

High throughput screening of a new fluorescent G-quadruplex ligand having telomerase inhibitory activity in human A549 cells.

Authors' name and affiliations:

Sourav Ghosh^{‡a}, Debapriya De ^{‡a}, Victor Banerjee^a, Soumyajit Biswas^a, Utpal Ghosh^{a*}

^aDepartment of Biochemistry & Biophysics

University of Kalyani, Kalyani- 741235, India

[‡]These authors contributed equally.

***Corresponding author:** Utpal Ghosh, Ph.D.

Department of Biochemistry & Biophysics

University of Kalyani, Kalyani- 741235, India

Fax No. 091 033-2582-8282

Phone No. +91-8420520334

Email: utpal8ghosh@gmail.com

Highlights

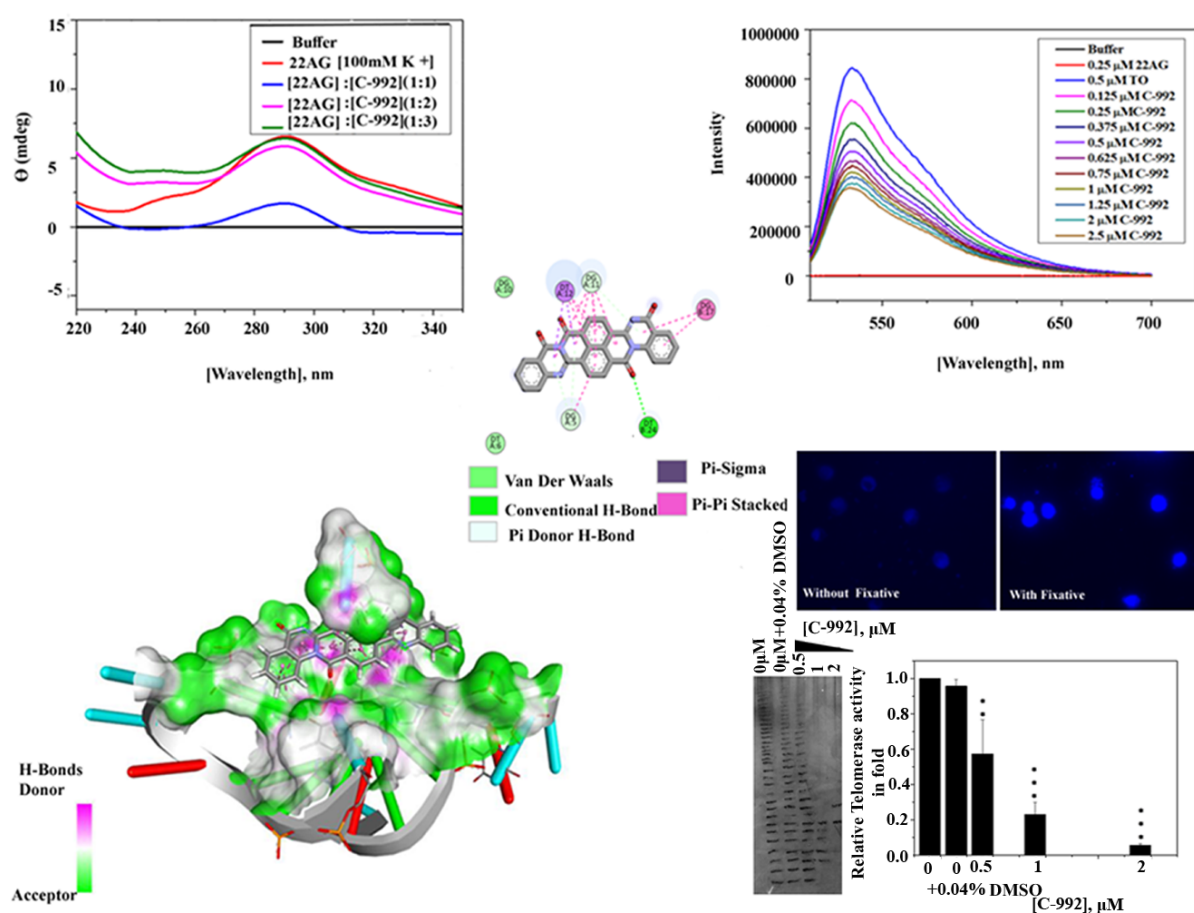
- A set of compounds were screened randomly by a High throughput method and Lipinski's rule of five from ZINC database to identify potential G4-binders.
- The compound (ZINC ID-05220992) was screened after docking with five G-quadruplex DNA
- It binds G-quadruplex DNA 22AG as detected by TO displacement and CD spectroscopy.
- It inhibits telomerase activity in A549 cells and also cytotoxic to this cell.
- It penetrates live A549 cell in culture and stains it with blue fluorescence

Abstract:

Genome-wide analysis showed that putative G-quadruplex DNA structures are prevalent in the human genome. The presence of G-quadruplex structure in the telomere and promoter region of certain oncogenes inspired people to use G-quadruplex ligand as anti-cancer agents. G-quadruplex structures, stabilized by ligand at telomere are resolved by telomerase making the cancer cells resistant to G-quadruplex ligand. So, identification of a new G-quadruplex ligand having anti-telomerase activity would be a promising strategy for cancer therapy as about 85% of human cancers are telomerase positive. A set of the drug-like compounds were screened from the ZINC database randomly and 2284 ligands were chosen following Lipinski's rule of five that were docked with five different G-quadruplex DNA sequences in idock. We screened 43 potential G-quadruplex binders using Z-score as a normalization scoring function. The compound (ZINC ID- 05220992) gave the best score (average idock = -10.17 kcal/mol, average normalized idock = -3.42). We performed G4 FID assay, CD analysis to understand its binding with three different G-quadruplex DNA sequences, and

checked its anti-telomerase activity in A549 cells using TRAP assay. We observed that this compound had an intrinsic fluorescence, capability to stain live cells with a blue fluorescence, and a specific affinity to only 22AG out of three different G-quadruplex DNA sequences under study. It showed cytotoxicity, good permeability to live cells, and a significant reduction of telomerase activity in human A549 cells at a very low dose. So, this compound has strong potential to be an anti-cancer drug.

Graphical Abstract:



Keywords:

ZINC database; Molecular Docking; G4 FID assay; Telomerase activity; TRAP Assay.

Introduction:

Despite the tremendous advancement of new molecular targets like proteins and RNA for cancer therapy, DNA targets are still the mainstay of cancer treatment. Since the discovery of the G-quadruplex DNA structure, designing and searching for novel molecules targeting G-quadruplex DNA has been increased for anti-cancer therapy [1]. The G-quadruplex structure is formed when the G-rich sequence of a DNA can fold back to form a specialized three-dimensional structure where four guanines (G) are interconnected by Hoogsteen base pairing in a plane in the presence of monovalent cations like Na^+ or K^+ and three or more such parallel planes are stacked one after another. Either strand or both the strands of DNA of a chromosome (intra-molecular) or DNA strands from two different chromosomes (intermolecular) can form a variety of G-quadruplex structures depending on the presence of different monovalent ion concentrations, length of G-track, and length of loop interspaced G-track [2,3]. There could be a large variety of G-quadruplex DNA concerning structural and topological conformations as reviewed [4]. Detection of *in-vitro* and *in-vivo* ligand binding with various topologically polymorphic G-quadruplex DNA can effectively be monitored by measuring various fluorescence properties like quantum yield, lifetime, anisotropy of fluorescence of the ligand [5,6]. The ligand which can stabilize this G-quadruplex structure produces DNA damage, apoptosis, and cell cycle arrest when treated with cells [7,8]. The first G-quadruplex DNA structure was discovered in the telomere region [9]. Since the replenishment of telomere erosion is urgently required for highly dividing cancer cells, the interference of replenishment of telomere shortening was exploited in anti-cancer therapy using G-quadruplex ligand which inhibits telomerase to reverse the shortening of telomere and thereby induce genome instability [10–12]. Later on, it has been observed that G-quadruplex DNA is not only found in the telomere region but also distributed throughout the

genome, especially at the promoter region of various oncogenes like c-MYC, c-KIT, K-RAS, hTERT, BCL-2, VEGF, HIF-1, C-MYB, PDGF-A, pRb which are responsible for cell proliferation, tumorigenesis, and anti-apoptotic effects [13]. This information makes this field more attractive for anti-cancer therapy. Even G-rich telomeric non-coding RNA (TERRA) can have various G-quadruplex structures [14]. Therefore, not only G-quadruplex DNA, G-quadruplex RNA, and G-quadruplex of DNA-RNA hybrid is also possible [4]. This makes the designing and screening of targeted G-quadruplex ligand very challenging. So far a large set of molecules have been synthesized targeting various G-quadruplex DNA structures but very few came into the clinical trial. The possible reasons are the existence of a wide variety of polymorphic G-quadruplex structures, environment-dependent DNA topology due to which *in-vitro* studies do not match with *in-vivo* studies, lack of knowledge of *in-vivo* conformation of G-quadruplex DNA, and ability of telomerase enzyme to resolve G-quadruplex DNA at telomere stabilized by ligand [15–17] etc. Therefore, identification of G-quadruplex ligand having good anti-telomerase activity *in-vivo* would be a more effective strategy for cancer therapy.

Here, we report a new G-quadruplex ligand (ZINC ID: 05220992) screened from the ZINC database based on Lipinski's rule of five and the Z-score value of its binding with several G-quadruplex DNA using docking studies. Then the binding of this compound with G-quadruplex DNA was experimentally verified. The compound has intrinsic fluorescence property and capable of fluorescence staining of living cultured human cancer cells, has a strong binding affinity with a specific G-quadruplex DNA present in telomere, potent anti-telomerase activity.

Materials and methods

Materials: The compound ZINC ID:05220992 henceforth named C-992 was purchased from MolProt, Latvia. Thiazole Orange (TO) was obtained from Sigma Aldrich. Tris base, DMSO, KCl were purchased from SRL India. Three different G-quadruplex DNA sequences such as 22AG (5'-AGGGTTAGGGTTAGGGTTAGGG-3'), Telo24 (5'-TAGGGTTAGGGTTAGGGTTAGGGT-3'), Telo28 (5'-TTAGGGTTTAGGGT TTAGGGTTTAGGGT- 3'), and the primers for TRAP assay such as CX primer (5'-CCCTTACCCTTACCCTTACCCTAA-3'), TS primer (5'-AATCCGTCGAGCAGAGTT-3') were procured from Eurofins, India.

Calf thymus DNA (CT-DNA) used as the control was obtained from Bangalore GeNei, India. Other fine chemicals were purchased locally.

Methods:

Screening of new G-quadruplex ligand from ZINC database using molecular docking:

Five telomeric G-quadruplex structures were selected for virtual screening targets from Protein Data Bank (PDB) as given in Table 1 [18]. Co-crystallized ligands and water atoms were removed manually from the PDB structure of these five receptor molecules and then converted to PDBQT format from PDB using AutoDockTools for docking purposes [19]. A set of compounds were screened from 78890 compounds from the ZINC database randomly. From this set, about 2284 compounds were chosen based on Lipinski's rule and saved in SDF format [19,20]. RdKit was used to add hydrogen and energy minimization of molecules for preprocessing of ligands using Knime workflow [21]. Then Open Babel was used to converting the SDF to PDBQT format [21]. The free and open-source docking software idock 2.2.3 was used for the docking of 2284 selected molecules at the reported ligands binding site of the five target receptor G-quadruplex DNA [22]. The details of the docking parameter of

idock are given in the supplementary file S1. The binding energy and binding affinity were predicted using RF-score (pKd) and idock score (kcal/mol). All the receptors were docked with 20x20x20 grid points in the x, y, and z directions at the region of the receptor where inhibitor binds as reported in PDB. The low idock score (binding energy) and high RF-score (estimation of intermolecular binding affinity) is an indication of the possible potential of the G-quadruplex ligand. For normalization of the all receptor binding affinity, we have used Z-score (standard score) ligand q to a given receptor p and was calculated using RF-score and idock scoring function by following equations:

$$Z_{pq} = \frac{E_{pq} - E_p}{sd_p}$$

Where E_{pq} = The idock score or RF score of ligand q docking to receptor p. E_p = Mean idock score or RF score value for the receptor p, sd_p = Standard deviation of idock score or RF score of the receptor p.

A large negative value of Z_{pq} signifies the higher affinity for the ligand to the receptor. Then we categorized ligands into four groups as per Z-score values. The high G4 binder having Z-score < -2.0, the medium G4 binder having Z-score between -2.0 to -1.0, the low G4 binder having Z-score between -0.1 to 0.0, and the non-G4 binder having Z-score 0.0 to 0.1 as shown in Table 2. Finally, the compounds were sorted in ascending order according to their calculated binding energy averaged across the five receptor molecules. Discovery Studio was used for the visualization of G-quadruplex-ligands interaction.

Cell Culture: A549 cells were obtained from NCCS, Pune and cultured in Dulbecco's modified Eagle's medium (HiMedia, India) supplemented with 10% heat-inactivated Fetal Bovine Serum (HiMedia, India) at 37 °C with 5% CO₂ under humid conditions.

Preparation of solutions of G-quadruplex DNA, ligand (C-992), and Thiazole Orange (TO):

The G-quadruplex DNA was dissolved in a Benchmark buffer containing 10 mM Tris-Cl, pH 7.2, and 100 mM KCl. The stock concentration was 500 μ M. Then several aliquots of 50 μ M in Benchmark buffer were prepared from this and stored at -20 $^{\circ}$ C for future use. Before the experiment, one aliquot was taken and heated at 95 $^{\circ}$ C for 5 min followed by chilled in ice to favor the intramolecular folding to form a specialized G-quadruplex structure. The final concentration of the DNA in the reaction mixture was 0.25 μ M. A stock solution of 1 mM Thiazole Orange (TO) (Sigma, India) was prepared in DMSO. Then a sub stock of 70 μ M was prepared in Benchmark buffer. The final concentration of TO in the reaction mixture was 0.5 μ M for G-quadruplex DNA and 0.75 μ M for CT-DNA during the G4-FID assay. Similarly, a stock solution of 5 mM of the compound C-992 was prepared in DMSO and stored at -20 $^{\circ}$ C. Then sub stocks were freshly prepared in Benchmark buffer before the experiment. The concentration range of C-992 used during titration was 0.125 μ M to 2.5 μ M.

Spectroscopic method:

G4 FID assay: The affinity of the compound C-992 towards G-quadruplex DNA can be measured by a simple fluorescence-based method known as Fluorescence Intercalator Displacement assay (FID) or G4-FID assay[23]. Here, the binding affinity of the C-992 was detected by the displacement of the fluorescent probe TO from the TO-saturated G-quadruplex DNA. The G-quadruplex DNA structure was prepared in presence of K^{+} ion as described above. TO has a strong affinity to this particular G-quadruplex structure. TO-saturated G-quadruplex DNA gives maximum fluorescence in the emission range 510 nm-750 nm with excitation at 501 nm. So, the gradual adding of C-992 to TO-saturated G-quadruplex DNA decreased fluorescence, and the percentage of TO displacement was determined by the following equation

$${}^{\text{TO}}D_x = 100 - [(F_{A_x}/F_{A_l}) \times 100]$$

Where $^{TO}D_x$ = Percentage of TO displacement for each concentration of compound; FA_x = fluorescence area of spectrum recorded after the addition of various concentrations of the ligand; FA_1 = fluorescence area of the spectrum recorded after the addition of TO.

Intrinsic fluorescence of ligand C-992: We checked first whether the ligand C-992 has any intrinsic fluorescence property or not. First, the absorption spectra of the C-992 in Benchmark buffer were recorded in Cary Series UV-Vis spectrophotometer, (Agilent Technologies, India). Then emission spectra of the C-992 were recorded in the range 420 nm to 600 nm after excitation at its absorption maxima 370 nm in Cary Eclipse Fluorescence Spectrophotometer (Agilent Technologies, India).

Circular dichroism (CD spectra): CD spectra of G-quadruplex DNA 22AG with and without C-992 were recorded in the Applied Photophysics Chirascan spectrophotometer at a scan speed of 50 nm/min and over the range of 220 nm to 400 nm. The sample was taken in a quartz cuvette with a 2 mm path length. CD spectra of G-quadruplex in the presence of ligands (at [DNA] : [compound] ratios 1:1, 1:2 and 1:3) in Benchmark buffer (10 mM Tris-Cl, pH 7.2 and 100 mM KCl) was recorded.

Cell permeability and cell imaging by fluorescence microscope: The penetrations of C-992 in living and fixed cultured human cells were checked as per our earlier report [24]. In brief, human lung cancer A549 cells were grown on cover slips placed inside the Petri plate for 24 h. Then one set of the cover slip containing the adherent living cells was treated with 20 nM C-992 for 30 min in dark (without fixation). Another set of cover slips having adherent cells was first fixed with 3:1 methanol and glacial acetic acid at 4 °C for 2 h (with fixation). Then it was treated with the same concentration of 20 nM of C-992 for 30 min in dark. Further,

cells after dislodgement from the plate by trypsinization were also tested for permeability experiment. A set of cells were dislodged from the plate by trypsinization divided into two groups. One group was treated with 20 nM of C-992 for 30 min in dark (without fixation). Another group was fixed first with 70% ethanol at 4 °C for 2 h (with fixation). Then both adherent and dislodged cells with and without fixation were observed under the 40X objective of Axio Scope A1 microscope (Carl Zeiss, Germany). Bright-field and fluorescence images were captured by Axio Scope A1 (data not shown).

MTT assay: Cell viability assay was done by standard method [25]. In brief, about 3000 cells were seeded in each well of a 96 well plate in triplicate. The cells were treated with different concentrations of compound C-992 after 24 h of cell seeding. Cells were allowed to incubate at 37 °C for 46 h in presence of C-992. Then the medium was replaced by fresh DMEM containing 0.2 mg/ml MTT and allowed to incubate in dark for a further 2 h. Then the medium was discarded and DMSO was added to it. The absorbance was measured in a Hitachi U-290 spectrophotometer at 595 nm.

Telomerase Repeat Amplification Protocol or TRAP Assay: Telomerase activity assay was performed as per our earlier report [26].

In brief, A549 cells were grown at 60-80% confluence. Then cells were treated with 0.5-2 μ M of C-992 for 24 h. A control set was also made where cells remained untreated. Since the compound stock solution was prepared in DMSO, so another set was prepared by treating the cells with 0.04% DMSO for 24 h. After treatment, cells were trypsinized and washed twice with cold DEPC treated 1X PBS. Then cells were again washed with wash buffer followed by centrifugation at 3000 rpm for 2 min at 4 °C. Then the whole-cell extract was prepared using lysis buffer (0.5% CHAPS, 10 mM Tris-Cl, pH=7.5, 1 mM MgCl₂, 1 mM EGTA, 5 mM β -mercaptoethanol, 10% glycerol) and was kept at 0 °C for 30 min. Next, the whole-cell extract

was cleared by centrifugation at 13500 g for 15 min. Then, the supernatant was collected in a fresh tube. Protein content was measured by Lowry's method.

Telomerase substrate (TS) [5'-AATCCGTCGAGCAGAGTT-3'] was allowed to extend by 0.9 µg of cell extract for 40 minutes at 30 °C temperature in TRAP reaction mix (20 mM Tris-HCl, 1.5 mM MgCl₂, 78 mM KCl, 0.005% Tween-20, 1 mM EGTA) in a total volume of 25 µl. The telomeric repeat was amplified by PCR using CX primer [5'-(CCCTTA)₃CCCTAA-3']. The hot start Taq pol was activated at 95 °C for 5 min followed by 35 cycles of 30 s at 95 °C, 45 s at 50 °C, 1 min 30 s at 72 °C and final extension at 72 °C for 10 min. Then telomeric repeats were resolved in 10% non-denaturing PAGE and visualized after silver staining of the bands.

The intensity of each band in the control lane was added and considered as 100% telomerase activity. Accordingly, the bands of each lane were measured and normalized for the control lane. The experiment was repeated at least thrice and the statistical significance of the treated samples was evaluated by IBM SPSS Statistics version 21 software using ANOVA with Dunnett's test. The p-values were denoted as '*' (0.01 < p ≤ 0.05), '**' (0.001 < p ≤ 0.01) and '***' (p ≤ 0.001).

Result & Discussion:

Virtual screening of compounds having the potential to bind G-quadruplex DNA from ZINC database by molecular docking using idock:

As described in the method section, we started from 78890 compounds in the ZINC database and screened a subset of compounds following Lipinski's rule of five and from this, a subset of about 2284 randomly chosen drug-like molecules was docked with five different G-quadruplex receptors DNA at specified region. A wide range of idock score (binding energy)

and RF score (intermolecular binding affinity) was obtained for all five receptor molecules as given in Table 1.

Table-1: Summary of iDock score and RF-Scores after binding G-quadruplex ligands with receptors G-quadruplex DNA

PDB	iiDock score (kcal/mol)			RF-Score (pKd)		
	Mean	Min	Max	Mean	Min	Max
1NZM - 0	-8.552371991	-11.6	-5.67	6.038148796	4.4	7.91
1NZM - 1	-10.43320788	-15.42	-3.28	6.918490153	4.91	8.17
2HRI - 0	-5.756652079	-7.99	-3.79	4.590310722	3.19	6.58
2HRI - 1	-4.57854267	-6.15	-3.13	3.94854267	3.04	5.27
2JWQ - 0	-7.402455142	-10.61	-5	5.646673961	3.99	7.34
2JWQ - 1	-8.406617068	-12.67	-4.92	6.680866521	4.41	8.29
3CDM - 0	-6.330568928	-10.28	-4.47	5.172984683	3.57	7.1
3CE5 - 0	-8.565006565	-13.3	-6.16	5.70873523	3.93	7.55

For example, 1NZM has two docking sites and showed idock score range -11.6 kcal/mol to -5.67 kcal/mol with a mean value at -8.55 kcal/mol and -15.42 kcal/mol to -3.28 kcal/mol with a mean value at -10.43 kcal/mol for site 1 and 2 respectively. The frequency distribution of all receptor's idock score, RF score, normalized idock score, and normalized RF score are shown in the matrix position [1,1], [2,2], [3,3], and [4,4] respectively in Fig 1. The normalized idock score [3,3] and normalized RF score [4,4] indicated normalized value distribution calculated by Z-score. The upper and lower off-diagonal positions of the figure showed a correlation between four scoring methods. A high negative correlation value (-0.85) between idock score and RF score for all five receptors was obtained. To bring the scores of all five G-quadruplex DNA receptors on the same scale, we used a Z-score which

normalized the score for all five receptors by making each of their mean to zero as shown in matrix position [3,3] and [4,4] in Fig 1.

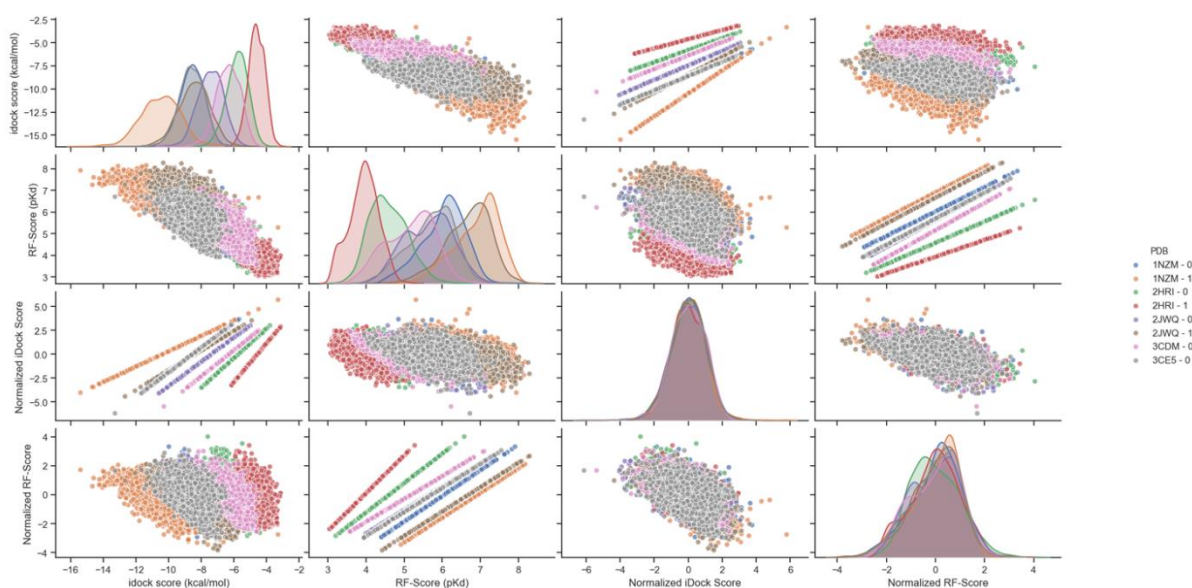


Figure 1: idock scores: The frequency distribution of idock score, RF score, normalized idock score, and normalized RF score of ligand-receptor interaction.

Then we grouped the ligands into four categories as described above in the method section as shown in Table 2.

Table 2: The ligands are categorized into four groups based on Z-score.

Results	Average idock score (kcal/mol)			Average RF score (pKd)			Count	Z-score
	Max	Mean	Min	Max	Mean	Min		
High G4 Binder	-8.2875	-	-	7.0425	6.160610465	5.4975	43	< -2.0
Medium G4Binder	-	-	-9.2625	6.79	5.911588235	5.0475	255	-2.0 to -1.0
Low G4 Binder	-7.195	-	-8.70125	6.89	5.764914773	4.76625	792	-1.0 to 0.0

Non G4	-	-	-8.05875	6.605	5.380951633	4.29625	1194	> 0
Binder	5.28875	7.000082705						

About 43 ligands having Z-score less than -2.0 categorized as High G-quadruplex DNA binders (G4 binder). The frequency distribution of average idock score and average RF score for all categories is shown in matrix positions [1,1] and [2,2] respectively in Fig 2. All the 43 High G4DNA binders showed an average idock score in the range of -8.2 kcal/mol to -10.17 kcal/mol and an average RF score in the range of 5.49 pKd to 7.04 pKd as shown in Table 2. The Scatter plot and 2D density distribution plot as shown in matrix position [2,1] and [1,2] respectively in Fig 2, distinctly indicated the distribution of four types of G-quadruplex binders with respective average idock and RF score.

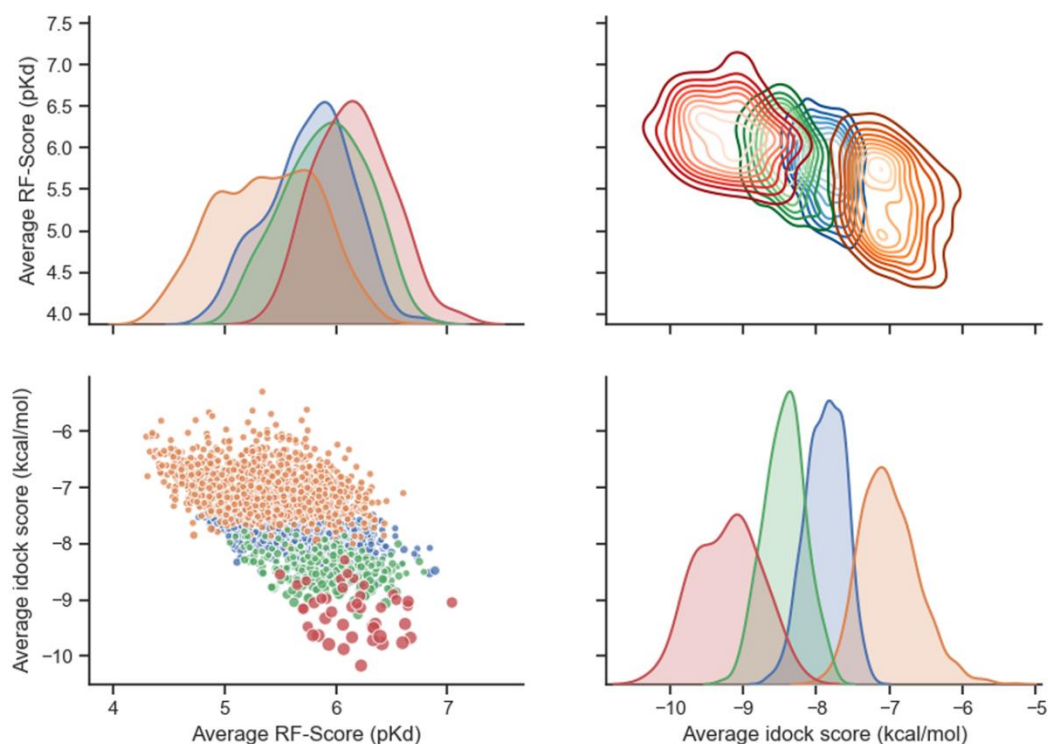


Figure 2: The frequency distribution of average idock score and average RF score for four category ligands. Each category ligand shows a distinct color spot or contour area.

Distinct chemical properties of the High G4DNA binder were observed among the four categories of ligands as shown graphically in Fig 3.

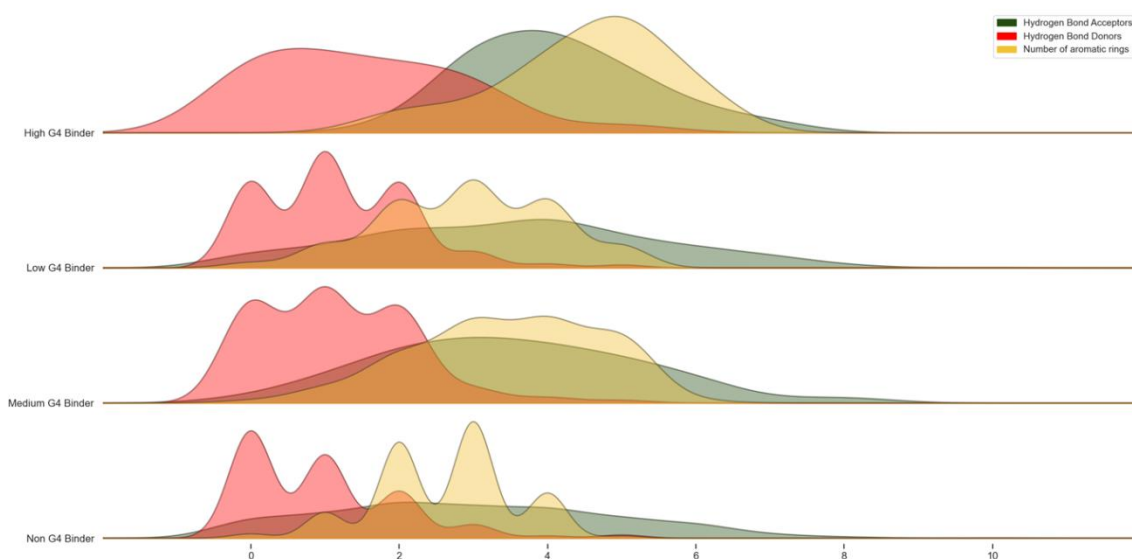


Figure 3: Chemical profiling of four category ligands in terms of H-bond acceptor, H-bond donor, and aromatic ring for all four category ligands.

Our result shows the presence of a high number of the aromatic ring and a comparatively higher number of H-bond acceptors than the number of H-bond donors is a prerequisite condition to be a high G4DNA binder ligand. Based on the average idock score and average RF score, three ligands such as ZINC000005220992, ZINC000013519380, and ZINC000013488698, as denoted by a circle, rhombus, and square respectively around the spot in scatter plot in matrix position [2,1] in Fig 2 were chosen. The structures of these three ligands are shown in Fig 4.

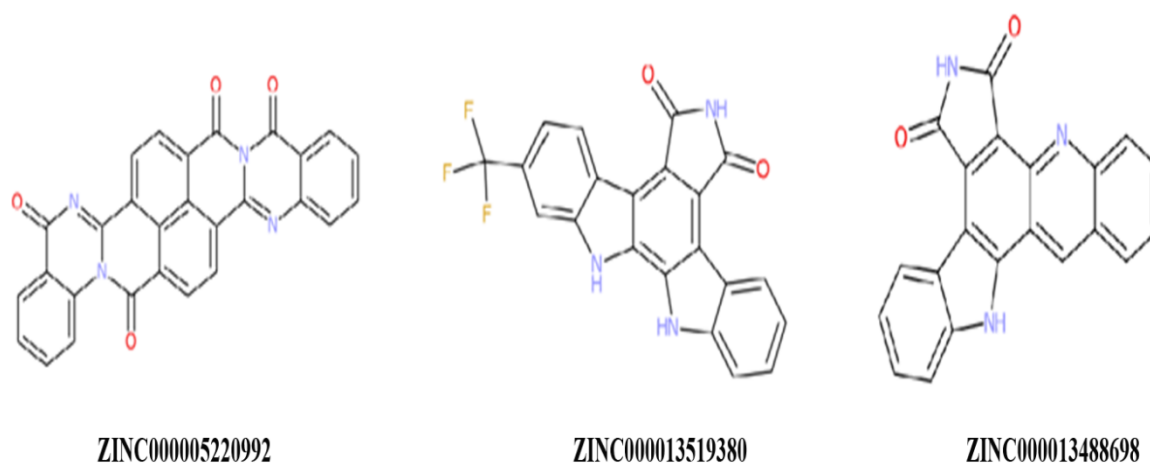


Figure 4: Chemical structure of three top-scored ligands. ZINC ID is given below each molecule.

The average idock score and average RF score for these ligands are (-10.17 kcal/mol, 6.22 pKd), (-9.87 kcal/mol, 6.22 pKd) and (-9.79 kcal/mol, 5.93 pKd). The best ligand as observed from the result is ZINC ID:05220992 (named as C-992) which was used for experimental validation. Several reports showed that using idock score people screened inhibitor (idock score \leq -8 kcal/mol) molecules against various receptors and worked well as per experimental validation report [27,28]. In this context, our screened compound showing a better idock score and RF score (-10.17 kcal/mol, 6.22 pKd).

The C-992 follows Lipinski's rule of five such as molecular weight (460 Da) $<$ 500, lipophilicity measure LogP $<$ 5, hydrogen bond donor $<$ 5, hydrogen bond acceptors $<$ 10, rotatable bonds $<$ 10, and net charges = 0. The molecular docking result of C-992 with G-quadruplex DNA (PDB ID 3CE5, which is the same as 22AG except for a single base at the end) by Discovery Studio is shown in Fig 5.

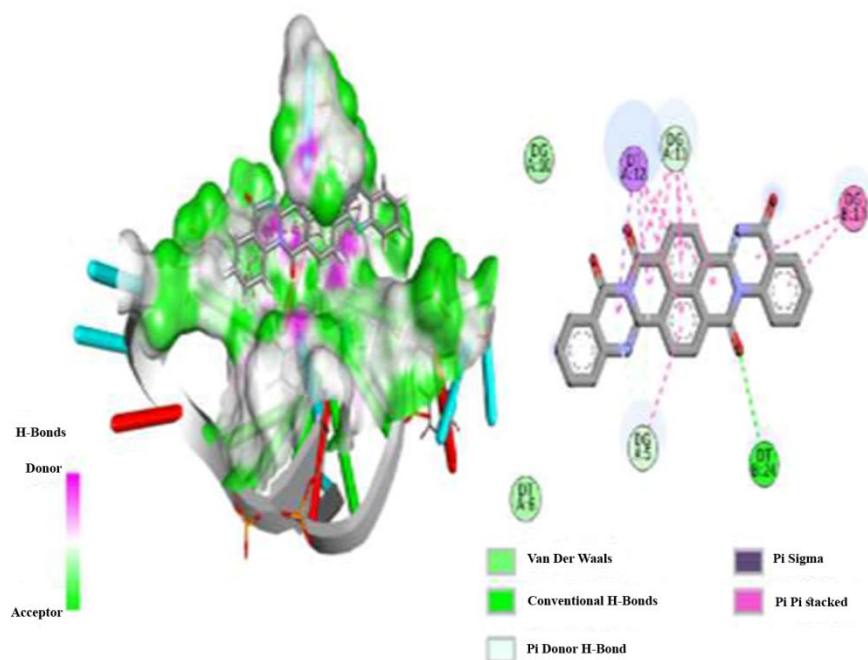


Figure 5: Molecular Docking results of G-quadruplex DNA (PDB ID: 3CE5) with ligand ZINC000005220992 obtained from Discovery Studio. 2D interaction diagram is shown on the right side with the bond formation of ligand molecules with receptors binding sites.

The docking of this C-992 with other G-quadruplex DNA is not shown here.

G4-FID assay:

We performed the G4-FID assay to check the capability of binding of C-992 to various G-quadruplex structures. We have taken three different G-quadruplex structures like 22AG (the same as PDB 3CE5 except for single base change at the end), Telo24 (the same as PDB 3CDM except one T is absent at 5' end), and Telo28 (PDB 1NZM, 2JWQ) as given in Table 1. Each type of G-quadruplex DNA (0.25 μM) was saturated with TO (0.5 μM) and titrated with various concentrations of C-992 to check TO replacement. We observed that 50% TO replacement (DC_{50}) from 22AG G-quadruplex DNA occurred at 0.99 μM of C-992, whereas that from CT-DNA was occurred at 1.56 μM as evident from Fig 6a. We also observed that

the DC₅₀ value for Telo24 and Telo28 was more than 27.12 and 21.94 μ M respectively as shown in Fig 6b.

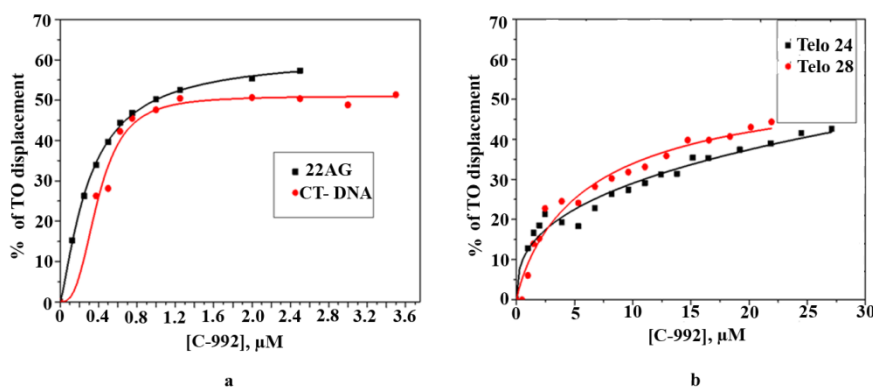


Figure 6: G4 FID assay: (a) TO displacement by C-992 from 22AG G-quadruplex DNA and CT-DNA; (b) TO displacement by C-992 from Telo24 and Telo28 G-quadruplex DNA.

The raw data for all these titrations are given in supplementary file S2. This data implicates that the compound C-992 has a stronger affinity to a specific G-quadruplex DNA 22AG among the three different types of G-quadruplex DNA under study. The affinity of this compound to 22AG is higher than double-stranded (CT-DNA). The 50% TO replacement (DC₅₀ value) from various DNA by the compound C-992 is given in Table 3.

Table 3: DC₅₀ values of TO replacement from various receptors DNA molecules

Serial No.	Receptor G-quadruplex DNA or CT-DNA	DC ₅₀ Value (μ M)
1	22AG	0.99
2	Telo24	>27.12
3	Telo28	>21.94
4	CT-DNA	1.56

All three G-quadruplex DNA sequences have human telomeric repeats TTAGGG but they are different in the loop sequence of G-quadruplex structures. Our data implicate that loop sequence can alter binding affinity because C-992 has a very strong affinity to 22AG only out

of the three G-quadruplex DNA. This observation corroborates an earlier report [29]. Moreover, the calculated DC_{50} value of C-992 for 22AG is close to reported DC_{50} values of some known G-quadruplex ligands [30,31] and even better than some of the reported compounds [32,33].

Absorption and fluorescence emission spectra of C-992:

The compound C-992 has intrinsic fluorescence property. We have taken various concentrations of C-992 (4.13-32.62 μ M) in Benchmark buffer (10 mM Tris-Cl, pH 7.2 and 100 mM KCl) and scanned wavelength range from 300 nm to 600 nm to get absorption spectra as shown in Fig 7a. The maximum absorption was observed at 363 nm. We have excited our compound at a different wavelength (320 nm - 370 nm) and fluorescence emission was recorded as shown in Fig 7b. This data suggests that the molecule has an added property – it has good intrinsic fluorescence and higher intensity was observed when excited at 320 nm.

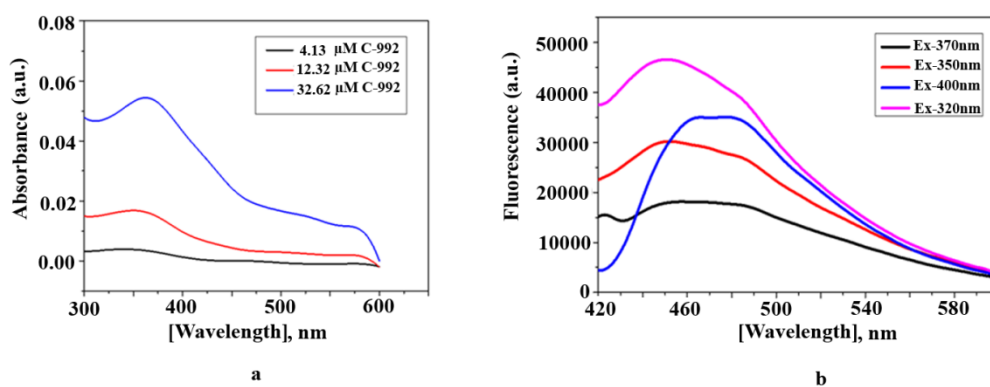


Figure 7: (a) Absorption spectra of different concentrations of C-992 (4.13-32.62 μ M) in Benchmark buffer (10 mM Tris-Cl, pH 7.2 and 100 mM KCl); (b) intrinsic fluorescence spectra of C-992 in the same buffer with different excitation wavelength.

CD Experiments:

CD Experiment was done to determine the G-quadruplex (22AG) binding property of the compound in K^+ containing buffer as mentioned in materials and methods. 22AG DNA forms a hybrid structure in 100 mM K^+ containing buffer with a strong positive peak at 295 nm with a shoulder at 265 nm, a small positive peak at 250 nm, and a small negative peak at 235 nm as per earlier report [34]. Generally, 22AG takes parallel or hybrid anti-parallel structure in K^+ buffer. In our case, it takes the same conformation as evident from the typical positive peak at 295 and 250 nm as shown in Fig 8.

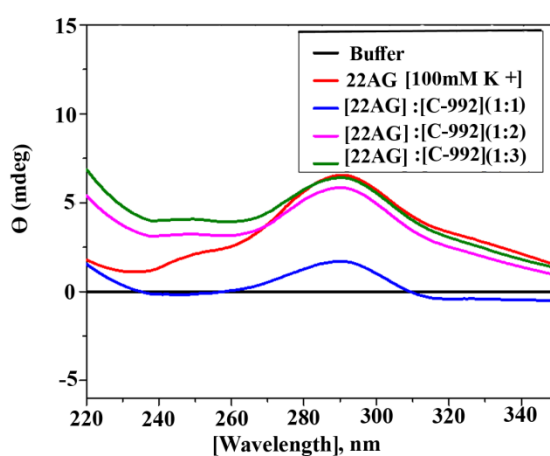


Figure 8: CD spectra of free 22AG and C-992 bound 22AG in Benchmark buffer (10 mM Tris-Cl, pH 7.2 and 100 mM KCl)

There were a drastic decrease in both 295 and 250 nm peaks during titration of 22AG with the C-992 at the ratio [DNA]:[compound] = 1:1. However, both the peaks dramatically increased along with an overall change in ellipticity at the ratio [DNA]:[compound] equal to 1:2 and 1:3. This data implicates that our compound C-992 has the capability of binding the G-quadruplex structure of 22AG and thereby can alter the three-dimensional conformation of the DNA.

Cellular uptake & cell imaging by fluorescence:

To observe the cellular uptake of this compound by the live-cell A549 under culture condition with and without using a fixative, we incubated the A549 cells in presence of a very low concentration (20 nM) of C-992 in PBS for 30 min at 37 °C as per our previous report [24,35,36]. Then the cell images were captured under bright field mode (data not shown) and fluorescence mode using a filter set 49 of an Axioscope A1 fluorescence microscope (excitation 365 nm and Emission 445/50). It is evident from Fig 9a that C-992 is capable of entering into cells through the live cell membrane and label the cell with sky blue fluorescence color. We also incubated the trypsinized cells in presence of C-992 after fixation with 70% ethanol. This fixation makes the cell porous and hence increases the permeability of the membrane. In addition to binding G-quadruplex DNA, it also binds double-stranded DNA with comparatively lower affinity. So, we got a brighter fluorescence image for porous cells as shown in Fig 9b.

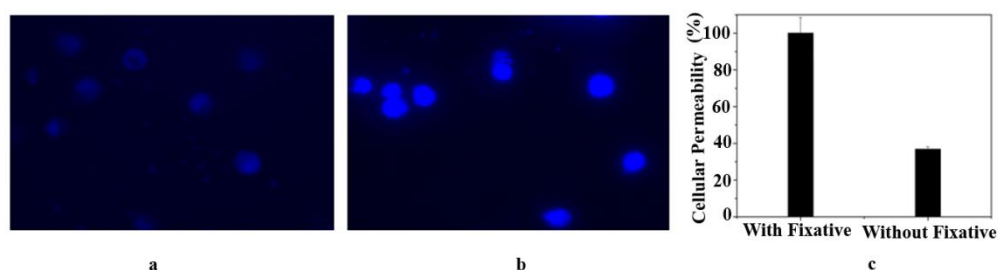


Figure 9: Cell permeability and live-cell imaging using C-992; A549 cell imaging under Axioscope A1 in fluorescence mode without fixation (a) and after fixation (b) of trypsinized cells; (c) percentage of cellular permeability of the compound C-992 (with and without fixative).

Taking the image intensity of the porous cells (after fixation) as 100% permeable, we calculated the intensity of cell image without fixation and observed about 40% permeability of the compound in live cells as shown in Fig 9c. We also did the permeability of the

compound in A549 cells in adherent conditions and got a similar result (data not shown). Our data suggest that this compound has a strong potential to enter into a living cell and capable of lighting up the nucleus.

Cell viability assay with MTT assay:

We have checked cell viability by MTT assay. The A549 cells were treated with various doses of C-992 (10-100 μM) and incubated for 46 h. The cell cytotoxicity is low with the dose range we have used. About 87% and 65% cell viability was observed at 10 μM and 100 μM of C-992 respectively. The data is shown in Fig 10.

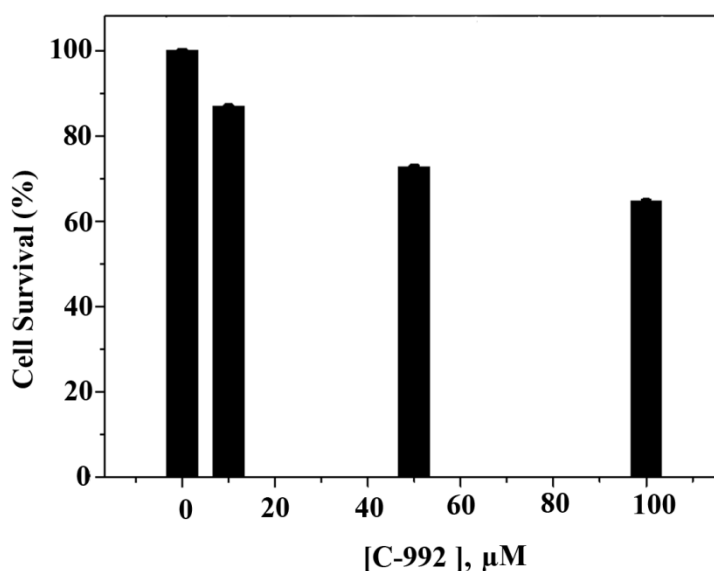


Figure 10: Percentage of A549 cell survival after treatment with various concentrations of C-992 for 46 h by MTT assay.

Telomerase inhibition by C-992 in A549 cells:

We observed that C-992 is a potential G-quadruplex sequence binder and it has an affinity for the telomeric G-quadruplex sequence. Hence, we measured the telomerase activity by TRAP assay. A549 cells were treated with different doses of C-992. An untreated control set and

another set treated with 0.04% DMSO were also taken into consideration because the highest dose of the compound contains 0.04% DMSO in culture media. There was no difference in the number of bands between the untreated control and 0.04% DMSO sets (Fig 11a). This proves that 0.04% DMSO does not affect telomerase activity. However, a dose-dependent reduction of telomerase activity was observed in A549 cells treated with C-992 (0 – 2 μ M) as shown in Fig 11a.

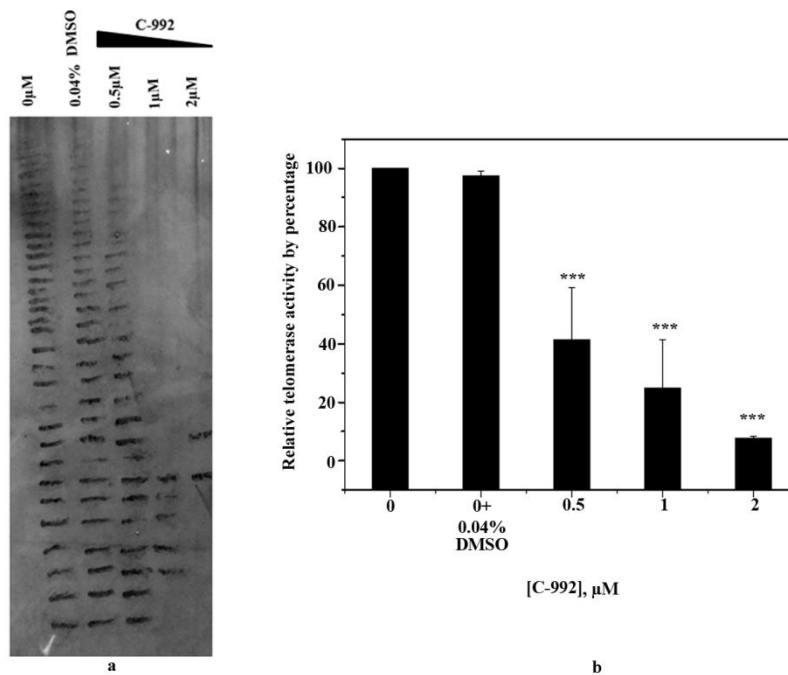


Figure 11: TRAP assay of A549 with C-992. a. Gel image shows C-992 can inhibit telomerase activity in a dose-dependent manner in A549 cells. b. Quantification of TRAP assay in the 11a. Student T-test was performed, p values are as follows ‘*’ ($0.01 < p \leq 0.05$), ‘**’ ($0.001 < p \leq 0.01$) and ‘***’ ($p \leq 0.001$).

Notably, a significant ($0.001 < p \leq 0.01$) reduction of telomerase activity was observed at a dose as low as 0.5 μ M as shown in Fig 11b. The telomerase activity further diminished below 10% at a dose of 2 μ M. This data implicates that C-992 is a potent telomerase inhibitor. Several reports showed that the G-quadruplex ligand can inhibit telomerase activity [35–40]. The proposed mechanism by which the G-quadruplex ligand inhibits telomerase is by stabilizing the G-quadruplex sequence of the telomere, by stabilizing the G-quadruplex

sequence of the promoter region of hTERT or reduction of hTERT expression and by direct interaction with the catalytic domain of hTERT [41,42].

A different strategy was opted to investigate the actual mechanism of telomerase inhibition; moreover, it was required to determine whether inhibition of Taq DNA pol is responsible for the TRAP assay result. So, the same TRAP assay was performed with control A549 cell lysate where the cells were not treated with C-992. C-992 was added at a concentration of 2 μM to the TRAP reaction mix during the reaction since it was the highest dose of C-992 used for the earlier TRAP assay (Fig 11). A control set was also prepared where C-992 was not added and another set was prepared with 0.2% DMSO since C-992 was dissolved in DMSO. After the polyacrylamide gel electrophoresis, the same number of bands were observed in each lane. Moreover, there was not much difference in the total intensity of the bands (Fig 12).



Fig 12: TRAP assay of A549 was performed with indicated concentration where C-992 was added to the cell-free system. The result showed no inhibition of telomerase activity.

We hypothesized that if no band was observed in C-992 lane it might be due to the formation of the G-quadruplex structure in the cell-free system or due to the inhibition of Taq DNA polymerase by the compound C-992. On the contrary to our hypothesis, bands were observed in each lane. This result is contradictory to our earlier result where telomerase activity was almost inhibited by 2 μ M C-992 (Fig 11). The possible reason behind this contradictory result is that 2 μ M C-992 was not sufficient to form or stabilize a G-quadruplex structure in the cell-free system and also could not inhibit Taq DNA pol activity. Whatever be the mechanism of telomerase inhibition by C-992, it is evident from our whole data that C-992 is a very good telomerase inhibitor and it can not affect the activity of Taq DNA pol.

Conclusion

In conclusion, we report that compound C-992 from the ZINC database has a strong binding affinity with a specific G-quadruplex structure present in the telomere. Compound C-992 complies with Lipinski's rule of five and is permeable to live cells implicating its potential to be a drug. It inhibits telomerase activity significantly *in-vivo* at a very low dose and hence it could be promising in cancer therapy. It has an added intrinsic fluorescence property and can stain the cell with blue color fluorescence.

Funding Sources:

This work was supported by UGC-SAP (DRS-II). UG thank ICMR (ICMR-SRF No.3/2/2/36/2018/Online OncoFship/NCD-III and No.3/1/3/JRF-2018/HRD-055(66816)) and DST (DST/INSPIRE Fellowship/2014/37), for providing the manpower for this work.

References

- [1] S. Asamitsu, S. Obata, Z. Yu, T. Bando, H. Sugiyama, molecules Recent Progress of Targeted G-Quadruplex-Preferred Ligands Toward Cancer Therapy, (n.d.). doi:10.3390/molecules24030429.
- [2] V. Víglaský, L. Bauer, K. Tlučková, Structural Features of Intra- and Intermolecular G-Quadruplexes Derived from Telomeric Repeats, *Biochemistry*. 49 (2010) 2110–2120. doi:10.1021/bi902099u.
- [3] I.M. Pedroso, L.F. Duarte, G. Yanez, K. Burkewitz, T.M. Fletcher, Sequence specificity of inter- and intramolecular G-quadruplex formation by human telomeric DNA, *Biopolymers*. 87 (2007) 74–84. doi:10.1002/bip.20790.
- [4] Y. Xu, Chemistry in human telomere biology: Structure, function and targeting of telomere DNA/RNA, *Chem. Soc. Rev.* 40 (2011) 2719–2740. doi:10.1039/c0cs00134a.
- [5] T. Kimura, K. Kawai, M. Fujitsuka, T. Majima, Fluorescence properties of 2-aminopurine in human telomeric DNA, *Chem. Commun.* 4 (2004) 1438–1439. doi:10.1039/b403913k.
- [6] T. Kimura, K. Kawai, M. Fujitsuka, T. Majima, Detection of the G-quadruplex-TMPyP4 complex by 2-aminopurine modified human telomeric DNA, *Chem. Commun.* (2006) 401–402. doi:10.1039/b514526k.
- [7] C. Douarre, X. Mergui, A. Sidibe, D. Gomez, P. Alberti, P. Mailliet, C. Trentesaux, J.-F. Riou, DNA damage signaling induced by the G-quadruplex ligand 12459 is modulated by PPM1D/WIP1 phosphatase., *Nucleic Acids Res.* 41 (2013) 3588–99. doi:10.1093/nar/gkt073.
- [8] Y.C. Tsai, H. Qi, C.P. Lin, R.K. Lin, J.E. Kerrigan, S.G. Rzuczek, E.J. LaVoie, J.E.

- Rice, D.S. Pilch, Y.L. Lyu, L.F. Liu, A G-quadruplex stabilizer induces M-phase cell cycle arrest, *J. Biol. Chem.* 284 (2009) 22535–22543. doi:10.1074/jbc.M109.020230.
- [9] D. Sen, W. Gilbert, Formation of parallel four-stranded complexes by guanine-rich motifs in DNA and its implications for meiosis, *Nature.* 334 (1988) 364–366. doi:10.1038/334364a0.
- [10] A.M. Zahler, J.R. Williamson, T.R. Cech, D.M. Prescott, Inhibition of telomerase by G-quartet DNA structures, *Nature.* 350 (1991) 718–720. doi:10.1038/350718a0.
- [11] G. Zhou, X. Liu, Y. Li, S. Xu, C. Ma, X. Wu, Y. Cheng, Z. Yu, G. Zhao, Y. Chen, Telomere targeting with a novel G-quadruplex-interactive ligand BRACO-19 induces T-loop disassembly and telomerase displacement in human glioblastoma cells, *Oncotarget.* 7 (2016). www.impactjournals.com/oncotarget.
- [12] D.J. Patel, A.T. Phan, V. Kuryavyy, Human telomere, oncogenic promoter and 5'-UTR G-quadruplexes: diverse higher order DNA and RNA targets for cancer therapeutics., *Nucleic Acids Res.* 35 (2007) 7429–55. doi:10.1093/nar/gkm711.
- [13] S. Balasubramanian, L.H. Hurley, UKPMC Funders Group Targeting G-quadruplexes in gene promoters : a novel anticancer, *Nat. Rev. Drug Discov.* 10 (2011) 261–275. doi:10.1038/nrd3428.Targeting.
- [14] H. Martadinata, A.T. Phan, Structure of human telomeric RNA (TERRA): Stacking of two G-quadruplex blocks in K⁺ solution, *Biochemistry.* 52 (2013) 2176–2183. doi:10.1021/bi301606u.
- [15] A.L. Moye, K.C. Porter, S.B. Cohen, T. Phan, K.G. Zyner, N. Sasaki, G.O. Lovrecz, J.L. Beck, T.M. Bryan, ARTICLE Telomeric G-quadruplexes are a substrate and site of localization for human telomerase, (2015). doi:10.1038/ncomms8643.
- [16] Z. Schrank, N. Khan, † Id, C. Osude, S. Singh Id, R.J. Miller, C. Merrick, A. Mabel, A. Kuckovic, N. Puri, molecules Oligonucleotides Targeting Telomeres and

- Telomerase in Cancer, (n.d.). doi:10.3390/molecules23092267.
- [17] M. Ivancich, Z. Schrank, L. Wojdyla, B. Leviskas, A. Kuckovic, A. Sanjali, N. Puri, antioxidants Treating Cancer by Targeting Telomeres and Telomerase, (2017). doi:10.3390/antiox6010015.
- [18] H.M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T.N. Bhat, H. Weissig, I.N. Shindyalov, P.E. Bourne, The Protein Data Bank, *Nucleic Acids Res.* 28 (2000) 235–242. doi:10.1093/nar/28.1.235.
- [19] G.M. Morris, H. Ruth, W. Lindstrom, M.F. Sanner, R.K. Belew, D.S. Goodsell, A.J. Olson, Software news and updates AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility, *J. Comput. Chem.* 30 (2009) 2785–2791. doi:10.1002/jcc.21256.
- [20] C.A. Lipinski, F. Lombardo, B.W. Dominy, P.J. Feeney, Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings, *Adv. Drug Deliv. Rev.* 46 (2001) 3–26. doi:10.1016/S0169-409X(00)00129-0.
- [21] N.M. O’Boyle, M. Banck, C.A. James, C. Morley, T. Vandermeersch, G.R. Hutchison, Open Babel: An Open chemical toolbox, *J. Cheminform.* 3 (2011) 33. doi:10.1186/1758-2946-3-33.
- [22] H. Li, K.S. Leung, M.H. Wong, Idock: A multithreaded virtual screening tool for flexible ligand docking, in: 2012 IEEE Symp. Comput. Intell. Comput. Biol. CIBCB 2012, 2012: pp. 77–84. doi:10.1109/CIBCB.2012.6217214.
- [23] D. Monchaud, C. Allain, H. Bertrand, N. Smargiasso, F. Rosu, V. Gabelica, A. De Cian, J.-L. Mergny, M.-P. Teulade-Fichou, Ligands playing musical chairs with G-quadruplex DNA: a rapid and simple displacement assay for identifying selective G-quadruplex binders., *Biochimie.* 90 (2008) 1207–23. doi:10.1016/j.biochi.2008.02.019.

- [24] A. Majumdar, S. Mondal, C.G. Daniliuc, D. Sahu, B. Ganguly, S. Ghosh, U. Ghosh, K. Ghosh, α -Amino Acid Derived Benzimidazole-Linked Rhodamines: A Case of Substitution Effect at the Amino Acid Site toward Spiro Ring Opening for Selective Sensing of Al^{3+} Ions, *Inorg. Chem.* 56 (2017) 8889–8899. doi:10.1021/acs.inorgchem.7b00835.
- [25] R. Sadhukhan, P. Chowdhury, S. Ghosh, U. Ghosh, Expression of Telomere-Associated Proteins is Interdependent to Stabilize Native Telomere Structure and Telomere Dysfunction by G-Quadruplex Ligand Causes TERRA Upregulation, *Cell Biochem. Biophys.* 76 (2018) 311–319. doi:10.1007/s12013-017-0835-0.
- [26] U. Ghosh, N.P. Bhattacharyya, Benzamide and 4-amino 1,8 naphthalimide treatment inhibit telomerase activity by down-regulating the expression of telomerase associated protein and inhibiting the poly(ADP-ribosyl)ation of telomerase reverse transcriptase in cultured cells, *FEBS J.* 272 (2005) 4237–4248. doi:10.1111/j.1742-4658.2005.04837.x.
- [27] D.A.E. Pitaloka, D.S.F. Ramadhan, Arfan, L. Chaidir, T.M. Fakhri, Docking-Based Virtual Screening and Molecular Dynamics Simulations of Quercetin Analogs as Enoyl-Acyl Carrier Protein Reductase (InhA) Inhibitors of Mycobacterium tuberculosis, *Sci. Pharm.* 2021, Vol. 89, Page 20. 89 (2021) 20. doi:10.3390/SCIPHARM89020020.
- [28] B. LR, L. DJ, N. DT, A. MO, H. M, G. AT, E. S, Identification of potential Zika virus NS2B-NS3 protease inhibitors via docking, molecular dynamics and consensus scoring-based virtual screening, *J. Mol. Model.* 25 (2019). doi:10.1007/S00894-019-4076-6.
- [29] K. Takahama, C. Sugimoto, S. Arai, R. Kurokawa, T. Oyoshi, Loop Lengths of G-Quadruplex Structures Affect the G-Quadruplex DNA Binding Selectivity of the RGG

- Motif in Ewing's Sarcoma, *Biochemistry*. 50 (2011) 5369–5378.
doi:10.1021/bi2003857.
- [30] S. S, H. HL, G. X, Y. JL, L. CY, Z. J, S. WL, Y. TM, J. LN, A comparative study of the interaction of two structurally analogous ruthenium complexes with human telomeric G-quadruplex DNA, *J. Inorg. Biochem.* 121 (2013) 19–27.
doi:10.1016/J.JINORGBIO.2012.12.011.
- [31] B. Maji, K. Kumar, K. Muniyappa, S. Bhattacharya, New dimeric carbazole-benzimidazole mixed ligands for the stabilization of human telomeric G-quadruplex DNA and as telomerase inhibitors. A remarkable influence of the spacer, *Org. Biomol. Chem.* 13 (2015) 8335–8348. doi:10.1039/c5ob00675a.
- [32] J. Medeiros-Silva, A. Guédin, G.F. Salgado, J.L. Mergny, J.A. Queiroz, E.J. Cabrita, C. Cruz, Phenanthroline-bis-oxazole ligands for binding and stabilization of G-quadruplexes, *Biochim. Biophys. Acta - Gen. Subj.* 1861 (2017) 1281–1292.
doi:10.1016/j.bbagen.2016.11.024.
- [33] K. Duskova, S. Sierra, M.S. Arias-Pérez, L. Gude, Human telomeric G-quadruplex DNA interactions of N-phenanthroline glycosylamine copper(II) complexes, *Bioorg. Med. Chem.* 24 (2016) 33–41. doi:10.1016/J.BMC.2015.11.037.
- [34] J. Medeiros-Silva, A. Guédin, G.F. Salgado, J.L. Mergny, J.A. Queiroz, E.J. Cabrita, C. Cruz, Phenanthroline-bis-oxazole ligands for binding and stabilization of G-quadruplexes, *Biochim. Biophys. Acta - Gen. Subj.* 1861 (2017) 1281–1292.
doi:10.1016/j.bbagen.2016.11.024.
- [35] S. Bhattacharya, P. Chaudhuri, A.K. Jain, A. Paul, Symmetrical bisbenzimidazoles with benzenediyl spacer: The role of the shape of the ligand on the stabilization and structural alterations in telomeric g-quadruplex dna and telomerase inhibition, *Bioconjug. Chem.* 21 (2010) 1148–1159. doi:10.1021/bc9003298.

- [36] B. Maji, S. Bhattacharya, Advances in the molecular design of potential anticancer agents via targeting of human telomeric DNA, *Chem. Commun.* 50 (2014). doi:10.1039/c4cc00611a.
- [37] Q.P. Qin, J.L. Qin, T. Meng, W.H. Lin, C.H. Zhang, Z.Z. Wei, J.N. Chen, Y.C. Liu, H. Liang, Z.F. Chen, High in vivo antitumor activity of cobalt oxoisoaporphine complexes by targeting G-quadruplex DNA, telomerase and disrupting mitochondrial functions, *Eur. J. Med. Chem.* 124 (2016) 380–392. doi:10.1016/j.ejmech.2016.08.063.
- [38] S.K. Noureini, H. Esmaili, F. Abachi, S. Khiali, B. Islam, M. Kuta, A.A. Saboury, M. Hoffmann, J. Sponer, G. Parkinson, S. Haider, Selectivity of major isoquinoline alkaloids from *Chelidonium majus* towards telomeric G-quadruplex: A study using a transition-FRET (t-FRET) assay, *Biochim. Biophys. Acta - Gen. Subj.* 1861 (2017) 2020–2030. doi:10.1016/j.bbagen.2017.05.002.
- [39] C. Wang, B. Carter-Cooper, Y. Du, J. Zhou, M.A. Saeed, J. Liu, M. Guo, B. Roembke, C. Mikek, E.A. Lewis, R.G. Lapidus, H.O. Sintim, Alkyne-substituted diminazene as G-quadruplex binders with anticancer activities, *Eur. J. Med. Chem.* 118 (2016) 266–275. doi:10.1016/j.ejmech.2016.04.030.
- [40] Z.Z. Wei, Q.P. Qin, J.N. Chen, Z.F. Chen, Oxoisoaporphine as potent telomerase inhibitor, *Molecules.* 21 (2016). doi:10.3390/molecules21111534.
- [41] S.M. Gowan, J.R. Harrison, L. Patterson, M. Valenti, M.A. Read, S. Neidle, L.R. Kelland, A G-Quadruplex-Interactive Potent Small-Molecule Inhibitor of Telomerase Exhibiting in Vitro and in Vivo Antitumor Activity, 61 (2002) 1154–1162.
- [42] S.M.L. Palumbo, S.W. Ebbinghaus, L.H. Hurley, Formation of a unique end-to-end stacked pair of G-quadruplexes in the hTERT core promoter with implications for inhibition of telomerase by G-quadruplex-interactive ligands, *J. Am. Chem. Soc.* 131 (2009) 10878–10891. doi:10.1021/ja902281d.

High throughput screening of a new fluorescent G-quadruplex ligand having telomerase inhibitory activity in human A549 cells.

Authors' name and affiliations:

Sourav Ghosh^{‡1}, Debapriya De ^{‡1}, Victor Banerjee¹, Soumyajit Biswas¹, Utpal Ghosh^{1*}

¹Department of Biochemistry & Biophysics

University of Kalyani, Kalyani- 741235, India

[‡]These authors contributed equally.

***Corresponding author:** Utpal Ghosh, Ph.D.

Department of Biochemistry & Biophysics

University of Kalyani, Kalyani- 741235, India

Fax No. 091 033-2582-8282

Phone No. +91-8420520334

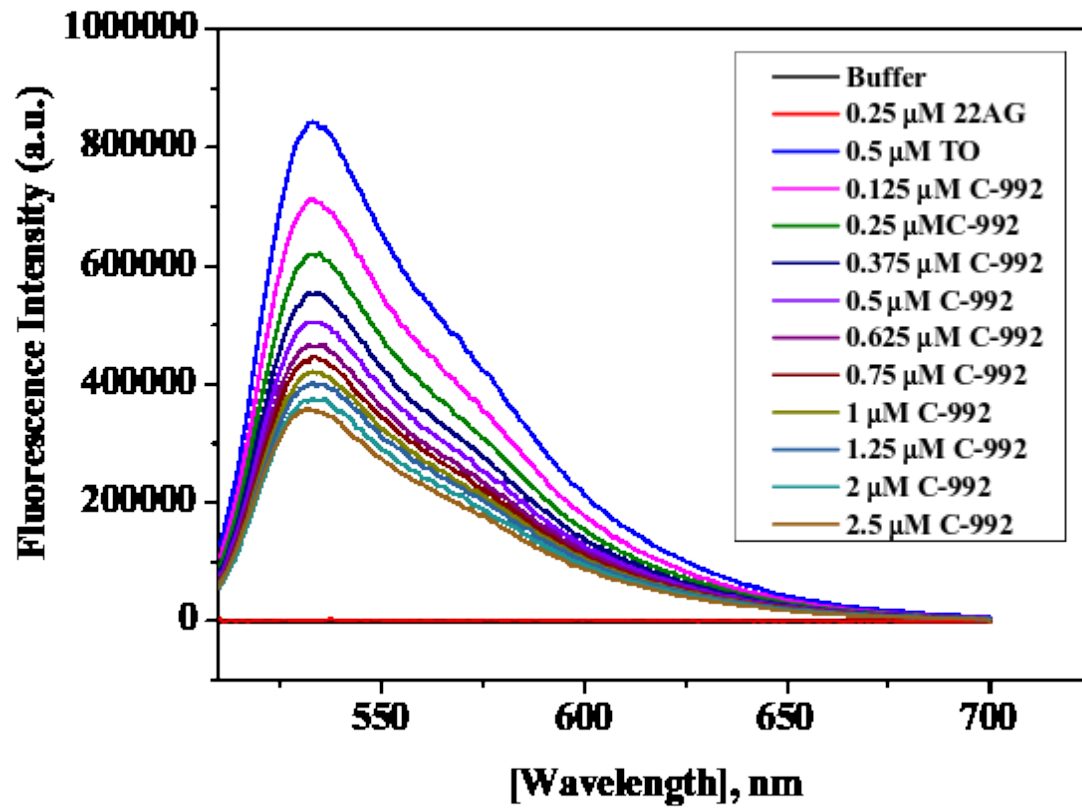
Email: utpal8ghosh@gmail.com

Supporting Information

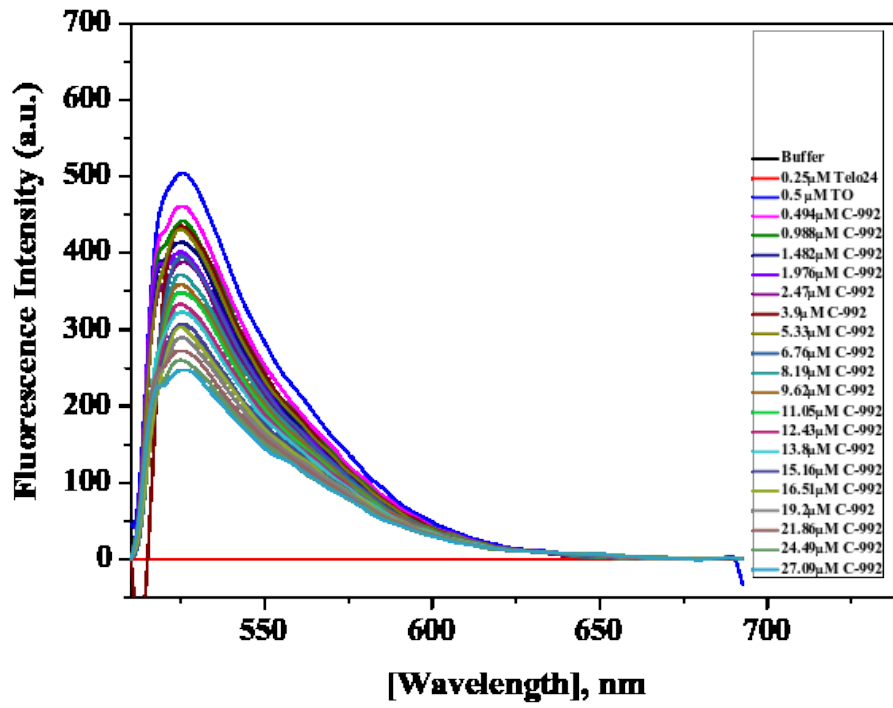
S1: Details of docking parameter of receptor G-quadruplex DNA

PDB	center_x	center_y	center_z	size_x	size_y	size_z
receptor						
1NZM_0	29.171186	35.135953	18.012209	20	20	20
1NZM_1	27.728512	29.371535	30.69614	20	20	20
2HRI_0	-18.602135	-11.650558	-15.250596	20	20	20
2HRI_1	-9.872596	0.1565	-0.082	20	20	20
2JWQ_0	-5.531532	-2.952129	21.318581	20	20	20
2JWQ_1	-4.194823	1.860935	9.099758	20	20	20
3CDM_0	2.629881	0.820333	-14.777619	20	20	20
3CE5_0	17.984886	18.087727	9.447205	20	20	20

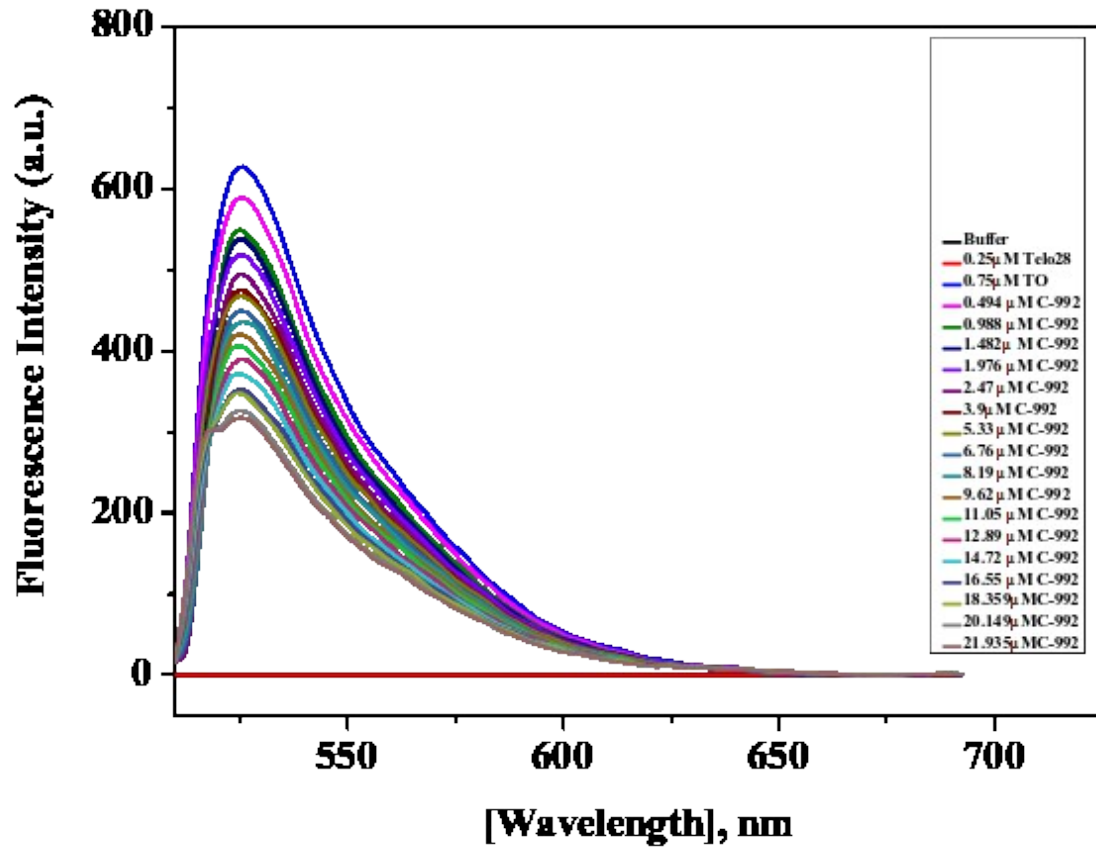
S2: Raw spectrum of TO displacement from three G-quadruplex DNA and CT-DNA.



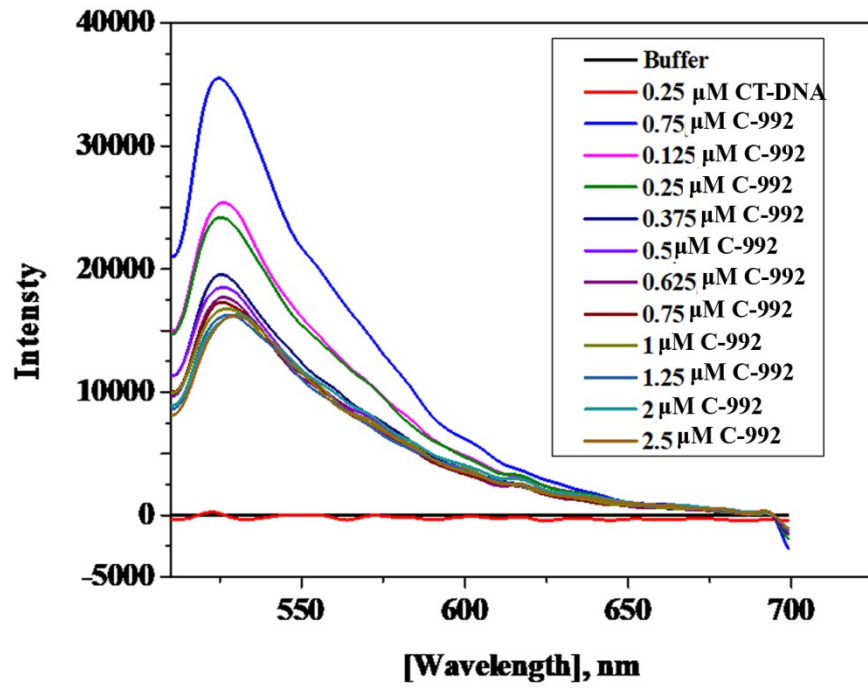
TO displacement by C-992 from TO-saturated 22AG G-quadruplex DNA



TO displacement by C-992 from TO-saturated Telo24 G-quadruplex DNA.



TO displacement by C-992 from TO-saturated Telo28 G-quadruplex DNA



TO displacement by C-992 from TO-saturated CT-DNA