Isolation of G-quadruplex structures from human cells using small molecular probes

Nitheesh K.

(Reg. No.20101072)

Department of Biology

A thesis submitted in partial fulfilment of the requirements

for the BS-MS dual degree programme in IISER Pune.

Thesis Guide

Dr. Shantanu Chowdhury

Proteomics and Structural Biology Unit

CSIR- Institute of Genomics and Integrative Biology,

New Delhi, India

[1]

Certificate

This is to certify that this dissertation entitled "**Isolation of G-quadruplex structures from human cells using small molecular probes**" towards the partial fulfilment of the BS-MS dual degree programme at the Indian Institute of Science Education and Research, Pune represents original research carried out by Nitheesh K. (Reg. No. 20101072) at the CSIR-Institute of Genomics and Integrative Biology, New Delhi, India, under the supervision of Dr. Shantanu Chowdhury, Proteomics and structural Biology Unit, during the academic year 2015-2016.

Dr. Shantanu Chowdhury

Declaration

I hereby declare that the matter embodied in the report entitled **"Isolation G-quadruplex structures from human cells using small molecular probes"** are the results of the investigations carried out by me at the Proteomics and Structural Biology Unit, CSIR-Institute of Genomics and Integrative Biology, New Delhi, under the supervision of Dr. Shantanu Chowdhury and the same has not been submitted elsewhere for any other degree.

Nitheesh K.

(Reg. No. 20101072)

Abstract

Guanine rich nucleotide sequences of specific pattern are capable of forming unusual secondary structures called G-quadruplexes (G4s). *In silico* studies have shown that putative G-quadruplex forming motifs (pG4s) are frequently present in functionally important regions of the genome. But formation of G-quadruplex structures by these pG4s in cells remains elusive. In this study, we made an attempt to isolate intrinsic G-quadruplex structure from human cells using biotinylated quadruplex interacting small organic molecules. The ligands showed significant affinity towards structurally different G4s *in vitro.* This interaction was observed to be selective towards G4s over duplex DNA. Biotinylated ligands were then used to perform Chromatin immuno-precipitation from human cells to isolate cellular G-quadruplex structures. Even though the ligands showed selective binding to G-quadruplex structures *in vitro,* the methodology used in this study failed to pull down G4s selectively from cells. This demands the need of alternative approaches to identify cellular G-quadruplexes. The ligands have shown significant inhibition of growth and proliferation of cancer cells. This implicates their potential applications in cancer therapeutics.

Table of Contents

List of figures

List of tables

Acknowledgements

I would like to express my sincere gratitude to Dr. Shantanu Chowdhury for giving me an opportunity to work in his lab and for his continuous support and suggestions.

I would like to thank Dr. Dhurjhoti Saha for all his motivation and encouragement throughout the project and for helping me bring this project to the current form.

I would also like to thank all the members of SC lab for their suggestions and fruitful discussions.

I thank Dr. Satyaprakash Pandey, Amrita Singh and Smita Nahar from Dr. Souvik Maiti's lab for their help with all the Biophysical experiments.

I express my thanks to my TAC member Dr. Kundan Sengupta for his valuable suggestions on the Project. I thank

Finally, I would like to thank all my friends in CSIR-IGIB and IISER-Pune and my family for their encouragement and continuous moral support.

Introduction

Guanine rich nucleotide sequences can fold into four stranded, non-canonical secondary structures called G-quadruplexes (G4s)(Burge et al., 2006). These are self-assembled structures of guanines from a single strand or multiple strands. The structural unit of a G-quadruplex is called a G-quartet which is a planar structure formed from four guanines connected together by the non-canonical Hoogsteen hydrogen bonding (Yaku et al., 2012). In a typical G-quadruplex structure, two or more G-quartets are stacked on top of each other by π-π stacking interaction (figure 1). These interactions are further stabilized by monovalent cations such as K^+, Na^+ and Li^+.

.
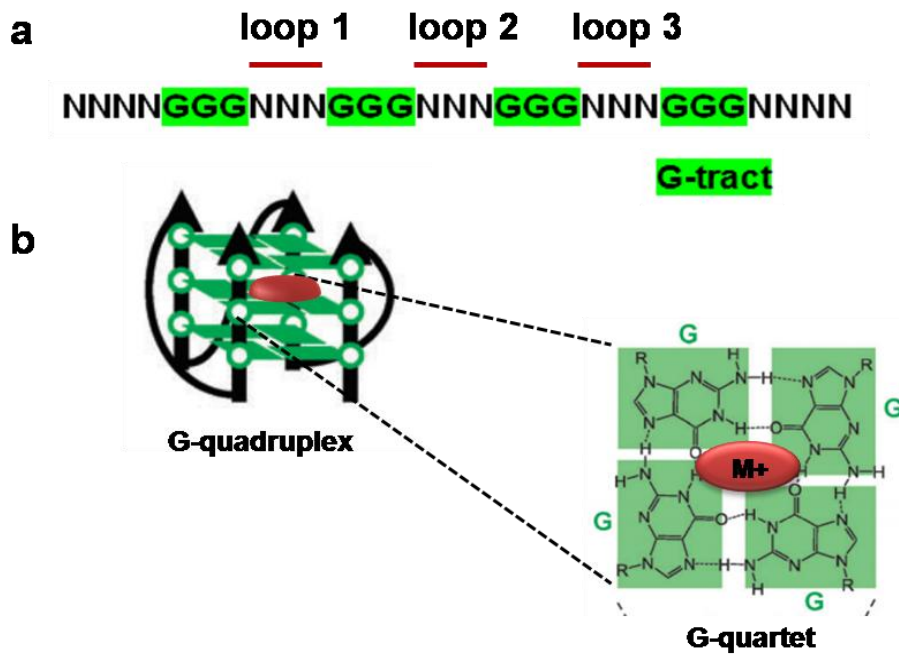
Figure 1 G-quadruplex structure.

Four Guanines connected together by Hoogsteen hydrogen binding forms a planar structure which is the structural unit of a G-quadruplex called a G-quartet. Two or more G-quartets stacked on top of each other by π- π stacking forms a G-quadruplex. Monovalent cations like K^+ coordinate to the O_6 Oxygen atom of guanine and stabilize these structures. Figure modified from (Yaku et al., 2012) with permission.

A G-quadruplex structure can be formed from a single strand of nucleotides or from multiple strands. A quadruplex formed from a single strand of DNA or RNA is called an intra-molecular G-quadruplex while an inter-molecular G-quadruplex involves multiple strands (figure 2a) (Burge et al., 2006).

Depending on the nucleotide sequence and the cation present, a quadruplex can adopt different topologies such as parallel, anti-parallel and mixed (Burge et al., 2006). In a parallel conformation, the individual strands orient themselves in the same direction (figure 2b). These structures are shown to be stable under near physiological conditions and are stabilized by monovalent cations such as K^+, Na^+ and Li^+ (Burge et al., 2006; Lam et al., 2013)
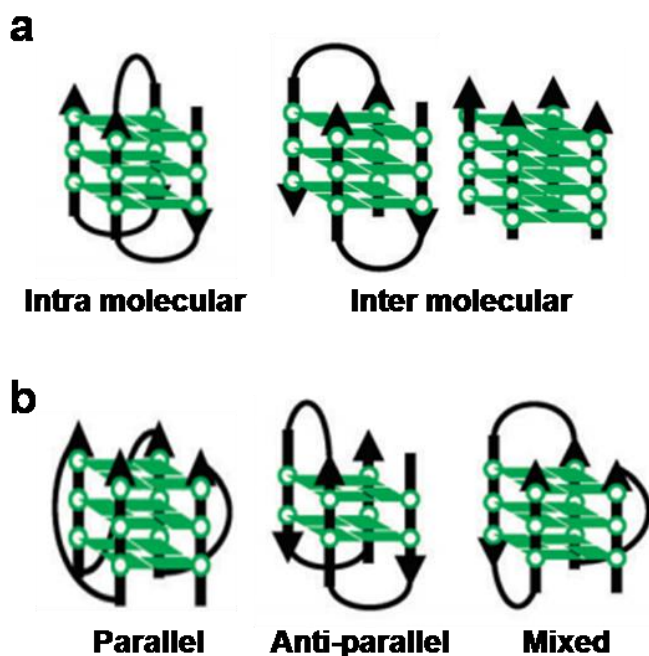
Figure 2: Different types of G-quadruplex structures.

(a) Quadruplex structure formed from a single strand of DNA is referred to as an intra molecular quadruplex and quadruplex formed from multiple strands is called an inter molecular G4 (b) different conformations of G-quadruplexes. Figure modified from (Yaku et al., 2012) with permission

In silico studies which search for motifs such as on given in figure 1a have shown that putative G-quadruplex (pG4s) forming motifs are frequently present in the human genome and the distribution of such motifs across the genome is non-random (Bochman et al., 2012; Huppert and Balasubramanian, 2005; Murat and Balasubramanian, 2014; Verma et al., 2009). Functionally relevant regions of the genome such as the telomeres, promoters of various genes and 5' un-translated region of the mRNA shows higher frequency of pG4s (Basundra et al., 2010; Huppert and Balasubramanian, 2007; Rhodes and Lipps, 2015; Wang et al., 2011; Yuan et al., 2013).

Quadruplex structures are shown to be stable under near-physiological conditions and various reports suggest the possible roles of G-quadruplex structures in different cellular processes such as transcription, telomere maintenance and replication stalling (Rhodes and Lipps, 2015). Different small organic molecules are shown to stabilize G-quadruplex structures *in vitro* and many such molecules have been used to target telomere quadruplexes as well as those present in the promoters of oncogenes such as *c-MYC, c-KIT, VEGF* and *SRC* (Agrawal et al., 2013; Burger et al., 2005; Incles et al., 2004; Islam et al., 2014; Kim et al., 2002; Lam et al., 2013; Muller et al., 2012; Riou et al., 2002). Stabilization of telomere quadruplex has shown to inhibit telomerase as well as ALT mediated telomere maintenance, eventually resulting in genomic instability and growth arrest in cancer cells (Tauchi et al., 2006; Wang et al., 2011; Zhou et al., 2006). Similarly, stabilization of oncogene promoter G4s showed down-regulation of these genes *in vitro*. Different analogues of quadruplex binding ligand Pyridostatin have shown to cause growth arrest and cellular senescence in human cancer cells (Muller et al., 2012). Similar effects are observed with the telomeric G-quadruplex interacting ligand BRACO-19 which inhibits telomerase mediated telomere maintenance and induces senescence and growth arrest in human uterus carcinoma cells (Burger et al., 2005; Neidle, 2010). Stabilization of quadruplex at the promoter of oncogene c-MYC has also shown to cause apoptosis in cancer cells by promoting telomere DNA damage as well as down regulation of DNA damage response proteins (Islam et al., 2014). Targeting G-quadruplex structures are now being used as therapeutic strategy in cancer

since treatment with quadruplex stabilizing molecules showed significant inhibition of growth and proliferation and induction of cellular senescence in many cancer cells(Lipps and Rhodes, 2009; Siddiqui-Jain et al., 2002; Verma et al., 2009).

But how many of the computationally predicted G4 motifs form a G-quadruplex *in vivo* is not known. Majority of the studies on the formation and functional importance of G-quadruplex structures were done under *ex vivo* experimental conditions (Di Antonio et al., 2012). So their existence in cells needs to be supported with substantial evidences. Till date there are only a very few studies published which address the cellular existence of G4s.The first report on the use of small organic molecules to isolate cellular quadruplex structures was from Muller, kumari et al (2010)(Muller et al., 2010). They employed chromatin immunoprecipitation using biotin tagged G4 interacting ligand Pyridostatin to isolate G-quadruplex structures from isolated genomic DNA from human cancer cells and mapped their positions in the chromosome. But the methodology employed in this study was as good as an *in vitro* experiment(Muller et al., 2010).

In our study, we will use a similar approach to isolate cellular G-quadruplex structures using different biotin labelled quadruplex ligands and map their positions in the chromosome. Instead of using genomic DNA or cell extracts, we thought of treating the cells with biotinylated ligands followed by precipitation of the bound sequences using streptavidin beads to isolate G-quadruplex structures from cells. This approach will be 'more cellular' in nature compared to the previous methods used to identify cellular G-quadruplex structure. Besides mapping their positions in the chromosome, we will also look for novel G-quadruplex interacting proteins. We are also interested in looking at the effect of different G-quadruplex interacting ligands in various biological processes.

Aims and objectives of this study

The main aim of this study is to isolate DNA G-quadruplex structures from human cancer cells using small organic molecules.

The specific objectives of the study

1. Checking the affinity and selectivity of biotinylated ligands towards G-quadruplex structures
2. Chromatin Immuno-precipitation using biotinylated ligands to isolate cellular G-quadruplex structures from human cells
3. Biological attributes of different G-quadruplex interacting ligands

Materials and Methods

Materials

All the oligonucleotides were purchased from Sigma-Aldrich. Sodium Cacodylate, Potassium chloride, Sodium chloride, Thiazole Orange, Dimethyl sulphoxide, Minimun Essential Media (MEM), Formaldehyde, Glycine, Sodium dodecyl sulphate, Ethylenediaminetetraacetic acid, Trizma base®, Triton-X 100, Lithium chloride, IGEPAL®, Deoxycholic acid, Phenol-Chloroform isoamyl alcohol, Isopropanol were all purchased from Sigma-Aldrich. Fetal Bovine Serum (FBS) was from Invitrogen. Proteinase K was from Fermentas.

Oligo nucleotides and Ligands

All the oligonucleotides were purchased from Sigma-Aldrich. Oligo stocks were prepared in molecular biology grade nuclease free water at a concentration of 200µM. These were diluted to required concentration with 10mM Sodium Cacodylate with or without 100mM Potassium Chloride for further experiments.

Biotinylated ligands Compound1, Compound b, Compound 4, Compound 5 and non-biotinylated ligands Compound 2 and Compound 3 were synthesized in collaboration with Dr. Shantanu Bhattacharya's lab at the Indian Institute of Science, Bangalore. All the ligands were stored at -20^{0}C.

Circular Dichroism

CD spectra were recorded in Jasco-810 CD spectropolarimeter in a quartz cuvette with 1cm path length. Quadruplex and mutant oligos were annealed at a concentration of 5µM in 10mM Sodium Cacodylate buffer containing 100mM Potassium Chloride by heating at 95^{0}C for 5 minutes and slow cooling to RT overnight. All the CD spectra represent the average of 3 scans recorded from 220 nm to 320 nm with 1nm bandwidth, 1nm data pitch and 1 sec D.I.T at 25^{0}C and zero corrected at 320 nm. Double stranded

oligos were annealed in 10mM Sodium Cacodylate buffer in the absence of potassium ions as mentioned above.

UV melting

UV melting experiments were performed in UV 2700 UV-Vis spectrometer. Oligo nucleotides were used at a concentration of 2µM in 1mL quartz cuvette with path length of 10mm. Absorbance at 295nm was recorded by changing the temperature from 20^0C to 95^0C at a rate of 0.5^0C/minute. Graphs were plotted in Graphpad Prism.

G4 Fluorescent Intercalator Displacement Assay

FID experiments were performed in Fluoromax-4 spectro-fluorometer as described earlier (Monchaud and Teulade-Fichou, 2010). In brief, Annealed quadruplex and double stranded oligonucleotides were used at a concentration of 0.5µM in a 100µl cuvette containing 10mM Sodium Cacodylate with an without 100mM Potassium Chloride respectively. Fluorescence of 1µM Thiazole Orange in buffer was recorded from 510nm to 700nm after excitation at 501nm. Preformed G-quadruplex or double stranded oligonucleotides at a final concentration of 0.5µM were then added to buffer containing two or three molar equivalent TO respectively. Fluorescence spectra were recorded after an incubation time of 5 minutes. Now, small amount of the ligand of interest was added to displace TO from the TO-DNA complex. Fluorescence spectra were recorded after a short incubation of three minutes after each addition. Then the percentage TO displacement values were calculated using the following formula.

$$\%TO\ displacement = [(F_0 - F_x)/F_0] \times 100$$

where,

F_0- Fluorescence intensity at 530nm without the addition of ligand

F_x- Fluorescence intensity at 530nm with the of x amount of ligand

Percentage TO displacement values were plotted against the concentration of the ligand used and DC_{50} values, which corresponds the concentration of ligand require to

displace 50% of TO from the DNA, were used to compare the relative affinity of a ligand towards different quadruplex structures. Graphs were plotted in Graphpad Prism 6.

Cell culture conditions

HT1080 fibro sarcoma cells were grown in T25 culture Flasks in Minimum Essential Media supplemented with 10% (v/v) Fetal Bovine Serum (FBS). Cells were grown at 37^0C in humidified incubators with 5% CO2 and were split at 70 to 80% confluency.

Cell viability assay

HT1080 fibro-sarcoma cells were seeded at a density of 5000 cells per well of a 96-well plate in a total volume of 100µl complete media. Cells were then grown at 37^0C in a humidified incubator with 5% CO_2 for 24 hours. Cells were then treated with different concentrations of compounds used in this study for 48 hours. Cell viability was estimated using CelltiterGlo cell viability kit according to the manufacturer's protocol. Graphs were plotted using Graphpad Prism 6.

Chromatin Immuno-precipitation

HT 1080 cells were seeded at a density of 10^5 cells per mL of Minimum Essential Media supplemented with 10% Fetal Bovine Serum in T25 culture flasks. Cells were then grown at 37^0C in a humidified incubator with 5%CO_2 for 24 hours. Cells were then treated with two different concentrations of either DMSO or the compound of interest. Cells were harvested after 48 hours of treatment and re-suspended in 5mL culture media and were fixed with 1% Formaldehyde for 10 minutes at 37^0C.Fixation was stopped by adding 500µl of 1.25M Glycine and incubated at 37^0C for 5 minute. Pellet down the cells at 4^0C at 1000 rpm for five minutes. The supernatant was discarded and the cells were washed in 1X Phosphate Buffered Saline (PBS) containing 1X Protease Inhibitor Cocktail (PIC) twice. Cells were then lysed using 1X SDS lysis buffer (1% (w/v) SDS, 10mM EDTA, 50mM Tris PH 8.1) containing 1X PIC on ice for 30 to 45 minutes. Genomic DNA of lysed cells were sheared in Diagenode Bioruptor to obtain a fragment

size of 500-1000 base pairs at a pulse of 30 sec ON, 45 sec OFF at high power. 10% volume of the sonicated fragments was further processed for the input DNA and the remaining sample was diluted with equal volume of ChIP dilution buffer (0.01% SDS (w/v), 1.1%Triton X-100 (v/v), 1.2mM EDTA, and 16.7mM Tris PH 8.1, 167mM NaCl). Diluted samples were incubated with 30 to 50μg blocked Streptavidin C magnetic beads at 4^0C on a rotor for 14 to 16 hours. Bead bound DNA fragments were separated by magnetic separation and washed 1 to 2 times with 200μl of low salt wash buffer (0.1% (w/v) SDS, 1.0% Triton X-100 (v/v), 2mM EDTA, 20mM Tris PH 8.1 and 150mM NaCl), high salt wash buffer (0.1% (w/v) SDS, 1.0% Triton X-100 (v/v), 2mM EDTA, 20mM Tris PH 8.1 and 500mM NaCl) and LiCl wash buffer (0.25M LiCl, 1% (v/v) IGEPAL, 1% (w/v) Deoxycholic acid, 1mM EDTA and 10mM Tris) respectively. The samples were then re-suspended in 200μl TE buffer containing 0.01% SDS and 2μl Proteinase K and were hated for 30 minutes at 55^0C. The DNA fragments were then isolated by standard phenol-chloroform isoamyl alcohol mixture and precipitated with Isopropanol. Samples were washed with ice-cold ethanol and dissolved in molecular biology grade nuclease free water. Recovered DNA samples were quantified using Nanodrop 1000 Spectrophotometer.

Results

Characterization of G-quadruplex structures by Circular Dichroism spectroscopy

G-quadruplex and mutant oligonucleotides were annealed in 10mM Sodium Cacodylate buffer containing 100mM Potassium Chloride at a concentration of 5µM. Samples were heated at 95^0C for 5 minutes and slowly cooled down to room temperature. The structures of the quadruplex formed were confirmed using Circular Dichroism spectroscopy and UV melting experiments. The list and the sequences of all oligonucleotides used in this study are given in Table 2.

Sl. No.	Name	Sequence (5' to 3')
1	Telomere G4	TTAGGGTTAGGGTTAGGGTTAGGG
2	Telomere G4 mut	TTAGAGTTAAGATTAGGGTTAAGG
3	c-MYC G4	TGGGGAGGGTGGGGAGGGTGGGGAAGG
4	c-MYC G4 mut	TGGGGAGGGTGAAGAGAGTGAAGAAGG
5	hras1 G4	TCGGGTTGCGGGCGCAGGGCACGGGCG
6	hras1 G4 mut	TCGTGTTGCGTTCGCAGTGCACGTGCG
7	hras2 G4	CGGGGCGGGGCGGGGGCGGGGGCG
8	hras2 G4 mut	CGGTGCGTGGCGGGTGCGTGGGCG
9	Thymidine Kinase1 G4	GGTCGGCGCGGGAACCAGGGG
10	Thymidine Kinase1 G4 mut	GGTCGGCGCAAGAACCAGGGG
11	Thymidine Kinase2 G4	GGCCCCATGGCGGCGGGGCCGG
12	Thymidine Kinase2 G4 mut	GGCCCCATGACGGCGGGGCCGG
13	ds DNA 17	CCAGTTCGTAGTAACCC
14	ds DNA 26	CAATCGGATCGAATTCGATCCGATTG

Table 1. List of oligos used in the study.

All the oligonucleotide stocks were prepared in molecular biology grade nuclease free water at a concentration of 200µM and were stored at -20^0C.

c-MYC and hras2 promoter G-quadruplexes adopt a parallel conformation in K+ solution which is characterized by a negative peak at 240nm and a positive peak around 265nm in their CD signature (figure 3b and 3f) (Ambrus et al., 2005; Membrino et al., 2011; Yang and Hurley, 2006). The anti-parallel structure of the hras1 promoter quadruplex was confirmed by a negative and a positive peak at 260nm and 295 nm respectively (figure 3c) (Membrino et al., 2011). Telomeric repeat sequence, CDKN1A, hTERT and Thymidine Kinase (TKQ1 and TKQ2) promoter quadruplexes fold into a mixed quadruplex structure (figure 3a,3d and 3e) (Basundra et al., 2010; Micheli et al., 2010). The formation of G-quadruplex structures was further confirmed by comparing the CD signatures of respective oligonucleotides with their mutants (figure 3).
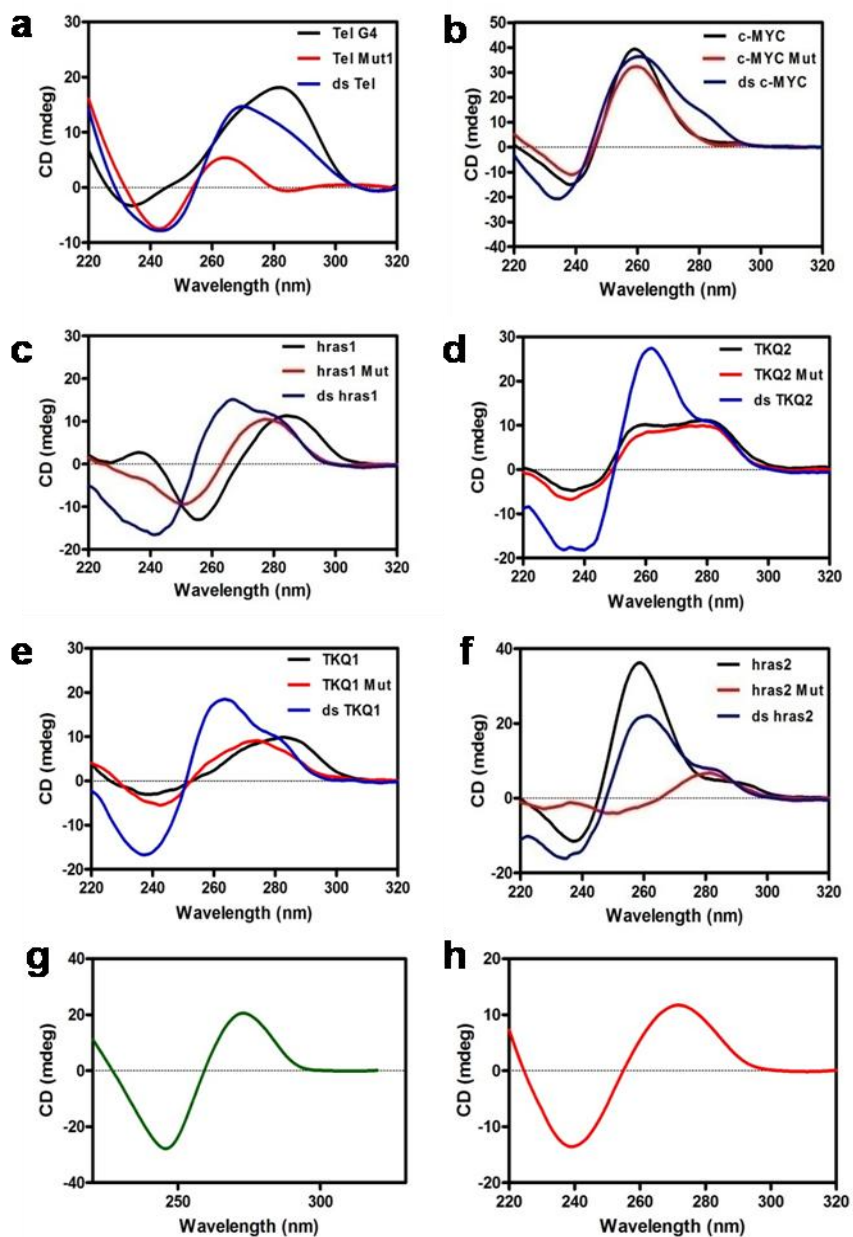
Figure 3: Confirmation of different G-quadruplex structures using circular dichroism Spectroscopy.

All the oligos were annealed at a concentration of 5µM in 10mM Sodium Cacodylate buffer containing 100mM Potassium Chloride by heating at 95^0C for 5 minutes and slow cooling to RT overnight. Double stranded oligos were annealed in 10mM Sodium Cacodylate buffer without any added salts. All CD spectra represent the average of 3 scans recorded from 220 nm to 320 nm at 25^0C and zero corrected at 320 nm. CD

[21]

signatures of (a) Hybrid Telomere (b) parallel c-MYC (c) anti-parallel hras1 (d) parallel hras2 (e) hybrid Thymidine Kinase 1 (TKQ1) and (f) hybrid Thymidine Kinase 2 (TKQ2) quadruplexes and their mutants. CD spectra of (g) ds 17 and (h) ds 26.

Biotinylated ligands compound1, compound4, compound 5 and compound b interact with structurally different G-quadruplexes *in vitro*.

The affinity of Biotinylated ligands towards structurally different quadruplexes were checked by G-quadruplex Fluorescent Intercalator Displacement Assay (G4-FID) (Monchaud and Teulade-Fichou, 2010). The principle behind the method is illustrated in figure 4. Annealed G-quadruplex, mutant or double stranded oligonucleotides were incubated with with Thiazole Orange (TO) till saturation and the samples were excited at 501 nm and the fluorescent spectra were recorded from 510 to 700 nm. Thiazole Orange in its free state, shows negligible fluorescence but when bound to nucleic acids it shows increased quantum yield of fluorescence. This property of TO was exploited here to measure the relative affinity of the ligands of interest towards structurally different G-quadruplexes and double stranded DNA.

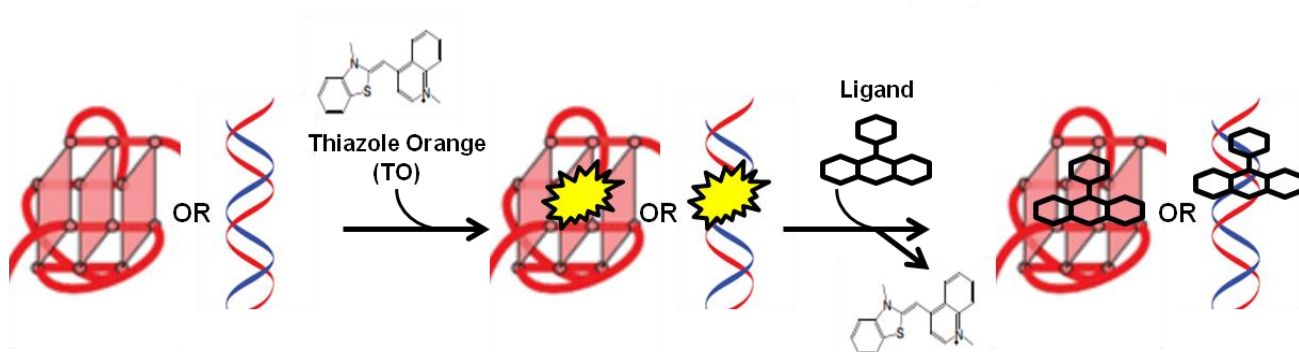**Figure 4: Principle of G-quadruplex Fluorescent Intercalator Displacement Assay.**
Thiazole Orange (TO) in its unbound state, shows negligible fluorescence. When bound to nucleic acids, the quatum yield of fluorescence is increased to nearly 10^4 folds. When a nucleic acid binding compound is introduced in to the solution, it replaces TO and binds to the nucleic acid structure. This results in a decrease of mean TO fluorescence

which is proportional to the affinity of the ligand towards the nucleic acid structure. Figure modified from (Kendrick and Hurley, 2010; Yaku et al., 2012)

Now, to dermine the relative affinity of the biotinylated compounds 1, 4, 5 and b towards structurally different G-quadruplexes, G4-FID with preformed quadruplexes was performed (figure 5 and 6). Fluorescence spectra of TO-G4 complex was recorded from 510 to 710nm after excitation at 501nm. Small amounts of the biotinylated ligands were then added to this complex and incubated for 3 minutes after each addition and the fluorescence was again measured. Addition of all the four ligands showed significant decrease in the TO fluorescence at low concentrations (figure 5 and 6). Thiazole orange has its emission maximum at 530nm. The change in the fluorescence intensity at 530nm was measured after each successive addition of the ligand in order to calculate the percentage TO displacement. The percentage TO displacement values were calculated as mentioned and was plotted against the concentrations of the ligands used. The DC50 value, which is the concentration of the ligand at which there is 50% displacement of TO from the TO-DNA complex, was calculated to compare the relative affinity of ligands towards different G-quadruplex structures. Representative displacement curves for Compound 1 and Compound b with structurally different G-quadruplexes are given in figure 5 and figure 6 respectively. The summary of G4-FID assay with biotinylated ligands compound 1 and compound b with structurally different quadruplexes is given in table 3 and table 4 respectively.

Figure 5: G4-FID displacement curves of compound 1 with structurally different G-quadruplexes.

Pre-formed quadruplex and double stranded oligonucleotides were used at a final concentration of 0.5µM in a 100µM cuvette. Two and three molar equivalents of

Thiazole orange was incubated with G4 and double stranded oligonucleotides respectively for five minutes at 25^0C. Small amounts of biotinylated ligands compound 1 was added to the TO-DNA complex and incubated for 3 minutes. Samples were excited at 501nm and the fluorescence spectra were recorded from 510 to 700nm after each addition. Percentage TO displacement was plotted against the concentration of ligand used and DC_{50} values were calculated. Representative G4-FID displacement curves ND TO displacement plots for Compound 1 with structurally different G4s: **(a)** Hybrid Telomeric G-quadruplex **(b)** Anti-parallel hras 1 promoter G-quadruplex **(c)** Parallel hras 2 promoter G-quadruplex and **(d)** double stranded DNA.

		DC 50 value (µM)		
Name	**Type of G4**	**Quadruplex**	**ds DNA**	**Mutant**
Telomere	Hybrid	1.00µM	4.00µM	2.50µM
c-MYC	Parallel	0.50µM	7.50µM	1.50µM
hras 1	Anti-parallel	1.00µM	3.00µM	1.50µM
hras 2	Parallel	1.00µM	4.50µM	1.25µM
TKQ1	Hybrid	1.50µM	3.00µM	1.50µM
TKQ2	Hybrid	1.50µM	4.00µM	2.25µM
ds 26	ds DNA	-	4.00µM	-

Table 2: Summary of G-quadruplex-FID assay of Compound 1 with structurally different G-quadruplexes.

G4-FID assay was performed with pre-formed structurally different G-quadruplexes, their mutants and double stranded DNA as described earlier. Displacement curves were plotted and percentage TO displacements were calculated. DC_{50} values were calculated and compared across samples to compare the affinity and selectivity of Compound 1 towards G-quadruplex structures.
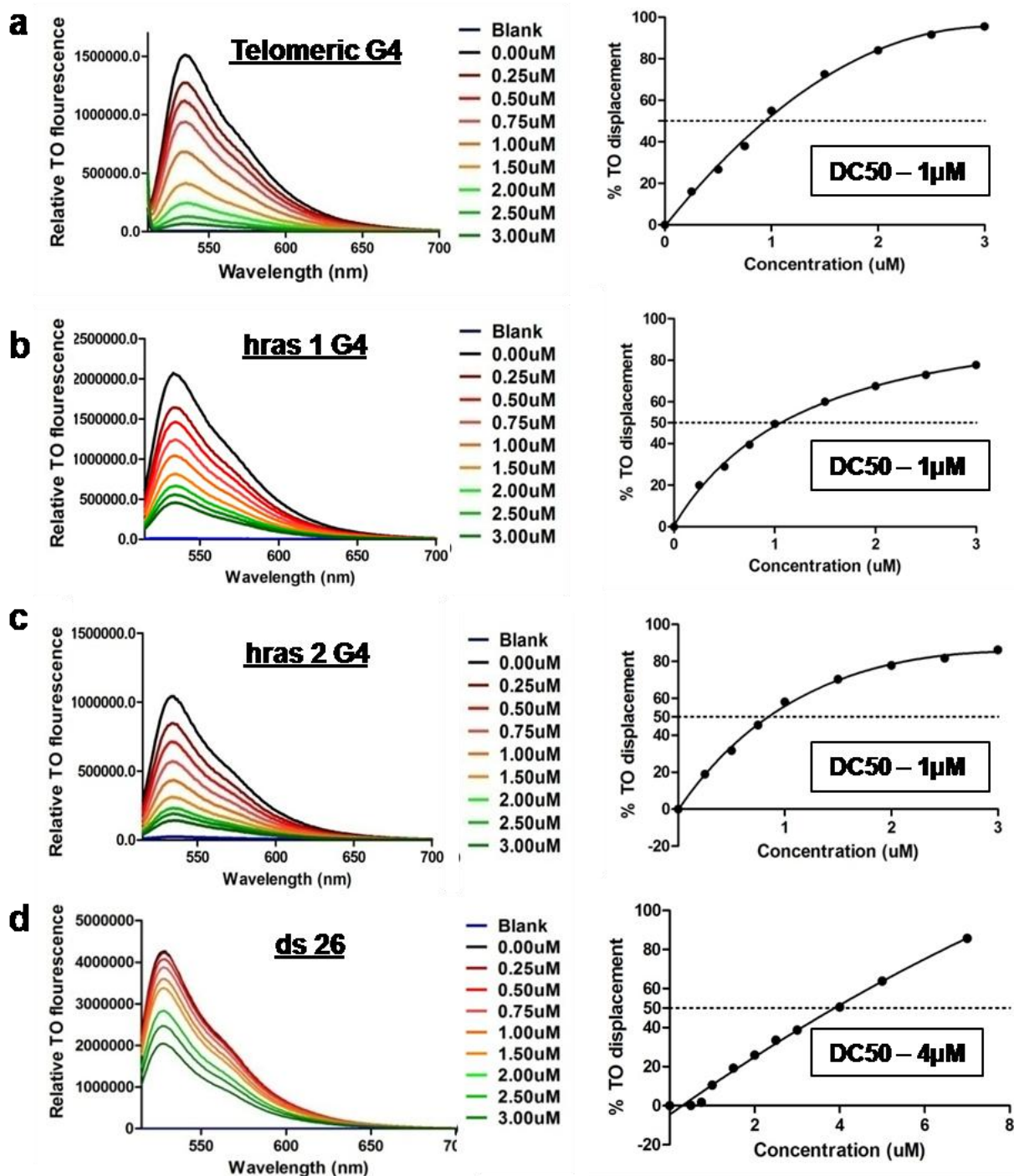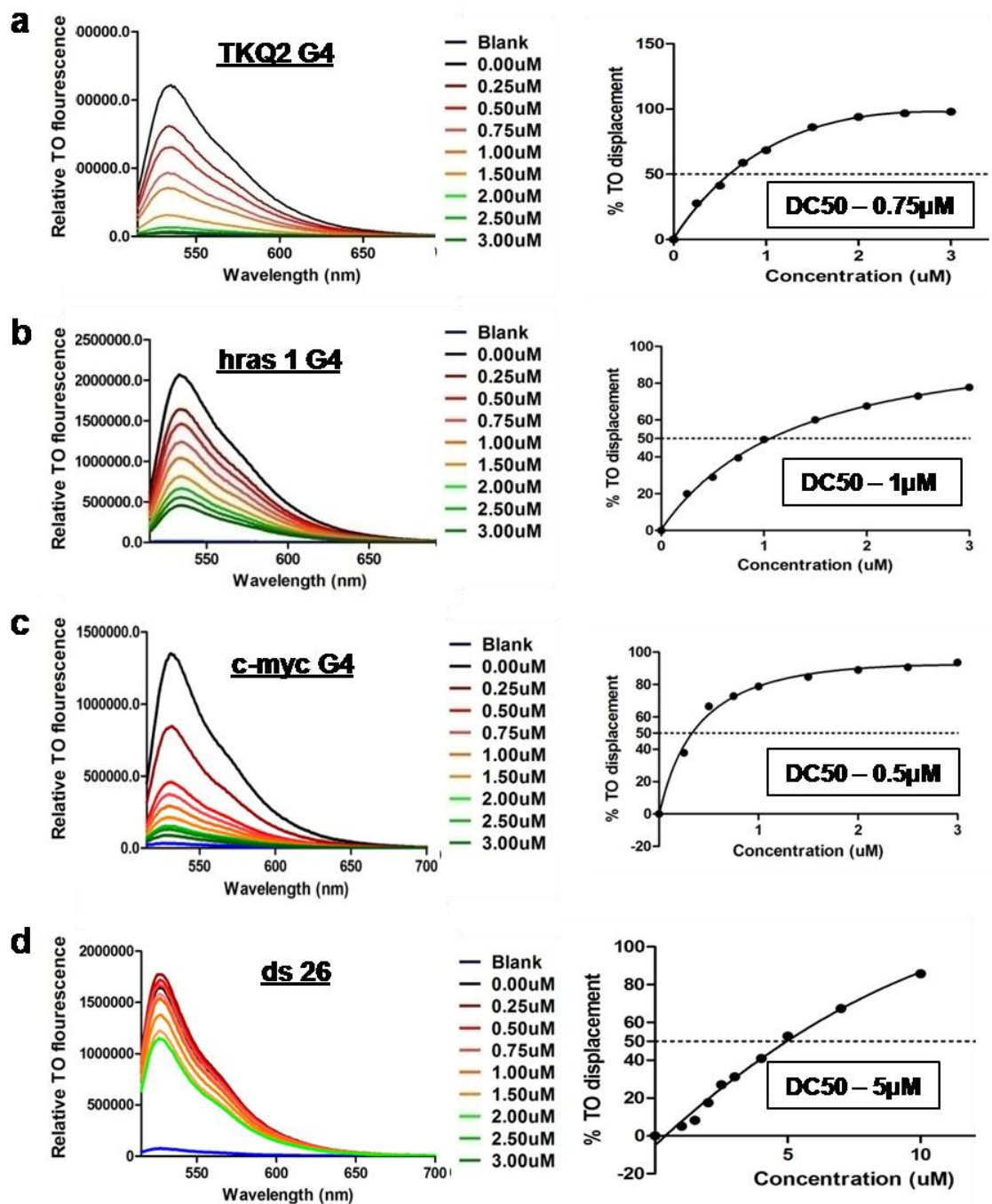
Figure 6: G4-FID displacement curves of compound b with structurally different G-quadruplexes.

Pre-formed quadruplex and double stranded oligonucleotides were used at a final concentration of 0.5µM in a 100µM cuvette. Two and three molar equivalents of

[26]

Thiazole orange was incubated with G4 and double stranded oligonucleotides respectively for five minutes at 25^0C. Small amounts of biotinylated ligands compound 1 was added to the TO-DNA complex and incubated for 3 minutes. Samples were excited at 501nm and the fluorescence spectra were recorded from 510 to 700nm after each addition. Percentage TO displacement was plotted against the concentration of ligand used and DC_{50} values were calculated. Representative G4-FID displacement curves and TO displacement plots for Compound 1 with structurally different G4s: **(a)** Hybrid Thymidine Kinase 2 promoter G-quadruplex **(b)** Anti-parallel hras 1 promoter G-quadruplex **(c)** Parallel c-MYC promoter G-quadruplex and **(d)** double stranded DNA.

		DC 50 value (µM)		
Name	**Type of G4**	**Quadruplex**	**ds DNA**	**Mutant**
Telomere	Hybrid	1.00µM	3.00µM	1.00µM
c-MYC	Parallel	0.50µM	3.00µM	0.50µM
hras 1	Anti-parallel	0.50µM	2.75µM	0.50µM
hras 2	Parallel	1.00µM	3.00µM	0.75µM
TKQ1	Hybrid	0.75µM	2.00µM	1.00µM
TKQ2	Hybrid	1.00µM	4.00µM	1.00µM
ds 26	ds DNA	-	5.00µM	-

Table 3: Summary of G-quadruplex-FID assay of Compound b with structurally different G-quadruplexes.

G4-FID assay was performed with pre-formed structurally different G-quadruplexes, their mutants and double stranded DNA as described earlier. Displacement curves were plotted and percentage TO displacements were calculated. DC_{50} values were calculated and compared across samples to compare the affinity and selectivity of Compound b towards G-quadruplex structures.

[27]

Biotinylated ligands decrease the viability of HT 1080 fibro sarcoma cells.

The effect of biotinylated ligands on the viability of HT 1080 fibro sarcoma cell cells were checked with CelltiterGlo cell viability assay. Cells were seeded at density of 4000 cells per well of a 96-well plate in 100µl complete media and were grown for 24 hours. Cells were treated with increasing concentrations of the ligands for 48 hours. Cell viability was measured after 48 hours using Cell titerglo luminescent cell viability assay kit. All the ligands showed significant inhibition of cell proliferation with IC_{50} values falling in the range of 50µM to 200µM (figure 7). The approximate IC_{50} values for all the biotinylated ligands in HT 1080 cells are summarized in table 5.
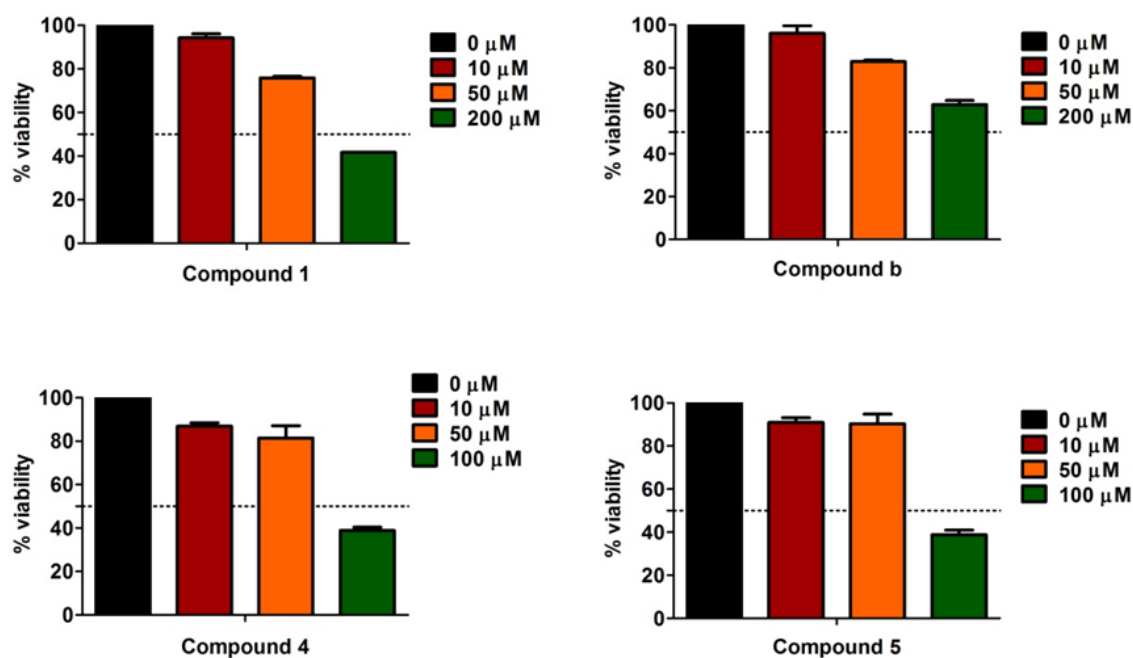
Figure 7: Cell viability of HT 1080 cells in presence of biotinylated G-quadruplex ligands.

HT 1080 cells were seeded at a density of 4000 cell per well of 96-well plate in 100µl complete medium and grown for 24 hours. Cells were then treated with serial dilutions of biotinylated ligands compound1, compound b, compound 4 and compound 5 for 48 hours. Cell viability after 48 hours was measured using Cell titerglo cell viability assay

as described. Graphs were plotted using Graphpad Prism 6. Data represents the average and standard deviation values from three or more independent experiments.

Compound	IC 50 range(µM)
Compound 1	150-200
Compound b	>200
Compound 4	75-100
Compound 5	75-100

Table 4: Approximate IC 50 values for biotinylated compounds in HT 1080 cells.

HT 1080 cells were treated with serial dilutions of biotinylated G-quadruplex binding ligands Compound 1, Compound b, Compound 4 and Compound 5 for a duration of 48 hours. Cell viability after 48 hours was measured using CellTiter-Glo® cell viability assay as described and the approximate IC 50 values were calculated.

Chromatin Immuno-precipitation with biotinylated G4 ligands

HT 1080 cells were seeded at a density of 10^5 cells /mL in a T25 culture flask in complete culture medium and were grown for 24 hours. Cells were then treated with either biotinylated ligands or DMSO for 48 hours. Cells were harvested, fixed with formaldehyde, lysed and the chromatin was sheared by sonication to generate fragments of size of approximately 500-1000 base pairs. The ligand bound DNA fragments were then pulled down using streptavidin coated magnetic beads and DNA fragments were recovered from the ligands. These fragments were then validated by PCR with primers for known putative G-quadruplex forming sequences such as the telomeric repeats, hTERT and p21 promoter G-quadruplexes. The methodology is summarised in the flowchart given in figure 8**.**

**Biotinylated
G4 ligand**

HT1080 cells

48 hours

**Formaldehyde fixation
and
Sonication**

**Fragments size
~500-1000 bp**

Streptavidin beads

**Magnetic separation of
bound fragments**

**Recovery of DNA fragments
From beads**

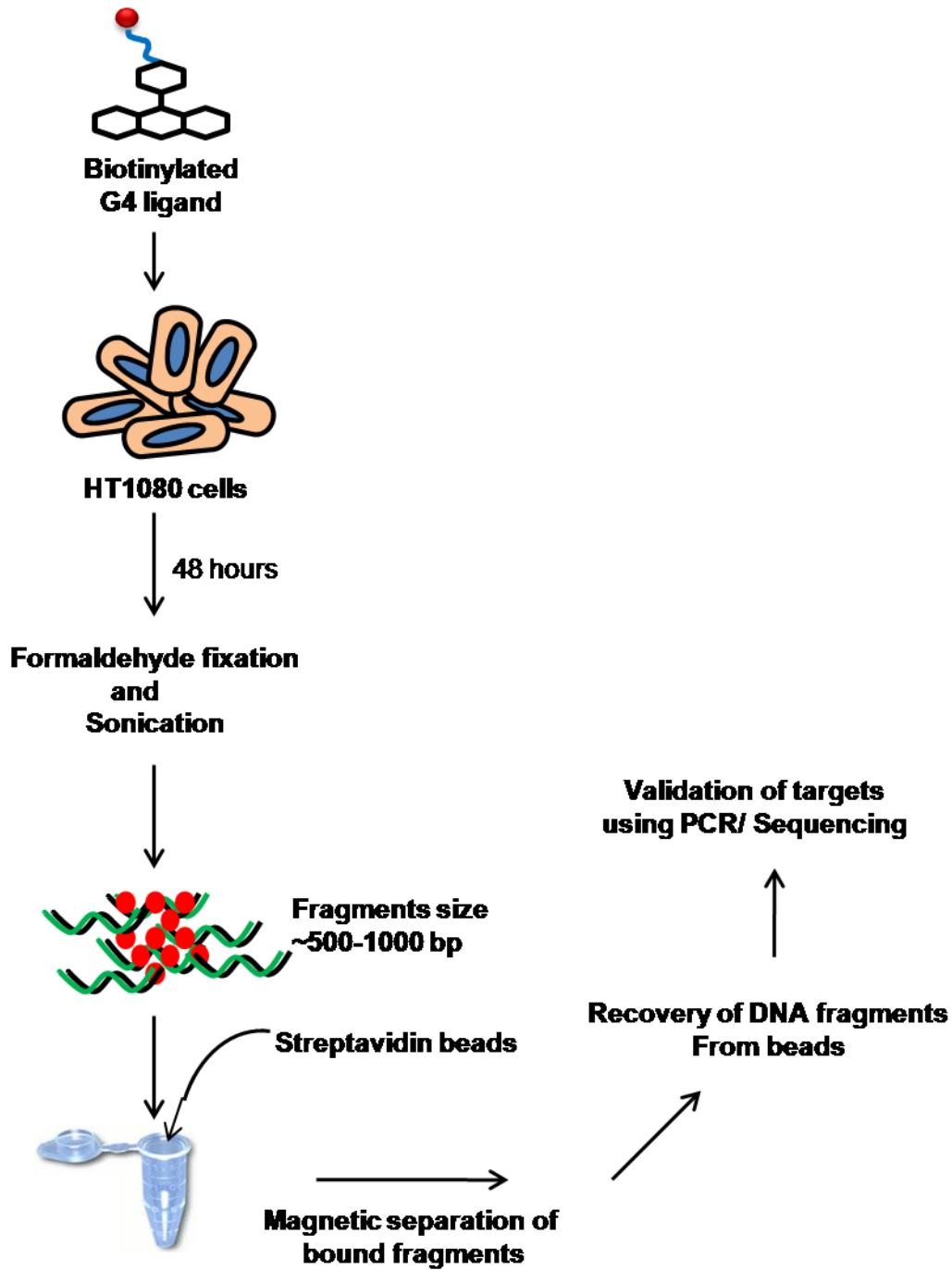**Validation of targets
using PCR/ Sequencing**

Figure 8: Schematic diagram showing the methodology of Chromatin Immuno-precipitation with biotinylated ligands.

All the four biotinylated Compounds 1, b, 4 and 5 showed significant selectivity towards different G-quadruplexes over double stranded DNA *in vitro* (figure 9 and 10). Now, to isolate the intrinsic G-quadruplex structures present in HT 1080 cells, chromatin immuno-precipitaion using biotinylated ligands was performed as described. The pulled down sequences were validated by PCR using primers for known potential G-quadruplex forming sequences in the genome such as the telomere repeats, p21 and hTERT promoter G-quadruplexes which were shown to interact with biotinylated ligands (figure 9 and 10).

Figure 9: Chromatin immuno-precipitation from HT 1080 cells with biotinylated compounds 4 and 5.

HT 1080 cells were treated with either DMSO or biotinylated G-quadruplex interacting ligands compound 4 and 5 for 48 hours. Cells were harvested after 48 hours and ChIP was performed as described earlier. **(a)** 1% agarose gel showing sonicated chromatin fragments of approximately 500 to 1000 base pair size **(b)** Agarose gel showing the

validation of ChIP using PCR primers for telomeric repeats, p21 and hTERT promoter G-quadruplexes.

Biotinylated compounds 4 and 5 were able to pull down telomeric repeats, p21 and hTERT quaduplexes from HT 1080 cells which was shown by significant amplification of the targets from the pulled down fragments (figure 9). But the amplification was also observed in the negative control for ChIP which is the fragments pulled down by only streptavidin beads. Thus, no significant enrichment of potential G-quadruplex structures was observed with compounds 4 and 5 over streptavidin beads (figure 9).
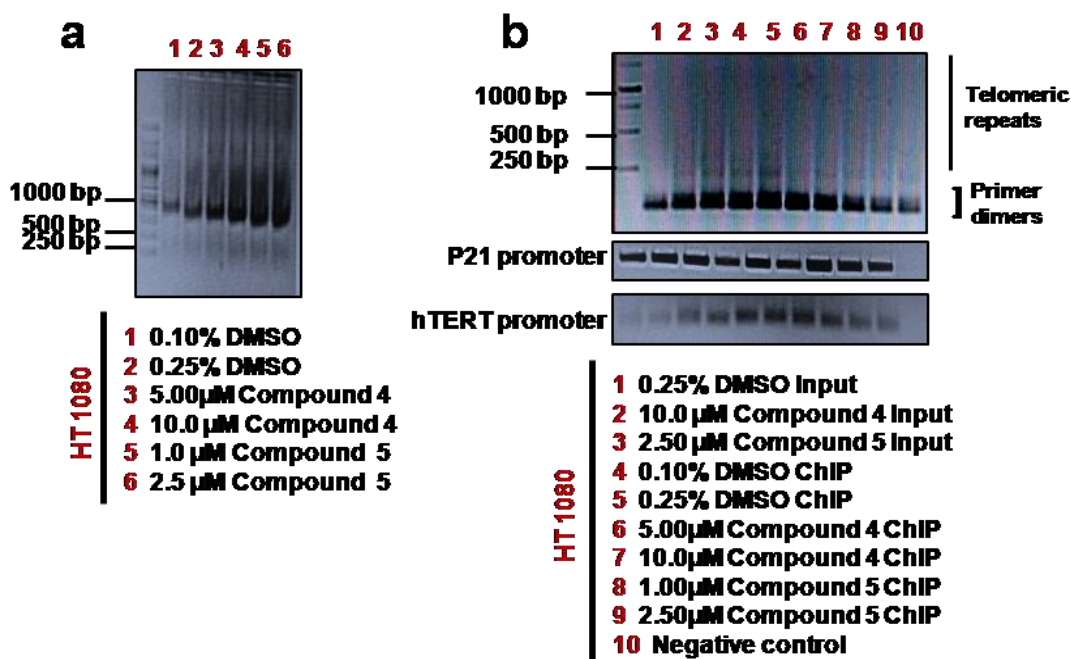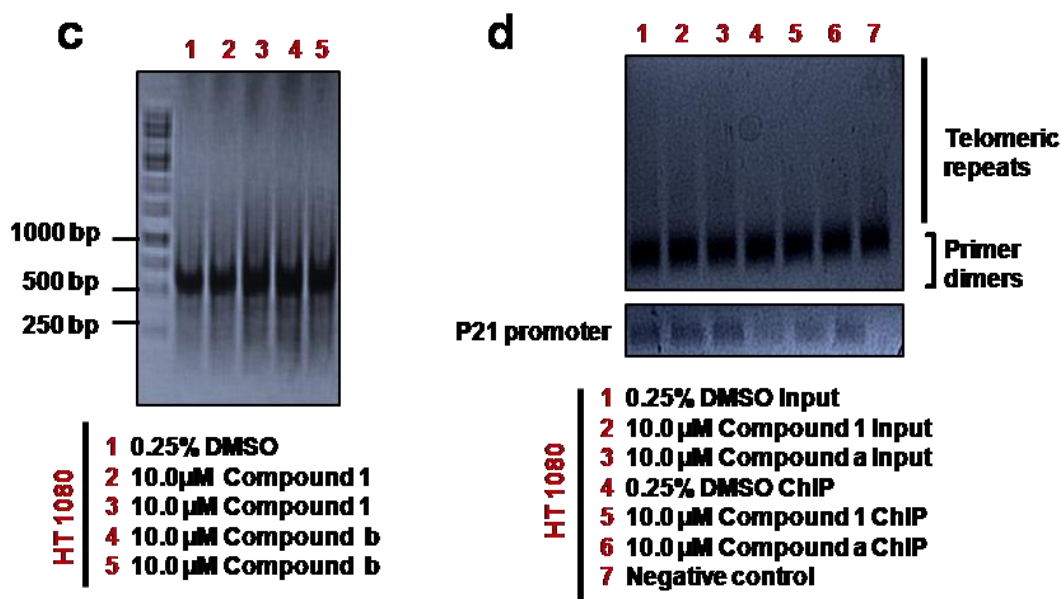

Figure 10: Chromatin immuno-precipitation from HT 1080 cells with biotinylated compounds 1 and b.

 HT 1080 cells were treated with either DMSO or biotinylated G-quadruplex interacting ligands compound 4 and 5 for 48 hours. Cells were harvested after 48 hours and ChIP was performed as described earlier. **(a)** 1% agarose gel showing sonicated chromatin fragments of approximately 500 to 1000 base pair size **(b)** Agarose gel showing the validation of ChIP using PCR primers for telomeric repeats, p21 promoter G-quadruplexes.

[32]

Chromatin precipitation was also performed with biotinylated G-quadruplex binding ligands compound 1 and compound b (figure 10). The pulled down fragments were then subjected to PCR with primers for telomeric repeats and p21 promoter G-quadruplex. No significant enrichment of the telomeric repeats was observed with the ligands as compared to the control. But a slight enrichment of p21 was observed in the fragments pulled down by compound1 and compound b (figure 10).

Effects of non-biotinylated ligands compound 2 and compound 3 in HT 1080 cells.

Compound 2 and compound 3 binds to structurally different G-quadruplexes.

Before checking the biological effects of non-biotinylated ligands in HT 1080 fibro-sarcoma cell line, the *in vitro* interaction of these ligands towards structurally different G-quadruplexes were checked using G4-FID assay. Both compound 2 and 3 displaces Thiazole orange from TO- G4 DNA complex at lower concentrations compared to that of TO- double stranded DNA complex (figure 11). Thus both the compounds interact selectively with pre-formed G-quadruplex structures *in vitro*.

Name	Type of G4	DC 50 (µM)
Telomere	Hybrid	1.50µM
c-MYC	Parallel	0.75µM
hras 1	Anti-parallel	1.50µM
ds 26	ds DNA	3.00µM

Table 5: Summary of G-quadruplex-FID assay of Compound 2 with structurally different G-quadruplexes.

G4-FID assay was performed with pre-formed structurally different G-quadruplexes, their mutants and double stranded DNA as described earlier. Displacement curves were plotted and percentage TO displacements were calculated. DC_{50} values were calculated and compared across samples to compare the affinity and selectivity of Compound 2 towards G-quadruplex structures.
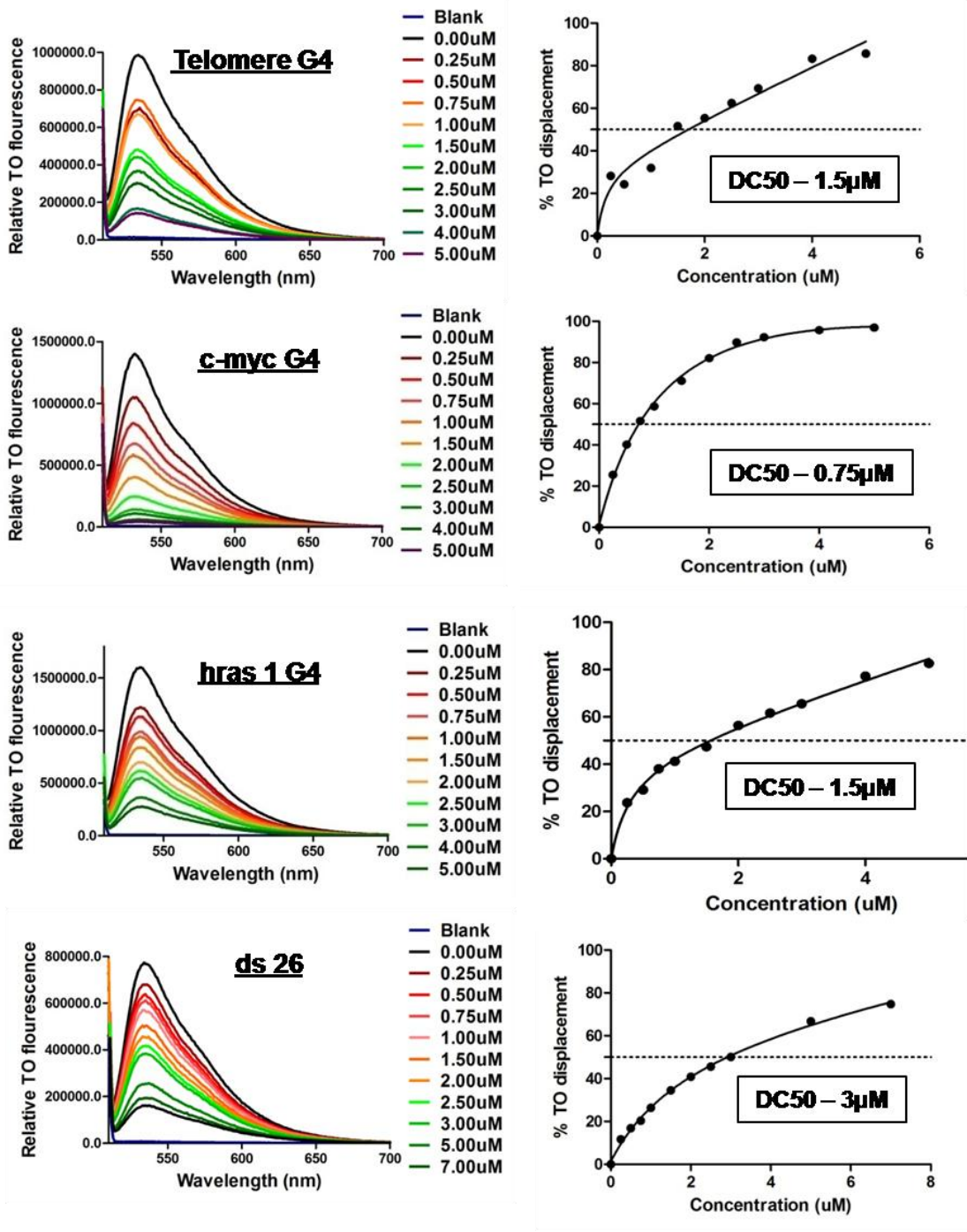
Figure 11: G4-FID displacement curves of compound 2 with structurally different G-quadruplexes.

Pre-formed quadruplex and double stranded oligonucleotides were used at a final concentration of 0.5µM in a 100µM cuvette. Two and three molar equivalents of

[35]

Thiazole orange was incubated with G4 and double stranded oligonucleotides respectively for five minutes at 25^0C. Small amounts of non-biotinylated ligand compound 2 was added to the TO-DNA complex and incubated for 3 minutes. Samples were excited at 501nm and the fluorescence spectra were recorded from 510 to 700nm after each addition. Percentage TO displacement was plotted against the concentration of ligand used and DC_{50} values were calculated. Representative G4-FID displacement curves and TO displacement plots for Compound 2 with structurally different G4s: **(a)** Hybrid Telomeric G-quadruplex **(b)** Parallel c-MYC promoter G-quadruplex **(c)** Anti-parallel hras 1 promoter G-quadruplex and **(d)** double stranded DNA, ds 26.

Compound 2 and 3 decreases the cell viability of HT 1080 cells.

We first checked whether the ligands are getting internalized into HT 1080 cells using phase contrast microscopy (figure 12). Cells were seeded at a density of 10^5 cells per ml of media supplemented 10% FBS and were grown for 24 hours. Cells were the treated with either DMSO or compounds 2 and 3 for a period of 24 hours. Ligand treated cells were imaged using blue light filter, since the excitation-emission wavelengths of the ligands fall in the range of blue light. Both the ligands get internalized into HT 1080 cells and previous studies done in the lab show that they localize in the nucleus (figure 12).
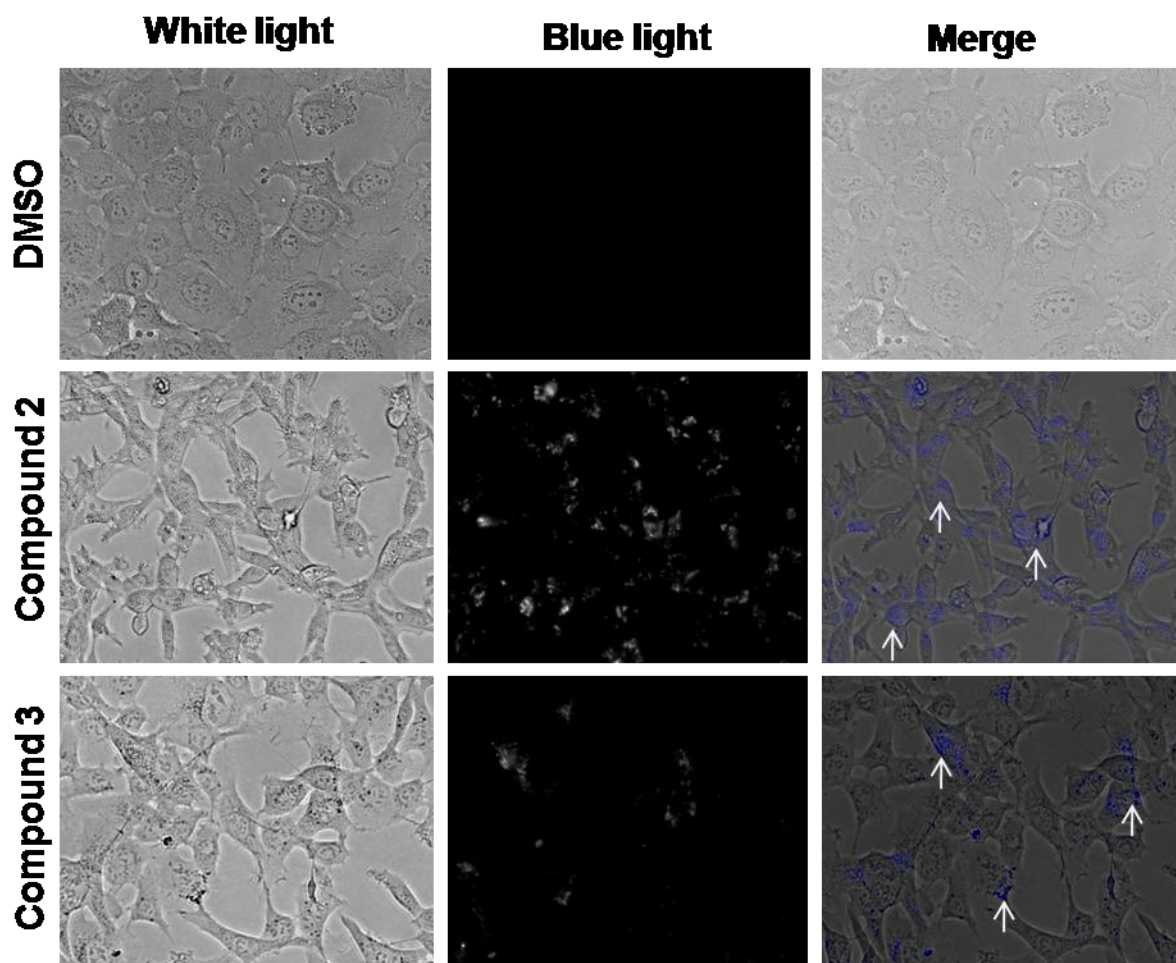
Figure 12: Image showing the cellular internalization of Compound 2 and Compound 3 in HT 1080 cells.

Now to check the effects of quadruplex interacting compounds 2 and 3 on cell viability of HT 1080 fibro-sarcoma cells, Celltiterglo cell viability assay was performed (figure 13). Cells were seeded at a density of 4000 cells per well of a 96 well plate in complete culture media for 24 hours. cells were then treated with serial dilutions of compound 2 and 3 for 48 hours and the cell viability was measured. Both the ligands showed significant decrease in the viability of HT 1080 cells after 48 hours of treatment. The approximate IC_{50} concentrations of compound 2 and compound 3 are 100µM and 50µM respectively.
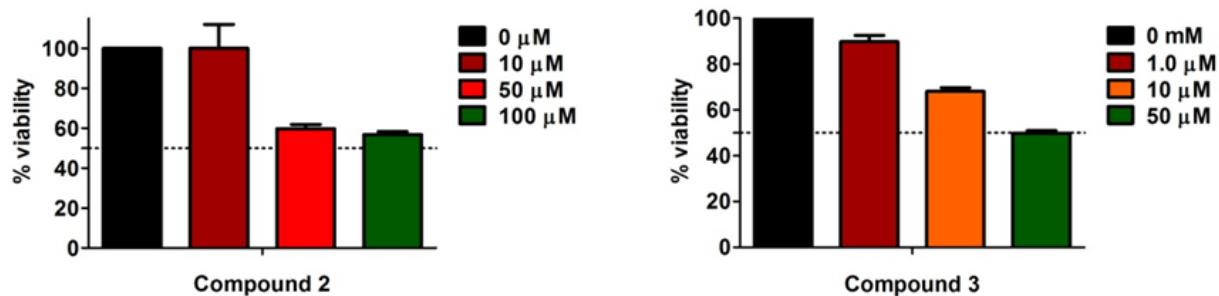
Figure 13: Compound 2 and 3 decreases the viability of HT 1080 cells.

HT 1080 cells were seeded at a density of 4000 cell per well of 96-well plate in 100µl complete medium and grown for 24 hours. Cells were then treated with serial dilutions of ligands compound 2 and compound 3 for 48 hours. Cell viability after 48 hours was measured using CellTiter-Glo® cell viability assay as described. Graphs were plotted in Graphpad Prism 6. Data represents the average and standard deviation values from three or more independent experiments.

Discussion

The non-canonical secondary structures formed from guanine rich nucleotide sequences called the G-quadruplex structures have attracted the attention of many chemists and biologists in the last decade (Bochman et al., 2012). Many computational algorithms which search for specific guanine rich motifs in the human genome has revealed that potential quadruplex forming sequences are present throughout the human genome (Huppert and Balasubramanian, 2005). The non random distribution and the higher frequency of occurrence in functionally important regions of the genome such as telomeres and promoters of various genes of such motifs suggests their possible roles in various biological processes (Murat and Balasubramanian, 2014). Many *in vitro* studies have shown that these structures are stable under near physiological conditions (Lam et al., 2013).

Studies have been carried out to elucidate the possible roles of G-quadruplex structures in different cellular processes but majority of them are done under *ex vivo* experimental conditions (Di Antonio et al., 2012). Interestingly a very good part of the scientific community does not believe in the existence of such structures in cells. In silico studies have shown that there are more than 3,70,000 potential G-quadruplex forming motifs present across the human genome (Huppert and Balasubramanian, 2005). But how many of such motifs actually form a G-quadruplex *in vivo* remains elusive. Many approaches have been used to identify intrinsic G-quadruplexes in cells which include the use quadruplex specific antibodies, small organic molecules selectively binding to quadruplexes (Biffi et al., 2013; Di Antonio et al., 2012; Muller et al., 2010; Rodriguez et al., 2012Rodriguez et al., 2012). But there is a huge discrepancy in the number of quadruplexes identified in comparison to the computational data. More over many of such studies were as good as *in vitro* experiments since they were carried out on genomic DNA isolated from human cells.

We started with checking the affinity and selectivity of biotinylated ligands towards structurally different G-quadruplexes using G4-FID. Different G-quadruplex structures were prepared and the conformations were confirmed using circular dichroism spectroscopy (figure 3). These pre-formed G4s were then used for G4-FID assay. All

[39]

the four ligands used in the study showed strong binding and selectivity towards G-quadruplex structures compared to double stranded DNA. The ligands But only a marginal difference was observed in the affinity of these ligands towards G4s and their respective mutants.

We then went ahead with isolation of intrinsic G-quadruplexes from HT 1080 cells. Cells were treated with biotinylated ligands for 48 hours and chromatin immuno-precipitation was done as described. In order to validate the targets of the ligands, PCR with primers for known potential G-quadruplexes was carried out. We used primers for G-quadruplex motif present in the telomeric repeats, p21 and hTERT gene promoters. ChIP with compounds 4 and 5 did not show any significant enrichment of target sequences as compared to the control pull down. Amplification of the targets were observed in DNA fragments bound to streptavidin beads (negative control). This could possibly be a non-specific interaction. We could not confirm this because of the poor yield of the pull down.

But ChIP with the biotinylated ligands Compound 1 and Compound b showed the enrichment of p21 promoter G4 over streptavidin pull down. But no significant enrichment of telomeric repeats was observed. Even though the pull downs were done multiple times, the results were not consistent.

All the ligands showed significant binding and selectivity towards structurally different G-quadruplexes over double stranded DNA but failed to pull down them from cells. This might be either because of the inability of these ligands to bind specifically to G-quadruplexes *in vivo* or because of technical problems in DNA isolation as the yields of the pull downs were consistently poor. Another study from Shankar Balasubramanian and colleagues has also mentions about the technical problems regarding isolation of G4 from complete cell extracts using biotinylated ligands (Di Antonio et al., 2012). This demands the need of alternative approaches for isolation of cellular G4s.

References

Agrawal, P., Hatzakis, E., Guo, K.X., Carver, M., and Yang, D.Z. (2013). Solution structure of the major G-quadruplex formed in the human VEGF promoter in K+: insights into loop interactions of the parallel G-quadruplexes. Nucleic Acids Res *41*, 10584-10592.

Ambrus, A., Chen, D., Dai, J.X., Jones, R.A., and Yang, D.Z. (2005). Solution structure of the biologically relevant g-quadruplex element in the human c-MYC promoter. implications for g-quadruplex stabilization. Biochemistry-Us *44*, 2048-2058.

Basundra, R., Kumar, A., Amrane, S., Verma, A., Phan, A.T., and Chowdhury, S. (2010). A novel G-quadruplex motif modulates promoter activity of human thymidine kinase 1. Febs Journal *277*, 4254-4264.

Biffi, G., Tannahill, D., McCafferty, J., and Balasubramanian, S. (2013). Quantitative visualization of DNA G-quadruplex structures in human cells. Nat Chem *5*, 182-186.

Bochman, M.L., Paeschke, K., and Zakian, V.A. (2012). DNA secondary structures: stability and function of G-quadruplex structures. Nature reviews Genetics *13*, 770-780.

Burge, S., Parkinson, G.N., Hazel, P., Todd, A.K., and Neidle, S. (2006). Quadruplex DNA: sequence, topology and structure. Nucleic acids research *34*, 5402-5415.

Burger, A.M., Dai, F.P., Schultes, C.M., Reszka, A.P., Moore, M.J., Double, J.A., and Neidle, S. (2005). The G-quadruplex-interactive molecule BRACO-19 inhibits tumor growth, consistent with telomere targeting and interference with telomerase function. Cancer Res *65*, 1489-1496.

Di Antonio, M., Rodriguez, R., and Balasubramanian, S. (2012). Experimental approaches to identify cellular G-quadruplex structures and functions. Methods *57*, 84-92.

Huppert, J.L., and Balasubramanian, S. (2005). Prevalence of quadruplexes in the human genome. Nucleic Acids Res *33*, 2908-2916.

Huppert, J.L., and Balasubramanian, S. (2007). G-quadruplexes in promoters throughout the human genome. Nucleic acids research *35*, 406-413.

[41]

Incles, C.M., Schultes, C.M., Kempski, H., Koehler, H., Kelland, L.R., and Neidle, S. (2004). A G-quadruplex telomere targeting agent produces p16-associated senescence and chromosomal fusions in human prostate cancer cells. Molecular cancer therapeutics *3*, 1201-1206.

Islam, M.A., Thomas, S.D., Murty, V.V., Sedoris, K.J., and Miller, D.M. (2014). c-Myc Quadruplex-forming Sequence Pu-27 Induces Extensive Damage in Both Telomeric and Nontelomeric Regions of DNA. J Biol Chem *289*, 8521-8531.

Kendrick, S., and Hurley, L.H. (2010). The role of G-quadruplex/i-motif secondary structures as cis-acting regulatory elements. Pure and applied chemistry Chimie pure et appliquee *82*, 1609-1621.

Kim, M.Y., Vankayalapati, H., Kazuo, S., Wierzba, K., and Hurley, L.H. (2002). Telomestatin, a potent telomerase inhibitor that interacts quite specifically with the human telomeric intramolecular G-quadruplex. J Am Chem Soc *124*, 2098-2099.

Lam, E.Y., Beraldi, D., Tannahill, D., and Balasubramanian, S. (2013). G-quadruplex structures are stable and detectable in human genomic DNA. Nat Commun *4*, 1796.

Lipps, H.J., and Rhodes, D. (2009). G-quadruplex structures: in vivo evidence and function. Trends in cell biology *19*, 414-422.

Membrino, A., Cogoi, S., Pedersen, E.B., and Xodo, L.E. (2011). G4-DNA Formation in the HRAS Promoter and Rational Design of Decoy Oligonucleotides for Cancer Therapy. Plos One *6*.

Micheli, E., Martufi, M., Cacchione, S., De Santis, P., and Savino, M. (2010). Self-organization of G-quadruplex structures in the hTERT core promoter stabilized by polyaminic side chain perylene derivatives. Biophys Chem *153*, 43-53.

Monchaud, D., and Teulade-Fichou, M.P. (2010). G4-FID: a fluorescent DNA probe displacement assay for rapid evaluation of quadruplex ligands. Methods in molecular biology *608*, 257-271.

Muller, S., Kumari, S., Rodriguez, R., and Balasubramanian, S. (2010). Small-molecule-mediated G-quadruplex isolation from human cells. Nat Chem *2*, 1095-1098.

Muller, S., Sanders, D.A., Di Antonio, M., Matsis, S., Riou, J.F., Rodriguez, R., and Balasubramanian, S. (2012). Pyridostatin analogues promote telomere dysfunction and

long-term growth inhibition in human cancer cells. Organic & Biomolecular Chemistry *10*, 6537-6546.

Murat, P., and Balasubramanian, S. (2014). Existence and consequences of G-quadruplex structures in DNA. Curr Opin Genet Dev *25*, 22-29.

Neidle, S. (2010). Human telomeric G-quadruplex: The current status of telomeric G-quadruplexes as therapeutic targets in human cancer. Febs J *277*, 1118-1125.

Rhodes, D., and Lipps, H.J. (2015). G-quadruplexes and their regulatory roles in biology. Nucleic Acids Res *43*, 8627-8637.

Riou, J.F., Guittat, L., Mailliet, P., Laoui, A., Renou, E., Petitgenet, O., Megnin-Chanet, F., Helene, C., and Mergny, J.L. (2002). Cell senescence and telomere shortening induced by a new series of specific G-quadruplex DNA ligands. P Natl Acad Sci USA *99*, 2672-2677.

Rodriguez, R., Miller, K.M., Forment, J.V., Bradshaw, C.R., Nikan, M., Britton, S., Oelschlaegel, T., Xhemalce, B., Balasubramanian, S., and Jackson, S.P. (2012). Small-molecule-induced DNA damage identifies alternative DNA structures in human genes. Nature chemical biology *8*, 301-310.

Siddiqui-Jain, A., Grand, C.L., Bearss, D.J., and Hurley, L.H. (2002). Direct evidence for a G-quadruplex in a promoter region and its targeting with a small molecule to repress c-MYC transcription. Proceedings of the National Academy of Sciences of the United States of America *99*, 11593-11598.

Tauchi, T., Shin-ya, K., Sashida, G., Sumi, M., Okabe, S., Ohyashiki, J.H., and Ohyashiki, K. (2006). Telomerase inhibition with a novel G-quadruplex-interactive agent, telomestatin: in vitro and in vivo studies in acute leukemia. Oncogene *25*, 5719-5725.

Verma, A., Yadav, V.K., Basundra, R., Kumar, A., and Chowdhury, S. (2009). Evidence of genome-wide G4 DNA-mediated gene expression in human cancer cells. Nucleic acids research *37*, 4194-4204.

Wang, Q., Liu, J.Q., Chen, Z., Zheng, K.W., Chen, C.Y., Hao, Y.H., and Tan, Z. (2011). G-quadruplex formation at the 3' end of telomere DNA inhibits its extension by telomerase, polymerase and unwinding by helicase. Nucleic Acids Res *39*, 6229-6237.

Yaku, H., Fujimoto, T., Murashima, T., Miyoshi, D., and Sugimoto, N. (2012). Phthalocyanines: a new class of G-quadruplex-ligands with many potential applications. Chemical communications *48*, 6203-6216.

Yang, D.Z., and Hurley, L.H. (2006). Structure of the biologically relevant G-quadruplex in the c-MYC promoter. Nucleos Nucleot Nucl *25*, 951-968.

Yuan, L.B., Tian, T., Chen, Y., Yan, S.Y., Xing, X.W., Zhang, Z.A., Zhai, Q.Q., Xu, L., Wang, S.O., Weng, X.C.*, et al.* (2013). Existence of G-quadruplex structures in promoter region of oncogenes confirmed by G-quadruplex DNA cross-linking strategy. Sci Rep-Uk *3*.

Zhou, J.M., Zhu, X.F., Lu, Y.J., Deng, R., Huang, Z.S., Mei, Y.P., Wang, Y., Huang, W.L., Liu, Z.C., Gu, L.Q.*, et al.* (2006). Senescence and telomere shortening induced by novel potent G-quadruplex interactive agents, quindoline derivatives, in human cancer cell lines. Oncogene *25*, 503-511.