PCCP

ROYAL SOCIETY OF CHEMISTRY

PAPER

View Article Online
View Journal | View Issue

Cite this: *Phys. Chem. Chem. Phys.*, 2022, **24**, 6238

Stabilization and fluorescence light-up of G-quadruplex nucleic acids using indolyl-quinolinium based probes†

Annyesha Biswas, Sushma B. Singh, ‡ Chaitra S. Todankar,‡ Sruthi Sudhakar, Sushree Prangya Priyadarshinee Pany and P. I. Pradeepkumar *

Received 15th October 2021,
Accepted 7th February 2022

DOI: 10.1039/d1cp04718c

rsc.li/pccp

G-Quadruplexes (G4s) are four-stranded motifs formed by G-rich nucleic acid sequences. These structures harbor significant biological importance as they are involved in telomere maintenance, transcription, and translation. Owing to their dynamic and polymorphic nature, G4 structures relevant for therapeutic applications need to be stabilized by small-molecule ligands. Some of these ligands turn on fluorescence upon binding to G4 structures, which provides a powerful detection platform for G4 structures. Herein, we report the synthesis of fluorescent ligands based on the indolyl-quinolinium moiety to specifically stabilize G4 structures and sense DNA. CD titration and melting experiments have shown that the lead ligand induces the formation of parallel G4 with preferential stabilization of the *c-MYC* and *c-KIT1* promoter G4s over the telomeric, *h-RAS1* G4, and duplex DNA. Fluorimetric titration data revealed fluorescence enhancement when these ligands interact with G4 DNA structures. The fluorescence lifetime experiment of the ligand with different DNAs revealed three excited state lifetimes (ns), which indicates more than one binding site. MD studies showed that the ligand exhibits 3 : 1 stoichiometry of binding with *c-MYC* G4 DNA and revealed the unique structural features, which impart selectivity toward parallel topology. The ligand was found to have low cytotoxicity and exhibited preferential staining of DNA over RNA. Collectively, the results presented here offer avenues to harness indolyl-quinolinium scaffolds for sensing and selective stabilization of G4 structures.

Introduction

G-Quadruplexes (G4s) are non-canonical secondary structures of DNA. These structures are four-stranded motifs formed by π–π stacking of two or more G-quartets (Fig. 1).[1] G-Quartet is an organization of four guanine bases arranged in a square planar fashion, held by Hoogsteen hydrogen bonding between adjacent guanine bases.[2] The negative electrostatic potential created at the inner core of the G4 structure due to the presence of O^6 atoms of the guanine bases is stabilized by suitable monovalent cations such as Na^+ and K^+.[3] G4 structures are formed by G-rich sequences present in the end part of chromosomes called telomeres,[4] promoter regions[5] of proto-oncogenes (*c-MYC*,[6] *c-KIT*,[7] *BCL-2*,[8] *VEGF*,[9] *HIF1α*,[10] *PDGF-A*,[11] *etc.*), introns,[12] immunoglobin switch regions,[13] and untranslated regions of mRNAs.[14] Because of their conformational dynamics, these structures can display various topologies depending upon the strand orientation, size and sequence of loops, metal ions in the medium and torsional angle of glycosidic bonds (Fig. 1).[15–17] Telomeric G4 DNA shows structural polymorphism and exists in various topologies such as parallel, antiparallel and (3+1) hybrid,[18] whereas promoter G4 DNA exists mostly in parallel topology in the presence of K^+ ions.[19,20]

G4 structures have gained immense prominence in scientific research because they play a pivotal role in regulating a plethora of biological processes such as maintenance of telomere length by telomerase inhibition, DNA replication, and regulation of gene expression.[21–23] Since G4 structures are dynamic in nature, visualization of relevant G4 structures is crucial to validate their *in vivo* existence and corresponding biological significance. Although *in vivo* visualization of G4 structures was reported by immunofluorescent staining using BG4 antibodies, this approach is limited because it requires fixed cells.[24] However, recently, Balasubramaniam and co-workers developed a G4-specific fluorescent probe (SiR-PyPDS) and established single-molecule visualization of G4 formation in live cells.[25] This demands the development of cell-permeable small molecule optical probes that would enable the detection of G4 structures in

Department of Chemistry, Indian Institute of Technology Bombay, Powai, Mumbai 400 076, India. E-mail: pradeep@chem.iitb.ac.in

† Electronic supplementary information (ESI) available. See DOI: 10.1039/d1cp04718c

‡ These authors contributed equally.

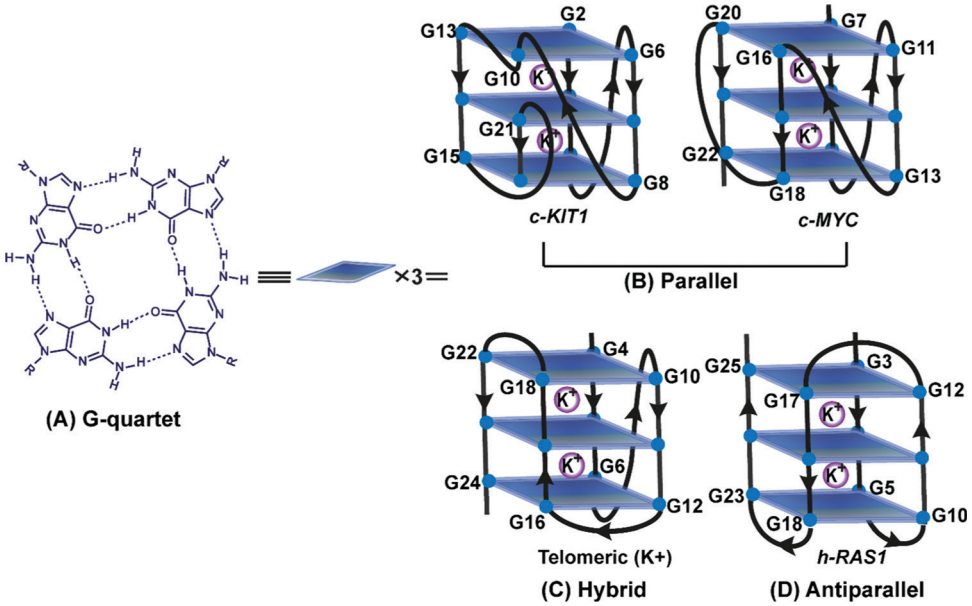


Fig. 1 Schematic showing G-quartet and *c-KIT1, c-MYC,* telomeric and *h-RAS1* G-quadruplex structures showing various topologies in the presence of K^+.

live cells.[26,27] Probes such as cyanine dyes,[28–30] triphenylmethane dyes,[31] thioflavin-T and its analogs,[32,33] imidazole analogs,[34] bis-quinolinium derivatives,[35,36] triphenylamine derivatives,[35,37] and thiazole orange analogs[38,39] exhibit strong fluorescence (light-up probes) upon binding to G4 nucleic acids. However, only a few are selective or specific to a particular G4 structure. This opens up avenues to design target-specific G4 fluorescence light-up probes that exhibit strong fluorescence emission upon target binding.

Herein, we designed ligands with a common indolyl-quinolinium core with ethyl and propyl side chains and different cationic groups (Fig. 2). We explored their fluorescence properties, binding efficiency, selectivity, and stabilization with various G4 structures. Biophysical studies such as CD titration, CD melting, absorption and fluorescence titrations, and TCSPC studies, in addition to cell-based assays, were employed to evaluate the interaction of these ligands with G4 structures. Molecular modeling and dynamics studies were used to delineate the binding mode of ligands to target G4 structures.[22]

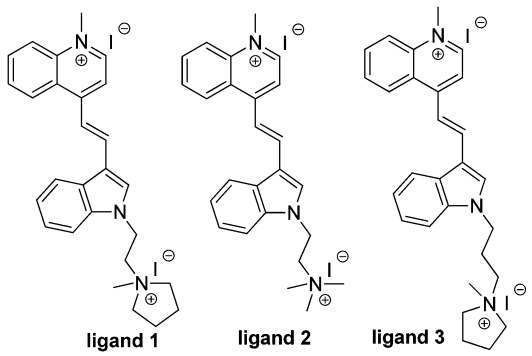


Fig. 2 Structure of ligands **1**, **2** and **3** containing an indolyl-quinolinium core with cationic amine side chains.

Experimental section

General methods

All chemicals and solvents used were obtained from commercial sources (Merck, India; Spectrochem, India; Sigma-Aldrich, Germany; and Alfa Aesar, UK). Dry solvents ACN and EtOH were dried using calcium hydride and magnesium, respectively. Commercially obtained DMF was dried overnight under a nitrogen atmosphere by incubation in activated molecular sieves (4 Å). Thin-layer chromatography (TLC) was performed using silica gel plates (Merck, Germany) pre-coated with a fluorescent indicator and further visualized under UV light (260 nm). Silica gel (100–200 mesh) and basic alumina (60–325 mesh BSS) were used for column chromatography to purify the compounds, wherever required. ^{1}H and ^{13}C NMR spectra were recorded on Bruker-AvanceIII 400 MHz and 500 MHz instruments. The chemical shifts in parts per million (ppm) were recounted downfield from the TMS signal (0 ppm) and referenced to the TMS signal or residual proton signal of deuterated solvents: $CDCl_3$ (7.26 ppm), CD_3OD (3.31 ppm), and DMSO-d_6 (2.5 ppm) for ^{1}H NMR spectra; and $CDCl_3$ (77.2 ppm), CD_3OD (49.1 ppm), and DMSO-d_6 (39.5 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 doublets), and (q) quintet or m (multiplet and overlapping spin systems). The apparent coupling constants (J) are reported in Hz. High resolution mass spectra (HRMS) were obtained on a Maxis Impact (Bruker) instrument in positive ion electrospray ionization (ESI) mode using a Q-TOF analyzer. The Discover-SP (CEM, USA) instrument was used to carry out the microwave reaction.

Synthetic procedures

1,4-Dimethylquinolin-1-ium iodide (5). Commercially obtained lepidine (100 mg, 0.69 mmol) was dissolved in dry

ACN (2 ml) and refluxed with MeI (0.2 ml, 2.79 mmol) overnight. The solvent was evaporated under reduced pressure and the yellow precipitate was separated by washing with cold diethyl ether to obtain the desired pure methylated iodide salt. Yield: (158 mg) 80%; R_f = 0.2 (5% methanol in DCM); ^{1}H NMR (400 MHz, DMSO-d_6): δ 9.37 (d, J = 6 Hz, 1H), 8.54 (d, J = 9.3 Hz, 1H), 8.49 (d, J = 8.8 Hz, 1H), 8.29–8.25 (m, 1H), 8.08–8.05 (m, 2H), 4.58 (s, 3H), 3.01 (s, 3H); ^{13}C NMR (100 MHz, DMSO-d_6): δ 158.1, 148.9, 137.6, 134.9, 129.6, 128.4, 126.7, 122.4, 119.5, 45.0, 19.2; HRMS (ESI): m/z calcd for $C_{11}H_{12}N$ $[M]^+$ 158.0964; found, 158.0967 (Δm = +0.0003 and error = +1.8 ppm).

1-(2-Bromoethyl)-1*H*-indole-3-carbaldehyde (7). Commercially available indole-3-carbaldehyde **6** (1 g, 6.88 mmol) was dissolved in dry DMF (10 ml) at RT followed by addition of anhydrous K_2CO_3 (1.9 g, 13.76 mmol). Then 1,2-dibromoethane (1.2 ml, 13.76 mmol) was added at 0 °C under a nitrogen atmosphere and the reaction mixture was stirred for 24 h at RT. After completion of the reaction, the compound was extracted with EtOAc. The organic layer was dried over anhydrous Na_2SO_4, evaporated under reduced pressure and purified by column chromatography (20% ethyl acetate in petroleum ether) using silica gel to afford the corresponding brominated compound **7** as a tan coloured solid. Yield: (550 mg) 32%; R_f = 0.5 (30% ethylacetate in petroleum ether); ^{1}H NMR (400 MHz, $CDCl_3$): δ 10.02 (s, 1H), 8.34–8.32 (m, 1H), 7.78 (s, 1H), 7.37–7.32 (m, 3H), 4.59 (t, J = 6.4 Hz, 2H), 3.72 (t, J = 6.5 Hz, 2H); ^{13}C NMR (100 MHz, $CDCl_3$): δ 184.8, 138.7, 136.8, 125.6, 124.4, 123.4, 122.6, 118.7, 109.6, 48.7, 29.3; HRMS (ESI): m/z calcd for $C_{11}H_{10}BrNO$ $[M + Na]^+$ 273.9838; found, 273.9840 (Δm = +0.0002 and error = +0.6 ppm).

1-(2-(Pyrrolidin-1-yl)ethyl)-1*H*-indole-3-carbaldehyde (9a). Compound **7** (400 mg, 1.59 mmol) was dissolved in dry ACN (4 ml) followed by addition of pyrrolidine (0.5 ml, 6.35 mmol). The reaction mixture was refluxed for 12 h in an inert atmosphere. After completion of the reaction, the solvent was evaporated under reduced pressure. The resulting residue was purified by column chromatography (0–1% MeOH in DCM) using basic alumina as the stationary phase to afford compound **9a**; yield: (90 mg) 23%; R_f = 0.5 (5% methanol in DCM); ^{1}H NMR (400 MHz, $CDCl_3$): δ 9.99 (s, 1H), 8.30 (d, J = 7.3 Hz, 1H), 7.81 (s, 1H), 7.40–7.29 (m, 3H), 4.30 (t, J = 6.4 Hz, 2H), 2.95 (t, J = 6.8 Hz, 2H), 2.57 (s, 4H), 1.80 (s, 4H); ^{13}C NMR (100 MHz, $CDCl_3$): δ 184.7, 139.0, 137.3, 125.5, 124.1, 123.0, 122.3, 118.3, 110.0, 55.3, 54.4, 46.4, 23.7; HRMS (ESI): m/z calcd for $C_{15}H_{19}N_2O$ $[M + H]^+$ 243.1492; found, 243.1494 (Δm = −0.0002 and error = −0.9 ppm).

1-(2-(3-Formyl-1*H*-indol-1-yl)ethyl)-1-methylpyrrolidin-1-ium iodide (11a). Compound **9a** (190 mg, 0.78 mmol) was dissolved in ACN (4 ml) and refluxed with MeI (0.2 ml, 3.14 mmol) for 12 h. The solvent was evaporated under reduced pressure and the crude compound was purified using basic alumina column chromatography (1–2% MeOH in DCM) to afford the methylated iodide salt **11a**; m.p. −90 to 92 °C; yield: (182 mg) 61%; R_f = 0.3 (5% methanol in DCM); ^{1}H NMR (400 MHz, DMSO-d_6): δ 9.96 (s, 1H), 8.48 (s, 1H), 8.14 (d, J = 7.8 Hz, 1H), 7.78 (d, J = 8.3 Hz, 1H), 7.41–7.32 (m, 2H), 4.85 (t, J = 7.4 Hz, 2H), 3.88 (t, J = 7.4 Hz, 2H), 3.62–3.54 (m, 4H), 3.18 (s, 3H), 2.12 (s, 4H); ^{13}C NMR (100 MHz, DMSO-d_6): δ 184.8, 140.7, 136.7, 124.6, 123.9, 122.9, 121.1, 117.8, 111.2, 64.1, 60.6, 47.9, 40.7, 21.0; HRMS (ESI): m/z calcd for $C_{16}H_{21}N_2O$ $[M]^+$ 257.1648; found, 257.1648.

Ligand 1. Compound **11a** (80 mg, 0.20 mmol) and compound **5** (58 mg, 0.20 mmol) were dissolved in EtOH (4 ml) along with 2–3 drops of piperidine and heated at 50 °C. After 30 min, the mixture was subjected to microwave irradiation at 280 W for 10 min at a temperature of 50 °C. The solvent was evaporated under reduced pressure and the precipitate obtained was washed with a minimum amount of DCM and an excess of diethyl ether to get the pure red coloured solid product; m.p. −265 to 267 °C; yield: (42 mg) 32%; R_f = 0.3 (10% methanol in DCM); ^{1}H NMR (500 MHz, DMSO-d_6): δ 9.19 (d, J = 6.7 Hz, 1H), 8.97 (d, J = 8.2 Hz, 1H), 8.55 (s, 1H), 8.55–8.50 (m, 3H), 8.39 (d, J = 8.9 Hz, 1H), 8.32 (d, J = 7.6 Hz, 1H), 8.25 (t, J = 7.5 Hz, 1H), 8.07–8.01 (m, 2H), 7.80 (d, J = 8.1 Hz, 1H), 7.43–7.36 (m, 2H), 4.88 (t, J = 7.3 Hz, 2H), 4.48 (s, 3H), 3.92 (t, J = 7.4 Hz, 2H), 3.64–3.57 (m, 4H), 3.21 (s, 3H), 2.14 (s, 4H); ^{13}C NMR (100 MHz, DMSO-d_6): δ 153.2, 147.0, 138.8, 137.0, 136.9, 134.7, 134.5, 128.7, 126.1, 126.0, 125.5, 123.5, 122.0, 120.7, 119.2, 114.4, 114.1, 114.0, 111.2, 64.1, 60.7, 47.9, 44.1, 40.6, 21.0; HRMS (ESI): m/z calcd for $C_{27}H_{31}N_3$ $[M/2]^+$ 198.6254; found, 198.6259 (Δm = +0.0006 and error = +2.9 ppm).

1-(2-(Dimethyl amino)ethyl)-1*H*-indole-3-carbaldehyde (9b). Compound **7** (400 mg, 1.58 mmol) was dissolved in dry ACN (4 ml) followed by addition of dimethylamine (0.3 ml, 6.34 mmol). The reaction mixture was refluxed for 12 h under a nitrogen atmosphere. After completion of the reaction, the solvent was evaporated under reduced pressure. The resulting residue was purified by column chromatography (16% EtOAc in petroleum ether) using basic alumina as the stationary phase to afford compound **9b**; yield: (100 mg) 35%; R_f = 0.5 (5% methanol in DCM); ^{1}H NMR (400 MHz, $CDCl_3$): δ 9.99 (s, 1H), 8.32–8.30 (dd, J = 2.6 Hz, J = 6.5 Hz, 1H), 7.82 (s, 1H), 7.38–7.29 (m, 3H), 4.23 (t, J = 6.5 Hz, 2H), 2.73 (t, J = 6.7 Hz, 2H), 2.29 (s, 6H); ^{13}C NMR (100 MHz, $CDCl_3$): δ 184.7, 139.1, 137.3, 125.5, 124.0, 123.0, 122.3, 118.3, 109.9, 58.5, 45.7, 45.2; HRMS (ESI): m/z calcd for $C_{13}H_{17}N_2O$ $[M + H]^+$ 217.1335; found, 217.1337 (Δm = +0.0001 and error = +0.6 ppm).

1-(2-(3-Formyl-1*H*-indol-1-yl)ethyl)-*N*,*N*,*N*-trimethylethanaminium iodide (11b). Compound **9b** (100 mg, 0.46 mmol) was dissolved in ACN (2 ml) and refluxed with MeI (0.1 ml, 1.84 mmol) for 12 h. The solvent was evaporated under reduced pressure and the precipitate was washed with cold diethyl ether to afford the methylated iodide salt; m.p. −200 to 202 °C; yield: (135 mg) 82%; R_f = 0.2 (10% methanol in DCM); ^{1}H NMR (400 MHz, DMSO-d_6): δ 9.97 (s, 1H), 8.47 (s, 1H), 8.14 (d, J = 7.8 Hz, 1H), 7.77 (d, J = 8.0 Hz, 1H), 7.40 (t, J = 7.2 Hz, 1H), 7.32 (t, J = 7.4 Hz, 1H), 4.85 (t, J = 7.3 Hz, 2H), 3.85 (t, J = 7.4 Hz, 2H), 3.22 (s, 9H); ^{13}C NMR (100 MHz, DMSO-d_6): δ 184.7, 140.6, 136.6, 124.6, 123.9, 122.9, 121.1, 117.8, 111.1, 62.6, 52.7; HRMS (ESI): m/z calcd for $C_{14}H_{19}N_2O$ $[M]^+$ 231.1492; found, 231.1498 (Δm = −0.0006 and error = −2.8 ppm).

Ligand 2. Compound **11b** (38 mg, 0.10 mmol) and compound **5** (30 mg, 0.10 mmol) were dissolved in EtOH (4 ml) along with 2–3 drops of piperidine and heated at 50 °C. After 30 min, the mixture was subjected to microwave irradiation at 280 W for 10 min at a temperature of 50 °C. The solvent was evaporated under reduced pressure and the precipitate obtained was washed with a minimum amount of DCM and an excess of diethyl ether to get the red coloured solid pure product; m.p. −274 to 275 °C; yield: (30 mg) 48%; R_f = 0.3 (10% methanol in DCM); ^{1}H NMR (400 MHz, DMSO-d_6): δ 9.19 (d, J = 6.8 Hz, 1H), 8.97 (d, J = 8.3 Hz, 1H), 8.55–8.50 (m, 2H), 8.47 (s, 1H) 8.39 (d, J = 8.8 Hz, 1H), 8.32 (d, J = 7.3 Hz, 1H), 8.25 (t, J = 7.8 Hz, 1H), 8.08–8.01 (m, 2H), 7.78 (d, J = 7.8 Hz, 1H), 7.44–7.36 (m, 2H), 4.87 (t, J = 7.4 Hz, 2H), 4.49 (s, 3H), 3.89 (t, J = 7.5 Hz, 2H), 3.25 (s, 9H); ^{13}C NMR (100 MHz, DMSO-d_6): δ 153.2, 147.0, 138.8, 137.0, 136.9, 134.7, 134.4, 128.8, 126.1, 126.0, 125.5, 123.6, 122.1, 120.7, 119.2, 114.5, 114.1, 114.1, 111.2, 62.7, 52.8, 44.2, 40.0; HRMS (ESI): m/z calcd for $C_{25}H_{29}N_3$ $[M/2]^+$ 185.6175; found, 185.6181 (Δm = +0.0005 and error = +3.0 ppm).

1-(2-Bromopropyl)-1*H*-indole-3-carbaldehyde (8). Commercially available indole-3-carbaldehyde **6** (2 g, 13.79 mmol) was dissolved in dry ACN (30 ml) followed by addition of anhydrous K_2CO_3 (3.8 g, 27.58 mmol). Then 1,3-dibromopropane (2.8 ml, 27.58 mmol) was added at 0 °C under a nitrogen atmosphere and the reaction mixture was stirred for 24 h at RT. After completion of the reaction, water was added and the compound was extracted with EtOAc. The organic layer was dried over anhydrous Na_2SO_4, evaporated under reduced pressure and purified by column chromatography (10% DCM in petroleum ether) using silica gel to afford the corresponding brominated compound **8**. Yield: (1.8 g) 49%; R_f = 0.6 (30% DCM in petroleum ether); ^{1}H NMR (500 MHz, $CDCl_3$): δ 9.92 (s, 1H), 8.28 (d, J = 6.7 Hz, 1H), 7.70 (s, 1H), 7.37 (d, J = 6.1 Hz, 1H), 7.31–7.27 (m, 2H), 4.31 (t, J = 6.6 Hz, 2H), 3.26 (t, J = 6.0 Hz, 2H), 2.35–2.30 (m, 2H); ^{13}C NMR (125 MHz, $CDCl_3$): δ 184.5, 138.7, 136.8, 125.2, 124.0, 122.9, 122.0, 118.1, 110.0, 44.8, 31.9, 29.8.

1-(2-(Pyrrolidin-1-yl)propyl)-1*H*-indole-3-carbaldehyde (10). Compound **8** (2 g, 7.51 mmol) was dissolved in dry ACN (10 ml) followed by addition of pyrrolidine (1.2 ml, 15.03 mmol). The reaction mixture was refluxed for 12 h in an inert atmosphere. After completion of the reaction, the solvent was evaporated under reduced pressure. The resulting residue was purified by column chromatography (8% EtOAc in petroleum ether) using basic alumina as the stationary phase to afford compound **10**; yield: (1.2 g) 62%; R_f = 0.5 (5% methanol in DCM); ^{1}H NMR (400 MHz, $CDCl_3$): δ 9.91 (s, 1H), 8.27–8.25 (m, 1H), 7.69 (s, 1H), 7.36–7.35 (m, 1H), 7.29–7.27 (m, 1H), 7.25–7.23 (m, 1H), 4.21 (t, J = 6.9 Hz, 2H), 2.43–2.40 (m, 4H), 2.35 (t, J = 6.9 Hz, 2H), 2.02–1.95 (m, 2H), 1.75–1.72 (m, 4H); ^{13}C NMR (100 MHz, $CDCl_3$): δ 184.5, 139.0, 137.3, 125.4, 123.8, 122.8, 122.0, 118.0, 110.2, 53.9, 52.5, 44.9, 28.7, 23.5; HRMS (ESI): m/z calcd for $C_{16}H_{21}N_2O$ $[M + H]^+$ 257.1648; found, 257.1648.

1-(2-(3-Formyl-1*H*-indol-1-yl)propyl)-1-methylpyrrolidin-1-ium iodide (12). Compound **10** (100 mg, 0.39 mmol) was dissolved in ACN (2 ml) and refluxed with MeI (0.1 ml, 1.56 mmol) for 12 h. The solvent was evaporated under reduced pressure and washed with cold diethyl ether to afford the methylated iodide salt **12**; yield: (90 mg) 58%; R_f = 0.2 (10% methanol in DCM); ^{1}H NMR (500 MHz, MeOD): δ 9.80 (s, 1H), 8.24 (s, 1H), 8.09 (d, J = 7.6 Hz, 1H), 7.66 (d, J = 8.2 Hz, 1H), 7.30–7.27 (m, 1H), 7.24–7.21 (m, 1H), 4.37 (t, J = 7.3 Hz, 2H), 3.56–3.50 (m, 6H), 3.04 (s, 3H), 2.41–2.35 (m, 2H), 2.12 (s, 4H); ^{13}C NMR (125 MHz, MeOD): δ 186.8, 141.9, 138.4, 126.1, 125.1, 124.0, 122.5, 119.0, 111.9, 65.6, 62.1, 49.4, 44.6, 25.5, 22.4; HRMS (ESI): m/z calcd for $C_{17}H_{23}N_2O$ $[M]^+$ 271.1805; found, 271.1804 (Δm = −0.0001 and error = −0.3 ppm).

Ligand 3. Compound **12** (194 mg, 0.49 mmol) and compound **5** (139 mg, 0.49 mmol) were dissolved in EtOH (4 ml) along with 2–3 drops of piperidine and heated at 50 °C. After 30 min, the mixture was subjected to microwave irradiation at 280 W for 10 min at a temperature of 50 °C. The solvent was evaporated under reduced pressure and the precipitate obtained was washed with a minimum amount of DCM and an excess of diethyl ether to afford the pure red coloured solid product; m.p. −250 to 252 °C; yield: (145 mg) 45%; R_f = 0.2 (10% methanol in DCM); ^{1}H NMR (500 MHz, DMSO-d_6): δ 9.15 (d, J = 6.6 Hz, 1H), 8.98 (d, J = 8.5 Hz, 1H), 8.56 (d, J = 15.7 Hz, 1H), 8.48 (d, J = 6.6, 1H), 8.45 (s, 1H), 8.37 (d, J = 8.9 Hz, 1H), 8.31 (d, J = 7.5 Hz, 1H), 8.23 (t, J = 7.8 Hz, 1H), 8.06–8.00 (m, 2H), 7.75 (d, J = 7.9 Hz, 1H), 7.40–7.34 (m, 2H), 4.47 (s, 3H), 4.41 (t, J = 7.1 Hz, 2H), 3.54–3.46 (m, 6H), 3.01 (s, 3H), 2.36–2.33 (m, 2H), 2.09 (s, 4H); ^{13}C NMR (125 MHz, DMSO-d_6): δ 153.2, 146.6, 138.7, 137.4, 137.0, 134.5, 128.5, 126.1, 126.0, 125.3, 123.2, 121.8, 120.5, 119.0, 114.0, 113.6, 113.2, 111.1, 63.7, 60.3, 47.7, 44.0, 43.4, 24.2, 21.0; HRMS (ESI): m/z calcd for $C_{28}H_{33}N_3$ $[M/2]^+$ 205.6332; found, 205.6338 (Δm = +0.0006 and error = +3.2 ppm).

Stock solution of ligands. Stock solutions (5 mM) of all the ligands were prepared in DMSO.

Oligonucleotides

The DNA oligonucleotides used for all the studies are listed in Table S1 (ESI†). The oligonucleotide sequences were procured from Sangon (China) and dissolved in water as received. The oligonucleotides were further purified by 20% PAGE (7 M urea, 40% acrylamide-bisacrylamide solution and 10× TBE) through standard protocols and were further desalted using the C18 Sep-Pak cartridge. The concentration of all the oligonucleotides was measured using a PerkinElmer Lambda Bio$^+$ UV-Vis spectrophotometer at 260 nm wavelength employing suitable molar extinction coefficients (ε).

CD titration studies

The CD spectra were recorded on a Jasco-815 CD spectrometer and scanned in the 200–700 nm wavelength range using a quartz cuvette of 1 mm path length. The scanning speed was fixed at 50 nm s^{-1} with a response time of 8 s at 25 °C. The strand concentration of oligonucleotides used for the study was 10 μM. For the titration study of the quadruplex forming sequences in the absence of salt, the measurements were carried out with 10 μM DNA in 10 mM lithium cacodylate buffer, pH 7.2. The DNA quadruplex used for the study in the presence of salt was annealed by heating *c-KIT1* and telomeric

DNA sequences in 10 mM KCl and 90 mM LiCl, *c-MYC* DNA in 1 mM KCl and 99 mM LiCl, and *h-RAS1* DNA in 50 mM KCl and 50 mM LiCl, along with 10 mM lithium cacodylate buffer, pH 7.2, at 95 °C for 5 min followed by gradual cooling to RT over 3–4 h. After the addition of the ligand to DNA, each time the solution was equilibrated for 3 min. Each spectrum was taken as an average of three measurements. All spectra were baseline corrected and further analyzed using Origin 8.0 software.

CD melting studies

For melting studies, 10–15 μM strand concentration of DNA in 10 mM lithium cacodylate buffer, pH 7.2, along with the required amount of LiCl and KCl, and 3 molar equivalents of ligands were used. *c-KIT1*, 22AG telomeric DNA and duplex DNA in 10 mM KCl and 90 mM LiCl, *c-MYC* DNA in 1 mM KCl and 99 mM LiCl, and *h-RAS1* in 50 mM KCl and 50 mM LiCl were annealed by heating at 95 °C for 5 min followed by gradual cooling to RT over 3–4 h. 3 molar equivalents of ligand were added to the annealed DNA sample and kept at 4 °C overnight. Thermal melting was monitored at 263 nm for c-*KIT1* and c-*MYC*, at 295 nm for telomeric, at 290 nm for *h-RAS1* and at 242 nm for duplex DNA using a 1 mm path length quartz cuvette at a heating rate of 1 °C min^{-1}. The melting temperature was determined by plotting the sigmoidal curve using the Boltzmann function fit in Origin 8.0 software.

Absorption and emission studies

Absorption experiments were carried out using a UV-Vis-NIR spectrophotometer (UV-3600 SHIMADZU). Absorption spectra were measured in the wavelength range of 200–700 nm using a 1 ml quartz cuvette of 10 mm path length. Absorption was recorded with 30 μM ligands in various solvents. Emission spectra were recorded in the wavelength range of 200–700 nm using a slit width of 2 nm on a HORIBA Fluoromax-4 spectrofluorometer using a 1 ml quartz cuvette of 10 mm path length.

Fluorescence quantum yields (Φ_f) were determined by comparison with a standard solution using the following equation, to calculate the relative fluorescence quantum yields:

$$\Phi_f = \Phi_s \left(\frac{n_X}{n_s}\right)^2 \frac{A_S}{A_X}\frac{F_X}{F_s} \quad (1)$$

where n_X and n_s are the refractive indexes of the sample and reference, respectively, F_X and F_s are the integrated fluorescence spectra for the sample and reference, respectively, and A_X and A_s are the absorbance for the sample and reference at the excitation wavelength, respectively. Absorption and fluorescence spectra were recorded using a 10 mm path length quartz cuvette. The standard used in this study is Rhodamine 6G in EtOH having a fluorescence quantum yield of 0.94.

Fluorimetric titration studies

Emission spectra were recorded on a HORIBA Fluoromax-4 spectrofluorometer using a 1 ml quartz cuvette of 10 mm path length. The quadruplex forming DNA sequences were annealed by heating *c-KIT1*, telomeric, and duplex forming DNA in 10 mM KCl and 90 mM LiCl, *c-MYC* forming DNA in 1 mM KCl and 99 mM LiCl, and *h-RAS1* DNA in 50 mM KCl and 50 mM LiCl along with 10 mM lithium cacodylate buffer, pH 7.2, at 95 °C for 5 min followed by gradual cooling to RT over 3–4 h. Fluorescence emission spectra were recorded in the range of 470–700 nm, using a slit width of 2 nm and an excitation wavelength of 465 nm for ligands **1** and **2** and 472 nm for ligand **3**. To 5 μM concentration of ligand solution, increasing concentration of DNA solution was added. The emission spectra were recorded after equilibrating the ligand–DNA solution for 3 min. The data from the fluorimetric titrations were analyzed according to the independent site model by nonlinear fitting to eqn (2),[40] in which F_0 is the fluorescence intensity of the ligand in the absence of G4 DNA, n is the putative number of the ligand binding sites to a given DNA matrix, Q is the fluorescence enhancement upon saturation, $A = 1/[K_a C_{ligand}]$ and $x = nC_{DNA}/C_{ligand}$. The parameters Q and A were found by the Levenberg–Marquardt fitting routine in Origin 8.0 software to derive the binding constants.

$$\frac{F}{F_0} = 1 + \frac{Q-1}{2}\left[A + 1 + x - \sqrt{(A+1+x)^2 - 4x}\right] \quad (2)$$

Time-resolved fluorescence lifetime studies

Fluorescence lifetime decays were recorded using the DeltaFlex time-correlated single-photon counting (TCSPC) technique using an excitation wavelength of 465 nm for ligands **1** and **2** and 472 nm for ligand **3** with the full-width at half maximum (FWHM) of the instrument response function (IRF) being 77 ps. The photomultiplier hybrid (PMT-hybrid) was used as a detector. The fluorescence decay was collected at an emission polarizer set at a magic angle (54.7 °C). Emission decays were fit with the instrument response function (IRF), collected by the ludox solution. Decay spectra were fit by using Ez time software based on an iterative reconvolution method:

$$I'^{(t)} = I_0(t) \sum_i a_i e^{-t/\tau_i} \quad (3)$$

where I_0 and $I'^{(t)}$ are the photoluminescence (PL) intensity at $t = 0$ and $t = t$, respectively, τ is the lifetime and a_i is the amplitude of the ith component.

Cell culture

HeLa (human cervical cells) and Hep G2 (human liver cancer cells) cells obtained from (NCCS Pune) were grown in minimal essential medium Eagle (MEM Eagle) and Lenti-X (human embryonic kidney cell line from (NCCS Pune) cells were grown in Dulbecco's modified Eagle's medium (DMEM) with Earle's salts, L-glutamine and non-essential amino acids with sodium bicarbonate and 10% fetal bovine serum (HiMedia), and 1% of antibiotic antimycotic Solution 100× liquid (HiMedia) in an enriched 5% v/v CO_2 atmosphere at 37 °C in an incubator (Memmert ICO 240). For all experiments, cells were cultured and maintained in a T25 flask.

Cell viability study

Cell viability assay was performed using 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide (MTT, MP Biomedicals) as per the manufacturer's protocol. Cells were seeded at a density of 10 000 cells per well in a 96 well plate for 24 h. After 24 h, cells were treated with different concentrations of ligand **1** (0 to 105 μM) and incubated for 24 h. The next day, the medium was removed, and cells were washed with PBS. 20 μl of MTT (5 mg ml^{-1}) was added to each well and incubated for 4 h. The crystals were then dissolved with 150 μl of DMSO and absorbance was recorded at 570 nm using a PerkinElmer LAMBDA Bio + UV-Vis spectrophotometer. The percentage cell viability was quantified by using eqn (4). IC_{50} was calculated by plotting % cell viability *versus* log [ligand] using GraphPad Prism 5 software. The curve was fit using the non-linear dose–response curve using eqn (5):

$$\%\ \text{Cell viability} = \frac{\text{treated cells}}{\text{untreated cells}} \times 100 \quad (4)$$

$$Y = \text{Bottom} + \frac{(\text{Top} - \text{Bottom})}{1 + 10^{(\text{LogIC50} - X) \times \text{HillSlope})}} \quad (5)$$

Fixed-cell confocal fluorescence imaging

For confocal imaging, about 2×10^6 cells were seeded in a confocal Petri dish (SPL) and fixed with 3.7% *p*-formaldehyde for 15 min at RT and then rinsed thrice with PBS. The cells were then incubated with ligand **1** (20 μM) for 4 h. The medium was removed and the cells were washed with PBS. Cellular nuclei co-staining was done with DAPI (5 μg ml^{-1}) for 5 min. Then cells were imaged under a confocal microscope (Carl Zeiss, LSM 780). Ligand **1** was excited at 496 nm and emission was monitored at 519 nm, and for DAPI the excitation wavelength was 358 nm and emission was monitored at 461 nm.

Flow cytometry analysis

For flow cytometry experiments, HeLa cells were cultured in a T25 flask for 24 h. After 70% confluency was reached, cells were trypsinized using trypsin–EDTA 0.25% (Gibco), washed with PBS, and fixed with chilled 70% ethanol for 1 h on ice. After washing with PBS, cells were incubated with DNase 1 (0.1 g ml^{-1}, Sigma Aldrich) or RNase (0.1 g ml^{-1}, Sigma Aldrich) for 30 min at 37 °C before treatment with the ligand (5 μM) for 2 h. Cells were analyzed using the green channel (561 nm) on a FACS instrument (Becton Dickinson, Aria SORP). Statistical analysis to quantify percentage mean fluorescence intensity was performed using Origin 8.5 software and the *p*-value was calculated using unpaired Student's *t*-test. For all analyses, *p*-values less than 0.01 were considered significant.

TD-DFT studies

The structures of ligands **1**, **2** and **3** were optimized at HF/6-31G* level of theory in Gaussian 16, revision B.01. The ESP charges were also calculated for these ligands in Gaussian 16. Quantum chemical calculations were performed at the TD-DFT level by evaluating 6 excitation states using water as the solvent. The isosurface plots of the Frontier Molecular Orbitals were rendered using Chemcraft.[41]

Molecular modeling and dynamics methodology

The structure of ligand **1** was energy-optimized in Gaussian 16 at theoretical level HF/6-31G*. The optimized ligand **1** was then docked with *c-MYC* (PDB ID: 1XAV), telomeric hybrid G4 (PDB ID: 2JPZ)[42] and telomeric antiparallel G4 (PDB ID: 143D)[43] structures using AutoDock 4.2.6. 250 independent conformations were generated after docking using the genetic algorithm. Based on the docking results, electrostatic potential charges (ESP) were calculated for the docked conformations using Gaussian 16 and the restained electrostatic potential (RESP)[44] was fitted using antechamber[45] of AMBER 18. The system for simulation was prepared in tleap. The MD simulation was done using the procedure reported by Heider and Neidle.[46] The generalized AMBER force field (GAFF)[47] and OL15 AMBER DNA force field, the latter of which includes all the previous modifications,[48–52] were used for ligand **1** and DNA, respectively. The systems were neutralized by adding K^+ ions. The complex system was then solvated using TIP3P water molecules extending up to 10 Å in a rectangular box. These complexes were then subjected to 10 000 steps of minimization, 100 ps of heating, and 100 ps of density equilibration, followed by 800 ps of equilibration. Finally, the systems were subjected to 500 ns of unrestrained dynamics using the GPU accelerated version of PMEMD of AMBER18. The coordinates were saved every 2 ps in all the simulation steps. Binding free energies were calculated using the MM-PBSA method over the last 20 ns. The trajectory analysis was carried out using CPPTRAJ[53] and VMD.[54] RMSD values of heavy atoms, per nucleotide-RMSF values, and H-bonds were all calculated for every 5th frame (every 10 ps) with a total of 50 000 frames. Cutoff values of 3.5 Å and 135° were used for defining H-bonds. Hoogsteen H-bond was also found using similar methodology. UCSF Chimera[55] was used for visualization of MD trajectories. Pictures were rendered using PyMOL (www.pymol.org).

Results and discussion

Ligand design and synthesis

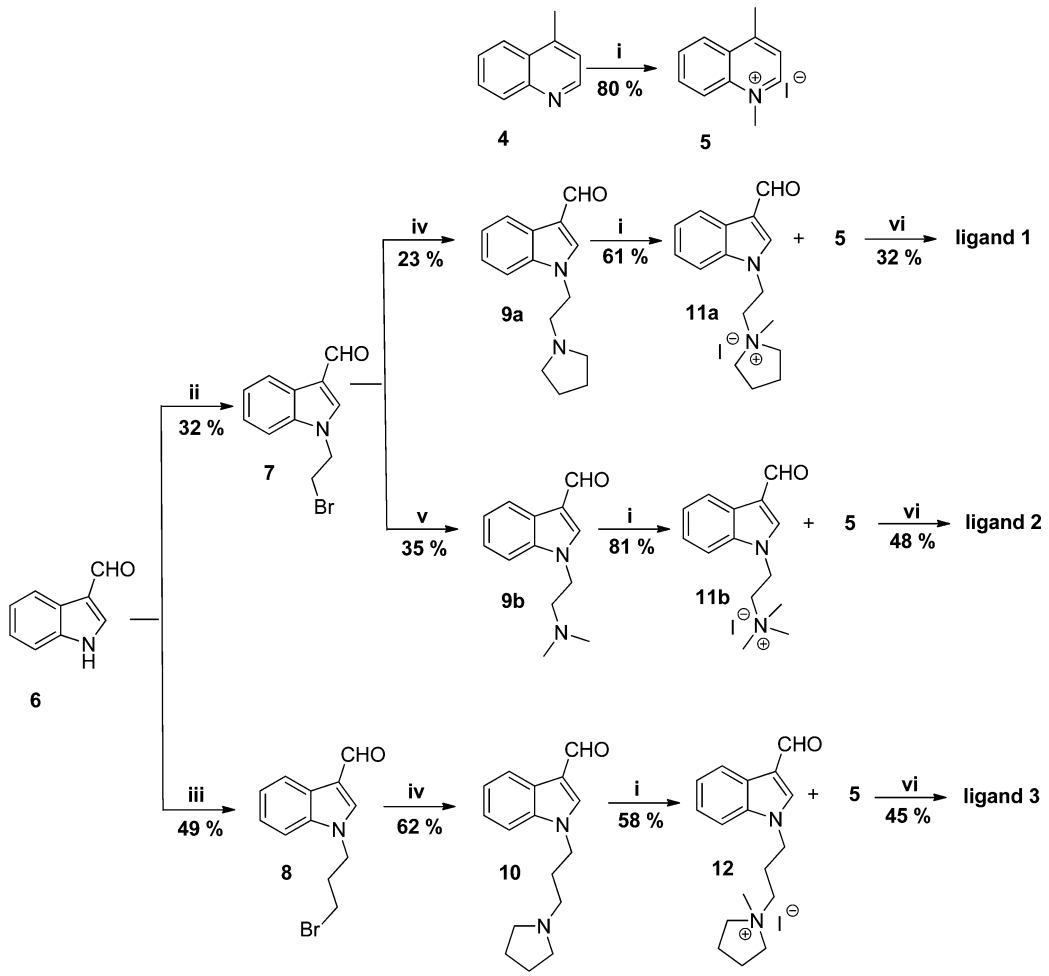
Lu and co-workers reported pyridinium-based fluorescent probes having C_1, C_2 and C_3-symmetry conjugated with different styryl-like groups to understand the effect of symmetry and side groups for G4 DNA binding and imaging studies.[56] Further, Lu and co-workers developed a scaffold consisting of 1-methylquinolinium along with the indole ring and revealed the importance of systematically introducing the indolyl moiety at different positions when conjugated with the 1-methylquinolinium scaffold *via* an ethylene bridge. These considerations were discovered as the critical factors for the good fluorescence signal switch-on property for discriminating G4 DNA from other DNA structures.[57] Sun and co-workers reported planar and rotatable π-conjugated indole and pyridinium molecules linked by an ethylene bridge to improve the structural flexibility of the

conjugated two ring co-planar system, demonstrating remarkable fluorescence discrimination between G4 DNA and other DNA tertiary structures. Xie and co-workers[58] reported indole- or pyrrole-substituted distyryl and mono-styryl dyes and studied their fluorescent light-up properties for G4 DNAs, duplex DNA, or single-stranded RNA.[59] However, none of these probes were found to have a preference for a particular G4 topology. Considering all these, we have designed ligands based on the indolyl-quinolinium moiety as donor and acceptor substituents (Fig. 2). The indole and quinoline rings are connected through the ethylene skeleton to impart extended conjugation and structural flexibility to the system. This D–π–A approach prompts strong intramolecular charge transfer (ICT), which enables absorption and emission in the longer wavelength regions.[60] The basic planar aromatic core with extended delocalization is intended to offer a large π-surface, for effective π–π stacking with G-tetrad. We further modified the structure of the ligands by using ethyl and propyl chains, to study the effect of length of the side chain on target binding. The incorporation of positively charged side chains can help support G4 binding by interacting with the negative phosphodiester backbone of the loops/grooves of G4 structures.

The synthetic path engaged to access the designed probes is shown in Scheme 1. The starting compound **5** was prepared using a previously reported method.[61] Indole-3 carbaldehyde was alkylated with 1,2-dibromoethane and 1,3-dibromopropane in the presence of a base to afford the corresponding alkyl bromides **7** and **8** in 32% and 49% yields, respectively. Bromine was then displaced using pyrrolidine and dimethylamine to provide compound **9a** in 23%, **9b** in 35%, and **10** in 62% yields.[62] The amino groups were further methylated by refluxing with methyl iodide to obtain compound **11a** in 61%, **11b** in 81%, and **12** in 58% yields. These compounds were then condensed with compound **5** under basic conditions in a microwave environment to give the desired ligands **1**, **2** and **3** in 32%, 48% and 45% yields, respectively.

CD titration studies

CD spectroscopic studies are used as a qualitative tool to obtain information about various topologies of G4 structures. CD



Scheme 1 Synthesis of ligands **1**, **2** and **3**. Reagents and conditions: (i) MeI, ACN, reflux, 12 h; (ii) 1,2-dibromoethane, K_2CO_3, DMF, RT, 24 h; (iii) 1,3-dibromopropane, K_2CO_3, DMF, RT, 24 h; (iv) pyrrolidine, ACN, reflux, 12 h; (v) dimethylamine, ACN, reflux, 12 h; (vi) piperidine, EtOH, 50 °C, 30 min and then MW (280 W), 50 °C, 10 min.

titration studies of G4 DNAs in the absence of metal ions illuminate on the competence of ligands to induce a specific G4 topology. CD titration of G4 DNAs with increasing concentration of ligand in the presence of metal ions helps to investigate the structural influence of the ligand on a particular topology of G4 DNA.[63] CD spectra of promoter G4s, *c-MYC* and *c-KIT1*, in the absence of any added metal ions, show a characteristic positive band at 260 nm and a negative band at 240 nm, suggestive of parallel G4 topology (Fig. 3A, B and Fig. S1 and S3, ESI†). Upon increasing the concentration of ligands **1**, **2** and **3**, further enhancement in the intensity of the 260 nm band was observed. The band at 240 nm fluctuated in intensity up to the addition of ~5 molar equivalents of ligands. This suggests retention of the parallel topology of *c-MYC* and *c-KIT1* DNA in the presence of ligands. For *c-MYC* G4, the onset of a bisignate ICD signal was observed in a higher wavelength region, which may be attributed to the coupling of the transition dipoles of the bound ligand and the chiral DNA (Fig. 3A).[64] With the increase in ligand concentration from 0 to 5 molar equivalents of *c-KIT1* G4, initially, a negative ICD band was seen at the higher wavelength range (Fig. 3B). The negative ICD band eventually emerged to form a bisignate ICD signal upon addition of more than 5 molar equivalents of ligand (Fig. 3B inset). As expected, in the presence of K^+ ions, *c-MYC*, and *c-KIT1*, DNA exists in parallel topology (Fig. S2 and S4, ESI†).[65] Further addition (0–5 molar equivalents) of ligands **1**, **2**, and **3** produced negligible change in the intensity of the bands. Bisignate ICD bands in the higher wavelength region were observed, due to coupling of the transition dipoles of the bound ligand and the chiral DNA, as previously seen in the absence of K^+ ions for both *c-MYC* and *c-KIT 1* G4s.

Telomeric DNA displays a positive band at 254 nm and a small band at 290 nm in CD spectra in the absence of any added metal ion in the medium, which does not correspond to any specific G4 topology (Fig. 3C and Fig. S5, ESI†). Upon addition of up to 9 molar equivalents of ligands, in all the three cases, a high-intensity band at 260 nm was observed with a trough at 240 nm, and a band with diminishing intensity at 290 nm was observed. Collectively, these observations indicate the induction of parallel topology. In the higher wavelength region, upon addition of 9 molar equivalents of ligands, a negative ICD signal was observed initially (Fig. 3C), which gradually increased in intensity and gave rise to a bisignate ICD signal (Fig. 3C inset). In the presence of K^+ ions, telomeric DNA folds into a mixture of parallel and antiparallel topology, shows a positive band at 290 nm, a shoulder band at 270 nm, a slight positive band at 250 nm and a trough at 239 nm (Fig. S6, ESI†).[66] Upon increasing the concentration of ligands **1**, **2** and **3**, broadening of the 290 nm band was observed accompanied by a decrease in intensity and a band around 256–257 nm

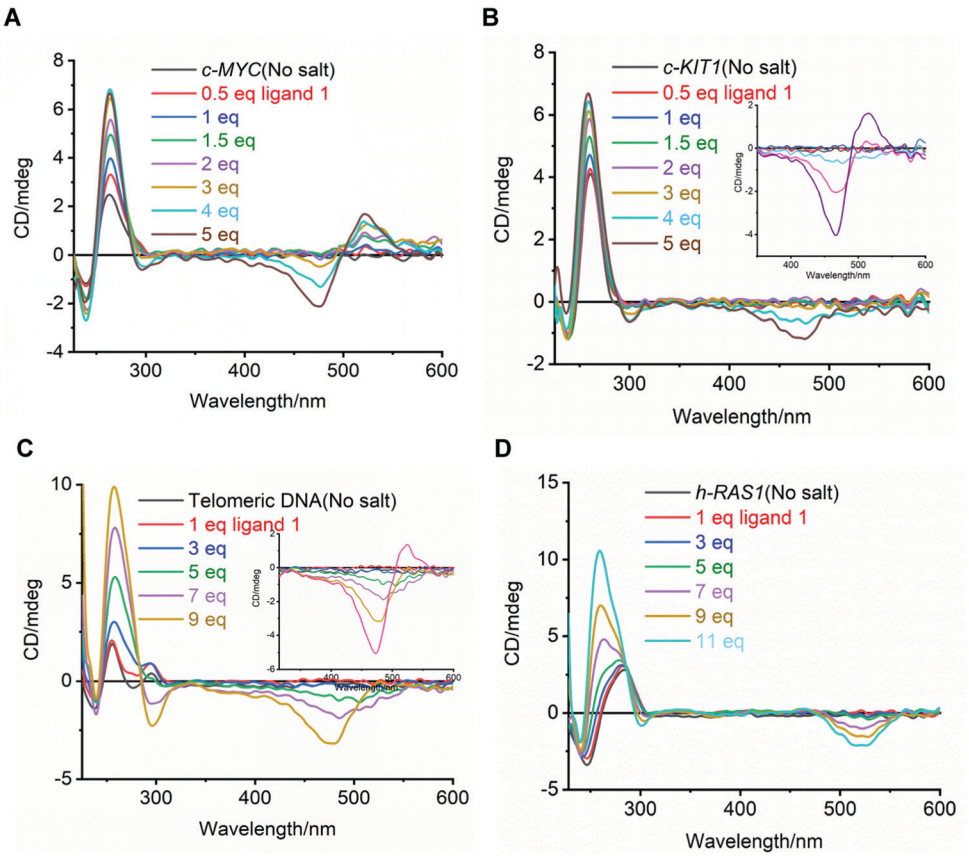


Fig. 3 CD titration of ligand **1** with quadruplex DNAs (10 μM in 10 mM lithium cacodylate buffer, pH 7.2) in the absence of added metal ions; (A) *c-MYC*; (B) *c-KIT1*; inset: 5 and more molar equivalents of ligand; (C) telomeric; inset: 9 and more molar equivalents of ligand; and (D) *h-RAS1* DNAs.

appeared, which does not account for any particular G4 topology. This may be indicative of the unfolding of the G4 structure. Also, in a higher wavelength region, a bisignate ICD signal was observed.

Induction of parallel topology for telomeric DNA in the absence of K^+ ions led us to explore the influence of ligands on *h-RAS1* DNA both in the absence and in the presence of K^+ ions. In the absence of any metal ions, *h-RAS1* exhibited a positive band at 280 nm and a negative band at 246 nm, which do not attribute to any specific G4 topology. Upon addition of up to 11 molar equivalents of ligands **1** (Fig. 3D) and **2**, and 7 molar equivalents of ligand **3**, a highly intense positive band at 260 nm appeared along with a negative band at 240 nm (Fig. S7, ESI†), indicating the presence of parallel topology. Further addition of ligands disrupted the parallel topology. In a higher wavelength region, a negative ICD band was observed in the presence of all the ligands. In the presence of K^+ ions, *h-RAS1* folds into an antiparallel structure with a sharp positive band at 290 nm and a negative band at 260 nm (Fig. S8, ESI†).[67] Upon addition of up to 11 molar equivalents of ligands, the 290 nm band diminishes in intensity. It gives rise to a new band at around 256–257 nm, which does not account for any particular G4 topology (Fig. S8, ESI†). This indicates that the ligand disrupts the antiparallel structure of *h-RAS1* G4. Overall, the results of CD titration studies suggest that ligands **1**, **2**, and **3** induce parallel topology from telomeric and *h-RAS1* G-rich sequences in the absence of K^+ ions. In the presence of K^+ ions, for both telomeric and *h-RAS1* G4, the characteristic CD bands corresponding to their mixed and antiparallel topology are diminished. Thus, both structures were seen to be unfolded at high concentrations of ligands.

CD melting studies

CD melting was used to study the effect of ligand binding on the stability of G4 structures by monitoring the CD spectra at the maximum intensity wavelength in the presence of ligands.[68] The CD titration experiment revealed that the addition of up to ~3 molar equivalents of ligands did not have any effect on the topology of G4 DNAs, both in the presence and in the absence of salt. Therefore, 3 molar equivalents of ligands were used to study the effect of ligands on thermal stabilization of the DNAs. Addition of any of the three ligands to the telomeric DNA resulted in a modest increase in the melting temperature (ΔT_m ~ 4–5 °C) when monitored at 295 nm (Table 1 and Fig. 4B). Thermal denaturation profiles for parallel promoter G4 DNAs, *c-KIT1* and *c-MYC*, were evaluated at 263 nm wavelength. In the presence of ligands **1** and **2**, *c-MYC* DNA displayed a significant enhancement in the melting temperature (ΔT_m ~ 16–18 °C) (Table 1 and Fig. S9A, ESI†). In the presence of ligand **3**, however, comparatively less enhancement in the melting temperature (ΔT_m ~ 10 °C) was observed. Similarly, the addition of ligands **1** and **2** to *c-KIT1* DNA led to a remarkable increase in the melting temperature (ΔT_m ~ 24 °C). In contrast, ligand **3** showed relatively less stabilization of *c-KIT1* DNA (ΔT_m ~ 17 °C) (Table 1 and Fig. 4A). Similarly, the CD melting experiment of *h-RAS1* DNA was monitored in the presence of K^+ at 290 nm. A moderate increase in the melting temperature was observed (ΔT_m ~ 6–7 °C) upon addition of ligands **1** and **2**, whereas the addition of ligand **3** caused a modest increase (ΔT_m ~ 3 °C) (Table 1 and Fig. S9B, ESI†).

Table 1 Thermal stability of G4 and duplex DNAs with ligands measured by CD melting experiments

	ΔT_m^a (°C)				
Ligands	*c-MYC*	*c-KIT1*	*h-RAS1*	Telomeric (K^+)	Duplex (ds 17)
Ligand **1**	17.6 ± 0.6	24.2 ± 0.8	7.2 ± 0.6	5.0 ± 0.1	1.2 ± 0.2
Ligand **2**	16.5 ± 0.4	24.1 ± 0.1	6.5 ± 0.7	4.9 ± 0.3	1.7 ± 0.2
Ligand **3**	10.6 ± 0.1	16.9 ± 0.6	2.7 ± 0.8	3.9 ± 0.1	0.6 ± 0.1

ΔT_m^a represents the difference in thermal melting [$\Delta T_m = T_m$ (DNA + 3 molar equivalents of ligands) − T_m (DNA)]. The experiments were carried out using 10 mM lithium cacodylate buffer, pH 7.2. 10 μM concentration for G4 DNAs and 15 μM concentration for duplex DNA were used in the experiments. T_m values in the absence of ligands are: 56.2 ± 0.4 °C (*c-MYC* DNA in 1 mM KCl and 99 mM LiCl); 48.6 ± 0.3 °C (*c-KIT1* DNA in 10 mM KCl and 90 mM LiCl); 52.3 ± 0.2 °C (*h-RAS1* in 50 mM KCl and 50 mM LiCl); 52.4 ± 0.3 °C (telomeric DNA in 10 mM KCl and 90 mM LiCl); and 62.1 ± 0.2 °C (duplex DNA in 10 mM KCl and 90 mM LiCl).

To further evaluate the thermal stabilization of G4 DNAs over duplex DNA by the ligands, CD melting of duplex DNA was carried out. The ellipticity of duplex DNA, monitored at 242 nm, demonstrated a negligible increase in the melting temperature, with the addition of any ligands (Table 1 and Fig. S9C, ESI†). Overall, CD melting studies revealed that all three ligands preferentially stabilize the parallel promoter G4 DNAs *c-KIT1* and *c-MYC* over antiparallel promoter *h-RAS1*, telomeric and duplex DNAs. Additionally, the ligand with the ethyl side chain emerged as a more promising candidate.

Absorption and emission studies

The absorption and emission properties of the synthesized ligands were evaluated in the presence of different solvent systems. The ligands showed the maximum blueshift of absorbance in H_2O, whereas the maximum redshift was observed in DMSO. The absorption spectra of ligand **1** showed an absorption maximum of around 485 nm in EtOH (Fig. 5A). In protic solvents, with the increase in solvent polarity, the spectra exhibited a blue shift. This effect can be explained on the basis of interactions between ligands and solvents, which stabilize the ground state of the ligand due to hydrogen bonding.[69,70] The solvent effect in aprotic solvents is, however, dictated by the dipole–dipole interactions between the dye and the solvent molecules.[30] Generally, the ground state of the ligand is known to have more dipole moment. With the increase in solvent polarity, the energy of the excited state decreases, thus, leading to a redshift of the absorption maxima.[30,71] A similar observation was made for ligands **2** and **3** as well (Fig. S10A and S11A, ESI†).

The emission spectra showed that the ligands emit in the higher wavelength region (~572–606 nm) in the presence of different solvents and thus displayed remarkable Stokes shifts

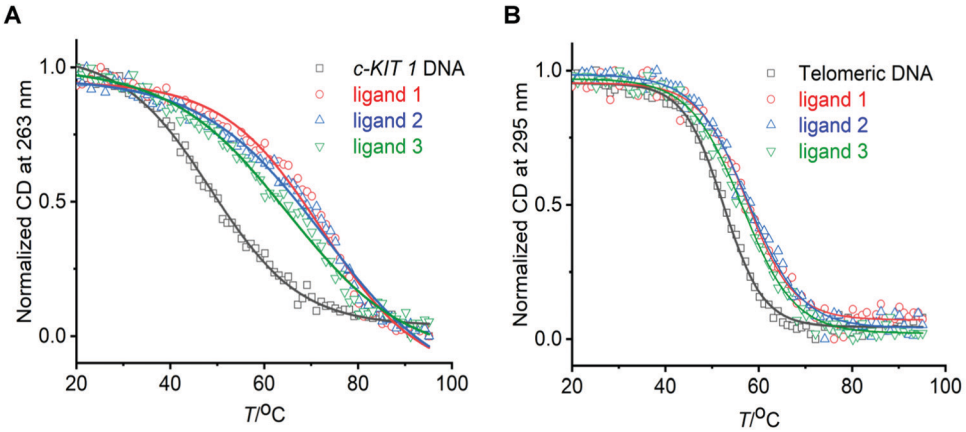


Fig. 4 Normalized CD melting curve for (A) *c-KIT1* and (B) telomeric DNAs (10 μM in 10 mM lithium cacodylate buffer, pH 7.2) in the absence and presence of 3 molar equivalent of ligands in 10 mM KCl and 90 mM LiCl.

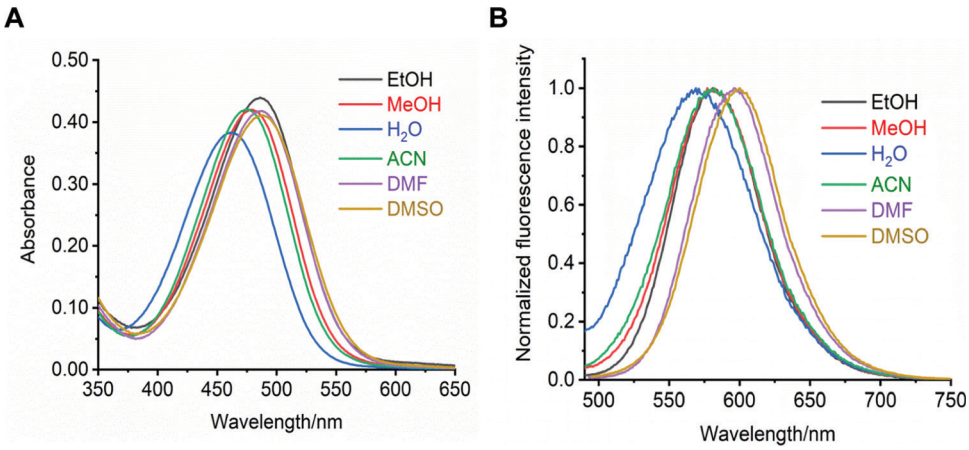


Fig. 5 Absorbance and emission spectra of 30 μM ligand **1** in the presence of various solvents.

(~88–107 nm) (Table 2). This can be attributed to the significant solvent reorganization accompanying a strong change of the dipole moment of the dye upon excitation. A large Stokes shift is observed in the case of H_2O (~100–107 nm), which indicates that these ligands can be effectively used in biological environments. The emission spectra of all the ligands displayed a red shift in λ_{em} with the increase in solvent polarity in the case of aprotic solvents (Fig. 5B and Fig. S10B and S11B, ESI†). Fluorescence quantum yield (Φ_f) calculations (Table 2) of the ligands were performed using Rhodamine 6G in EtOH as the reference.[30,72] Ligands **1**, **2** and **3** have comparable quantum yields in the respective solvents. It was further revealed that the quantum yields of ligands **1**, **2** and **3** increased in the presence of *c-MYC* G4 DNA (Table 2).

The maximum quantum yield for all the ligands was observed in DMSO (0.043–0.058) and the lowest quantum yield in H_2O (0.007–0.016) compared to other solvents, suggesting increased non-radiative decay processes in the presence of H_2O molecules.[70]

Job plot analysis

Continuous variation analysis (Job's plot) was performed to estimate the binding stoichiometry between ligand **1** and DNAs, using the fluorescence property of the ligand. In this method, the mole fraction of the ligand and DNA was varied, whereas the total concentration of the ligand and DNA was kept constant at 10 μM. The plot of fluorescence intensity *versus* mole fraction of ligand **1** showed an inflection point at 0.7–0.75 indicating a stoichiometry of ~3 : 1 for ligand–DNA interaction (Fig. 6).

Fluorimetric titration studies

In order to investigate the binding and the fluorescence properties of the ligands in the presence of various G4 and duplex DNAs, fluorescence titration experiments were performed (Fig. 7 and Fig. S12, S14, S16, ESI†). Ligand **1** exhibited an absorption maximum at 465 nm and emission at ~575 nm. Upon addition of *c-MYC* G4 DNA (~8 μM), a ~9-fold fluorescence enhancement was observed (Fig. 7A). When increasing concentration of *c-KIT1* G4 DNA was added to a fixed ligand concentration (5 μM), ~4-fold fluorescence emission was observed at the saturation concentration (~3 μM) of DNA (Fig. 7C). Further, a ~1.7-fold and ~2.6-fold increase in fluorescence emission was observed with *h-RAS1* G4 (~7 μM) (Fig. 7E) and telomeric G4 (~4 μM) DNAs (Fig. S13A, ESI†),

Table 2 Absorption and emission data of ligands in various solvents

Ligand	Solvents	λ_{max} (nm)	λ_{em} (nm)	Stokes shift[a]	Φ_f[b]
Ligand **1**	EtOH	485	580	95	0.033
	H_2O	465	572	107	0.016
	MeOH	480	580	100	0.029
	CH_3CN	477	580	103	0.019
	DMF	488	595	107	0.028
	DMSO	490	597	107	0.058
	Buffer	465	572	107	0.0039
	c-MYC	465	572	107	0.051
Ligand **2**	EtOH	485	580	95	0.034
	H_2O	465	572	107	0.013
	MeOH	480	580	100	0.027
	CH_3CN	477	580	103	0.019
	DMF	488	594	106	0.027
	DMSO	490	597	107	0.058
	Buffer	465	572	107	0.0039
	c-MYC	465	572	107	0.057
Ligand **3**	EtOH	500	588	88	0.026
	H_2O	472	572	100	0.007
	MeOH	495	588	93	0.016
	CH_3CN	490	585	95	0.011
	DMF	500	598	98	0.021
	DMSO	502	606	104	0.043
	Buffer	472	572	100	0.0037
	c-MYC	472	572	100	0.079

[a] Stokes shift. [b] Relative fluorescence quantum yield of the ligands in various solvents; the standard used is Rhodamine 6G in EtOH, Φ_f = 0.94; buffer – 10 mM lithium cacodylate buffer, pH 7.2.

respectively. But in the case of duplex DNA, ~5-fold fluorescence light-up was observed at ~5 μM concentration of DNA (Fig. S13B, ESI†). From the F/F_0 plot of the ligands with different DNAs, it was found that the maximum fluorescence enhancement was observed for *c-MYC* G4. The least fluorescence enhancement was observed for ligands **1** and **2** with *h-RAS1*, whereas for ligand **3**, it was with duplex DNA (Fig. 7G). The binding constant was obtained by plotting F/F_0 against nC_{DNA}/C_{ligand} and was fitted to an independent site model.[40] The apparent binding constants (K_a) obtained for ligand **1** with *c-MYC* and *c-KIT1* DNA were $(0.5 \pm 0.07) \times 10^6$ M^{-1} and $(4.0 \pm 0.01) \times 10^6$ M^{-1}, respectively (Table 3 and Fig. 7B, D). The K_a value of *h-RAS1* DNA was found to be $(0.3 \pm 0.04) \times 10^6$ M^{-1} (Table 3 and Fig. 7F), whereas for the duplex DNA it was $(1.1 \pm 0.01) \times 10^6$ M^{-1} (Table 3 and Fig. S13B, ESI†). This suggests that *c-KIT1* DNA has ~8-fold, ~13-fold and ~4-fold preference for binding with the ligand over *c-MYC*, *h-RAS1* and duplex DNA, respectively. Also, the maximum binding constant was obtained for *c-KIT1* G4 with ligand **1**.

Time-resolved fluorescence studies

The time-domain of the various excited state processes of the system can be probed using the time correlated single photon counting (TCSPC) technique.[73] An insight into excited state lifetimes and decay processes provides information on the

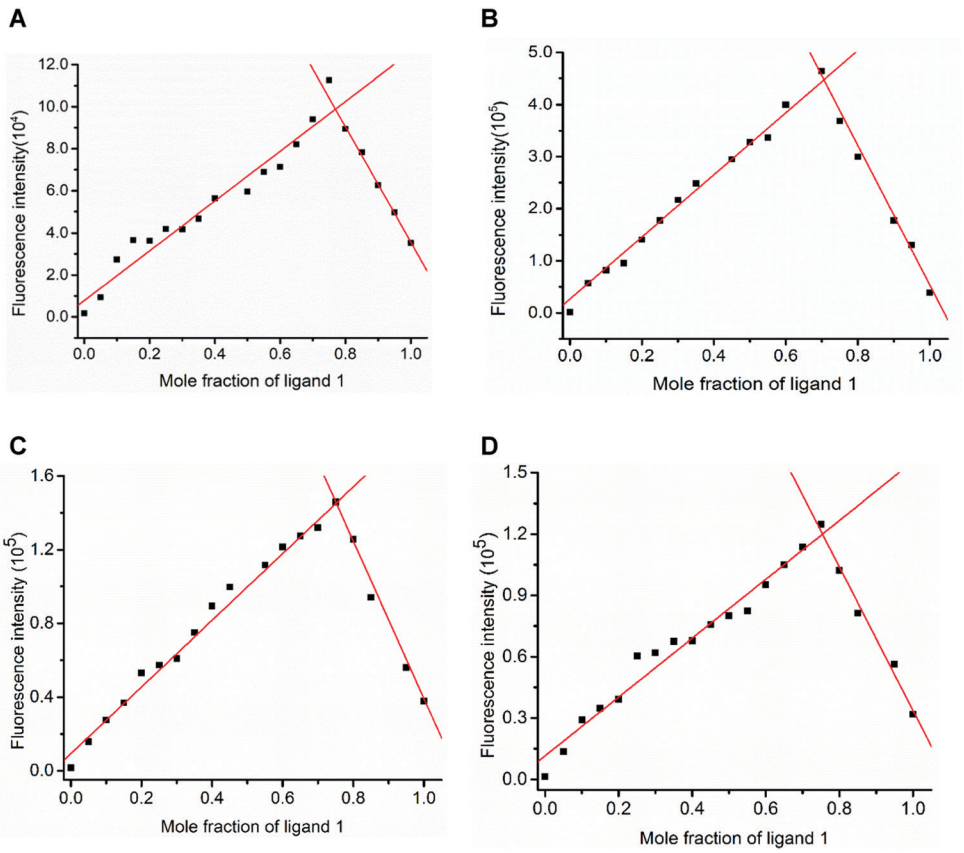


Fig. 6 Job's plot analysis of ligand **1** with (A) *c-KIT1*; (B) *c-MYC*; (C) *h-RAS1*; and (D) duplex DNA. The total molar concentration (ligand **1** + DNA) was kept constant at 10 μM. The excitation wavelength used was 465 nm.

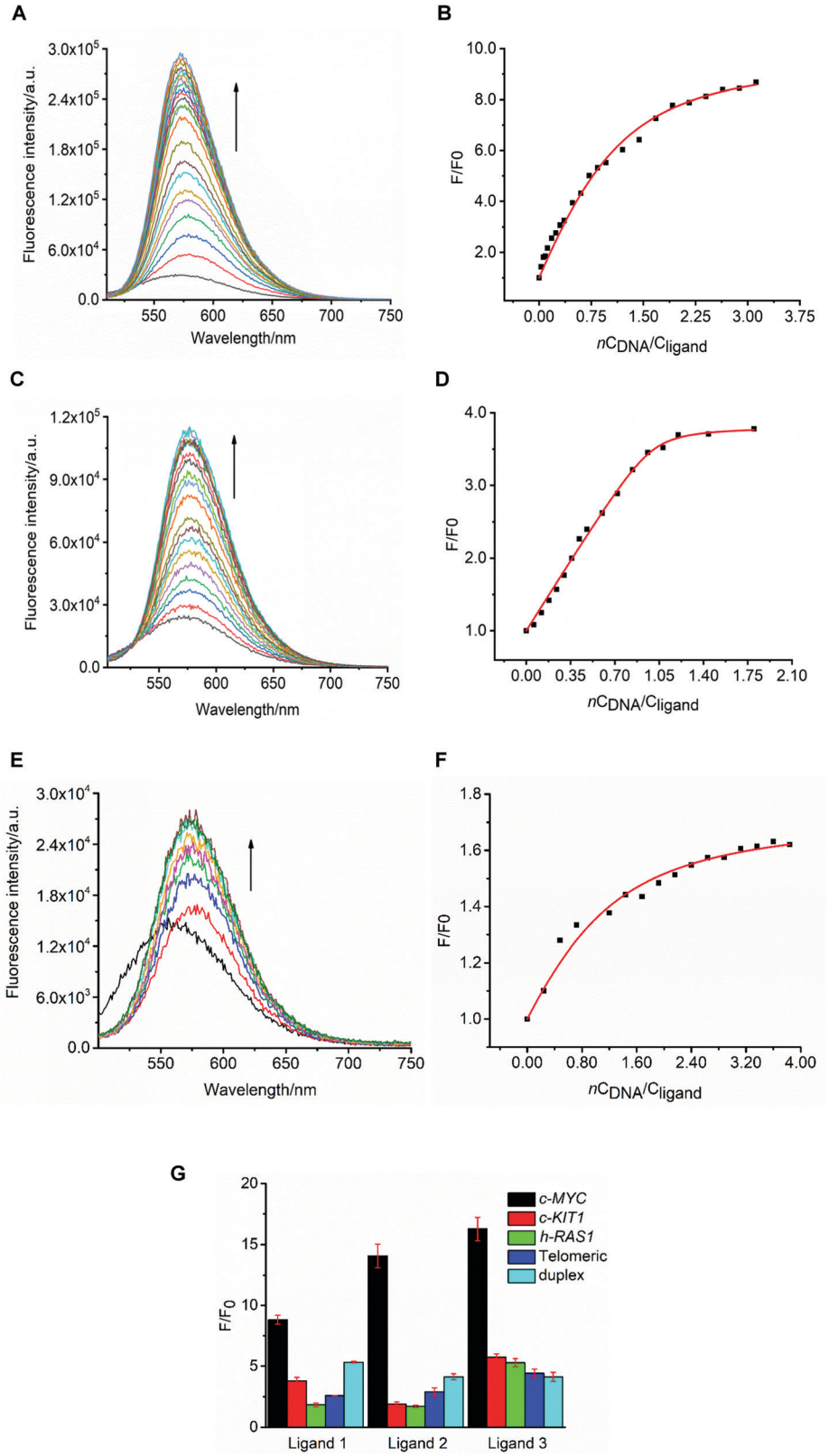


Fig. 7 Fluorescence emission spectra of ligand **1** (5 μM in 10 mM lithium cacodylate buffer, pH 7.2) along with (A) *c-MYC* G4 (0-8 μM in 1 mM KCl and 99 mM LiCl); (C) *c-KIT1* G4 (0–3 μM in 10 mM KCl and 90 mM LiCl); (E) *h-RAS1* G4 (0–7 μM in 50 mM KCl and 50 mM LiCl); plot of F/F_0 *vs.* nC_{DNA}/C_{ligand} with n = 3, for (B) *c-MYC*; (D) *c-KIT1*; and (F) *h-RAS1* and (G) bar graph of F/F_0 plots of ligands **1**, **2** and **3** with different DNAs (saturation concentration).

Table 3 Binding constants of ligand **1** with G4 and duplex DNAs measured by fluorimetric titration

	K_a^a (10^6 M^{-1})			
Ligands	*c-MYC*	*c-KIT1*	*h-RAS1*	Duplex (ds 17)
Ligand **1**	0.5 ± 0.07	4.0 ± 0.01	0.31 ± 0.04	1.1 ± 0.01
Ligand **2**	0.6 ± 0.04	1.0 ± 0.15	0.13 ± 0.01	4 ± 0.02

K_a^a represents the apparent binding constant. The experiments were carried out using 5 μM ligand **1** in 10 mM lithium cacodylate buffer, pH 7.2, along with *c-MYC* DNA (1 mM KCl and 99 mM LiCl); *c-KIT1* DNA (10 mM KCl and 90 mM LiCl); *h-RAS1* (50 mM KCl and 50 mM LiCl); telomeric DNA (10 mM KCl and 90 mM LiCl); and duplex 17 DNA (10 mM KCl and 90 mM LiCl). The emission data were then analyzed using a non-linear curve fitting equation (eqn (2))[40] (see Experimental section). The binding constant was derived from the plot of F/F_0 *vs.* nC_{DNA}/C_{ligand} with n = 3.

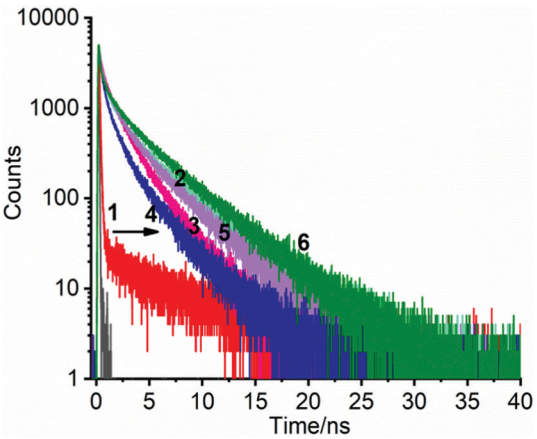


Fig. 8 Fluorescence decay traces of ligand **1** (5 μM) at 575 nm in the absence and presence of G4 DNAs in 10 mM lithium cacodylate buffer, pH 7.2, at different metal ion concentrations: (1) ligand **1**; (2) *c-MYC* DNA in 1 mM KCl and 99 mM LiCl; (3) *c-KIT1* DNA in 10 mM KCl and 90 mM LiCl; (4) *h-RAS1* in 50 mM KCl and 50 mM LiCl; (5) telomeric DNA in 10 mM KCl and 90 mM LiCl; and (6) duplex DNA in 10 mM KCl and 90 mM LiCl.

interaction of the ligand with DNA. In the unbound state, the ligands showed fast fluorescence decay, with an average lifetime of <1 ns. The lifetimes increased in the presence of different DNAs at different metal ion concentrations (Fig. 8 and Fig. S17, ESI†). The best fit was obtained using tri-exponential fitting. The decay profiles displayed three lifetimes, which are relatively slow decay time-constants, in the order of few nanoseconds (Table 4). These results indicate the presence of more than one binding site between ligand **1** and DNA which is in agreement with Job's plot analysis. Similar results were obtained with ligands **2** and **3** (Fig. S17, ESI†).

Cell based studies

Low cytotoxicity is one of the important criterion for stabilizing/sensing G4 ligands. Cytotoxicity measurements of ligand **1** in HeLa, Hep G2 and Lenti-X cells were done using the MTT assay.[74] The percentage cell viability after 24 h was calculated and plotted against concentration (Fig. S18, ESI†). Low cytotoxicity was found in HepG2 and Lenti-X (IC_{50} = 57.7 ± 11.5 μM)

Table 4 Fluorescence lifetime of ligands with different DNA forms

Ligand	DNA	τ_1 (ns) a_1 (%)	τ_2 (ns) a_2 (%)	τ_3 (ns) a_3 (%)	χ^2
Ligand **1**	*c-MYC*	1.18 (30)	4.39 (54)	0.26 (16)	1.10
	c-KIT 1	1.08 (40)	2.77 (48)	0.21 (12)	1.07
	h-RAS 1	0.93 (43)	3.09 (36)	0.18 (21)	1.06
	Telomeric	1.05 (33)	3.68 (55)	0.19 (12)	1.11
	Duplex	0.87 (17)	3.99 (74)	0.12 (9)	1.10
Ligand **2**	*c-MYC*	0.97 (29)	3.96 (56)	0.22 (15)	1.07
	c-KIT 1	1.08 (39)	2.66 (50)	0.23 (11)	1.07
	h-RAS 1	0.88 (44)	3.13 (34)	0.17 (22)	1.04
	Telomeric	1.04 (34)	3.66 (52)	0.19 (14)	1.09
	Duplex	0.92 (18)	4.20 (72)	0.13 (10)	1.12
Ligand **3**	*c-MYC*	1.49 (37)	4.34 (54)	0.38 (9)	1.02
	c-KIT 1	1.22 (24)	3.98 (72)	0.21 (4)	1.05
	h-RAS 1	1.26 (35)	3.82 (56)	0.25 (9)	1.05
	Telomeric	1.29 (33)	4.11 (61)	0.23 (6)	1.07
	Duplex	1.04 (20)	4.41 (72)	0.10 (8)	1.12

λ_{ex} = 465 nm for ligands **1** and **2**; λ_{ex} = 472 nm for ligand **3**; λ_{em} = 575 nm for ligands **1**, **2** and **3**. All the experiments were performed using 5 μM concentration of the ligands in 10 mM lithium cacodylate buffer, pH 7.2. Saturation concentration of *c-MYC* DNA in 1 mM KCl and 99 mM LiCl; *c-KIT1* DNA in 10 mM KCl and 90 mM LiCl; *h-RAS1* in 50 mM KCl and 50 mM LiCl; telomeric DNA in 10 mM KCl and 90 mM LiCl; and duplex 17 DNA in 10 mM KCl and 90 mM LiCl. The decay profiles were analyzed using tri-exponential fitting by eqn (3).

cell lines. However, moderate cytotoxicity (35.9 ± 4.9 μM) was found in the HeLa cell line.

Fluorescence enhancement of ligand **1** in the presence of DNA was observed in biophysical studies. This inspired us to investigate the intracellular localization of ligand **1** in HeLa cells by confocal fluorescence microscopy. Fixed HeLa cells were stained with ligand **1** (Fig. 9B) and nuclear staining was done using DAPI (Fig. 9A).75 As seen, ligand **1** exhibited strong localized fluorescence in the nucleoli and the cytoplasm (Fig. 9B). The stained region (Fig. 9B) corresponds to the nucleoli where rDNA undergoes transcription. Guanine-rich rDNA can also adopt temporal G4 conformations.[39] The fluorescence signal from the cytoplasm may be attributed to the binding of ligand **1** with RNA. Cytotoxicity and confocal fluorescence microscopy studies showed that the toxicity of ligand **1** is only apparent after a long period of incubation.

To verify the binding and specificity of ligand **1** with DNA or RNA, we further performed flow cytometry experiment by detecting fluorescence intensity.[76] Following DNase treatment, fluorescence due to ligand **1** binding was markedly reduced compared to that in untreated cells (Fig. 9F and Fig. S19, ESI†). However, fluorescence was attenuated by only 20% upon RNase treatment indicating preferential binding to DNA inside the cells (Fig. 9F and Fig. S19, ESI†).

TD-DFT studies

To gain more insight into the photophysical properties of the ligands, quantum chemical calculations were carried out. The three ligands were optimized at the HF/6-31G* level (Fig. S20–S22, ESI†) using Gaussian 16, revision B.01.[77] The optimized structures were utilized for TD-DFT[78] calculations using water as solvent. Six excited states were calculated and the energy and shapes of the orbitals are shown in Fig. S23 (ESI†). In all three ligands, the lowest unoccupied molecular

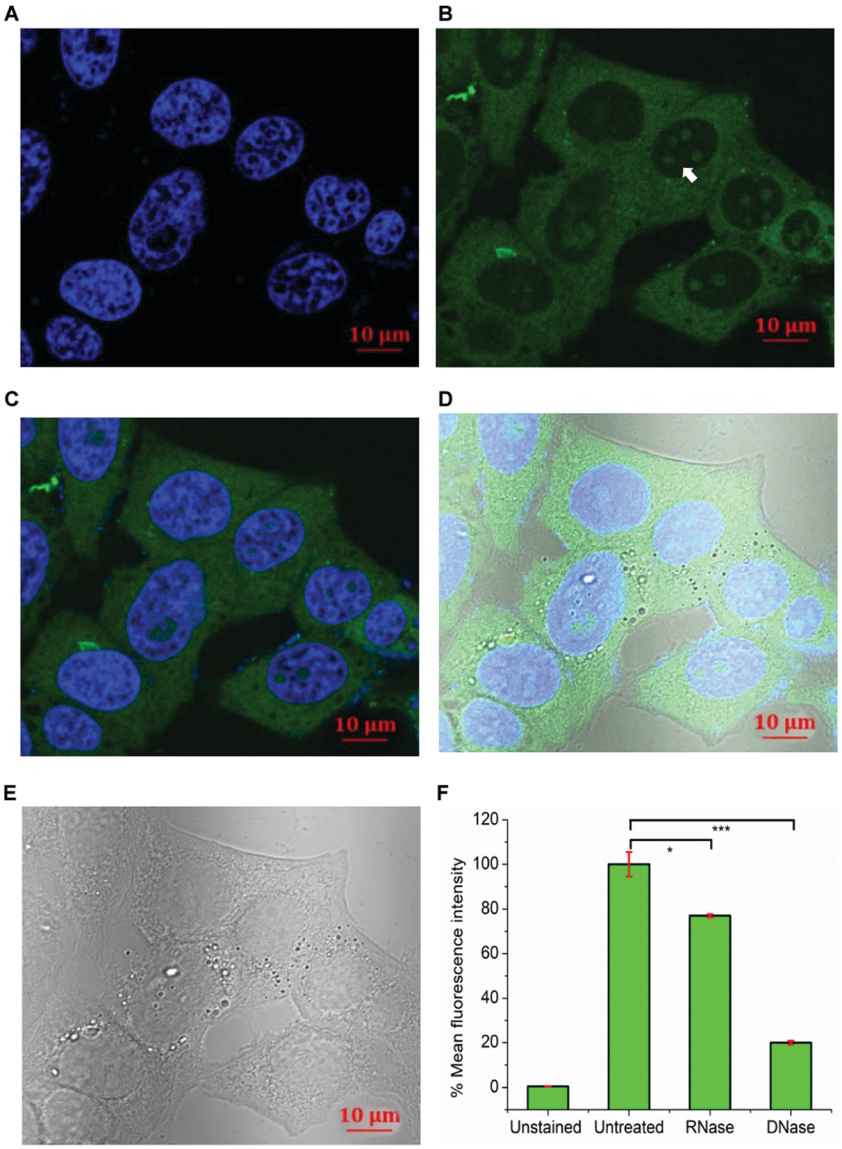


Fig. 9 Confocal fluorescence microscopy images of fixed HeLa cells co-stained with (A) DAPI (0.5 mg ml^{-1}, 5 min); (B) ligand **1** (20 μM, 4 h); (C) ligand **1**-DAPI merge; (D) ligand **1**-DAPI merge with a bright field image; (E) bright-field image; and (F) % mean fluorescence intensity of ligand **1** from flow cytometry analysis. Cells were treated independently with DNase and RNase for 30 min. The error bar denotes standard deviation (SD) for $n = 3$; (*) – p-value < 0.01 and (***) – p-value < 0.0001.

Table 5 The values obtained from TD-DFT calculations in Gaussian

Ligand	λ_{H_2O} (nm)	λ_{max} (nm)	Excitation energy (eV)
Ligand **1**	464.86	465	2.6671
Ligand **2**	465.57	465	2.6630
Ligand **3**	470.56	472	2.6348

The values were calculated at the B3LYP/6-31G* level with H_2O as solvent. λ_{H_2O} represents the wavelength of the first excitation state obtained from the quantum chemical calculation and λ_{max} represents the experimental value of the maximum absorption wavelength.

orbital (LUMO) is mainly spread out on the quinolinium ring, while the highest occupied molecular orbital (HOMO) is spread out on the indole ring. This indicates that any electronic transition can cause a redistribution of electron density and the transitions in these ligands involve electron transfer from the indole to quinolinium ring. The TD-DFT study revealed that the HOMO–LUMO gap for ligands **1**, **2**, and **3** is almost similar (Fig. S23, ESI†). The trend is also repeated in the overall excitation energy values (Table 5) The wavelengths of excitation show an excellent agreement with the experimental values (Table 5).

Molecular modeling and dynamics studies

To study the mode of binding of ligand **1** with *c-MYC* G4, molecular modeling and dynamics studies were conducted. The energy of ligand **1** was previously optimized (Fig. S20, ESI†) in the Gaussian 16 at HF/6-31G* level of theory and used for docking studies with *c-MYC* G4 DNA (PDB ID: 1XAV)[79] using

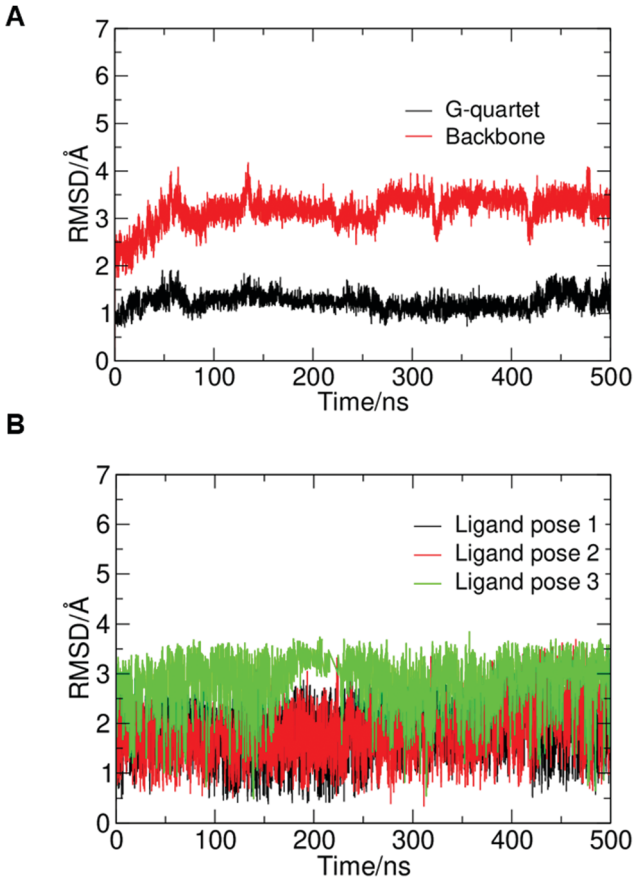


Fig. 10 Time dependent RMSD values of the *c-MYC*–ligand **1** complex: (A) RMSD of the G4 quartets and backbone and (B) RMSD of the ligand.

AutoDock 4.2.6.[80] A total of 250 independent conformers were generated after the docking. The clusters of docked conformations were analyzed to identify the various binding modes. The major modes of binding were found to be stacking at the 5′ end, partial groove binding and complete groove binding. This is in line with the experimental observation of 3 : 1 ligand binding with the *c-MYC* G4 DNA by the Jobs plot analysis. Three low energy conformers showing these binding modes were chosen as starting structures for molecular dynamics studies. The selected docked conformers of ligand **1** (Fig. S24, ESI†) after the electrostatic potential calculations in Gaussian 16 were subjected to 500 ns unrestrained simulations using the GPU accelerated version of AMBER18[81] using PMEMD.[82–84]

The conformational stability of G4 and the conformational flexibility of the ligand were investigated from the RMSD values of the heavy atoms of the ligand, backbone and G-quartets separately (Fig. 10). RMSD values were calculated with the first frame of the simulation as reference. The RMSD values of *c-MYC* quartets showed marginal variations, which indicate the stability induced in the quartets due to the ligand. The RMSD value of the DNA backbone is high and showed fluctuations, which indicate that the DNA backbone becomes flexible upon ligand binding. The RMSF values (Fig. S25, ESI†) per nucleotide in the *c-MYC* sequence showed a value greater than 1.8 Å for all the nucleotides other than the quartet forming dG nucleotides. The quartet forming dG nucleotides in *c-MYC* showed an RMSD value around 1 Å, suggesting minimal fluctuations within the G4 structure. The RMSF values agree with the RMSD fluctuations observed (Fig. 10). The ligand RMSD was also found to be stable throughout the simulation. This shows that the conformation of the ligand does not change significantly during the course of the simulation.

The 500 ns simulation trajectory was then clustered into 5 ensembles. The cluster images as well as the visual inspection of the trajectory revealed that the ligand in the complex showed 3 different binding modes – stacking at the 5′ end (pose 1), partial stacking over the G4 quartet at the 3′ end (pose 2) and groove binding (pose 3). For *c-MYC* G4, one major ensemble contributed to ~50% of the simulation time and revealed that pose 1 stabilized the top quartet (Fig. 11 and Fig. S26A, ESI†).

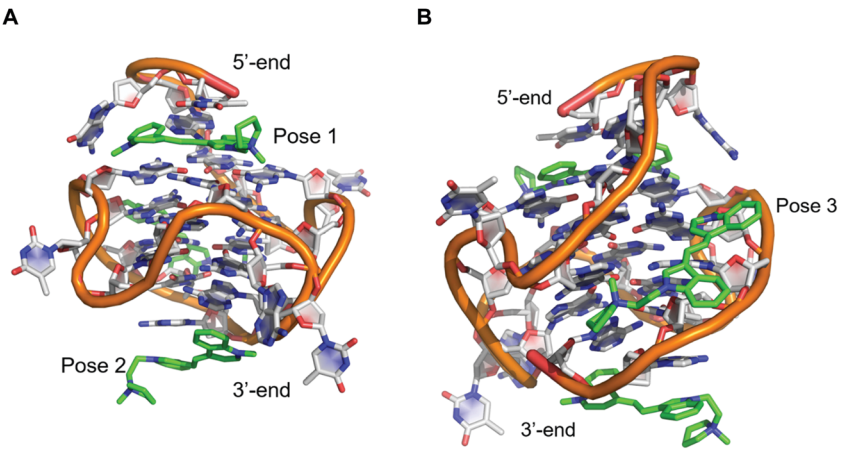


Fig. 11 Representative snapshot of the major cluster of *c-MYC* G4 DNA from MD simulations. (A) The 5′ end stacking is represented as pose 1, 3′ stacking is represented as pose 2 and (B) the groove binding is represented as pose 3 along with the stacking modes. The same snapshot is represented in two different orientations for clarity. Carbon atoms of DNA are represented using white and carbon atoms of the ligand are represented using green, nitrogen atoms using blue, oxygen atoms using red and phosphorus using orange-red.

The 5′ flanking nucleotides reoriented to accommodate the ligand and the flanking nucleotide dT1 stacked over the ligand in pose 1 (Fig. S27A, ESI†). In pose 2, the ligand stacked partially over the flanking nucleotides of the 3′ end and partially over the bottom quartet (Fig. 11). The flanking nucleotides dT20, dA21 and dA22 interacted with the ligand in this pose (Fig. S26B and S27B, ESI†). Pose 3 of the ligand, which is the groove binding mode, was stabilized by multiple weak electrostatic interactions between the positively charged nitrogen atom in the pyrrolidine ring and the negatively charged backbone atoms of G4 DNA (Fig. S27C, ESI†).

The MD trajectory was explored to account for the non-covalent interactions that exist in the ligand **1**–G4 complex. The electrostatic interaction distance between the positively charged nitrogen atoms (N1 and N3) of the ligand and the negatively charged atoms of the DNA backbone was explored. All the electrostatic interactions existed only for a short period of the total simulation time. The N1 atom of the ligand in pose 1 has a short interaction with the O6 atom of dG2 ($\sim$20% of the simulation time) (Fig. S27A, ESI†). In pose 2, the N1 atom of the ligand showed weak interactions with the OP1 and OP2 of dT11 ($\sim$10% of the simulation time) (Fig. S27B, ESI†). As mentioned earlier, the positively charged N3 atom of ligand **1** in pose 3 exhibited some weak electrostatic interactions with the OP1 and OP2 of dG6 ($\sim$10% of the simulation time) (Fig. S27C, ESI†).

The binding energy of the complex was computed using the MM-PBSA[85] module of AMBER18 (Table S2, ESI†). The total value of the binding energy was found to be -48.66 ± 10.54 kcal mol^{-1}. The total enthalpic contribution was -77.81 ± 6.08 kcal mol^{-1}, which includes the electrostatic, van der Waals and solvation energies, and the entropic contribution was -29.15 ± 9.05 kcal mol^{-1}. These values indicate that ligand binding to *c-MYC* G4 is an enthalpy driven process.

To find the selectivity of the ligand towards parallel topology, ligand **1** was docked with telomeric hybrid 2 (PDB ID: 2JPZ) and telomeric antiparallel (PDB ID: 143D) G4 structures. 250 docked conformers were generated similar to the *c-MYC* G4 DNA. Out of the docked structures, the lowest energy structures were chosen as the starting structures for the MD studies. In the case of antiparallel G4 (Fig. S28, ESI†), within the first 5 ns of simulations, the ligand moved away from the G4 groove immediately. For the next 100 ns, the ligand had almost negligible contact with the G4 structure. Though the ligand did not completely move away from the vicinity of the quadruplex, it was not interacting with the quartet or altering the loop arrangements of the G4 structure. In the case of telomeric hybrid G4 (Fig. S29, ESI†), by the end of 10 ns, the ligand moved away from the system and at around 100 ns, ligand **1** was not found in the proximity of the hybrid G4. These results are in full agreement with the CD melting studies where the ligand show preferential stabilization toward the parallel G4 structure. One reason for the selectivity can be attributed to the structure of ligand **1**. When it binds to the promoter G4, both indole and quinolinium rings reorient to adapt parallel conformation with respect to each other (Fig. 11), which facilitates stacking over the G quartet. In the case of antiparallel and hybrid G4s, the topology of the structures including the arrangement of the loops and terminal nucleotides do not allow the pyrrolidine ring to pass and reach the quartets. As a result, the fused rings of the ligand are not entirely parallel and do not stack over the quartet.

Summary and conclusions

We synthesized a series of indolyl-quinolinium based ligands with different side chains and performed various biophysical and biological studies with different G4 and duplex DNAs. CD titration and melting studies revealed that the ligands induce parallel topology of G4 structures and preferentially stabilize parallel *c-MYC* and *c-KIT1* promoter G4s, over the telomeric, *h-RAS1*, and duplex DNAs. Among the ligands studied, ligand **1** with the ethyl side chain emerged as a potent G4 stabilizer. Fluorimetric titration studies confirmed a $\sim$9 fold fluorescence enhancement of ligand **1** upon binding with *c-MYC* G4. The ligands exhibited more than one possible binding mode as shown by the TCSPC study. MD studies revealed 3 : 1 binding of ligand **1** with *c-MYC*, with both stacking and groove binding as the preferred binding modes. The unique structural features of the ligand, which impart selectivity toward parallel G4 topology, also emerged from MD studies. Cellular studies showed that ligand **1** was the least toxic and that it preferentially stained the nucleoli and the cytoplasm. Although the ligands preferentially stabilize *c-MYC* and *c-KIT1*, they are non-specific in their fluorescence properties. Our current focus is devoted to structural fine-tuning of these ligands to achieve specificity for the cellular visualization of G4 structures.

Author contributions

All authors have contributed to the manuscript. All authors have given approval to the final version of the manuscript.

Conflicts of interest

The authors declare no conflict of interest.

Acknowledgements

This work is financially supported by Science and Engineering Research Board (SERB-DST), Government of India (grant no. EMR/2016/003268) and IRCC-IIT Bombay. We acknowledge DST-FIST (grant no. SR/FST/CS-II/2017/37) for NMR and HRMS facilities; BSBE-IIT Bombay for the confocal microscope facility; and SAIF-IIT Bombay for FACS and TCSPC facilities. We acknowledge Spacetime High performance computing resources, IIT Bombay, for the computational facilities. We thank Prof. Ruchi Anand for providing access to her lab facilities. We are grateful to Prof. G. Krishnamoorthy and Prof. Nisanth N. Nair for valuable discussions. We acknowledge Dr S. Harikrishna, Praneeth Bommisetti and Dr Shubhandra

Tripathi for their extensive help in the MD studies. A. B. thanks DST-INSPIRE for the PhD fellowship. S. S. thanks PMRF for the PhD fellowship. S. P. P. thanks the Council of Scientific and Industrial Research (CSIR) for the PhD fellowship. C. S. T. thanks SERB for the fellowship and S. B. S. thanks IIT Bombay for the institute post-doctoral 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 W. E. Wright, V. M. Tesmer, K. E. Huffman, S. D. Levene and J. W. Shay, *Genes Dev.*, 1997, **11**, 2801–2809.
5 J. L. Huppert and S. Balasubramanian, *Nucleic Acids Res.*, 2007, **35**, 406–413.
6 A. Siddiqui-Jain, C. L. Grand, D. J. Bearss and L. H. Hurley, *Proc. Natl. Acad. Sci. U. S. A.*, 2002, **99**, 11593–11598.
7 S. Rankin, A. P. Reszka, J. Huppert, M. Zloh, G. N. Parkinson, A. K. Todd, S. Ladame, S. Balasubramanian and S. Neidle, *J. Am. Chem. Soc.*, 2005, **127**, 10584–10589.
8 J. Dai, D. Chen, R. A. Jones, L. H. Hurley and D. Yang, *Nucleic Acids Res.*, 2006, **34**, 5133–5144.
9 D. Sun, K. Guo, J. J. Rusche and L. H. Hurley, *Nucleic Acids Res.*, 2005, **33**, 6070–6080.
10 R. De Armond, S. Wood, D. Sun, L. H. Hurley and S. W. Ebbinghaus, *Biochemistry*, 2005, **44**, 16341–16350.
11 Y. Qin, E. M. Rezler, V. Gokhale, D. Sun and L. H. Hurley, *Nucleic Acids Res.*, 2007, **35**, 7698–7713.
12 J. Eddy and N. Maizels, *Nucleic Acids Res.*, 2008, **36**, 1321–1333.
13 D. Sen and W. Gilbert, *Nature*, 1988, **334**, 364–366.
14 J. L. Huppert, A. Bugaut, S. Kumari and S. Balasubramanian, *Nucleic Acids Res.*, 2008, **36**, 6260–6268.
15 A. Risitano and K. R. Fox, *Nucleic Acids Res.*, 2004, **32**, 2598–2606.
16 M. Cevec and J. Plavec, *Biochemistry*, 2005, **44**, 15238–15246.
17 J. L. Huppert, *Chem. Soc. Rev.*, 2008, **37**, 1375–1384.
18 J. Dai, M. Carver and D. Yang, *Biochimie*, 2008, **90**, 1172–1183.
19 Y. Chen and D. Yang, *Current protocols in nucleic acid chemistry*, 2012, vol. 50, p. 17.
20 A. T. Phan, Y. S. Modi and D. J. Patel, *J. Am. Chem. Soc.*, 2004, **126**, 8710–8716.
21 A. M. Zahler, J. R. Williamson, T. R. Cech and D. M. Prescott, *Nature*, 1991, **350**, 718–720.
22 S. Balasubramanian, L. H. Hurley and S. Neidle, *Nat. Rev. Drug Discovery*, 2011, **10**, 261–275.
23 D. Rhodes and H. J. Lipps, *Nucleic Acids Res.*, 2015, **43**, 8627–8637.
24 G. Biffi, D. Tannahill, J. McCafferty and S. Balasubramanian, *Nat. Chem.*, 2013, **5**, 182.
25 M. Di Antonio, A. Ponjavic, A. Radzevičius, R. T. Ranasinghe, M. Catalano, X. Zhang, J. Shen, L.-M. Needham, S. F. Lee and D. Klenerman, *Nat. Chem.*, 2020, **12**, 832–837.
26 A. C. Bhasikuttan and J. Mohanty, *Chem. Commun.*, 2015, **51**, 7581–7597.
27 M. Di Antonio, A. Ponjavic, A. Radzevičius, R. T. Ranasinghe, M. Catalano, X. Zhang, J. Shen, L.-M. Needham, S. F. Lee and D. Klenerman, *Nat. Chem.*, 2020, 1–6.
28 X. Chen, J. Wang, G. Jiang, G. Zu, M. Liu, L. Zhou and R. Pei, *RSC Adv.*, 2016, **6**, 70117–70123.
29 P. Chilka, P. R. Patlolla and B. Datta, *RSC Adv.s*, 2016, **6**, 87400–87404.
30 D. Lin, X. Fei, Y. Gu, C. Wang, Y. Tang, R. Li and J. Zhou, *Analyst*, 2015, **140**, 5772–5780.
31 J.-W. Yan, S.-B. Chen, H.-Y. Liu, W.-J. Ye, T.-M. Ou, J.-H. Tan, D. Li, L.-Q. Gu and Z.-S. Huang, *Chem. Commun.*, 2014, **50**, 6927–6930.
32 J. Mohanty, N. Barooah, V. Dhamodharan, S. Harikrishna, P. Pradeepkumar and A. C. Bhasikuttan, *J. Am. Chem. Soc.*, 2013, **135**, 367–376.
33 S. Zhang, H. Sun, L. Wang, Y. Liu, H. Chen, Q. Li, A. Guan, M. Liu and Y. Tang, *Nucleic Acids Res.*, 2018, **46**, 7522–7532.
34 S.-B. Chen, W.-B. Wu, M.-H. Hu, T.-M. Ou, L.-Q. Gu, J.-H. Tan and Z.-S. Huang, *Chem. Commun.*, 2014, **50**, 12173–12176.
35 H. Lai, Y. Xiao, S. Yan, F. Tian, C. Zhong, Y. Liu, X. Weng and X. Zhou, *Analyst*, 2014, **139**, 1834–1838.
36 P. Yang, A. De Cian, M. P. Teulade-Fichou, J. L. Mergny and D. Monchaud, *Angew.Chem., Int. Ed.*, 2009, **48**, 2188–2191.
37 M.-Q. Wang, S. Liu, C.-P. Tang, A. Raza, S. Li, L.-X. Gao, J. Sun and S.-P. Guo, *Dyes Pigm.*, 2017, **136**, 78–84.
38 Y.-J. Lu, Q. Deng, D.-P. Hu, Z.-Y. Wang, B.-H. Huang, Z.-Y. Du, Y.-X. Fang, W.-L. Wong, K. Zhang and C.-F. Chow, *Chem. Commun.*, 2015, **51**, 15241–15244.
39 Y.-J. Lu, Q. Deng, J.-Q. Hou, D.-P. Hu, Z.-Y. Wang, K. Zhang, L. G. Luyt, W.-L. Wong and C.-F. Chow, *ACS Chem. Biol.*, 2016, **11**, 1019–1029.
40 M.-Q. Wang, J. Xu, L. Zhang, Y. Liao, H. Wei, Y.-Y. Yin, Q. Liu and Y. Zhang, *Bioorg. Med. Chem.*, 2019, **27**, 552–559.
41 G. Zhurko, ChemCraft software, version 1.6, 2014.
42 J. Dai, M. Carver, C. Punchihewa, R. A. Jones and D. Yang, *Nucleic Acids Res.*, 2007, **35**, 4927–4940.
43 Y. Wang and D. J. Patel, *Structure*, 1993, **1**, 263–282.
44 T. Fox and P. A. Kollman, *J.Phys. Chem. B*, 1998, **102**, 8070–8079.
45 J. Wang, W. Wang, P. A. Kollman and D. A. Case, *J. Mol. Graph. Model.*, 2006, **25**, 247–260.
46 S. Haider and S. Neidle, *G-Quadruplex DNA*, Springer, 2010, pp. 17–37.
47 J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman and D. A. Case, *J.Comput. Chem.*, 2004, **25**, 1157–1174.
48 J. Wang, P. Cieplak and P. A. Kollman, *J.Comput. Chem.*, 2000, **21**, 1049–1074.
49 A. Pérez, I. Marchán, D. Svozil, J. Sponer, T. Cheatham III, C. Laughton and M. Orozco, *Biophys. J.*, 2007, **92**, 3817–3829.

50 M. Zgarbová, J. Sponer, M. Otyepka, T. E. Cheatham III, R. Galindo-Murillo and P. Jurecka, *J. Chem. Theory Comput.*, 2015, **11**, 5723–5736.
51 M. Krepl, M. Zgarbová, P. Stadlbauer, M. Otyepka, P. Banáš, J. Koca, T. E. Cheatham III, P. Jurecka and J. Sponer, *J. Chem. Theory Comput.*, 2012, **8**, 2506–2520.
52 W. D. Cornell, P. Cieplak, C. I. Bayly, I. R. Gould, K. M. Merz, D. M. Ferguson, D. C. Spellmeyer, T. Fox, J. W. Caldwell and P. A. Kollman, *J. Am. Chem. Soc.*, 1995, **117**, 5179–5197.
53 D. R. Roe and T. E. Cheatham III, *J. Chem. Theory Comput.*, 2013, **9**, 3084–3095.
54 W. Humphrey, A. Dalke and K. Schulten, *J. Mol. Graph*, 1996, **14**, 33–38.
55 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.
56 Y.-J. Lu, D.-P. Hu, K. Zhang, W.-L. Wong and C.-F. Chow, *Biosens. Bioelectron.*, 2016, **81**, 373–381.
57 Y.-J. Lu, X.-L. Guo, M.-H. Xu, W.-W. Chen, W.-L. Wong, K. Zhang and C.-F. Chow, *Dyes Pigm.*, 2017, **143**, 331–341.
58 N. Sun, C. Wang, M.-H. Xu, Y.-J. Lu, Y.-Y. Zheng, Y. Yan, X.-L. Guo, J. Hou, K. Zhang and L. G. Luyt, *Sens. Actuators, B*, 2017, **250**, 543–551.
59 X. Xie, M. Zuffo, M.-P. Teulade-Fichou and A. Granzhan, *Beilstein J. Org. Chem.*, 2019, **15**, 1872–1889.
60 B. Kumari, A. Yadav, S. P. Pany, P. Pradeepkumar and S. Kanvah, *J. Photochem. Photobiol., B*, 2019, **190**, 128–136.
61 Y. Li, D. Xu, S.-L. Ho, H.-W. Li, R. Yang and M. S. Wong, *Biomaterials*, 2016, **94**, 84–92.
62 F.-C. Meng, F. Mao, W.-J. Shan, F. Qin, L. Huang and X.-S. Li, *Bioorg. Med. Chem. Lett.*, 2012, **22**, 4462–4466.
63 S. Paramasivan, I. Rujan and P. H. Bolton, *Methods*, 2007, **43**, 324–331.
64 B. Nordén and T. Kurucsev, *J. Mol. Recognit.*, 1994, **7**, 141–155.
65 J. Dash, P. S. Shirude, S.-T. D. Hsu and S. Balasubramanian, *J. Am.Chem. Soc.y*, 2008, **130**, 15950–15956.
66 E. M. Rezler, J. Seenisamy, S. Bashyam, M.-Y. Kim, E. White, W. D. Wilson and L. H. Hurley, *J. Am. Chem. Soc.*, 2005, **127**, 9439–9447.
67 A. Membrino, S. Cogoi, E. B. Pedersen and L. E. Xodo, *PLoS One*, 2011, **6**(9), e24421.
68 A. Guédin, L. Lacroix and J.-L. Mergny, *Drug–DNA interaction protocols*, Springer, 2010, pp. 25–35.
69 A. Mishra, R. K. Behera, P. K. Behera, B. K. Mishra and G. B. Behera, *Chem. Rev.*, 2000, **100**, 1973–2012.
70 B. Song, Q. Zhang, W.-H. Ma, X.-J. Peng, X.-M. Fu and B.-S. Wang, *Dyes Pigm.*, 2009, **82**, 396–400.
71 M. Pordel, H. Chegini, S. Ramezani and M. Daee, *J. Mol. Struct.*, 2017, **1129**, 105–112.
72 A. M. Brouwer, *Pure Appl. Chem.*, 2011, **83**, 2213–2228.
73 N. T. Dao, R. Haselsberger, M.-E. Michel-Beyerle and A. T. Phan, *FEBS Lett.*, 2011, **585**, 3969–3977.
74 D. M. Morgan, *Polyamine protocols*, Springer, 1998, pp. 179–184.
75 A. Laguerre, J. M. Wong and D. Monchaud, *Sci. Rep.*, 2016, **6**, 32141.
76 J. Figueiredo, J. Lopes-Nunes, J. Carvalho, F. Antunes, M. Ribeiro, M. P. C. Campello, A. Paulo, A. Paiva, G. F. Salgado and J. A. Queiroz, *Int. J. Pharm.*, 2019, **568**, 118511.
77 M. Frisch, G. Trucks, H. Schlegel, G. Scuseria, M. Robb, J. Cheeseman, G. Scalmani, V. Barone, G. Petersson and H. Nakatsuji, *Gaussian 16*, 2016.
78 E. Runge and E. K. Gross, *Phys. Rev. Lett*, 1984, **52**, 997.
79 A. Ambrus, D. Chen, J. Dai, R. A. Jones and D. J. B. Yang, *Biochemistry*, 2005, **44**, 2048–2058.
80 G. M. Morris, R. Huey, W. Lindstrom, M. F. Sanner, R. K. Belew, D. S. Goodsell and A. J. Olson, *J. Comput. Chem.*, 2009, **30**, 2785–2791.
81 D. A. Case, I. Y. Ben-Shalom, S. R. Brozell, D. S. Cerutti, T. E. Cheatham III, V. W. D. Cruzerio, T. A. Darden, R. E. Duke, D. Ghoreishi, M. K. Gilson, H. Gohlke, A. W. Goetz, D. Greene, R. Harris, N. Homeyer, T. Kurtzman, A. Kovalenko, T. S. Lee, S. LeGrand, P. Li, C. Lin, J. Liu, T. Luchko, D. J. Luo, R. Mermelstein, K. M. Merz, Y. Miao, G. Monard, C. Nguyen, H. T. Nguyen, I. Omelyan, A. Onufriev, F. Pan, R. Qi, D. R. Roe, A. Roitberg, C. Sagui, S. Schott-Verdugo, J. Shen, C. L. Simmerling, J. Smith, R. Salomon-Ferrer, J. Swails, R. C. Walker, J. Wang, H. Wei, R. M. Wolf, X. Wu, L. Xiao, D. M. York and P. A. Kollman, *AMBER*, University of California, San Francisco, 2018.
82 R. Salomon-Ferrer, A. W. Götz, D. Poole, S. Le Grand and R. C. Walker, *J. Chem. Theory Comput.*, 2013, **9**, 3878–3888.
83 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.
84 S. Le Grand, A. W. Götz and R. C. Walker, *Comput. Phys. Commun.*, 2013, **184**, 374–380.
85 P. A. Kollman, I. Massova, C. Reyes, B. Kuhn, S. Huo, L. Chong, M. Lee, T. Lee, Y. Duan and W. Wang, *Acc. Chem. Res.*, 2000, **33**, 889–897.