Targeting a noncanonical, hairpin-containing G-quadruplex structure from the MYCN gene

Mo Yang1, Sakereh Carter2, Shaifaly Parmar1, Desta D. Bume1, David R. Calabrese1, Xiao Liang1, Kamyar Yazdani1, Man Xu2, Zhihui Liu2, Carol J. Thiele2 and John S. Schneekloth, Jr1,*

1Chemical Biology Laboratory, Center for Cancer Research, National Cancer Institute, Frederick, MD 21702-1201, USA and 2Pediatric Oncology Branch, Center for Cancer Research, National Cancer Institute, Bethesda, MD 20892-1928, USA

Received March 05, 2021; Revised June 10, 2021; Editorial Decision June 24, 2021; Accepted July 20, 2021

ABSTRACT

The MYCN gene encodes the transcription factor N-Myc, a driver of neuroblastoma (NB). Targeting G-quadruplexes (G4s) with small molecules is an attractive strategy to control the expression of undruggable proteins such as N-Myc. However, selective binders to G4s are challenging to identify due to the structural similarity of many G4s. Here, we report the discovery of a small molecule ligand (4) that targets the noncanonical, hairpin containing G4 structure found in the MYCN gene using small molecule microarrays (SMMs). Unlike many G4 binders, the compound was found to bind to a pocket at the base of the hairpin region of the MYCN G4. This compound stabilizes the G4 and has affinity of 3.5 ± 1.6 μM. Moreover, an improved analog, MY-8, suppressed levels of both MYCN and MYCNOS (an lncRNA embedded within the MYCN gene) in NBEB neuroblastoma cells. This work indicates that the approach of targeting complex, hybrid G4 structures that exist throughout the human genome may be an applicable strategy to achieve selectivity for targeting disease-relevant genes including protein coding (MYCN) as well as non-coding (MYCNOS) gene products.

INTRODUCTION

The MYC family of genes encode transcription factors that broadly govern and amplify gene expression (1). Three Myc proteins, c-Myc, N-Myc and L-Myc, contain basic helix-loop-helix (bHLH) regions that bind to DNA and directly regulate transcription (2). Among these genes, MYCN has been shown to be critical to fetal development and is highly expressed in neural tissue (3). MYCN is often overexpressed or mutated in cancers and is considered an oncogene, particularly in neuroblastoma and small cell lung cancers (4). Embedded within the MYCN locus is a second transcript that is produced through antisense transcription and originally called ncyM which was shown to be within a larger long noncoding RNA (lncRNA) called MYCN opposite-strand...
(MYCNOS) (5–7). Expression of MYCNOS has been shown to decrease promoter occupancy of MYCN and control the stability of the MYCN transcript through sequence complementarity. More recent studies have shown that NCYM, the variant 2 of the MYCNOS transcript can also be translated to produce a small protein (8). This polypeptide, known as NCYM, has been characterized to have diverse regulatory activities, including controlling the stability of N-Myc protein, inhibiting GSK3β, and influencing Wnt/β-catenin signaling (8,9). The specific detailed regulatory roles of MYCNOS and NCYM remain under active investigation.

It is clear that a complex interplay exists between MYCN, NCYM and OCT4, and the expression of these genes has a direct impact on N-Myc-driven cancers (10). Given the important role of N-Myc in cancer, molecules that control or inhibit the N-Myc protein could potentially be of interest as anticancer agents. However, as a transcription factor, the N-Myc protein is classically considered ‘undruggable’ due to the flexibility of its helix-loop-helix structure, and attempts to develop ligands that target N-Myc itself have not been successful due to a lack of defined small molecule binding pockets. An alternative approach involves the development of molecules that target Aurora Kinase A (11). These molecules block the interaction of Aurora A and N-Myc, enabling N-Myc to be degraded by a ubiquitin-dependent pathway. So far, there are no examples of molecules that target MYCNOS or NCYM. One approach to target ‘undruggable’ genes is to develop molecules that control their expression, for example by targeting the gene itself (12,13). In this area, non-B DNA structures represent an important class of targets for small molecules that influence gene expression (14–16). Among the most studied examples in this class are G-quadruplexes (G4s): non-canonical structures that form by Hoogsteen-bonded guanines in G-rich sequences of DNA and RNA (17,18). G4s have been shown to broadly exist in human cells (19–21), and some of them play pivotal roles in modulating the transcription and translation of genes, especially cancer-driving genes (22,23). While extensive work has gone into discovering small molecules that target G4s (24–29), the development of selective ligands has remained a challenge.

Many canonical G4 structures are relatively simple: they contain tetrads of guanines that are stabilized by central potassium ions, and small 1–7 nucleotide loops (30). However, recent examples of more complex G4s have also been shown to contain additional structural elements, such as hairpins (31–33). Among the reported G4s in MYCN (34–36), one example of G4 with such a structure is found in intron 1 of MYCN (37). Importantly, genome-wide probing studies have experimentally confirmed the existence of a G4 in the MYCN gene, though it does not occur in the promoter region of MYCN like many other regulatory G4s (38). Hairpin containing G4 structures can be considerably more structurally complex than simple G4s, and thus may represent opportunities for selective targeting with small molecules as they can contain unique folds and pockets. Indeed, it has been observed that highly selective small molecule binding to RNA G4s in aptamers can be achieved when more complex or unique structural elements are incorporated near the G4. This is particularly evident in the structures of fluorescent RNA aptamers Spinach and Mango (39), which rely on G4-mediated recognition of small molecule fluorophores. Although considerably more rare than simple G4s, hairpin containing G4s have been calculated and predicted throughout the genome (40).

Here, we report multiple classes of small molecule ligands that target the hairpin containing G4 structure found in the MYCN gene. We used a small molecule microarray (SMM) approach to identify several classes of small molecules that bind directly to the MYCN G4 with preference over other nucleic acid structures. In-depth biophysical analyses demonstrated that these molecules have distinct modes of recognition to the G4. Of particular interest are a class of heterocycles that bind near the G4-hairpin junction. Treatment of NBEB neuroblastoma (NB) cells with this compound resulted in decreases in both the MYCN and MYCNOS transcripts as well as the N-Myc protein levels. Thus, this represents an intriguing case of impacting the expression of both an oncogenic mRNA and associated lncRNA simultaneously with a small molecule. This work demonstrates that targeting complex, hairpin containing G4 structures such as the one found in intron 1 of MYCN may be a broadly applicable strategy to identify selective DNA G4 binders by targeting their unique structures.

MATERIALS AND METHODS

General information

MYCN G4 oligonucleotides (5′-AGG GGG TGG GAG GGG GCA TGC AGA TGC AGG GGG T-3′) with or without labeling were purchased from Integrated DNA Technology. Other sequences of DNA/RNA samples used for binding specificity study are summarized in supplementary material (Supplementary Table S1 and S2).

Compounds for binding validation were purchased from ChemDiv and ChemBridge. The RNA extraction kit RNeasy Plus Mini kit (#74134) was purchased from Qiagen, CA. The cDNA synthesis kit High-Capacity RNA-to-cDNA™ Kit (#4387406) was purchased from Life Technologies (Invitrogen, New York). The Fast SYBR™ Green Master Mix was purchased from Life Technologies (Invitrogen, NY). The anti-N-Myc (B8.4.B): sc-53993 antibody (1:4000 dilution; Santa Cruz Biotechnology (SCB), USA), the anti-GAPDH (0411); sc-47724 antibody (1:2000 dilution; SCB, USA), and the goat anti-mouse IgG-HRP: sc-2005 antibody (1:1000 dilution; SCB, USA) were used for immunoblots.

Small molecule microarray screening

Small molecule microarrays (SMMs) were prepared on 2D epoxy glass slides (Schott). Briefly, ~15 000 compounds (10 mM in DMSO) from ChemDiv and ChemBridge libraries as well as control dyes (Alexa Fluor 647 and Alexa Fluor 488, 10 μM in DMSO, Invitrogen) were prepared in 384 well plates (ArrayJet Jetstar™). Each of the compounds (10 mM in DMSO) was printed in duplicate blocks by ArrayJet robotic microarray printer. After printing, glass slides were incubated in the arrayer overnight for immobilization, followed by vacuum drying for 24 h. To quench
the unreacted epoxy groups, the slides were incubated in 1 M ethanolamine aqueous solution (pH 8.5) for 2 h, then thoroughly rinsed with DMF and ddH2O, followed by N2 drying.

To begin with the screening against the target of interest, 5′-Cy5-labeled MYCN G4 DNA sample was firstly prepared in the annealing buffer (10 mM Tris, pH 7.0, 100 mM KCl) at 10 μM concentration. The DNA sample was incubated in a heating block at 95°C for 5 min, then slowly cooled down to room temperature (rt) for more than 1 h. Then the folded DNA was diluted into the screening buffer (10 mM Tris, pH 7.0, 100 mM KCl, 0.005% Tween 20) to meet the final concentration of 50 nM. A SMM slide was placed into a four-well slide holder and incubated with 3 ml 10× tRNA (500 nM in screening buffer) for 2 h. Then, the slide was carefully washed by screening buffer and incubated with 3 ml MYCN G4 DNA solution (50 nM DNA with 500 nM tRNA) for another 2 h. After incubation, the slide was transferred into a 50 ml conical tube and gently washed with PBST and ddH2O for 3 times in each buffer solution. Then, the slide was dried by centrifuging at 1700 g for 2 min. Finally, the slide with SMM was imaged by a fluorescence scanner (InnoScan 1100 AL) at 635 nm, with a resolution of 5 μm. The fluorescent intensity of each spot was quantified by Innopsys Mapix software and the hits were identified based on the criteria reported previously (41,42). To identify hits, another slide only incubated with screening buffer was also tested and imaged using the same method as a negative control.

Fluorescence intensity assay (FIA)
In this study, FIA was used for both validation of hit binding at a single dose and binding affinity determination. For validating hit compounds from SMM screening, 100 μM solution of each compound was prepared in triplicate in a 96-well plate (Costar, black side clear bottom), resulting in 5% DMSO concentration. 5′-Cy5-labeled MYCN G4 DNA was annealed as described above, then added into the well plate resulting in a final concentration of 100 nM. The plate was incubated upon shaking for 30 min, followed by centrifuging at 500 rpm for 2 min. The fluorescence intensity was then measured on a Synergy Mx microplate reader (BioTek) at Ex 649 nm/Em 670 nm. TMPyP4/DMSO were used as positive and negative controls, respectively. Compounds were considered binders when a change of >10% change in fluorescence intensity was observed compared to that of negative control (DMSO solution). For binding affinity measurement, small molecule solutions were prepared as serial dilutions in DMSO. Then a final test plate was prepared by adding 5 μl small molecule solutions and 10 μl annealed DNA samples into 85 μl buffer solutions (10 mM Tris, pH 7.0, 100 mM KCl, 0.005% Tween 20), resulting in the final small molecule concentrations from 0 to 250 μM (5% final DMSO concentration). The fluorescence intensities were normalized, and the binding affinity was calculated by fitting the curve using one-site total model in GraphPad Prism 8.3.1 software. The fluorescence titration using 3′-Cy5-labeled oligonucleotides was also performed the same way.

Fluorescence displacement assay
In this study, fluorescence displacement assays were carried out by using two classical minor groove binders (Hoechst 33258 (Combi-Blocks) and netropsin (Sigma-Aldrich)). Hoechst 33258, as a fluorophore, was prepared at 5 μM concentration in the buffer (10 mM Tris, pH7.0, 100 mM KCl), and then titrated with different concentrations of unlabeled MYCN G4 oligonucleotides. For displacement, compound to be tested was added into the solution resulting in a final concentration of 5 or 50 μM. Fluorescence signals were obtained with an excitation wavelength of 352 nm and an emission wavelength of 500 nm.

Surface plasmon resonance (SPR) analysis
The SPR binding assays in this study were performed with a BIAcore 3000 (GE Healthcare) instrument. A CM5 SPR biochip was loaded into the system and primed with running buffer (10 mM Tris, pH7.0, 100 mM KCl, 0.005% Tween 20, 5% DMSO). For surface making, the flow rate was set as 5 μl/min. Then both Flow Cell (Fc) 1 and 2 were activated by EDC/NHS (0.4 M/0.1 M) aqueous solution for 15 min, followed with an injection of streptavidin (SA) solution (0.2 mg/ml in 10 mM sodium acetate buffer, pH4.5) for 30 min. After the immobilization amount of SA reached 8000–10 000 RU, the surface was deactivated by flowing 1 M ethanolamine aqueous solution (pH 8.5) for 10 min and regenerated with 10 mM NaOH for 2 min to remove the unbound SA. In the meantime, biotinylated MYCN G4 DNA was prepared at 5 μM in the annealing buffer and heated up to 95°C for 5 min and slowly cooled down. After annealing, a total of 150 μl solution was injected in the Fe 2 of SPR system for 30 min to immobilize DNA onto the chip surface. The small molecule solutions were tested once the baseline was stable.

To detect the binding signal as well as the binding affinity, a higher flowing rate (25 μl/min) was used in both Fc 1 (reference) and Fc 2 (DNA). Each of the compound solution was prepared at 20× designed the concentrations in DMSO, and then diluted into non-DMSO running buffer, resulting in a final concentration of 5% DMSO. Then, a total of 50 μl compound solution was injected in Fc 1–2 flow path for 120 s for association, followed with 200 s running of buffer for dissociation. An injection of 50 μl regeneration buffer (1 M KCl) could be performed between two samples if necessary. The final binding curve was obtained by reference subtraction. To determine the binding affinity (K_D), a series of diluted compound solutions were injected and K_D was calculated by BIAevaluation 4.0 software (GE Healthcare) using Langmuir 1:1 binding model.

Circular dichroism (CD) characterization and melting assay
The folded G-quadruplex structure was characterized by circular dichroism using a J-1500 circular dichroism spectrometer (Jasco). Unlabeled MYCN oligonucleotide (5′-AGGGGGTGGGAGGGGGCATGCAGATGCAGG GGT3′) was prepared in annealing buffer (10 mM Tris, pH 7.0, 100 mM KCl) at 5 μM concentration. As a contrast, DNA samples diluted in Li+ -containing buffer (10 mM Tris, pH 7.0, 100 mM LiCl) and H2O were also used as negative controls.
controls. The annealing procedure was the same as mentioned above. CD spectra were recorded from 320 to 200 nm at 25 °C with a step of 1 nm. Each spectrum was obtained by averaging the signals of three replicate samples.

For CD melting assay, MYCN oligonucleotides were prepared in KCl buffer (10 mM Tris, pH 7.0, 5 mM KCl) at 5 μM concentration. To test the stabilization effect, G4 samples were incubated with and without 20 μM compound (final solution containing 5% DMSO). Then a total of 300 μl solution was added in a cuvette and heated from 20 to 95 °C in a CD spectrometer with an interval of 1°C. To calculate the melting temperature (T_m), the peak of CD spectrum at 263 nm was tracked, and ellipsometry vs temperature was plotted and fitted using a nonlinear sigmoidal dose-response model in GraphPad Prism 8 software (27,43). The shifting of melting temperature was plotted and fitted using a nonlinear sigmoidal dose-response model with a variable slope in GraphPad Prism 8 software. The melting temperature (ΔT_m) was calculated using T_m (compound) – T_m (DMSO).

2-Aminopurine (2-AP) fluorescence titration

Fluorescence titrations based on 2-aminopurine labeling were performed using a protocol reported previously (41). Briefly, MYCN G4 oligonucleotide with a 2-AP substitution at either A11 or A18 position was annealed by heating the sample to 95°C for 5 min and cooling down to rt slowly. The oligonucleotide with folded structure was diluted to 10 μM in Tris buffer (10 mM Tris, pH7.0, 100 mM KCl, 0.005% Tween 20) and prepared in a black 96-well plate (Costar) in triplicate as designed. Small molecules were diluted in DMSO to obtain a series of solutions with concentrations ranging from 0.1 to 5 mM, and then diluted by 20 times added into the plate (5% DMSO final concentration). The final concentration of 2-AP DNA was 1 μM in the well plate. To obtain the background fluorescence, small molecule solutions without DNA were also prepared in the same plate. After incubation for 30 min at rt, the plate was briefly centrifuged and scanned in Synergy Mx microplate reader (BioTek) with Ex310nm/Em365 nm. Fluorescence signals were calculated by averaging the intensities in triplicate wells after subtracting the references. Finally, the binding signals were normalized and the KD value was determined by fitting the curve using a nonlinear sigmoidal dose-response model with a variable slope in GraphPad Prism 8.3.1 software.

Job plot analysis

To determine the stoichiometry of binding, we used continuous variation method by changing the fraction of the compound in the solution (44). Briefly, stock solutions (5% DMSO containing) of 5 μM annealed A11 2-AP labeled MYCN G4 DNA sample and 5 μM compound were prepared, respectively. Two series of solutions were then prepared in 96 black well plates for the experiment: one with a varying fraction of small molecules by mixing the DNA samples with compound stocks, keeping the total concentration constant (5 μM); the other with a varying concentration of DNA samples diluted with buffer, resulting in a track of samples as references. After scanning, the difference between two series of solutions in fluorescence intensity is calculated to generate a Job plot. Then, linear regression analysis was performed using GraphPad Prism 8.3.1 software.

Microscale thermophoresis (MST)

MST experiments were carried out on a Monolith NT.115 system (NanoTemper Technologies). The 3′-Cy5 MYCN G4 DNA solutions were prepared in 10 mM Tris (pH 7.0), 100 mM KCl, 0.005% T20, annealed (according to the above-mentioned method), and diluted to 100 nM (2×). The DMSO solutions of small molecules were prepared in the buffer with a series 1:1 dilution, resulting in 2× of designed concentrations (10% DMSO). Then, DNA samples were 1:1 (v/v) mixed with corresponding small molecule solutions, resulting in 50 nM DNA and 5% DMSO. Finally, MST signals were detected in triplicate capillaries, and the dissociation constant was determined by fitting the curve using a single-site model (MO.Affinity Analysis v2.3).

DMS-footprinting

Footprinting of 5′-Cy5-labeled DNA was performed as described in the literature (45,46). MYCN G4 DNA samples were prepared in appropriate buffer (10 mM Tris, pH 7.0, 5 mM KCl, 0.005% T20) with 5 μM concentration, followed by annealing using above-mentioned method. Then, DNA solutions were diluted to 50 nM in corresponding buffers (2 ml final volume in each tube) and compound stock solutions (in DMSO) were added resulting in solutions with designed concentrations (resulting in 5% DMSO). After incubation for 30 min, the folded DNA samples were treated with 0.5% DMS for 10 min at rt and the reaction was stopped by adding 200 μl stop buffer (2.5 M NH4OAc, 0.1 M β-mercaptoethanol, 1 mg/ml calf thymus DNA). After phenol/chloroform/isoamyl alcohol extraction and ethanol precipitation, the DNAs were dissolved in 50 μl of nuclease free water. An equal volume of 10% piperidine was added to each tube, and the mixed solutions were heated at 90°C for 30 min, followed by quick chilling on ice. The DNAs were again subjected to phenol/chloroform/isoamyl alcohol extraction and ethanol precipitation. The precipitated DNAs were dissolved in 10 μl of nuclease free water, denatured at 95°C for 5 min, and resolved on a 17% denaturing polyacrylamide gel. After gel running, Cy5-labeled DNA fragments were visualized on a Typhoon Imager (Amersham) and digitized using ImageJ software.

Cell lines and cell culture

The neuroblastoma cell line NBEB (single copy MYCN) cells were used to evaluate the efficacy of the MYCN/MYCNOS G-quadruplex binding molecules. NBEB came from stocks in Pediatric Oncology Branch, National Cancer Institute, National Institutes of Health, Bethesda, MD and have been STS verified. Cells were cultured in RPMI-1640, supplemented with 10% FBS (Atlanta Biologicals, Atlanta), 2 mM glutamine (Life Technologies, New York), and antibiotics (penicillin 100 μg/ml, streptomycin 100 μg/ml; Life Technologies, New York) at 37°C in a 5% CO2 incubator.
To examine the effects of G4-binding compounds on NBEB cells, live-cell imaging was performed using an Incucyte interface (described below). NBEB (5K/well), cells were seeded into 96-well plates in triplicate, treated with compound MY-8 at various concentrations (0–45 μM), and cultured until control wells reached confluence. For RT-qPCR and western blotting, NBEB (300K) cells were treated with MY-8 in six-well dish with different time and doses. To validate cell viability, a Trypan Blue exclusion test was performed following treatment at 48 and 96 h post treatment and the proportion of live versus dead cells was calculated using a hemocytometer.

Cell survival
Following live cell image analysis, the CellTiter colorimetric cell viability assay (MTS) (Promega, USA) or CellTiter-Glo luminescent cell viability assay (Promega, USA) was performed to examine cell survival. Absorbance was measured with a microplate reader (BIO-RAD, America). To calculate the proportion of viable cells, MY-8 treated cells were divided by the control samples from each group.

Incucyte live cell imaging system
Live cell imaging was performed using the Incucyte Zoom Live-cell Imaging System from Essen Bioscience (Ann Arbor, MI, USA). Incucyte measured cell confluence from predefined processing definitions for NBEB cells. The Incucyte Zoom Live-cell Imaging System scanned three phase contrast images per well every 6 h for the duration of treatment (0–120 h).

Quantitative RT-PCR
Total RNA was extracted from NBEB cells according to manufacturer’s instructions (RNeasy Plus Mini Extraction kit (#74134) (Qiagen, CA)). RNA (1 μg) was reverse transcribed into cDNA using the cDNA synthesis High-Capacity RNA-to-cDNA™ Kit. Quantitative PCR was performed using Fast SYBR™ Green Master Mix according to the manufacturer’s protocol. Beta-actin was used as a housekeeping gene. Relative expression was calculated using the 2^−ΔΔCt method. The qPCR primer sequences can be found in supplementary materials.

Protein assay
Forty-eight hours post-MY-8 treatment in six-well plates, single copy MYCN NBEB cells were lysed with Radioimmunoprecipitation assay buffer (Beyotime, China), pelleted, and chilled at −80°C. Total protein concentrations were determined using a Bradford Assay using the Bradford reagent (Beyotime, China). Each protein sample (10 μg) was loaded onto a 12% gel (Bio-Rad, USA), electrophoresed at 90 V for 90 min and transferred to a nitrocellulose membrane (Immobilon-P, Millipore, Bedford, MA, USA) using a Bio Rad Trans-Blot Turbo Transfer System (Bio-Rad, USA). Nitrocellulose membranes were incubated in 5% milk in Tris-buffered saline supplemented with 0.5% Tween 20 for 1 h at room temperature and incubated with anti-N-Myc: sc-53993 (1:4000) and anti-GAPDH: sc-47724 (1:2000) antibodies overnight at 4°C. Following incubation, the membranes were washed with Tris-buffered saline supplemented with 0.5% Tween 20 and probed with a goat anti-mouse (1:1000) antibody conjugated to peroxidase and incubated for 1 h at room temperature. Chemiluminescent signal was detected using Clarity™ Western ECL Substrate (Bio-Rad, USA) or Supersignal™ West Femto Maximum Sensitivity Substrate (ThermoFisher Scientific, USA) with a ChemiDoc gel imaging system (Bio-Rad, USA).

RESULTS
Analysis of G-quadruplexes in the MYCN gene
As shown in Figure 1A, the MYCN G4 discussed in this study is located in the negative strand and near the transcription start site (TSS) of MYCNOS. However, unlike simple G4s, the MYCN G4 contains a hairpin flanking the tetrads (37). This class of hairpin-containing G4 structures have been recently characterized by both genomic data mining and experimental approaches (31,32,40), and represent intriguing targets for small molecules due to their unique structures (47). Although the MYCN G4 has been structurally characterized in vitro, questions frequently arise about the existence of G4s in biological contexts. In order to confirm that the MYCN gene contains sequences that fold into G4s, we analyzed the genome-wide G4-seq data reported by the Balasubramanian group (38). This approach consisted of two rounds of sequencing before and after stabilizing the G4s. A higher mismatch rate between the experiments represented a higher chance to form G4 structure in the corresponding region. Within these data, we found that the G4 sequence of interest to this work had a much higher mismatch rate in both K+ and pyridostatin (PDS)-stabilized data sets, confirming the presence of the G4 in MYCN gene (Figure 1B, Supplementary Figure S1). Further, by calculating the G-scores within this sequence using QGRS mapper (48), we found that our target of interest showed a promising G-score of 33 (Supplementary Figure S2).

Small molecule microarray screening and hit identification
To identify a small molecule binder of the MYCN G4, we used a small molecule microarray (SMM) screening approach. In this study, glass slides modified with epoxy groups were prepared according to an analogous protocol reported previously (49). Then, compounds were printed by a robotic arrayer (Arrayjet™) and covalently immobilized on the glass surfaces. Then the slides were incubated with 50 nM MYCN G4 DNA with Cy5 labeling at 5′ and buffer in parallel. After washing and drying, the slides were imaged by a microarray scanner. For each compound, a composite Z-score was calculated, and hits were identified (41). Of note, the immobilization of compounds through a specific nucleophilic group may result in false negatives due to this site being blocked. However, this inherent bias does not prevent the discovery of all hit molecules (as described below). Compounds that lacked promiscuous binding to other nucleic acids in other SMM screens were prioritized. In total, 14 hits were identified for follow up studies (numbered from 1 to 14, see Supplementary Figure S3).
Figure 1. High-throughput screening using small molecule microarray (SMM) toward MYCN G4 DNA and hit validation by fluorescence intensity assay (FIA) and surface plasmon resonance (SPR). (A) G-quadruplex forming sequence identification in MYCN gene; (B) MYCN G4 forming probability obtained by genome-wide G4-seq analysis; the highlighted region represents the MYCN G4 target in this research; (C) four hits identified from SMM screening and single-concentration re-screening by FIA and SPR; (D) binding affinities ($K_d$) measured by FIA and SPR titrations.
To validate the 14 hit compounds as genuine binders, we performed single-concentration binding assays by fluorescence intensity assay (FIA) and surface plasmon resonance (SPR). Compounds were evaluated at 100 μM concentration in each assay. A positive binding was defined with a criteria of >10% quenching in FIA study, while signal-to-noise ratio (SNR) > 3 was used as the cut-off in SPR experiments. From these assays we identified four compounds (1, 2, 4 and 5) that showed positive responses in both binding assays (Figure 1C, Supplementary Figure S4). Besides, all the four compounds also showed promising selectivity on SMM (Supplementary Figure S5). We next measured equilibrium dissociation binding constant (K_D) values by FIA and SPR (Figure 1D, Supplementary Figure S6). By titrating MYCN G4 DNA sample with the gradient concentrations of compounds, we found that compound 4 showed the best binding among the four candidates (K_D = 3.5 ± 1.6 μM by FIA, and K_D = 3.6 ± 2.4 μM by SPR). Compound 1 and 2 shared a similar structure, but 1 had a better binding affinity (20.8 ± 2.2 μM by FIA, and 11.7 ± 2.6 μM by SPR). Thus, compounds 1 and 4 were selected for further study.

We next evaluated the effect of 1 and 4 on MYCN G4 thermal stability by CD. First, G4 formation was confirmed by CD spectroscopy. Samples of MYCN G4 DNA were prepared in 3 different buffer conditions (10 mM Tris with 100 mM KCl, 10 mM Tris with 100 mM LiCl, and H2O). As shown in Figure 2A, the MYCN G4 DNA exhibited a positive peak at 263 nm and a negative peak at 240 nm, characteristic of a properly folded parallel G-quadruplex structure in KCl buffer. In contrast, spectra of DNA samples in LiCl buffer and H2O (which do not promote proper folding of the G4) showed a much weaker peak at 263 nm while the peak at 240 nm was negligible.

Next, we demonstrated a melting assay on CD instrument using MYCN G4 DNA incubated without/with small molecule binders. The 263 nm peak in the CD spectrum decreased gradually during heating (Figure 2B), and T_m was determined by fitting the normalized temperature-dependent intensities. The MYCN G4 was first melted in 10 mM Tris buffer (pH 7.0) with 100 mM KCl, and T_m was determined as 78.0 ± 1.9°C (Supplementary Figure S7), which was consistent with a previously reported value (37). To evaluate the impact of molecules on thermal unfolding, we performed assays using reduced KCl concentration (50), after optimizing the buffer/incubation conditions and confirming the folding of G4 was not affected by lower KCl (Supplementary Figure S8-10). Similar effects on melting by 4 were observed in both phosphate buffer and Tris buffer (Supplementary Figure S11). In 5 mM KCl buffer, the melting curve of MYCN G4 was better fitted, with a T_m of 54.6 ± 0.4°C. The T_m was not affected by the addition of 5% DMSO into the buffer. 20 μM 1 or 4 was added into MYCN G4 solution and incubated for 15 min before unfolding. As a result, both of the compounds induced significant shift in melting curves, increasing T_m by 4.7 ± 0.7°C (for 1) and 3.7 ± 0.5°C (for 4), respectively (Figure 2C). Neither of the compounds (20 μM) showed a measurable effect on dsDNA melting (Supplementary Figure S12, Figure 2D), indicating that those hit compounds were not B-DNA binders.

The binding selectivity of two compounds was also evaluated. In a separate experiment, fluorescence-based assays were performed by titrating the compounds into G4 solutions with the presence/absence of excess tRNA. As with dsDNA, excess tRNA had no effect on the binding affinity of 1 or 4, again suggesting a specific mode of interaction (Supplementary Figure S13). In addition, another six 5′-Cy5-labeled DNA/RNA G4s from cancer-relevant genes (Supplementary Table S1) were also introduced in selectivity profiling. By titrating the two hit compounds, we observed that compound 1 showed some binding to all the G4s with binding affinities ranging from 25.2 to 127.9 μM, but weaker than that of MYCN G4 (Supplementary Figure S14). However, four showed no significant quenching behavior and K_D values could not be measured for any of the tested G4s besides MYCN (Supplementary Figure S15 and S16). This result indicated that 4 bound to MYCN G4 via its unique structure. Hence, these two hit compounds were identified as MYCN G4 stabilizers and were used for further study.

Identification of the binding mode of 4

By analyzing SPR sensograms, we observed significantly different binding levels (R_max) between compounds 1 and 4, which suggested that the binding modes of the two compounds might be different. To investigate, we performed a competitive assay on SPR by step-wise injection of the compounds in a designated order or injecting a mixture (51,52). By injecting 100 μM of 1 followed by 100 μM of 4, we observed that the binding level of 4 was not significantly changed, compared with that of individual injection (Figure 3A). Even at a higher concentration (500 μM) of 1, the level of 4 binding was maintained (Supplementary Figure S17), suggesting that 1 does not compete with 4. Next, TMPyP4, a classical stacking G4-binder (53), was also tested together with 4. The addition of 4 (20 μM) into 500 nM TMPyP4 solution resulted in an increase of SPR binding signals (Supplementary Figure S18A), showing a non-competitive manner. Similar results were observed when using other G4-binders such as PhenDC3 and PDS, suggesting that 4 bound to MYCN G4 in a mode distinct from classical G4 stacking (Supplementary Figure S18B and C). In a related assay, no binding of 4 to dsDNA was observed using SPR up to a concentration of 50 μM (Supplementary Figure S19). In addition, a CD melting assay was also performed using either 4 or a mixture of two hits (1 and 4). A mixture of solution containing 20 μM 1 and 20 μM 4 increased T_m of MYCN G4 by 9.7 ± 0.7°C, which was significantly higher than that of 20 μM 4 only (Figure 3B and C, Supplementary Figure S20), suggesting non-overlapping modes of interaction. In contrast, the ΔT_m of MYCN G4 with 40 μM 4 addition was not significantly increased, presumably due to the saturation of binding sites.

To evaluate the binding stoichiometry of 4 with MYCN G4 we performed a Job plot analysis. Briefly, by varying the component fraction during fluorescent titration, the changes in fluorescent intensity were plotted and fitted, resulting in a maximum at 0.47, indicating a 1:1 binding stoichiometry (Figure 3D). The binding between 4 and MYCN G4 was studied by microscale thermophoresis (MST). 3′-
Figure 2. MYCN G4 structure stabilization by small molecule and binding selectivity study. (A) CD spectra of MYCN G4 DNA in different buffers (10 mM Tris containing 100 mM KCl or 100 mM LiCl) and in water. (B) CD curve recording of 5 μM MYCN G4 during melting in low KCl buffer (10 mM Tris, pH 7.0, 5 mM KCl, 5% DMSO). (C) MYCN G4 melting with/without 1 or 4. (D) CD melting test (based on the peak at 250 nm) of dsDNA incubated with 1 or 4.

Cy5 labeled DNA sample was used since there was no quenching observed (Supplementary Figure S21). The MST curve was well fitted with 1:1 binding model, and the binding affinity was calculated as 8.1 ± 1.7 μM which agreed with the results of FIA and SPR experiments (Figure 3E and F).

**Binding site identification using FIA and DMS-footprinting**

To further understand the binding details between 4 and MYCN G4, a series of FIA studies were designed and performed. Given that the hit compound bound to G4 with 1:1 stoichiometry, we hypothesized that the unique binding behavior was related to the hairpin structure. We first investigated the binding region by introducing environment-sensitive 2-aminopurine (2-AP) fluorophores at distinct positions (A11, A18, and A24) of the G4 (Figure 4A). After incubating 100 μM 4 with labeled DNA samples, fluorescence intensities were measured and compared to DMSO controls. Fluorescence of the A11 2-AP DNA sample was quenched by almost 100%, while the quenching percentages were 80% and 70% for A18 and A24 2-AP DNAs, respectively (Figure 4B, Supplementary S22), while K_D values were mostly unchanged. Based on the previous observation of different quenching behaviors between 5′ and 3′-Cy5 labeled oligos (15% and 0% quenching, Figure 4C), we hypothesized that 4 was binding to the region near the hairpin. Next, we explored the necessity of hairpin in this binding event by designing two DNA constructs including mutated long loop (incapable of base-pairing) and truncated sequence (lacking a hairpin) (see Supplementary Table S2). By titrating 4 into the two DNAs, the binding affinity was decreased considerably to >100 μM (at least ~30-fold weaker than that of the wild type MYCN G4). In contrast, 1 was also tested against these G4s, and affinities were comparable to wild type (WT) constructs (Figure 4D). These results indicated that the hairpin structure in MYCN G4 was important to maintain the binding with 4, but not with 1.

In addition, to explore whether the compound bound to minor groove of the hairpin, 4 was tested together with other two classical minor groove binders (Hoechst 33258 and netropsin) by fluorescence displacement assay (54). Since Hoechst 33258 has been reported as a fluorophore...
Figure 3. Identification of the binding mode of 4. (A) SPR curve of competitive binding study between 1 and 4. (B) CD melting curve of G4 stabilized by individual compound (4) or compound mixture (1 and 4) in 5% DMSO containing buffer. (C) Melting temperature of MYCN G4 stabilized by corresponding compound combination. (D) Stoichiometry of 4 binding with MYCN G4 determined by Job Plot analysis. (E and F) MST study on 4 binding with 3′-Cy5-MYCN G4.

and widely used as DNA staining reagent, it was first incubated and titrated with different concentrations of unlabeled MYCN G4 DNAs. By utilizing excitation wavelength of 352 nm and an emission wavelength of 500 nm, the fluorescence intensity was enhanced with the increasing DNA concentration as expected (Figure 4E). Next, either 4 or netropsin was added as a competitor into the solutions during DNA titration. Netropsin, as a well-known minor groove binder, showed a significant displacement of Hoechst 33258 as measured by corresponding decrease in fluorescence intensity. However, for 4, there was no measurable change observed at 5 μM, while partial fluorescence decrease was observed at 50 μM. Thus, although the hairpin is required for 4 binding, 4 does not appear to be a classical minor groove binder.

In order to better understand the mode of binding of compounds to the MYCN G4, a dimethyl sulfate (DMS) footprinting assay was developed. This assay exploits the ability of DMS to methylate the free N7 within guanines (Gs), which can subsequently be cleaved using piperidine (55,56). In contrast, Hoogsteen bonded Gs in the quadruplex structure remain protected from modification. Further, changes in modification upon incubation with small molecules can be observed near binding sites. 5′-Cy5-labeled MYCN G4 was annealed in the presence of 5 mM KCl, and subsequently compounds were added at concentrations according to observed KDs. As shown in Figure 4F, bands corresponding to G10-G12 were dim due to protection indicating that those Gs were incorporated within tetrads in the quadruplex. However, G13-G16 remained unprotected showing much brighter bands, due to hairpin formation via Watson-Crick base pairing. In the presence of compound 4, concentration-dependent protection of various Gs was observed, indicating that the compound stabilized the G4 structure. Upon the addition of 4, G8, G9, G16, G17 were protected considerably, while G13-G15 were slightly protected. This result suggested that 4 was binding to the junction between quadruplex and hairpin, in agreement with multiple results described above (Figure 4B).

Structure–activity relationship (SAR) study of 4

To obtain a better binder for MYCN G4 target, we performed a structure–activity relationship (SAR) study by using a series of analogs of 4 (see Table 1). Based on the precursors (15 and 16), a focused library containing 4 and 14 derivatives (compound 17–30) was either purchased or synthesized (Supplementary Methods, Supplementary Figures S29–S44). Each compound was evaluated by SPR and 2-AP titration as described above. In our preliminary study, the pyrrolidine group in 4 played a role not only by maintaining the binding with G4, but also helping with the solubility. Meanwhile, the heterocyclic core of 4 (shown in Table 1) acted as a key part for MYCN G4 recognition. Thus, most analogs contained altered side chain R groups.

Then, a systematic binding assay between 14 analogs and MYCN G4 was performed by SPR and 2-AP (A11) fluorescence titration (Supplementary Figure S23 and S24). Observed binding affinities ranged from 1.4 to 23.5 μM (SPR), and 0.2 to 13.2 μM (2-AP titration). Most of the analogs exhibited good binding behavior toward the MYCN G4, except 19 and 24, both of which had poor solubility in suitable buffers. In SPR experiments, 4 analogs (17, 20, 23 and 25) were identified as stronger binders than the parent compound, among which 20 and 23 displayed the tightest bind-
Figure 4. Binding site study using fluorescence quenching-based methods and DMS-foot printing. (A) Schematic prediction of MYCN G4 folded structure. (B) Quenching study by adding 100 μM 4 into solutions of 5′/3′-Cy5 labeled MYCN G4s. (C) Quenching study by adding 100 μM 4 into solutions of 2-AP MYCN G4s labeled at A11, A18 and A24 positions, respectively. (D) Fluorescence titration using Wild type/mutated/truncated MYCN G4 DNA samples with 5′-Cy5 labeling. Fluorescence intensities during titration were recorded and Kd values were calculated by fitting the curves. (E) Fluorescence replacement assay using minor groove binders (Hoechst 33258 and netropsin). (F) DMS foot-printing result of MYCN G4 DNA incubated with different concentrations of compound 4. G-tracts involved in quadruplex tetrads are underlined. Protected Gs affected by compound 4 were labeled with red dots (highly affected) or red circles (slightly affected).

Effect on MYCN/MYCNOS expression after MY-8 treatment

In order to evaluate the effects of MY-8 on MYCN gene expression, NBEB cells were cultured and treated with different concentrations of MY-8. We evaluated the effects of MY-8 on cell viability using an Incucyte live cell imaging system. Based on a 4-day measurement of cell confluence after single treatment of MY-8, a significant inhibition of NBEB cell growth was observed (Figure 5A–C). With the increase of compound concentration up to 45 μM, an IC50 of 19 μM was confirmed (Supplementary Figure S27). In contrast, in HEK293T cells MY-8 had an IC50 value of 28 μM. Treatment of cells with 22.5 μM MY-8 resulted in a more than 90% reduction of viable NBEB cells but only 50% reduction of viable cells in HEK293T. Next, mRNA levels of genes including MYCN, as well as two MYCNOS transcripts (MYCNOS001 and MYCNOS002), were measured at different time points (24, 36 and 48 h) after MY-8 treat-
Table 1. SAR study of lead compound (4)

| Name | R          | SPR K<sub>D</sub> (µM) | A11 2AP K<sub>D</sub> (µM) | Name | R          | SPR K<sub>D</sub> (µM) | A11 2AP K<sub>D</sub> (µM) |
|------|------------|------------------------|-----------------------------|------|------------|------------------------|-----------------------------|
| 4    |            | 3.6 ± 2.4              | 3.2 ± 0.6                   | 23   |            | 1.5 ± 0.3              | 0.6 ± 0.2                   |
| 17   |            | 2.6 ± 1.9              | 0.20 ± 0.04                 | 24   |            | N/A                    | N/A                         |
| 18   |            | 5.9 ± 2.7              | 1.2 ± 0.6                   | 25   |            | 2.2 ± 0.6              | 0.7 ± 0.2                   |
| 19   |            | N/A                    | N/A                         | 26   |            | 9.2 ± 1.9              | 1.2 ± 0.2                   |
| 20   |            | 1.4 ± 0.2              | 1.0 ± 0.1                   | 27   |            | > 100                  | N/A                         |
| 21   |            | 4.9 ± 2.6              | 1.0 ± 0.2                   | 28   |            | 8.2 ± 2.4              | 3.2 ± 0.7                   |
| 22   |            | 23.5 ± 4.7             | 13.2 ± 4.6                  | 29   |            | 3.6 ± 1.0              | 0.9 ± 0.2                   |

DISCUSSION

In this study, we demonstrate a strategy to target a unique hairpin-containing G4 DNA structure in the MYCN gene with small molecules. By using high-throughput SMM screening, several distinct classes of druglike small molecule binders were successfully identified. Among these hits, one of the candidates showed promising binding affinity as well as good selectivity. These compounds generally have affinities that are weaker than other known quadruplex binders, but are in line with affinity values for hits frequently seen from unbiased high throughput screening efforts of nucleic acids or proteins. In-depth biophysical analyses using multiple orthogonal techniques revealed that the lead compound (4) bound to the MYCN G4 in an atypical manner by interacting with the unique fold formed by the G-tetrad and hairpin. Additionally, a preliminary SAR study resulted in the discovery of MY-8, an improved analog with a submicromolar binding affinity as measured by 2-AP titration. Biological evaluation of MY-8 revealed that it decreased levels of MYCN as well as MYCNOS, downregulating levels of both gene products at the RNA and protein level. More extensive biological evaluation will provide further insight into the role this G4 plays in MYCN and MYCNOS expression but will necessitate the development of improved small molecules with superior affinity and potency.

This work demonstrates that compounds with multiple distinct modes of recognition can be identified using a single high throughput screening approach and can be stratified as hits using rigorous biophysical analysis. Based on the synergistic binding behavior of different classes of hit compounds, complex structures such as this G4 could represent an opportunity to develop next-generation binders with improved binding affinity as well as selectivity by designing a bivalent compound employing a proper linker, a strategy that has also been proposed by other researchers (57).

Considerable effort has been invested in developing G4 ligands, though highly selective ligands have proven mostly elusive. Targeting higher order structure within DNA G4s could provide a new avenue for developing selective inhibitors that control the expression of undruggable oncogenes. Indeed, higher complexity RNA G4s have been successful frameworks for highly selective small molecule binding. While these examples are mostly evolved fluorescent RNA aptamers, they are an elegant demonstration that complex structures containing G4s can provide unique pockets for highly selective small molecule recognition. Further investigation into small molecules that recognize complex DNA G4s is therefore warranted.

Given the potential for G4s containing hairpins as targets, we further explored the existence of other G4s with similar structures in human genome. Using data from genome-wide G4-seq reported by Balasubramanian (38), we identified ~2 million unique loops (> 7 nt) between G-tracts of G4s embedded in 407,491 regions of observed quadruplexes (OQs). After a hairpin-folding test using UNAfold (58) (hybrid-ss-min package at 37°C), we found 33,912 of OQs contained a hairpin structure (~8.3% of the
total OQs, see Supplementary Figure S28), in agreement with a previous computational prediction (40). Among the hairpin-G4 regions, 58% were associated with protein coding genes while 13% were associated with non-coding RNAs. Several of the hairpin-containing G4s were located in genes of cancer-relevant proteins (such as FOXA3, KRAS, MYCL and BRD4), (see Supplementary Table S3 for representative examples). This work indicates that higher complexity G4s that contain embedded hairpins are prevalent throughout the genome and are often associated with oncogenes or important lncRNAs.

In the broader sense, this study indicates that small molecule recognition of G4 elements may impact lncRNA expression. While lncRNAs represent important drivers in multiple cancer types, there are few examples of small molecules capable of controlling their expression or function. Our informatic analysis of publicly available datasets indicates that hairpin-G4 complex DNA structures are frequently found throughout the human genome. The efforts here indicate that the unique folds formed by these complex structures could be valuable functional targets for small molecules as well, facilitating the targeting of disease-
relevant genes including protein coding as well as non-coding gene products.

**SUPPLEMENTARY DATA**

Supplementary Data are available at NAR Online.

**ACKNOWLEDGEMENTS**

The authors thank Philip Homan for helpful discussions regarding bioinformatics. This work utilized the computational resources of the NIH HPC Biowulf cluster (http://hpc.nih.gov). We also thank members from biophysics resource for helpful comments and suggestions including Dr Sergey G. Tarasov and Marzena Dyba for help with CD and MST experiments. We thank Dr James A. Kelley from NCI/CBL for help with HRMS experiments.

**FUNDING**

National Institutes of Health, National Cancer Institute, Center for Cancer Research [Z01 BC011585 07 (PI) to J.S.S., Z01 BC 010788 (PI), to C.J.T.]. Funding for open access charge: Publication charges will be paid from the lab budget of the corresponding author, John Schneekloth, Jr. Conflict of interest statement. None declared.

**REFERENCES**

1. Dang, C.V. (2012) MYC on the path to cancer. *Cell*, **149**, 22–35.
2. Murre, C., McCaw, P.S., Vaessen, H., Caudy, M., Jan, L.Y., Jan, Y.N., Cabrera, C.V., Buskin, J.N., Hauschka, S.D., Lassar, A.B. et al. (1989) Interactions between heterologous helix-loop-helix proteins generate complexes that bind specifically to a common DNA sequence. *Cell*, **58**, 537–544.
3. Rickman, D.S., Schulte, J.H. and Eilers, M. (2018) The expanding world of N-MYC-driven tumors. *Cancer Discov.*, **8**, 150–163.
4. Grunblatt, E., Wu, N., Zhang, H., Liu, X., Norton, J.P., Ohol, Y., Leger, P., Hiatt, J.B., Eastwood, E.C., Thomas, R. et al. (2020) MYCN drives chemoresistance in small cell lung cancer while USP7 inhibition can restore chemosensitivity. *Genes Dev.*, **34**, 1210–1226.
5. Vaidie, N., Saamyan, S., Lenox, A., Ackley, A., Clemson, M., Burdach, J., Hart, J., Vogt, P.K. and Morris, K.V. (2015) MYC/NS functions as an antisense RNA regulating MYCN. *RNA Biol.*, **12**, 893–899.
6. Krystal, G.W., Armstrong, B.C. and Battey, J.F. (1990) N-myc mRNA forms an RNA-RNA duplex with endogenous antisense transcripts. *Mol. Cell. Biol.*, **10**, 4180–4191.
7. Armstrong, B.C. and Krystal, G.W. (1992) Isolation and characterization of complementary DNA for N-cym, a gene encoded by the DNA strand opposite to N-myc. *Cell Growth Differ.*, **3**, 385–390.
8. Suenaga, Y., Islam, S.M., Alaguz, J., Kaneko, Y., Kato, M., Tanaka, Y., Kawanami, H., Hossain, S., Matsumoto, D., Yamamoto, M. et al. (2014) NCYM, a Cis-antisense gene of MYCN, encodes a de novo evolved protein that inhibits GSK3beta resulting in the stabilization of MYCN in human neuroblastomas. *PLoS Genet.*, **10**, e1003996.
9. Suenaga, Y., Nakatani, K. and Nakagawa, A. (2020) De novo evolved gene product NCYM in the pathogenesis and clinical outcome of human neuroblastomas and other cancers. *Jpn. J. Oncol.*, **50**, 839–846.
10. Kaneko, Y., Suenaga, Y., Islam, S.M., Matsumoto, D., Nakamura, Y., Ohira, M., Yokoi, S. and Nakagawa, A. (2015) Functional interplay between MYCN, NCYM, and OTC4 promotes aggressiveness of human neuroblastomas. *Cancer Sci.*, **106**, 840–847.
11. Gustafsson, W.C., Meyerowitz, J.G., Nekritz, E.A., Chen, J., Benes, C., Charron, E., Simonds, E.E., Seeger, R., Matthay, K.K., Hertz, N.T. et al. (2014) Drugging MYCN through an allosteric transition in Aurora kinase A. *Cancer Cell*, **26**, 414–427.
12. Warren, C.L., Kratochvil, N.C., Hauschild, K.E., Foister, S., Brezinski, M.L., Dervan, P.B., Phillips, G.N. Jr and Ansari, A.Z. (2006) Defining the sequence-recognition profile of DNA-binding molecules. *PNAS*, **103**, 867–872.
13. Chenoweth, D.M., Harki, D.A., Phillips, J.W., Dose, C. and Dervan, P.B. (2009) Cyclic pyrrole-imidazole polyamides targeted to the androgen response element. *J. Am. Chem. Soc.*, **131**, 7182–7188.
14. Francisco, A.P. and Paula, A. (2017) Oncogene expression modulation in cancer cell lines by DNA G-quadruplex-interactive small molecules. *Curr. Med. Chem.*, **24**, 4873–4904.
15. Neidle, S. (2016) Quadruplex nucleic acids as novel therapeutic targets. *J. Med. Chem.*, **59**, 5987–6011.
16. Neidle, S. (2017) Quadruplex nucleic acids as targets for anticancer therapeutics. *Nat. Rev. Chem.*, **1**, 0041.
17. Bochman, M.L., Paeuschke, K. and Zatkin, V.A. (2012) DNA secondary structures: stability and function of G-quadruplex structures. *Nat. Rev. Genet.*, **13**, 770–780.
18. Bianco, M. and Ferre-D’Amare, A. (2021) The emerging structural complexity of G-quadruplex RNAs. *RNA*, **27**, 390–402.
19. Di Antonio, M., Ponjavic, A., Radzevicius, A., Ransinghe, R.T., Catalano, M., Zhang, X., Shen, J., Needham, L.M., Lee, S.F., Klenerman, D. et al. (2020) Single-molecule visualization of DNA G-quadruplex formation in live cells. *Nat. Chem.*, **12**, 832–837.
20. 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.
21. Yang, S.Y., Lejault, P., Chevrier, S., Boidot, R., Robertson, A.G., Wong, J.M.Y. and Monchaud, D. (2018) Transcriptome-wide identification of transient RNA G-quadruplexes in human cells. *Nat. Commun.*, **9**, 4730.
22. Lejault, P., Mitteaux, J., Sperti, F.R. and Monchaud, D. (2021) How to untie G-quadruplex knots and why? *Cell Biol. Chem.*, **28**, 436–455.
23. Tian, T., Chen, Y.-Q., Wang, S.-R. and Zhou, X. (2018) G-Quadruplex: a regulator of gene expression and its chemical targeting. *Chem., 4*, 1314–1344.
24. Raguero, F., Chowdhury, S., Minard, A. and Di Antonio, M. (2020) Chemical-biology approaches to probe DNA and RNA G-quadruplex structures in the genome. *Chem. Commun.*, **56**, 1317–1324.
25. Carvalho, J., Mergny, J.L., Salgado, G.F., Queiroz, J.A. and Cruz, C. (2020) G-Quadruplex, friend or foe: the role of the G-quartet in anticancer strategies. *Trends Mol. Med.*, **26**, 848–861.
26. Xu, H., Di Antonio, M., McKinsey, S., Mathew, V., Ho, B., O’Neil, N.J., Santos, N.D., Silvester, J., Wei, V., Garcia, J. et al. (2017) CX-5461 is a DNA G-quadruplex stabilizer with selective lethality in BRCA1/2 deficient tumours. *Nat. Chem.*, **8**, 14432.
27. Calabrese, D.R., Chen, X., Leon, E.C., Gaikwad, S.M., Phyó, Z., Hewitt, W.M., Alden, S., Hilimire, T.A., He, F., Michalowski, A.M. et al. (2017) Chemical and structural studies provide a mechanistic basis for recognition of the MYC G-quadruplex. *Nat. Commun.*, **9**, 4229.
28. Ruggiero, E. and Richter, S.N. (2018) G-quadruplexes and G-quadruplex ligands: targets and tools in antiviral therapy. *Nucleic Acids Res.*, **46**, 3270–3283.
29. Pirota, V., Nadai, M., Doria, F. and Richter, S.N. (2019) Naphthalene imidazoles as multimodal G-quadruplex-selective ligands. *Molecules*, **24**, 426.
30. Varshney, D., Spiegel, J., Zyner, K., Tannahill, D. and Balasubramanian, S. (2020) The regulation and functions of DNA and RNA G-quadruplexes. *Nat. Rev. Mol. Cell Biol.*, **21**, 459–474.
31. Onel, B., Carver, M., Wu, G., Timonina, D., Kalarr, S., Larriva, M. and Yang, D. (2016) A new G-quadruplex with hairpin loop immediately upstream of the human BCL2 P1 promoter modulates transcription. *J. Am. Chem. Soc.*, **138**, 2563–2570.
32. Ngoc Nguyen, T.Q., Lim, K.W. and Phan, A.T. (2020) Duplex formation in a G-quadruplex bulge. *Nucleic Acids Res.*, **48**, 10567–10575.
33. Tan, D.J.Y., Winndery, F.R., Lim, K.W. and Phan, A.T. (2020) Coexistence of two quadruplex-duplex hybrids in the PIM1 gene. *Nucleic Acids Res.*, **48**, 11162–11171.
34. Trajkovski, M., da Silva, M.W. and Plavec, J. (2012) Unique structural features of interconverting monomeric and dimeric G-quadruplexes adopted by a sequence from the intron of the N-myc gene. *J. Am. Chem. Soc.*, **134**, 4132–4141.
