Electronic Supplementary Material (ESI) for Organic & Biomolecular Chemistry. This journal is © The Royal Society of Chemistry 2016

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

Figure S1	CD spectra of ligands with telomeric and <i>c-MYC</i> DNAs in the absence of added metal cations
Figure S2	CD spectra of ligands with <i>c-KIT1</i> and <i>h-RAS1</i> DNA in the absence of added metal cations
Figure S3	Normalised CD melting curves of telomeric, c-KIT1, h-RAS1 and duplex DNAsPage S3
Figure S4	DC ₅₀ plots for <i>c-KIT1</i> , telomeric and duplex DNAs from the FID assayPage S4
Figure S5	Fluorimetric titration curves of ligand 2 complexed with telomeric and duplex DNAs Page S5
Figure S6	Denaturing PAGE of Taq DNA polymerase stop assayPage S6
Figure S7	IC ₅₀ plots of <i>Taq</i> DNA polymerase stop assay for ligands with <i>c-MYC</i> DNAPage S7
Figure S8	Energy optimized structure of the ligand 2Page S8
Figure S9	Root mean square fluctuations of residues of <i>c-MYC</i> and <i>c-KIT1</i> DNAsPage S8
Figure S10	Percentage occupancies of the Hoogsteen H-bonds in <i>c-MYC</i> DNAPage S9
Figure S11	Percentage occupancies of the Hoogsteen H-bonds in c-KIT1 DNAPage S9
Figure S12	Reorientation of the ligand during the course of dynamics with <i>c-MYC</i> DNAPage S10
Figure S13	Orientations of the flanking nucleotides of <i>c-MYC</i> and <i>c-KIT1</i> DNAs with respect to the ligand 2
Figure S14	Major electrostatic interactions between ligand 2 and <i>c</i> - <i>MYC</i> DNA Page S11
Figure S15	Distance plots of major electrostatic interactions during the simulation of <i>c-MYC</i> -ligand 2 complexPage S12
Figure S16	Reorientation of the ligand during the course of dynamics with <i>c-KIT1</i> DNA Page S12
Figure S17	Major electrostatic interactions between ligand 2 and <i>c-KIT1</i> DNA Page S13
Figure S18	Distance plots of the major electrostatic interactions during the simulation of <i>c-KIT1</i> -ligand 2 complex
Table S1	Stacking distances and angles calculated for the complexes of <i>c</i> - <i>MYC</i> and the <i>c</i> - <i>KIT1</i> with ligand 2

Table S2	Binding free energy components of <i>c-MYC</i> and <i>c-KIT1</i> DNAs and ligand 2	Page S14
Table S3	DNA sequences used for various experiments	Page S15
¹ H NMR sj	pectrum of compound 5	Page S16
¹³ C NMR s	spectrum of compound 5	Page S16
¹⁹ F NMR s	spectrum of compound 5	Page S17
¹ H NMR s	spectrum of compound 6	Page S17
¹³ C NMR s	spectrum of compound 6	Page S18
¹ H NMR sp	pectrum of compound 7	Page S18
¹³ C NMR s	spectrum of compound 7	Page S19
¹ H NMR sp	pectrum of ligand 1	Page S19
¹³ C NMR s	spectrum of ligand 1	Page S20
¹⁹ F NMR s	spectrum of ligand 1	Page S20
¹ H NMR s	pectrum of ligand 2	Page S21
¹³ C NMR s	spectrum of ligand 2	Page S21
¹⁹ F NMR s	spectrum of ligand 2	Page S22
¹ H NMR sj	pectrum of ligand 3	Page S22
¹³ C NMR s	spectrum of ligand 3	Page S23



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 μ M 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 μ M 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 μ M) and duplex DNA (15 μ M) 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);



DC50 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 μ M in 100 mM KCl, 10 mM lithium cacodylate buffer, pH 7.2). (A) *c-KIT1* DNA (against 0-14 μ M ligand concentration); (B) Telomeric DNA (against 0-30 μ M ligand concentration); and (C) Duplex DNA (against 0-30 μ M 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 μ M in 100 mM KCl and 10 mM lithium cacodylate buffer, pH 7.2 and 0-50 μ M 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 telomeric quadruplex DNA against the logarithm of increasing concentration of ligand **2**. The Hill1 equation was used for the curve fitting.

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



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 μ M). (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 MgCl₂, 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.

IC50 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 μ M) 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 MgCl₂, 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 5^{th} 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 5^{th} 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.

	c-MYC		c-KIT1			
	BnT	Furan ring	Bn ring	BnT ring	Furan ring	Bn ring
	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

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

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.

	<i>c-MYC</i> (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	$-3.49\pm.0.18$
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
$\Delta H_{\rm PB}(\Delta E_{\rm MM} + \Delta PB_{\rm solv})$	-48.79 ± 4.20	-35.58 ± 4.07
$\Delta S_{\mathrm{TRANS}}$	-13.19 ± 0.00	-13.13 ± 0.00
$\Delta S_{ m ROTA}$	-11.14 ± 0.04	-11.39 ± 0.04
$\Delta S_{ m VIBR}$	6.66 ± 5.09	6.93 ± 4.35
$\mathbf{T}\Delta S$	-17.61 ± 5.09	-17.55 ± 4.35
$\Delta \boldsymbol{G} (\Delta \boldsymbol{H}_{\rm PB} - \mathbf{T} \Delta \boldsymbol{S})$	-31.18 ± 6.74	$\textbf{-18.02} \pm \textbf{6.00}$

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

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

DNA sequences used for various experiments

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

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



¹³C NMR spectrum of compound 5



¹⁹F NMR spectrum of compound 5





¹H NMR spectrum of compound 6



¹³C NMR spectrum of compound 6



¹³C NMR spectrum of compound 7



¹H NMR spectrum of ligand 1



¹³C NMR spectrum of ligand 1



¹⁹F NMR spectrum of ligand 1



---74.46

¹H NMR spectrum of ligand 2



¹³C NMR spectrum of ligand 2



¹⁹F NMR spectrum of ligand 2



¹H NMR spectrum of ligand 3



¹³C NMR spectrum of ligand 3

