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 c-MYC DNAs in the absence of added metal cations .......................................................... Page S1
Figure S2 CD spectra of ligands with c-KIT1 and h-RAS1 DNA in the absence of added metal cations .......................................................... Page S2
Figure S3 Normalised CD melting curves of telomeric, c-KIT1, h-RAS1 and duplex DNAs........ Page S3
Figure S4 DC50 plots for c-KIT1, telomeric and duplex DNAs from the FID assay................. Page 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 assay ........................................ Page S6
Figure S7 IC50 plots of Taq DNA polymerase stop assay for ligands with c-MYC DNA......... Page S7
Figure S8 Energy optimized structure of the ligand 2....................................................... Page S8
Figure S9 Root mean square fluctuations of residues of c-MYC and c-KIT1 DNAs .............. Page S8
Figure S10 Percentage occupancies of the Hoogsteen H-bonds in c-MYC DNA............... Page S9
Figure S11 Percentage occupancies of the Hoogsteen H-bonds in c-KIT1 DNA............. Page S9
Figure S12 Reorientation of the ligand during the course of dynamics with c-MYC DNA..... Page S10
Figure S13 Orientations of the flanking nucleotides of c-MYC and c-KIT1 DNAs with respect to the ligand 2 .................................................................. Page S10
Figure S14 Major electrostatic interactions between ligand 2 and c-MYC DNA............. Page S11
Figure S15 Distance plots of major electrostatic interactions during the simulation of c-MYC-ligand 2 complex ................................................................. Page S12
Figure S16 Reorientation of the ligand during the course of dynamics with c-KIT1 DNA..... Page S12
Figure S17 Major electrostatic interactions between ligand 2 and c-KIT1 DNA............. Page S13
Figure S18 Distance plots of the major electrostatic interactions during the simulation of c-KIT1-ligand 2 complex ................................................................ Page S13
Table S1 Stacking distances and angles calculated for the complexes of c-MYC and the c-KIT1 with ligand 2 .................................................................. Page S14
Table S2  Binding free energy components of c-MYC and c-KITI DNAs and ligand 2 ............ Page S14
Table S3  DNA sequences used for various experiments ......................................................... Page S15
1H NMR spectrum of compound 5 .......................................................... Page S16
13C NMR spectrum of compound 5 .......................................................... Page S16
19F NMR spectrum of compound 5 .......................................................... Page S17
1H NMR spectrum of compound 6 .......................................................... Page S17
13C NMR spectrum of compound 6 .......................................................... Page S18
1H NMR spectrum of compound 7 .......................................................... Page S18
13C NMR spectrum of compound 7 .......................................................... Page S19
1H NMR spectrum of ligand 1 .............................................................. Page S19
13C NMR spectrum of ligand 1 .............................................................. Page S20
19F NMR spectrum of ligand 1 .............................................................. Page S20
1H NMR spectrum of ligand 2 .............................................................. Page S21
13C NMR spectrum of ligand 2 .............................................................. Page S21
19F NMR spectrum of ligand 2 .............................................................. Page S22
1H NMR spectrum of ligand 3 .............................................................. Page S22
13C NMR 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 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 \(c\text{-}KIT1\) and \(h\text{-}RAS1\) DNA in the absence of added metal cations

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

<table>
<thead>
<tr>
<th>c-MYC</th>
<th>c-KIT1</th>
</tr>
</thead>
<tbody>
<tr>
<td>BnT ring</td>
<td>Furan ring</td>
</tr>
<tr>
<td>Average distance</td>
<td>3.9 ± 0.4</td>
</tr>
<tr>
<td>Average angle</td>
<td>9.4 ± 4.9</td>
</tr>
<tr>
<td>% of simulation time</td>
<td>~ 79</td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th>c-MYC (PDB entry:2L7V)</th>
<th>c-KIT1 (PDB entry: 2O3M)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ΔE_ELEC</td>
<td>-441.35 ± 23.78</td>
</tr>
<tr>
<td>ΔE_VDW</td>
<td>-63.33 ± 3.05</td>
</tr>
<tr>
<td>ΔE_MM(ΔE_ELEC+ΔE_VDW)</td>
<td>-504.69 ± 25.31</td>
</tr>
<tr>
<td>ΔPBNP</td>
<td>-4.44 ± 0.19</td>
</tr>
<tr>
<td>ΔPB_cal</td>
<td>460.33 ± 22.89</td>
</tr>
<tr>
<td>ΔPB_solv(ΔPBNP + ΔPB_cal)</td>
<td>455.90 ± 22.79</td>
</tr>
<tr>
<td>ΔH_PB(ΔE_MM + ΔPB_solv)</td>
<td>-48.79 ± 4.20</td>
</tr>
<tr>
<td>ΔS_TRANS</td>
<td>-13.19 ± 0.00</td>
</tr>
<tr>
<td>ΔS_ROTA</td>
<td>-11.14 ± 0.04</td>
</tr>
<tr>
<td>ΔS_VIBR</td>
<td>6.66 ± 5.09</td>
</tr>
<tr>
<td>TΔS</td>
<td>-17.61 ± 5.09</td>
</tr>
<tr>
<td>ΔG(ΔH_PB − TΔS)</td>
<td>-31.18 ± 6.74</td>
</tr>
</tbody>
</table>

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. ΔPBNP 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. TΔS is the solute entropic contribution, where ΔS is the sum of translational, rotational and vibrational entropies. ΔG(ΔH_PB − TΔS) 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

<table>
<thead>
<tr>
<th>DNA</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Telomeric DNA</td>
<td>5’-AGGGTTAGGGTTAGGGTTAGG-3’</td>
</tr>
<tr>
<td>c-MYC DNA</td>
<td>5’-TGAGGGTGGGTAGGGTTAGG-3’</td>
</tr>
<tr>
<td>c-KITI DNA</td>
<td>5’-GGGAGGGCGCTGGGAGGAGGG-3’</td>
</tr>
<tr>
<td>h-RAS1 DNA</td>
<td>5’-TGAGGGTGGGTAGGGTGGTAA-3’</td>
</tr>
<tr>
<td>Duplex DNA</td>
<td>5’-CCAGTTCTGATGTAACCC-3’</td>
</tr>
<tr>
<td></td>
<td>5’-GGGTACTACGAATGG-3’ (complementary strand)</td>
</tr>
<tr>
<td>Primer for stop assay</td>
<td>5’-ACGACTCACTATAGCAATTGCG-3’</td>
</tr>
<tr>
<td>Template of c-MYC DNA</td>
<td>5’-TGAGGGTGGGTAGGGTTAGGTTAGGGTTAGGCTATAGGTAGGCTG-3’</td>
</tr>
<tr>
<td></td>
<td>GCTATAGGTAGTCGT-3’</td>
</tr>
<tr>
<td>Template of mutated c-MYC DNA</td>
<td>5’-TGAGGGTGGGTAGGTTAGGGTTAGGCTATAGGTAGGCTG-3’</td>
</tr>
<tr>
<td></td>
<td>GCTATAGGTAGTCGT-3’</td>
</tr>
</tbody>
</table>

**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**

![19F NMR spectrum of ligand 2](image)

**$^1$H NMR spectrum of ligand 3**

![1H NMR spectrum of ligand 3](image)
$^{13}$C NMR spectrum of ligand 3