molecular G-quadruplex using molecular dynamics[60]. These studies provide invaluable insight into the model systems which strongly suggest the importance of using atomistic simulations to rationalize biologically relevant phenomena.[39, 58] Even more, many biological studies have been performed beyond computation that highlight the potential of targeting DNA G-quadruplexes with small molecules like CX-5461, however there are very limited complex structures available.

With these facts in mind, the goal of this study is to use free ligand all atom molecular dynamics binding simulations to study the binding of CX-5461 to the human telomeric, c-KIT1 and c-Myc G4s and a DNA duplex. Our post simulation analysis will identify the major binding pose, binding mechanisms and binding pathways of the CX-5461 complexes and provide novel insight as to how CX-5461 has been experimentally shown to selectively bind to and stabilize these G4s. Through our analysis we also address the order of stability of each system and features that differentiate the binding of CX-5461 to the G4's and the DNA duplex. These findings help understand the experimentally determined binding affinity and selectivity of CX-5461 to the G4 structures over the duplex. With these interaction insights, we propose optimizations to CX-5461 that may increase its interactions with G4s but decrease its interactions with the DNA duplex structure, which may improve its anti-cancer efficacy with less adverse side effects.

MATERIAL AND METHODS

Table 1. Molecular dynamics simulation runs.

System ID[1]	DNA	No. of Ligands	No. of run	No. of Water molecules	ions	Box Size (Å)*	Drug Initial Pose	NPT eq. (ns)	NVT (ns)	Total time (μs)
1	0	1	2	1420	1 Cl-	40.6	Free	1	1999	4
2	1(h-Tel)	0	2	4671	21 K+	60.9	N/A	1	1999	4
3	1(c-KIT1)	0	2	3954	21 K+	58.1	N/A	1	1999	4
4	1(c-Myc)	0	2	4843	23 K+	61.9	N/A	1	1999	4
5	1(Duplex)	0	2	4515	18 K+	55.4	N/A	1	1999	4
6	1(h-Tel)	1	27/3	8261	20 K+	72.4	Free	1	499 /1999	19.5
7	1(c-KIT1)	1	28/2	6371	20 K+	67.0	Free	1	499 /1999	18.0
8	1(c-Myc)	1	28/2	5958	22 K+	65.8	Free	1	499 /1999	18.0
9	1(Duplex)	1	28/2	5282	17 K+	57.0	Free	1	499 /1999	18.0

[1]Systems 1-4 refer to the free DNA-only systems, system 5 refers to the CX-5461 free ligand simulation, systems 6-9 refer to the free DNA plus free ligand simulations (6:9: Human telomeric (PDB ID: 1KF1), c-KIT1 (PDB ID: 4WO3), c-Myc (PDB ID: 2MGN) and Duplex complexes, respectively).
*Triclinic box equivalent to the true truncated octahedral box

Simulation Protocol. A full description of the methods used in this study is provided in the Supporting Information. In brief, a total of nine systems were constructed: a free ligand quadruplex-ligand complex system for each of the following PDB ID's: 1KF1, 4WO3, 2MGN, as well as one free ligand DNA duplex-ligand complex system, the apo form of each respective system, as well as a CX-5461 only system (**Table 1**). The DNA duplex-ligand system was constructed using a B-DNA duplex structure (sequence: $d([GC]_{10})_2$), using the Maestro program. The three DNA quadruplex-ligand systems were solvated inside a water box of truncated octahedron with 10 Å water buffer. Cl^- or K^+ counter ions were used to neutralize the system. The DNA fragments were represented using a refined version of the AMBER DNA OL15 (i.e. parm99bsc0[61] $+\chi_{OL4}$[62]$+ \varepsilon/\zeta_{OL1}$[63]$+ \beta_{OL1}$[64] updates). The water was represented using the TIP3P and the K^+/Na^+ model developed by Cheatham group was used to represent the K^+ ions.[65] Using this model, the K^+ ions remain tightly bound in the G4 ion pore during the length of the simulations. The standard AMBER protocol was used to create the force field for CX-5461. This procedure included calculating the molecular electrostatic potential (MEP) of the CX-5461 molecule at the HF/6-31G* level after geometry optimization at the same theory level. Along with other parameters collected from the AMBERGAFF2[66] force field, the MEP was used to identify the partial charges of the CX-5461 atoms using Restrained Electrostatic Potential/RESP method with two stage fitting[67]. Using AMBER DNA force fields are a highly effective and widely used in nucleic acid simulations.[68-71] This experiment was able to simulate the binding process of the DNA G-Quadruplex (G4) stabilizer CX-5461, to a human telomeric G4 (PBD ID: 1KF1), a human c-KIT1 G4 (PBD ID: 4WO3), and a c-MYC promoter G4 (PBD ID: 2MGN),[30-32] as well as a B-DNA fragment. [72] The AMBER GAFF2 force field of CX-5461 in Mol2 format is provided at the end of the supporting document. The detailed protocol for these simulations follow

our earlier studies;[33, 73-75] where the AMBER 16 simulation package was used for the production runs of all systems.[66] Following the Maxwell-Boltzmann distribution, atoms of the system were assigned different initial velocities by use of random seeds after the energy minimization. Thirty independent trajectories were run for each of the four complex systems to allow for better sampling of binding poses and pathway. In order to equilibrate the system density, a 500 ns production run at 300 K included a 1.0 ns molecular dynamics in the NPT ensemble mode (constant pressure and temperature). During this production run the DNA and the ligand were under Cartesian restraints (1.0 kcal/mol/Å), and 499.0 ns molecular dynamics in the NVT ensemble mode (constant volume and temperature). Two or three representative trajectories for each of four complex system were further extended into 1999.0 ns. A 2.0 fs time step in the simulations was created using SHAKE[76], which was able to constrain any bond connecting hydrogen atoms. Long-range electrostatic interactions under periodic boundary conditions (charge grid spacing of ~1.0 Å, the fourth order of the B-spline charge interpolation; and direct sum tolerance of 10^{-5}) were treated with the particle-mesh Ewald method.[77] The long range van der Waals interactions were based on a uniform density approximation; the cutoff distance for short-range non-bonded interactions was 10 Å. Non-bonded forces were calculated using a two-stage RESPA approach.[78] During this approach, the short-range forces were updated every step whereas the long range forces were updated every two steps. Using a Langevin thermostat with a coupling constant of 2.0 ps the temperature was controlled and the trajectories were saved at 50.0 ps intervals for analysis.

Order parameters characterize the DNA-drug binding pathway. The DNA-drug binding process was characterized using seven order parameter calculations: hydrogen bond analysis, drug-base dihedral angle which measures the angle between the two intersecting planes, drug center to ligand center distance (R*)*, K+-K+ cation distance, DNA and ligand RMSD and MM-GBSA binding energy (ΔE). The distance cutoff between H-donor and H-acceptor was set to be 3.5Å and the donor-H-acceptor angle cutoff was set to be 120°. The hydrogen bonds were calculated for the top/first, middle/second and bottom/third base layers, and other base pairing when applicable (**Figure 2 E-H**). A visual representation of the hydrogen bond networks are presented in the supporting document. The definition for the quadruplexes is the three guanine layers with the 5' side as the first layer. The center-to-center distance (R) was defined in two ways: as the length from the DNA center to the drug molecule center and the length between the two K^+ cations within the G-quadruplex structure. The dihedral angle was defined as the dihedral angle between the plane of the stable unbroken base-layer of the DNA that is closest to the drug binding site and the drug's ring plane. MM-PBSA[79] (Molecular Mechanics Poisson Boltzmann Surface Area) module in the AMBER package (PB1 model with mBondi radii set, salt concentration of 0.2 M, and surface tension of 0.0378 kcal/$Å^2$) was used to analyze the energetics of the bound complexes. The MM-PBSA binding energy for a system was calculated based on three simulations: the ligand only, the DNA only and the DNA-ligand complex using equation 1. The equation is made of four components Eq2: Van der Waals interaction energy (VDW), hydrophobic interaction energy (SUR), electrostatic interaction (PBELE) and the change of the conformation energy for DNA and ligand which are calculated using equation 3 and 4. MM-PBSA binding energy is an effective tool for ranking ligand binding affinities proven by up to 1864 crystal complexes tested in systematic benchmarking studies. [80-84]

Eq 1: $\Delta E = E_{complex} - E_{DNA\_free} - E_{lig\_free}$

Eq 2: $\Delta E = \Delta E_{vdw} + \Delta E_{SUR} + \Delta E_{GBELE} + \Delta E_{conformation}$

Eq 3: $\Delta E_x = E_{x\_complex} - E_{x\_DNA\_complex} - E_{x\_lig\_complex}$, *x= vdw, sur and pbele*

Eq 4: $\Delta E_{Conformation} = E_{DNA\_complex} + E_{lig\_complex} - E_{DNA\_free} - E_{lig\_free}$

Virtual Screening. The top derivatives were chosen from a virtual screening of a CX-5461 derivate library using Maestro 10.3 [85]. First, a combinatorial library of 64 ligands prepared using the Interactive Enumeration program. The variants were defined by establishing substitution sites where there were three possible points of substitution to CX-5461. At each possible point of substitution there were 4 functional groups that can be substituted which included hydrogen, fluorine, chlorine, bromine. This suggests that there are 4^3 possible modified versions of CX-5461, and a combinatorial library of these 64 ligands was generated. The charge of each compound at pH=7 was determined by Epik (an empirical pKa prediction program)[86] followed by a geometry optimization that minimized the potential energy using the default parameters. Using the active receptor structure from the most abundant conformation of each system, a grid file was generated using the Receptor Grid Generation program to prepare the complex for the subsequent docking calculation. In each system, CX-5461 was selected and a grid box was generated around the ligand with a Van der Waals radius scaling factor of 1.0 and a partial cutoff of 0.25. Then, these 64 compounds were docked using Glide with Extra Precision (XP) scoring function, and then filtered using QikProp package[87], to predict the absorption, distribution, metabolism, and excretion (ADME) properties. QikProp ranks the full molecular structure based on pharmaceutically relevant properties by giving each compound a number of stars; compounds with no starts are predicted to be the most drug-like. Finally, four potential compounds were manually chosen based on XP scores that were more negative than the docking of CX-5461 into the active conformation of the G-quadruplex from the most abundant clusters, along with the compounds' synthesizability containing fewer substitution groups.

RESULTS

The DNA only simulation runs showed that the overall DNA scaffold was maintained. To validate the force field used in our simulations, independent 2μs stability simulation runs carried out for each apo form. We performed a root mean square deviation (RMSD) analysis which compared the deviation of the DNA backbone in each snapshot to the initial structure. Since the RMSD values plateaued, this showed that the ligand only or the apo DNA remained stable in each simulation run and after taking the average of each run per system a figure is presented in **Figure S1**. Nonetheless, the three G4 systems showed a larger structure deviation from the initial conformation (i.e. ~4Å of human telomeric G-quadruplex, ~3Å of c-KIT1 G-quadruplex and ~5Å of c-Myc G-quadruplex) due to the fluctuation of non-G-tetrad parts. Next, we analysed the last snapshots of each apo form simulation run (**Figure S2**) and found that each the scaffold of each system was visually maintained when compared to the initial structure. A deeper understanding of the RMSD values were obtained

through visual inspection of the apo form trajectories. For the human telomeric system, there appears to be high flexibility at the 5' terminal residue as well as for each of the three, three residue, connecting loops. The c-KIT1 system showed high flexibility at the 5' terminal residues as well as the four connecting loops, ranging from 2 to 5 residues long. The c-Myc system showed the greatest flexibility at the longer 5' terminal segment and also at the varied length connecting loops. The most notable difference of the last snap shots was for the c-Myc system, where some residues of the longer and highly flexible 5' terminal segment stabilized on the G4 core which was consistent in both runs, which is likely because the original PDB structure of this G4 contained a ligand and our simulation runs allowed the DNA to relax into an apo form. Following this analysis, we took representative snapshots (**Figure S3**) for each of the apo DNA simulation systems: human telomeric (A), c-KIT1 (B), c-Myc (C) G-quadruplexes, and DNA duplex (D). Along with each representative trajectory is an order plot which shows that each system maintained a backbone RMSD, the potassium ions in each G4 system maintained their initial positions, and each systems MMPBSA energy relative to the initial snapshot showed small fluctuations but the average remained the same overall.

The free ligand binding simulations converged. A variety of post simulation analyses were performed to ensure proper sampling and convergence was reached in our simulations. First, we generated a plot showing the position of a single atom of CX-5461 in each trajectory (**Figure S4**). Due to the clear distribution of binding around the DNA in each system, we concluded there was a good sampling in each system. Following this, the root mean square deviation (RMSD) of the DNA backbone was calculated against the initial structure for the complex systems and the average of each system was calculated and presented (**Figure S5**). The flat and small receptor RMSDs in each system indicated the stability of the DNA structures during the simulation period. Next, atom contacts between the DNA structure and CX-5461 were calculated using a 3.0 Å cut-off (**Figure S6**). Here we defined a stable complex as a complex with a number of atom contacts between the DNA and ligand greater than 30. The stable contact number in this figure indicates the simulation systems reached a steady state in all runs. With our analyses suggesting proper sampling and convergence, we started looking at the binding poses. The last snap shots and a table summarizing each system's final binding poses are provided in the supporting document (**Figures S7-S10**; **Table S1)**. An additional stability analysis characterizing each systems geometry was performed which is discussed in the supporting results (**Table S2-S6**).

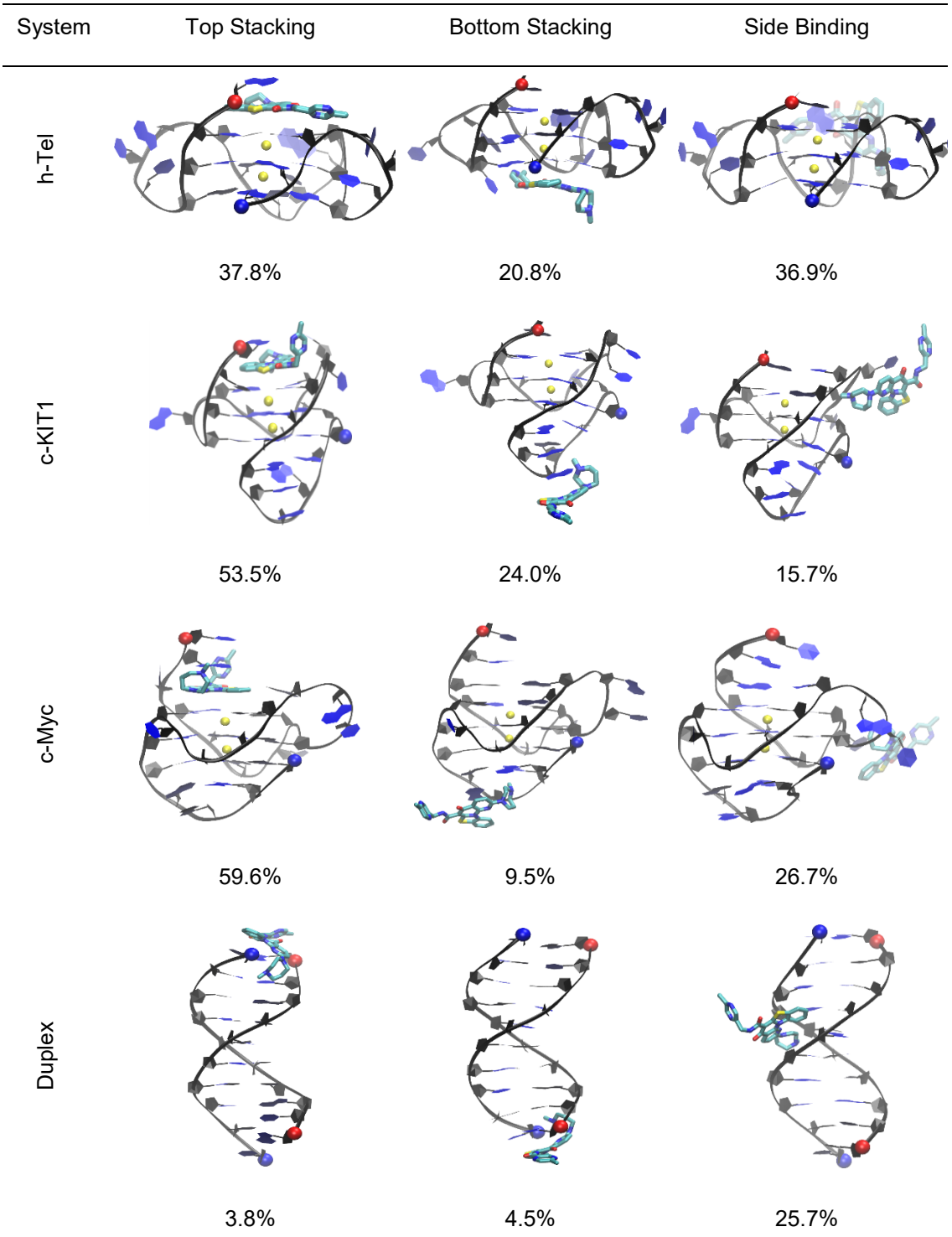


Figure 3. Major binding modes of CX-5461 binding to the human telomeric G4 (PDB ID: 1KF1), c-KIT1 G4 (PDB ID: 4WO3), c-Myc G4 (PDB ID: 2MGN) and duplex DNA. 5' and 3' of the DNA chain are indicated by a red and blue spheres, respectively. K^+ ions are indicated by yellow spheres. Overall population abundance of each binding mode is annotated.

Three binding poses were identified from the clustering analysis for each of the four complex systems. The centroid structure of each structural family and percentages are presented in **Figure 3**. In addition, two dimensional interaction diagrams along with top and side views of the most abundant clusters of each system are presented in the supporting document (**Figure S11 A-D**) and **Figures S12-S13** compare the top and groove clusters for each system, respectively. Evident from our results, the top binding mode was most abundant for the G4 systems and the groove binding mode was most abundant for the duplex system. Interestingly, CX-5461 appears to show selectivity to the G4s over the duplex structure here, where CX-5461 is bound to the duplex in only ~35% the simulation, and the remained of the simulation period CX-5461 is in an unbound state. A detailed discussion on the most abundant conformation can be found in the supporting results.

The MM-PBSA binding energy analysis revealed that the top intercalating binding mode for the human telomeric DNA G-Quadruplex was most energetically favorable of all systems. Of the four systems, the top intercalating mode of the human telomeric DNA G-quadruplex (1KF1) was shown to be the most energetically favorable binding mode at -37.0 kcal/mol (**Table 2**), closely followed by the bottom binding mode of the human telomeric system -33.6 kcal/mol. A more detailed discussion on the MM-PBSA results can be found in the supporting results. CX-5461's average MM/PBSA binding energy over the major poses were calculated for the human telomeric (-28.6 kcal/mol), c-KIT1 (-23.9 kcal/mol), c-Myc (-22.0 kcal/mol) G4s, and DNA duplex (-15.0 kcal/mol) systems where each systems propagation of error was calculated, using the standard equation for error propagation for addition of measured quantities, as 4.97, 2.72, 5.41, and 5.07 respectively. The duplex groove binding mode was the least energetically favorable, measuring -15.0 kcal/mol. These results suggest that CX-5461 is selective to DNA G-quadruplex systems over DNA duplex because each of the most favorable binding poses for the G4 systems is at least twice as energetically favorable as the duplex groove binding mode. Further the less favorable MM-PBSA binding energy of the G4 systems suggests that side binding may be an intermediate state. The binding energy was broken down into van der Waals (VDW), hydrophobic interactions (SUR), electrostatic interactions (GBELE), and the conformational energy change induced from the complex formation (CONF) (**Table 2**). From the table appears that the hydrophobic interactions ($\mathbf{\Delta_{SUR}}$) contribute the most to the total energy so it was not surprising that the top intercalation binding modes have the most energetically favorable hydrophobic interactions and are much more favorable than the duplex systems groove binding mode. However, the binding free energy calculation remains a challenging task and MMPBSA generally overestimates the binding free energy due to lack of the entropy term.

Table 2. MM-PBSA binding energy of CX-5461 binding to the human telomeric G-quadruplex (1KF1), c-KIT1 G-quadruplex (4WO3), c-Myc G-quadruplex (2MGN), and duplex DNA in each of the major binding modes.

System	Mode	Δ_{VDW}[1]	Δ_{SUR}	Δ_{PBELE}	Δ_{CONF}	Δ_{TOT}	$\Delta\Delta E$[2]	ΔT_m[3]
h-Tel	Top	-9.3±6.3	-32.9±3.1	10.5±1.3	-5.4±3.8	**-37.0±3.9**	0	30
	Bottom	-3.5±5.9	-22.3±2.4	6.0±3.8	13.8±4.6	-33.6±2.7	3.4	
	Side	-4.5±3.6	-23.7±2.0	11.3±2.3	1.5±8.0	-15.3±1.5	21.7	
c-KIT1	Top	-4.4±0.9	-25.4±0.5	8.2±0.6	-11.5±2.7	-33.1±1.7	3.9	27
	Bottom	-2.0±1.6	-11.0±0.7	5.1±1.0	-14.1±2.3	-22.0±1.4	15	
	Side	-3.0±1.9	-12.8±0.7	7.1±1.4	-7.8±11	-16.5±1.6	20.5	
c-Myc	Top	-13.7±4.3	-35.2±1.5	16.4±0.7	-0.1±3.9	-32.6±4.7	4.4	25
	Bottom	-4.3±0.8	-16.0±0.2	9.0±0.5	-4.3±4.7	-15.5±1.0	21.5	
	Side	-5.3±6.2	-25.7±3.0	15.9±1.6	-2.7±1.4	-17.8±2.5	19.2	
Duplex	Groove	-2.1±4.1	-23.1±2.1	8.2±2.3	3.2±3.9	-13.8±3.9	23.2	10
	Top	-4.3±0.8	-15.4±0.7	3.8±0.9	0.1±0.9	-16.1±2.7	20.9	
	Bottom	-3.8±1.2	-15.9±1.8	4.6±2.3	0.02±0.8	-15.1±1.8	21.9	

[1] The parameters in this table are reported in units of kcal/mol.
ΔVDW = Change of VDW energy in gas phase upon complex formation
ΔSUR = Change of energy due to surface area change upon complex formation
ΔEBELE = Change of GB reaction field energy + gas phase Elec. energy upon complex formation
ΔCONF = Change of conformational energy upon complex formation
ΔTOT = ΔVDW +ΔSUR + ΔEBELE + ΔCONF Change of potential energy in water upon complex formation
[2] $\Delta\Delta E^2 = (\Delta_{TOT} - (\mathbf{-37.0}))$
[3] Experimental melting temperature increase (in Kelvin) of each system with 10 μM CX-5461 from a FRET melting temperature assay[4]. Values are estimated based on the figure in the literature.

A Markovian State Model Analysis was performed to gain kinetic insight into the binding pathways. To characterize the binding of CX-5461 to the DNA G4s and duplex we performed a Markovian State Model (MSM) analysis. **Table 3** summarizes the transition times between each state for each system. The ratios of the on and off rate constants estimated are off (underestimate of the equilibrium constant) due to the fact that the limited simulation period, although extensive for a study of this kind, did not reach a true equilibrium. Thus we focus on the qualitative trend. We present the MSM along with representative trajectories and order plots for each of the thermodynamically favorable state in each system: human telomeric (**Figures 4-6), c-KIT1 (Figures 7-8), c-Myc (Figures 9-10), and duplex (Figure 11;S14)**. In addition to the order parameter plots, we also characterized the overall

geometry of the systems which can be found in the supporting results and **Tables S2-S5**. MSMs can be built from MD simulation data and are a comprehensive statistical approach used to create understandable yet high-resolution models of the intrinsic kinetics of a system.[88] Our MSM analysis follows a similar procedure to that of our previous work[74] which also examined the binding pathways and kinetic information of G-quadruplex structures. Our implied timescales of each cluster for all lag times of each system are presented in **Figure S15**. Due to choosing to cluster into a handful of "macrostates" and directly and skipping over the experimentally unverifiable thousand "microstates", the expected convergence time of the implied timescales should be significantly greater than that of a model with a greater number of clusters. This results in a coarser grained model that trades finer detail for greater experimental testability and easier human understanding [88, 89]. It is likely that directly clustering into "macrostates" still maintains the integrity of the MSM as verification through the Chapman-Kolmogorov test (**Figure S16A-D**) indicates that the model closely resembles the observed simulation data. Two-dimensional network models are also presented in the supporting document for each system (**Figure S17**).

Table 3. Parallel pathways of each systems Markovian State Model.

Parallel Pathways[1]	**Human telomeric[2]**		**c-KIT1[2]**		**c-Myc[2]**		**Duplex[2]**	
	Forward	Reverse	Forward	Reverse	Forward	Reverse	Forward	Reverse
U-T	3.3	5.1	1.2	3.7	1.4	3.9	0.1	3.0
U-B	4.0	2.9	3.4	2.2	16.7	2.3	0.1	2.7
U-S	1.2	3.8	1.6	3.0	1.0	2.4	2.0	5.7
U-S-T	4.3	6	2.7	5.2	2.4	4.9	-	-
U-S-B	6.9	5.8	3.9	7.9	-	-	-	-
U-B-S-T	-	-	6.8	9.3	-	-	-	-

[1]Each parallel pathway described as unbound (U), top (T), bottom (B), or side (S) with approximate interstate flux listed.

[2]Values reported in μs

The MSM revealed multiple parallel pathways toward the most thermodynamically favorable end stacking modes from unbound state in the human telomeric system. For the human telomeric G4 system, our calculated first mean passage times indicated that the pathway from unbound directly to the top binding state was the most favorable and slightly faster (3.3 μs) than unbound directly to the bottom binding state (4 μs) as well as both transition states. We present two representative trajectories to characterize the binding of CX-5461 to the top binding mode and one showing the side binding transition state ending in a top binding mode. The representative top binding mode of CX-5461 appears to undergo an induced fit binding mechanism (**Figure 5**). Detailed descriptions of the representative trajectory and order parameter plot can be found in the supporting results as well as an additional trajectory of CX-4561 in the bottom binding mode (**Figure S18**).

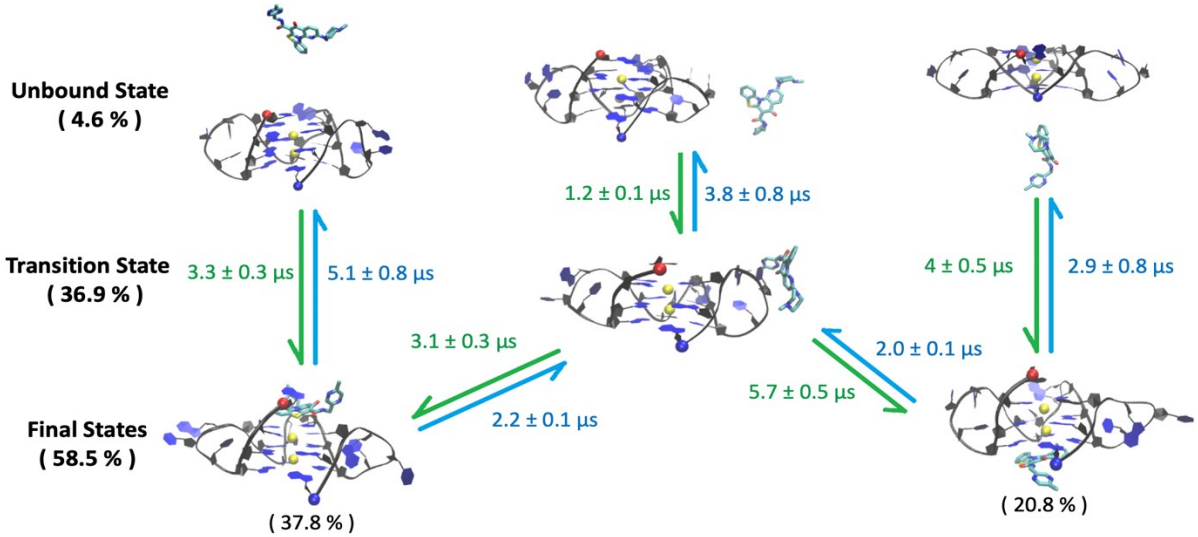


Figure 4. The mean first passage times between the four states (unbound, side transition, top, and bottom) of the human telomeric DNA G-quadruplex and CX-5461 complex system.

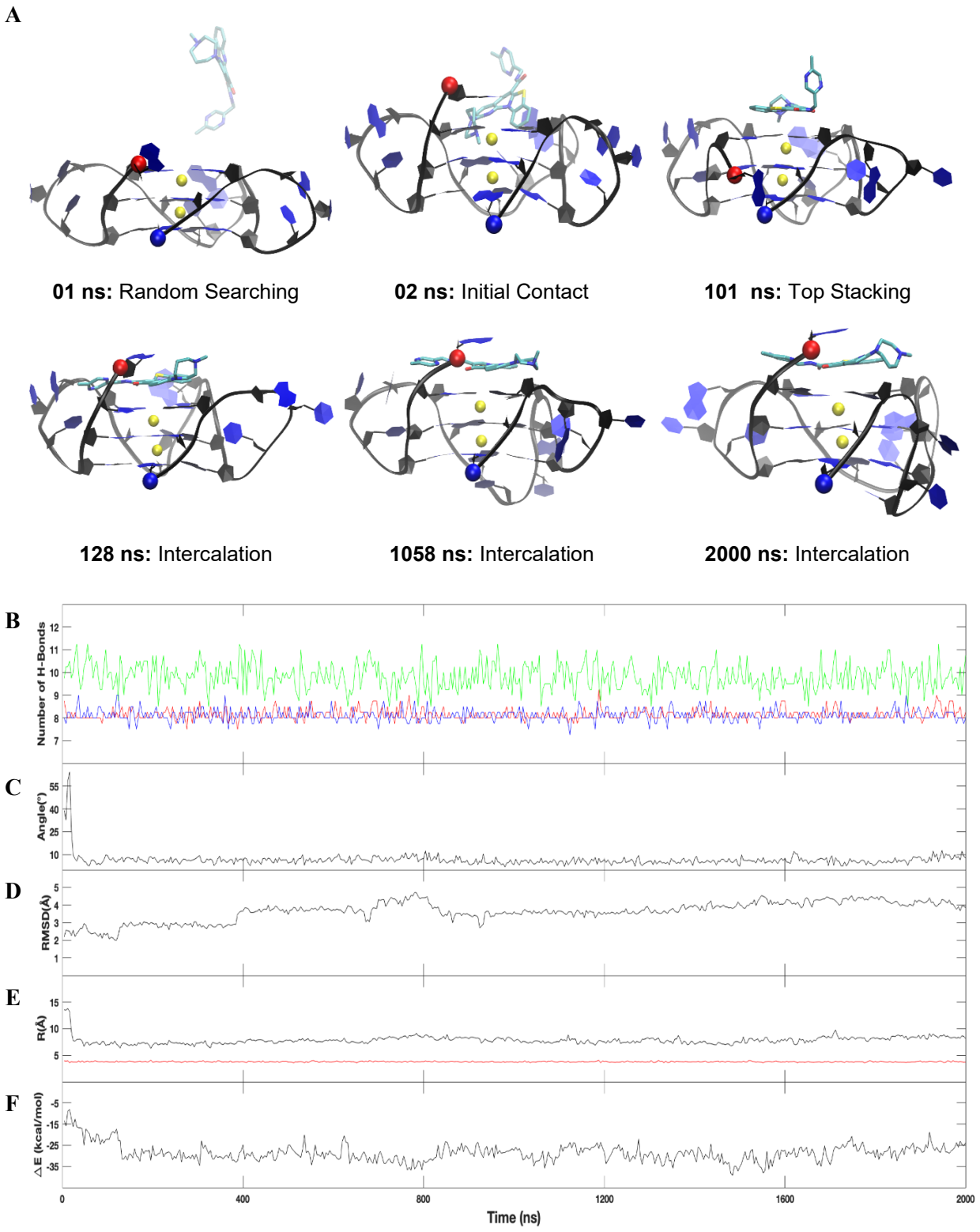
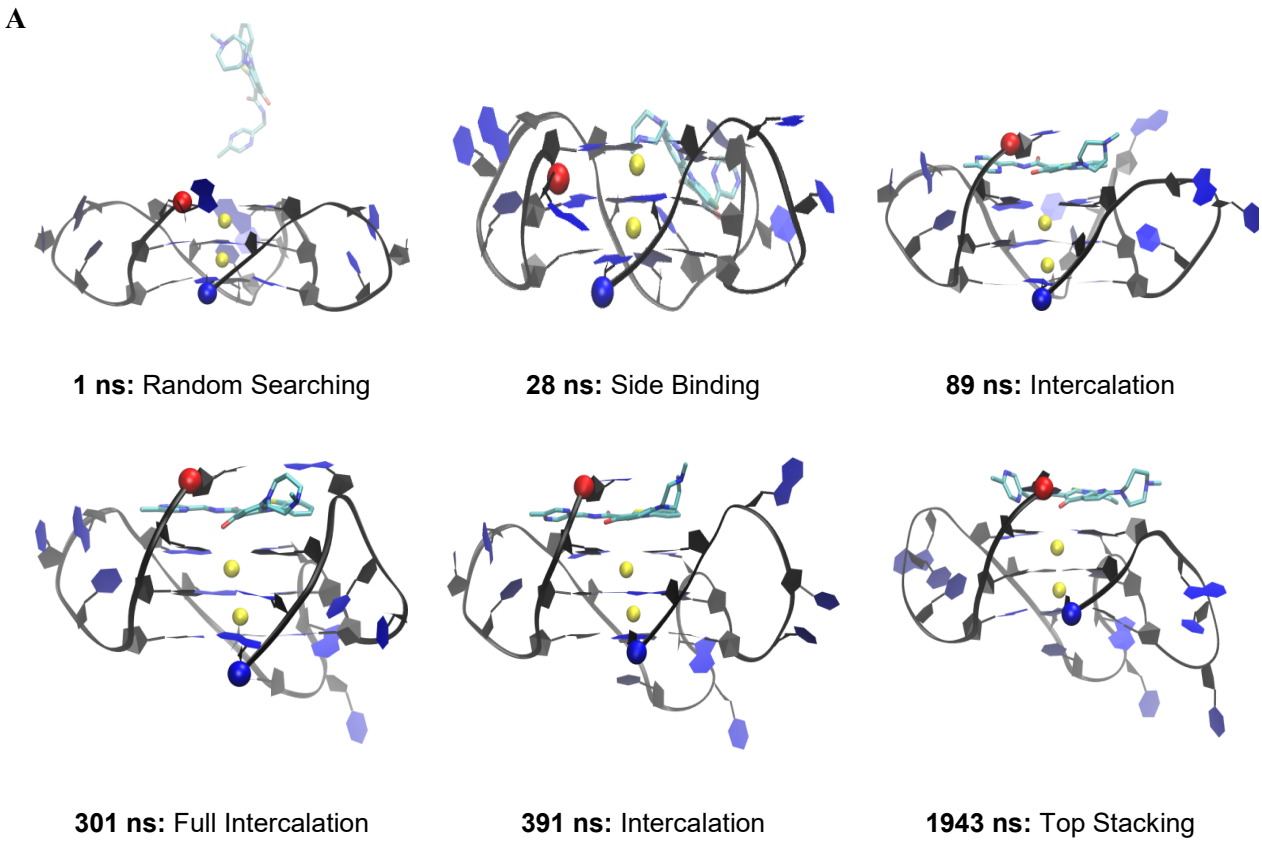


Figure 5. Snapshots from the representative trajectory of the top binding mode of the human telomeric G4 (A), run 13, including an order parameter plot (B-F). Illustrated in the plot is the breaking and reforming of hydrogen bonds per layer (5' Face in Red, Middle G-Tetrad in Blue, 3' Face in Green) (B), drug-base dihedral angle (C), DNA RMSD (Å) with reference to the initial structure (D), ligand center to DNA center distance (black) and K^+ to K^+ distance (red) (E), and the MM-GBSA binding energy (ΔE in kcal/mol) (F). 5' and 3' ends of the DNA chain are indicated by a red and blue ball, respectively. K^+ ions are indicated by yellow spheres.



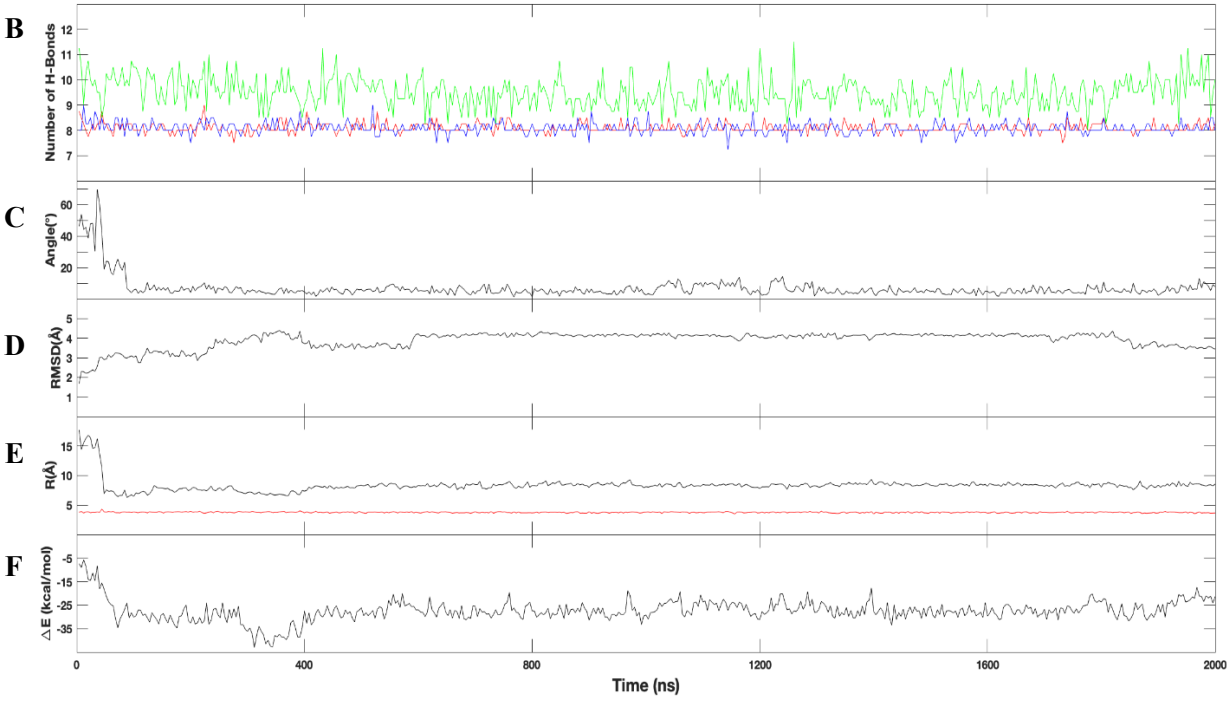


Figure 6. Snapshots from the representative trajectory of the top binding mode of the human telomeric G4 (A), run 05, including an order parameter plot (B-F). Illustrated in the plot is the breaking and reforming of hydrogen bonds per layer (5' Face in Red, Middle G-Tetrad in Blue, 3' Face in Green) (B), drug-base dihedral angle (C), DNA RMSD (Å) with reference to the initial structure (D), ligand center to DNA center distance (black) and K^+ to K^+ distance (red) (E), and the MM-GBSA binding energy (ΔE in kcal/mol) (F). 5' and 3' ends of the DNA chain are indicated by a red and blue ball, respectively. K^+ ions are indicated by yellow spheres.

The MSM revealed multiple parallel pathways toward the most thermodynamically favorable top binding mode from unbound state in the c-KIT1 system. Each of the pathways led to a thermodynamically favorable top binding state which accounted for 53.5 % of the simulation period and occurred in 1.2 μs. We believe the MSM determined that the bottom binding mode is not a thermodynamically favorable state in this system because of the limited simulation period, however we expect that if the simulations were extended further, a thermodynamically favorable bottom binding site would be seen. Since the only thermodynamically favorable state is the top binding mode, we present one representative trajectory of this mode (**Figure 8**) in the main text. Detailed descriptions of the representative trajectory and order parameter plot can be found in the supporting results as well as additional trajectories for the top (**Figure S19**) and bottom binding modes (**Figure S20-21**) which show CX-5461 binding to both the outside of the bottom loop as well as inside of the bottom loop, but not actually interacting with the G4 core, and the side to top binding mode (**Figure S22**).

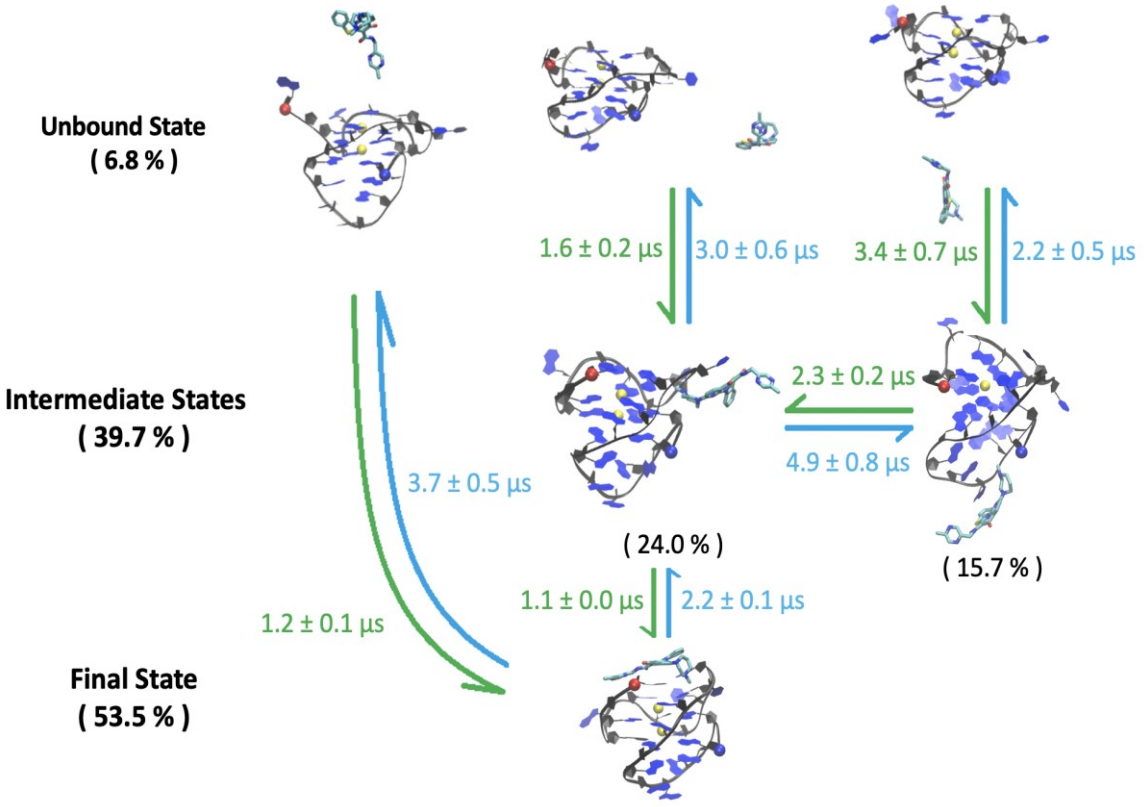
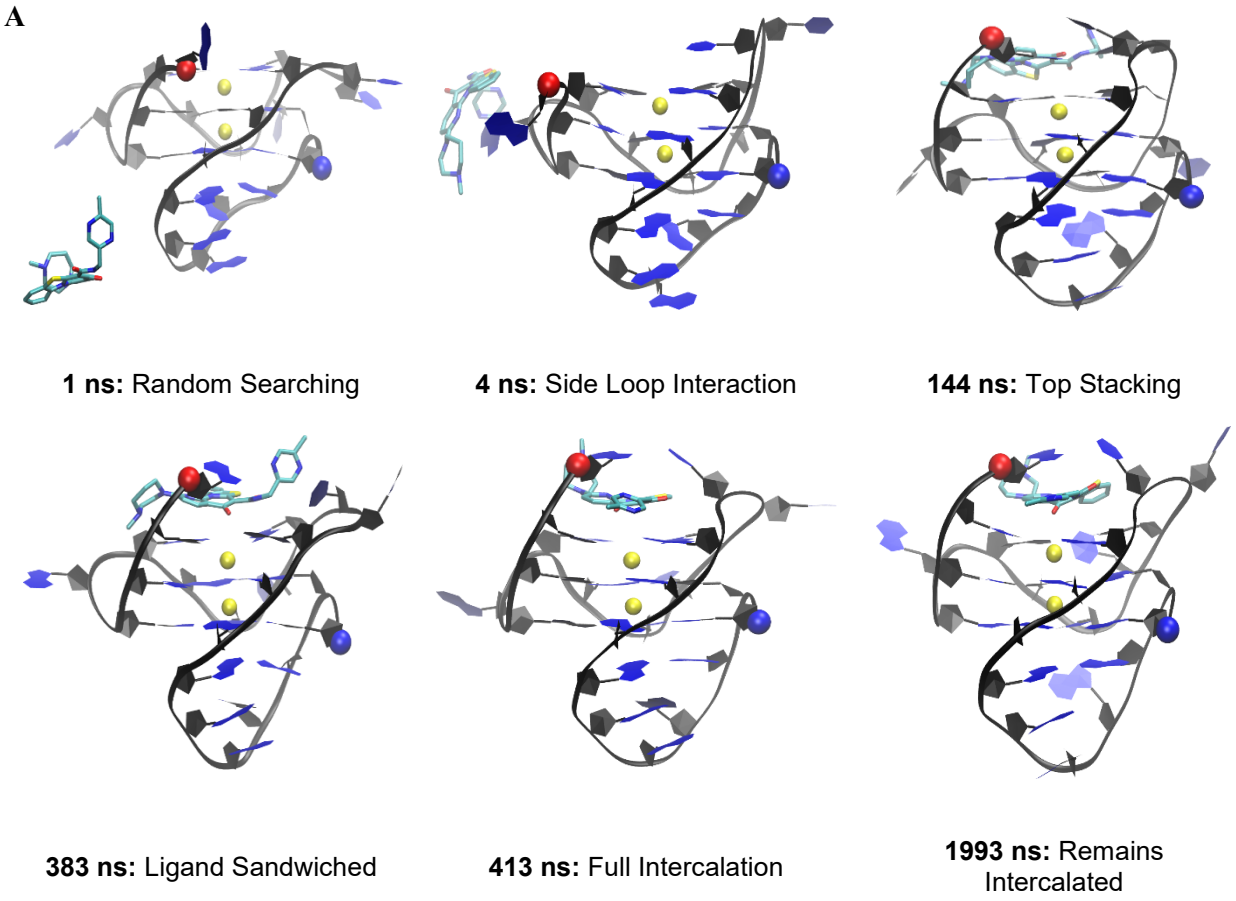


Figure 7. The mean first passage times between the four states (unbound, side transition, top, and bottom) of the c-KIT1 DNA G-quadruplex and CX-5461 complex system.



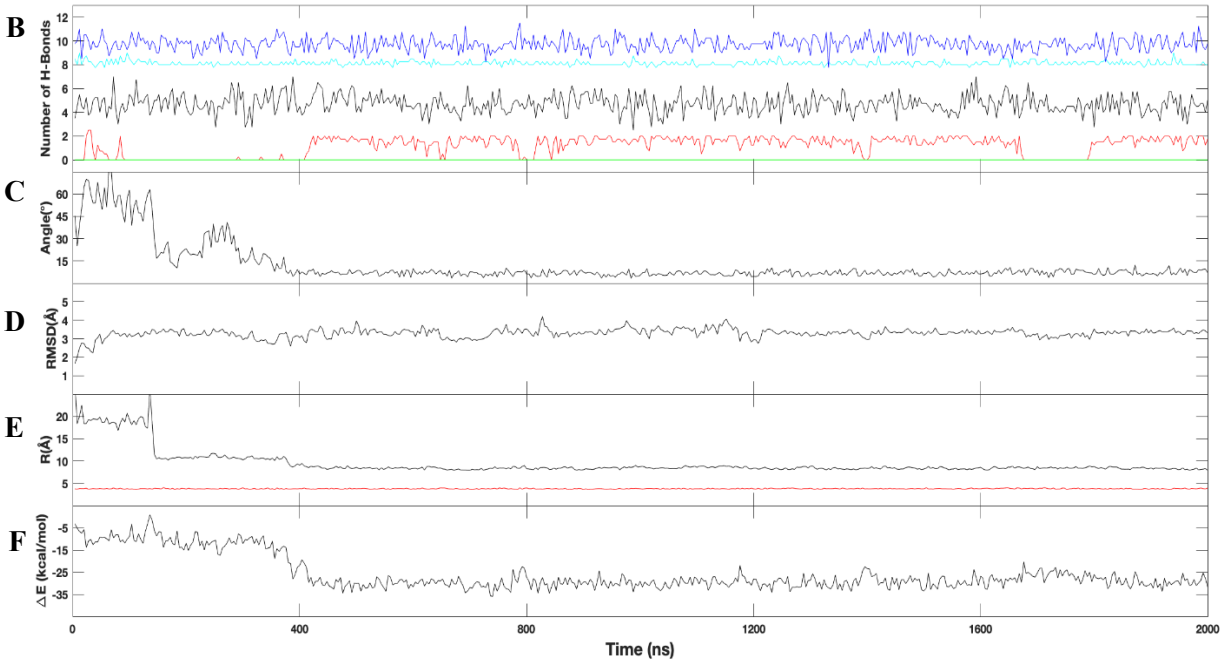


Figure 8. Snapshots from the representative trajectory of the top intercalating binding mode of the c-KIT1 G4 (A), run 19, including an order parameter plot (B-F). Illustrated in the plot is the breaking and reforming of hydrogen bonds per layer (5' Face in Red, Top G-Tetrad in Cyan, Middle G-Tetrad in Blue, Bottom G-Tetrad in Black, and 3' Face in Green) (B), drug-base dihedral angle (C), DNA RMSD (Å) with reference to the initial structure (D), ligand center to DNA center distance (black) and K^+ to K^+ distance (red) (E), and the MM-GBSA binding energy (ΔE in kcal/mol) (F). 5' and 3' ends of the DNA chain are indicated by a red and blue ball, respectively. K^+ ions are indicated by yellow spheres.

The MSM revealed multiple parallel pathways toward the most thermodynamically stable top binding mode from unbound state in the c-Myc system, where as the bottom stacking mode appears to be an off pathway intermediate. For the 2MGN system, three major parallel pathways were observed (**Figure 9)**: unbound to top, unbound to bottom, and unbound to side transition and ending in a top binding mode. Unique to this system, the bottom binding pose appears to be highly unstable and likely acts as an off pathway intermediate state where CX-5461 binds to the bottom from an unbound state and once again goes back to the unbound state and follows one of the other pathways leading to the thermodynamically favorable top binding mode. The interaction between the c-MYC G4 and CX-5461 is particularly interesting because of the clear use of a base flipping mechanism by both the terminal residues, T1, G2, and A3, but also the loop residue A12 (**Figure 10**). Detailed descriptions of the representative trajectory and order parameter plot can be found in the supporting results. In addition, there are additional trajectories presented in the supporting document for the top (**Figure S23**) and bottom binding modes (**Figure S24-25**) which show CX-5461 binding to both the outside of the bottom loop as well as inside of the bottom loop, but not actually interacting with the G4 core. In addition to this, a trajectory is presented in the supporting document to show the side to top binding mode (**Figure S26**).

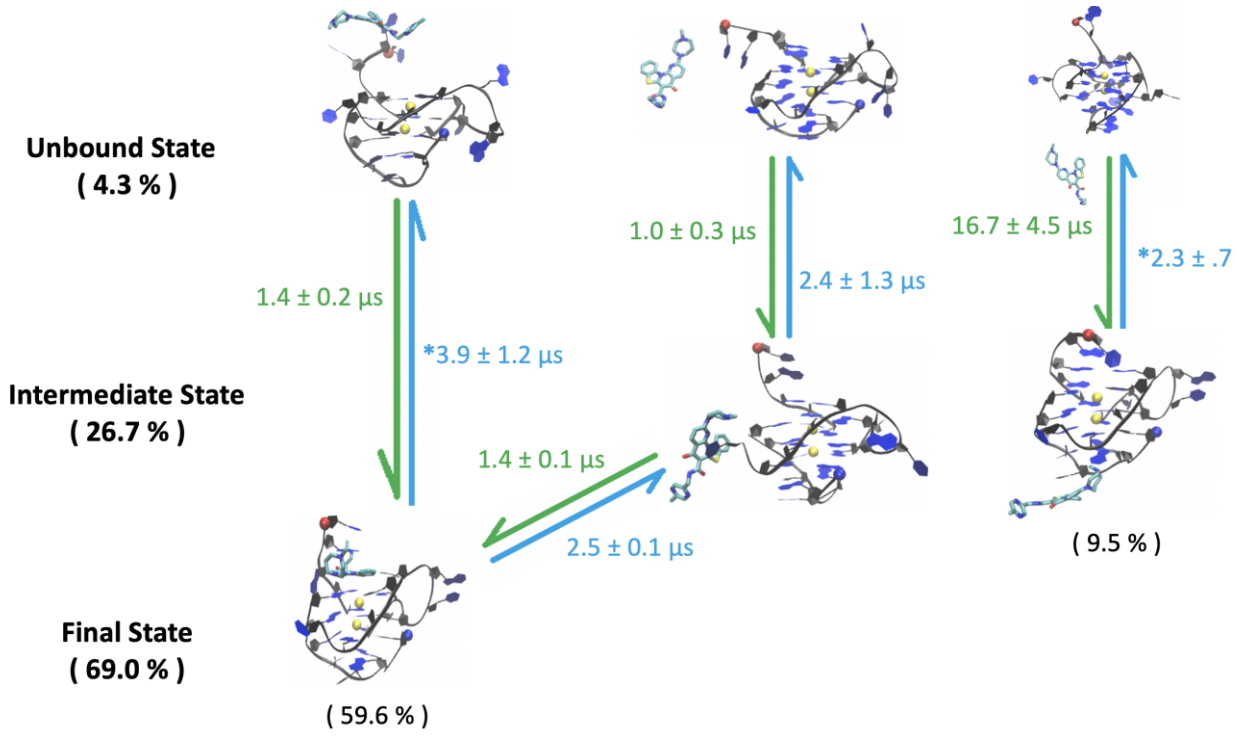


Figure 9. The mean first passage times between the four states (unbound, side transition, top, and bottom) of the c-MYC DNA G-quadruplex and CX-5461 complex system.

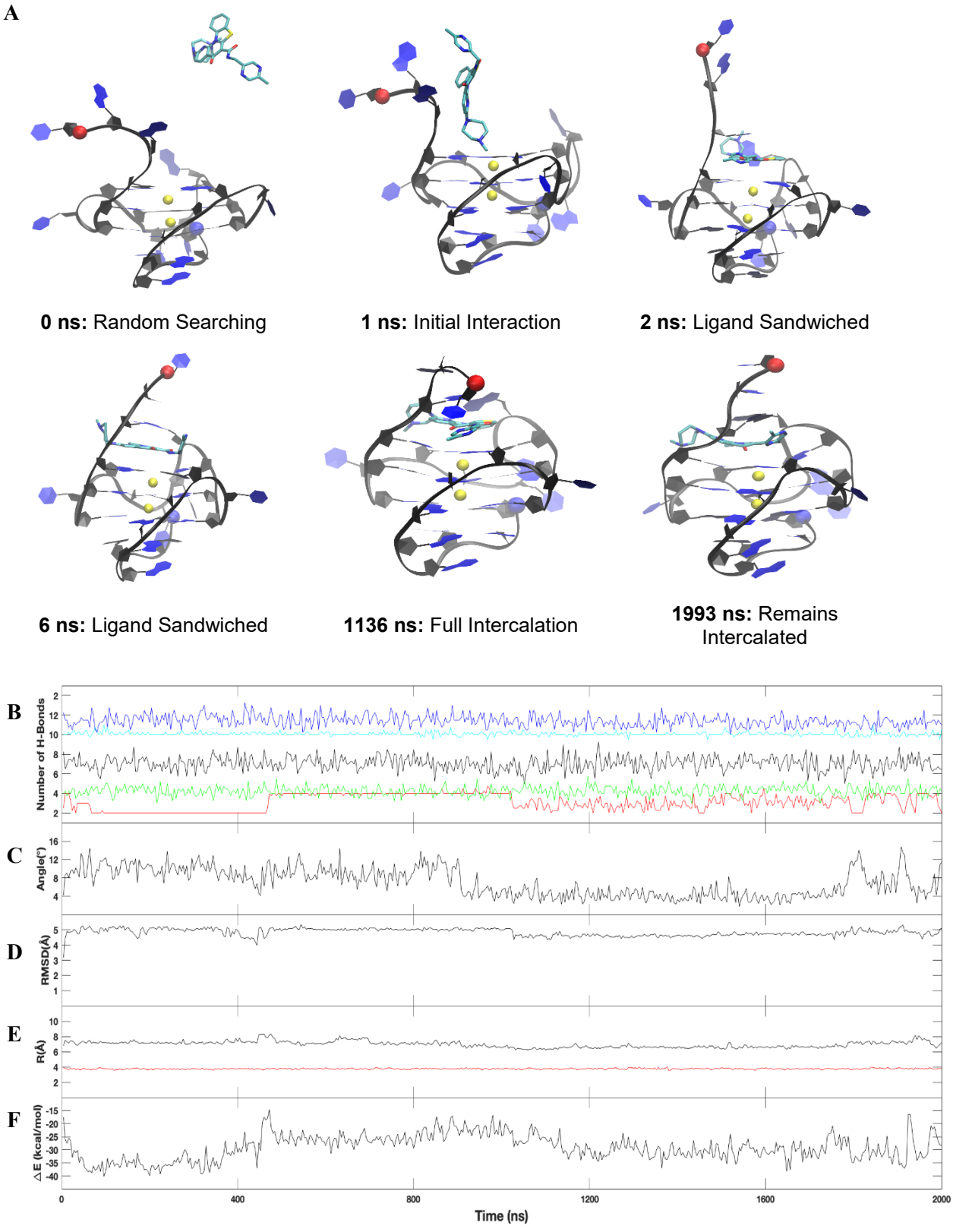


Figure 10. Snapshots from the representative trajectory of the top intercalating binding mode onto the c-Myc G4 (A), run 14, including an order parameter plot (B-F). Illustrated in the plot is the breaking and reforming of hydrogen bonds per layer (5' Face in Red, Top G-Tetrad in Cyan, Middle G-Tetrad in Blue, Bottom G-Tetrad in Black, and 3' Face in Green) (B), drug-base dihedral angle (C), DNA RMSD (Å) with reference to the initial structure (D), ligand center to DNA center distance (black) and K^+ to K^+ distance (red) (E), and the MM-GBSA binding energy (ΔE in kcal/mol) (F). 5' and 3' ends of the DNA chain are indicated by a red and blue ball, respectively. K^+ ions are indicated by yellow spheres.

The MSM revealed multiple parallel pathways toward the most thermodynamically stable groove binding mode from unbound state in the duplex system. The major pathways for the duplex system include one thermodynamically favorable groove biding state (25.7%) as well top and bottom states that end up in a groove binding mode (**Figure S14**). Since the groove binding mode is the only one of physiological relevance for a long chromosome DNA, pathways leading to this mode are discussed here. The representative trajectory for the groove binding mode of CX-5461 to the DNA duplex is shown in **Figure 11**. Interestingly in that simulation run we observed CX-5461 scaling the surface of the DNA duplex, unable to find a stable binding pose, repositioning around the grooves of the DNA duplex while maintaining interaction during the entire binding process. Three additional trajectories for this binding mode are presented in the supporting document (**Figure S27-S29**). In addition, hydrogen bond networks for each system are also presented (**Figures S30-S32**).

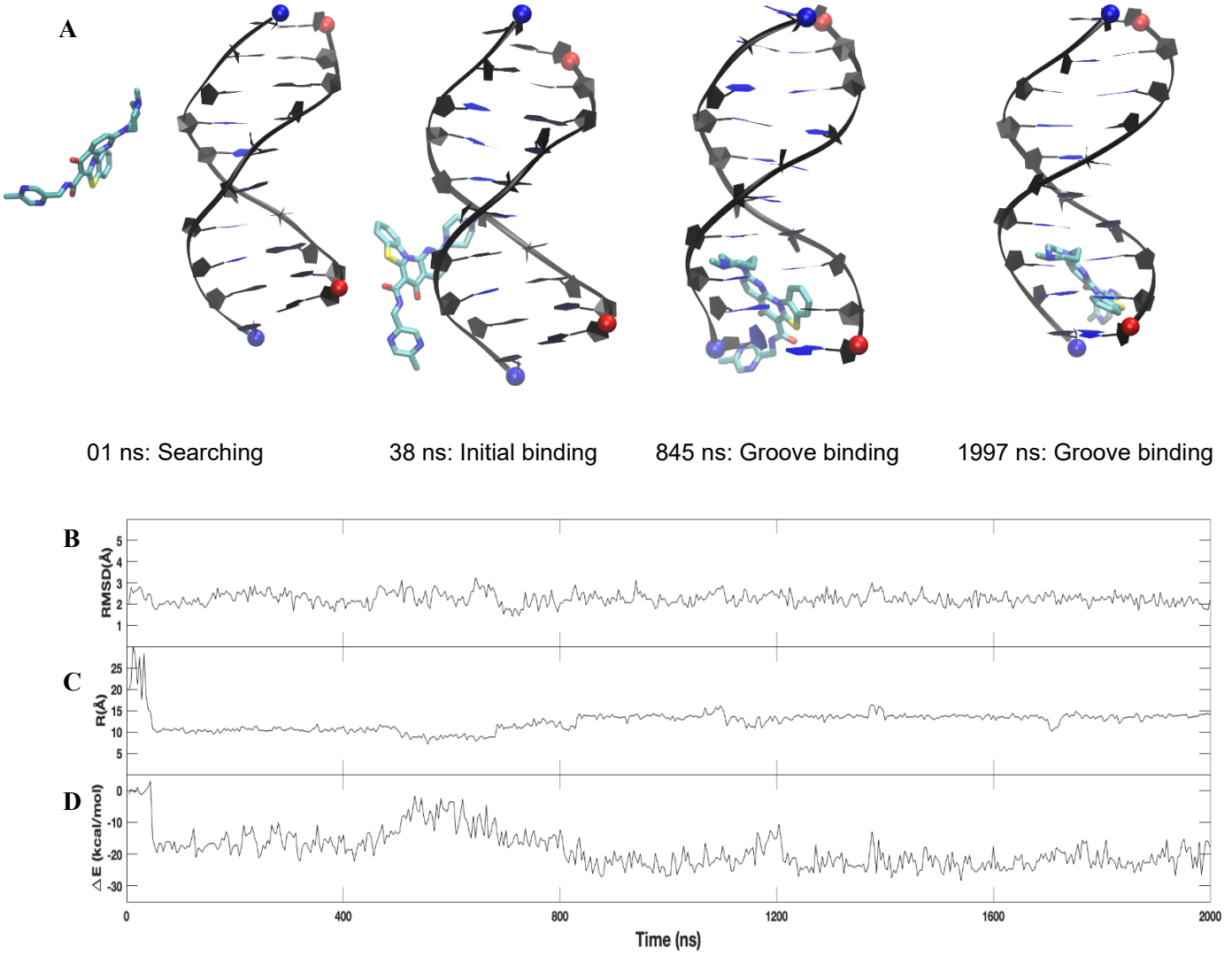


Figure 11. Snapshots from the representative trajectory of the bottom binding mode of the DNA duplex system (A), run 12, including an order parameter plot (B-D). Illustrated in the plot is the DNA RMSD (Å) with reference to the initial structure, ligand center to DNA center distance (C) and the MM-GBSA binding energy (ΔE in kcal/mol) (D). 5' and 3' ends of the DNA chain are indicated by a red and blue sphere, respectively.

The root mean square fluctuation (RMSF) data showed the connecting loops of each G4 and the termini of the duplex DNA had larger fluctuations than the G-quadruplex core; CX-5461 binding slightly reduced overall fluctuations. The RMSF plots for each system as well as a topology model which shows the position of each residue for which the RMSF was calculated and presented in **Figure 12**. For the human telomeric system three large (~5 Å) peaks were identified which correspond to three loops connecting the G-quartets: $T^{5}T^{6}A^{7}$, $T^{11}T^{12}A^{13}$, and $T^{17}T^{18}A^{19}$. Evident form the data here, the apo form has higher overall fluctuations when compared to the complex systems. Overall, there are two pronounced differences in the data series. The first is the lower fluctuations of the $T^{11}T^{12}A^{13}$ loop in the apo form simulation run, which may be due to this loop frequently flipping outward to clear room for CX-5461 binding. The second is in the bottom binding pose where the fluctuation of the $T^{17}T^{18}A^{19}$ loop has significantly decreased fluctuation in residues T^{17} and T^{18}, which may be due to a stabilizing effect because of CX-5461 directly interacting with residue G^{16} in this binding mode. In addition we see here that the terminal residues have roughly equal fluctuations due the structural similarity on each face of the G4. In the c-KIT1 systems the major peaks corresponded to residues of the connecting loops: A^{5}, C^{9}, C^{11}, T^{12} and G^{17}, G^{18} and A^{19}. Overall, the apo form trajectory had higher fluctuations than the complex systems indicating that CX-5461 binding slightly reduced the overall fluctuations at these peaks. The once exception was at loop C^{11}, T^{12} where the top binding mode has this region with higher fluctuations. This is due to this loop flipping upward and forming a hydrogen bond between R1 of CX-5461. Here we clearly see a difference in the terminal residues where the 5' terminal is significantly higher than the 3' terminal, which is likely due to the fact that the 3' terminal is actually a residue of the 3' G-tetrad whereas the 5' terminal is a free residue. There were 5 major peaks identified for the c-Myc systems corresponding to residues $T^{1\text{-}A3}$, T^{7}, $T^{10\text{-}A12}$, T^{16}, and A^{21}. Each residue showing significant fluctuation was either a part of the large 5' terminal segment (T^{1}, G^{2}, A^{3}), or a part of a connecting loop. In the c-Myc systems, the RMSF values were comparable in each peak except for the T^{10}, G^{11} and A^{12} region. The top binding mode had a high fluctuation at residue T^{10} but lower fluctuations at G^{11} and A^{12}, this is likely due to this loop flipping upward to stabilize the binding of CX-5461. Where the higher fluctuations of T^{10} are a result of the further distance travelled to flip upward compared to G^{11} which flips outward and A^{12} which directly binds to CX-5461. The bottom system had the largest fluctuation at T^{10} and G^{11} which flip outward during the simulation and the receptor only system had equal fluctuations at residues T^{10}, G^{11} and A^{12} which are also flipped outward. The terminal residues are most notably different here where the long 5' segment is highly flexible (5-6 Å) and the 3' terminal, which is a part of the 3' G-quartet has very low fluctuations (~1 Å). The duplex DNA systems (grey and black) showed peaks at residues G^{1}, C^{10}, G^{11}, and C^{20}, which are the terminal residues of the double helix. Overall, we see that the human telomeric G4 loops have the highest fluctuations of all systems which may help to explain the less favorable binding energy of the side binding modes compared to the other systems and the fluctuations of the apo G4 forms are larger than the complex systems, suggesting that CX-5461 stabilizes the G4 DNA upon binding.

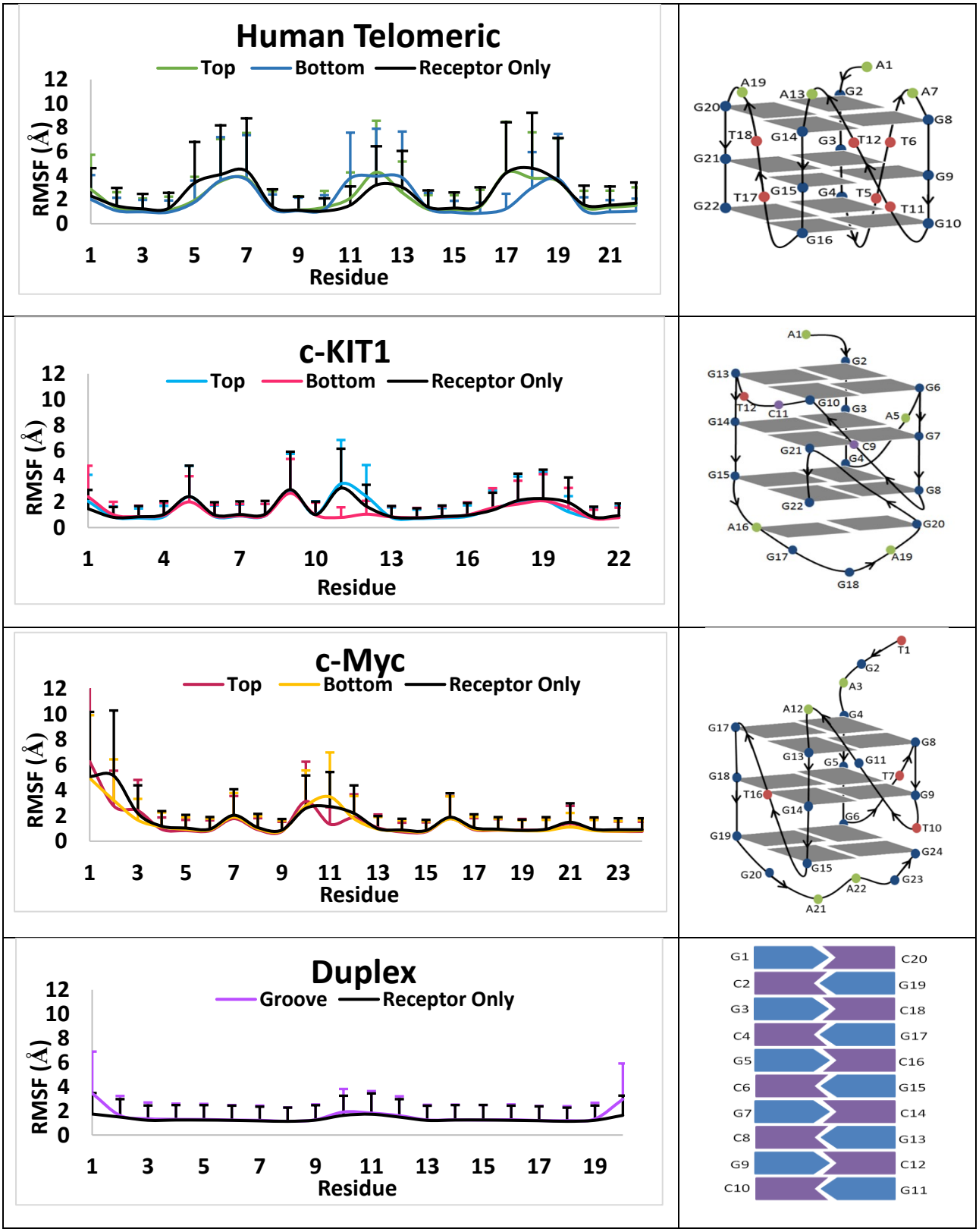


Figure 12. The root mean square fluctuation (RMSF) plot of the 11 systems (receptor only, plus top and bottom binding modes of 1KF1, 2MGN, and 4WO3, groove binding mode of the duplex system) with topology include on the right for reference, where error bars represent standard deviation.

DISCUSSION

DNA G-quadruplexes (G4s) are overrepresented in a wide variety of cancers making them a prime therapeutic target. [90] Stabilizing G4s of cancerous cells has become a novel therapeutic technique that has been shown to inhibit cancerous cell growth and replication leading to the cancerous cells death. [91] A promising new anti-cancer medication, CX-5461 is in in Phase I/II clinical trials for advanced hematologic malignancies and cancers with BRCA1/2 deficiencies [4]. Experimental evidence identified CX-5461's specific roles at G4s as the ability to bind to and stabilize G4 structures in human cancer cell lines, increase the number of G4 sites, as well as selectivity bind to G4 structures on the human genome.[4] These properties are extremely advantageous for cancer therapeutics and evident from recent work CX-5461 is a promising therapeutic agent for a variety of targets, however the detailed binding mode and mechanism for which CX-5461 interacts with the G4 structures remains elusive. Our analyses sought to characterize the binding of CX-5461 to the human telomeric, c-KIT1, c-Myc G4s as well as a DNA duplex to provide insight into the binding process and help to explain the experimentally reported order of binding induced stability.

Our simulation analysis identified three major binding features conserved in each G4 system. First, we observed that 5' intercalation was the most thermodynamically favorable binding mode. This was evident from the MM/PBSA binding energy (**Table 2**) and trajectory clustering (**Figure 3**). Second, the Markovian State Model (MSM) of each G4 system revealed that there were multiple parallel pathways all leading to the thermodynamically favorable top intercalation mode (**Figures 4, 7, 9**). Each of which include a direct pathway from the unbound state to top binding as well as indirect pathways which could be using side binding as an on-pathway intermediate state. The lower MM/PBSA binding energies of the side binding modes (**Table 2**) provide thermodynamic support of the side binding intermediate states. The third common binding feature was that the top binding mode of each system clearly was not achieved through the use of a lock-key mechanism. Instead, each G4 system demonstrated the use of a combination of the induced fit and conformational selection mechanisms. In more detail, our apo form simulations (**Figure S2-3**) show that because of the short flexible segments at the 5' face of the G4's there are some instances of base flipping observed for both the 5' terminal and 5' face loop residues. However, when CX-5461 binds, the population is shifted toward the population which these bases stabilize and stack onto the G4 core or CX-5461 which clearly shows evidence of a conformational-selection mechanism. In addition, the most thermodynamically stable state is only achieved by a base flipping insertion mechanism where terminal and side loop residues flip outward to make room for CX-5461 insertion followed by the bases flipping back to make contact with CX-5461, which demonstrated the use of an induced fit binding mechanism not observed in the apo form. A clear example is the interaction between the c-Myc G4 and CX-5461 which shows the use of a base flipping mechanism by both the terminal residues, T1, G2, and A3, and also the loop residue A12 (**Figure 10**), which is supported by the RMSF (**Figure 12**) and 2D interaction diagrams (**Figure S11**). **Therefore,** although there are structural differences between the G4s, these same general binding characteristics of CX-5461 are conserved which we suggest could be applicable to other targets not studied here.

Three notable differences were also extracted from our simulation analysis. It was evident that the structural features of each system caused subtle differences in binding modes (**Table S7**), binding energies (**Table 2**), and MSMs (**Figures 4, 7, 9**). First, the human telomeric system's structural similarities on the 5' and 3' faces allowed equal binding opportunities for both sites. The human telomeric system was the only MSM to have parallel pathways leading to the bottom binding pose as a thermodynamically favorable state (**Figure 4**), and the only system to have the binding energy of the bottom binding mode be comparable to the top binding mode (**Table 2**). The conclusion made by comparing the bottom binding modes for each G4 system leads to the second difference: each 3' site is different and as a result has a different effect on CX-5461 binding, which we discuss in more detail further in the discussion. The third major difference is the trend of the MSMs. As mentioned, the structural similarities on each face of the human telomeric system clearly differentiated it from the c-KIT1 and c-Myc, but the two G4s without the structural similarities on each face also had stark differences from each other. Most notably, the c-Myc systems bottom binding pose was determined to be the most thermodynamically unstable by the MM/PBSA binding energy analysis. This is demonstrated in the MSM where the bottom binding mode is a lowly abundant state which we suggest likely acts as an off pathway intermediate state. We believe this happens by CX-5461 binding to the bottom of the G4 from an unbound state and once again goes back to the unbound state and follows one of the other pathways which leads to the thermodynamically favorable top binding mode. In contrast to the c-KIT1 G4 system which utilizes the bottom binding mode as one of its on-pathway intermediates before leading to the more thermodynamically favorable top binding mode. Together, understanding these similarities and differences allowed us to provide insight into the binding mode and mechanism of CX-5461 and make sense of the experimental phenomenon published on CX-5461's affinity and selectivity to these targets.

Our simulation protocols and final binding poses were validated by comparing the binding of CX-5461 to the c-Myc promoter G4 in our study to the only experimentally solved structure of these G4s in complex with a small molecule to date. This was the binding pose of Phen-DC_3 in complex with Pu24T (PDB ID: 2MGN) which was solved using NMR spectroscopy [92]. Our study used Pu24T, the intramolecular G-quadruplex formed from the c-Myc promoter, from the Phen-DC3-Pu24T complex and simulated this G4s interactions with CX-5461. The major binding mode of CX-5461 identified in our study closely matches the binding pose of Phen-DC_3 solved using NMR spectroscopy (**Figure 13**). The highly flexible 5' terminal segment binds to and stabilizes both Phen-DC_3 and CX-5461 into an intercalation to the top G-quartet where the 5' terminal residue is bound to the center aromatic core of both ligands. This comparison provides support that the key interactions for this G4 are π-stacking with the guanine bases of the top G-tetrad, as well as validation for the simulation methods used in our study. Our study also supports the prediction made by Xu and coworkers[4] which suggests that CX-5461 may also be an end stacking G4 stabilizer due the structural relationship between CX-5461 and QQ58 (**Figure S33**), which was biophysically determined to be an end stacking G4 stabilizer in one study.

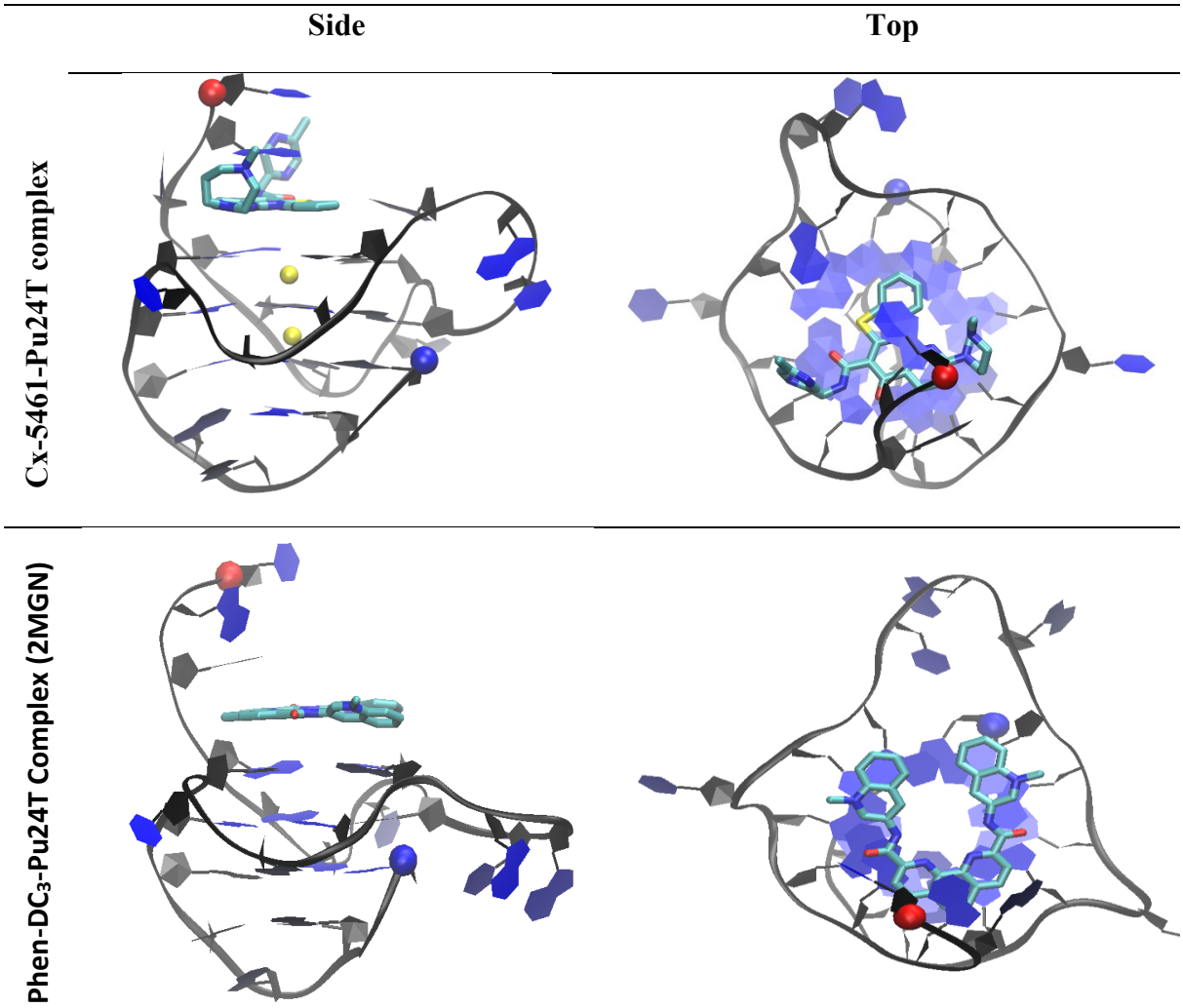


Figure 13. Major binding mode comparison of the CX-5461-Pu24T complex and Phen-DC_3-Pu24T Complex (PDB ID: 2MGN). 5' and 3' of the DNA chain are indicated by a red and blue sphere, respectively. K^+ cations are represented by yellow spheres.

As mentioned in the introduction, Xu and co-workers FRET-melting temperature increase assay [4] was performed to determine the affinity of CX-5461 binding to the G4s and duplex structure. A higher binding affinity was correlated to a higher melting temperature increase which indicated more energy was required to break the bonds formed upon CX-5461 binding and was essentially a measurement of enthalpy. The results of their assay indicated that with 10 μM CX-5461 the highest melting temperature increase, and thus the stabilizing effect due to CX-5461 binding, was demonstrated by the human telomeric system (~30 K) followed by the c-KIT1 G4 (~27 K), the c-Myc G4 (~25 K) and the DNA duplex (~10 K). Evident from this data, CX-5461 stabilizes the three G4 systems much more than the DNA duplex (~15-20 K). But what was causing marginally different melting temperature increases (~2-5 K) for the three G4 systems? In our study, each system's MM-PBSA binding energy calculations, which are also a measurement of binding enthalpy, were consistent with the order of stabilizing effect upon CX-5461 binding determined in the FRET-melting temperature assay (**Table 2**). CX-5461's average MM/PBSA over the major poses were calculated for the human telomeric (-28.6 kcal/mol), c-KIT1 (-23.9 kcal/mol), c-Myc (-22.0 kcal/mol) G4s, and DNA duplex (-15.0 kcal/mol) systems. Evident from this, the three G4 systems much higher MM-PBSA binding energies as compared to the DNA duplex.

Thus, it appears that the binding energy differences lead to the observed difference in the melting temperature change for these systems.

The unique structural differences at the 3' region provide insight into the binding modes and the binding energy differences. Due to the structural similarity on each face of the human telomeric G4, the binding poses and binding energies of the top (-37 kcal/mol) and bottom (-33 kcal/mol) modes were very comparable. However, this was not true for the c-KIT1 and c-Myc G4s, which contain 3 (GGA) and 4 (GAAG) residue long diagonal loops, respectively, on the 3' end. Instead, the bottom binding mode of c-Myc had a very unfavorable MMPBSA binding energy (-15.5 kcal/mol) which was comparable to the duplex system (~-15 kcal/mol), and the c-KIT1 system was not much more favorable (-22 kcal/mol). In these systems we observed a decreased ability for CX-5461 to interact directly with the third G-quartet due to hindrance from the diagonal loops which led to unequal binding opportunities when compared to the top binding mode. From our analysis we suggest that the more ordered diagonal loop of c-KIT1, which contains a A^{16}-G^{20} base pair, may contribute to the marginally higher stability determined from the FRET-melting temperature assay, when compared to the c-Myc G4 which contains the 4 residue diagonal loop (GAAG). Although CX-5461 was not able to intercalate onto the 3' G-quartet for the c-KIT1 and c-Myc G4's there was one case for both the c-KIT1 and c-Myc G4s where CX-4561 was able to intercalate into the 3' diagonal loop connecting the 3' G-tetrad, but our MMPBSA binding energy analysis revealed this binding mode was even less favorable in both the c-KIT1 (-11.66 kcal/mol) and c-Myc (-13.17 kcal/mol) systems. These facts help to explain why both of these G4s have a lower binding energy and melting temperature increase than the human telomeric G4, but also why the c-KIT1 G4 has a marginally higher affinity and binding energy when compared to the c-Myc G4. Thus, we propose that the structural differences at the 3' region of each G4 causes unequal CX-5461 binding opportunities which made for differences in our observed binding modes and may also play a critical role in the experimentally reported affinities from the melting temperature increase assay.

Along with characterizing the binding features of CX-5461's in each system to understand CX-5461s major role as a G4 binder and stabilizer observed in experiments, our study aimed to provide insight into the selectivity of CX-5461 to G4 structures over DNA duplex. The motivation for this is based on the major finding from the FRET-melting temperature assay performed by Xu and co-workers research showed CX-5461 selectively binds to and stabilizes G4 structures over duplex DNA[4]. This is an extremely desirable feature for a G4 stabilizer, since a major limiting factor for the therapeutic use of this class of drugs is the lack of selectivity to G4s over DNA duplex structures. Encouragingly, our trajectory analysis supports the selectivity of CX-5461 to G4s. The major finding used to conclude this was that for each of our G4 systems, an intercalation mode was observed for CX-5461, however for our duplex DNA system no groove intercalation mode was observed. The effects of this lack of highly stable binding pose in the duplex system was supported by MMPBSA binding energy analysis which showed that the duplex system had the lowest binding energy. The most energetically favorable binding pose observed in our duplex DNA system (-15.0 kcal/mol) was only comparable to the least favorable binding poses of the G4 systems, supporting that CX-5461 is selective to G4 structures over duplex DNA. In addition, our simulation analysis which supports that CX-5461 is not a DNA duplex intercalator

is consistent with an intercalation assay comparing CX-5461's ability to intercalate into calf thymus DNA to a known DNA intercalator Actinomycin-D[1]. This assay even showed that CX-5461 was even a weak minor groove binder at concentrations as high as 50 μM, which is consistent with our less favorable binding energy calculations for CX-5461 to the duplex when compared to the G4s.

To further compare the binding of CX-5461 to a known intercalator we compared the binding pose of CX-5461 to RHPS4, another G4 ligand with an aromatic core (**Figure S34**) from our previous work. [93] Evident from comparing the chemical structures, CX-5461 is longer and contains two flexible side chains whereas RHPS4 maintains a very planar aromatic structure. In the previous work, RHPS4 was modeled with the same duplex DNA and human telomeric G4 structure (1KF1) used in this work, among other G4s. RHPS4 was able to fully intercalate into the duplex DNA with a binding energy of -46.8±4.6 kcal/mol and in the human telomeric simulations the most favorable binding mode was bottom stacking with a MMPBSA binding energy of -48.9±2.4 kcal/mol. RHPS4's lack of selectivity toward a DNA structure indicated modifications were needed for it to be a successful G4 stabilizer. In this study, CX-5461 demonstrates an advantage in G4 selectivity over RHPS4 which we attribute to the two aromatic side chains of CX-5461. These two side chains limit CX-5461's planarity and therefore ability to fit into the narrow grooves of the DNA duplex structure. From our observations we believe that the positive charge on the methyl diazepane side chain, coupled with the flexibility of the methylpyrazine side chain, contributes to the selectivity of CX-5461 to G4 structures over duplex DNA.

Taking a deeper look at the overall binding pattern of CX-5461 to the DNA duplex, since it was clear CX-5461 was not a DNA duplex intercalator, we observed trajectories which showed an interesting mechanism comparable to a recent study.[94] It is conventional that a ligand may completely unbind and begin researching for a binding site in cases where ligand binding is not favorable. In our duplex system however, in cases when ligand binding was not favorable, we observed CX-5461 diffusing the surface of the DNA duplex searching for a stable binding site, maintaining a partial interaction throughout the entire binding process. This was made even clearer when the trajectory was extended from 500 ns to 2000 ns and provides unique insight into the binding interactions of CX-5461 to the duplex. We see that the drugs side chains (R1 and R2) are able to maintain interactions with the DNA, however a favorable binding pose is not achieved due to the lack of intercalation into the grooves of the duplex. Since the major factor limiting the therapeutic use of DNA G4 stabilizers to date has been a lack of selectivity to the DNA duplex, we kept this mechanism in mind when suggesting possible optimizations to this drug.

We developed a hypothesis to increase the binding of CX-5461 to the G4 structures and decrease the binding of CX-5461 to the duplex structure based on our analysis of the detailed interactions of each binding mode at the G4 and duplex structures. The detailed interactions (**Figures S11-S13**) of the most thermodynamically favorable binding poses (**Figures 3; Table 2**) indicate that CX-5461's rigid core (**Figure 1**) binding to the G4s is essential in stable binding. One level of support for this conclusion is that the core of CX-5461 is not exposed to solvent in our two-dimensional interaction diagrams of the most thermodynamically favorable top binding poses (**Figure S12**). Further, the second most thermodynamically favorable pose, human telomeric G4 bottom binding

(**Figure S11A**), also shows that the rigid core shows very little exposure to solvent. However, in the c-KIT1 and c-Myc bottom mode and all three G4 systems groove binding mode, the amount of solvent the rigid core of CX-5461 is exposed to is far greater (**Figure S11 A-D**). With these facts in mind we were able to suggest minor modifications be made to the benzothiazole-based scaffold of CX-5461 with the hypothesis that increasing the length of the core could increase the binding energy and stability. At each possible point of substitution, we picked 2 function groups (fluorine or chlorine) that could be substituted for hydrogen (**Figure 14; Figure S35**) leaving a total 56 compounds generated in our combinatorial library. Each new compound was docked to the same orthosteric binding site as the most abundant cluster for the most thermodynamically favorable top binding mode, and showed similar hydrogen bonds, π-π and hydrophobic interactions. Of these compounds we chose four derivates (**Figure 14**) based on their synthesizability determined by a minimal number of functional groups added, and their more negative XP docking scores for the G4 but less negative XP docking scores for the Duplex DNA.

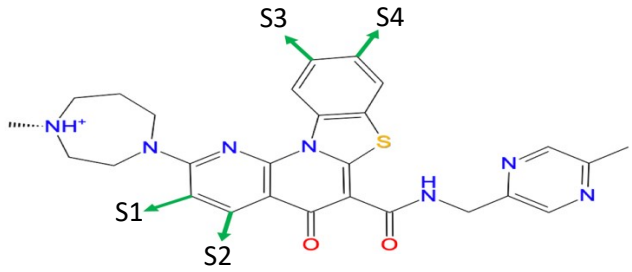
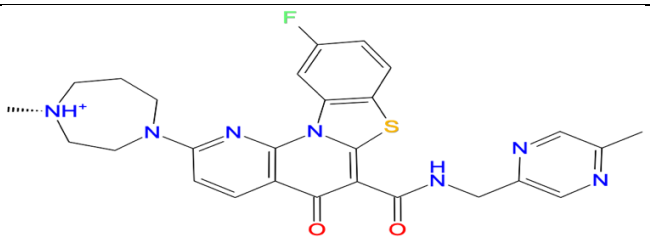
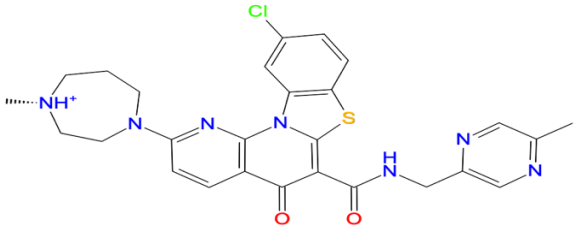
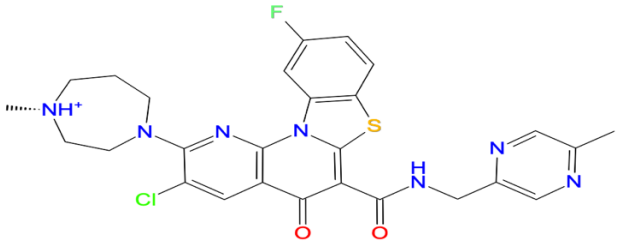
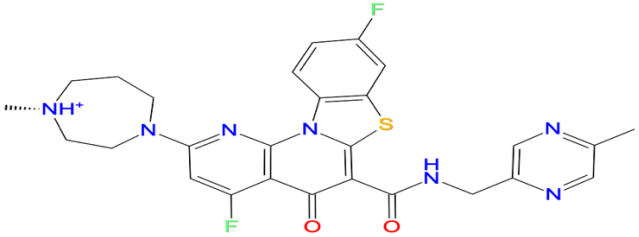
Chemical Structure of CX-5461: Substitution Sites Marked	Docking Score	
	h-Tel: -8.386 CKIT1: -8.112 c-Myc: -8.567 Duplex: -5.944	
Chemical Structure of New Ligand	**Docking Score**	**△CX-5461**
	h-Tel: -8.930 CKIT1: -8.552 c-Myc: -9.165 Duplex: -4.241	h-Tel: -0.544 CKIT1: -0.440 c-Myc: -0.598 Duplex: +2.703
	h-Tel: -9.272 CKIT1: -8.771 c-Myc: -8.818 Duplex: -3.326	h-Tel: -0.886 CKIT1: -0.659 c-Myc: -0.251 Duplex: +2.618
	h-Tel: -8.913 CKIT1: -8.179 c-Myc: -8.808 Duplex: -4.334	h-Tel: -0.527 CKIT1: -0.067 c-Myc: -0.241 Duplex: +1.610
	h-tel: -9.179 c-KIT1: -8.175 c-Myc: -8.667 Duplex: -5.273	h-Tel: -0.793 CKIT1: -0.063 c-Myc: -0.100 Duplex: +0.671

Figure 14. Chemical structure of CX-5461 and derivates identified through virtual screening, including docking scores. For CX-5461 green arrows indicate substitution sites for the derivates on this table. For the derivates, docking scores are provided as well as the difference between their docking scores and CX-5461.

While picking possible substitution sites, we closely compared our work to the original structure activity relationship study for which CX-5461 was identified[1]. Of note, their work found that

the addition of a F at our position S3 significantly increased the cellular activity of the compound, to a level even higher than reported for CX-5461. However, this addition was not made to CX-5461 itself. Rather than using the methylpyrazine based flexible side chain R2 (**Figure 1**) for their trial at this site they used a pyrrolidinoethylamine side chain. They discovered that this combination, although highly beneficial in terms of cellular activity, had an extremely low oral absorption. They later discovered that the addition of the methylpyrazine based side chain (R2) drastically increased the oral availability (~40 fold increase: ~2,300 to ~92,000 ng h/mL), however failed to revisit the possibility of using F at substitution site S3 with the addition of R2 to the aromatic core. Which led us to believe that because R2 provides such a drastic increase in oral availability, and the F at site S3 leads to a drastic increase in cellular activity, this combination could provide the resolve for both important factors. Our docking results further supported the promise of this combination where F at site S3 increased the binding to all G4's and provided the largest decrease in duplex binding (+2.703 kcal/mol) of all our derivatives. A closely following second was a Cl substitution at site S3 which also increased G4 binding and decreased duplex binding (+2.618 kcal/mol). A third ligand was identified which substituted F at site S3 and Cl at site S1 which increased G4 binding and decrease duplex binding (+1.610 kcal/mol). In addition to these ligands, we identified a number of other derivates which could be promising. This includes a ligand with F substitutions at sites S2 and S4 which increased binding for each the human telomeric, c-KIT1 and c-Myc G4s and decreased duplex binding (**Figure 14**) and a ligand which had a Cl substitution at site S2 which increased binding for each the human telomeric and c-Myc G4s but decreased binding for both the c-KIT1 G4 and duplex (**Figure S35**). All together suggesting substitutions at site S3 may address both the necessary biological properties as well as well as the intended decrease in duplex binding.

Furthermore, an interesting trend was observed for the human telomeric system. First, of the 56 ligands the intended effect (i.e. increase G4 binding or decrease duplex binding) was most observed for the human telomeric G4 system (48%) followed by the duplex (44%). Looking closer at the docking results, there were many new derivatives which produced the intended effect in just the human telomeric and duplex systems, six of which are exemplified in the supporting document (**Figure S35**). Putting together the most thermodynamically favorable system in our study was the human telomeric G4, the stability results of the melting temperature assay, and the results of our docking study, we suggest that using CX-5461 derivatives to specifically target the human telomeric G4 could be a promising therapeutic approach. This hypothesis is consistent with limited experimental testing such as a telomere FISH assay[4] which showed an increased frequency of telomere defects in BRCA -/- HCT116 cells after exposure to CX-5461 which they used to provide support of CX-5461's G4 stabilizing and ability to induce genome instability specifically at G4s in human cells. However, only targeting human telomeric G4 provides a limited solution that ties back into a major limiting factor of G4 stabilizers to date, whereby, the structure of the G4s overall do not provide a large binding pocket, as seen in proteins and other targets. This has implications on the reported efficacies where a G4 stabilizer has yet to bind in the nM concentration range. One approach to remedy this is to generate a hybrid G4 binder that has both the pharmacophores of a G4 intercalator and groove binder. The intention of this ligand design is that the intercalator core of this ligand will intercalate onto

a G-quartet and, that when connected by a flexible linker, the groove binder like side chains may wrap around the G4 core like arms and further stabilize the interactions. As far as we know, there has been no such ligand developed with this binding pose to date. Of course, further experimentation is required to support this hypothesis. Furthermore, given the promise of CX-5461 and its derivates to the parallel scaffold, a future study should include the binding of CX-5461 to multiple telomeric scaffolds (i.e. parallel, anti-parallel, and hybrid), because the insight gained from studying CX-5461 binding to each of the interconverting topologies may help to further optimize its structure.

The AMBER GAFF2 force field of CX-5461 in Mol2 format is also available in the supporting document (**Figure S36**).

CONCLUSION

Small molecule CX-5461 is a DNA G-quadruplex (G4) stabilizer currently in Phase I/II clinical trials for advanced hematologic malignancies and cancers with BRCA1/2 deficiencies. Although biologically promising, the detailed binding interactions of CX-5461 to the G4s, remains elusive due to a lack of an experimentally solved structure of any CX-5461-G4 complex. A FRET melting temperature increase assay measured CX-5461's binding affinity to the canonical DNA duplex structure and three DNA G-quadruplex (G4) forming sequences that are implicated in the cancerous complications resulting from BRCA1/2 mutations (human telomeric, c-KIT1, and c-Myc). The results of the FRET melting temperature increase assay indicated that with 10 μM CX-5461 the highest melting temperature increase, and therefore highest affinity, was demonstrated by the human telomeric system (~30 K) followed by the c-KIT1 (~27 K) and c-Myc (~25 K) G4s and the DNA duplex (~10 K). Extending upon these experimental findings, this study used free ligand molecular dynamics binding simulations of CX-5461 binding to the three DNA G-quadruplex forming sequences: human telomeric, c-KIT1 and c-Myc, and a DNA duplex. Through our detailed analysis of the 18+ μs simulation time per system, we probed the thermodynamic, kinetic and structural nature of each complex at the molecular level. Our MM/PBSA binding energy and Markovian State Mode (MSM) analyses revealed CX-5461 bound most favorably to the human telomeric (-28.6 kcal/mol), followed by the c-KIT1 (-23.9 kcal/mol), c-Myc (-22.0 kcal/mol) G4s, and then the DNA duplex (-15.0 kcal/mol), which explains the order of stabilizing effect due to CX-5461 binding from the melting temperature assay. Our thermodynamic and kinetic analysis suggested that the top binding mode could be the major pose responsible for the biological activity of CX-5461; and that the structural differences in the 3' sites of each G4 caused differences in the bottom binding modes which helped to explain the differences observed for the order of binding stabilization for the G4 systems. We also concluded that it was CX-5461's lack of intercalation to the duplex that explained the differences in melting temperature between the three G4s and duplex. Further characterizing specific binding features, we observed each G4 system undergoing a base flipping insertion induced fit binding mechanisms to achieve the most thermodynamically favorable top binding pose and evidence of a conformation-selection mechanism for both the human telomeric and c-KIT1 complex system. Further, our dynamic insight on the binding of CX-5461 to each of these structures revealed the rigid core of CX-5461 binding to the G4 core was essential for stable binding. Using this fact, halogen substitutions were made on CX-

5461's rigid core and docking into each G4 and duplex structure where we identified a number of derivates that increased G4 binding and decreased duplex binding, further enhancing the selectivity of CX-5461 to the G4s. Thus, this study provides insight that may aid in the further rational design of novel G-quadruplex stabilizers with increased selectivity to G-quadruplex structures over DNA duplexes.

ASSOCIATED CONTENT

This includes: supporting methods, RMSD and contact number plots for each trajectory, the last snapshots of quadruplex and duplex simulations, representative structures of the most populated structural families, detailed two dimensional interaction diagrams, hydrogen bond networks, additional trajectories and order parameter plots of each system, supporting figures for our kinetic analysis, the AMBER GAFF2 force field of the ligand, and a movie for the representative trajectory of each system are also provided.

FUNDING

This work was supported by the National Science Foundation under Grants RUI-1904797/MRI-1429467 and XSEDE MCB170088.

CONFLICT OF INTEREST

The authors declare no conflicts of interest.

REFERENCES

1. Haddach, M.; Schwaebe, M. K.; Michaux, J.; Nagasawa, J.; O'Brien, S. E.; Whitten, J. P.; Pierre, F.; Kerdoncuff, P.; Darjania, L.; Stansfield, R.; Drygin, D.; Anderes, K.; Proffitt, C.; Bliesath, J.; Siddiqui-Jain, A.; Omori, M.; Huser, N.; Rice, W. G.; Ryckman, D. M., Discovery of CX-5461, the First Direct and Selective Inhibitor of RNA Polymerase I, for Cancer Therapeutics. *ACS Med. Chem. Lett.* **2012**, 3, 602-606.
2. Andrews, W. J.; Panova, T.; Normand, C.; Gadal, O.; Tikhonova, I. G.; Panov, K. I., Old Drug, New Target: Ellipticines Selectively Inhibit RNA Polymerase I Transcription. *J. Biol. Chem.* **2013**, 288, 4567-4582.
3. Ye, F. B.; Hamza, A.; Singh, T.; Flibotte, S.; Hieter, P.; O'Neil, N. J., Characterization of the Phototoxicity, Chemigenetic Profile, and Mutational Signatures of the Chemotherapeutic CX-5461 in Caenorhabditis Elegans. *bioRxiv* **2019**, 2019.12.20.884981.
4. Xu, H.; Di Antonio, M.; McKinney, S.; Mathew, V.; Ho, B.; O'Neil, N. J.; Dos Santos, N.; Silvester, J.; Wei, V.; Garcia, J.; Kabeer, F.; Lai, D.; Soriano, P.; Banath, J.; Chiu, D. S.; Yap, D.; Le, D. D.; Ye, F. B.; Zhang, A. N.; Thu, K.; Soong, J.; Lin, S. C.; Tsai, A. H. C.; Osako, T.; Algara, T.; Saunders, D. N.;

Wong, J.; Xian, J.; Bally, M. B.; Brenton, J. D.; Brown, G. W.; Shah, S. P.; Cescon, D.; Mak, T. W.; Caldas, C.; Stirling, P. C.; Hieter, P.; Balasubramanian, S.; Aparicio, S., CX-5461 is a DNA G-Quadruplex Stabilizer with Selective Lethality in BRCA1/2 Deficient Tumours. *Nat. Commun.* **2017**, 8, 14432
5. Hilton, J.; Cescon, D. W.; Bedard, P.; Ritter, H.; Tu, D.; Soong, J.; Gelmon, K.; Aparicio, S.; Seymour, L., CCTG IND. 231: A Phase 1 Trial Evaluating CX-5461 in Patients with Advanced Solid Tumors. *Annals of Oncology* **2018**, 29, III8.
6. Wang, T. T.; Shatara, M.; Liu, F. B.; Knight, T.; Edwards, H.; Wang, G.; Lin, H.; Wang, Y.; Taub, J. W.; Ge, Y. B., Simultaneous Cotargeting of ATR and RNA Polymerase I Transcription Demonstrates Synergistic Antileukemic Effects on Acute Myeloid Leukemia. *Signal Transduction Targeted Ther.* **2019**, 4, 44.
7. Leung, A. W. Y.; Anantha, M.; Prosser, K. E.; Wehbe, M.; Walsby, C. J.; Bally, M. B., A Novel Formulation of CX-5461, a Small-Molecule Inhibitor of rRNA Synthesis, and Its Use for Treatment of Acute Myeloid Leukemia Models. *Cancer Res.* **2016**, 76.
8. Lee, H.; Wang, H.; Lin, H.; Baladandayuthapani, V.; He, J.; Jones, R. J.; Kuiatse, I.; Gu, D. M.; Wang, Z. Q.; O'Brien, S.; Keats, J. J.; Yang, J.; Davis, R. E.; Orlowski, R. Z., RNA Polymerase I Inhibition with CX-5461 As a Novel Strategy to Target MYC in Multiple Myeloma. *Blood* **2015**, 126.
9. Maclachlan, K. H.; Cuddihy, A.; Hein, N.; Cullinane, C.; Harrison, S. J.; Hannan, R.; Poortinga, G., Novel Combination Therapies with the RNA Polymerase I Inhibitor CX-5461 Significantly Improve Efficacy in Multiple Myeloma. *Blood* **2017**, 130.
10. Taylor, J. S.; Zeki, J.; Ornell, K.; Coburn, J.; Shimada, H.; Ikegaki, N.; Chiu, B., Down-Regulation of MYCN Protein by CX-5461 Leads to Neuroblastoma Tumor Growth Suppression. *Journal of Pediatric Surgery* **2019**, 54, 1192-1197.
11. Rebello, R. J.; Kusnadi, E.; Cameron, D. P.; Pearson, H. B.; Lesmana, A.; Devlin, J. R.; Drygin, D.; Clark, A. K.; Porter, L.; Pedersen, J.; Sandhu, S.; Risbridger, G. P.; Pearson, R. B.; Hannan, R. D.; Furic, L., The Dual Inhibition of RNA Pol I Transcription and PIM Kinase as a New Therapeutic Approach to Treat Advanced Prostate Cancer. *Clin. Cancer Res.* **2016**, 22, 5539-5552.
12. Li, L. M.; Li, Y.; Zhao, J. S.; Fan, S. L.; Wang, L. G.; Li, X., CX-5461 Induces Autophagy and Inhibits Tumor Growth via Mammalian Target of Rapamycin-Related Signaling Pathways in Osteosarcoma. *OncoTargets Ther.* **2016**, 9, 5985-5997.
13. Negi, S. S.; Brown, P., rRNA Synthesis Inhibitor, CX-5461, Activates ATM/ATR Pathway in Acute Lymphoblastic Leukemia, Arrests Cells in G2 Phase and Induces Apoptosis. *Oncotarget* **2015**, 6, 18094-18104.
14. Negi, S. S.; Brown, P., Transient rRNA Synthesis Inhibition with CX-5461 is Sufficient to Elicit Growth Arrest and Cell Death in Acute Lymphoblastic Leukemia Cells. *Oncotarget* **2015**, 6, 34846-34858.
15. Cornelison, R.; Dobbin, Z. C.; Katre, A. A.; Jeong, D. H.; Zhang, Y. F.; Chen, D. Q.; Petrova, Y.; Llaneza, D. C.; Steg, A. D.; Parsons, L.; Schneider, D. A.; Landen, C. N., Targeting RNA-Polymerase I in Both Chemosensitive and Chemoresistant Populations in Epithelial Ovarian Cancer. *Clin. Cancer Res.* **2017**, 23, 6529-6540.
16. Cornelison, R.; Llaneza, D. C.; Petrova, Y.; Dobbin, Z. C.; Schneider, D. A.; Landen, C. N., Inhibition of Autophagy Potentiates Cytotoxicity of CX-5461 Treatment in Chemoresistant Epithelial Ovarian Cancer. *Cancer Res.* **2017**, 77.
17. Cornelison, R.; Dobbin, Z. C.; Katie, A. A.; Jeong, D. H.; Petrova, Y.; Llaneza, D. C.; Steg, A. D.; Zhang, Y.; Schneider, D. A.; Landen, C. N., Targeting RNA-polymerase I Using CX-5461 as a Mechanism for Treating Chemotherapy Resistant Epithelial Ovarian Cancer. *Cancer Res.* **2016**, 76.
18. Ye, Q.; Pang, S.; Zhang, W. J.; Guo, X. T.; Wang, J. L.; Zhang, Y. T.; Liu, Y.; Wu, X.; Jiang, F., Therapeutic Targeting of RNA Polymerase I With the Small-Molecule CX-5461 for Prevention of Arterial Injury-Induced Neointimal Hyperplasia. *Arterioscler., Thromb., Vasc. Biol.* **2017**, 37, 476-484.
19. Kostopoulou, O. N.; Wilhelmi, V.; Raiss, S.; Ananthaseshan, S.; Lindstrom, M. S.; Bartek, J.; Soderberg-Naucler, C., Human Cytomegalovirus and Herpes Simplex Type I Virus Can Engage RNA Polymerase I for Transcription of Immediate Early Genes. *Oncotarget* **2017**, 8, 96536-96552.

20. Westdorp, K. N.; Terhune, S. S., Impact of RNA polymerase I Inhibitor CX-5461 on Viral Kinase-Dependent and Independent Cytomegalovirus Replication. *Antiviral Res.* **2018**, 153, 33-38.
21. Kerry, L. E.; Pegg, E. E.; Cameron, D. P.; Budzak, J.; Poortinga, G.; Hannan, K. M.; Hannan, R. D.; Rudenko, G., Selective Inhibition of RNA Polymerase I Transcription as a Potential Approach to Treat African Trypanosomiasis. *PLoS Neglected Trop. Dis.* **2017**, 11.
22. Neidle, S., Quadruplex Nucleic Acids as Novel Therapeutic Targets. *J. Med. Chem.* **2016**, 59, 5987-6011.
23. Cosconati, S.; Rizzo, A.; Trotta, R.; Pagano, B.; Iachettini, S.; De Tito, S.; Lauri, I.; Fotticchia, I.; Giustiniano, M.; Marinelli, L.; Giancola, C.; Novellino, E.; Biroccio, A.; Randazzo, A., Shooting for Selective Druglike G-quadruplex Binders: Evidence for Telomeric DNA Damage and Tumor Cell Death. *J. Med. Chem.* **2012**, 55, 9785-92.
24. Uziel, O.; Yerushalmi, R.; Zuriano, L.; Naser, S.; Beery, E.; Nordenberg, J.; Lubin, I.; Adel, Y.; Shepshelovich, D.; Yavin, H.; Ben Aharon, I.; Pery, S.; Rizel, S.; Pasmanik-Chor, M.; Frumkin, D.; Lahav, M., BRCA1/2 Mutations Perturb Telomere Biology: Characterization of Structural and Functional Abnormalities In Vitro and In Vivo. *Oncotarget* **2016**, 7, 2433-2454.
25. Rosen, E. M., BRCA1 in the DNA Damage Response and at Telomeres. *Front. Genet.* **2013**, 4, 85-85.
26. Sekido, Y.; Ohigashi, S.; Takahashi, T.; Hayashi, N.; Suzuki, K.; Hirota, S., Familial Gastrointestinal Stromal Tumor with Germline KIT Mutations Accompanying Hereditary Breast and Ovarian Cancer Syndrome. *Anticancer Res.* **2017**, 37, 1425-1431.
27. Waisbren, J.; Uthe, R.; Siziopikou, K.; Kaklamani, V., BRCA 1/2 Gene Mutation and Gastrointestinal Stromal Tumours: A Potential Association. *BMJ Case Rep* **2015**, 2015, bcr2014208830.
28. Grushko, T. A.; Dignam, J. J.; Das, S.; Blackwood, A. M.; Perou, C. M.; Ridderstråle, K. K.; Anderson, K. N.; Wei, M.-J.; Adams, A. J.; Hagos, F. G.; Sveen, L.; Lynch, H. T.; Weber, B. L.; Olopade, O. I., MYC Is Amplified in BRCA1-Associated Breast Cancers. *Clin. Cancer Res.* **2004**, 10, 499.
29. Chen, Y.; Xu, J.; Borowicz, S.; Collins, C.; Huo, D.; Olopade, O. I., c-Myc Activates BRCA1 Gene Expression Through Distal Promoter Elements in Breast Cancer Cells. *BMC Cancer* **2011**, 11, 246-246.
30. Zhang, X.; Xu, C. X.; Di Felice, R.; Sponer, J.; Islam, B.; Stadlbauer, P.; Ding, Y.; Mao, L.; Mao, Z. W.; Qin, P. Z., Conformations of Human Telomeric G-Quadruplex Studied Using a Nucleotide-Independent Nitroxide Label. *Biochemistry* **2016**, 55, 360-72.
31. Collie, G. W.; Campbell, N. H.; Neidle, S., Loop Flexibility in Human Telomeric Quadruplex Small-Molecule Complexes. *Nucleic Acids Res.* **2015**, 43, 4785-99.
32. Tawani, A.; Mishra, S. K.; Kumar, A., Structural Insight for the Recognition of G-Quadruplex Structure at Human c-Myc Promoter Sequence by Flavonoid Quercetin. *Sci. Rep.* **2017**, 7, 3600.
33. Machireddy, B.; Sullivan, H.-J.; Wu, C., Binding of BRACO19 to a Telomeric G-Quadruplex DNA Probed by All-Atom Molecular Dynamics Simulations with Explicit Solvent. *Molecules* **2019**, 24, 1010.
34. Shankar, U.; Jain, N.; Mishra, S. K.; Sharma, T. K.; Kumar, A., Conserved G-Quadruplex Motifs in Gene Promoter Region Reveals a Novel Therapeutic Approach to Target Multi-Drug Resistance Klebsiella pneumoniae. *Frontiers in Microbiology* **2020**, 11.
35. Galindo-Murillo, R.; Robertson, J. C.; Zgarbova, M.; Sponer, J.; Otyepka, M.; Jurecka, P.; Cheatham, T. E., Assessing the Current State of Amber Force Field Modifications for DNA. *J. Chem. Theory Comput.* **2016**, 12, 4114-4127.
36. Hou, J. Q.; Chen, S. B.; Tan, J. H.; Luo, H. B.; Li, D.; Gu, L. Q.; Huang, Z. S., New Insights From Molecular Dynamic Simulation Studies of the Multiple Binding Modes of a Ligand with G-quadruplex DNA. *J. Comput.-Aided Mol. Des.* **2012**, 26, 1355-68.
37. Xuan-Yu M., H.-X. Z., Mihaly M., and Meng C., Molecular Docking: A Powerful Approach for Structure-Based Drug Discovery. *Curr. Comput.-Aided Drug Des.* **2011**, 7, 12.
38. 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.
39. Di Leva, F. S.; Novellino, E.; Cavalli, A.; Parrinello, M.; Limongelli, V., Mechanistic Insight Into Ligand Binding to G-quadruplex DNA. *Nucleic Acids Res.* **2014**, 42, 5447-55.
40. Federica Moraca, J. A., Francesco Ortuso, Anna Artese, Bruno Pagano, Ettore Novellino,; Stefano Alcaro, M. P. a. V. L., Ligand Binding to Telomeric G-Quadruplex DNA Investigated by Funnel-Metadynamics Simulations. *Proc. Natl. Acad. Sci. U. S. A.* **2017**.
41. 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, 227-42.
42. Hong Z., S. X., Haojun L., Structural Dynamics of Human Telomeric G-Quadruplex Loops Studied by Molecular Dynamics Simulations. *PloS one* **2013**, 8, e71380.
43. Catalano, R.; Moraca, F.; Amato, J.; Cristofari, C.; Rigo, R.; Via, L. D.; Rocca, R.; Lupia, A.; Maruca, A.; Costa, G.; Catalanotti, B.; Artese, A.; Pagano, B.; Randazzo, A.; Sissi, C.; Novellino, E.; Alcaro, S., Targeting Multiple G-quadruplex-Forming DNA Sequences: Design, Biophysical and Biological Evaluations of Indolo-Naphthyridine Scaffold Derivatives. *Eur. J. Med. Chem.* **2019**, 182.
44. Haider, S., Computational Methods to Study G-Quadruplex-Ligand Complexes. *J. Indian Inst. Sci.* **2018**, 98, 325-339.
45. Ivani, I.; Dans, P. D.; Noy, A.; Perez, A.; Faustino, I.; Hospital, A.; Walther, J.; Andrio, P.; Goni, R.; Balaceanu, A.; Portella, G.; Battistini, F.; Gelpi, J. L.; Gonzalez, C.; Vendruscolo, M.; Laughton, C. A.; Harris, S. A.; Case, D. A.; Orozco, M., Parmbsc1: a Refined Force Field for DNA Simulations. *Nature Methods* **2016**, 13, 55-+.
46. Jamroskovic, J.; Doimo, M.; Chand, K.; Obi, I.; Kumar, R.; Brannstrom, K.; Hedenstrom, M.; Das, R. N.; Akhunzianov, A.; Deiana, M.; Kasho, K.; Sato, S. S.; Pourbozorgi, P. L.; Mason, J. E.; Medini, P.; Ohlund, D.; Wanrooij, S.; Chorell, E.; Sabouri, N., Quinazoline Ligands Induce Cancer Cell Death through Selective STAT3 Inhibition and G-Quadruplex Stabilization. *J. Am. Chem. Soc.* **2020**, 142, 2876-2888.
47. Prasad, B.; Das, R. N.; Jamroskovic, J.; Kumar, R.; Hedenstrom, M.; Sabouri, N.; Chorell, E., The Relation Between Position and Chemical Composition of Bis-Indole Substituents Determines Their Interactions with G-Quadruplex DNA. *Chem. - Eur. J.*
48. Lemkul, J. A.; MacKerell, A. D., Polarizable force field for RNA based on the Classical Drude Oscillator. *J. Comput. Chem.* **2018**, 39, 2624-2646.
49. Rebic, M.; Laaksonen, A.; Sponer, J.; Ulicny, J.; Mocci, F., Molecular Dynamics Simulation Study of Parallel Telomeric DNA Quadruplexes at Different Ionic Strengths: Evaluation of Water and Ion Models. *J. Phys. Chem. B* **2016**, 120, 7380-91.
50. 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.
51. Deng, N.; Xia, J.; Wickstrom, L.; Lin, C.; Wang, K.; He, P.; Yin, Y.; Yang, D., Ligand Selectivity in the Recognition of Protoberberine Alkaloids by Hybrid-2 Human Telomeric G-Quadruplex: Binding Free Energy Calculation, Fluorescence Binding, and NMR Experiments. *Molecules* **2019**, 24, 1574.
52. 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.
53. Bhat, J.; Mondal, S.; Sengupta, P.; Chatterjee, S., In Silico Screening and Binding Characterization of Small Molecules Toward a G-Quadruplex Structure Formed in the Promoter Region of c-MYC Oncogene. *ACS Omega* **2017**, 2, 4382-4397.
54. Salsbury, A. M.; Lemkul, J. A., Molecular Dynamics Simulations of the c-kit1 Promoter G-Quadruplex: Importance of Electronic Polarization on Stability and Cooperative Ion Binding. *J. Phys. Chem. B* **2019**, 123, 148-159.

55. Salsbury, A. M.; Lemkul, J. A., Influence of Monovalent Cations on the Dynamics of the C-KIT1 Promoter G-Quadruplex using Polarizable Molecular Dynamics Simulations. *Biophys. J.* **2020**, 118, 220a.
56. Dean, T.; Salsbury, A. M.; Lemkul, J. A., Dynamics of the 1:2:1 and 1:6:1 C-Myc G-Quadruplexes with the Drude Polarizable Force Field. *Biophys. J.* **2020**, 118, 65a.
57. Pany, S. P. P.; Bommisetti, P.; Diveshkumar, K. V.; Pradeepkumar, P. I., Benzothiazole Hydrazones of Furylbenzamides Preferentially Stabilize c-MYC and c-KIT1 Promoter G-Quadruplex DNAs. *Org. Biomol. Chem.* **2016**, 14, 5779-5793.
58. Read M., W. A., Harrison J., Gowan S., Kelland L., Dosanjh H., and Neidle S., Molecular Modeling Studies on G-Quadruplex Complexes of Telomerase Inhibitors: Structure−Activity Relationships. *J. Med. Chem.* **1999**, 42, 4538-4546.
59. Mulholland, K.; Wu, C., Binding of Telomestatin to a Telomeric G-Quadruplex DNA Probed by All-Atom Molecular Dynamics Simulations with Explicit Solvent. *J. Chem. Inf. Model.* **2016**, 56, 2093-2102.
60. Wang, Z.; Chen, R.; Hou, L.; Li, J.; Liu, J.-P., Molecular Dynamics and Principal Components of Potassium Binding with Human Telomeric Intra-Molecular G-Quadruplex. *Protein Cell* **2015**, 6, 423-433.
61. Pérez, A.; Marchán, I.; Svozil, D.; Sponer, J.; Cheatham, T. E. r.; Laughton, C. A.; Orozco, M., Refinement of the AMBER Force Field for Nucleic Acids: Improving the Description of α/γ Conformers. *Biophys. J.* **2007**, 92, 3817–3829.
62. Krepl, M.; Zgarbova, M.; Stadlbauer, P.; Otyepka, M.; Banas, P.; Koca, J.; Cheatham, T. E.; 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. Chem. Theory Comput.* **2012**, 8, 2506-2520.
63. Zgarbova, M.; Luque, F. J.; Sponer, J.; Cheatham, T. E.; Otyepka, M.; Jurecka, P., Toward Improved Description of DNA Backbone: Revisiting Epsilon and Zeta Torsion Force Field Parameters. *J. Chem. Theory Comput.* **2013**, 9, 2339-2354.
64. Zgarbova, M.; Sponer, J.; Otyepka, M.; Cheatham, T. E.; 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. Chem. Theory Comput.* **2015**, 11, 5723-5736.
65. 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.
66. Case, D. A.; Betz, R. M.; Botello-Smith, W.; Cerutti, D. S.; Cheatham, I., T.E. ; Darden, T. A.; Duke, R. E.; Giese, T. J.; Gohlke, H.; Goetz, A. W.; Homeyer, N.; Izadi, S.; Janowski, P.; Kaus, J.; Kovalenko, A.; Lee, T. S.; LeGrand, S.; Li, P.; Lin, C.; Luchko, T.; Luo, R.; Madej, B. *AMBER 2016*, University of California, San Francisco, 2016.
67. Bayly, C. I.; Cieplak, P.; Cornell, W. D.; Kollman, P. A., A Well-Behaved Electrostatic Potential Based Method Using Charge Restraints for Deriving Atomic Charges - the Resp Model. *J. Phys. Chem.* **1993**, 97, 10269-10280.
68. Lavery, R.; Zakrzewska, K.; Beveridge, D.; Bishop, T. C.; Case, D. A.; Cheatham, T. E. I.; Dixit, S.; Jayaram, B.; Lankas, F.; Laughton, C.; Maddocks, J. H.; Michon, A.; Osman, R.; Orozco, M.; Perez, A.; Singh, T.; Spackova, N.; Sponer, J., A Systematic Molecular Dynamics Study of Nearest-Neighbor Effects on Base Pair and Base Pair Step Conformations and Fluctuations in B-DNA *Nucleic Acids Res.* **2010**, 38, 299-313.
69. Cosconati, S.; Marinelli, L.; Trotta, R.; Virno, A.; De Tito, S.; Romagnoli, R.; Pagano, B.; Limongelli, V.; Giancola, C.; Baraldi, P. G.; Mayol, L.; Novellino, E.; Randazzo, A., Structural and Conformational Requisites in DNA Quadruplex Groove Binding: Another Piece to the Puzzle. *J. Am. Chem. Soc.* **2010**, 132, 6425-6433.
70. Fadrna, E.; Spackova, N. a.; Sarzynska, J.; Koca, J.; Orozco, M.; Cheatham, T. E., III;; 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

71. Mukherjee, A.; Lavery, R.; Bagchi, B.; Hynes, J. T., On the Molecular Mechanism of Drug Intercalation into DNA: A Simulation Study of the Intercalation Pathway, Free Energy, and DNA Structural Changes. *J. Am. Chem. Soc.* **2008**, 130, 9747-9755.
72. Lei, H.; Wang, X.; Wu, C., Early Stage Intercalation of Doxorubicin to DNA Fragments Observed in Molecular Dynamics Binding Simulations. *J. Mol. Graphics Modell.* **2012**, 38, 279-89.
73. Mulholland, K.; Wu, C., Binding of Telomestatin to a Telomeric G-Quadruplex DNA Probed by All-Atom Molecular Dynamics Simulations with Explicit Solvent. *Journal of Chemical Information and Modeling* **2016**, 56, 2093-2102.
74. Mulholland, K.; Sullivan, H.-J.; Garner, J.; Cai, J.; Chen, B.; Wu, C., Three-Dimensional Structure of RNA Monomeric G-Quadruplex Containing ALS and FTD Related G4C2 Repeat and Its Binding with TMPyP4 Probed by Homology Modeling based on Experimental Constraints and Molecular Dynamics Simulations. *ACS Chem. Neurosci.* **2020**, 11, 57-75.
75. Sullivan, H.-J.; Readmond, C.; Radicella, C.; Persad, V.; Fasano, T. J.; Wu, C., Binding of Telomestatin, TMPyP4, BSU6037, and BRACO19 to a Telomeric G-Quadruplex–Duplex Hybrid Probed by All-Atom Molecular Dynamics Simulations with Explicit Solvent. *ACS Omega* **2018**, 3, 14788-14806.
76. Ryckaert, J.; Ciccotti, G.; Berendsen, H. J. C., Numerical Integration of the Cartesian Equations of Motion of a System with Constraints: Molecular Dynamics of n-Alkanes. *J. Comput. Phys.* **1977**, 23, 327-341.
77. Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T. A.; Lee, H.; Pedersen, L. G., A Smooth Particle Mesh Ewald Method. *J. Comput. Phys.* **1995**, 103, 8577-8593.
78. Procacci, P.; Berne, B. J., Multiple Time-Scale Methods For Constant-Pressure Molecular-Dynamics Simulations Of Molecular-Systems. *Molecular Physics* **1994**, 83, 255-272.
79. Kollman, P. A.; Massova, I.; Reyes, C.; Kuhn, B.; Huo, S.; Chong, L.; Lee, M.; Lee, T.; Duan, Y.; Wang, W.; Donini, O.; Cieplak, P.; Srinivasan, J.; Case, D. A.; Cheatham, T. E. I., Calculating Structures and Free Energies of Complex Molecules: Combining Molecular Mechanics and Continuum Model. *Accounts of Chemical Research* **2000**, 33, 889-897.
80. Kollman, P.; Massova, I.; Reyes, C.; Kuhn, B.; Huo, S.; Chong, L.; Lee, M.; Lee, T.; Duan, Y.; Wang, W.; Donini, O.; Cieplak, P.; Srinivasan, J.; Case, D.; Cheatham, T., Calculating Structures and Free Energies of Complex Molecules: Combining Molecular Mechanics and Continuum Models. *Accounts of Chemical Research* **2000**, 33, 889-897.
81. Hou, T.; Wang, J.; Li, Y.; Wang, W., Assessing the Performance of the Molecular Mechanics/Poisson Boltzmann Surface Area and Molecular Mechanics/Generalized Born Surface Area Methods. II. The Accuracy of Ranking Poses Generated from Docking. *Journal of Computational Chemistry* **2010**, 32, 866-877.
82. Hou, T.; Wang, J.; Li, Y.; Wang, W., Assessing the Performance of the MM/PBSA and MM/GBSA Methods. 1. The Accuracy of Binding Free Energy Calculations Based on Molecular Dynamics Simulations. *Journal of Chemical Information and Modeling* **2011**, 51, 69-82.
83. Xu, L.; Sun, H.; Li, Y.; Wang, J.; Hou, T., Assessing the Performance of MM/PBSA and MM/GBSA Methods. 3. The Impact of Force Fields and Ligand Charge Models. *Journal of Physical Chemistry B* **2013**, 117, 8408-8421.
84. Sun, H.; Li, Y.; Tian, S.; Xub, L.; Hou, T., Assessing the Performance of MM/PBSA and MM/GBSA Methods. 4. Accuracies of MM/PBSA and MM/GBSA Methodologies Evaluated by Various Simulation Protocols Using PDBbind Data Set. *Physical Chemistry and Chemical Physics* **2014**, 16, 16719-16729.
85. Friesner, R. A.; Banks, J. L.; Murphy, R. B.; Halgren, T. A.; Klicic, J. J.; Mainz, D. T.; Repasky, M. P.; Knoll, E. H.; Shelley, M.; Perry, J. K.; Shaw, D. E.; Francis, P.; Shenkin, P. S., Glide: A New Approach for Rapid, Accurate Docking and Scoring. 1. Method and Assessment of Docking Accuracy. *Journal of Medicinal Chemistry* **2004**, 47, 1739-1749.

86. Sastry, G. M.; Adzhigirey, M.; Day, T.; Annabhimoju, R.; Sherman, W., Protein and Ligand Preparation: Parameters, Protocols, and Influence on Virtual Screening Enrichments. *Journal of Computer Aided Molecular Design* **2013**, 27, 221-234.
87. Ioakimidis, L.; Thoukydidis, L.; Mirza, A.; Naeem, S.; Reynisson, J., Benchmarking the Reliability of QikProp. Correlation Between Experimental and Predicted Values. *QSAR Combinatorial Sciences* **2008**, 27, 445-456.
88. Pande, V. S.; Beauchamp, K.; Bowman, G. R., Everything You Wanted to Know About Markov State Models but Were Afraid to Ask. *Methods* **2010**, 52, 99-105.
89. Jiang, H.; Sheong, F. K.; Zhu, L.; Gao, X.; Bernauer, J.; Huang, X., Markov State Models Reveal a Two-Step Mechanism of miRNA Loading into the Human Argonaute Protein: Selective Binding Followed by Structural Re-arrangement. *PLoS Comput. Biol.* **2015**, 11, e1004404.
90. Biffi, G.; Tannahill, D.; Miller, J.; Howat, W. J.; Balasubramanian, S., Elevated Levels of G-Quadruplex Formation in Human Stomach and Liver Cancer Tissues. *PloS one* **2014**, 9, e102711.
91. Neidle, S., Quadruplex Nucleic Acids as Novel Therapeutic Targets. *Journal of Medicinal Chemistry* **2016**, 59, 5987-6011.
92. Chung, W. J.; Heddi, B.; Hamon, F.; Teulade-Fichou, M. P.; Phan, A. T., Solution Structure of a G-quadruplex Bound to the Bisquinolinium Compound Phen-DC3. *Angewandte Chemie-International Edition* **2014**, 53, 999-1002.
93. 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.
94. Shan, Y. B.; Kim, E. T.; Eastwood, M. P.; Dror, R. O.; Seeliger, M. A.; Shaw, D. E., How Does a Drug Molecule Find Its Target Binding Site? *J. Am. Chem. Soc.* **2011**, 133, 9181-9183.

Table of Contents Graphic