Organic & Biomolecular Chemistry

PAPER

View Article Online
View Journal

Cite this: DOI: 10.1039/c6ob00138f

Benzothiazole hydrazones of furylbenzamides preferentially stabilize *c-MYC* and *c-KIT1* promoter G-quadruplex DNAs†

Sushree Prangya Priyadarshinee Pany, Praneeth Bommisetti, K. V. Diveshkumar and P. I. Pradeepkumar*

The stabilization of G-quadruplex DNA structures by using small molecule ligands having simple structural scaffolds has the potential to be harnessed for developing next generation anticancer agents. Because of the structural diversity of G-quadruplexes, it is challenging to design stabilizing ligands, which can specifically bind to a particular quadruplex topology. To address this, herein, we report the design and synthesis of three benzothiazole hydrazones of furylbenzamides having different side chains (ligands **1**, **2** and **3**), which show preferential stabilization of promoter quadruplex DNAs (*c-MYC* and *c-KIT1*) having parallel topologies over telomeric and duplex DNAs. The CD melting study revealed that all the ligands, in particular ligand **2**, exhibit higher stabilization toward parallel promoter quadruplexes (ΔT_m = 10–15 °C) as compared to antiparallel promoter quadruplex (*h-RAS1*), telomeric quadruplex and duplex DNAs (ΔT_m = 0–3 °C). FID assay and fluorimetric titration results also reveal the preferential binding of ligands toward *c-MYC* and *c-KIT1* promoter quadruplex DNAs over telomeric and duplex DNAs. Validating these results further, *Taq* DNA polymerase stop assay showed IC_{50} ~ 6.4 μM for ligand **2** with the *c-MYC* DNA template, whereas the same for the telomeric DNA template was found to be >200 μM. Molecular modeling and dynamics studies demonstrated a 1 : 1 binding stoichiometry in which stacking and electrostatic interactions play important roles in stabilizing the *c-MYC* G-quadruplex structure. Taken together, the results presented here provide new insights into the design of structurally simple scaffolds for the preferential stabilization of a particular G-quadruplex topology.

Received 16th January 2016,
Accepted 14th March 2016

DOI: 10.1039/c6ob00138f

www.rsc.org/obc

Introduction

G-quadruplexes are secondary DNA structures formed by the stacking of two or more planar G-quartets in the presence of suitable metal ions such as K^+ and Na^+.[1–3] G-quartets are formed by the H-bond association of adjacent guanine bases by utilizing both Hoogsteen and Watson–Crick faces.[2] Putative G-quadruplex forming sequences are highly prevalent in the human genome and they are mostly present at the telomere, gene promoters, introns and at the immunoglobulin switch regions.[4,5] Recently, the *in vivo* existence of G-quadruplex structures has been visualized by immunofluorescence staining using structure specific BG4 antibodies as well as fluorescent light up probes.[6–8] The G-quadruplex structures can adopt various topologies and their structural polymorphism depends on the strand orientation, size/sequence of the loops, glycosidic torsional angles and nature of the metal ions present in the medium.[9] For the human telomeric G-quadruplex DNA, various conformations such as antiparallel, (3 + 1) hybrid and parallel topologies are observed under different metal ion and molecular crowding conditions.[10] G-quadruplex structures formed at the promoter regions are well characterized to adopt mostly parallel topologies under K^+ conditions.[11,12]

G-quadruplex structures present in the telomeric and promoter regions were explored for their roles in telomerase inhibition and the regulation of gene expression, respectively, in the past few decades.[13–15] Therefore, the stabilization of these quadruplex structures by using small molecule ligands has emerged as an attractive area in the anticancer drug development.[16–19] Most of the reported ligands provide a planar aromatic surface for π-stacking interactions with the external surface of the G-quartet. Along with this, recognition was also provided by the electrostatic as well as H-bonding interactions of the neutral/cationic side chain with the loops/grooves and phosphate backbone of the quadruplex structure.[20] A wide range of G-quadruplex stabilizing ligands such as telomestatin,[21] diarylethynylamides,[22] naphthyridine

Department of Chemistry, Indian Institute of Technology Bombay, Powai, Mumbai-400076, India. E-mail: pradeep@chem.iitb.ac.in

†Electronic supplementary information (ESI) available. See DOI: 10.1039/c6ob00138f

derivatives,[23] DPC derivatives,[24] xanthene and xanthone derivatives[25] were reported for their selective stabilization toward G-quadruplex over duplex DNAs. Piperazinylquindoline derivatives,[26] ellipticine derivatives,[27] *N*-substituted berberine derivatives,[28] bisaryldiketene derivatives[29] and furan based cyclic homooligopeptides[30] were reported for the selective stabilization of the *c-MYC* quadruplex over duplex DNAs and the down regulation of *c-MYC* gene expression. Though these ligands selectively stabilize G-quadruplex over duplex DNA, most of them fail to achieve specificity or preference toward a particular G-quadruplex topology. Therefore, it is highly challenging to design ligands in a structure specific fashion to discriminate between different quadruplex topologies. Such ligands may be desirable to achieve a target specific therapeutic impact on malignant cells. The structure specific ligands should be designed by utilizing unique recognition elements of the loops and grooves of the quadruplex along with the recognition of varying dimensions of G-quartets.[19] There are only a handful of topology specific ligands reported in the literature. Our group has recently reported indenopyrimidine derivatives,[31] and bisbenzimidazole carboxamide derivatives,[32] which show specific stabilization of the parallel topology of a promoter over telomeric G-quadruplex and duplex DNAs. Pyridyl peptidomimetic ligands have been reported for their specific stabilization toward *c-KIT1* DNA over other quadruplex and duplex DNAs.[33] Recently, disubstituted benzofuran derivatives have been identified using small molecule micro array screens as specific inhibitors for *c-MYC* gene expression.[34]

In our continued efforts to develop topology specific G-quadruplex stabilizing ligands having simple structural scaffolds harbouring drug-like properties, herein, we report a new series of ligands based on the benzothiazole hydrazone moiety coupled with the furylbenzamide group anchoring different side chains (Fig. 1). The binding interaction, stability and selectivity of these ligands toward G-quadruplex (*c-MYC*, *c-KIT1*, *h-RAS1* and telomeric) and duplex DNAs have been explored by various biophysical and biochemical methods such as CD titration, CD melting, fluorescence intercalator displacement (FID) assay and *Taq* DNA polymerase stop assay. The results show that these ligands preferentially stabilize the parallel topology of *c-MYC* and *c-KIT1* promoter quadruplex DNAs over the antiparallel topology of *h-RAS1* promoter DNA, various topologies of telomeric quadruplex and duplex DNAs. The binding mode and interactions between the ligand and quadruplex DNA were explored by harnessing molecular modeling and dynamics studies.

Results and discussion

Ligand design and synthesis

Benzothiazole moieties are considered as one of the foremost building blocks in medicinal chemistry due to their prevailing biological relevance.[35] These entities show a very wide range of biological activities including cytotoxicity against tumorigenic cell lines.[36] Arylfuran groups are reported for their antiproliferative activity in human leukemia cells.[37,38] Recently, a series of compounds containing furyl-benzothiazole hydrazone derivatives are reported for their selective inhibition of *BCL-X*$_L$ protein.[39] These widely explored biological applications inspired us to utilize such a heteroaromatic core, which can be tuned into G-quadruplex stabilizing ligands. Owing to the structural similarity of benzothiazoles with the guanine base of G-quartets, they are expected to provide the stacking interaction with the G-quartets. Thus, we envisioned that the incorporation of suitable side chains such as benzamides into the furyl-benzothiazole hydrazone scaffold can enhance the recognition and stabilization properties of these ligands. To achieve this, we have designed three ligands having a common aromatic core with guanidinium, ethyl and propyl side chains (Fig. 1). These positively charged side chains of varying nature and lengths were exploited to discriminate between topologically different loops and grooves present in the quadruplex structures.

The synthetic route utilized to access all the three ligands is depicted in Scheme 1. The synthesis was embarked from compound **4**, which was prepared by following a previously reported procedure.[39] Surprisingly, the direct amide coupling of **4** with the side chains employing different coupling agents such as EDCI, HATU and BOP yielded products along with inseparable impurities. This forced us to employ a two-step procedure.[40] First, compound **4** was converted to 2,3,5,6-tetrafluorophenolic ester **5** in 70% yield. Further, the ester **5** was

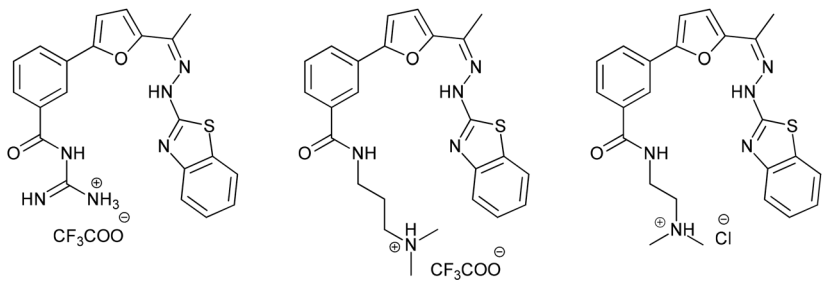
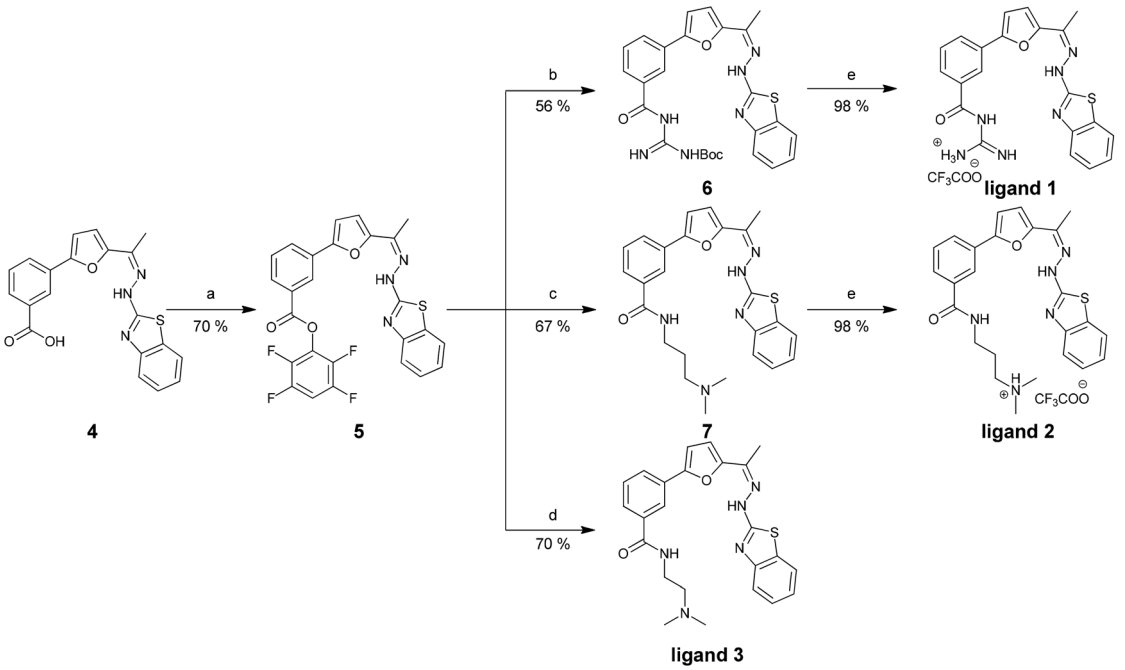


Fig. 1 Structure of G-quadruplex stabilizing ligands based on benzothiazole hydrazones of furylbenzamide scaffolds.



Scheme 1 Synthesis of ligands **1**, **2** and **3**. Reagents and conditions: (a) EDCI, 2,3,5,6-tetrafluorophenol, DMF, RT, 6 h; (b) *N*-Boc-guanidine, DCM : MeOH (1 : 1), RT, 3 h; (c) *N*,*N*-dimethylpropane-1,3-diamine, DCM, RT, 3 h; (d) *N*,*N*-dimethylethane-1,2-diamine, DCM, RT, 3 h; (e) trifluoroacetic acid, DCM, RT, 1 h.

converted to various amides **6**, **7** and **3** using the corresponding amines in 56%, 67% and 70% yields respectively. Finally, the amides **6** and **7** were treated with trifluoroacetic acid to furnish the protonated ligands **1** and **2**. However, we found that purification of the trifluoroacetate salt of **3** was cumbersome, and therefore a stock solution of the same in 10 mM HCl was prepared for various biophysical and biochemical studies.

Circular dichroism titration studies

Various topologies of G-quadruplex DNA structures can be qualitatively characterized by using circular dichroism (CD) spectroscopy.[41] CD titration spectra of telomeric DNA in the absence of any added metal ions showed two positive bands: one dominant at 254 nm and a small band around 290 nm, which are not attributed to any specific quadruplex topology. Upon titration of telomeric DNA with ligand **1**, the peak intensity at 290 nm was increased moderately, whereas with ligands **2** and **3**, a small increment was observed in a concentration dependant manner (Fig. 2A and S1, ESI†). Also, a new negative band at 260 nm with moderate intensity gradually started appearing along with the shifting of the peak from 254 nm to 244 nm (Fig. 2A and S1, ESI†). These moderate enhancements of ellipticities at 290, 260 and 244 nm indicate the weak induction of the antiparallel topology for the telomeric quadruplex DNA by the ligands.

Unlike telomeric quadruplex DNA, most of the promoter G-quadruplex DNAs are reported to exhibit a parallel topology even in the absence of any added monovalent metal ions.[22] The CD spectra for both *c-MYC* and *c-KIT1* DNAs without any added monovalent metal ions were found to have a positive peak at 260 nm and a negative peak at 240 nm, which reveals the pre-formed parallel topologies of these quadruplex DNAs (Fig. 2B, S1 and S2, ESI†). For the *c-MYC* DNA, the peak ellipticity at both 260 and 240 nm was found to be linearly increasing with the ligand concentration and attains saturation after the addition of 4 equivalents of ligands (Fig. 2B and S1, ESI†). A dramatic increase in the intensity of the characteristic peaks signifies further stabilization of the pre-folded parallel G-quadruplex topology for the *c-MYC* quadruplex DNA by all the ligands. Similarly, titration studies were performed with *c-KIT1* quadruplex DNAs and moderate stabilization of the existing parallel topology was observed with all the three ligands (Fig. S2, ESI†).

Moderate induction of the antiparallel topology for the telomeric DNA by the ligands prompted us to investigate whether the ligands are able to stabilize the antiparallel topology of *h-RAS1* promoter DNA reported under K^+ conditions.[42] The molar ellipticity of *h-RAS1* DNA without any added monovalent metal ion showed a positive peak at 280 nm and a negative peak at 246 nm, which does not account for any particular quadruplex topology. The addition of up to 5 equivalents of ligand **2** could not induce any particular quadruplex topology for this DNA (Fig. S2, ESI†). Overall, the observations from CD titration studies suggest that the ligands are able to weakly induce an antiparallel topology in telomeric G-quadruplex DNA and can further stabilise the parallel topology of *c-MYC* and *c-KIT1* promoter quadruplex DNAs. Nevertheless, the

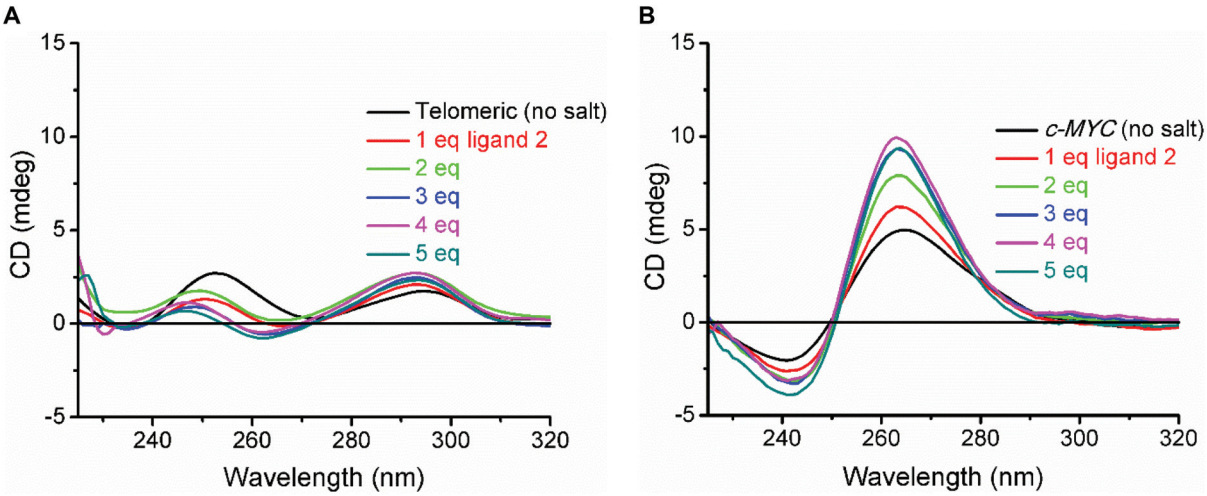


Fig. 2 CD titration spectra of telomeric and *c-MYC* DNAs (12.5 μM in 50 mM Tris-HCl, pH 7.2). (A) Telomeric DNA; and (B) *c-MYC* DNA with an increasing concentration of ligand **2** (0–5 equiv.) in the absence of added monovalent metal ions.

ligands fail to induce or stabilize an antiparallel topology in *h-RAS1* promoter DNA.

CD melting studies

CD melting experiments were carried out to assess the thermal stabilization properties of ligands with various G-quadruplex and duplex DNAs by following the reported experimental procedures.[43] The thermal stabilization of telomeric DNA as well as promoter and duplex DNAs was studied by monitoring at respective wavelengths of maximum intensities in the CD spectra. Melting experiments for the telomeric DNA were performed at 295 nm under K^+ conditions (10 mM) which yielded a melting temperature of 53 °C (Fig. 3A). After the addition of 5 equivalents of ligands, there was no considerable increase in the melting temperature with any of the three ligands (Table 1). Since the ligands were shown to weakly induce an antiparallel topology of the telomeric DNA, we examined the stabilization effect of ligands toward the antiparallel telomeric DNA structure by performing the melting experiment in the presence of Na^+ ions (30 mM).[44] The molar ellipticity of telomeric DNA was monitored at 295 nm to provide melting temperatures of 45 °C. The addition of 5 equivalents of ligands resulted only in a weak stabilization of the antiparallel topology ($\Delta T_m \sim 2.5$–3 °C) (Table 1 and Fig. S3, ESI†).

For parallel promoter G-quadruplex DNAs (*c-MYC* and *c-KIT1*), the melting temperatures (T_m) were monitored at 263 nm under K^+ conditions and high enhancement of thermal stabilizations were observed after the addition of ligands (Fig. 3B and S3, ESI†). The addition of ligand **2** to *c-MYC* and *c-KIT1* quadruplex DNAs furnished an increase in the melting temperature of 15.3 °C, whereas ligands **1** and **3** showed slightly less enhancement in melting temperatures ($\Delta T_m \sim 10$–13.5 °C) (Table 1, Fig. 3B and S3, ESI†). Additionally, the thermal stabilization effect of ligands on the

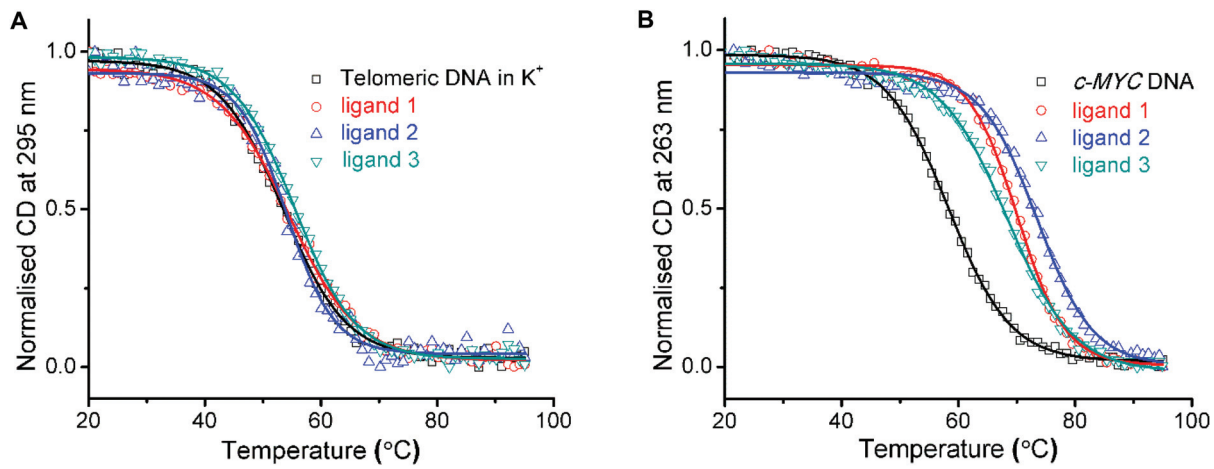


Fig. 3 Normalised CD melting curve of telomeric and *c-MYC* quadruplex DNAs (10 μM in 10 mM lithium cacodylate buffer, pH 7.2) in the absence and presence of 5 equivalents of ligands. (A) Telomeric DNA (10 mM KCl and 90 mM LiCl); and (B) *c-MYC* DNA (1 mM KCl and 99 mM LiCl).

Table 1 Thermal stabilization of various quadruplex and duplex DNAs determined from CD melting experiments

	ΔT_m [a]					
Ligands	Telomeric (K^+)	Telomeric (Na^+)	*c-MYC*	*c-KIT1*	*h-RAS1*	Duplex (ds 17)
1	1.2 ± 0.1	3.0 ± 0.1	11.8 ± 0.2	13.5 ± 0.9	2.3 ± 0.1	−0.5 ± 0.1
2	0.6 ± 0.1	2.5 ± 0	15.3 ± 0.2	15.3 ± 0.5	1.3 ± 0.1	−1.3 ± 0.1
3	1.7 ± 0.1	3.0 ± 0.2	10.0 ± 0.2	11.3 ± 0.5	2.4 ± 0.1	−1.0 ± 0.1

[a] ΔT_m stands for the difference in the thermal melting [$\Delta T_m = T_m$(DNA + 5 molar equivalent ligand) − T_m(DNA)]. The reported values are the average of three independent experiments with estimated standard deviations. All the experiments were carried out using 10 μM quadruplex DNA or 15 μM duplex DNA in 10 mM lithium cacodylate buffer, pH 7.2 for telomeric (K^+) and promoter DNAs or in 10 mM sodium cacodylate buffer, pH 7.2 for telomeric DNA (Na^+). The T_m values in the absence of ligands are 53.2 ± 0.2 for telomeric DNA in the presence of K^+ ions (10 mM KCl and 90 mM LiCl), 45.1 ± 0.5 for telomeric DNA in the presence of Na^+ ions (20 mM NaCl and 80 mM LiCl), 57.8 ± 0.4 for *c-MYC* DNA (1 mM KCl and 99 mM LiCl), 46.6 ± 0.6 for *c-KIT1* DNA (10 mM KCl and 90 mM LiCl), 53.8 ± 0.4 for *h-RAS1* DNA (50 mM KCl and 50 mM LiCl) and 64.0 ± 0.2 for ds 17 DNA (10 mM KCl and 90 mM LiCl).

antiparallel *h-RAS1* promoter DNA was investigated by monitoring the CD signal at 290 nm under K^+ conditions (50 mM). The addition of 5 equivalents of ligands provided a modest increase in the melting temperature (maximum $\Delta T_m \sim 2.5$ °C) suggesting very weak stabilization of this particular antiparallel topology (Table 1 and Fig. S3, ESI†). These results clearly show that ligand **2** with the propyl side chain is more efficient in achieving the preferential stabilization of parallel topologies of *c-MYC* and *c-KIT1* quadruplex DNAs.

To further investigate the selectivity of ligands toward G-quadruplex DNA over duplex DNAs, melting experiments were performed at 242 nm with the duplex DNA. The stabilizing effect of the ligands on duplex DNA was found to be negligible (Table 1 and Fig. S3, ESI†). Overall, results from CD melting studies reveal that all the three ligands preferentially stabilize the parallel topology of the promoter quadruplex (*c-MYC* and *c-KIT1*) over the telomeric quadruplex and duplex DNAs. Moreover, ligand **2** was found to be more potent in preferentially stabilizing the parallel promoter quadruplex DNAs when compared to the other two ligands.

Fluorescence intercalator displacement (FID) assay

To support the findings from the CD titration and melting results, FID assay was performed. FID assay is based on the efficiency of the ligands to displace thiazole orange (TO) from its complex with quadruplex DNA, and the percentage of displacement can be plotted against the concentration of the ligand to afford the DC_{50} values.[45] Low DC_{50} values represent the high binding affinity of the ligand toward quadruplex DNA. For the *c-MYC* DNA, all the three ligands were able to furnish low DC_{50} values (1.31–4.4 μM), indicating high stabilization of *c-MYC* quadruplex DNAs by the ligands (Table 2 and Fig. 4). Similarly, in the case of *c-KIT1* quadruplex DNA moderate DC_{50} (5.5–7.5 μM) values were observed for all the three ligands (Table 2 and Fig. S4, ESI†). However, for telomeric quadruplex DNA and duplex DNAs, high DC_{50} values (~15 μM to 27 μM) were obtained from FID studies. These results clearly indicate that the ligands are not able to strongly stabilize telomeric and duplex DNAs (Table 2 and Fig. S4, ESI†). For *c-MYC* and *c-KIT1* DNAs as observed from the CD melting experiments, ligand **2** with the propyl side chain, was found to be more potent with the lowest DC_{50} values compared to the other two ligands. In addition to that, all the three ligands showed a slight preference toward the *c-MYC* over the *c-KIT1* quadruplex DNA.

Table 2 DC_{50} values for telomeric, *c-MYC*, *c-KIT1* and duplex DNAs with all three ligands measured from FID assay

	TO displacement[a] (DC_{50} in μM)			
Ligands	Telomeric (K^+)	*c-MYC*	*c-KIT1*	Duplex (ds 17)
1	15.5 ± 0.5	3.36 ± 0.2	5.6 ± 0.1	17.6 ± 0.6
2	26.7 ± 0.1	1.31 ± 0.1	3.5 ± 0.2	25 ± 0.5
3	16.4 ± 0.3	4.43 ± 0.5	7.55 ± 0.9	21.3 ± 0.7

[a] DC_{50} denotes the concentration of ligands required to displace 50% of TO. All the experiments were triplicated and the values are shown with estimated standard deviations. For all the experiments, 0.25 μM DNA in 100 mM KCl and 10 mM lithium cacodylate buffer, pH 7.2 was used for the incubation (2 h) of 2 equivalents of TO for quadruplex DNAs and 3 equivalents of TO for duplex DNA.

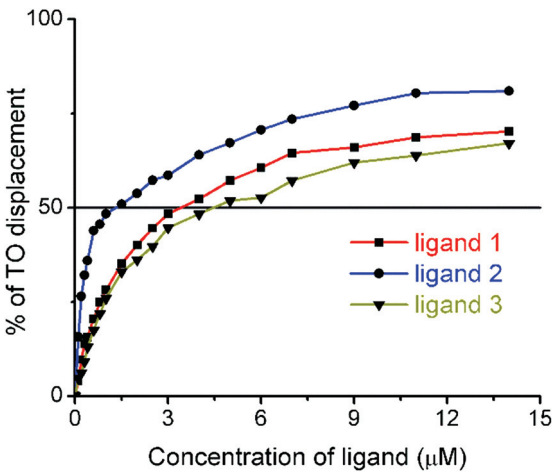


Fig. 4 Plot of percentage of TO displacement from *c-MYC* quadruplex DNA by ligands **1**, **2**, and **3**. FID titrations were carried out by using 0.25 μM DNA and 0.5 μM TO in buffer (100 mM KCl, 10 mM lithium cacodylate, pH 7.2). Ligand concentrations were varied from 0–14 μM.

Fluorimetric titration studies

To explore the preferential binding of ligands toward various quadruplex and duplex DNAs, fluorescence titration experiments were performed with ligand **2**, which evolved as the most effective quadruplex stabilizer. Ligand **2** showed an absorption maximum at 367 nm and strong fluorescence emission around 450 nm. An increasing concentration of the ligand (0–50 μM) was titrated against a fixed concentration of pre-annealed DNAs (5 μM) and emission spectra were recorded. Upon the addition of ligand **2** to *c-MYC* quadruplex DNA, a ~60 fold enhancement of the fluorescence intensity was observed in a concentration dependent manner and saturation was attained at ~24 μM ligand concentration (Fig. 5A). But, in the case of telomeric quadruplex DNAs similar enhancement of the fluorescence intensity was attained at a ligand concentration of 50 μM (Fig. S5, ESI†). As expected, the titration of ligand **2** with duplex DNA yielded negligible (~10 fold) enhancement in the fluorescence intensity at 50 μM saturation concentration (Fig. S5, ESI†). The binding constant was deduced by plotting the normalised fluorescence intensity against the logarithm of the concentration of the ligand and was fitted by using the Hill equation.[46,47] The apparent binding constant (K_a) of ligand **2** with *c-MYC* quadruplex DNA was found to be $(1.3 \pm 0.3) \times 10^5$ M^{-1} (Fig. 5B). Binding constants with telomeric quadruplex DNA and duplex DNA were found to be $(7 \pm 0.1) \times 10^4$ and $(2.7 \pm 0.1) \times 10^4$ M^{-1} respectively (Fig. S5, ESI†). The apparent K_a values emerging from the fluorimetric titration suggest that there is only a two and five fold preference for the ligand binding to *c-MYC* quadruplex over telomeric and duplex DNAs respectively. It should be noted here that irrespective of the moderate binding preference, ligands show strong stabilization of the parallel topologies of *c-MYC* and *c-KIT1* promoters, which was revealed from the CD melting, FID and *Taq* polymerase stop assays (see below).

Taq DNA polymerase stop assay

Having established the preferential stabilization and binding of ligands toward the *c-MYC* and the *c-KIT1* structures, we further investigated the quadruplex targeting efficiency of ligands with the aid of *Taq* DNA polymerase stop assay.[48,49] Experiments were performed with templates containing *c-MYC* and mutated *c-MYC* DNAs at 55 °C and with telomeric DNA at 40 °C. For the *c-MYC* DNA, there was a concentration dependant increase in the formation of stop products with all the three ligands (Fig. 6A and S6, ESI†). Among the three ligands, ligand **2** showed the lowest IC_{50} value of ~6.4 μM indicating higher stabilization of the *c-MYC* quadruplex DNA whereas, ligand **1** and ligand **3** showed moderate to high IC_{50} values (~14 and ~37 μM respectively) (Fig. 6B and S7, ESI†). As expected, there were no stop products observed for mutated *c-MYC* DNA under similar reaction conditions and ligand concentrations (Fig. 6A and S6, ESI†).

Furthermore, stop assay with the template containing telomeric DNA yielded very high IC_{50} values for ligands **1** and **3** (~140 μM and ~145 μM respectively) and there were only 16% stop products observed for ligand **2** even at 200 μM (Fig. S6, ESI†). These results show that all the three ligands are able to possess preferential stabilization toward the *c-MYC* promoter quadruplex over the telomeric quadruplex DNA. Moreover, the findings from stop assay are in good agreement with the outcomes from the biophysical studies, and ligand **2** is found to be more efficient toward the preferential stabilization of the *c-MYC* promoter quadruplex structure.

Molecular modeling and dynamics studies

To elucidate the mode of binding and the major factors governing the G-quadruplex recognition, molecular modeling and dynamics studies were carried out using ligand **2** and *c-MYC* and *c-KIT1* structures. The energy optimized ligand **2** at

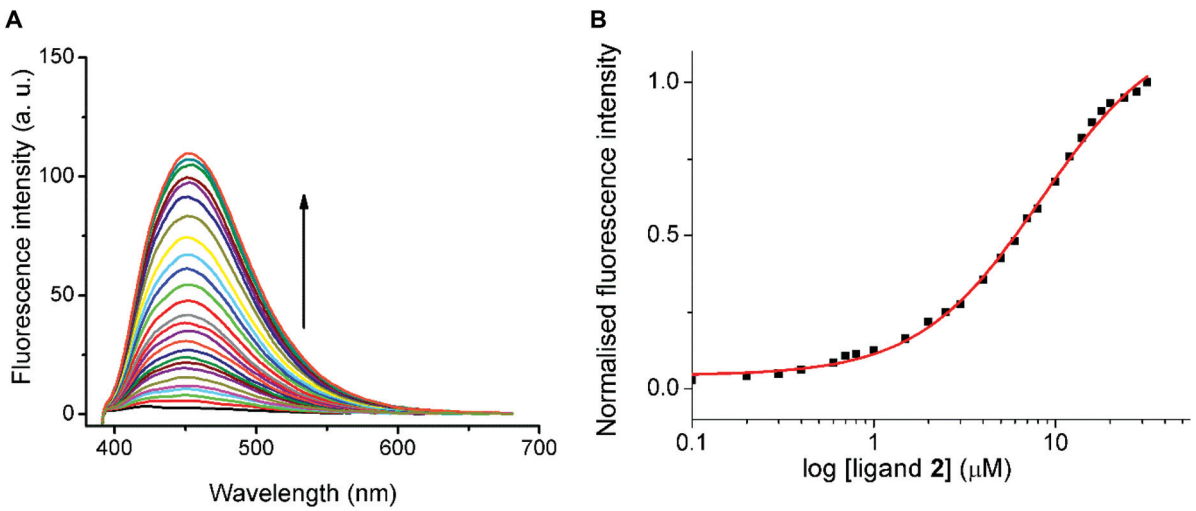


Fig. 5 (A) Emission spectra of ligand **2** complexed with *c-MYC* quadruplex DNA [5 μM in 100 mM KCl and 10 mM lithium cacodylate, pH 7.2 with an increasing concentration of ligand **2** (0–24 μM)] and (B) plot of normalised fluorescence intensity against the logarithm of an increasing concentration of ligand **2**. The Hill equation was used for curve fitting.

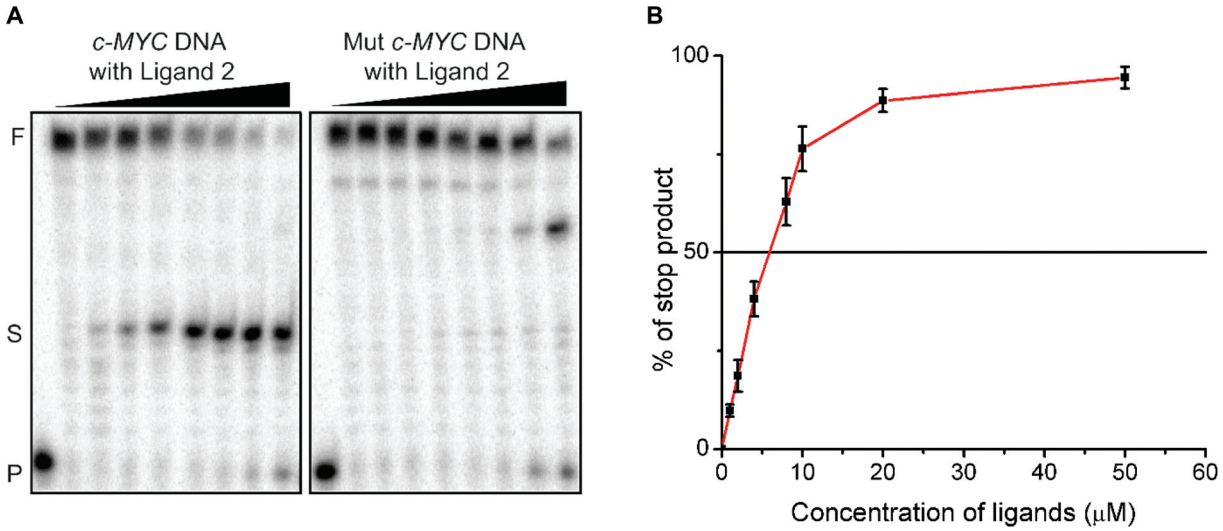


Fig. 6 (A) 15% denaturing PAGE (7 M urea) of *Taq* DNA polymerase stop assay of the *c-MYC* DNA template with an increasing concentration of ligand **2** (0–50 μM) and of the mutated *c-MYC* DNA template with ligand **2** (0–200 μM); and (B) plot of the percentage of the stop product against the increasing concentration of ligand **2**. Conditions: 100 nM template DNA, 50 nM primer, 0.2 mM dNTPs and 0.5 U *Taq* DNA polymerase enzyme in buffer (50 mM Tris-HCl, pH 7.2, 0.5 mM DTT, 0.1 mM EDTA, 5 mM $MgCl_2$, 5 mM KCl). P, S and F denote primer, stop product and full length product respectively.

HF/6-311+G** level having a *syn* conformation (Fig. S8, ESI†) was docked with *c-MYC* (PDB entry: 2L7V)[50] and *c-KIT1* (PDB entry: 2O3M)[51] DNAs using AutoDock 4.2.[52] While the whole DNA was considered in the case of docking with *c-MYC*, a restricted grid docking was carried out at the 5′-end of *c-KIT1*. The docking revealed the formation of 1 : 1 ligand–G-quadruplex complexes for both *c-MYC* and *c-KIT1*. The results show that ligand **2** preferably stacks at the 5′-end of *c-MYC* with the nitrogen atom on the propyl side chain involved in electrostatic interactions with the backbone phosphate atoms. With *c-KIT1*, it was observed that the furan and benzene rings of ligand **2** stack over the G-quartet core with the protonated nitrogen atom on the propyl side chain of the ligand involved in electrostatic interactions with the backbone phosphate atoms. Based on the docking results, a total of 500 ns unrestrained molecular dynamics (MD) simulations were performed with the complexes using PMEMD in AMBER 14.[53–56]

To verify the conformational stability of the complexes during the course of dynamics, the root mean square deviation (RMSD) values of the heavy atoms of ligand **2**, the DNA backbone, and the G-quartets with respect to the first frame were calculated for each of the complexes. The RMSD values (Fig. 7) of the G-quartets showed minimal fluctuations and suggest that ligand **2** stabilizes the G-quadruplexes. However, the backbone RMSD values fluctuate because of flanking and loop nucleotides. This is evident from the high root mean square fluctuation (RMSF) values of loop and flanking nucleotides compared to those of the guanines in the G-quartets (Fig. S9, ESI†). The ligand RMSD values fluctuate to some extent owing to the dynamic nature of the propyl side chain. The percentage occupancies of Hoogsteen H-bonds for the G-quartets were calculated and found to be >96% for both *c-MYC* and *c-KIT1* complexes (Fig. S10 and S11, ESI†). In the case of both the complexes, the top quartets (5′-end) where the ligand stacks showed >99% H-bond occupancies. This implies that the top G-quartets are being strongly stabilized by ligand **2**.

During the course of dynamics with the *c-MYC* DNA, ligand **2** undergoes a rigorous reorientation (Fig. S12, ESI†). Before starting the production run, the ligand completely stacks over the top G-quartet, but at a time frame of 380–400 ns, the ligand is slightly displaced by the flanking residue dA3, which in turn stacks on the G-quartet (Fig. S13A, ESI†). During this time, only the furan ring of ligand **2** exhibited stacking interactions with the G-quartet. The final MD snapshots are shown in Fig. 8, which clearly show that eventually ligand **2** replaces dA3 and the whole ligand stacks over the G-quartet core (Fig. 8C). Overall, the stacking interactions have been found to be crucial in stabilizing the ligand–*c-MYC* complex. The parameters, which contribute to stacking from each of the rings of ligand **2** are summarized in Table S1.† The stacking contribution from the benzothiazole ring of ligand **2** was observed to be maximum followed by the benzene ring and furan ring. To unravel the electrostatic interactions, the distance between the protonated nitrogen atom on the propyl side chain of ligand **2** and the phosphate atoms of the *c-MYC* backbone was probed. The major interactions observed were with the OP2 atom of the residue dG8 with an average distance of 2.9 ± 0.2 Å, and the OP1 atom of the residue dG13 with an average distance of 2.8 ± 0.2 Å (Fig. S14A and B, ESI†). Also, an ion-induced diploe interaction was observed between the nitrogen atom on the propyl side chain of the ligand and the O6 atom of the flanking residue dG2 for a very brief amount of simulation time (~4–5%) with an average distance of 3.0 ± 0.2 Å (Fig. S14C,

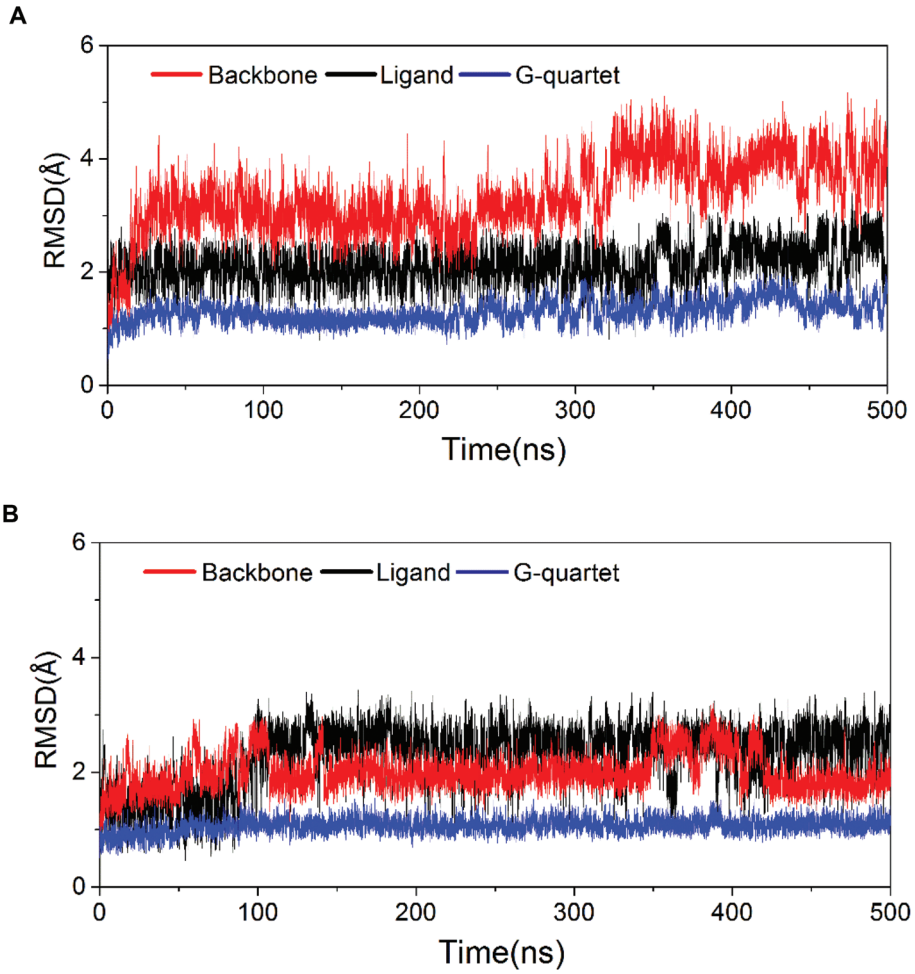


Fig. 7 Time dependent root mean square deviation (RMSD) graphs of (A) the ligand **2** bound *c-MYC* complex; and (B) the ligand **2** bound *c-KIT1* complex. RMSDs of the backbone (red), ligand (black) and G-quartet core (blue) are plotted against time. Only heavy atoms of a ligand were considered while calculating RMSD values. The G-quartet core includes the top, middle and bottom G-quartets and backbone atoms include C3′, O3′, C5′, O5′, P, OP1 and OP2.

ESI†). Overall, these electrostatic interactions were observed for ~12–15% of the simulation time (Fig. S15, ESI†).

It was also observed that the ligand undergoes reorientation during the course of dynamics with the *c-KIT1* DNA but less significantly compared to *c-MYC* DNA (Fig. 8B and S16, ESI†). Results show that the benzothiazole ring of ligand **2** is the major contributor for the stacking with G-quartet followed by the benzene ring (Table S1†). As seen from Fig. 8E, the benzothiazole ring of the ligand stacks over the dG12 residue, and the benzene ring on the dG5 residue. Electrostatic interactions also play a major role in stabilizing the ligand–*c-KIT1* complex. The protonated nitrogen atom on the propyl side chain of ligand **2** makes electrostatic contacts with the phosphate atom of the residue dG9 with an average distance of 2.8 ± 0.1 Å (Fig. S17A, ESI†). This interaction was present for ~45% of the simulation time (Fig. S18, ESI†), and was more predominant toward the end of dynamics. An ion-induced diploe interaction between the nitrogen atom on the propyl side chain of ligand **2** and O2 of dC10 (Fig. S17B, ESI†) was also observed but for a very brief amount of simulation time (~4–5%) with an average distance of 3.0 ± 0.2 Å. The loop residues dC10 and dT11 were found to be far away from the G-quartet core at the start of dynamics, but toward the end of dynamics, the residue dT11 tends to stack over the benzothiazole ring of ligand **2**, while the residue dC10 stacks well on dT11 (Fig. S13B, ESI†). Overall, the results reveal that stacking and electrostatic interactions play a major role in the stabilization of the *c-KIT1* G-quadruplex–ligand **2** complex. There were no prominent intermolecular H-bonds observed between the ligand and DNA in both the complexes.

Binding free energies (ΔG) for both G-quadruplex–ligand complexes were estimated using the MM-PBSA method.[57] The free energy components are summarized in Tables 3 and S2.† The binding free energy of ligand **2** with *c-KIT1* was −18.02 ± 6.00 kcal mol^{-1} and that with *c-MYC* was −31.18 ± 6.74 kcal mol^{-1}. Since the entropy loss of both the systems was found to be very similar, the binding is governed by gain in enthalpy, which was found to be high for *c-MYC* compared to *c-KIT1* quadruplex DNA (Table 3).

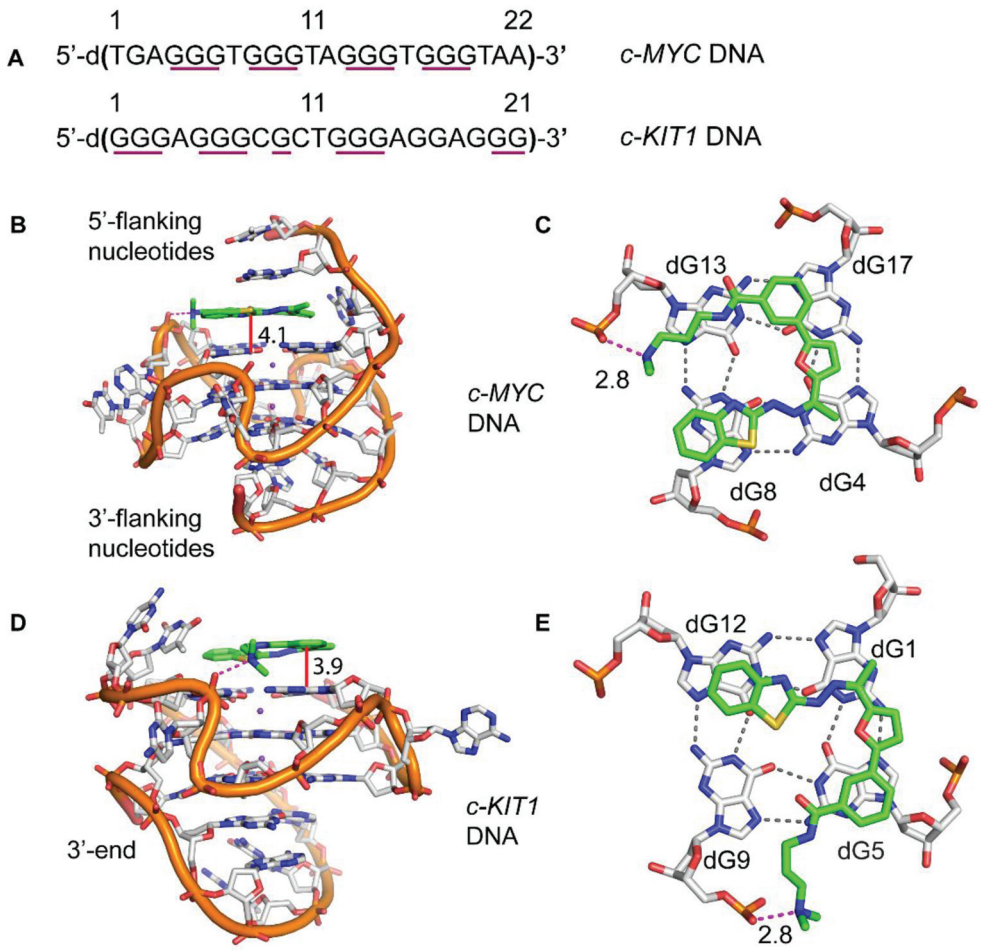


Fig. 8 Sequences of *c-MYC* and *c-KIT1* DNAs and averaged MD snapshots of ligand **2** bound *c-MYC* and *c-KIT1* from last 2 ns of 500 ns simulations. (A) Sequences of *c-MYC* and *c-KIT1* DNAs used for modeling studies; (B) ligand **2** bound to 5′-end face of *c-MYC* DNA (side view); (C) axial view of stacked ligand **2** with *c-MYC* DNA showing electrostatic interactions; (D) ligand **2** bound to the 5′-end face of *c-KIT1* DNA; and (E) axial view of stacked ligand **2** with *c-KIT1* DNA showing electrostatic interactions. The red lines in A and B indicate the average distance between nucleobases and the ligand. The dashed magenta lines represent the electrostatic interactions. The K^+ ions are shown as purple spheres. G-quartet forming sequences are underlined in purple. Distances are mentioned in Å.

Table 3 Binding free energy components of ligand **2**–quadruplex complexes

DNA	ΔH	$T\Delta S$	$\Delta G(\Delta H - T\Delta S)$
c-MYC (2L7V)	−48.8 ± 4.2	−17.6 ± 5.1	−31.2 ± 6.7
c-KIT1 (2O3M)	−35.6 ± 4.1	−17.6 ± 4.3	−18.0 ± 6.0

Enthalpy values are calculated using the MM-PBSA method and entropy values were obtained through nmode calculations. All values are in kcal mol^{-1}.

Summary and conclusions

In summary, we have synthesized a new series of benzothiazole hydrazones of furylbenzamide based potent G-quadruplex stabilizing ligands with different side chains and studied their interactions with various G-quadruplex and duplex DNAs. Based on the results from CD melting, FID assay and *Taq* DNA polymerase stop assay we conclude that all the ligands can preferentially stabilize the parallel topology of *c-MYC* and *c-KIT1* promoter G-quadruplex DNAs over the antiparallel topology of the *h-RAS1* promoter, and various topologies of telomeric quadruplex and duplex DNAs. Among the three ligands studied, ligand **2** with a propyl side chain emerged as the most promising candidate. Molecular modeling and dynamics studies revealed that along with the π–π stacking of the aromatic units of the molecular scaffold with the G-quartet, the electrostatic interaction of the positively charged side chains play an important role in stabilizing the parallel quadruplex DNA topology of c-*MYC* and *c-KIT1* promoter DNAs. Overall, our studies show that preferential binding toward a particular G-quadruplex topology can be achieved with the help of suitably designed simple structural scaffolds, which satisfy the desired criteria for small molecule drugs. Our current efforts are directed towards engineering these scaffolds to enhance the target specificity as well as to screen their biological

activities. It should be noted here that since RNA and some of the other promoter quadruplex DNAs can adopt parallel structures, it is indeed a challenging task to design ligands, which are capable of achieving discrimination between various parallel G-quadruplex topologies.

Experimental section

General

All the non-aqueous reactions were performed under the nitrogen atmosphere using oven-dried glasswares. MeOH and DCM were dried using calcium hydride and commercially available DMF was dried by sequential incubation with molecular sieves under the nitrogen atmosphere. Thin layer chromatography (TLC) was performed on silica gel 60 F_{254} plates pre-coated with a fluorescent indicator with visualization by UV light (254 nm). Silica gel (100–200 mesh) was used for column chromatography. ^{1}H NMR (400 MHz and 500 MHz), ^{13}C NMR (100 MHz and 125 MHz) and ^{19}F NMR (376 MHz and 470 MHz) were recorded on 400 MHz and 500 MHz instruments. The chemical shifts in parts per million (δ) were reported downfield from TMS (0 ppm) and referenced to the TMS signal (0 ppm) or residual proton signal of deuterated solvents as follows: $CDCl_3$ (7.26 ppm), DMSO-d_6 (2.50 ppm) and CD_3OD (3.31 ppm) for ^{1}H NMR spectra; $CDCl_3$ (77.2 ppm), DMSO-d_6 (39.5 ppm) and CD_3OD (49.1 ppm) for ^{13}C NMR spectra. Multiplicities of ^{1}H NMR spin couplings are reported as s (singlet), d (doublet), t (triplet), q (quartet), dd (doublet of doublet), dt (doublet of triplet), tt (triplet of triplet), br. s. (broad singlet) or m (multiplet and overlapping spin systems). Values for apparent coupling constants (J) are reported in Hz. High resolution mass spectra (HRMS) were obtained in positive ion electron spray ionization (ESI) mode using a Q-TOF analyser.

Synthetic procedures

(*Z*)-2,3,5,6-Tetrafluorophenyl-3-(5-(1-(2-(benzo[*d*]thiazol-2-yl)-hydrazono)ethyl)furan-2-yl)benzoate (5). Compound **4** (500 mg, 1.32 mmol) was dissolved in dry DMF (5 mL) with gentle heating. EDCI (504 mg, 2.64 mmol) and 2,3,5,6-tetrafluorophenol (422 mg, 2.64 mmol) were added and the reaction mixture was stirred at room temperature for 6 h. After completion of the reaction, DCM (3 × 50 mL) was added and the excess DMF and EDCI were removed by extraction with water (3 × 50 mL). The organic layer was dried over anhydrous Na_2SO_4, evaporated under reduced pressure, and purified by column chromatography on silica gel using 20% ethyl acetate in petroleum ether to give compound **5** as a yellow solid. Yield: 70% (501 mg); R_f = 0.3 (50% ethyl acetate in petroleum ether); mp = 189–192 °C. ^{1}H NMR (500 MHz, $CDCl_3$): δ 8.53 (t, J = 1.7 Hz, 1 H), 8.13 (dt, J = 7.9, 1.3 Hz, 1 H), 8.07 (dt, J = 7.9, 1.3 Hz, 1 H), 7.67 (d, J = 7.9 Hz, 1H), 7.62 (t, J = 7.8 Hz, 1H), 7.54 (d, J = 7.9 Hz, 1H), 7.34 (t, J = 7.5 Hz, 1H), 7.17 (t, J = 7.6 Hz, 1H), 7.07 (tt, J = 9.9, 7.0 Hz, 1H), 6.90 (q, J = 3.6 Hz, 2H), 2.31 (s, 3H). ^{13}C NMR (125 MHz, $CDCl_3$): δ 168.4, 162.3, 152.9, 152.1, 149.4, 147.4, 140.6, 131.1, 130.1, 129.8, 129.6, 129.4, 127.8, 126.0, 126.0, 122.3, 121.4, 118.2, 111.6, 108.7, 103.6, 103.4, 103.2, 12.8. ^{19}F NMR (470 MHz, $CDCl_3$): δ −138.7, −152.5. HRMS (ESI): calcd for $C_{26}H_{15}F_4N_3O_3S$, $[M + Na]^+$ 548.0662; found, $[M + Na]^+$ 548.0663 (Δm = −0.0001 and error = −0.1 ppm).

(*Z*)-3-(5-(1-(2-(Benzo[*d*]thiazol-2-yl)hydrazono)ethyl)furan-2-yl)-*N*-Boc- carbamimidoylbenzamide (6). Compound **5** (150 mg, 0.27 mmol) was dissolved in DCM (2 mL) and *N*-Boc-guanidine (220 mg, 1.38 mmol) was added followed by the addition of MeOH (2 mL). The reaction mixture was stirred at room temperature for 6 h. After completion of the reaction, the solvent was evaporated and purified by column chromatography on silica gel using 30% ethyl acetate in petroleum ether to give compound **6** as a yellow solid. Yield: 56% (77 mg); R_f = 0.5 (50% ethyl acetate in petroleum ether); mp = 205–210 °C. ^{1}H NMR (400 MHz, DMSO-d_6): δ 11.38 (s, 1H), 9.66 (br. s., 1H), 8.65 (br. s., 1H), 8.48 (s, 1H), 8.03 (d, J = 7.8 Hz, 1H), 7.92 (d, J = 7.5 Hz, 1H), 7.71 (d, J = 6.5 Hz, 1H), 7.55 (t, J = 7.8 Hz, 1H), 7.25–7.30 (m, 1H), 7.10 (d, J = 3.3 Hz, 1H), 7.07 (t, J = 7.6 Hz, 1H), 7.01 (d, J = 3.3 Hz, 1H), 2.33 (s, 3H), 1.48 (s, 9H). ^{13}C NMR (100 MHz, DMSO-d_6): δ 176.3, 167.8, 159.2, 154.6, 153.7, 152.3, 138.4, 130.1, 129.2, 128.5, 126.9, 126.4, 124.5, 122.2, 121.9, 112.6, 108.8, 82.2, 28.1, 14.3. HRMS (ESI): calcd for $C_{26}H_{26}N_6O_4S$, $[M + Na]^+$ 541.1628; found, $[M + Na]^+$ 541.1625 (Δm = +0.0003 and error = +0.6 ppm).

(*Z*)-3-(5-(1-(2-(Benzo[*d*]thiazol-2-yl)hydrazono)ethyl)furan-2-yl)-*N*-(3-(dimethylamino)propyl)benzamide (7). Compound **5** (80 mg, 0.15 mmol) was dissolved in DCM (3 mL) and to this *N*,*N*-dimethyl-1,3-propanediamine (0.2 mL, 1.5 mmol) was added and the reaction mixture was stirred at room temperature for 3 h. After completion of the reaction, DCM (3 × 50 mL) was added and the excess amine was removed by extraction with water (3 × 50 mL). The organic layer was dried over anhydrous Na_2SO_4, concentrated to obtain an orange coloured solid compound **7**. Yield: 67% (47 mg); R_f = 0.6 (15% MeOH in DCM). ^{1}H NMR (400 MHz, $CDCl_3$): δ 8.56 (t, J = 4.4 Hz, 1H), 8.13 (br. s., 1H), 7.85 (d, J = 7.6 Hz, 1H), 7.63–7.67 (m, 2H), 7.44–7.49 (m, 2H), 7.30 (t, J = 7.5 Hz, 1H), 7.14 (t, J = 7.5 Hz, 1H), 6.86 (d, J = 3.4 Hz, 1H), 6.80 (d, J = 2.4 Hz, 1H), 3.6 (q, J = 5.7 Hz, 2H), 2.50–2.56 (m, 2H), 2.31 (s, 6H), 2.24 (s, 3H), 1.78–1.83 (m, 2H). ^{13}C NMR (100 MHz, $CDCl_3$): δ 168.9, 167.2, 154.1, 152.1, 149.2, 141.7, 135.4, 130.9, 130.1, 129.2, 126.7, 126.3, 126.2, 123.1, 122.4, 121.7, 118.0, 111.8, 108.4, 58.6, 45.2, 40.1, 25.4, 13.3. HRMS (ESI): calcd for $C_{25}H_{27}N_5O_2S$, $[M + Na]^+$ 484.1778; found, $[M + Na]^+$ 484.1771 (Δm = +0.0007 and error = +1.3 ppm).

(*Z*)-1-(3-(5-(1-(2-(Benzo[*d*]thiazol-2-yl)hydrazono)ethyl)furan-2-yl)benzoyl)guanidinium 2,2,2-trifluoroacetate (ligand 1). Compound **6** (30 mg, 0.06 mmol) was dissolved in DCM (2 mL). To this, trifluoroacetic acid (0.05 mL, 0.6 mmol) was added and was allowed to stir for 1 h. Then the solvent was evaporated and dried under high vacuum to remove the excess of trifluoroacetic acid. The crude product was dissolved in a minimum amount of 1% MeOH in DCM and then precipitated with the addition of petroleum ether (2 mL). The supernatant

was decanted and the same precipitation procedure was repeated 6–7 times to obtain a pure yellow coloured solid. Yield: 98% (23.5 mg); R_f = 0.6 (20% MeOH in DCM); mp = 235–240 °C. ^{1}H NMR (400 MHz, DMSO-d_6): δ 11.59 (br. s., 1H), 8.66 (br. s., 2H), 8.43 (br. s., 2H), 8.29 (br. s., 1H), 8.09 (d, J = 7.8 Hz, 1H), 7.86 (d, J = 8.0 Hz, 1H), 7.67–7.73 (m, 2H), 7.27–7.32 (m, 2H), 7.24 (d, J = 3.5 Hz, 1H), 7.07 (t, J = 7.9 Hz, 1H), 7.03 (d, J = 3.5 Hz, 1H), 2.32 (br. s., 3H). ^{13}C NMR (100 MHz, DMSO-d_6): δ 167.4, 158.7, 158.4, 158.0, 155.1, 152.4, 152.1, 132.6, 130.5, 129.7, 128.3, 127.1, 126.0, 123.1, 121.8, 121.5, 117.5, 114.6, 112.3, 109.5, 13.8. ^{19}F NMR (376 MHz, DMSO-d_6): δ −74.4. HRMS (ESI): calcd for $C_{21}H_{19}N_6O_2S$, $[M]^+$ 419.1285; found, $[M]^+$ 419.1289 (Δm = −0.0004 and error = −1.0 ppm).

(*Z*)-3-(3-(5-(1-(2-(Benzo[*d*]thiazol-2-yl)hydrazono)ethyl)furan-2-yl)benzamido)-*N*,*N*-dimethylpropan-1-aminium 2,2,2-trifluoroacetate (ligand 2). Compound 7 (35 mg, 0.07 mmol) was dissolved in DCM (2 mL). To this, trifluoroacetic acid (0.05 mL, 0.7 mmol) was added and it was allowed to stir for 1 h. Then the solvent was evaporated and dried under high vacuum to remove the excess of trifluoroacetic acid. The crude mass was dissolved in a minimum amount of 1% MeOH in DCM and then precipitated with the addition of petroleum ether (2 mL) and finally the supernatant was decanted. The same precipitation procedure was repeated 6–7 times to obtain a pure yellow coloured sticky solid ligand **2**. Yield: 98% (14 mg); R_f = 0.6 (15% MeOH in DCM). ^{1}H NMR (500 MHz, CD_3OD): δ 8.23 (s, 1H), 7.91 (d, J = 7.9 Hz, 1H), 7.76 (d, J = 7.6 Hz, 1H), 7.70 (d, J = 7.6 Hz, 1H), 7.50 (t, J = 7.6 Hz, 1H), 7.45 (d, J = 7.9 Hz, 1H), 7.36 (t, J = 7.5 Hz, 1H), 7.20 (t, J = 7.5 Hz, 1H), 7.02 (d, J = 3.6 Hz, 1H), 6.96 (d, J = 3.4 Hz, 1H), 3.52 (t, J = 6.4 Hz, 2H), 3.20–3.24 (m, 2H), 2.92 (s, 6H), 2.35 (s, 3H), 2.07 (dt, J = 14.0, 6.7 Hz, 2H). ^{13}C NMR (125 MHz, CD_3OD): δ 170.3, 169.7, 155.8, 152.8, 146.6, 144.7, 135.9, 132.0, 130.4, 128.2, 128.0, 127.8, 124.5, 124.2, 123.3, 119.1, 116.8, 114.8, 109.7, 56.8, 48.8, 37.7, 26.3, 14.0. ^{19}F NMR (470 MHz, CD_3OD): δ −77.0. HRMS (ESI): calcd for $C_{25}H_{28}N_5O_2S$, $[M]^+$ 462.1958; found, $[M]^+$ 462.1962 (Δm = −0.0004 and error = +0.9 ppm).

(*Z*)-3-(5-(1-(2-(Benzo[*d*]thiazol-2-yl)hydrazono)ethyl)furan-2-yl)-*N*-(2-(dimethylamino)ethyl)benzamide (ligand 3). Compound **5** (170 mg, 0.32 mmol) was dissolved in DCM (5 mL) and *N*,*N*-dimethyl-1,2-ethanediamine (0.35 mL, 3.2 mmol) was added and the reaction mixture was stirred at room temperature for 3 h. After completion of the reaction, DCM (3 × 50 mL) was added and the excess amine was removed by extraction with water (3 × 50 mL). The organic layer was dried over anhydrous Na_2SO_4, evaporated under reduced pressure to give the yellow coloured solid compound **3**. Yield: 70% (102 mg); R_f = 0.5 (15% MeOH in DCM); mp = 193–196 °C. ^{1}H NMR (500 MHz, $CDCl_3$): δ 8.14 (br. s., 1H), 7.83 (d, J = 7.9 Hz, 1H), 7.62 (t, J = 6.4 Hz, 2H), 7.40–7.46 (m, 2H), 7.30 (t, J = 7.6 Hz, 1H), 7.13 (t, J = 7.6 Hz, 1H), 6.83 (d, J = 3.0 Hz, 1H), 6.79 (d, J = 2.7 Hz, 1H), 3.60 (q, J = 5.2 Hz, 2H), 2.60 (t, J = 5.8 Hz, 2H), 2.30 (s, 6H), 2.27 (s, 3H). ^{13}C NMR (125 MHz, $CDCl_3$): δ 168.7, 167.6, 153.8, 151.8, 148.7, 141.5, 135.1, 130.7, 129.8, 129.0, 126.6, 126.1, 126.1, 122.9, 122.2, 121.5, 117.6, 111.6, 108.3, 58.1, 45.0, 37.1, 13.1. HRMS (ESI): calcd for $C_{24}H_{25}N_5O_2S$, $[M + H]^+$ 448.1802; found, $[M + H]^+$ 448.1805 (Δm = +0.0003 and error = +0.8 ppm).

Ligand stock solution

Stock solutions (5 mM) of all the ligands were made in DMSO. For ligand **3**, the stock solution was further diluted to 1 mM using 10 mM HCl in water.

DNA oligonucleotides

The oligonucleotide sequences used for all the experiments are listed in Table S3 (ESI†). The oligonucleotides were synthesized on the 1 μM scale using a Mermade-4 DNA synthesizer. After synthesis, oligonucleotides were deprotected and purified by 20% PAGE using standard protocols. Oligonucleotides used for the fluorimetric titration studies were commercially purchased and purified by 20% PAGE. The concentration of all DNAs was calculated at 260 nm using the Perkin Elmer Lambda Bio$^+$ UV spectrophotometer using appropriate molar extinction coefficients (ε).

CD titration studies

CD titration experiments were performed using a Jasco 815 CD spectrometer. All the CD spectra were recorded in the wavelength range of 220–320 nm using a 1 mm path length quartz cuvette. The scanning speed was 100 nm min^{-1} with a response time of 2 s at 25 °C. The strand concentration of DNA was 12.5 μM in 50 mM of Tris-HCl buffer, pH 7.2. Initially, the baseline was recorded with 50 mM Tris-HCl buffer, pH 7.2, which was used for baseline subtraction. After each addition of the ligand (1 eq. each) to DNA, the solution was thoroughly mixed and kept for 5 minutes to attain equilibrium. Each spectrum was taken as an average of 3 measurements and analysed by using Origin 8.0 software.

CD melting studies

All the melting experiments were performed using a Jasco 815 CD spectrometer. For melting studies, a 10 μM strand concentration of quadruplex and 15 μM of duplex DNAs in 10 mM lithium cacodylate, pH 7.2 or sodium cacodylate, pH 7.2, the required amount of monovalent salts (KCl or NaCl and LiCl), and 5 molar equivalents of ligands (50–75 μM) were used. Telomeric DNA in K^+ ions (10 mM KCl and 90 mm LiCl), telomeric DNA in Na^+ ions (20 mM NaCl and 80 mM LiCl), *c-KIT1* DNA (10 mM KCl and 90 mM LiCl), *c-MYC* DNA (1 mM KCl and 99 mM LiCl), *h-RAS1* DNA (50 mM KCl and 50 mM LiCl) and duplex DNA (10 mM KCl and 90 mM LiCl) were annealed by heating at 95 °C for 5 min, followed by gradual cooling to room temperature over 3–4 h. Ligands (5 equivalents) were added to the annealed DNAs and kept at 4 °C for overnight. Thermal melting was monitored at 295, 290, 263, and 242 nm for telomeric, *h-RAS1*, promoter, and duplex DNAs respectively, at a heating rate of 1 °C min^{-1}. The melting temperatures were determined from the sigmoidal curve fit using the Boltzmann function in Origin 8.0 software.

Fluorescence intercalator displacement assay

FID assay was performed using a Cary Eclipse spectrofluorimeter using a 1 mL quartz cuvette having a 1 cm path length. The quadruplex and duplex DNA (0.25 μM) in 100 mM KCl and 10 mM lithium cacodylate buffer, pH 7.2 were annealed by heating at 95 °C for 5 minutes followed by slow cooling to room temperature for 3–4 h. To the annealed quadruplex DNA, 0.50 μM and to the duplex DNA, 0.75 μM of thiazole orange dye was added and allowed to equilibrate for 2 h at 4 °C. The emission spectrum was recorded in the range of 505–700 nm with an excitation wavelength of 500 nm with a slit width of 5 nm. A ligand (0–30 μM) was added to the DNA–TO complex and emission spectra were recorded for each addition after 3 minutes of equilibration. The fluorescence area under each curve was calculated and further converted to the percentage of the TO displacement value using the equation:

$$\text{Percentage of TO displacement} = 100 - \left[\left(\frac{F}{F_0}\right) \times 100\right]$$

where F_0 = fluorescence area without a ligand, F = fluorescence area in the presence of a ligand. The percentage of TO displacement was plotted against the concentration of ligands and the DC_{50} values were determined. All spectra were analysed by origin 8.0 software.

Fluorimetric titration studies

Fluorimetric titration experiments were performed using a Cary Eclipse spectrofluorimeter using a 1 mL quartz cuvette having a 1 cm path length. Quadruplex and duplex DNAs (5 μM) in 100 mM KCl and 10 mM lithium cacodylate buffer, pH 7.2 were annealed by heating at 95 °C for 5 minutes followed by gradual cooling to room temperature for 3–4 h. The DNAs were excited at an excitation wavelength of 367 nm and the emission spectra were recorded in the range of 372–700 nm using a slit width of 5 nm. Ligand **2** (0–24 μM for *c-MYC* DNA and 0–50 μM for telomeric and duplex DNAs) was titrated to the DNA sample and the emission spectra were recorded after equilibration for 3 minutes. A spectral blank was measured with only ligand **2** in the buffer and the corresponding fluorescence intensity was deducted from the original intensities of ligand **2** with DNA. The normalised fluorescence intensity (F_N) was plotted against the logarithm of the concentration of the ligand and the following equation (Hill equation) was used for curve fitting to derive the apparent binding constant (K_d) values:

$$F_N = F_0 + (F_S - F_0)\left(\frac{[L]^n}{[K_d]^n + [L]^n}\right)$$

Normalised fluorescence intensity (F_N) was derived using the equation:

$$F_N = \frac{(F_i - F_0)}{(F_s - F_0)}$$

where F_i = fluorescence intensity of the bound ligand at each titration point, F_0 = fluorescence intensity in the absence of a ligand, and F_s = fluorescence intensity at the saturation concentration of a ligand. L = concentration of the ligand, n = Hill coefficient *i.e.*, the degree of cooperative binding, K_d = dissociation constant ($1/K_a$). All the experiments were triplicated and the spectra were analysed by using Origin 8.0 software.

5′-End labelling

5′-End of the DNA primer (10 pmol) was labelled with [γ^{32}-P] ATP (10 μCi) by using T4 polynucleotide kinase (PNK, 5 U) in 1× PNK buffer (50 mM Tris-HCl, pH 7.6, 10 mM $MgCl_2$, 5 mM DTT, 0.1 mM spermidine and 0.1 mM EDTA). The total volume of the reaction mixture was maintained at 10 μL. The reaction mixture was incubated at 37 °C for 1 h for forward reaction, and the enzyme activity was stopped by heating at 70 °C for 3 minutes. Labelled DNA was purified by using the QIA quick nucleotide removal kit using the protocol provided by the manufacturer.

Taq DNA polymerase stop assay

Appropriate amounts of the radiolabelled primer, unlabelled primer (50 nM) and template (100 nM) in buffer (5 mM Tris-HCl, pH 8.0, 15 mM NaCl, 0.1 mM EDTA) were annealed by heating at 95 °C for 5 minutes followed by gradual cooling to room temperature for 3–4 h. The annealed primer-template DNA was mixed with 1× polymerase buffer (50 mM Tris-HCl, pH 8.0, 0.5 mM DTT, 0.1 mM EDTA, 5 mM $MgCl_2$, 5 mM KCl for *c-MYC* and 10 mM KCl for telomeric DNA template), 1 μg μL^{-1} BSA in 5% glycerol and 0.2 mM dNTPs. Appropriate amounts of ligands (0–50 μM for *c-MYC* and 0–200 μM for telomeric and mutated DNAs) were added in a total volume of 10 μL and the reaction mixture was incubated for 30 minutes at room temperature. *Taq* DNA polymerase enzyme (0.5 U) was added to initiate the primer extension reaction, and heated at 55 °C for *c-MYC* and 40 °C for the telomeric DNA template for 30 minutes. After completion of the reaction, the extension reaction was stopped by adding 10 μL of 2× stop dye (80% formamide, 1× TBE, 50 mM EDTA, pH 8.0, 0.025% each bromophenol blue and xylene cyanol FF) to the reaction mixture and the DNA was denatured at 95 °C for 3 minutes. The extension products were analysed by 15% denaturing PAGE (7 M urea) in 1× TBE (89 mM of Tris-HCl and boric acid each, 2 mM EDTA, pH 8.2) running buffer. The gels were kept inside the cassette for 2 h and analysed by autoradiography using a phosphorimager. The bands were quantified by using Image Quant 5.2 software and the percentage of the stop product was plotted against the concentration of ligands using Origin 8.0 software to derive the IC_{50} values.

Molecular modeling and dynamics studies

Ligand **2** was optimized in Gaussian09 [58] at different stages. The first step of optimization at a low theoretical level (PM3) was followed by optimization at a finer theoretical level (HF/6-311+G**). The final step involved the rotation of the N2–N1 bond to orient the benzothiazole moiety in the plane of the ligand, and then the ligand is optimized at the theoretical level HF/6-311+G** to obtain the final structure (Fig S8, ESI†). Optimized ligand **2** was docked with *c-MYC* (PDB entry: 2L7V)

and *c-KIT1* (PDB entry: 2O3M) DNAs using the Lamarckian genetic algorithm in AutoDock 4.2 [52] to generate 250 independent docked conformations. Ligands were removed from the *c-MYC* PDB file (2L7V) and the 5′-end residue dA was removed from the *c-KIT1* PDB file (2O3M) to facilitate docking. Also, the nitrogen atom on the propyl side chain of ligand **2** was protonated to explore the electrostatic interactions. A restricted grid docking was carried out to explore the conformational space available at the 5′-end of *c-KIT1*, whereas whole DNA was considered in the case of *c-MYC*. The generalized amber force field (GAFF)[59] and ff09bsc0 [60] force fields were used for the ligand and the DNA respectively. RESP[61] charge fitting was done using the antechamber[62] of AMBER 14. The system was solvated using TIP3P water molecules extending up to 10 Å in a truncated octahedron. K^+ ions were added to neutralize the system. The solvated system was subjected to 10 000 steps of minimization, 100 ps of heating with restriction on the ligand, 100 ps of density equilibration and 800 ps of final equilibration followed by 500 ns of unrestrained dynamics with coordinates saved every 2 ps. The time step for dynamics was 2 fs and the temperature was maintained at 300 K using a Langevin thermostat. The dynamics was run using GPU accelerated implementation of PMEMD in AMBER 14.[53–56] Binding free energies were estimated using the MM-PBSA[57] (MMPBSA.py) method over last 20 ns. Every 5th frame was considered, and a total of 2000 frames were considered for binding energy calculations. RMSDs of the heavy atoms, Hoogsteen bond occupancies and distances were calculated for every 10 ps (every 5th frame using CPPTRAJ and PTRAJ[63] modules). A cut off value of 3.5 Å between heavy atoms is considered for defining the H-bond and a cut off value of 3.5 Å is considered for defining electrostatic interactions. Cut off values of 5 Å between centroids of aromatic rings and an angle of 20° between planes of aromatic rings were considered for defining stacking interactions. RMSF values were calculated residue wise for every 5th frame and a total of 50 000 frames were considered. Every 25th frame (every 50 ps) was considered for the calculation of stacking distances and angles. UCSF Chimera[64] was used for the visualization of MD simulations, and calculation of stacking distances and angles. PyMOL (http://www.pymol.org) was used for rendering images.

Acknowledgements

We are thankful to Department of Biotechnology (DBT), Government of India (Pilot Project Grants for Young Investigators in Cancer Biology, Grant No. 6242-P4/RGCB/PMD/DBT/PKPI/2015) and IRCC IIT Bombay for financial support. We thank Professor Ruchi Anand and Professor Krishna P. Kaliappan for providing access to their laboratory facilities. We are grateful to the central facility supported by IRCC-IIT Bombay for MALDI spectra, and the computer centre, IIT Bombay, for the HPC facility. We also thank Ashil S. for his support in the synthesis of ligands, and S. Harikrishna and Ruhee D'cunha for their suggestions and assistance in molecular modelling studies. S. P. P. and K. V. D. thank the Council of Scientific and Industrial Research (CSIR), Government of India for Ph.D. fellowships, and P. B. thanks the Department of Science and Technology (DST), Government of India for the INSPIRE fellowship.

References

1 G. W. Collie and G. N. Parkinson, *Chem. Soc. Rev.*, 2011, **40**, 5867–5892.
2 S. M. Haider, S. Neidle and G. N. Parkinson, *Biochimie*, 2011, **93**, 1239–1251.
3 M. L. Bochman, K. Paeschke and V. A. Zakian, *Nat. Rev. Genet.*, 2012, **13**, 770–780.
4 N. Maizels and L. T. Gray, *PLoS Genet.*, 2013, **9**, e1003468.
5 J. L. Huppert and S. Balasubramanian, *Nucleic Acids Res.*, 2007, **35**, 406–413.
6 G. Biffi, D. Tannahill, J. McCafferty and S. Balasubramanian, *Nat. Chem.*, 2013, **5**, 182–186.
7 A. Shivalingam, M. A. Izquierdo, A. L. Marois, A. Vysniauskas, K. Suhling, M. K. Kuimova and R. Vilar, *Nat. Commun.*, 2015, **6**, 8178.
8 A. Henderson, Y. Wu, Y. C. Huang, E. A. Chavez, J. Platt, F. B. Johnson, R. M. Brosh, D. Sen and P. M. Lansdorp, *Nucleic Acids Res.*, 2014, **42**, 860–869.
9 S. Burge, G. N. Parkinson, P. Hazel, A. K. Todd and S. Neidle, *Nucleic Acids Res.*, 2006, **34**, 5402–5415.
10 N. An, A. M. Fleming, E. G. Middleton and C. J. Burrows, *Proc. Natl. Acad. Sci. U. S. A.*, 2014, **111**, 14325–14331.
11 Y. Chen and D. Yang, *Curr. Protoc. Nucleic Acid Chem.*, 2012, **50**, 17.15.11–17.15.17.
12 A. T. Phan, Y. S. Modi and D. J. Patel, *J. Am. Chem. Soc.*, 2004, **126**, 8710–8716.
13 A. M. Zahler, J. R. Williamson, T. R. Cech and D. M. Prescott, *Nature*, 1991, **350**, 718–720.
14 S. Balasubramanian, L. H. Hurley and S. Neidle, *Nat. Rev. Drug Discovery*, 2011, **10**, 261–275.
15 D. Rhodes and H. J. Lipps, *Nucleic Acids Res.*, 2015, **43**, 8627–8637.
16 J. Bidzinska, G. Cimino-Reale, N. Zaffaroni and M. Folini, *Molecules*, 2013, **18**, 12368–12395.
17 S. A. Ohnmacht and S. Neidle, *Bioorg. Med. Chem. Lett.*, 2014, **24**, 2602–2612.
18 D. Monchaud and M. P. Teulade-Fichou, *Org. Biomol. Chem.*, 2008, **6**, 627–636.
19 S. Neidle, *J. Med. Chem.*, 2016, DOI: 10.1021/acs.jmedchem.5b01835.
20 S. Zhang, Y. Wu and W. Zhang, *ChemMedChem*, 2014, **9**, 899–911.
21 M. Y. Kim, H. Vankayalapati, K. Shin-ya, K. Wierzba and L. H. Hurley, *J. Am. Chem. Soc.*, 2002, **124**, 2098–2099.
22 J. Dash, P. S. Shirude, S.-T. D. Hsu and S. Balasubramanian, *J. Am. Chem. Soc.*, 2008, **130**, 15950–15956.
23 V. Dhamodharan, S. Harikrishna, C. Jagadeeswaran, K. Halder and P. I. Pradeepkumar, *J. Org. Chem.*, 2012, **77**, 229–242.

24 W. J. Chen, C. X. Zhou, P. F. Yao, X. X. Wang, J. H. Tan, D. Li, T. M. Ou, L. Q. Gu and Z. S. Huang, *Bioorg. Med. Chem.*, 2012, **20**, 2829–2836.
25 A. Altieri, A. Alvino, S. Ohnmacht, G. Ortaggi, S. Neidle, D. Nocioni, M. Franceschin and A. Bianco, *Molecules*, 2013, **18**, 13446–13470.
26 P. V. Boddupally, S. Hahn, C. Beman, B. De, T. A. Brooks, V. Gokhale and L. H. Hurley, *J. Med. Chem.*, 2012, **55**, 6076–6086.
27 R. V. Brown, F. L. Danford, V. Gokhale, L. H. Hurley and T. A. Brooks, *J. Biol. Chem.*, 2011, **286**, 41018–41027.
28 Y. Ma, T. M. Ou, J. H. Tan, J. Q. Hou, S. L. Huang, L. Q. Gu and Z. S. Huang, *Eur. J. Med. Chem.*, 2011, **46**, 1906–1913.
29 D. Peng, J. H. Tan, S. B. Chen, T. M. Ou, L. Q. Gu and Z. S. Huang, *Bioorg. Med. Chem.*, 2010, **18**, 8235–8242.
30 T. Agarwal, S. Roy, T. K. Chakraborty and S. Maiti, *Biochemistry*, 2010, **49**, 8388–8397.
31 K. V. Diveshkumar, S. Sakrikar, S. Harikrishna, V. Dhamodharan and P. I. Pradeepkumar, *ChemMedChem*, 2014, **9**, 2754–2765.
32 V. Dhamodharan, S. Harikrishna, A. C. Bhasikuttan and P. I. Pradeepkumar, *ACS Chem. Biol.*, 2015, **10**, 821–833.
33 A. Chauhan, S. Paladhi, M. Debnath, S. Mandal, R. N. Das, S. Bhowmik and J. Dash, *Bioorg. Med. Chem.*, 2014, **22**, 4422–4429.
34 K. M. Felsenstein, L. B. Saunders, J. K. Simmons, E. Leon, D. R. Calabrese, S. Zhang, A. Michalowski, P. Gareiss, B. A. Mock and J. S. Schneekloth, *ACS Chem. Biol.*, 2016, **11**, 139–148.
35 G. Lessene, P. E. Czabotar, B. E. Sleebs, K. Zobel, K. N. Lowes, J. M. Adams, J. B. Baell, P. M. Colman, K. Deshayes, W. J. Fairbrother, J. A. Flygare, P. Gibbons, W. J. A. Kersten, S. Kulasegaram, R. M. Moss, J. P. Parisot, B. J. Smith, I. P. Street, H. Yang, D. C. S. Huang and K. G. Watson, *Nat. Chem. Biol.*, 2013, **9**, 390–397.
36 M. Yoshida, I. Hayakawa, N. Hayashi, T. Agatsuma, Y. Oda, F. Tanzawa, S. Iwasaki, K. Koyama, H. Furukawa, S. Kurakata and Y. Sugano, *Bioorg. Med. Chem. Lett.*, 2005, **15**, 3328–3332.
37 S. Chandrappa, C. V. Kavitha, M. S. Shahabuddin, K. Vinaya, C. S. Ananda Kumar, S. R. Ranganatha, S. C. Raghavan and K. S. Rangappa, *Bioorg. Med. Chem.*, 2009, **17**, 2576–2584.
38 B. S. Holla, B. S. Rao, B. K. Sarojini and P. M. Akberali, *Eur. J. Med. Chem.*, 2004, **39**, 777–783.
39 B. E. Sleebs, W. J. A. Kersten, S. Kulasegaram, G. Nikolakopoulos, E. Hatzis, R. M. Moss, J. P. Parisot, H. Yang, P. E. Czabotar, W. D. Fairlie, E. F. Lee, J. M. Adams, L. Chen, M. F. van Delft, K. N. Lowes, A. Wei, D. C. S. Huang, P. M. Colman, I. P. Street, J. B. Baell, K. Watson and G. Lessene, *J. Med. Chem.*, 2013, **56**, 5514–5540.
40 M. K. Smalley and S. K. Silverman, *Nucleic Acids Res.*, 2006, **34**, 152–166.
41 A. Randazzo, G. Spada and M. W. Silva, *Top. Curr. Chem.*, 2013, **330**, 67–86.
42 A. Membrino, S. Cogoi, E. B. Pedersen and L. E. Xodo, *PLoS One*, 2011, **6**, e24421.
43 A. Guedin, L. Lacroix and J. L. Mergny, *Methods Mol. Biol.*, 2010, **613**, 25–35.
44 Y. Wang and D. J. Patel, *Structure*, 1993, **1**, 263–282.
45 D. Monchaud, C. Allain, H. Bertrand, N. Smargiasso, F. Rosu, V. Gabelica, A. De Cian, J. L. Mergny and M. P. Teulade-Fichou, *Biochimie*, 2008, **90**, 1207–1223.
46 S. G. Srivatsan and Y. Tor, *J. Am. Chem. Soc.*, 2007, **129**, 2044–2053.
47 A. A. Tanpure and S. G. Srivatsan, *Nucleic Acids Res.*, 2015, **43**, e149.
48 H. Han, L. H. Hurley and M. Salazar, *Nucleic Acids Res.*, 1999, **27**, 537–542.
49 D. Sun and L. H. Hurley, *Methods Mol. Biol.*, 2010, **608**, 65–79.
50 J. Dai, M. Carver, L. H. Hurley and D. Yang, *J. Am. Chem. Soc.*, 2011, **133**, 17673–17680.
51 A. T. Phan, V. Kuryavyi, S. Burge, S. Neidle and D. J. Patel, *J. Am. Chem. Soc.*, 2007, **129**, 4386–4392.
52 G. M. Morris, D. S. Goodsell, R. S. Halliday, R. Huey, W. E. Hart, R. K. Belew and A. J. Olson, *J. Comput. Chem.*, 1998, **19**, 1639–1662.
53 R. Salomon-Ferrer, A. W. Götz, D. Poole, S. Le Grand and R. C. Walker, *J. Chem. Theory Comput.*, 2013, **9**, 3878–3888.
54 A. W. Götz, M. J. Williamson, D. Xu, D. Poole, S. Le Grand and R. C. Walker, *J. Chem. Theory Comput.*, 2012, **8**, 1542–1555.
55 S. Le Grand, A. W. Götz and R. C. Walker, *Comput. Phys. Commun.*, 2013, **184**, 374–380.
56 D. A. Case, T. A. Darden, T. E. Cheatham, C. L. Simmerling, J. Wang, R. E. Duke, R. Luo, R. C. Walker, W. Zhang, K. M. Merz, B. Roberts, B. Wang, S. Hayik, A. Roitberg, G. Seabra, I. Kolossváry, K. F. Wong, F. Paesani, J. Vanicek, J. Liu, X. Wu, S. R. Brozell, T. Steinbrecher, H. Gohlke, Q. Cai, X. Ye, J. Wang, M. J. Hsieh, G. Cui, D. R. Roe, D. H. Mathews, M. G. Seetin, C. Sagui, V. Babin, S. Gusarov, A. Kovalenko and P. A. Kollman, *AMBER 14*, University of California, San Francisco, CA, USA, 2014.
57 P. A. Kollman, I. Massova, C. Reyes, B. Kuhn, S. Huo, L. Chong, M. Lee, T. Lee, Y. Duan, W. Wang, O. Donini, P. Cieplak, J. Srinivasan, D. A. Case and T. E. Cheatham, *Acc. Chem. Res.*, 2000, **33**, 889–897.
58 M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery Jr., J. E. Peralta, F. Ogliaro, M. J. Bearpark, J. Heyd, E. N. Brothers, K. N. Kudin, V. N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. P. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, N. J. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador,

J. J. Dannenberg, S. Dapprich, A. D. Daniels, Ö. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski and D. J. Fox, *Gaussian 09, Revision A. 02*, Gaussian Inc., Wallingford, CT, USA, 2009.

59 J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman and D. A. Case, *J. Comput. Chem.*, 2004, **25**, 1157–1174.

60 M. Zgarbová, M. Otyepka, J. Šponer, A. Mládek, P. Banáš, T. E. Cheatham and P. Jurečka, *J. Chem. Theory Comput.*, 2011, **7**, 2886–2902.

61 T. Fox and P. A. Kollman, *J. Phys. Chem. B*, 1998, **102**, 8070–8079.

62 J. Wang, W. Wang, P. A. Kollman and D. A. Case, *J. Mol. Graphics Modell.*, 2006, **25**, 247–260.

63 D. R. Roe and T. E. Cheatham, *J. Chem. Theory Comput.*, 2013, **9**, 3084–3095.

64 E. F. Pettersen, T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt, E. C. Meng and T. E. Ferrin, *J. Comput. Chem.*, 2004, **25**, 1605–1612.

Electronic Supplementary Material (ESI) for Organic & Biomolecular Chemistry.

Benzothiazole hydrazone of furylbenzamides preferentially stabilize *c-MYC* and *c-KIT1* promoter G-quadruplex DNAs

Sushree Prangya Priyadarshinee Pany, Praneeth Bommisetti, K. V. Diveshkumar, and P. I. Pradeepkumar*

Department of Chemistry, Indian Institute of Technology Bombay, Powai, Mumbai-400076, India

Email: pradeep@chem.iitb.ac.in

TABLE OF CONTENTS

CD spectra of ligands with telomeric and *c-MYC* DNAs in the absence of added metal cations

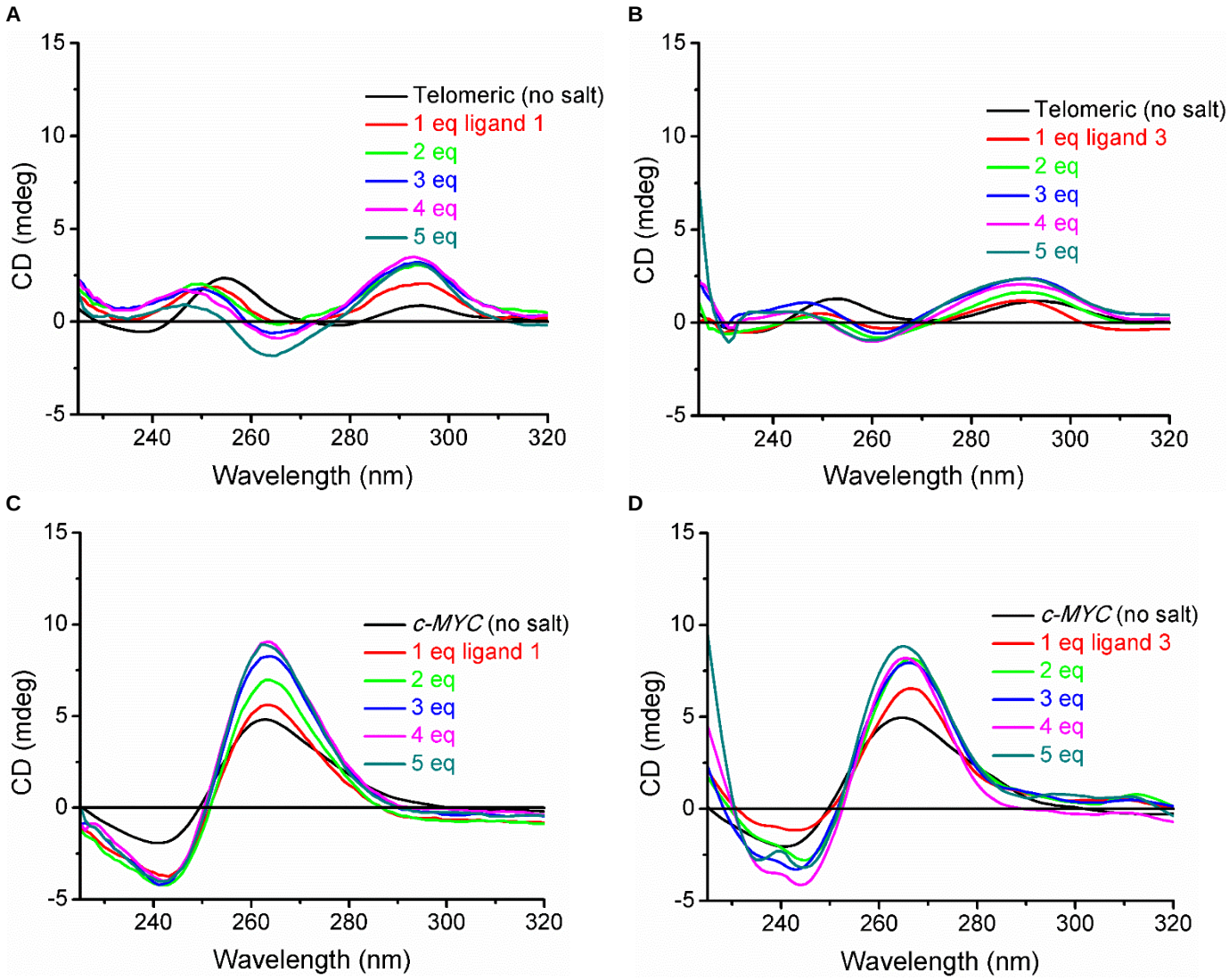


Figure S1. CD titration spectra of telomeric and *c-MYC* DNA (12.5 μM in 50 mM Tris-HCl buffer, pH 7.2). (A) Telomeric DNA with ligand **1**; (B) Telomeric DNA with ligand **3**; (C) *c-MYC* DNA with ligand **1**; and (D) *c-MYC* DNA with ligand **3** in the absence of added monovalent cations.

CD spectra of ligands with *c-KIT1* and *h-RAS1* DNA in the absence of added metal cations

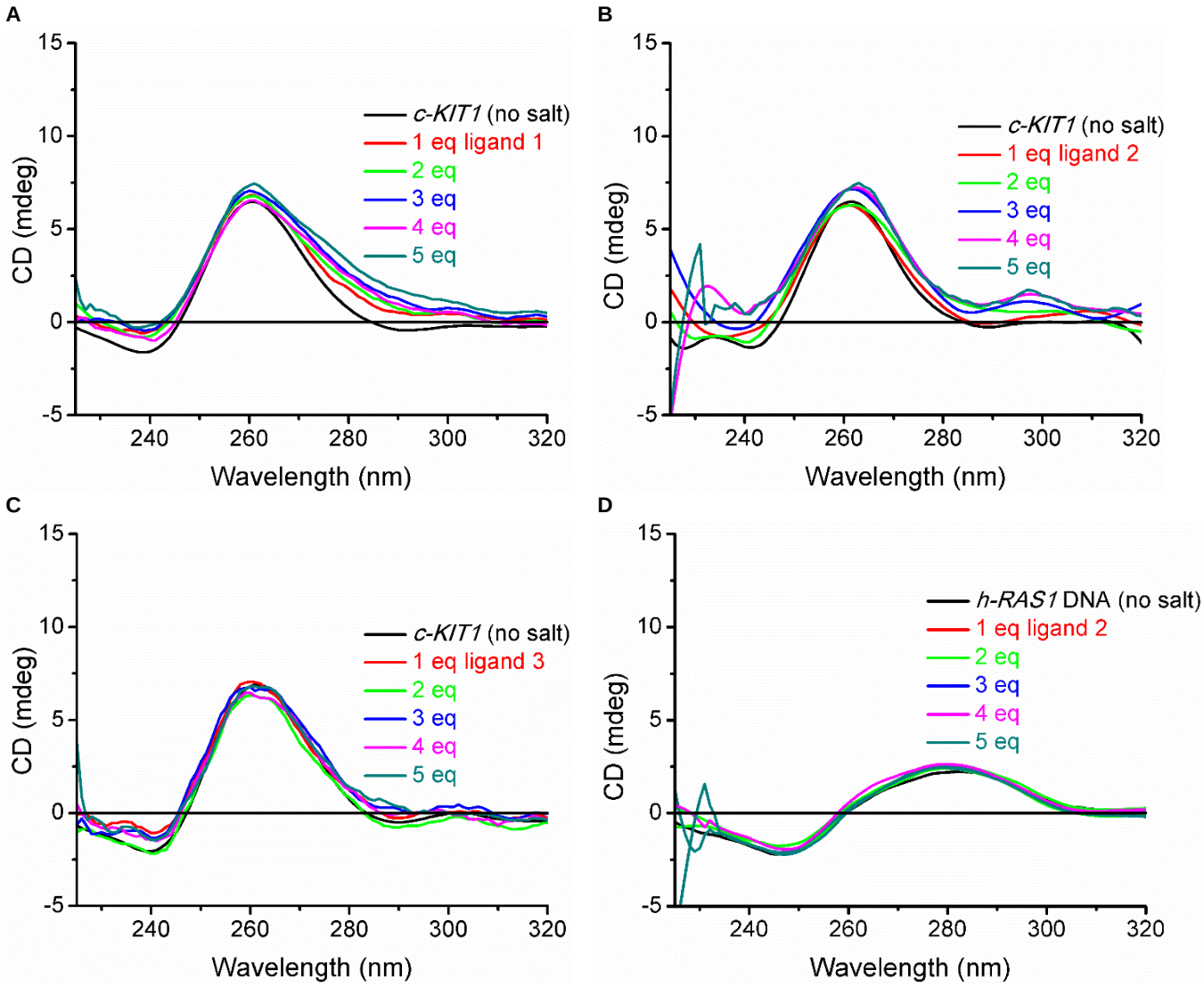


Figure S2. CD titration spectra of *c-KIT1* and *h-RAS1* DNA (12.5 μM in 50 mM Tris-HCl buffer, pH 7.2) in the absence of added monovalent cations (A) *c-KIT1* DNA with ligand **1**; (B) *c-KIT1* DNA with ligand **2**; (C) *c-KIT1* DNA with ligand **3**; and (D) *h-RAS1* DNA with ligand **2**

Normalised CD melting curves of telomeric, *c-KIT1*, *h-RAS1* and duplex DNAs

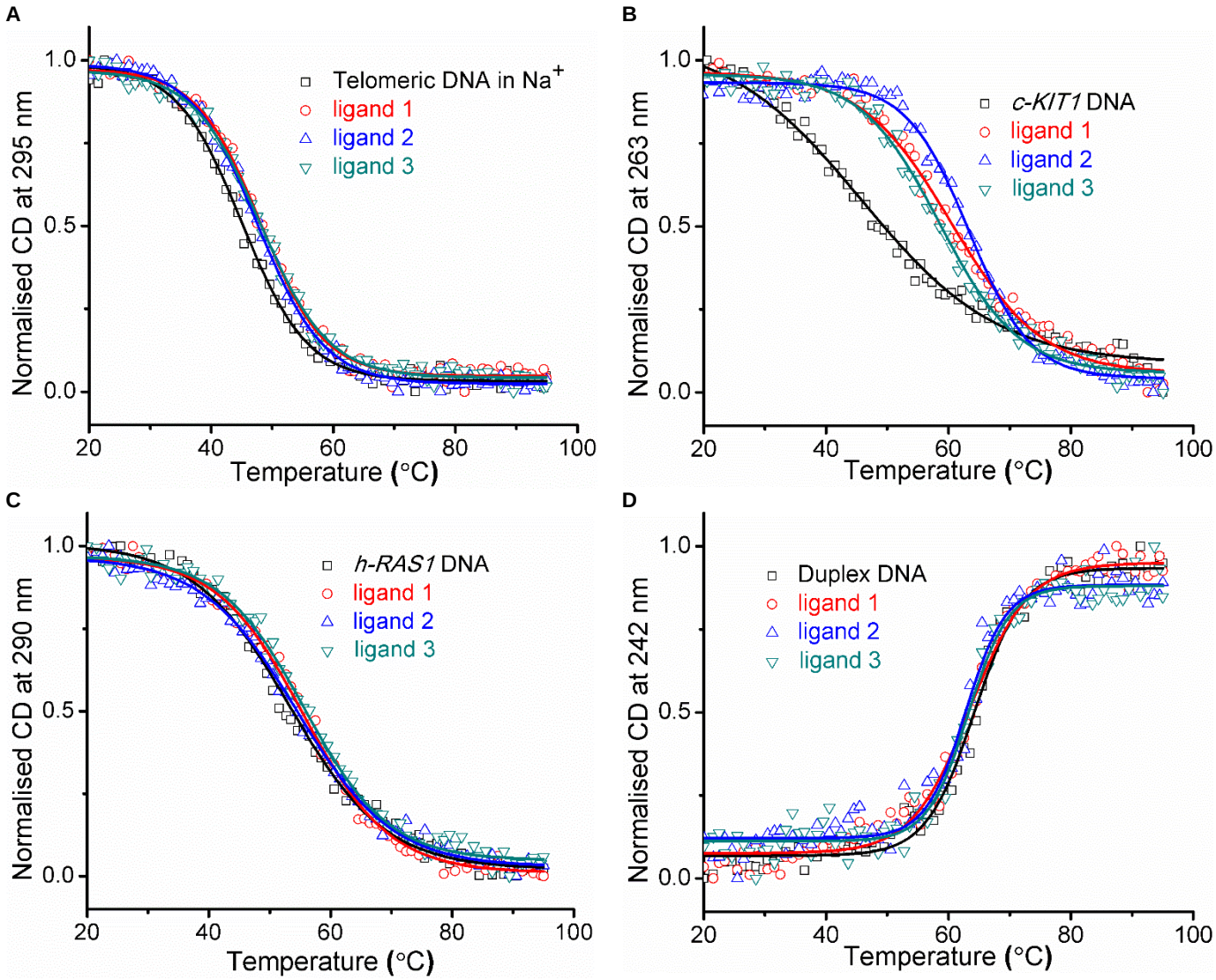


Figure S3. Normalised CD melting curves of various quadruplex DNAs (10 μM) and duplex DNA (15 μM) in the absence and presence of 5 equivalents of ligands. (A) Telomeric DNA in Na^+ ion (20 mM NaCl, 80 mM LiCl and 10 mM sodium cacodylate buffer, pH 7.2); (B) *c-KIT1* DNA (10 mM KCl, 90 mM LiCl and 10 mM lithium cacodylate buffer, pH 7.2); (C) *h-RAS1* DNA (50 mM KCl, 50 mM LiCl and 10 mM lithium cacodylate buffer, pH 7.2); and (D) Duplex DNA (10 mM KCl, 90 mM LiCl and 10 mM lithium cacodylate buffer, pH 7.2)

DC_{50} plots for *c-KIT1*, telomeric and duplex DNAs from the FID assay

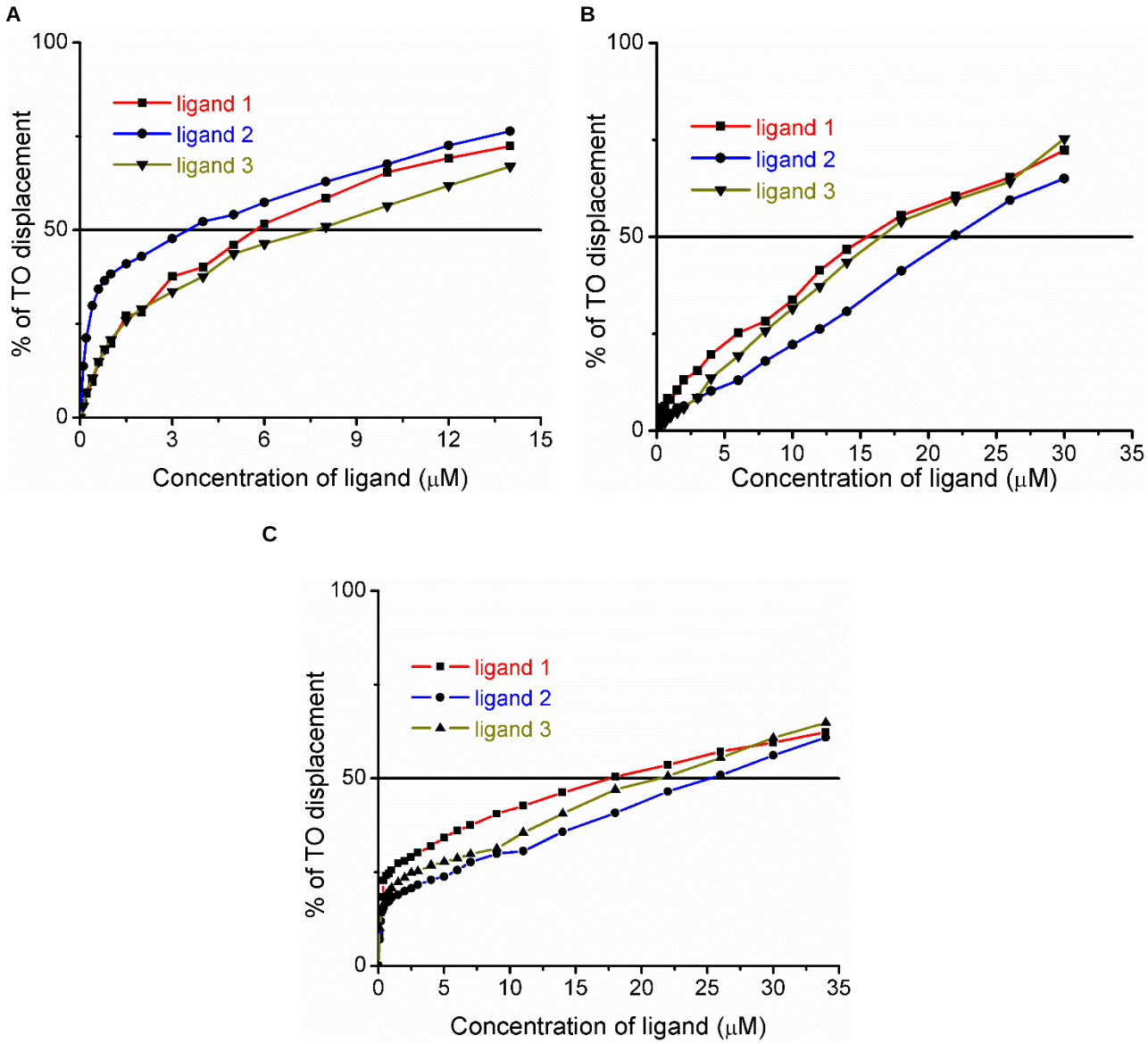


Figure S4. Plot of percentage of TO displacement for *c-KIT1*, duplex and telomeric DNAs (0.25 μM in 100 mM KCl, 10 mM lithium cacodylate buffer, pH 7.2). (A) *c-KIT1* DNA (against 0-14 μM ligand concentration); (B) Telomeric DNA (against 0-30 μM ligand concentration); and (C) Duplex DNA (against 0-30 μM ligand concentration). 2 equivalents of TO was used for quadruplex DNA and 3 equivalents of TO was used for duplex DNA.

Fluorimetric titration curves of ligand 2 complexed with telomeric and duplex DNAs

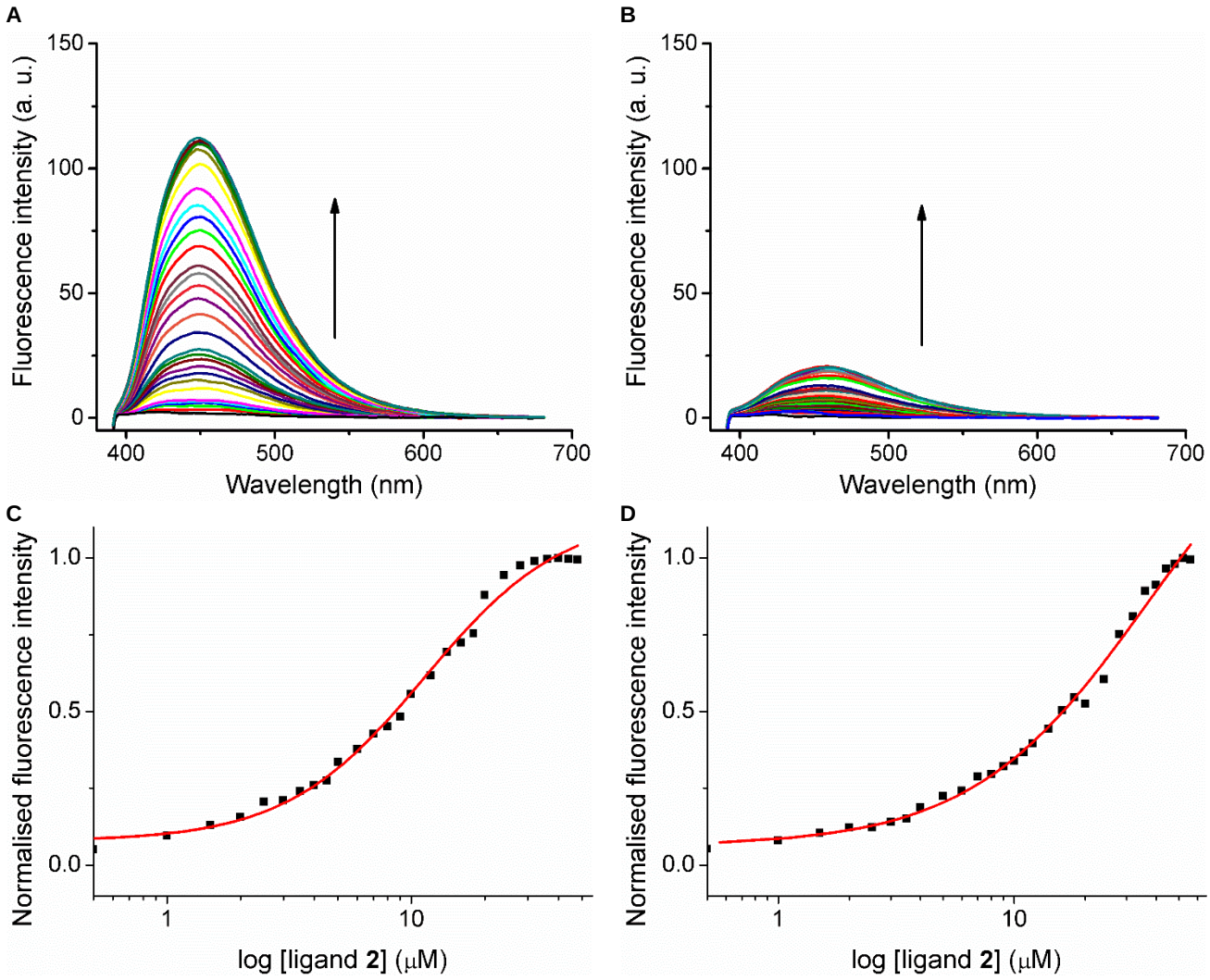


Figure S5. Fluorimetric titration of curves of ligand **2** complexed with telomeric and duplex DNA (5 μM in 100 mM KCl and 10 mM lithium cacodylate buffer, pH 7.2 and 0-50 μM ligand **2**). (A) Emission spectra of ligand **2** complexed with telomeric quadruplex DNA; and (B) Emission spectra of ligand **2** complexed with duplex DNA; (C) Plot of normalised fluorescence intensity for telomeric quadruplex DNA against the logarithm of increasing concentration of ligand **2**; and (D) Plot of normalised fluorescence intensity for duplex DNA against the logarithm of increasing concentration of ligand **2**. The Hill1 equation was used for the curve fitting.

Denaturing PAGE of *Taq* polymerase stop assay for ligands with template containing *c-MYC*, mutated *c-MYC* and telomeric DNA

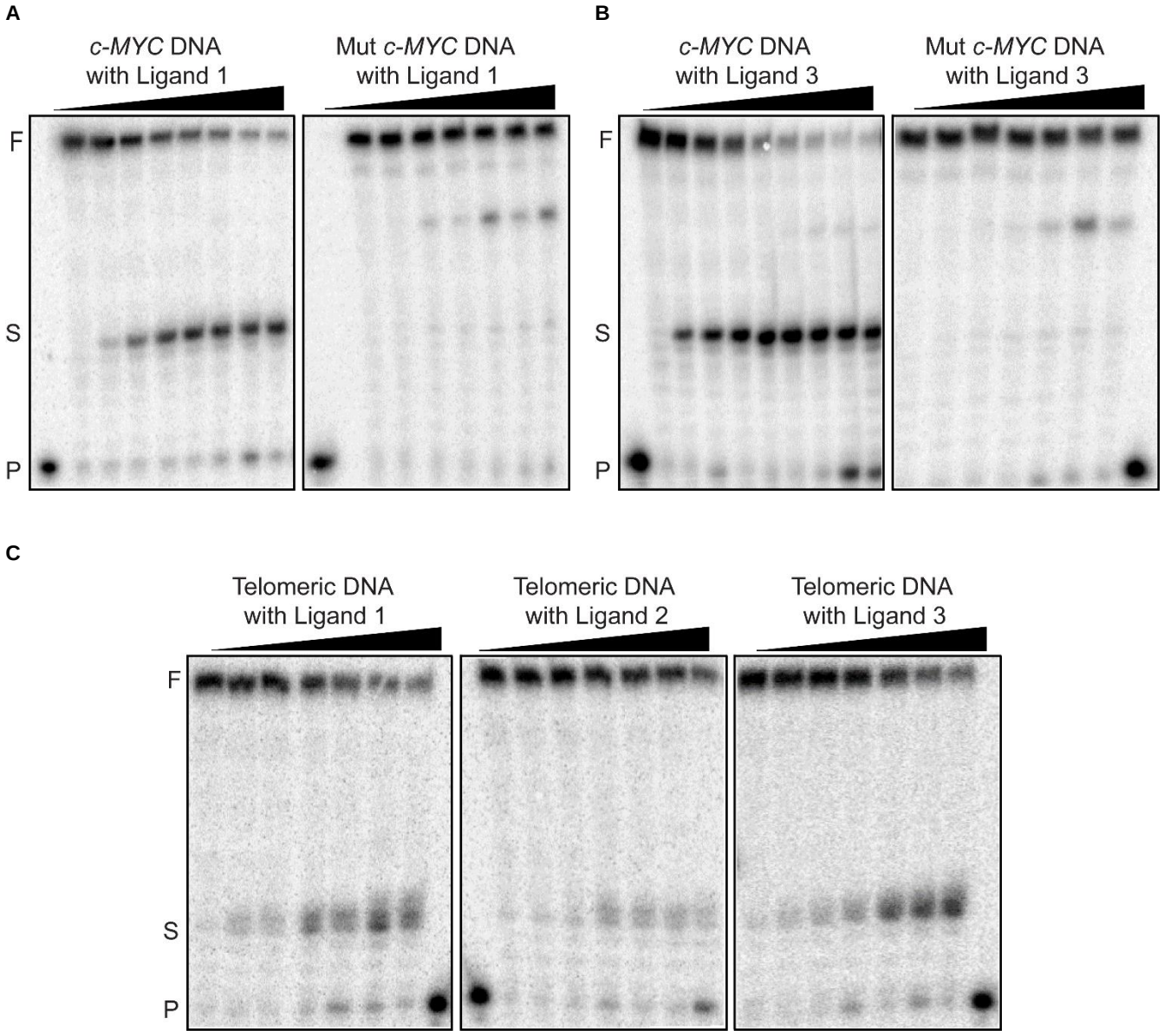


Figure S6. 15% denaturing PAGE (7 M urea) of *Taq* DNA polymerase stop assay for *c-MYC* DNA, mutated *c-MYC* DNA and telomeric DNA templates with increasing concentration of ligands (0-200 μM). (A) *c-MYC* and mutated *c-MYC* DNA template with ligand **1**; (B) *c-MYC* and mutated *c-MYC* DNA template with ligand **3**; and (C) Telomeric DNA template with ligands **1**, **2** and **3**. Conditions: 100 nM template DNA, 50 nM primer, 0.2 mM dNTPs, 0.5 U *Taq* DNA polymerase enzyme in buffer (50 mM Tris-HCl, pH 7.2, 0.5 mM DTT, 0.1 mM EDTA, 5 mM $MgCl_2$, 5 mM KCl for *c-MYC* DNA and 10 mM KCl for telomeric DNA). P, S and F denote primer, stop and full length products respectively.

IC_{50} plots of *Taq* polymerase stop assay for ligands with *c-MYC* DNA

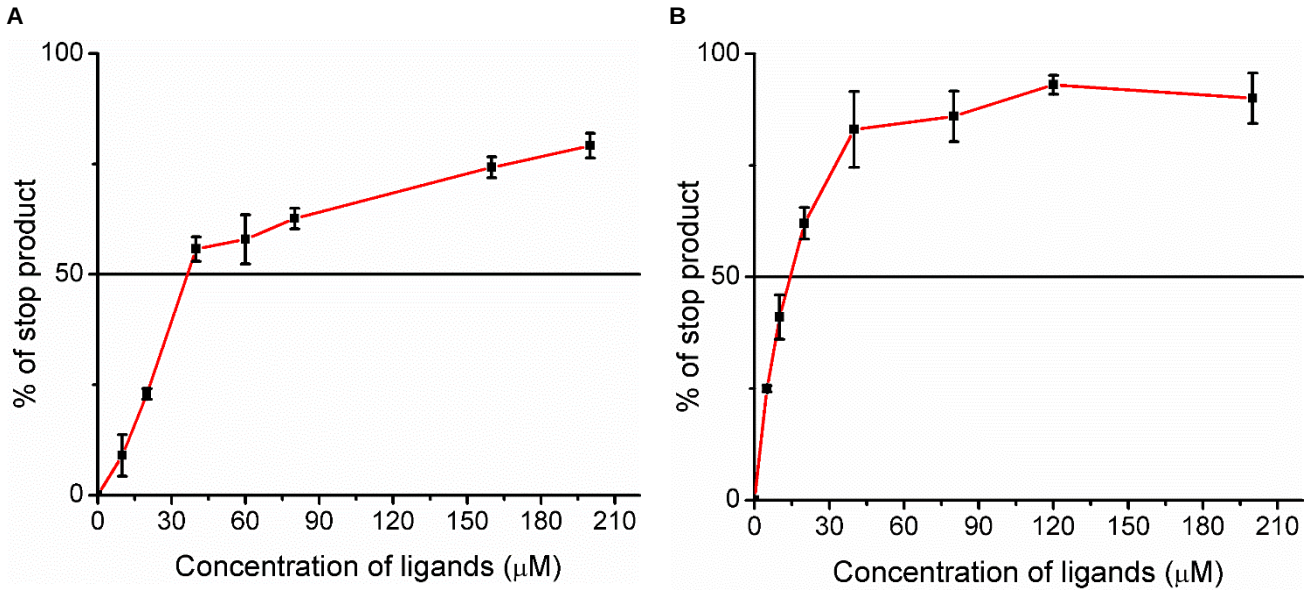


Figure S7. Plot of percentage of stop product against the increasing concentration of ligands **1** and **3** (0-200 μM) for *c-MYC* DNA template. (A) Ligand **1**; and (B) Ligand **3**. Conditions: 100 nM template DNA, 50 nM primer, 0.2 mM dNTPs, 0.5 U *Taq* polymerase enzyme in buffer (50 mM Tris, pH 7.2, 0.5 mM DTT, 0.1 mM EDTA, 5 mM $MgCl_2$, 5 mM KCl). Error bars represent the standard deviations derived from three independent experiments.

Energy optimized structure of the ligand 2

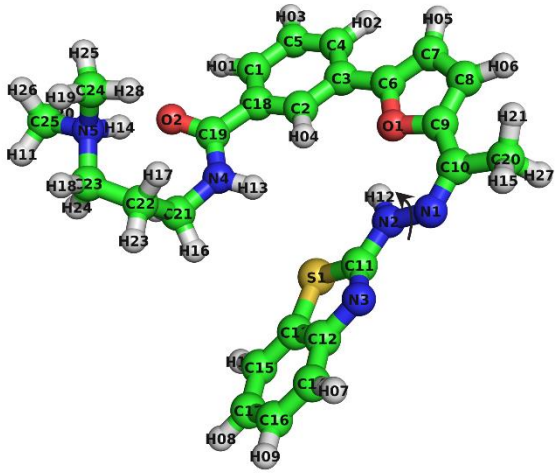


Figure S8. Energy optimized structure of ligand **2**. Black arrow represents the bond rotated prior to final step of optimization. Carbon atoms are represented in green, nitrogen atoms in blue, oxygen atoms in red, sulphur atom in yellow and hydrogen atoms in light grey. Ligand is optimized in three stages, initially at a lower theoretical level PM3 and then at a theoretical level HF/6-311+G**. Final step involved rotation of N2-N1 bond to orient the benzothiazole moiety in the plane of the rest of the molecule, and then the ligand is optimized at a theoretical level HF/6-311+G**. Two possibilities were explored one being *syn* and the other *anti* conformation. *Syn* conformation was found to be energy minimum and is shown in the above image.

Root mean square fluctuatitons of residues of *c-MYC* and *c-KIT1* DNAs

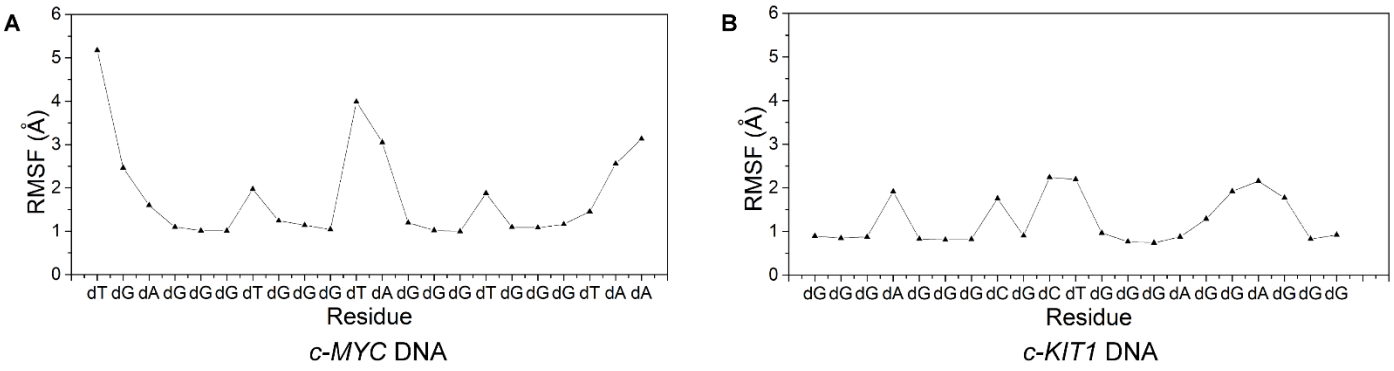


Figure S9. RMSF values of *c-MYC* and *c-KIT1* DNAs. (A) RMSF of residues in *c-MYC*; and (B) RMSF of residues in *c-KIT1*. Every 5th frame (every 10 ps) and a total of 50,000 frames were considered for calculations. The values were calculated using CPPTRAJ module of AMBER 14.

Percentage occupancies of the Hoogsteen H-bonds in *c-MYC* DNA

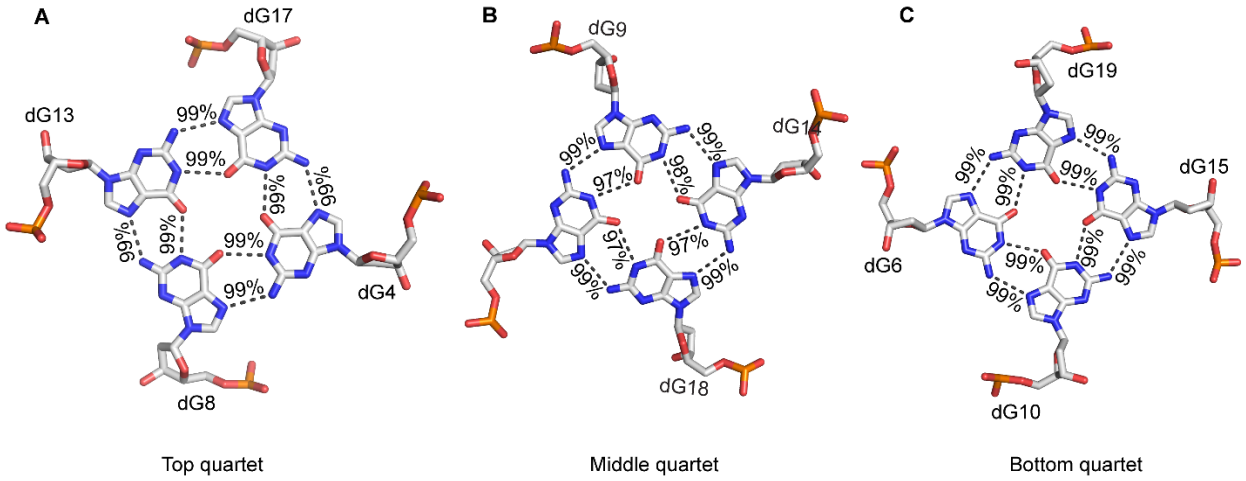


Figure S10. Percentage occupancies of the Hoogsteen H-bonds calculated for the ligand **2** bound *c-MYC* complex. The values were considered for every 10 ps (every 5th frame) and a total of 50,000 frames were considered. The values were calculated using PTRAJ module of AMBER 14.

Percentage occupancies of Hoogsteen H-bonds in the *c-KIT1* DNA

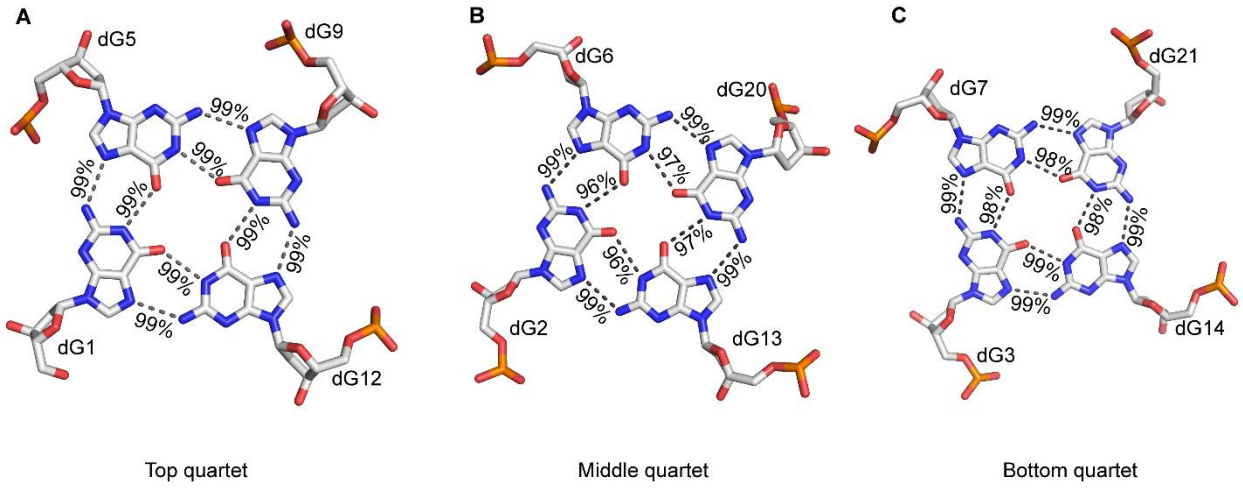


Figure S11. Percentage occupancies of the Hoogsteen H-bonds calculated for the ligand **2** bound *c-KIT1* complex. The values were considered for every 10 ps (every 5th frame) and a total of 50,000 frames were considered. The values were calculated using PTRAJ module of AMBER14.

Reorientation of the ligand during the course of dynamics with *c-MYC* DNA

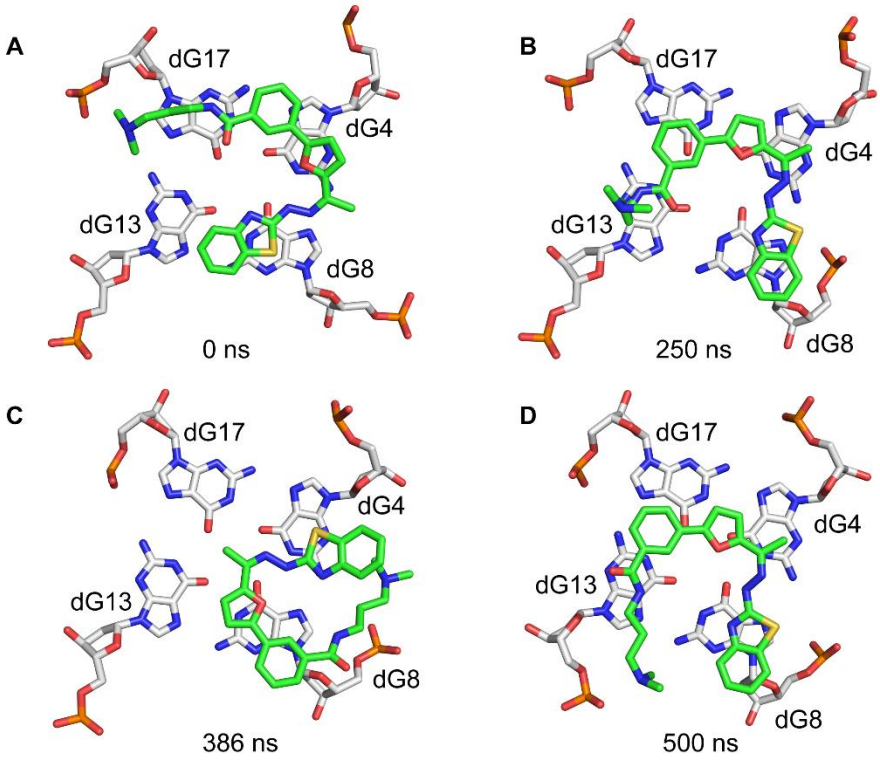


Figure S12. Top view of the ligand **2** and top G-quartet of *c-MYC* DNA showing the orientation of the ligand at various stages of MD simulations. (A) Initial frame before the production run; (B) a frame at 250 ns; (C) a frame at 386 ns; and (D) a frame at 500 ns. Ligand is highlighted in green. Black dashed lines represent the Hoogsteen H-bonds

Orientations of flanking nucleotides of *c-MYC* and *c-KIT1* DNAs with respect to the ligand 2

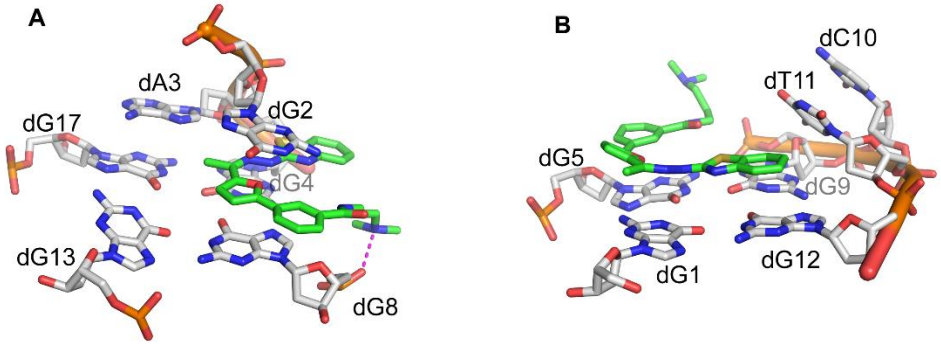


Figure S13. (A) In the *c-MYC* G-quadruplex simulations, at a time frame of 380-400 ns, the ligand **2** is sandwiched between residues dG2 and dG8, while the residue dA3 stacks over the G-quartet core; and (B) In *c-KIT1* simulations, the residues dC10 and dT11 are found to be far away from the quartet, but toward the end of the dynamics, the residue dC10 stacks over the residue dT11, while the residue dT11 tends to stack over benzothiazole ring of the ligand **2**. Electrostatic interactions are represented in dashed magenta lines. Ligand is represented in green.

Major electrostatic interactions between ligand 2 and *c-MYC* DNA

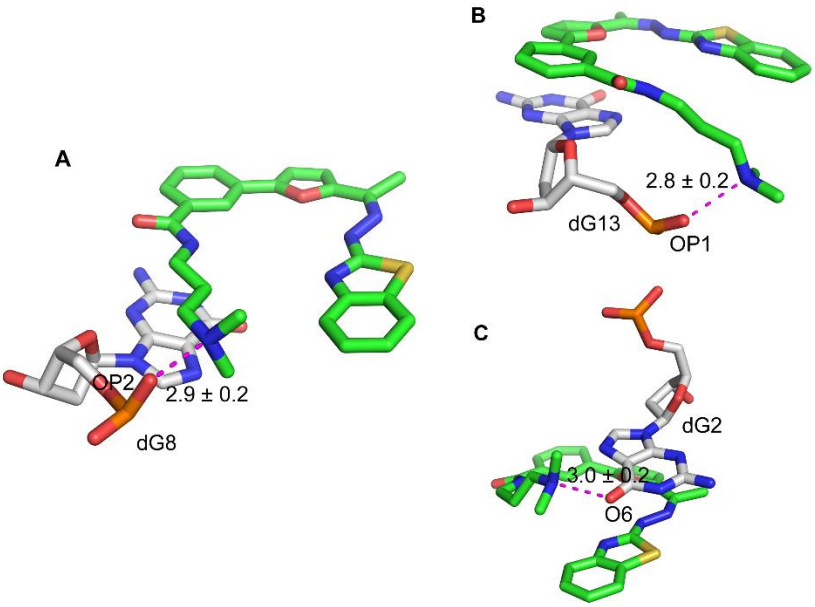


Figure S14. Electrostatic interactions between *c-MYC* DNA and the protonated nitrogen atom on the propyl side chain of the ligand **2**. (A) Dashed magenta lines represent the electrostatic interaction between OP2 of residue dG8 and the protonated nitrogen atom on the propyl side chain of the ligand; (B) Dashed magenta lines represent electrostatic interaction between OP1 of the residue dG13 and the protonated nitrogen atom on the side propyl side chain of the ligand **2**; and (C) Ion-induced dipole interaction between the protonated nitrogen atom on the propyl side chain of the ligand **2** and O6 of the residue dG2 is represented by dashed magenta lines. Ligand is represented in green colour. Average distances and standard deviations are shown in Å.

Distance plots of the major electrostatic interactions during the simulation of *c-MYC*-ligand 2 complex

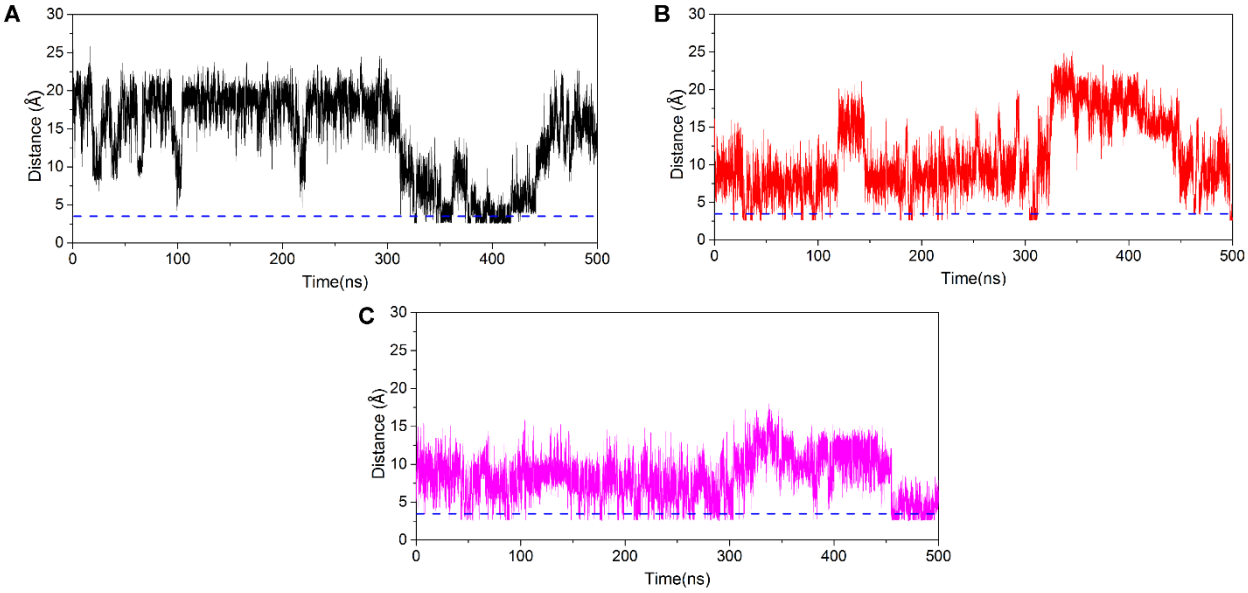


Figure S15. The distances between the atoms of *c-MYC* DNA and protonated nitrogen atom on the propyl side chain of the ligand **2** are plotted against time. Only the distance plots of major electrostatic interactions shown in **Figure S14** are presented here. (A) Black plot represents distance between OP2 of residue dG8 and protonated nitrogen atom on the propyl side chain of the ligand **2**; (B) Red plot represents distance between OP1 of residue dG13 and protonated nitrogen atom on the propyl side chain of ligand **2**; and (C) Magenta plot represents the distance between O6 of residue dG2 and protonated nitrogen atom on the propyl side chain of the ligand **2**, which corresponds to ion-induced dipole interaction. Blue dashed line represents a cut off value of 3.5 Å, which is used to define the electrostatic interactions. Overall, these interactions were observed for ~ 12-15% of simulation time. Every frame has been considered for the analysis.

Reorientation of the ligand during the course of dynamics with *c-KIT1* DNA

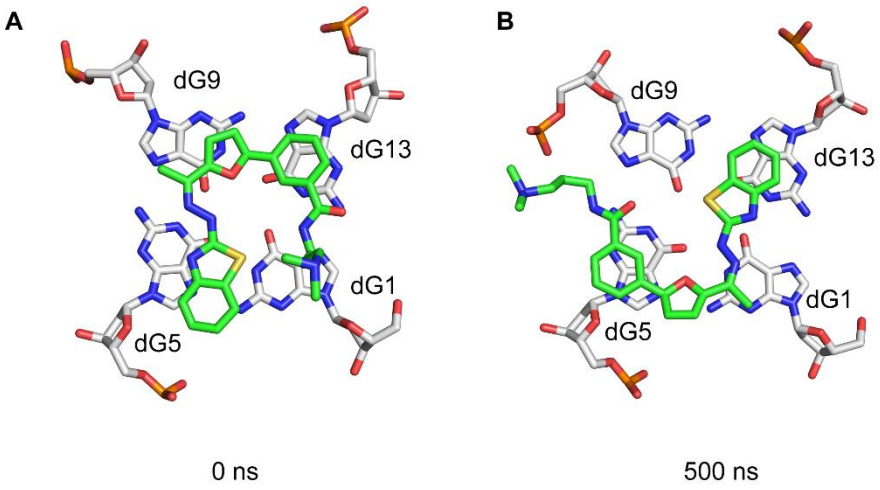


Figure S16. Top view of the ligand **2** and top G-quartet of *c-KIT1* DNA showing the orientation of the ligand over the top G-quartet. (A) Initial frame before production run; and (B) a frame at 500 ns of simulations. Ligand is highlighted in green. Black dotted lines represent the Hoogsteen H-bonds.

Major electrostatic interactions between ligand 2 and *c-KIT1* DNA

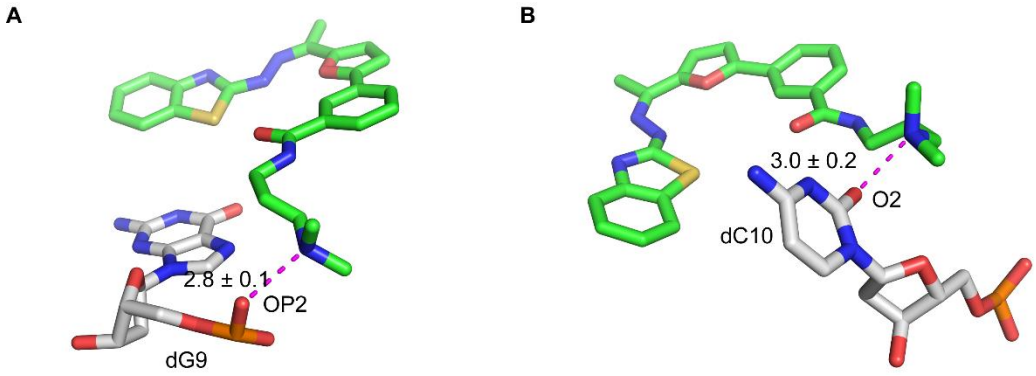


Figure S17. Electrostatic interactions between *c-KIT1* and the nitrogen atom on the propyl side chain of the ligand **2**. (A) Electrostatic interaction between ligand and OP2 atom of residue dG9 represented by dashed magenta lines; and (B) Induced dipole-ion interaction between O2 of loop residue dC10 and ligand represented by dashed magenta line. Ligand is represented in green colour. Average distances and standard deviations are denoted in Å.

Distance plots of major electrostatic interactions during the simulation of *c-KIT1*-ligand 2 complex

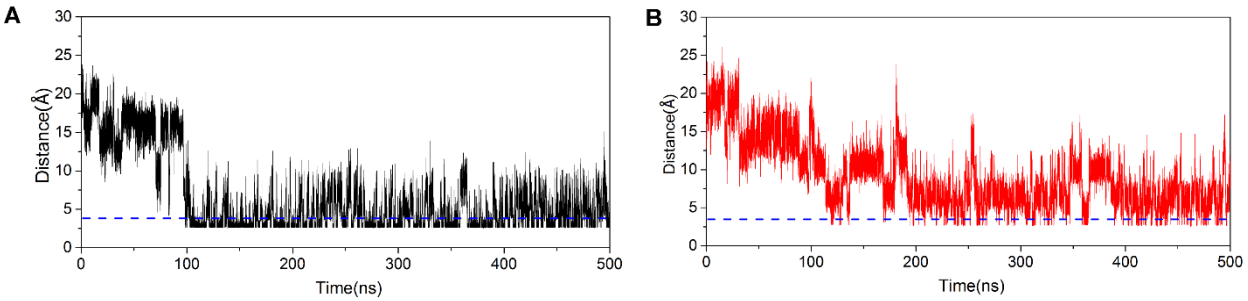


Figure S18. The distances between the atoms of *c-KIT1* G-quadruplex and protonated nitrogen atom on the propyl side chain of the ligand **2** are plotted against time. Only the distance plots of major electrostatic interactions shown **Figure S17** are presented here. (A) Black plot represents the distance between OP2 of residue dG9 and protonated nitrogen atom on the propyl side chain of the ligand; and (B) Red plot represents the distance between O2 of loop residue dC10 and protonated nitrogen atom on the propyl side chain of the ligand **2**. Dashed blue line represents a cut off value of 3.5 Å, which is used to define the electrostatic interactions. Overall, these interactions were observed for ~ 50 % of simulation time.

Stacking distances and angles calculated for complexes of *c-MYC* and the *c-KIT1* with ligand 2

	c-MYC			***c-KIT1***		
	BnT ring	**Furan ring**	**Bn ring**	**BnT ring**	**Furan ring**	**Bn ring**
Average distance	3.9 ± 0.4	4.3 ± 0.4	4.1 ± 0.4	3.8 ± 0.3	-	4.1 ± 0 .4
Average angle	9.4 ± 4.9	11.0 ± 4.9	9.3 ± 4.7	8.2 ± 4.5	-	8.5 ± 4.4
% of simulation time	~ 79	~ 60	~ 69	~ 83	-	~ 76

Table S1. Average stacking distances and angles of benzothiazole ring (BnT), furan ring and benzene (Bn) rings of the ligand **2** with the top G-quartet (5'end) of *c-MYC* and *c-KIT1* DNAs. Distances and their standard deviations are mentioned in Å. Angles and their standard deviations are mentioned in degrees. Cut-off values of 5 Å and 20° were considered for distances and angles respectively. Furan ring is not considered in *c-KIT1* because this ring doesn't stack on any of the quartet residues. Every 25th frame and a total of 10,000 frames were considered for this calculation. UCSF Chimera was used for calculating distances and angles.

Binding free energy components of *c-MYC* and *c-KIT1* DNAs and ligand 2

	***c-MYC* (PDB entry:2L7V)**	***c-KIT1* (PDB entry: 2O3M)**
ΔE_{ELEC}	−441.35 ± 23.78	−411.30 ± 27.29
ΔE_{VDW}	−63.33 ± 3.05	−46.63 ± 2.87
$\Delta E_{MM}(\Delta E_{ELEC}+\Delta E_{VDW})$	−504.69 ± 25.31	−457.90 ± 28.33
ΔPB_{np}	−4.44 ± 0.19	−3.49 ± .0.18
ΔPB_{cal}	460.33 ± 22.89	425.82 ± 26.57
$\Delta PB_{solv}(\Delta PB_{np} + \Delta PB_{cal})$	455.90 ± 22.79	422.32 ± 26.43
$\boldsymbol{\Delta H_{PB}(\Delta E_{MM} + \Delta PB_{solv})}$	**−48.79 ± 4.20**	**−35.58± 4.07**
ΔS_{TRANS}	−13.19 ± 0.00	-13.13 ± 0.00
ΔS_{ROTA}	−11.14 ± 0.04	-11.39 ± 0.04
ΔS_{VIBR}	6.66 ± 5.09	6.93 ± 4.35
$\boldsymbol{T\Delta S}$	**−17.61 ± 5.09**	**-17.55 ± 4.35**
$\boldsymbol{\Delta G(\Delta H_{PB} - T\Delta S)}$	**−31.18 ± 6.74**	**-18.02 ± 6.00**

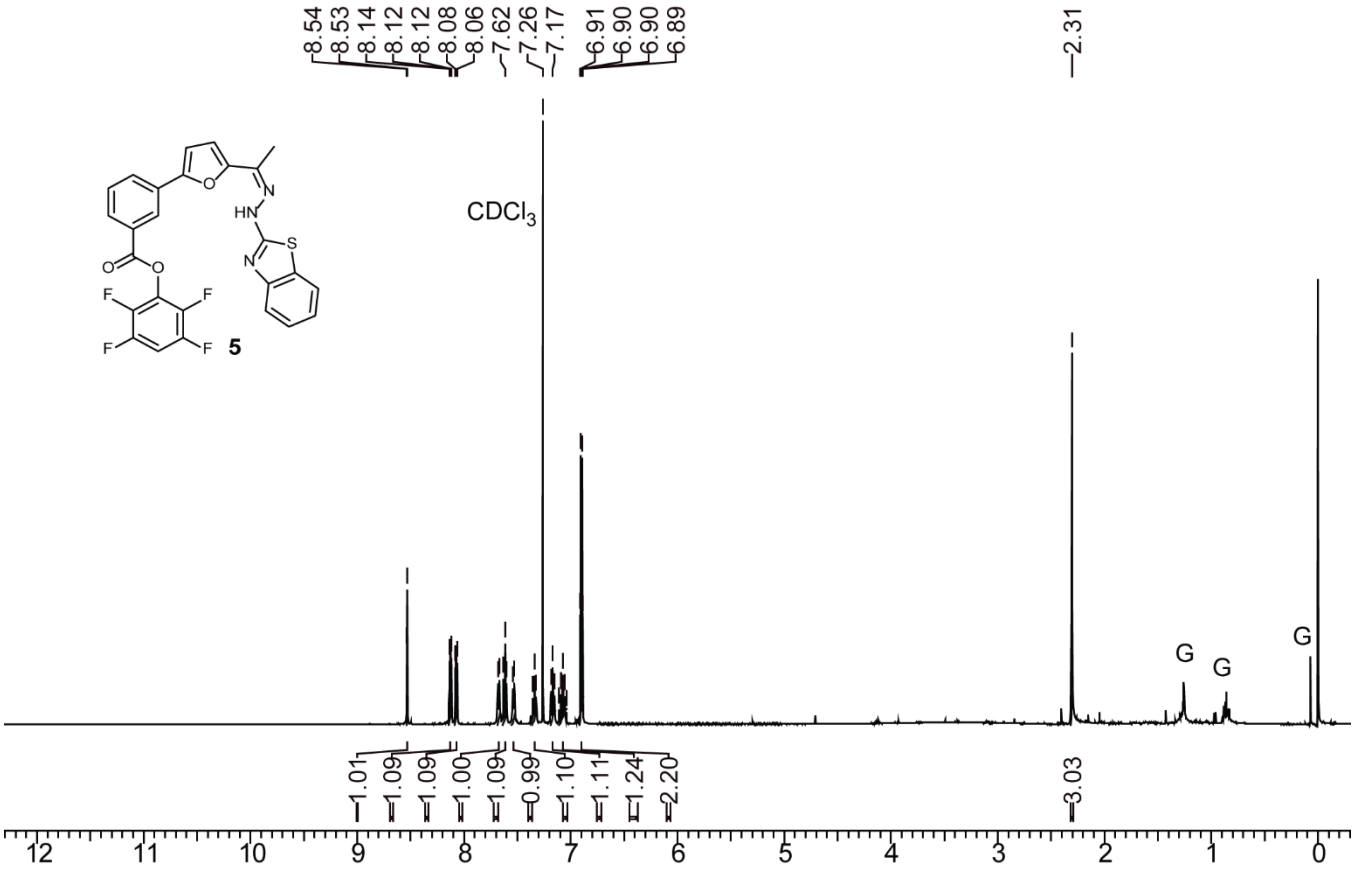
Table S2. Binding free energy components of *c-MYC* and *c-KIT1* G-quadruplex DNAs with the ligand **2** calculated from last 20 ns of 500ns simulations. Every 5th frame and a total of 2000 frames were considered for the calculations. The molecular-mechanical energy calculations are performed using MM/PBSA, and entropy calculations using nmode analysis. ΔE_{ELEC} is the electrostatic contribution. ΔE_{VDW} is the van der Waals contribution. ΔE_{MM} is the total molecular-mechanical energy. ΔPB_{np} is the non-polar contribution to the solvation energy. ΔPB_{cal} is the electrostatic contribution to the solvation energy. ΔPB_{solv} is the total solvation energy. $T\Delta S$ is the solute entropic contribution, where ΔS is the sum of translational, rotational and vibrational entropies. $\Delta G(\Delta H_{PB} - T\Delta S)$ is the estimated binding free energy. All the values are reported in kcal mol^{-1}. For nmode analysis parameters used are: drms = 0.5, dielec (distance dependent dielec) = 4, maxcyc = 10,000 and AMBER prescribed default values are used for PB calculations

DNA sequences used for various experiments

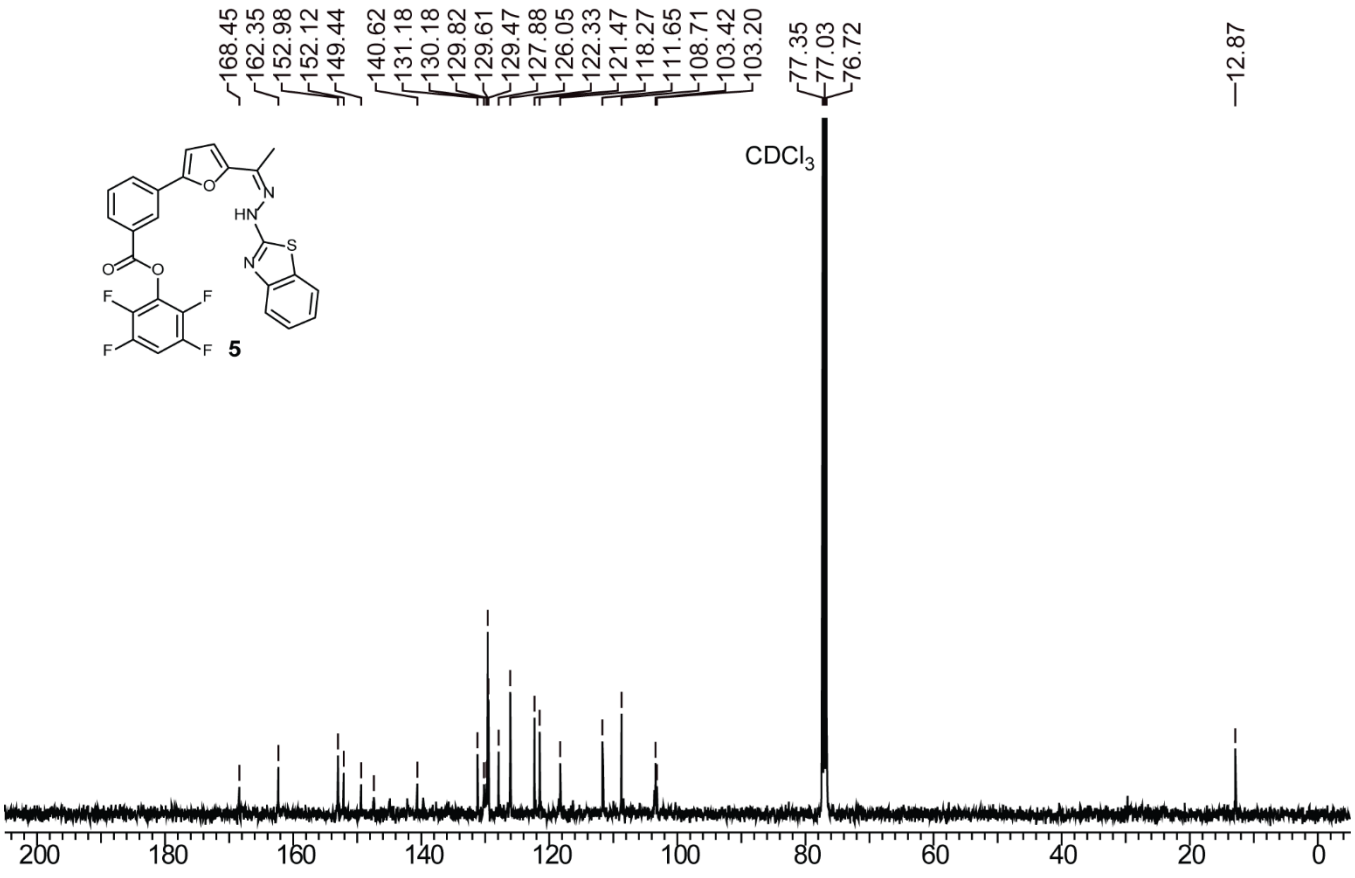
DNA	Sequence
Telomeric DNA	5'-AGGGTTAGGGTTAGGGTTAGGG-3'
c-MYC DNA	5'- TGAGGGTGGGTAGGGTGGGTAA-3'
c-KIT1 DNA	5'-GGGAGGGCGCTGGGAGGAGGG-3'
h-RAS1 DNA	5'-TCGGGTTGCGGGCGCAGGGCACGGGCG -3'
Duplex DNA	5'-CCAGTTCGTAGTAACCC-3'
	5'-GGGTTACTACGAACTGG-3' (complementary strand)
Primer for stop assay	5'-ACGACTCACTATAGCAATTGCG-3’
Template of *c-MYC* DNA	5'-TGAGGGTGGGTAGGGTGGGTAAGCCACCGCAATT GCTATAGTGAGTCGT-3'
Template of mutated *c-MYC* DNA	5'-TGAGGGTGGGTAGAGTGGGTAAGCCACCGCAATT GCTATAGTGAGTCGT-3'
Template of telomeric DNA	5'-AGGGTTAGGGTTAGGGTTAGGGGCCACCGCAATT GCTATAGTGAGTCGT-3'

Table S3. DNA sequences used for the experiments in 5' to 3' direction

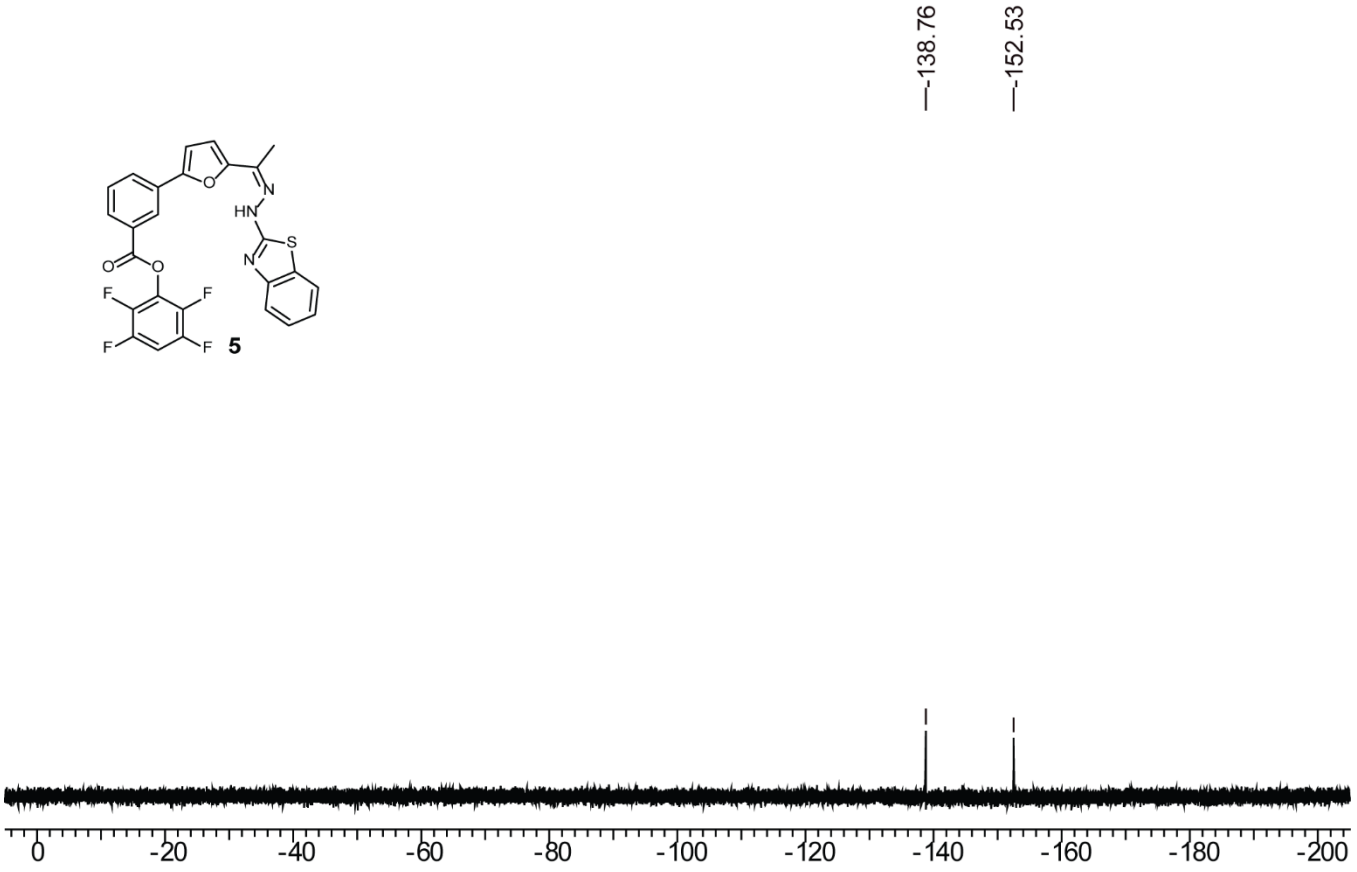
^{1}H NMR spectrum of compound 5 (G: Grease, Im: Impurity)



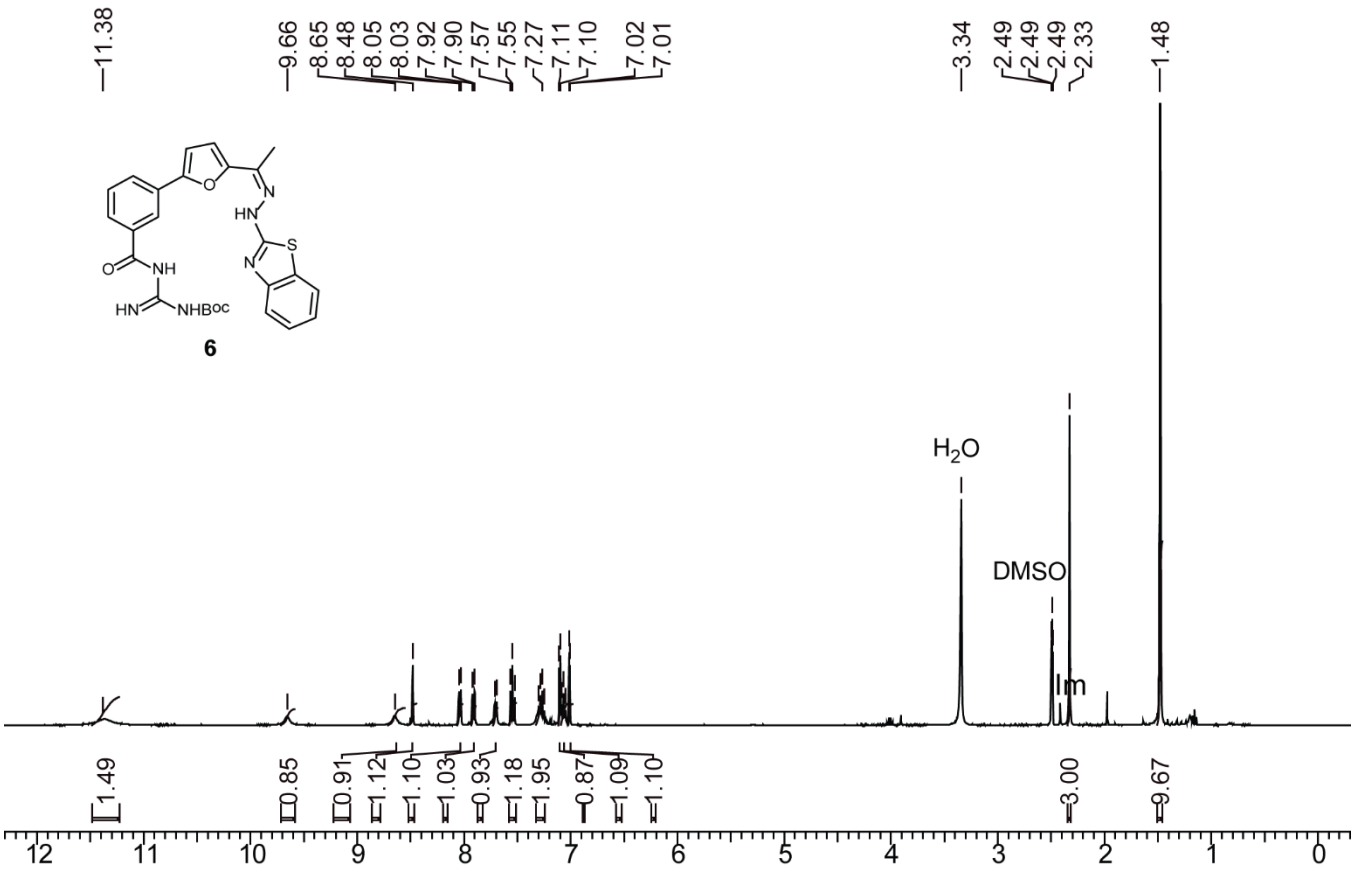
^{13}C NMR spectrum of compound 5



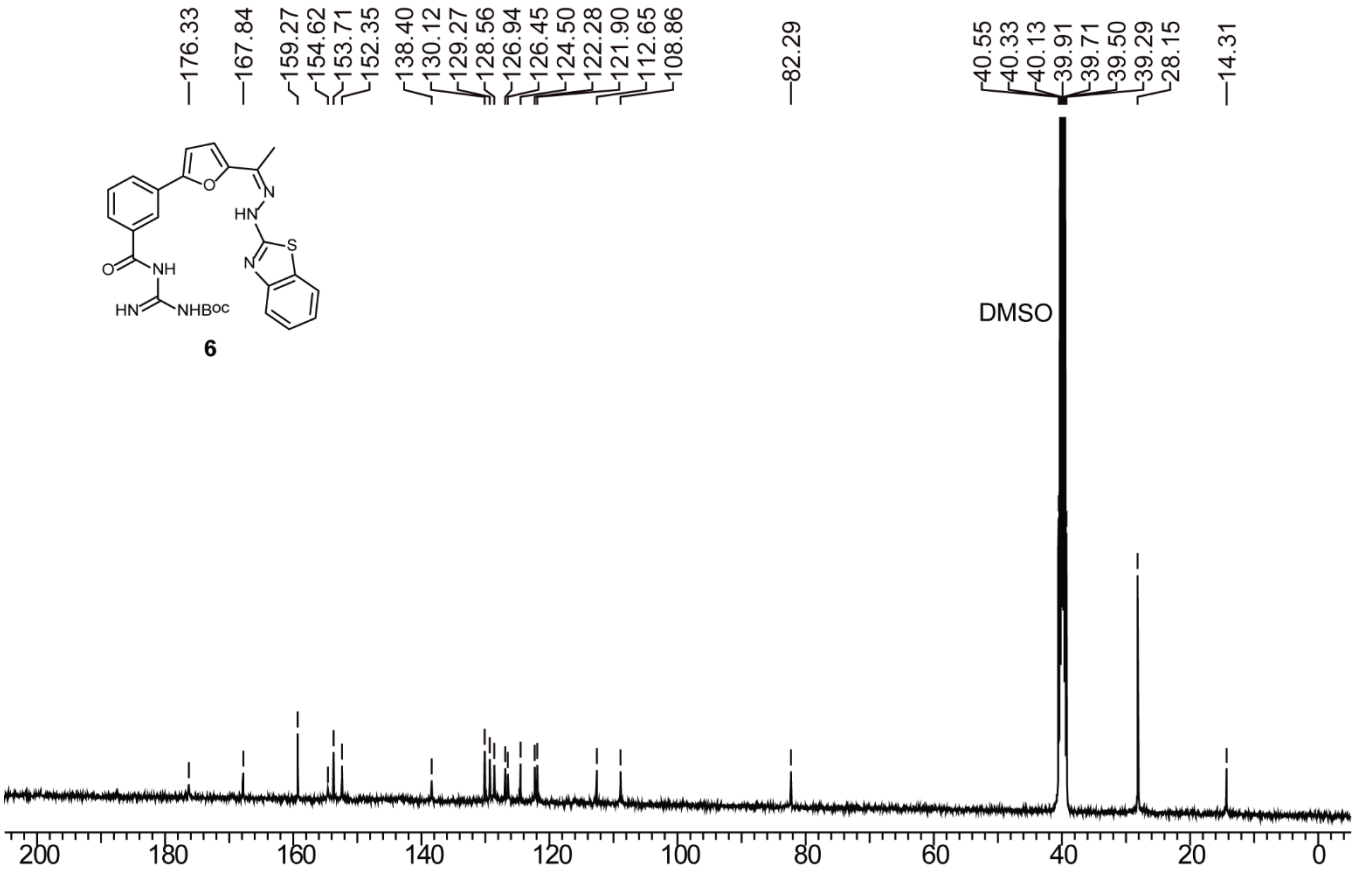
^{19}F NMR spectrum of compound 5



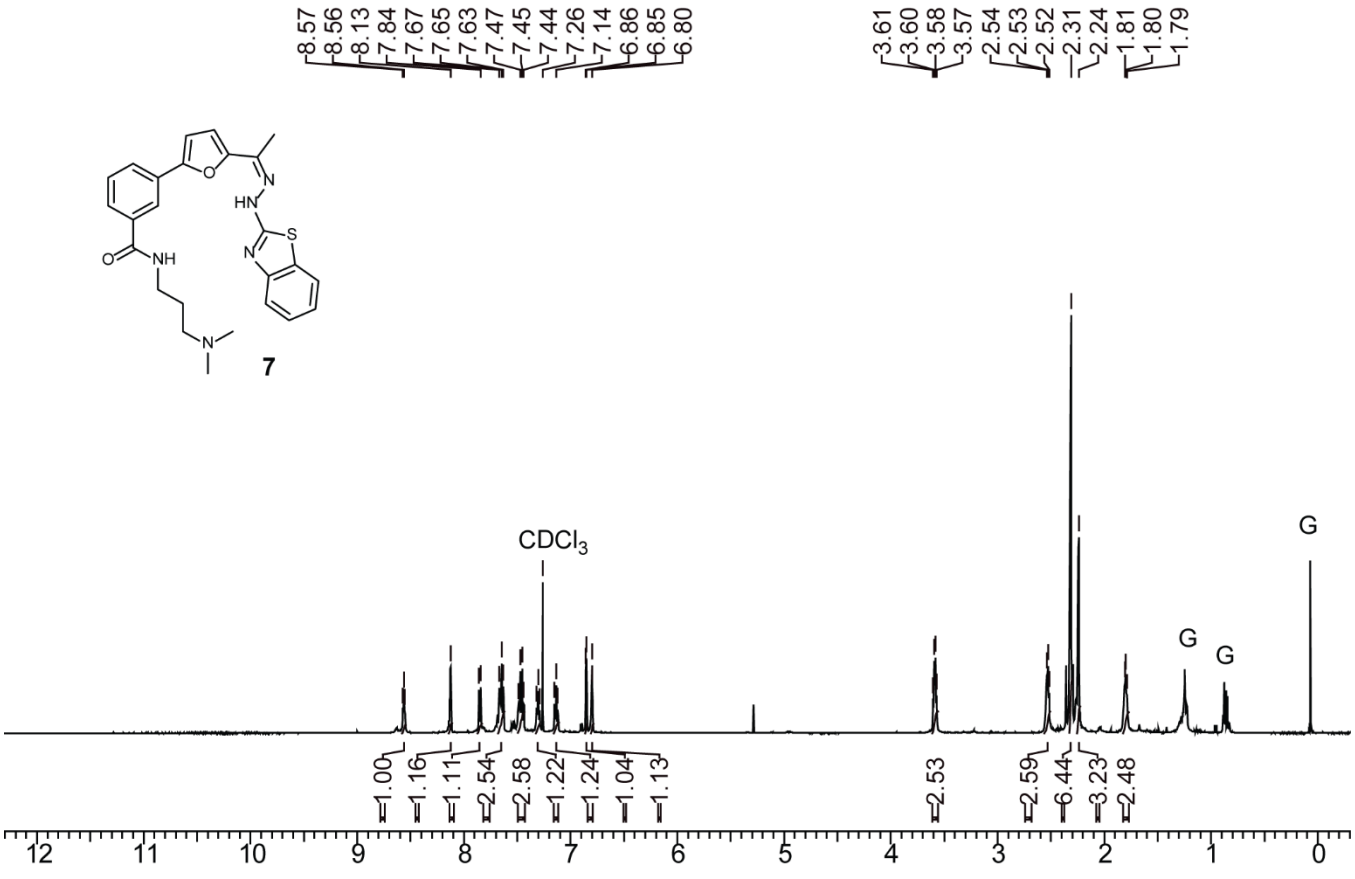
^{1}H NMR spectrum of compound 6



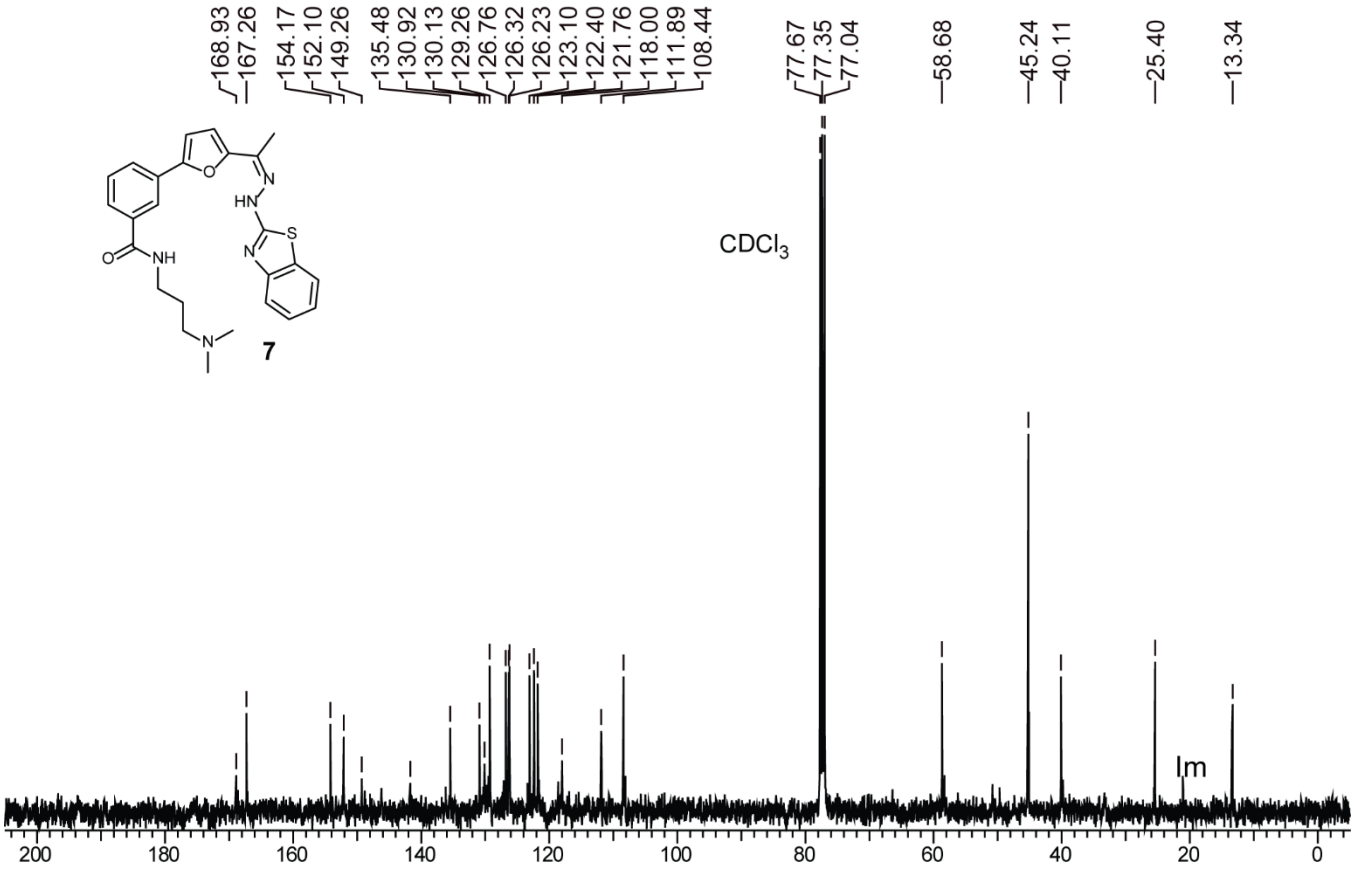
^{13}C NMR spectrum of compound 6



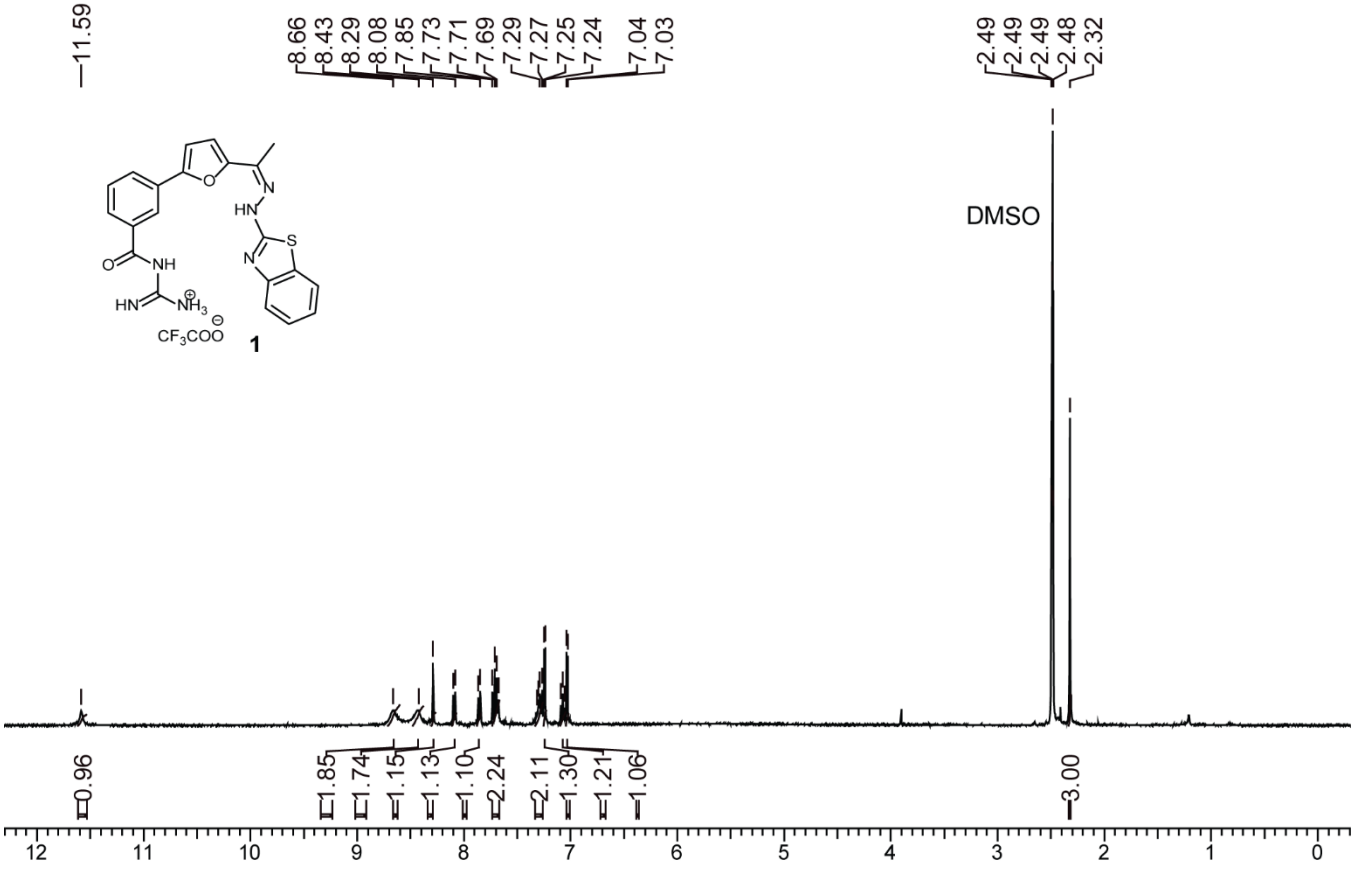
^{1}H NMR spectrum of compound 7



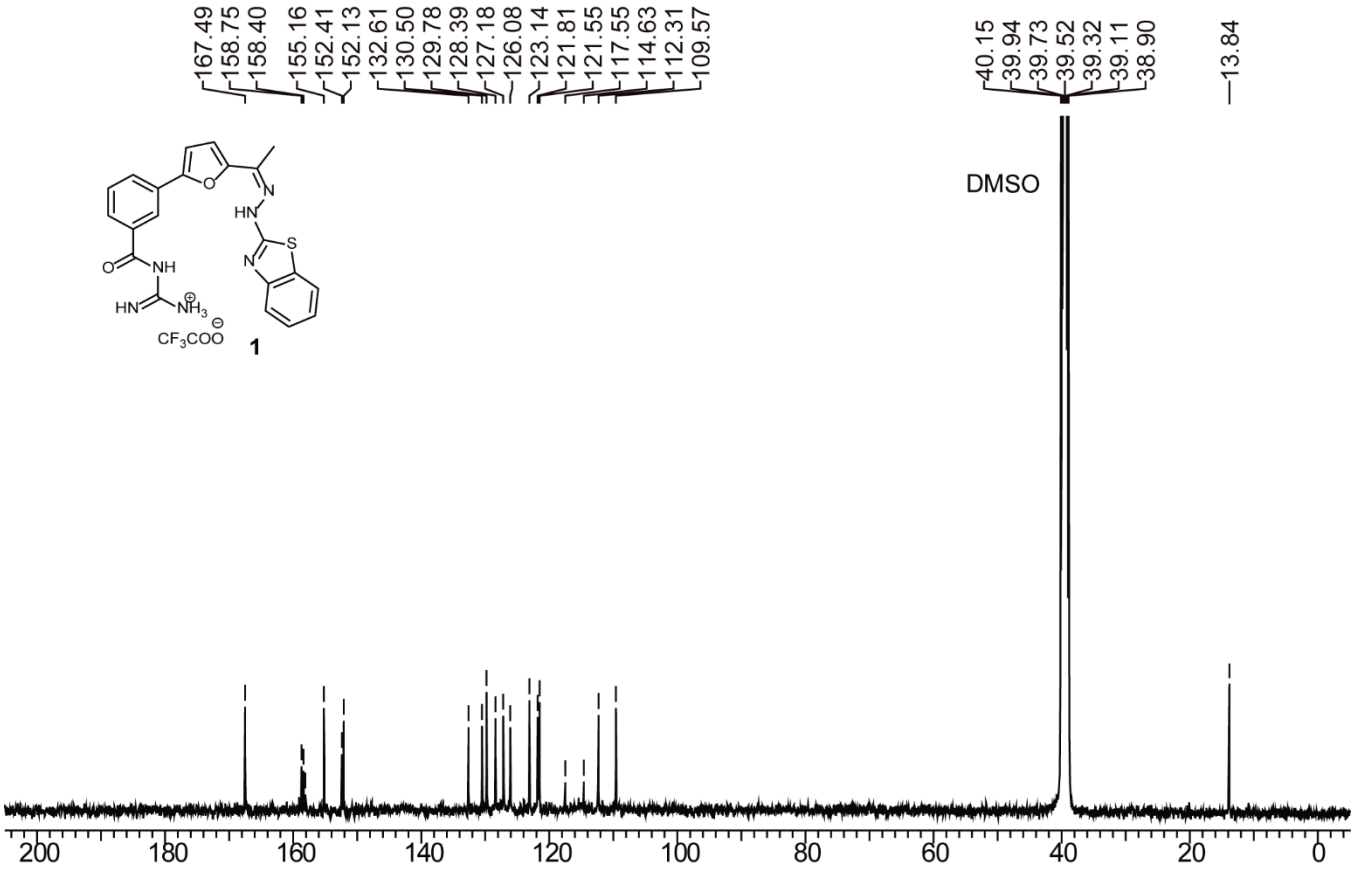
^{13}C NMR spectrum of compound 7



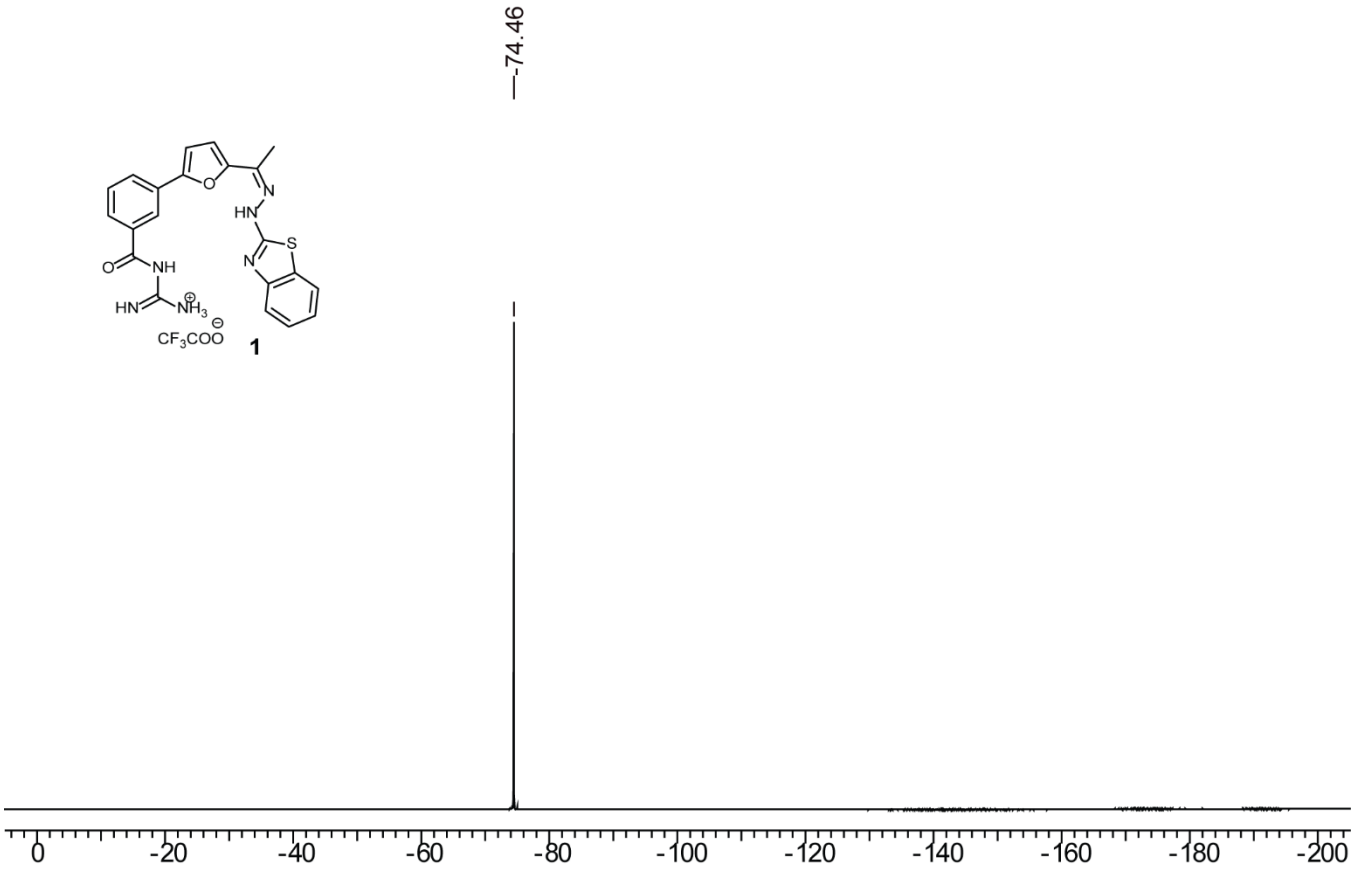
^{1}H NMR spectrum of ligand 1



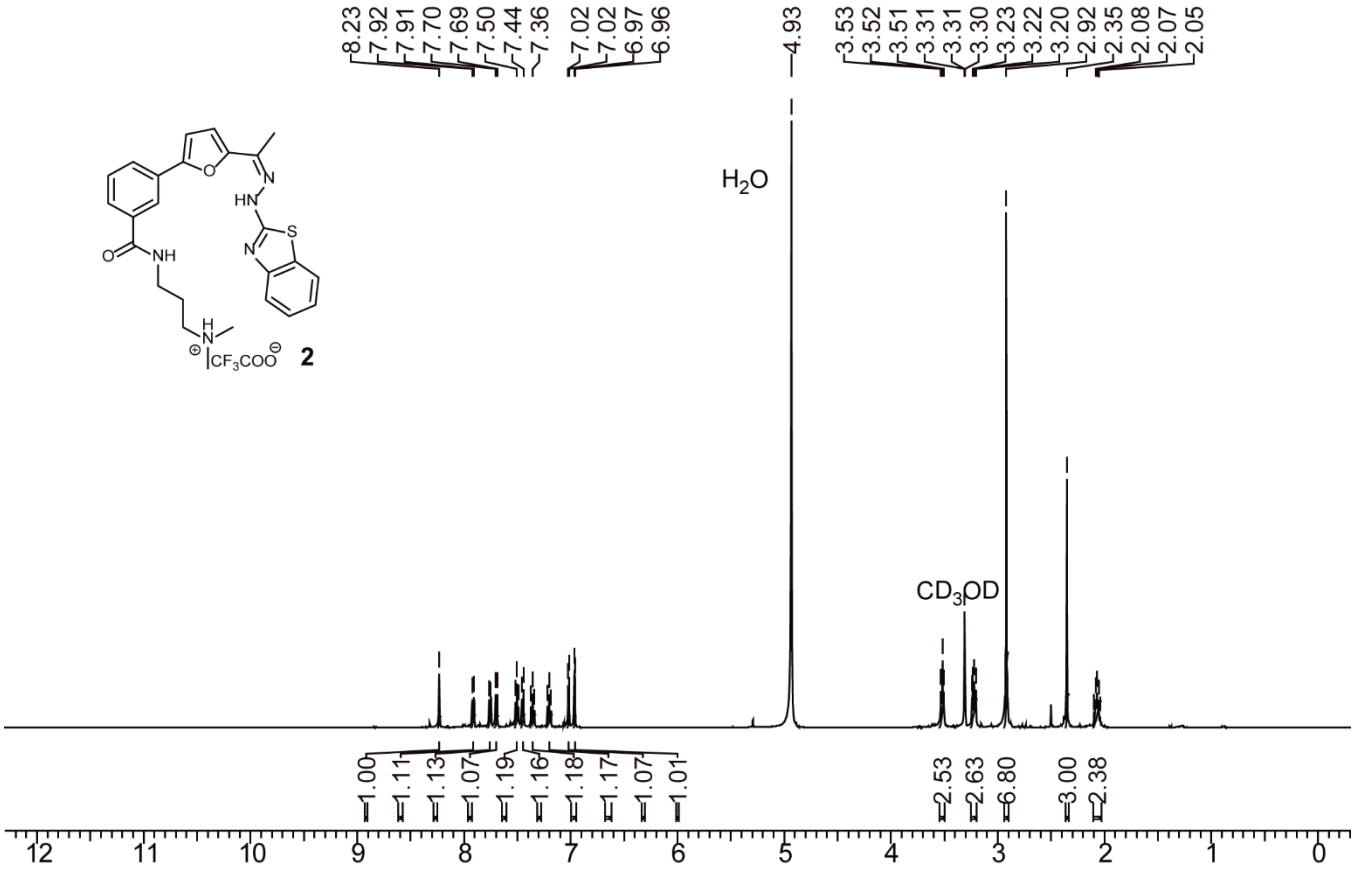
^{13}C NMR spectrum of ligand 1



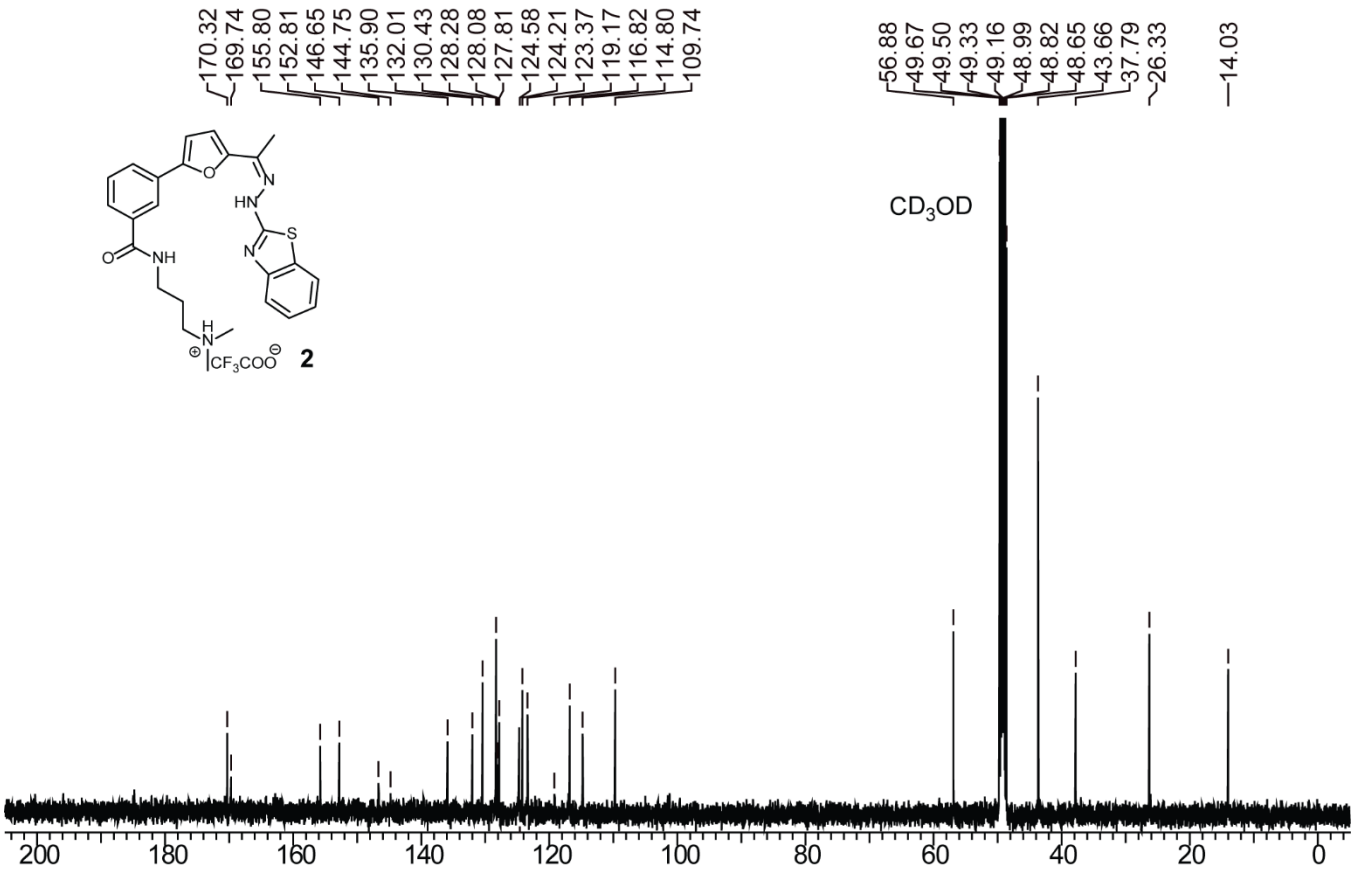
^{19}F NMR spectrum of ligand 1



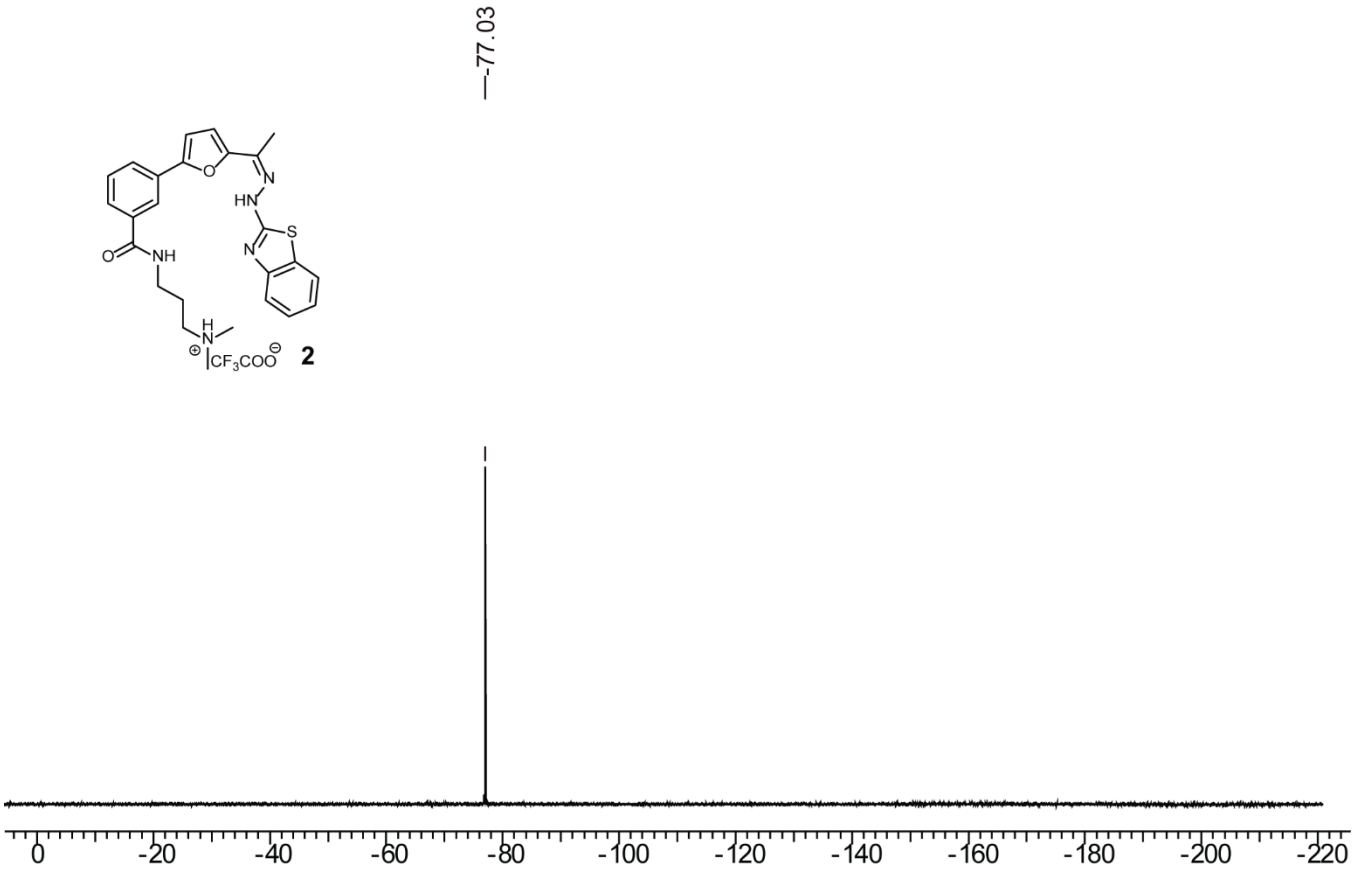
^{1}H NMR spectrum of ligand 2



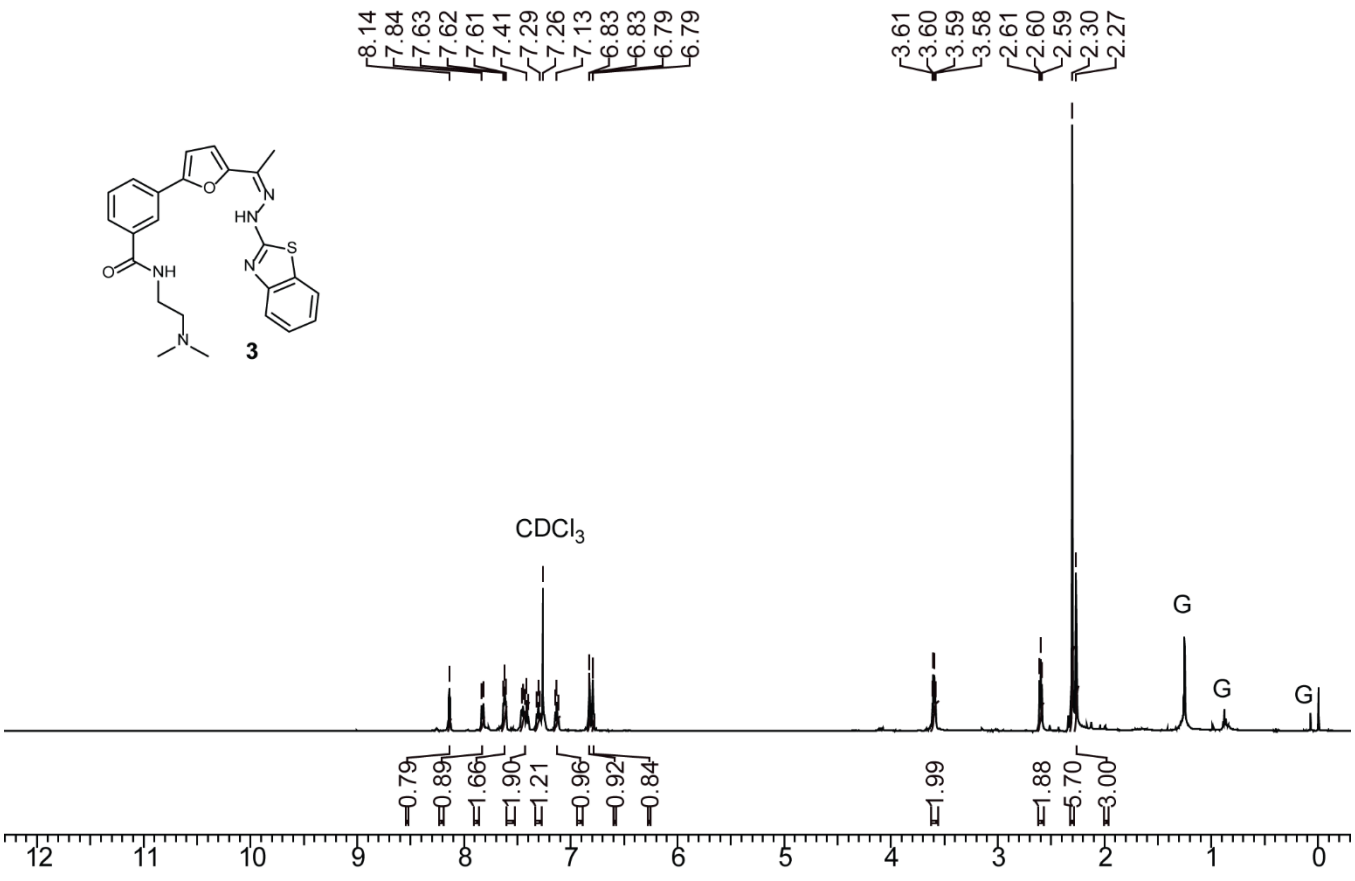
^{13}C NMR spectrum of ligand 2



^{19}F NMR spectrum of ligand 2



^{1}H NMR spectrum of ligand 3



^{13}C NMR spectrum of ligand 3