Electronic Supplementary Material (ESI) for Molecular BioSystems.

Supporting Information

Ligand Induced Conformational Preorganization of Loops of *c-MYC* G-Quadruplex DNA and its Implications in Structure Specific Drug Design

S. Harikrishna*, Saikiran Kotaru and P. I. Pradeepkumar*

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

*Email: harikrishna.s@iitb.ac.in or pradeep@chem.iitb.ac.in

TABLE OF CONTENTS

Energy optimized structures of ligands at HF/6-31G* level

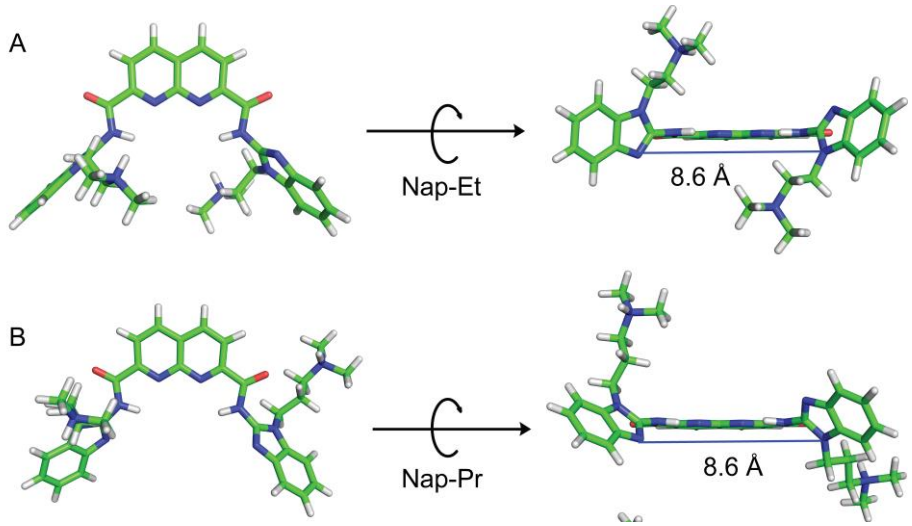

Figure S1. Energy optimized structures of ligands using HF/6-31G* theory level in Gaussian 09. (A) **Nap-Et** and (B) **Nap-Pr** optimized structures. Atoms are shown in stick representation. The solid blue lines between two benzimidazole rings specify the distance between benzimidazole side chains.

Stacking interaction between ligands and top-quartet of the G4

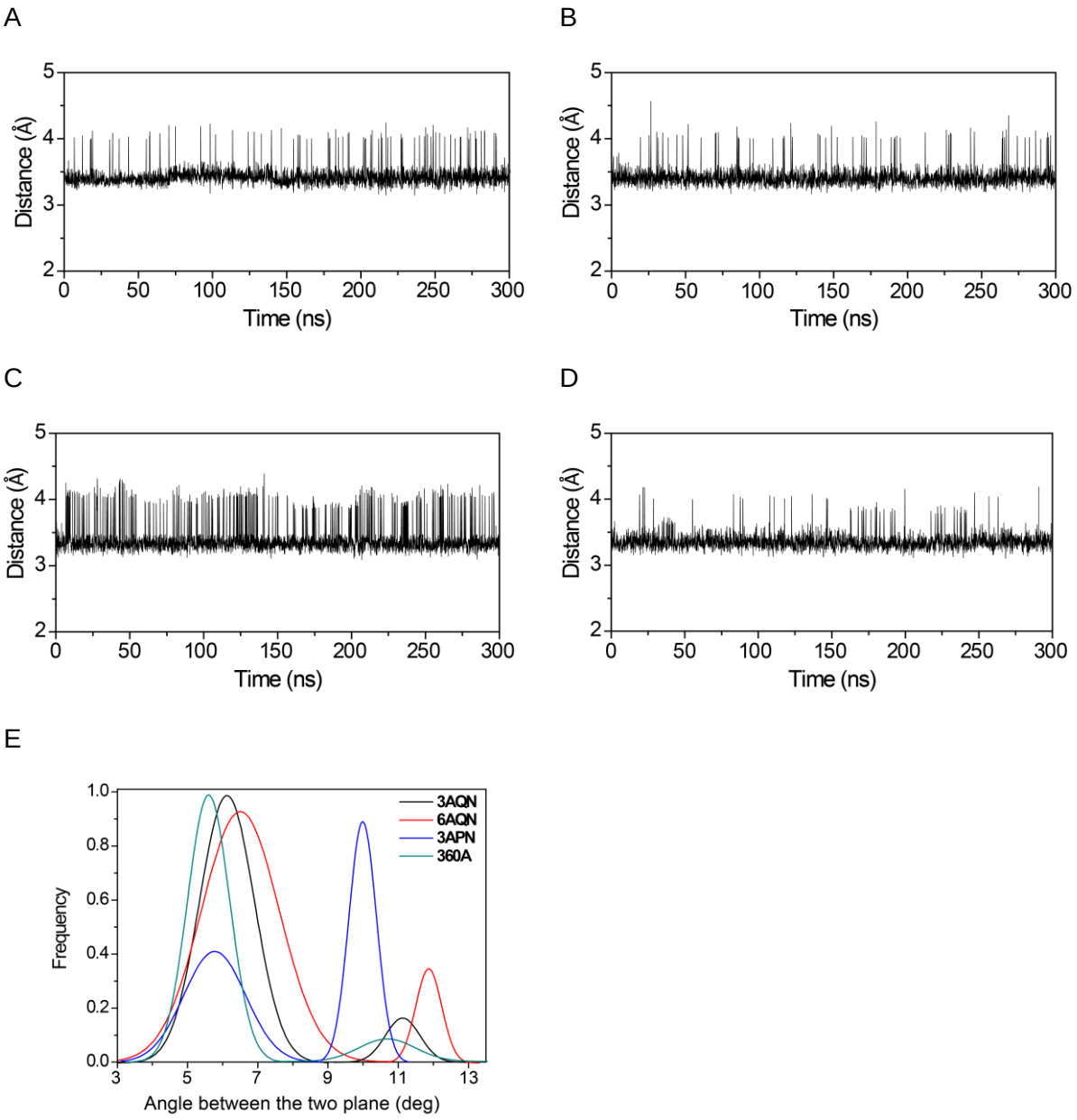

Figure S2. Stacking distance and angle between the plane of the aromatic moiety in the ligand and plane of the G-quartet during the course of MD simulations. Distance between the plane of (A) 3AQN and top quartet, (B) 6AQN and top quartet, (C) 3APN and top quartet (D) 360A and top quartet and (E) angle between the plane of ligand and quartet. These calculations were performed using PLUMED plugin in the UCSF Chimera.

Non-covalent interactions between 5’-flanking nucleotides and ligands

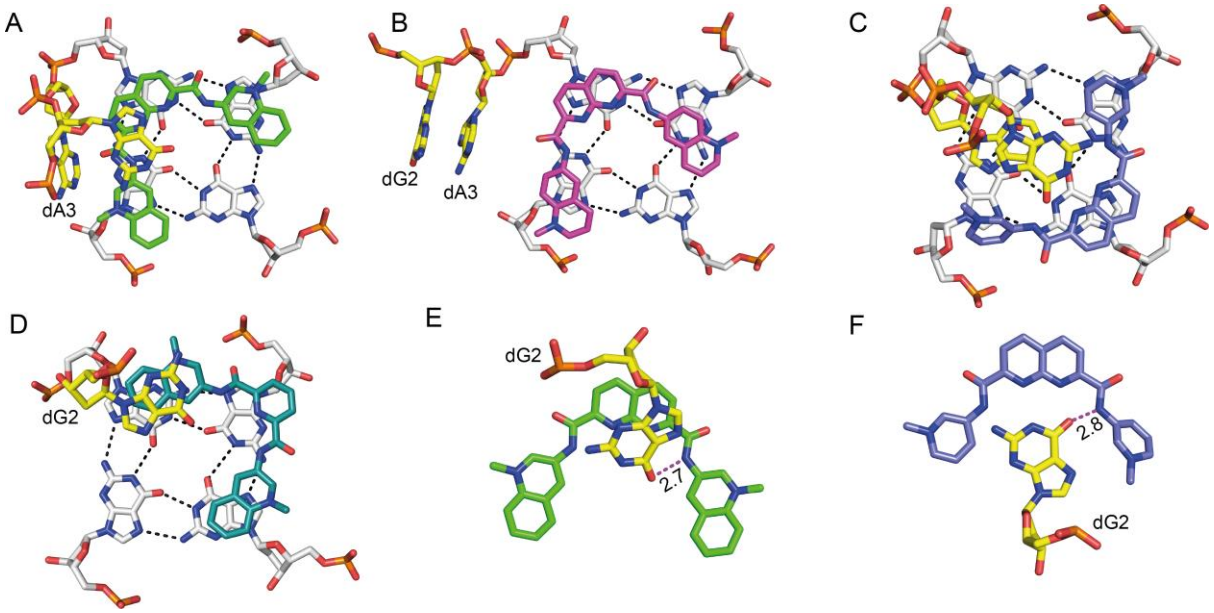

Figure S3. Non-covalent interactions between top quartet binding ligands and the 5’-flanking nucleotides. (A) The dG2 in the 5’-flanking nucleotide stacks on the **3AQN**. (B) Both dG2 and dA3 in the 5’-flanking nucleotides flipped out of the G-quartet surface and are not stacking on the **6AQN** (C) The 5’-flanking nucleotide dA2 stacks on the G-quartet and the dA3 stacks on the dG2 nucleotide in the **3APN** complex. (D) The dG2 nucleotide stacks on the **360A**. (E) The O^6 in the dG2 nucleotide make H-bond interaction with the NH group of the side chain in **3AQN**. (F) The O^6 in the dG2 nucleotide make H-bond interaction with the NH group of the side chain in **3APN**. All the ligands shown here stack on the top quartet of the G4 DNA. All the distances were mentioned in Å.

Non-covalent interactions between DNA and groove binding ligands

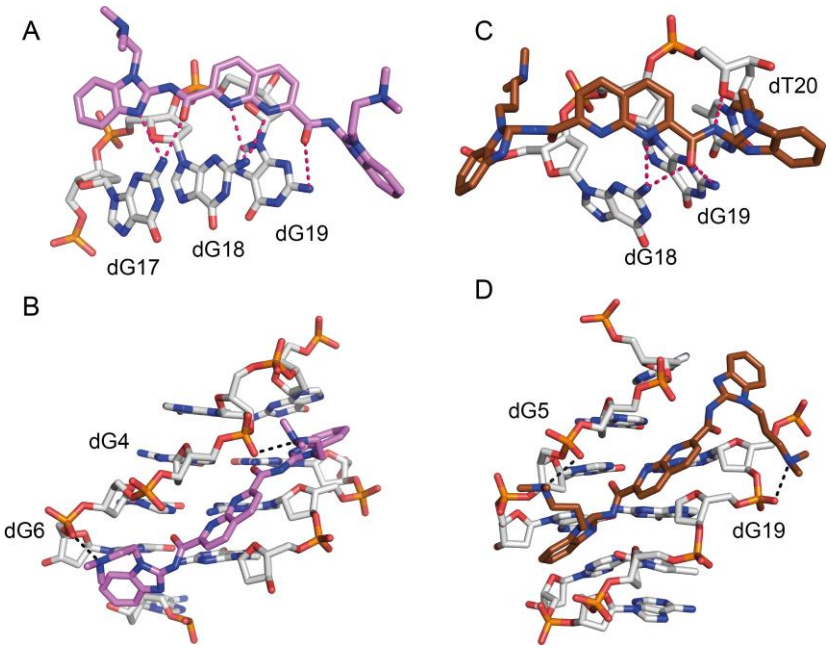

Figure S4. Non-covalent interactions between the G4 DNA and the groove binding ligands. (A) H-bond interactions between guanines including dG17, dG18 and dG19 in the G-quartet and **Nap-Et**, the distances of these H-bonds were between 2.7 and 3.1 Å, and the occupancies of these H-bonds are found to be >65 % of the total simulation time. (B) Electrostatic interactions between the positively charged side chain in **Nap-Et**, and the negatively charged phosphate backbone of dG4 and dG6 in the G-quartet. The distances of the two electrostatic contacts were between 2.5 and 3.3 Å. (C) H-bond interactions between guanines including dG18 and dG19 in the G-quartet, dT20 in the 3'-flanking nucleotide and **Nap-Pr**. The distances of these H-bonds were between 2.8 and 3.1 Å, and the occupancy of these H-bonds are found to be >62 % of the total simulation time. (D) Electrostatic interactions between the positively charged side chain in **Nap-Pr** and the negatively charged phosphate backbone of dG5 and dG19 in the G-quartet. The distances of the two electrostatic interactions are between 2.7 and 3.4 Å.

Dynamic cross-correlation map of ligand free *c-MYC* G4 DNA

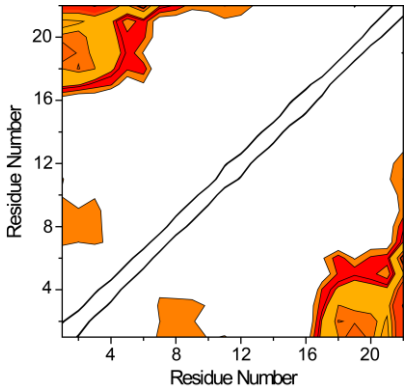

Figure S5. Dynamic cross-correlation map (DCCM) of ligand free *c-MYC* G4 DNA during 300 ns of MD simulations. Correlation between 0.75 to 0.98 and anticorrelation between −0.98 to −0.75 were considered to plot the graph. Red (0.80 to 0.98); Orange (0.78 to 0.88). The correlated motion between the G-quartet were discarded for clarity.

Percentage occupancies of the clusters of loop-2 conformers from MD simulations

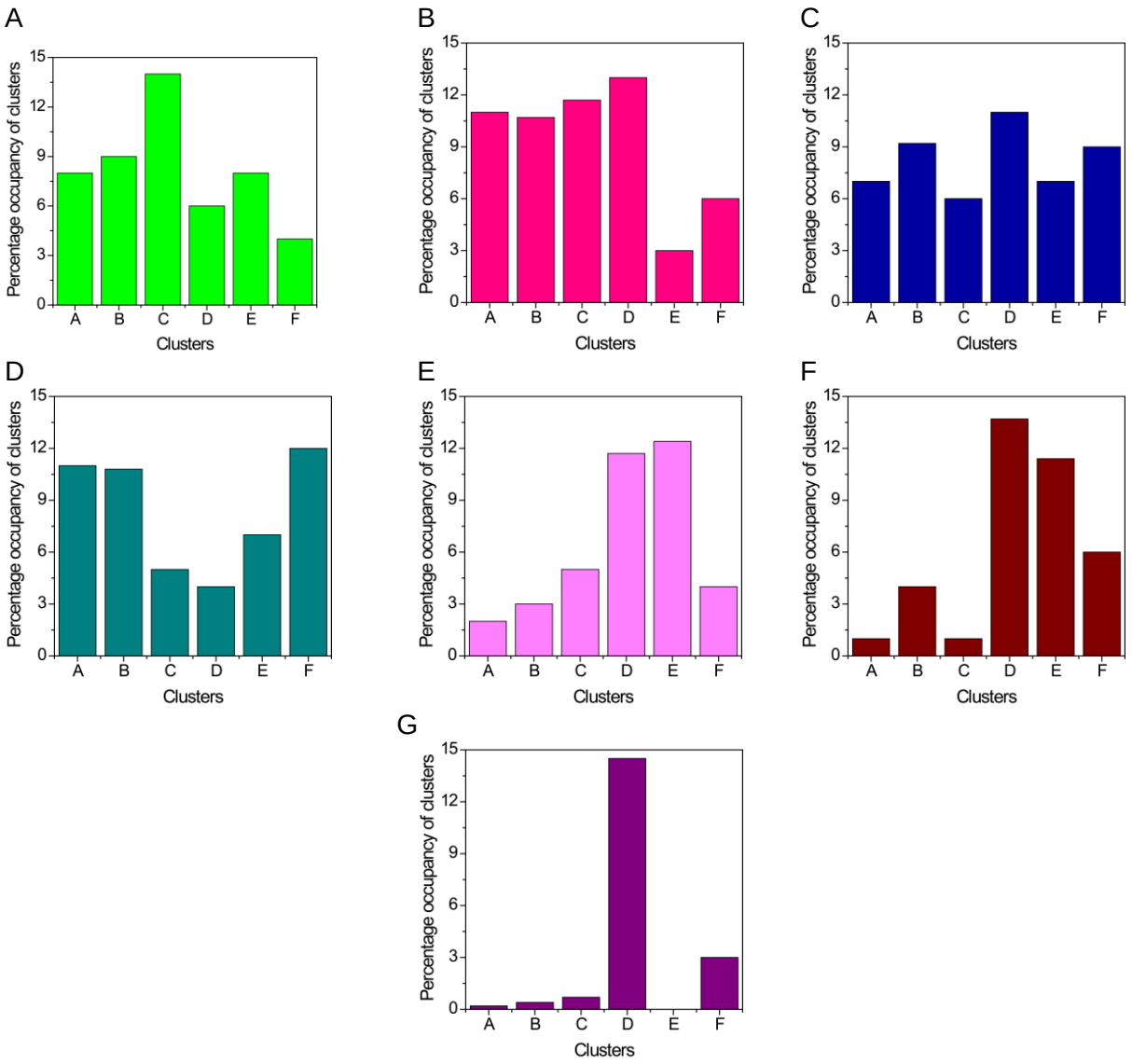

Figure S6. Percentage occupancies of the clusters (loop2) from the 300 ns MD simulations trajectories. Six conformational ensembles identified from the cluster analysis of the MD simulations in complex with six different ligands including (A) **3AQN**, (B) **6AQN**, (C) **3APN**, (D) **Nap-Et** (E) **Nap-Pr** and the (F) ligand free G4 DNA. The representative structures of the conformers A-F were shown in Figure 8 (Main text).

SASA values of the *c-MYC* G4 DNA in complex with ligands

c-MYC G4 DNA -Ligand	SASA (Å^2)	ΔSASA (Å^2)
3AQN	3671	397
6AQN	3715	441
3APN	3607	333
360A	3678	404
Nap-Et	3342	68
Nap-Pr	3346	72

Table S1. The solvent accessible surface area values of *c-MYC* G4 DNA in complex with ligands used. The SASA of native *c-MYC* G4 DNA is 3274 Å^2. The ΔSASA is calculated as the difference between the values of native *c-MYC* G4 DNA and ligand bound *c-MYC* G4 DNA complexes after 300 ns of MD simulations. SASA values were calculated using SURF tool in AMBER 14.