Individual G-quadruplex targeting with ligand-functionalized CRISPR-Cas9 uncovers transcriptional-dependent functional responses

Sabrina Pia Nuccio[1,2], Enrico Cadoni[1,3], Roxani Nikoloudaki[1], Silvia Galli[1], An-Jie Ler[1], Thomas E. Maher [1,2], Ella Fan[1], Ling Sum Liu[1], Christopher R. Fullenkamp[4], John S. Schneekloth, Jr[4], Marco Di Antonio[1,2,5] *

1 Imperial College London, Department of Chemistry, Molecular Sciences Research Hub, 82 Wood Lane, London W12 0BZ, UK
2 Imperial College London, Institute of Chemical Biology, Molecular Sciences Research Hub, 82 Wood Lane, London W12 0BZ, UK
3 Department of Organic and Macromolecular Chemistry - Organic and Biomimetic Chemistry Research Group (OBCR), Faculty of Sciences, Campus Sterre, Krijgslaan 281, Building S4, 9000 Gent, Belgium
4 Chemical Biology Laboratory, National Cancer Institute, Frederick MD 21702, USA
5 The Francis Crick Institute, 1 Midland Road, London NW1 1AT, UK

*To whom correspondence should be addressed. Email: m.di-antonio@imperial.ac.uk (MDA); Tel: +44 (0)20 7594 5866

Abstract

The development of selective ligands to target DNA G-quadruplexes (G4s) has been pivotal in revealing their role in transcriptional regulation. However, most of the ligands described to date lack intra-G4 selectivity, severely limiting their potential for uncovering the biological function of individual G4s across the genome. To overcome these limitations, we developed ATENA (Approach to Target Exact Nucleic Acid alternative structures). ATENA relies on the chemical modification of established G4-ligands to enable their conjugation onto a catalytically inactive Cas9 protein (dCas9) using HaloTag, allowing for the targeting of individual G4s in living cells. We have systematically screened the length of the PEG-linkers connecting the G4-ligands to the HaloTag and sgRNA sequences to attain optimal G4 engagement both *in vitro* and in cells. Using optimized conditions, we leveraged ATENA to demonstrate how the selective targeting of the well-studied G4 in the promoter of the oncogene *c-MYC* suppresses its transcription exclusively from the P1 promoter. We also show that positioning ligands in the proximity of regulatory elements suppresses *c-MYC* transcription in a G4-independent manner, highlighting the importance of appropriate design to measure genuine G4-mediated transcriptional changes. We also demonstrate that selective targeting of a G4 in the *PVT1* promoter can either stimulate or repress its transcription depending on the type of G4-ligand used, indicating that functional responses associated with G4-stabilization can highly depend on the type of ligand used. We further harnessed ATENA to study transcriptional perturbation associated with cell-specific G4s, revealing that the functional responses associated with these structures are tightly linked with the expression levels of the targeted gene. Our study provides critical insights into G4-based therapeutic design, offering an innovative platform to investigate G4 biology with high precision.

Introduction

G-quadruplex (G4) structures can promptly form within single-stranded DNA sequences rich in guanines by generating stacks of G-quartets held together by Hoogsteen H-bonding that are further stabilized by coordination of potassium ions (Fig.1a) [1]. The ability of G-rich sequences to form G4s under physiological conditions has been known for decades [2, 3]. Still, the biological relevance of these structures has been heavily disputed until recently. Over the past decade, the development of orthogonal approaches to detect and map G4s has provided robust evidence to support their formation in living cells. These include immuno-fluorescence [4], live-cell imaging [5, 6], and genome-wide mapping strategies [7-10].

While G4-formation in cells has been extensively validated, their direct contribution to regulating biological processes is yet to be fully demonstrated. Given the high enrichment of G4s at gene-promoters measured experimentally *in vitro* and in cells [9, 11], it has long been speculated that these structures may play an essential role in regulating gene expression [12]. Indeed, many G4-selective ligands have been developed to date, revealing that targeting G4s within gene promoters is generally associated with transcriptional suppression [13]. Since key oncogenes like *c-MYC, KRAS,* and *BCL-2* bear a G4 in their promoter, applying G4 ligands for cancer intervention has been extensively investigated [14]. However, the human genome displays more than 700,000 experimentally detected G4-structures [11, 15], making single-G4 targeting by small molecules extremely challenging. Indeed, most ligands recognize G4s by establishing π- π end-stacking with G-tetrads, a promiscuous interaction that hampers these molecules from displaying any significant intra-G4 selectivity [16]. For example, evidence supports that the transcriptional suppression of *c-MYC* observed upon treatment with certain G4-ligands is an indirect consequence of global G4-stabilization rather than a specific response regulated exclusively by the G4 in the *c-MYC* promoter [17].
Additionally, there are experimental discrepancies between what could be the endogenous biological function of G4s and what is instead observed when these structures are bound by ligands. Indeed, a series of independent genomic studies consistently indicated that G4-formation is strongly associated with active gene expression [18], contrasting the transcriptional repression observed upon G4-ligands treatment. To justify this, it has been postulated that ligand binding can prevent key transcriptional activators from recognizing G4s, leading to transcriptional repression upon G4-ligand treatment [19]. However, it remains challenging to discern whether transcriptional suppression at specific genes is caused by protein displacement at given G4 sites or is elicited by global G4-stabilization, as described for *c-MYC* [17].

It is evident that the lack of intra-G4 selectivity typical of most G4 ligands strongly hampers their use as therapeutic agents and their application as reliable tools to investigate G4 biology. Therefore, the development of G4-ligands displaying preferential binding toward a specific G4 structure (i.e., *c-MYC*) has represented a longstanding quest in this area of research. To this end, Schneekloth and co-workers recently designed a small-molecule ligand (DC-34) that displays preferential binding to the G4 in the *c-MYC* promoter [20]. Treatment of multiple myeloma cell lines with DC-34 revealed suppression of *c-MYC* transcription with minimal perturbation of other G4-bearing oncogenes, such as *KRAS* and *BCL-2* [20]. Similarly, our group and others have shown that short peptides and oligonucleotides can also display binding selectivity towards specific G4s[21-24]. Nevertheless, these remain isolated examples that require ligand optimization to achieve binding selectivity against any given G4 of interest and are not suitable for high-throughput screening of G4 function at scale. Besides the lack of intra-G4 selectivity, the highly diverse chemical and structural nature of G4-ligands represents an additional confounding

factor. Indeed, different G4-ligands might elicit different biological responses when binding to the same G4s, reflecting the different structural nature of the G4-ligand complexes formed. This might lead to experimental observations that are dependent on the type of ligand used, rather than reflecting the endogenous biological function of the targeted G4.

To overcome these limitations, we have developed ATENA (Approach to Target Exact Nucleic Acid alternative structures), a CRISPR-Cas9-based platform that enables selective localization of different G4-ligands in the proximity of a given G4 of interest (Fig.1b). To achieve this, we have exploited HaloTag technologies to selectively install modified G4-ligands onto a protein of interest in living cells [25]. More specifically, we have designed a small library of established G4-ligands (pyridostatin (PDS) [26] and PhenDC3 [27]) functionalized with chloroalkane side chains that ensure incorporation into a nuclease inactive Cas9 protein fused to a HaloTag (dCas9-Halo) [28]. We have synthesized analogs of PDS and PhenDC3 with different PEG-linkers (n = 0, 2, 4, Fig.1c,d) conjugating the G4-binding scaffold to the chloroalkane, demonstrating efficient conjugation to the dCas9-Halo *in vitro*. We have also optimized targeting sgRNA sequences and PEG-linkers to ensure effective G4 engagement *in vitro* by developing a dedicated FRET-based assay. Next, we leveraged this knowledge in cells by selectively targeting the G4 in the proto-oncogene *c-MYC* using dCas9-Halo functionalized with either PDS or PhenDC3. Our results revealed that selective downregulation of transcripts controlled by the P1 promoter was observed upon treatment with either ligand, while a negligible reduction of global *c-MYC* expression was detected. These observations were also confirmed by treatment with the *c-MYC* selective ligand DC-34, validating the ability of the dCas9-Halo platform to target individual G4s. We have also observed that guiding G4-ligands in the proximity of TATA-boxes within gene promoters can lead to unspecific *c-MYC* downregulation, suggesting that careful design of the sgRNA sequence is required to ensure that observations made with ATENA are genuinely representative of G4-engagement. While downregulation of P1-mediated *c-MYC* expression was observed with PDS and PhenDC3, ATENA also revealed that other G4s could respond differently to ligand binding. These findings confirmed that the biological response elicited by G4-binding might reflect more the binding nature of the ligand rather than providing direct insights into the endogenous function of the targeted G4. This highlights the importance of considering the nature of the ligand used when assessing the biological function and the therapeutic potential of a given G4-structure. Finally, we showcase the ability of ATENA to infer the biological relevance of cell line-specific G4s, revealing a strong dependence of the functional response observed upon ligand treatment with the expression level of the gene targeted.

Altogether, our study provides robust evidence supporting ligand and transcriptional-dependent responses to G4-targeting. ATENA provides a modular platform to target individual G4s in living cells, enabling the study of G4 biology with high precision and control. We anticipate that ATENA will offer unprecedented potential for screening cell and ligand-specific responses to G4-targeting in a high-throughput fashion, which can be utilized for therapeutic design and development.

Results

Chemical labelling of CRISPR-Cas9 proteins with ligands for individual G4-targeting

Catalytically inactive CRISPR-Cas9 (dCas9) fused to functional proteins have been widely used in biology to attain site-selective perturbation of gene expression [29]. This strategy takes advantage of the selectivity provided by dCas9 bound to a short-guiding RNA (sgRNA) in recognizing a specific genomic site by base-pairing, which can be used to recruit an effector protein (i.e., a transcription factor or an epigenetic enzyme) at the targeted site [30]. We reasoned that a similar strategy could have been devised to attain single G4-targeting by decorating chemically dCas9 with G4-ligands. Chemical

functionalization of dCas9 proteins has been previously achieved using commercially available chloroalkane-modified fluorophores to label a dCas9-Halo fusion protein in living cells [30]. Therefore, we hypothesized that generating modified G4-ligands functionalized with chloroalkane moieties could have been exploited to attain site-selective G4-targeting under physiological conditions. To achieve this, we designed analogs of the widely established G4-ligands PDS based on the previously described PyPDS scaffold [5]. Unlike PDS, PyPDS presents a single primary amine within its structure, which can be selectively functionalized with chloroalkane sidechains by peptide coupling (Fig.1c) [5]. After successful synthesis of PyPDS following previously established methods [5], we functionalized the primary amine of the molecule with chloroalkane side chains of different lengths, enabling systematic investigation of the ideal distance between the G4-binding scaffold PyPDS and the HaloTag protein to achieve optimal G4-engagement. Specifically, we have used linkers containing different PEG repeats (n = 0, 2, 4) to vary the distance between PyPDS and the chloroalkane (Cl-PDS$_n$, Fig.1c, SI-1).
To avoid limiting the use of ATENA to a single G4-ligand, we decided to functionalize with chloroalkane moieties another widely characterized G4-ligand called PhenDC3 (Cl-PhenDC3$_n$) [27]. Unlike PDS, PhenDC3 has a cationic side chain and is structurally bulkier, displaying a phenanthroline core wider than the pyridine one present in the PDS scaffold (Fig.1c, d). While PDS and PhenDC3 have been extensively validated as selective G4-ligands, the structural differences between these two ligands can lead to differential biological responses when applied in cells. Therefore, we decided to compare these two ligands with ATENA when recruited to a single-G4 site. To achieve this, we have synthesized a previously reported PhenDC3 analog displaying an exocyclic primary amine [31], which can be functionalized by peptide coupling using the same synthetic strategy described for PyPDS to afford chloroalkane-modified PhenDC3 analogs that are compatible with HaloTag conjugation (Fig.1d, SI-1).

With both PyPDS and PhenDC3 analogs in hand, we first assessed if functionalization with the chloroalkane side chains could prevent the G4-binding properties of these molecules. To test this, we have subjected all the ligands to a standard FRET-melting assay to evaluate their ability to stabilize G4-structures. All analogs displayed good G4-stabilization, providing an increase in melting temperature (ΔT_m >10 ^oC at 4 μM) for three distinct G4 structures tested (hTelo, BCL-2 (Fig.1e) and c-KIT-2 (Extended Data Fig.1a). All the chloroalkane functionalized PhenDC3 analogs showed ΔT_m values comparable to those observed with the PhenDC3 (Extended Data Fig.1c, d), suggesting that the addition of the chloroalkane side chains had negligible effects on the G4-stabilization properties of the PhenDC3 scaffold. On the other hand, when functionalizing PyPDS with chloroalkane side chains, lower ΔT_m values than PyPDS are generally observed (Fig.1e Extended Data Fig.1c, d). This agrees with previous studies demonstrating that the functionalization of the primary amine of the PyPDS scaffolds leads to a reduction of the global charge of the molecule under physiological conditions, which is associated with slightly weaker stabilization of G4-structures compared to the parent compound [5]. Despite the reduced G4 stabilization, PyPDS analogs with a functionalized primary amine, such as SiR-PyPDS, display sufficient binding affinity and selectivity for reliable G4 detection in living cells [5].
Having confirmed that both PyPDS and PhenDC3 analogs functionalized with chloroalkane sidechains retained good G4-binding recognition properties, we next investigated if these molecules could be used to label the dCas9-Halo *in vitro*. To this end, we expressed and purified the dCas9-Halo protein (see methods) and incubated it for 45 minutes with increasing concentrations of the G4-ligands. This was followed by incubation with an excess (5 μM) of the commercially available HaloTag®TMR ligand to label dCas9-Halo with the TMR fluorophore. Since labeling of HaloTag is a covalent irreversible process [25], we reasoned that if the chloroalkane functionalized G4-ligands efficiently labeled the dCas9-Halo, they would prevent the subsequent incorporation of the fluorescent HaloTag®TMR, leading to a dose-dependent reduction of TMR incorporation, as schematically depicted in Fig.1f. All the analogs

tested induced a robust dose-dependent reduction of the dCas9-Halo TMR signal, demonstrating high efficiency in labeling dCas9-Halo irrespectively of the nature of G4-ligand (i.e., PyPDS or PhenDC3) or the PEG-linker used to connect the chloroalkane to the G4-binding scaffold (Fig.1g and Extended Data Fig.1d, e, f).

We next asked if dCas9-Halo functionalized with G4-ligands could be used to target individual G4s. To this end, we designed a dually fluorescent-labeled DNA template containing the c-KIT2 G4 forming sequence at its 3' end, which could be used to monitor G4-stabilization as a function of both the PEG-linkers and short guiding RNAs (sgRNAs) used (Fig.1h). To investigate this, we generated two sgRNAs to orient the dCas9-Halo complex towards the G4-structure sitting at either 18 ($sgRNA_{FRET-18}$) or 42 ($sgRNA_{FRET-42}$) base pairs from the targeted structure and whose Protospacer Adjacent Motifs (PAM) are located on the same strand of the G4 forming sequence. This is based on previous structural studies indicating that the C-terminal domain of the Cas9 protein, where the Halo protein is situated, will point towards the 3' end of the PAM sequence localized on the same strand of the targeting DNA sequence [32]. To control for any change in G4-stability induced by forming the dCas9-Halo complex on its own, we have also designed $sgRNA_{FRET-41}$ that also sits at 41 base pairs from the G4 but whose PAM is located on the reverse strand, orienting the dCas9-Halo toward the 5' end of the sequence and preventing any ligand-mediated G4-engagement (Fig.1h). Moreover, we have used a standard non-targeting RNA control (NTC) to rule out any potential unspecific binding that is not strictly mediated by sgRNA-driven proximity. As depicted in Fig.1i, we failed to detect any significant changes in the FRET signal for any of the sgRNAs tested when we decorated dCas9-Halo with Cl-PDS_0. Considering that Cl-PDS_0 can efficiently bind to dCas9-Halo (Fig.1g), the lack of G4-stabilization displayed by this molecule indicates that the linker connecting the PDS-scaffold to the dCas9-Halo is too short for engaging with the G4-structure. Conversely, when labeling dCas9-Halo with Cl-PDS_2 an increase in FRET-signal could be measured when using both $sgRNA_{FRET-18}$ and $sgRNA_{FRET-42}$ (Fig.1i and Extended Data Fig. 1g). Interestingly, the extent of FRET changes appeared to be stronger with $sgRNA_{FRET-18}$ compared to $sgRNA_{FRET-42}$, suggesting that placing the dCas9-PyPDS complex closer to the G4 facilitates ligand-structure engagement (Fig.1i). Additionaly, when using the control sgRNAs ($sgRNA_{FRET-41}$) no changes in the ΔFRET efficiency were detected (Fig.1i), strongly supporting that dCas9-PyPDS can induce G4-stabilization in a sgRNA dependent fashion. Similarly, when decorating dCas9-Halo with Cl-PDS_4, we observed an increase in the ΔFRET efficiency with $sgRNA_{FRET-18}$ and $sgRNA_{FRET-42}$, but to a lower extent than what was measured with Cl-PDS_2, which indicated that extensively long PEG-linkers can be detrimental to G4-recognition (Extended Data Fig. 1h). This demonstrated that dCas9-driven G4-ligand engagement is both molecule and sgRNA-dependent, with optimal G4-stabilization obtained when a PEG2 linker separated the PyPDS scaffold from the chloroalkane moiety and when the dCas9-Halo was placed ~18 base pairs from the targeted G4.

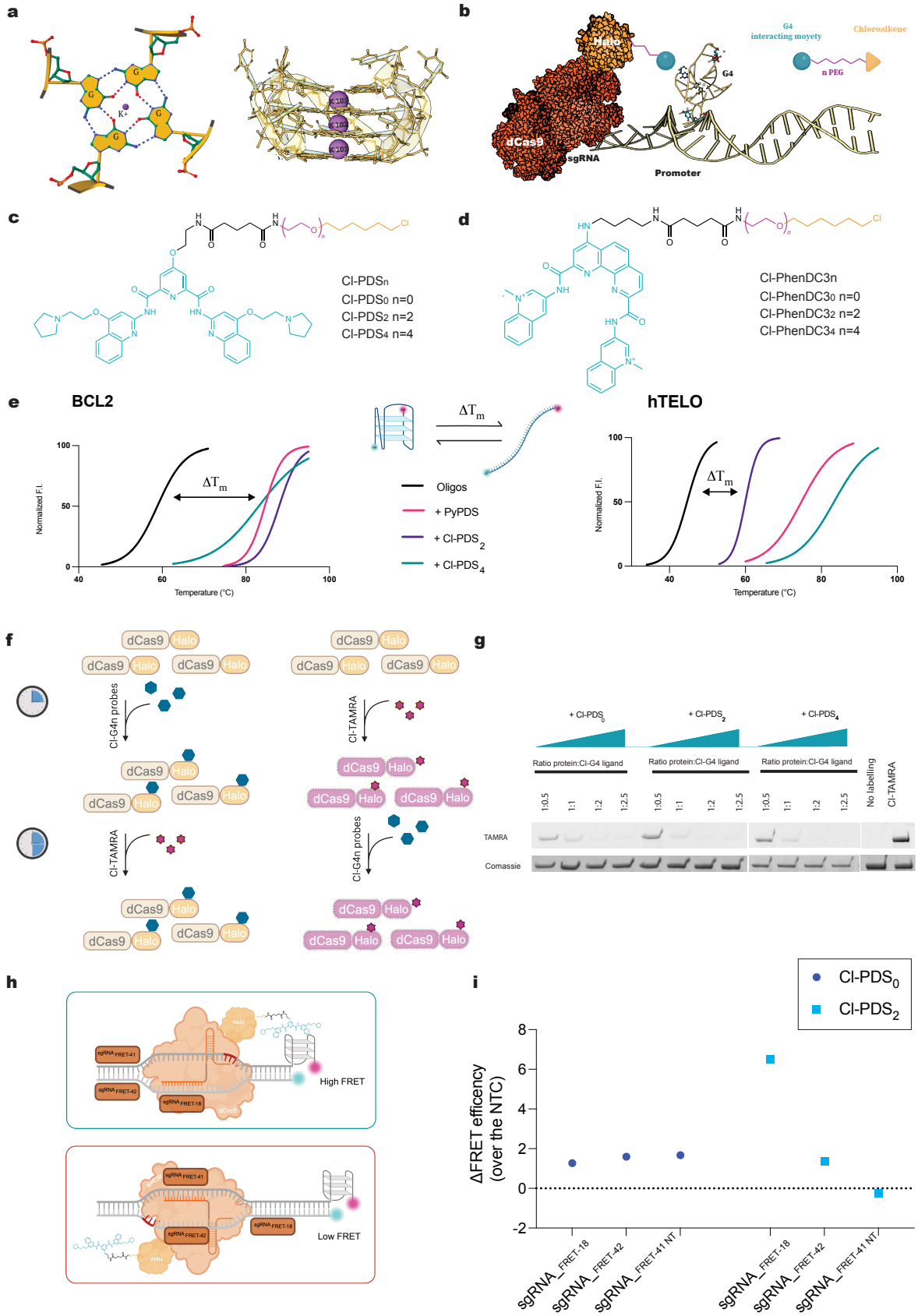

Figure 1: Chemical labeling of dCas9 for selective G4-targeting. **a**, Schematic representation of a G-Tetrad *via* Hoogsteen base pairing with central cation (K^+) (modified from PDB file:6W9P) (Left) and G-Tetrad stacking to form G4-structure (modified from PDB file:6W9P using Protein Imager [33]).

(Right). **b**, Schematic overview of ATENA: dCas9-Halo fusion protein functionalized with chloroalkane-modified G4-ligands enables single G4-targeting through sgRNA guidance. **c,** Chemical structure of chloroalkane-modified PyPDS where n indicates the different PEG linkers length (Cl-PDS_n). **d,** Chemical structure of chloroalkane-modified PhenDC3 (Cl-$PhenDC3_n$) where n indicates the different PEG linkers length (Cl-$PhenDC3_n$). **e,** FRET-melting assay of 4 μM Cl-PDS_2 or 4 μM Cl-PDS_4 with 0.4 μM of dually labeled BCL-2 G4-forming sequence (left) or hTelo –G4 forming sequence (right). **Δ**Tm indicates the difference in melting temperature between folded and unfolded G4. The mean is calculated from two independent experiments (n=2). **f,** Illustration of the competition assay workflow to test the binding ability of Cl-PDS_n probes to dCas9-Halo purified recombinant protein. **g,** SDS-Page gel of the Cl-PDS_n competition assay shows each sample's fluorescent level acquired in the TAMRA channel (542 nm) and the corresponding protein level (Coomassie staining). **h,** Schematic representation of the dually labeled FRET oligos containing c-kit–G4 forming sequence bound by dCas9-Halo labeled with Cl-PDS_n probes to study G4 stabilization with respect to the PEG-linker length and sgRNAs. **i,** ΔFRET efficiency. The values indicated were extrapolated from the band intensity (ImageStudio Software) measured in the Cy3 channel (Typhoon FLA 9500- 550 nm). Specifically, the ΔFRET efficiency was calculated as follows ((1-(Cy3 band intensity in presence of Cl-PDS_n/ Cy3 band intensity samples without molecule)) and normalized for the non-targeting sgRNA control (n=2).

Use of optimized ATENA for c-MYC targeting in cells.

Having established optimal conditions to attain dCas9-mediated G4-targeting *in vitro*, we next sought to uncover the potential of ATENA to promote single-G4 stabilization in living cells. The use of cell lines that constitutively express dCas9-Halo is essential to ensure that the cellular levels of this protein are consistent across different experiments, avoiding bias introduced by the significantly variable levels of protein expression typically associated with transient transfection. To this end, we decided to integrate dCas9-Halo into the genome of a breast cancer cell line (MCF7) using standard lentiviral integration approaches (see methods), given the highly diverse response in gene-expression changes previously reported upon treatment of these cells with PDS [34] that could be further investigated with ATENA. Successful integration of dCas9-Halo was confirmed by PCR-based genotyping and Western Blot (Extended Data Fig.2a, b).

Next, we evaluated the ability of the chloroalkane functionalized ligands to label dCas9-Halo in cells. To achieve this, we have leveraged an established chloroalkane penetration assay (CAPA) to evaluate the relative potency of the different ligands to label dCas9-Halo in living cells [35]. In short, during CAPA, cells are initially exposed to different concentrations of the chloroalkane-modified G4-ligands, which are then incubated with a Halo reactive fluorophore (Cl-Oregon Green). The efficiency of ligand incorporation can be measured as an inverse function of the Oregon Green emission, as labeling with the G4-ligands should prevent the incorporation of the fluorophore (Fig.2a). As displayed in Fig.2a, treatment of MCF7 cells stably expressing dCas9-Halo with both Cl-PDS_0 and Cl-PDS_4 revealed a modest reduction of Oregon Green emission, with 60% labeling obtained only at high concentrations of the ligands (10 μM). Conversely, Cl-PDS_2 could label ~90% dCas9-Halo at a concentration as low as 0.25 μM, reaching saturation at 2.5 μM (Fig.2a). This suggested that the cellular permeability and bioavailability of Cl-PDS_2 were particularly suitable for its application in ATENA.

Encouraged by these results, we initially focused on using Cl-PDS_2 to achieve single-G4 targeting in cells with ATENA. We started our investigation by considering the G4 present in the promoter of the *c-MYC* proto-oncogene, as this is one of the most studied G4 described in the literature. Many studies have linked the targeting of MYC-G4 by ligands to the downregulation of *c-MYC* expression [12]. To

address to what extent *c-MYC* expression is perturbed by ligands exclusively by stabilizing the G4 in its promoter, we have designed a series of sgRNAs to guide PDS selectively at the MYC-G4 using ATENA. To ensure the correct orientation of the Halo protein decorated with Cl-PDS_2 toward the MYC-G4, we designed sgRNAs targeting either the non-template strand (NT) bearing the G4-structure at its 3' end or the opposite strand at its 5' end (T) (Fig.2b). Based on our biophysical investigation, we reasoned that placing sgRNAs within a widow of ~50 base pairs would have ensured G4-stabilization by Cl-PDS_2 tethered to dCas9 (Fig.1i). Considering PAM sequences available for dCas9-Halo binding at either G4 ends, we designed $sgRNA_{MYC-19}$ and $sgRNA_{MYC-67}$ that would place the protein complex at respectively 19 and 67 base pairs away from the MYC-G4 at its 3' end on the T strand. Similarly, we generated $sgRNA_{MYC+22}$ and $sgRNA_{MYC+58}$, targeting the MYC-G4 from its 5' end and placing dCas9-Halo/sgRNA complex on the opposite strand at 22 and 58 base pairs, respectively. To further investigate the optimal distance to achieve G4 stabilization in cells with ATENA, we have also designed $sgRNA_{MYC-89}$ and $sgRNA_{MYC-93}$, along with $sgRNA_{MYC+75}$, $sgRNA_{MYC+77,}$ and $sgRNA_{MYC+119}$, also targeting the MYC-G4 at its 3' and 5' end, respectively (Fig. 2b). After cloning sequences encoding the various sgRNAs into a vector for mammal expression (see methods), we have transfected MCF7 cells stably expressing dCas9-Halo with individual sgRNAs, followed by treatment with Cl-PDS_2 for 48 hours. We then measured changes in *c-MYC* expression using RT-qPCR, normalizing the expression level against individual samples transfected with the respective sgRNAs but not exposed to Cl-PDS_2. This enabled us to consider any change in gene expression potentially triggered by the assembly of the dCas9-Halo complex not labeled with PDS on the promoter for each sgRNA used.
To our surprise, we observed that sgRNAs placing the dCas9-PyPDS complex within a ~50 bp window of the MYC-G4 on the T strand ($sgRNA_{MYC-19}$, $sgRNA_{MYC-67}$) led to negligible effects on *c-MYC* global expression (Fig.2c), which contrasts our biophysical predictions (Fig.1i). Conversely when placing the complex on the NT strand with $sgRNA_{MYC+22}$ and $sgRNA_{MYC+58}$, we observed reduction of *c-MYC* expression (~20%), which could be indicative G4-engagement mediated by ATENA (Fig.2c). To investigate this further, we placed the dCas9-PyPDS complex further away from the MYC-G4 using $sgRNA_{MYC+75}$, $sgRNA_{MYC+77,}$ and $sgRNA_{MYC+119}$, which should lead to abrogation of G4-engagement in a distance-dependent fashion. However, we observed an opposite trend by which *c-MYC* downregulation was increased the further away the complex was from the G4 (Fig.2c), which cannot be consistent with a ligand-meditated effect. By closer inspection of the promoter annotation, we noticed that $sgRNAs_{MYC}$ +58 to +119 overlapped with the P1 promoter and TATA-box sequences. In particular, $sgRNA_{MYC+75}$ and $sgRNA_{MYC+77}$ place the dCas9-PyPDS complex directly on the TATA-box sequence, which can justify the observed *c-MYC* suppression. This effect is exacerbated when $sgRNA_{MYC+119,}$ which targets the region next to the TSS of the P1 promoter, is used. Therefore, we reasoned that *c-MYC* suppression likely reflected perturbation of the transcriptional homeostasis regulated by the TATA-box domain caused by the dCas9-PyPDS complex rather than G4-stabilization. To test this hypothesis, we decorated dCas9 with Cl-Oregon Green or Cl-PDS_2 and monitored *c-MYC* expression changes when using $sgRNA_{MYC+58}$. Cl-Oregon Green is a fluorophore and is unable to stabilize G4s. Therefore, potential changes in *c-MYC* expression observed with dCas9-OregonGreen will be indicative of transcriptional dysregulation that is G4-independent. Crucially, we observed a suppression of *c-MYC* when using dCas9-Oregon-Green that was comparable to dCas9-PyPDS (Fig.2d). This strongly suggested that targeting the MYC-G4 at its 5' end using dCas9-conjugates is unable to provide reliable observations on G4-mediated transcriptional regulation due to the interference with key regulatory sites, such as the P1 promoter and TATA-boxes.

Considering that multiple promoters regulate *c-MYC* expression [36, 37], we investigated changes in transcriptional levels elicited by dCas9-PyPDS by focusing on the transcript generated by the G4-dependent P1 promoter. Indeed, it has been recently shown that deletion of the MYC-G4 from its

promoter only affects the expression of *c-MYC* controlled by the P1-promoter, while modest perturbation of global *c-MYC* transcription is observed [38]. To test this, we quantified *c-MYC* expression changes caused by dCas9-PyPDS using primers that exclusively amplify transcripts generated by the P1 promoter. Gratifyingly, we observed a distance-dependent suppression of P1-mediated *c-MYC* expression when targeting MYC-G4 with $sgRNA_{MYC-19}$ and $sgRNA_{MYC-67}$ (~6 and 4-fold, respectively, Fig.2e). Importantly, no significant changes in expression were measured when placing the dCas9-PyPDS complex further away from the G4 with $sgRNA_{MYC-89}$ and $sgRNA_{MYC-93}$, which is in agreement with our biophysical observations and consistent with G4-engagement (Fig.2e). Similarly, when targeting the G4 from its 5' end by placing the dCas9-PyPDS complex on the complementary strand (NT) with $sgRNA_{MYC+22}$, we also detected a reduction of P1-mediated *c-MYC* expression of 4-fold (Fig.2e). We also detected some transcriptional repression when targeting G4 with $sgRNA_{MYC+58}$ and $sgRNA_{MYC+75}$, although this might be G4-independent effects as discussed above (Fig.2e). However, P1-mediated *c-MYC* expression was entirely abrogated when the dCas9-PyPDS complex was placed at sites overlapping with the GC-box and close to the P1-TSS with $sgRNA_{MYC+119}$ (Fig.2e). These observations further confirmed that placing the dCas9 complex on the 5' end of the MYC-G4 cannot be used to reliably detect changes in gene expression that are strictly mediated by G4-targeting.

We next set out to validate if the changes in P1-dependent *c-MYC* expression observed upon treatment with $sgRNA_{MYC-19}$ and $sgRNA_{MYC-67}$ could indeed be ascribed to dCas9-PyPDS binding to the MYC-G4. Firstly, we assessed if selective targeting of the MYC-G4 with ATENA was associated with changes in the expression of another gene bearing a stable G4 in its promoter, such as *KRAS*. As displayed in Extended Data Fig.2c, no detectable change in *KRAS* expression could be detected when directing ATENA to MYC-G4, which supports selective MYC-G4 targeting. Conversely, when using free PyPDS, *KRAS* expression is significantly reduced (Extended Data Fig.2d), further supporting how ATENA can provide intra-G4 selectivity to ligands like PyPDS.

We further validated ATENA-mediated G4 stabilization by looking at protein accessibility at the G4 site upon targeting. It has been proposed that ligands bound to G4s can displace key transcription factors and regulatory proteins, leading to transcription suppression [19, 39]. Therefore, we reasoned that if dCas9-PyPDS complex was correctly positioned to enable ligand-G4 interaction, we would have observed reduced protein accessibility to the G4. To test this hypothesis, we have used the G4-selective antibody BG4 and performed CUT&Tag in the presence of dCas9-PyPDS in conjunction with either $sgRNA_{MYC-19}$ or a non-targeting control (NTC). We then used qPCR to measure the enrichment of the MYC-G4 signal against 3 independent validated G4-sites in MCF7 cells (MAZ, RPA3, RBBP4) to assess relative protein accessibility at these G4s under different conditions, as previously described [38]. As displayed in Fig.2f, when treating cells with dCas9-PyPDS using $sgRNA_{MYC-19}$, we observed a consistent reduction of BG4 accessibility that was not detected for the NTC, irrespective of the reference G4 used. This strongly indicates that PDS can bind the MYC-G4 guided by the dCas9, preventing the BG4 antibody from binding to the same target. These observations supported a model by which dCas9-PyPDS guided by accurately selected sgRNAs can lead to ligand-mediated G4-stabilization and P1-specific *c-MYC* downregulation.

To further demonstrate that *c-MYC* expression changes observed with ATENA could be confidently ascribed to G4-stabilization, we employed PhenDC3, as a validated G4-ligand that is structurally very different from PDS. To this end, we used the PEG-2 analog Cl-$PhenDC3_2$, which also showed good labeling of dCas9-Halo in cells by CAPA (Extended Data Fig.2e). Similarly to what was observed with PyPDS, also dCas9-PhenDC3 caused a significant reduction of P1-mediated *c-MYC* expression when used in conjunction with $sgRNA_{MYC-19}$ (Fig.2g). Interestingly, this combination led to a higher

suppression of transcription (~14-fold, Fig.2g), possibly reflecting the stronger G4-stabilization abilities of Cl-PhenDC3$_2$ compared to Cl-PDS$_2$ measured by FRET (Extended Data Fig.1b, c). This suggested that ATENA might also be used to assess the relative intensity of phenotypes that different G4-ligands elicit on the same target.

Altogether, our data indicate that ATENA can successfully target MYC-G4 in a ligand-independent fashion, leading to detectable G4 engagement and corresponding P1-specific *c-MYC* suppression.

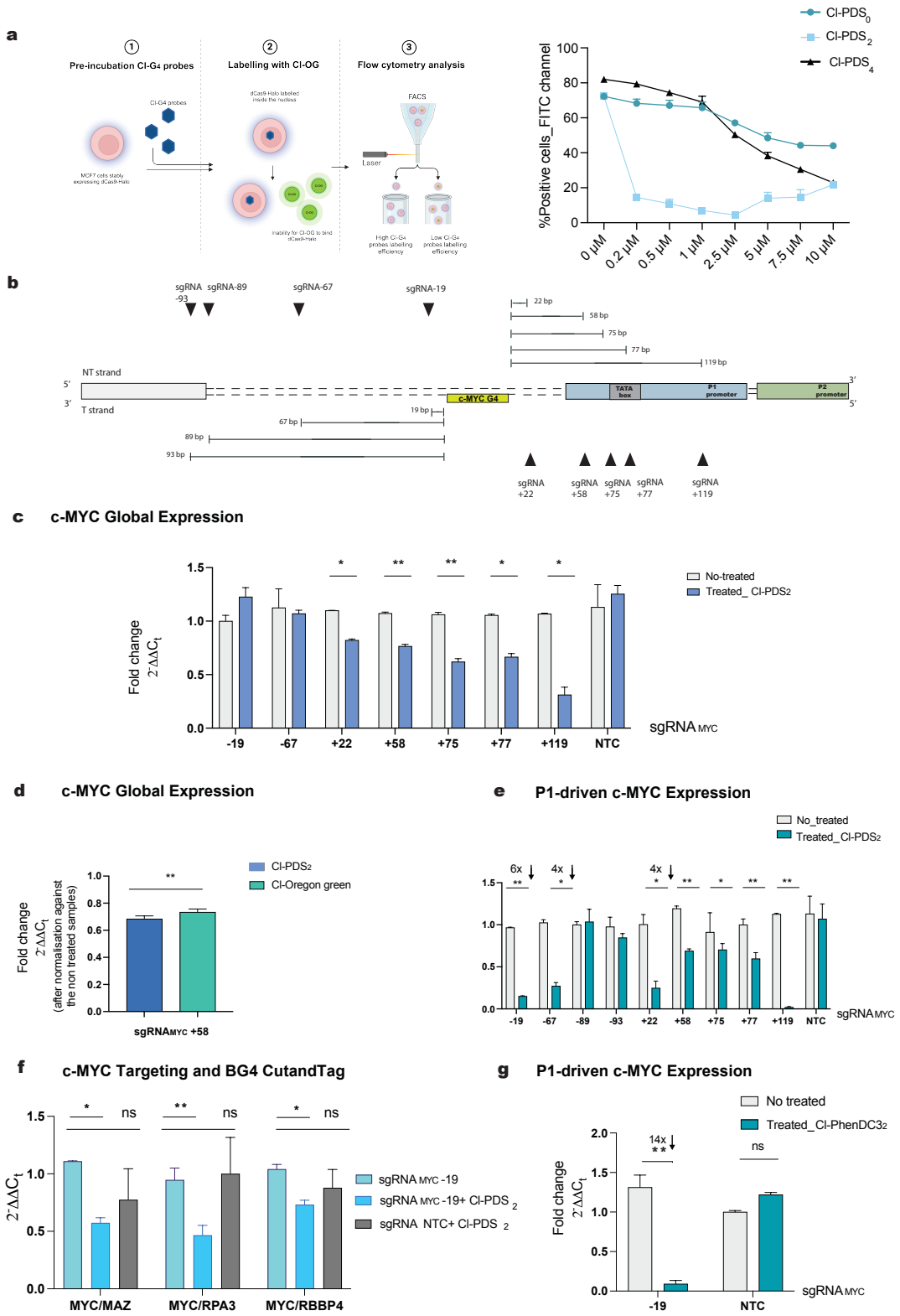

Figure 2: ATENA enables selective targeting of a G4 in the *c-MYC* promoter. a, Schematic representation of CAPA assay (left). CAPA assay results were obtained upon treatment of MCF7 cells

stably expressing dCas9-Halo to assess Cl-PDS_n label efficiency in cells at different concentrations (n=2). Flow cytometry was performed using Attune NxT Flow cytometer, and analysis was performed using FlowJo software (right). **b**, Schematic illustration of the *c*-MYC promoter with annotated G4, sgRNA targeting region (black triangles), and their relative distance in bp from the G4. **c**, RT-qPCR for *c-MYC* expression in MCF7 cells stably expressing dCas9-Halo transfected with the indicated sgRNAs and incubated for 48h in the presence of 2.5 μM of Cl-PDS_2. The expression values are represented as fold change ($2^{-\Delta\Delta Ct}$) with respect to the non-treated transfected samples and after normalization for a housekeeping gene (GAPDH). n=3, biological replicates, each of which included two technical replicates. **d,** RT-qPCR for *c-MYC* expression in MCF7 cells stably expressing dCas9-Halo transfected with $sgRNA_{MYC+58}$ and incubated for 24h in the presence of either Cl-PDS_2 (2.5 μM) or Cl-Oregon green (5 μM). The expression values are represented as fold change ($2^{-\Delta\Delta Ct}$) with respect to the non-treated transfected samples and after normalization for a housekeeping gene (GAPDH). n=2, biological replicates, each of which includes two technical replicates. **e,** RT-qPCR for P1-driven *c-MYC* expression in MCF7 cells stably expressing dCas9-Halo transfected with the indicated sgRNAs and incubated for 48h in the presence of 2.5 μM of Cl-PDS_2. The expression values are represented as fold change ($2^{-\Delta\Delta Ct}$) with respect to the non-treated but transfected samples and after normalization for a housekeeping gene (GAPDH). n=3, biological replicates, each of which includes two technical replicates. **f,** BG4 CUT&Tag-qPCR for MCF7 cells stably expressing dCas9-Halo transfected with either $sgRNA_{MYC-19}$ or sgRNA NTC and treated with DMSO or 2.5 μM Cl-PDS_2. BG4 accessibility was analyzed for *c-MYC* and normalized to three G4s in control gene sites (RPA3, MAZ, RBBP4). n=2, biological replicates for each sample three technical replicates for BG4 and one for the negative (no BG4 treatment). **g,** RT-qPCR for P1 dependent *c-MYC* expression in MCF7 cells stably expressing dCas9-Halo transfected with $sgRNA_{MYC-19}$ or sgRNA NTC and incubated for 48h in the presence of 2.5 μM of Cl-PDS_2. The expression values are represented as fold change ($2^{-\Delta\Delta Ct}$) with respect to the non-treated but transfected samples and after normalization for a housekeeping gene (GAPDH). n=3, biological replicates, each of which includes two technical replicates. Data presented are mean of n = number of independent biological samples. Statistical significance was calculated using a two-tailed t-test. Only significant values are represented P-value: ns > 0.05, * ≤0.05, ** ≤0.01, *** ≤0.001, **** ≤0.0001. Fold change (FC) was calculated on the normalized expression value.

ATENA mimics the MYC-G4 selective molecule DC-34

Encouraged by the P1-dependent *c-MYC* suppression observed with ATENA, we explored if a similar phenotype could be elicited when using ligands displaying intra-G4 selectivity. To this end, we leveraged the MYC-selective G4-ligand DC-34 (Fig.3a), demonstrating a strong binding affinity for the MYC-G4 with high selectivity over other G4-structures [20]. Indeed, DC-34 has been previously shown to cause G4-dependent *c-MYC* downregulation while displaying minimal interference with other genes bearing a G4 in their promoter, such as *KRAS* and *c-KIT* [20]. Therefore, we reasoned that treating MCF7 cells with DC-34 should have led to similar P1-specific *c-MYC* downregulation as observed with ATENA.

To test this hypothesis, we treated MCF7 cells with increasing concentrations of DC-34 for 48 hours before measuring changes in *c-MYC* expression employing RT-qPCR. Similarly to what was observed with ATENA, treatment with DC-34 caused negligible dose-dependent changes in global *c-MYC* expression (Fig.3b), supporting the notion that selective targeting of MYC-G4 does not impact the global expression of the gene in MCF7 cells. Conversely, when measuring P1-mediated expression, DC-34 revealed a strong dose-dependent suppression that plateaued at ~4-fold reduction measured at 10 μM (Fig.3c). This indicated that using an intra-G4 selective ligand as an orthogonal approach to ATENA for selective MYC-G4 targeting led to similar observations, further corroborating the validity of our platform for single G4-targeting.

We next decided to assess the relative selectivity toward MYC-G4 of DC-34. To achieve this, we have monitored dose-dependent changes of another proto-oncogene bearing at G4 in its promoter, *KRAS*. We have already demonstrated that using ATENA for MYC targeting does not lead to statistically significant changes in *KRAS* expression (Extended Data Fig 2.c). Conversely, treatment with DC-34 revealed a dose-dependent reduction of *KRAS* expression in treated cells (Fig.3d). The suppression of *KRAS* elicited by DC-34 treatment (1.6-fold) is not as pronounced as for *c-MYC*, suggesting that MYC-G4 is the primary target of the molecule but some residual binding to other G4 promoters can be observed at high doses. This reflects the high structural similarity of different G4s, which makes their individual targeting by small-molecule ligands challenging.
To further validate the G4-dependent *c-MYC* downregulation elicited by DC-34, we have also performed CUT&Tag enrichment at the MYC-G4 upon treatment with this molecule. As displayed in Fig.3e, DC-34 treatment led to a reduction of MYC-G4 enrichment by CUT&Tag, consistent with what was measured with dCas9-PyPDS. This further demonstrates that MYC-G4 targeting reduces protein accessibility and binding, which can be used indirectly to assess ligand-G4 engagement.

We next compared the intra-G4 selectivity provided by ATENA against DC-34 by looking at transcriptome-wide gene-expression changes by RNA-Seq. To achieve this, we generated RNA-Seq data upon treatment with DC-34 or dCas9-PyPDS conjugated to $sgRNA_{MYC-19}$ and compared these results to transcriptome-wide changes attained upon treatment with free PyPDS. We reasoned that higher intra-G4 selectivity provided to PyPDS by ATENA would result in a lower total number of differently expressed genes measured by RNA-Seq. When treating MCF7 cells with PyPDS (2.5 μM) for 6h, RNA-Seq data revealed that 3,041 genes were differentially expressed (429 down-regulated, 2612 upregulated, Extended Data Fig.3a, b). Conversely, when treating MCF7 cells with DC-34 (7.5 μM) under the same conditions, we observed a significantly lower number of differentially expressed genes (126 total, 40 down-regulated, 86 up-regulated, Figure 3f). This clearly indicated that the intra-G4 selectivity of DC-34 is reflected in a substantially reduced perturbation of the global MCF7 transcriptome. Moreover, RNA-Seq data obtained upon DC-34 treatment confirmed that the global expression of *c-MYC* is not affected by G4-stabilization under conditions in which we have demonstrated engagement of the molecule to the MYC-G4 (Fig.3e). This further indicates that G4-stabilization limits perturbation on transcripts under the control of the P1-promoter [38], which should be considered when devising G4-based therapies aiming at suppressing *c-MYC* expression. Considering the great reduction of differently expressed genes displayed by DC-34 when compared to a global G4-stabilizer like PyPDS, we hypothesized that conjugation of PyPDS to dCas9 in ATENA would have also resulted in a strong reduction of global transcriptional perturbation. To investigate this, we have transfected MCF7 cells with $sgRNA_{MYC-19}$ followed by treatment with $Cl\text{-}PDS_2$ for 6 hours before extracting poly-adenylated RNA and performing RNA-Seq. When comparing differentially expressed genes between sgRNA-transfected cells that were not treated with $Cl\text{-}PDS_2$ and those that were exposed to the ligand, we also observed a strong reduction of transcriptionally perturbed genes (125 total, 19 down-regulated, 106 up-regulated, Fig.3g). This clearly indicates that ATENA confers strong intra-G4 selectivity to PDS causing a reduction of a total of 2,916 genes that are no longer differently expressed when PyPDS is conjugated to dCas9. This further exemplifies how selective MYC-G4 targeting with ATENA mimics the behavior of a selective binder like DC-34, confirming the robustness of this approach to investigate the biology regulated by individual G4s.

Altogether, these observations validated the MYC-G4-dependent responses measured with ATENA, which are fully recapitulated by treatment with the selective MYC-G4 ligand DC-34. Moreover, our

data highlight the power of ATENA to provide intra-G4-selectivity to canonical ligands, enabling high-precision study of biological responses elicited by individual G4s.

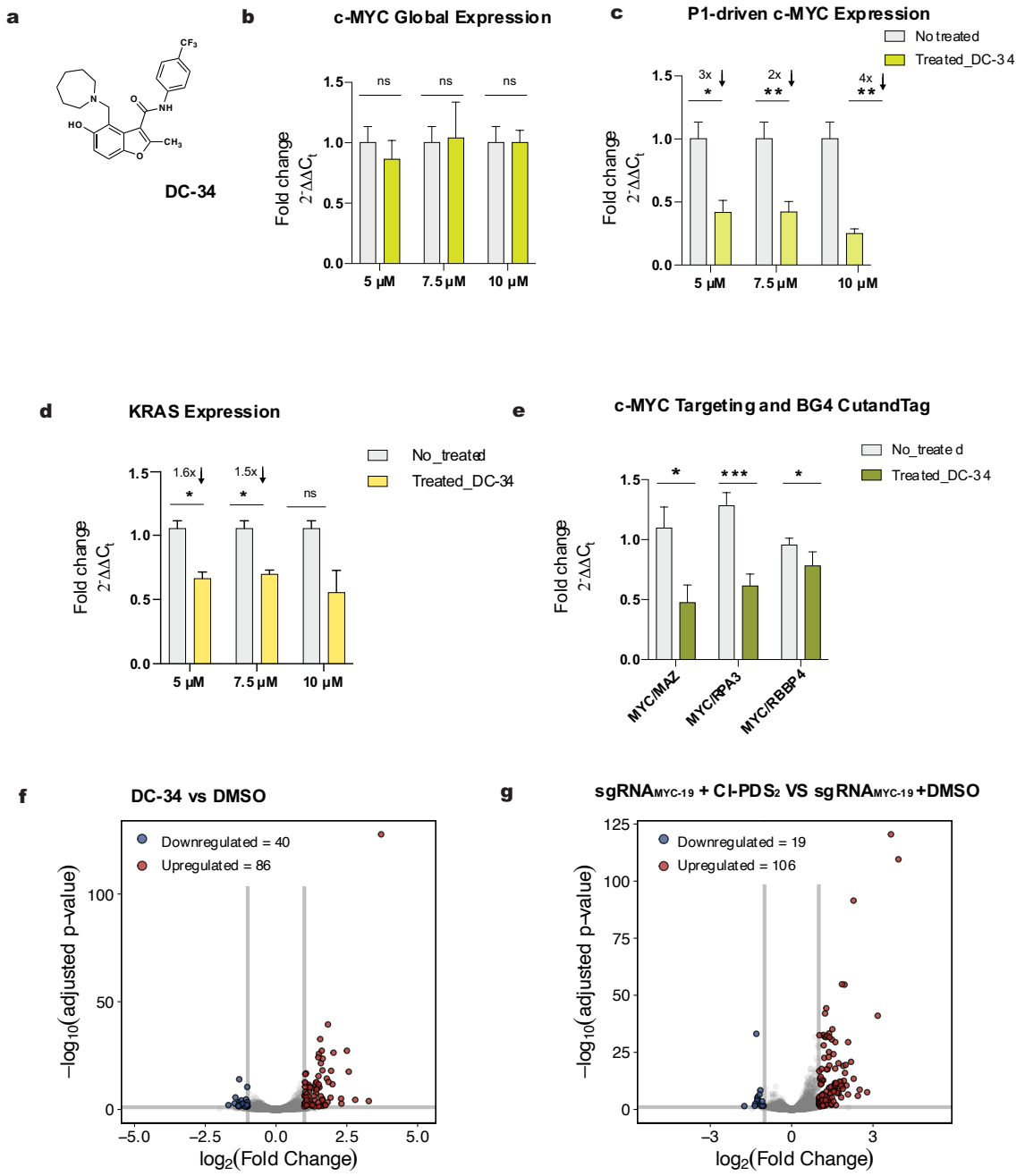

Figure 3: The selective G4-MYC ligand DC-34 matches ATENA. a, Chemical structure of DC-34. **b,** RT-qPCR for c-Myc expression in MCF7 cells treated with increasing concentration of the *c-MYC-G4* ligand DC-34 after 48h of treatment. Values are represented as fold change ($2^{-\Delta\Delta Ct}$) with respect to the DMSO samples and after normalization for a housekeeping gene (GAPDH). n=3, biological replicates, each of which included two technical replicates. **c**, RT-qPCR for P1-driven *c-MYC* expression in MCF7 cells treated with different concentrations of the *c-MYC-G4* ligand DC-34 after 48h of treatment. Values are represented as fold change ($2^{-\Delta\Delta Ct}$) with respect to the DMSO samples and

after normalization for a housekeeping gene (GAPDH). n=3, biological replicates, each of which includes two technical replicates. **d**, Evaluation of potential DC-34 off-targets by analysis of KRAS expression in MCF7 cells DC-34 treated samples. **e,** BG4 CUT&Tag-qPCR for MCF7 cells treated with DMSO or 7.5 μM DC-34. BG4 accessibility was analyzed for *c-MYC* and normalized to three G4s in control gene sites (RPA3, MAZ, RBBP4). n=2, biological replicates, including three technical replicates for BG4 and one for the negative (no BG4 treatment). **f,** Volcano plot of differentially expressed genes in MCF7 cells treated with DC-34 vs non-treated. **g,** Volcano plot of differentially expressed genes in MCF7-dCas9-Halo cells transfected with $sgRNA_{MYC-19}$ and treated with Cl-PDS_2 vs MCF7-dCas9-Halo cells transfected with $sgRNA_{MYC-19}$ and treated with DMSO. Data presented are mean of n = number of independent biological samples. Statistical significance was calculated using a two-tailed t-test. P-value: ns > 0.05, * ≤0.05, ** ≤0.01, *** ≤0.001, **** ≤0.0001. Fold change (FC) was calculated on the normalized expression value.

Biological responses observed at specific G-quadruplexes are ligand-dependent

After validating ATENA's power in underpinning biological responses elicited by individual G4s, we extended its application beyond *c-MYC*. Previous reports indicated that PDS treatment of MCF7 led to significant upregulation of the long non-coding RNA *PVT1*, which we also confirmed for PyPDS by RT-qPCR and RNA-seq (Extended data Fig.4a, b) [34]. Since G4-stabilization by ligands is typically linked with transcriptional suppression, we decided to deploy ATENA and investigate if *PVT1* over-expression was an indirect consequence of global G4-stabilization induced by PyPDS or if it was a response reflecting the individual stabilization of the G4 present in the *PVT1* promoter. Previous CUT&Tag experiments performed on the MCF7 cell line identified a G4-peak close to the *PVT1* promoter [40], which we decided to target by designing sgRNAs to position dCas9 in proximity to the PVT1-G4 at either 20 base pairs from its 5' end ($sgRNA_{PVT1-20}$) or 33 base pairs from its 3' end ($sgRNA_{PVT1+33}$), ensuring no overlap with any regulatory region of the promoter (Fig.4a). We then transfected MCF7 expressing dCas9-Halo with $sgRNA_{PVT1-20}$ and $sgRNA_{PVT1+33}$ before treating with Cl-PDS_2, as described earlier. In agreement with previous reports, we observed a ~4-fold increase of *PVT1* expression when targeting the G4 in its promoter with dCas9-PyPDS (Fig.4b). Given that gene-suppression associated with G4-ligands treatment is elicited by the displacement of key regulatory proteins from the G4, we questioned whether PyPDS could instead increase protein binding to the PVT1-G4, thus leading to the observed transcriptional increase. To test this hypothesis, we have performed CUT&Tag and measured enrichment of the PVT1-G4 when targeted with dCas9-PyPDS over previously characterized G4s in MCF7 cells (MAZ, RPA3, RBBP4). Strikingly, we observed that PyPDS binding to the PVT1-G4 and $sgRNA_{PVT1-20}$ causes an increase of BG4-mediated enrichment of ~ 2-fold (Fig.4c), contrasting the effect that the same ligand elicits when targeting the MYC-G4 (Fig.2f). This observation suggests that the previously described increase in *PVT1* expression elicited by PyPDS treatment in MCF7 cells reflects the genuine response mediated by the PVT1-G4, highlighting how varied the response of G4s can be when exposed to the same ligand.

We next asked if the increase in *PVT1* expression observed with PyPDS was a ligand-dependent effect or if it was indicative of a more general biological response to treatment with any G4-ligand. Interestingly, when exposing MCF7 to PhenDC3 as a free ligand not tethered to dCas9, we observed *PVT1* suppression rather than transcriptional enhancement (Extended Data Fig.4c). We, therefore, questioned if the response to G4-binding in the *PVT1* promoter was genuinely ligand-dependent by guiding PhenDC3 selectively at this promoter using ATENA. To investigate this, we have used Cl-$PhenDC3_2$ in conjunction with $sgRNA_{PVT1-20}$ and $sgRNA_{PVT1+33}$ to assess PhenDC3-mediated *PVT1* expression changes upon selective PVT1-G4 targeting. Under these conditions, we observed a suppression of *PVT1* expression (Fig.4d), which is opposite to what was measured with PyPDS and in

line with what was observed using PhenDC3 on its own (Extended Data Fig.4c). This suggested that G4-ligands might elicit a different response when bound to the identical G4, possibly reflecting binding modalities that can either increase or prevent protein accessibility.

To further investigate ligand-mediated responses to individual G4-targeting, we have synthesized a HaloTag-compatible analog of a third G4-ligand: Pyrido Dicarboxamide (PDC, Fig.4e) [41]. More specifically, we have functionalized the PCD scaffold with a PEG-2 chloroalkane side chain (Cl-PDC_2) to mimic the PyPDS and PhenDC3 analogs used in ATENA (Fig.4e, see methods SI-1). We then used Cl-PDC_2 in conjunction with $sgRNA_{-20}$ to target the *PVT1* promoter. Similarly to what was observed with PhenDC3, PDC also caused suppression of *PVT1* with a fold change of ~ 4-fold (Fig.4e), suggesting that the way PDC interacts with the PVT1-G4 is reminiscent of PhenDC3. Indeed, the PCD scaffold is similar to PhenDC3 by displaying methylated nitrogens on the quinolines that are oriented oppositely compared to PDS (Fig. 4e). Moreover, both PhenDC3 and PDC lack the amino-side chains present in PyPDS, further highlighting the structural similarity between these two scaffolds. Our result indicated that the biological response elicited by ligands at individual G4s reflects the interactions between the ligand and the structure and, therefore, cannot be used as a reliable proxy to infer the endogenous function of a targeted G4 but rather should be used to infer ligand-specific responses at targeted sites.

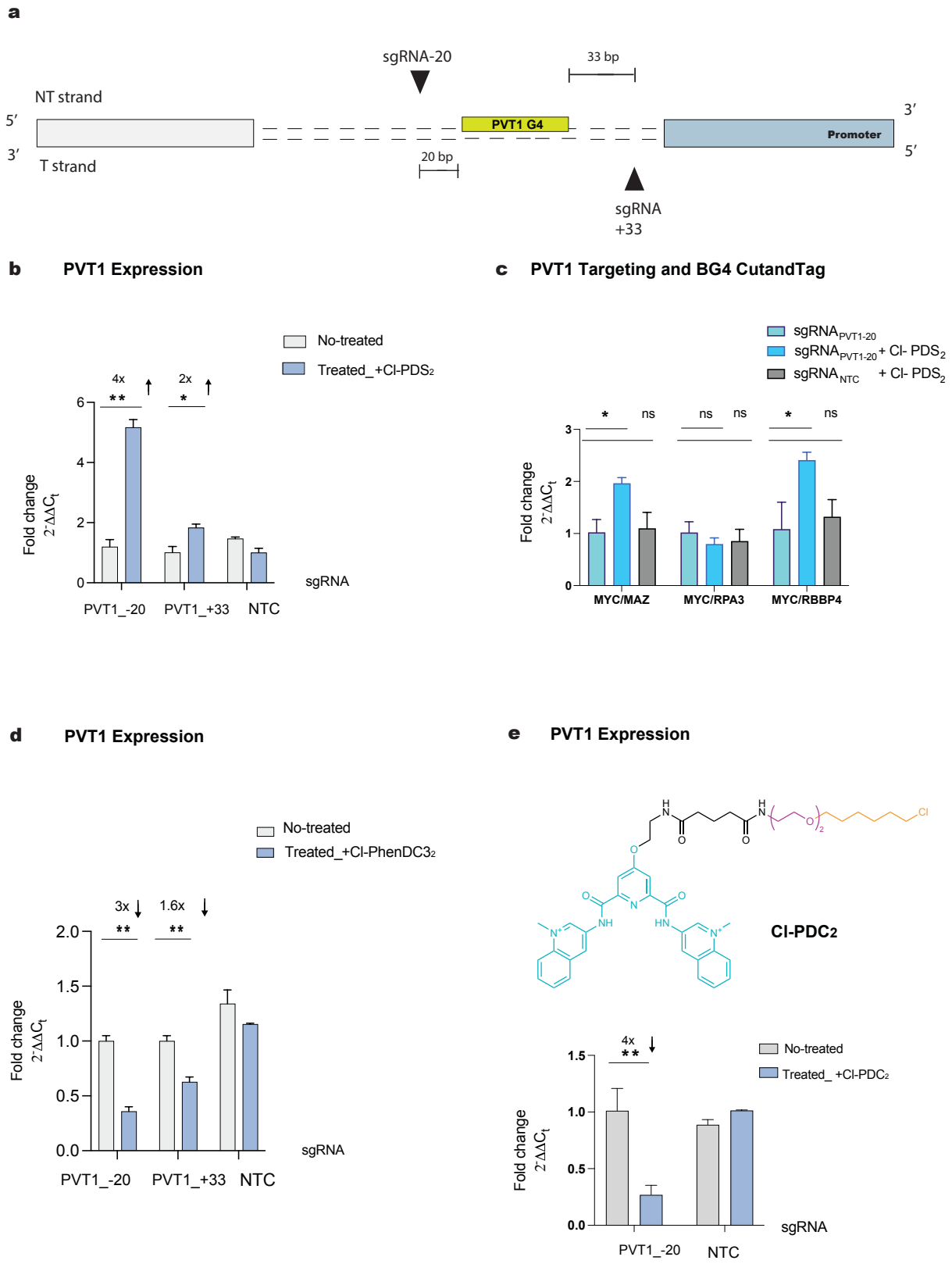

Figure 4: ATENA unveils ligand-dependent transcriptional response of the *lncPVT1*. a, Schematic overview of the *PVT1* promoter containing annotation of the predicted G4, sgRNA targeting region (black triangles) and their relative distance in bp from the G4 forming sequence **b,** RT-qPCR for *PVT1*

expression in MCF7 cells stably expressing dCas9-Halo, transfected with either $sgRNA_{PVT1-20}$,$sgRNA_{PVT1+33}$ or sgRNA NTC and treated with 2.5 μM of Cl-PDS_2 for 48h after transfection. The expression values are represented as fold change ($2^{-\Delta\Delta Ct}$) with respect to the non-treated transfected samples and after normalization for a housekeeping gene (GAPDH). n=3, biological replicates, each of which included two technical replicates. **c,** BG4 CUT&Tag-qPCR for MCF7 cells stably expressing dCas9-Halo transfected with either $sgRNA_{PVT1-20}$ or sgRNA NTC and treated with DMSO or 2.5 μM Cl-PDS_2. BG4 accessibility was analyzed for c-Myc and normalized to three G4s in control gene sites (RPA3, MAZ, RBBP4). n=2, biological replicates, including three technical replicates for BG4 and one for the negative (no BG4 treatment). **d**, RT-qPCR for *PVT1* expression in MCF7 cells stably expressing dCas9-Halo, transfected with either $sgRNA_{PVT1-20}$, $sgRNA_{PVT1+33}$ or sgRNA NTC and treated with 2.5 μM of Cl-$PhenDC3_2$ for 48h after transfection. Values are represented as fold change ($2^{-\Delta\Delta Ct}$) with respect to the DMSO samples and after normalization for a housekeeping gene (GAPDH). n=3, biological replicates, each of which includes two technical replicates. **e,** (Top) Chemical structure of chloroalkane-modified PDC (Cl-PDC_2) with two PEG linkers length. **e**, (bottom) RT-qPCR for *PVT1* expression in MCF7 cells stably expressing dCas9-Halo, transfected with either $sgRNA_{PVT1-20}$ or sgRNA NTC and treated with 2.5 μM of Cl-PDC_2 for 48h after transfection. Values are represented as fold change ($2^{-\Delta\Delta Ct}$) with respect to the DMSO samples and after normalization for a housekeeping gene (GAPDH). n=3, biological replicates each of which includes two technical replicates. Data presented are mean of n = number of independent biological samples. Statistical significance was calculated using a two-tailed t-test. P-value: ns > 0.05, * ≤0.05, ** ≤0.01, *** ≤0.001, **** ≤0.0001. Fold change (FC) was calculated on the normalized expression values.

Targeting cell-specific G4s with ATENA reveals transcriptional-dependent response to ligands.
CUT&Tag and other chromatin-compatible G4-mapping methods, like BG4-ChIP and Chem-Map, have revealed that G4 distribution in the genome is cell-dependent and typically prevalent at promoters of highly transcribed genes [7, 9, 10, 42]. Therefore, we decided to investigate biological responses attained when directing a ligand to previously unexplored G4-structures specifically detected in MCF7 cells. In particular, we wanted to investigate if the biological relevance of individual G4s and their response to ligand binding was linked to the transcriptional levels of the associated genes. To this end, we set out to apply ATENA to study *de novo* cell-specific promoter G4s, encouraged by the demonstrated ability of our method to target individual G4s within the genome of living cells. We leveraged the existing dataset on G4-distribution in MCF7 cells previously obtained with CUT&Tag [40]. This study identified a G4-peak in the promoter of the highly expressed *HMGN1* gene as an exclusive marker of MCF7 cells compared to other cell lines. *HMGN1* encodes for a non-histone chromosomal protein able to interact with nucleosomes and regulate chromatin structure [43]. We, therefore, questioned if we could leverage ATENA to target and stabilize the G4 in the *HMGN1* promoter selectively to modulate its transcriptional level in a G4-dependent fashion. To test this, we have designed $sgRNA_{HMGN1-22}$ and $sgRNA_{HMGN1+34}$ to target the HMGN1-G4 at 22 and 34 base pairs, respectively, at its 3' and 5' ends, ensuring no overlap with functional regions (Fig. 5a). After incorporation of $sgRNA_{HMGN1-22}$ and $sgRNA_{HMGN1+34}$ into an expressing vector, we have independently transfected MCF7 cells with both sgRNAs, before adding Cl-PDS_2 and incubating for further 48 hours, as per the optimized ATENA protocol. Strikingly, we have observed an extremely strong reduction of *HMGN1* expression of 76-fold when targeting its G4 at the closest distance of 22 base pairs with $sgRNA_{HMGN-22}$ (Fig.5b). *HMGN1* downregulation was partially attenuated when placing dCas9-PyPDS further away from the G4 with $sgRNA_{HMGN+34}$, consistent with the distance-dependent ligand engagement observed for other G4s (Fig.5b). This data indicated that a G4 that is detected specifically in MCF7 cells (*HMGN1*) could be targeted individually with ATENA to suppress its transcriptional levels up to the higher extent observed by G4-targeting in the context of this study. This indicates that maintaining G4-homeostasis at the promoter of highly transcribed genes is

key to sustaining elevated expression levels characteristic of specific cell lines, making these G4s particularly sensitive to G4 ligands. Altogether, our observations suggest that the varied transcriptional response observed upon G4-ligands treatment might reflect the relative relevance of different G4s in maintaining transcriptional homeostasis of the specific cell line studied.

To corroborate this hypothesis, we used ATENA to target a G4 present in the promoter region of a gene expressed at very low levels in MCF7 cells. To this end, we have noticed that CUT&Tag G4-mapping performed in MCF7 cells revealed a distinct G4-peak in the promoter region of *IL17RA* [40], a gene that is only marginally expressed in this cell line. *IL17RA* encodes for the interleukin 17A (IL17RA), a proinflammatory cytokine secreted by activated T-lymphocytes and, therefore, not essential for breast cancer cell homeostasis. We generated $sgRNA_{IL17RA-20}$ to target the IL17RA-G4 at 20 base pairs from its 5' end and within a region that does not overlap with other regulatory elements of this promoter. After transfection with $sgRNA_{IL17RA-20}$ and $sgRNA_{NTC}$ followed by incubation with Cl-PDS_2, we failed to detect any measurable changes in *IL17RA* expression levels (Fig.5c). This indicates that targeting a G4 in a promoter of a gene that is not transcriptionally active is not associated with gene-expression perturbation, linking tightly the functional relevance of G4s to the transcriptional levels of the genes associated. Considering that both G4s in the *HMNG1* and *IL17RA* promoters are equally detected in MCF7 by CUT&Tag [40] and targeted with similar sgRNA designs (within the limits imposed by the PAM availability), we reasoned that the transcriptional levels linked to the targeted G4 can be used to anticipate the extent of transcriptional perturbation associated with ligand treatment. This is also illustrated by the relatively modest transcriptional changes observed by the selective targeting of the *c-MYC* and *PVT1* G4s, which are only moderately expressed in MCF7 cells. Increasing evidence in the literature suggests a model by which G4s act as epigenetic factors to mark highly transcribed genes [12]. Our data support this model by showcasing how the extent of gene suppression/activation elicited by G4 ligands is linked to basal transcriptional levels, highlighting the relevance of maintaining G4 homeostasis to sustain elevated transcription characteristics of specific cell lines.

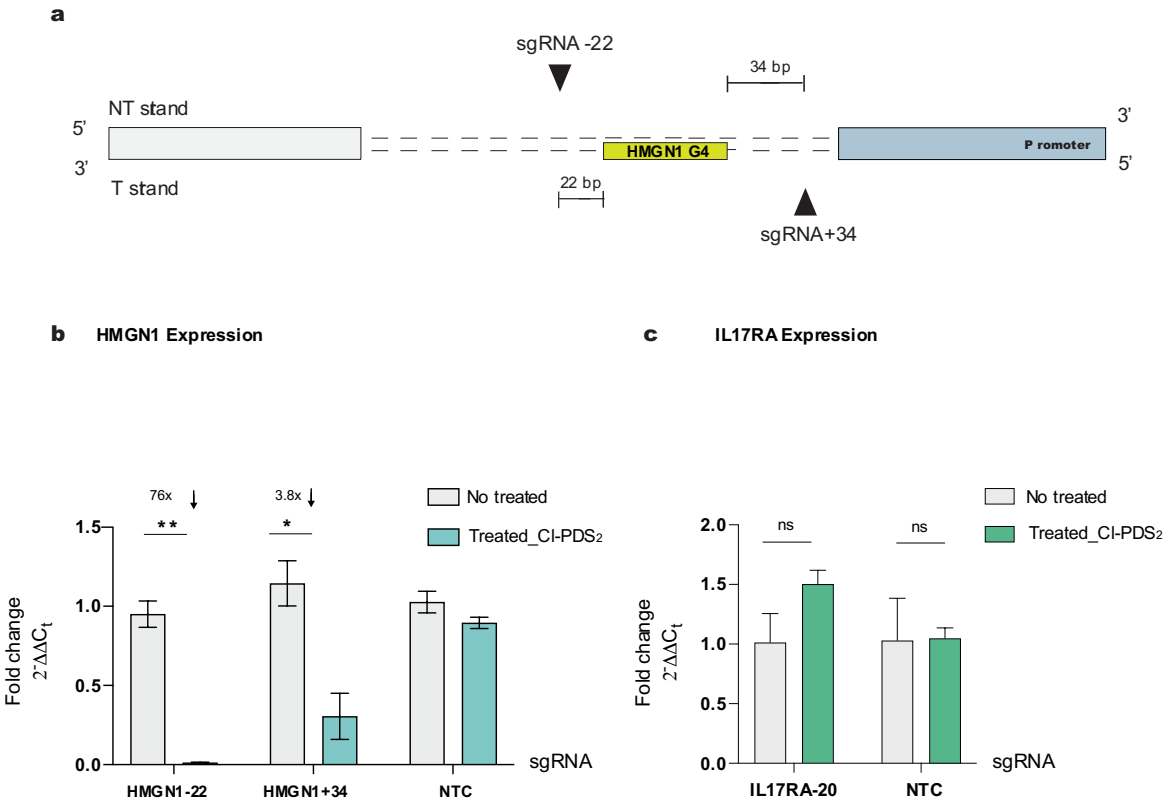

Figure 5: Targeting of *de novo* G4s with ATENA uncovers a transcriptional-dependent functional response. A schematic overview of the *HMGN1* promoter contains annotation of the predicted G4, two sgRNA designed to target HMGN1 G4 (complementary) (black triangles), and their relative distance in bp from the G4 forming sequence. **b,** RT-qPCR for *HMGN1* expression in MCF7 cells stably expressing dCas9-Halo, transfected with either $sgRNA_{HMGN1-22}$, $sgRNA_{HMGN1+34}$ or sgRNA NTC and treated with 2.5 µM of Cl-PDS_2 for 48h after transfection. Values are represented as fold change ($2^{-\Delta\Delta Ct}$). The expression values are represented as fold change ($2^{-\Delta\Delta Ct}$) with respect to the non-treated transfected samples and after normalization for a housekeeping gene (GAPDH). n=2, biological replicates, each of which includes two technical replicates. **c**, RT-qPCR for *IL17RA* expression in MCF7 cells stably expressing dCas9-Halo, transfected with either $sgRNA_{IL17RA-20}$ or sgRNA NTC and treated with 2.5 µM of Cl-PDS_2 for 48h after transfection. Values are represented as fold change ($2^{-\Delta\Delta Ct}$). The expression values are represented as fold change ($2^{-\Delta\Delta Ct}$) with respect to the non-treated transfected samples and after normalization for a housekeeping gene (GAPDH) n=2, biological replicates, each of which includes two technical replicates. Data presented are mean of n = number of independent biological samples. Statistical significance was calculated using a two-tailed t-test. P-value: ns > 0.05, * ≤0.05, ** ≤0.01, *** ≤0.001, **** ≤0.0001. Fold change (FC) was calculated on the normalized expression value.

Discussion

There is now substantial evidence to support that G4 structures are formed within endogenous chromatin and that their formation is strongly linked to active transcription [12]. For instance, G4s have been detected in the promoter regions of key oncogenes in cancer cells, and it has been recently demonstrated that they represent a key structural feature necessary to sustain elevated transcription of oncogenes, such as c-*MYC* [38]. Moreover, the development of selective G4 ligands over the past decades has revealed how G4-targeting by small molecules is typically associated with transcriptional suppression of such oncogenes [13]. For these reasons, G4s have represented a therapeutically appealing target for decades. However, using G4-ligands for clinical application has not yet gained traction, reflecting two intrinsic limitations. Firstly, the recognition mechanism leveraged by most G4 ligands relies on end-stacking interactions that are promiscuous, generating molecules that are excellent at discriminating G4s from duplex DNA but unable to discriminate amongst different G4s. Since G4 structures are highly abundant in the human genome and their formation is highly cell-dependent, this lack of intra-G4 selectivity leads to transcriptional perturbation of multiple genes simultaneously and to different responses depending on the cell line used. Secondly, G4-ligands might display differences in their relative binding affinity to different G4s, ultimately leading to the emergence of substantially different phenotypes when applied to cells and model organisms, making it hard to underpin distinctive G4 phenotypes.

For these reasons, tools to enable single-G4 targeting have been considered essential to unravel the fundamental biology regulated by specific G4s and to validate their therapeutical relevance. In this study, we have developed an innovative tool based on the chemical functionalization of dCas9 proteins with G4-ligands that allows the positioning of ligands in the proximity of an individual G4 of interest using short-guiding RNAs (ATENA). We leveraged ATENA to demonstrate that the positioning of a single G4-ligand is sufficient to observe transcriptional perturbation that can be confidently ascribed to G4-targeting. We have achieved this by extensive biochemical optimization of ATENA to ensure that ligands were covalently linked to the dCas9 protein under physiological conditions (using HaloTag) and optimized ideal ligand properties and sgRNA design to attain optimal G4-targeting by the dCas9-ligand conjugates *in vitro.*

With optimized conditions to attain single-G4 targeting, we have used ATENA to investigate the functional role of the G4-structure present in the promoter region of the proto-oncogene *c-MYC*. Many independent studies have previously reported that *c-MYC* suppression can be achieved upon treatment with G4-ligands [44]. However, it is less clear to what extent transcriptional suppression elicited by G4-ligand treatment directly results from MYC-G4 stabilization or is driven by global G4-targeting caused by the ligand used. Indeed, Hurley and co-workers have challenged the model they proposed in their seminal paper where *c-MYC* downregulation was exclusively attributed to the targeting of MYC-G4 [38], suggesting that transcriptional suppression was more likely a response to global G4-stabilization [17]. Using ATENA, we observed that G4-mediated *c-MYC* suppression in MCF7 cells is minimal and limited to P1-controlled transcription. These findings are not only in line with what has been shown by recent genomic studies [38] but were also fully recapitulated when using the MYC-G4 selective small molecule DC-34 [20]. This strongly indicates that the biological response obtained by selective MYC-G4 targeting is limited to P1-mediated transcription, irrespective of how selective MYC-G4 perturbation is achieved.

Importantly, we have noted that the MYC-G4 is located close to key regulatory regions of the promoter, including the P1-promoter and a TATA-box sequence. Therefore, using CRISPR-based tools for

selective MYC-G4 targeting can easily lead to misleading results when using sgRNAs targeting those regions and placing small molecules near key regulatory elements. A recent report published while this manuscript was in preparation has shown global *c-MYC* global downregulation when using dCas9-Nucleolin fusion or dCas9 poly-labeled with 10 G4-ligands in tandem [45]. However, this was observed when using the equivalent of our $sgRNA_{MYC+58}$ that targets a region 9-bp apart from the TATA-box sequence and causes unspecific *c-MYC* downregulation even when dCas9 is functionalized with the non-G4-binding molecule Oregon Green (Fig. 2d). This indicated that placing small-molecules in proximity of regulatory regions of promoters can lead to transcriptional changes that are independent of G4-stabilization. Therefore, our data strongly suggest that the use of multiple ligands on a single dCas9 protein to attain G4-stabilization should be avoided, as it might lead to false positives by creating local overcrowding of small molecules at promoters and elicit responses that cannot be ascribed solely to G4-targeting [45]. Similarly, using multiple sgRNA on a single gene-promoter is likely to perturb the homeostasis of regulatory elements in a ligand-dependent fashion but without necessarily reflecting changes associated with G4-targeting and, therefore, should be avoided [45]. Nevertheless, this showcases how the use of orthogonal chemistries to label dCas9 with ligands (biotin-streptavidin) can also be leveraged to attain single G4-targeting, validating ATENA and expanding the scope of both these strategies for potential multi-ligand targeting.

Using the selective G4-antibody BG4, we could also show that P1-specific *c-MYC* downregulation observed with ATENA is associated with reduced protein accessibility at the G4-structure, which is likely responsible for the observed transcriptional suppression. Importantly, a similar response can also be observed using DC-34, strongly suggesting that a reduction of BG4 binding, rather than an increase [45], should be observed as a consequence of G4 engagement at the *c-MYC* promoter. This allowed us to conclude that the selective targeting of the G4 in the *c-MYC* promoter can be leveraged therapeutically to suppress its expression by preventing protein binding to the G4 structure. However, this approach will depend highly on how much global *c-MYC* expression is driven by the P1-promoter in the model system investigated [20], which should be considered for therapeutic applications of G4-ligands. This is exemplified by the strong suppression of global *c-MYC* expression observed upon treatment with DC-34 in multiple myeloma cells, which is in line with the strong P1-driven c-MYC expression that is characteristic of certain myelomas [46, 47].

The development of ATENA has also enabled us to characterize ligand-dependent responses associated with targeting the same G4-structure. It has often been reported that G4s in different promoters can display varied responses to G4-ligands, leading to transcriptional increase or suppression of the G4-associated gene on a case-by-case basis. This has often been wrongly attributed to the endogenous biological function of the specific G4 investigated, which was thought to either stimulate or repress transcription depending on its genomic context. However, this model strongly contradicts genomics studies that clearly indicate a global association of G4-formation to transcriptional activation rather than suggesting a context-dependent function of these structures. Using ATENA in combination with CUT&Tag, we could demonstrate that the response variation observed at certain G4s reflects the different response to protein interacting with the targeted G4-site caused by the binding of structurally different ligands. We demonstrated this concept by characterizing the different transcriptional responses observed upon targeting the G4 in the promoter of the long non-coding RNA *PVT1* by using two established G4 ligands: PyPDS and PhenDC3. The previously reported increase in *PVT1* transcription observed upon PyPDS stabilization was recapitulated by individual targeting with ATENA, which was also associated with increased BG4-binding at this G4-site measured by CUT&Tag qPCR, indicating enhanced protein binding elicited by PDS binding. Conversely, PhenDC3 treatment was associated with the more commonly observed transcriptional suppression, demonstrating that structurally different

ligands elicit different responses at the same G4 site. Importantly, we could observe the same ligand-dependent response also when using free PyPDS or PhenDC3 not bound to dCas9, suggesting that the lack of intra-G4 selectivity does not necessarily prevent these molecules from providing meaningful information on the response that they elicit at a specific G4-site. However, our observations indicated that while G4-ligands are useful tools to perturb G4-homeostasis and investigate biological responses associated with the targeting of these structures, they cannot be used to infer the endogenous biological function regulated by G4s directly but should rather be used to gain insights on the responses triggered by their binding to these structures.

Finally, we used ATENA to target *de novo* G4s previously detected in MCF7 cells with CUT&Tag but never investigated for their biological functions. Since G4-formation is largely associated with active transcription, we reasoned that the biological response to G4-ligand treatment would be somewhat linked to the transcriptional levels of the gene targeted. We took advantage of ATENA to guide PyPDS to a G4-promoter of either a highly expressed (*HMGN1*) or a low-expressed gene (*IL17RA*). Importantly the G4 in the promoter of *HMGN1* is only detected in MCF7 cells but not in other cancer cell lines [40], suggesting that this G4 is somehow important to breast cancer cells to sustain elevated expression of *HMGN1*. Single targeting of the G4 in the *HMGN1* promoter by ATENA led to the higher transcriptional suppression measured in MCF7 (76-fold), indicating that G4-ligand response is not only dependent on the type of G4-targeted but also on the transcriptional levels of the gene investigated. This is particularly relevant considering that the G4 in the *HMGN1* promoter has been previously characterized as a distinctive feature of MCF7 cells, suggesting that highly expressed cell-specific G4-containing genes are more sensitive to G4-ligands treatment. This is also in line with our recent findings demonstrating how treatment with PDS in chemo-resistant ovarian cancer cells leads to re-sensitization to chemotherapy, given the high enrichment of G4s in highly transcriptionally active genes that are essential to establish a chemo-resistant state, such as *WNT* [48]. This notion is further corroborated by the lack of transcriptional changes observed when targeting G4s previously detected in MCF7 but not associated with actively transcribed genes (i.e., *IL17RA*). Altogether, this evidence strongly supports a model by which the biological significance of G4 structures within the genome is tightly linked to the epigenetic landscape of a given cell line, as previously suggested by independent genomic studies.

In conclusion, ATENA has enabled single-G4 targeting within the genome of living cells by distinct G4-ligands. This approach allowed us to compare the different biological responses elicited by different ligands at the same G4 site and the responses of various G4s when targeted by the same ligand with high precision. Altogether, we could infer that the chemical nature of the ligand used can differently perturb the local protein-binding homeostasis of promoter G4s, leading to either increased protein accessibility and activation of transcription or decreased accessibility and transcriptional suppression. Notably, this can be measured by BG4-ChIP or CUT&Tag, providing a direct measure to anticipate the response to the ligand binding. This can be valuable for developing therapeutic agents based on G4-targeting, which can be tailored chemically to either diminish or amplify transcription based on their binding modalities. Moreover, ATENA enabled us to disentangle local effects triggered by individual G4-targeting from responses driven by global G4-stabilization elicited by canonical G4-ligands. Interestingly, we concluded that G4-ligands lacking intra-G4 selectivity generally predict the response elicited by selective targeting of individual G4s and can be used to infer therapeutic potential of individual structures. Indeed, we observed that the main indicator of the response to G4-ligand treatment is the intrinsic level of transcription of the promoter targeted, making ligand response highly epigenetic and cell-dependent.

We envisage that further development of ATENA can be leveraged to generate genome-wide screening of G4-functional response to various ligands in different cell lines. Moreover, this technology can be used to generate Halo-functional ligands to screen their ability to perturb protein accessibility at a given G4 site that can be useful for therapeutic applications. We also anticipate that given the minimal perturbation of ligands required for Halo-functionalization (chloroalkane), ATENA can be further extended for gene-selective localization of other DNA-interacting molecules, such as Topoisomerase inhibitors or cross-linking agents.

Methods

Comprehensive synthetic protocols and purification methods for Cl-PDS$_0$, Cl-PDS$_2$, Cl-PDS$_4$ and Cl-PhenDC3$_2$, Cl-PhenDC3$_4$, and Cl-PDC$_2$ ligands are provided in the Supplementary Information (SI-1).

FRET-melting assay

Before each experiment, a solution containing 20 µM DNA ordered from Integrated DNA Technologies (IDT) (Table.S1), labeled at 5'- with FAM and at 3'- with TAMRA, was freshly prepared in 10 mM lithium cacodylate buffer, pH 7.4 supplemented with 20 mM KCl. For the non-G4-forming sequence dsDNA26 (26 mer), the solution was supplemented with 150 mM KCl. The DNA solution was placed on an Eppendorf Thermomixer, annealed by heating at 95°C for 10 minutes, and left cooling down at 0.5°C/min. The DNA was aliquoted in a 96-well RT-PCR plate (Thermo Fischer Scientific) and supplemented with increasing equivalent amounts of each ligand at a final oligonucleotide concentration of 0.4 µM. Fluorescence readings were performed on an Agilent Stratagene MX3000P, in the range of 25-95°C, in technical duplicate. The data points obtained were analyzed with GraphPad Prism 10, and the melting temperature was extrapolated from a Bolzmann Sigmoidal function.

dCas9-Halo purification

The expression vector for dCas9-Halo (Addgene #72269) was introduced into BL21 cells (D3,NEB-C2527H) by heat shock. Bacteria were then plated in LB agar plates and selected with Ampicillin (100 ug/ml). One bacteria colony was picked and grown overnight in a 50 ml flask at 37°C. Overnight precultures were diluted 1:100 into the main culture (LB media, Ampicillin 100 ug/ml) and incubated at 37°C, 200 rpm until an OD600~0.6. The expression of the recombinant dCas9-Halo was induced by using 1 mM of isopropyl β-D-1-thiogalactopyranoside (IPTG). Upon induction, the temperature was reduced to 16°C for dCas9-Halo expression. BL21-dCas9-Halo cells were harvested after overnight incubation at 16°C by centrifugation (3500 g, 20 min, 4°C). Cells were lysed in lysis buffer (50 mM sodium phosphate, 300 mM NaCl pH 7.2, 20% glycerol, 0.1% Tween-20, 1 mM TCEP, 1 mg/ml lysozyme, 1 mM PMSF, 1 mM $MgCl_2$ and 5U/g of culture Benzonase). Cell lysates were then homogenized, and cell debris was removed by centrifugation (16000xg, 15 min, 4°C). The soluble fraction of the lysate was filtered and then loaded onto a cobalt column, bound protein was eluted with 150 mM imidazole. After imidazole removal and concentration by using a 50,000-MWCO centrifugal filter (Millipore Amicon), fractions were loaded into 1-ml HiTrap SP XL (GE Healthcare), ÄKTA pure 25 for cationic exchange chromatography, and elution was performed with a gradient from 0% to 70% of Buffer (50 mM HEPES pH 7.5, 1 M KCl, 1mM TCEP). Purest fractions were pooled and then used for gel filtration chromatography. Gel filtration was performed using a Superdex 200 10/300 GL column, and isocratic elution was performed in a buffer containing 50 mM HEPES (pH 7.5), 150 mM

KCl, and 1 mM TCEP. Protein aliquots in storing buffer (50 mM HEPES-pH 7.5, 150 mM KCl, 20%glycerol) were snap frozen and stored at -80°C.

BG4 purification

BL21 *E. coli* competent cells (D3,NEB-C2527H) were transformed with a pSANG10-3F-BG4 plasmid (Addgene #55756). Single colonies were inoculated in 2xYT medium supplemented with 2% glucose and 50 µg/ml kanamycin and incubated at 30 °C at 250 rpm overnight. The starting culture was diluted 1:500 into ZYM-5052 autoinduction medium prepared following the previously described method (https://doi.org/10.1016/j.pep.2005.01.016) supplemented with 50 µg/ml kanamycin. The culture was then incubated at 37 °C for 6 hours at 250 rpm, followed by an overnight incubation at 25 °C at 280 rpm. The culture was centrifuged at 4,000 x *g*, 4 °C for 30 min, the collected pellet lysed in 80 ml ice-cold TES buffer (50 mM Tris-HCl pH 8.0, 1 mM EDTA pH 8.0, 20% sucrose) and incubated on ice for 10 minutes. This was followed by 15 incubation in ice-cold TES (diluted 1:5), supplemented with 15U/mL benzonase, 2 mM MgSO4, and a protease inhibitor cocktail. The lysate was centrifuged at 16,000 x g for 20 minutes. The supernatant was collected and incubated with Nickel resin (Sigma, P6611), previously equilibrated in PBS, at 4 °C for 1 hour in rotation. The complex was washed three times with wash buffer (PBS, 100 mM NaCl, 10 mM imidazole), then BG4 was eluted in 4 mL elution buffer (PBS pH 8.0, 250 mM imidazole) and dialyzed against PBS at 4 °C using GeBaflex tubes (Generon). Purified BG4 was snap-frozen in liquid nitrogen and stored at -80 °C.

dCas9-binding competition assay

dCas9-Halo purified protein (4 µM) was incubated with several dilutions of Cl-PDS$_n$ or Cl-PhenDC3$_n$ probes or with commercially available Halo-TMR for the positive control (Promega, G8252) for 30 minutes at room temperature in PBS. Pretreated dCas9-Halo proteins were then incubated with 5 µM Halo-TMR, or Cl-PDS$_n$ for the positive control, and samples were loaded in an SDS-PAGE gel (8% polyacrylamide) and run at 180V for 45 minutes. Gels were then imaged using Typhoon FLA 9500 (GE) using the TAMRA filter set (542 nm).

In vitro transcription for sgRNA synthesis

Oligonucleotides containing sgRNA sequences (Table.S2) were ordered from Integrated DNA Technologies (IDT), and a PCR was set up to generate the corresponding duplex following the previously described method (https://currentprotocols.onlinelibrary.wiley.com/doi/10.1002/cpmb.43). Following manufacturer instructions, the DNA product was transcribed using the HiScribe T7 High Yield RNA Synthesis Kit (NEB, E2040S). The resulting sgRNA was purified using the RNeasy Mini kit (74104) and stored at -80°C.

dCas9-Halo_FRET assay

Before the experiment, Cy5-3' end labeled and Cy3-5' end labeled oligos (Table S1) were annealed in 10 mM Tris HCl pH 7.5, 100 mM KCl at a final concentration of 1 µM (95 °C for 10 minutes and left for overnight cooling down at room temperature). dCas9-Halo purified protein at a final concentration of 4 µM was incubated with Cl-PDS$_n$ probes in a ratio of 1:2.5 for 45 minutes in binding buffer (20 mM HEPES, pH 7.5, 100 mM KCl, 1 mM MgCl2,1 mM TCEP, 10% glycerol). The pretreated dCas9-Halo was then incubated with 120 pmol of purified sgRNA for 20 minutes at room temperature, followed by incubation with 200 nM of pre-annealed oligos at 37°C for 1h. The complexes were then loaded into

an 8% polyacrylamide native gel and ran for 120 minutes at 4°C in 1x TBE supplemented with 2 mM MgCl2. Gels were then imaged using Typhoon FLA 9500 (GE) using the Cy3 and Cy5 filter set.

Cloning of lentiviral dCas9-Halo construct

The lentiviral dCas9-Halo plasmid was cloned using Gibson assembly. A lentiviral backbone containing already dCas9 (Addgene #61425) was digested using BamHI and BsrGI (New England Biolabs, R3575S, R3136S) and then assembled with the HaloTag sequence amplified (Table.S3) from pET302-6His-dCas9-Halo (Addgene #72269) in a 1:3 molar ratio of backbone: insert using HiFi DNA Assembly Mix (E2621S, New England Bioscience) following manufacturing protocols. Post incubation, assembled products were diluted with water, and 5 µL of the product was transformed by heat shock into 10-beta competent cells (New England Biolabs, C3019I). Cells were then plated on agarose plates (supplemented with ampicillin 100 µg/mL) for overnight outgrowth at 37 °C. Single clones were picked and grown in 5 mL of Amp LB media overnight at 37 °C while shaking. Plasmid DNA was purified from cells using the Promega PureYield plasmid miniprep system (Promega, A1223) and sequenced using the Geneweiz service.
The positive plasmid was further modified by substituting blasticidine resistance with mCherry coding sequence using Gibson assembly after digestion with BsrGI and EcoRI enzymes (Table.S3). Positive clones were then sequenced using the Geneweiz service.

Cloning of sgRNA in pLG1 plasmid

pLG1 backbone (Addgene #109003) was digested with BstXI and BlpI (FastDigest, ThermoFisher) for 1h at 37°C. Oligos containing the sgRNA sequence (Table S4) were ordered from IDT and annealed in water at 10 µM in a thermocycler (37°C, 30 minutes- 95°C, 5 minutes and ramp down 5 degrees/minute to 25°C). The annealed oligos were diluted (1:50) and then assembled with the digested plasmid in a 1:2 molar ratio of backbone: insert using HiFi DNA Assembly Mix (E2621S, New England Bioscience) following manufacturing protocols. Post incubation, assembled products were diluted ½ in water, and 5 µL of the product was transformed by heat shock into 10-beta competent cells (New England Biolabs, C3019I). Cells were then plated on agarose plates (supplemented with ampicillin 100 µg/mL) for overnight outgrowth at 37 °C. Single clones were picked and grown into 5 mL of Amp LB media overnight in a 37 °C rotating shaker. Plasmid DNA was purified from cells using the Promega PureYield plasmid miniprep system (Promega, A1223) and sequenced using the Geneweiz service.

Mammalian cell culture

HEK293FT cells were cultured in DMEM GlutaMAX™ (Thermo Fisher Scientific, 10566016) supplemented with 1x penicillin-streptomycin (Thermo Fisher Scientific, 15140122) and 10% (vol/vol) fetal bovine ser (FBS) and maintained at 37 °C and 5% CO2.
MCF7 cells were cultured in DMEM GlutaMAX™ (Thermo Fisher Scientific, 10566016) supplemented with 1x β-Estradiol (Sigma, 50-28-2), 1x penicillin-streptomycin (Thermo Fisher Scientific, 15140122) and 10% (vol/vol) fetal bovine serum (FBS) and maintained at 37 °C and 5% CO2.

Lentivirus production and transduction

For lentiviral production, HEK293T cells were seeded at $3.8x10^6$ in a 10 cm tissue culture plate the day before transfection. 8.4 ug of the envelope plasmids pCMV-VSV-G (Addgene #8454) and 6.4 ug of packaging plasmid R8.74 (Addgene #22036) were co-transfected along with 2.1 ug of the target plasmid

(dCas9-Halo-T2A-mCherry) using polyethyleneimine (PEI-Linear, MW 25000) in a ug DNA: ug PEI ratio of 1:3. Viral supernatant was harvested after 48h and 72h, and before usage, it was filtered using a 0.45 µm filter unit. MCF7 cells were plated on 6-well plates the day before infection. They were infected with lentiviruses in DMEM 10% FBS, β-Estradiol, in the presence of polybrene with a final concentration of 10 µg/mL. Half of the media was changed the day after, and cells were grown for one week and examined by flow cytometry (Attune NxT, ThermoFisher) to confirm successful transduction (YL2-Channel). After genotyping (Table.S3), cells were single-cell sorted using FACS BD FACS Diva 9.0.1 (Fig.SI-3).

Western blot

MCF7 stably expressing dCas9-Halo cells were harvested and processed using 1x RIPA buffer (Merck,20-188) supplemented with 1x Protein inhibitor cocktail. The total protein concentration of the cleared lysate was then measured by Bradford assay (Pierce™ Bradford Plus Protein Assay Kits, Thermo Fisher Scientific, 23236) at 595 nm using a ClarioStar plate reader. 20/25 ug total protein was then loaded into an 8% polyacrylamide SDS-Page gel and then transferred to a pre-assembled PVDF membrane (Trans-Blot Turbo Mini 0.2 µm PVDF Transfer Packs, BioRad) using a semi-dry method (BioRad). The membrane was blocked (5% milk in TBS-T) and incubated with primary antibodies (1:1000 Cas9 antibody- 14697T, 1:1000 GAPDH antibody- 2118T, Cell Signaling Technology) overnight at 4 degrees. The membrane was then washed in TBS-T and incubated with secondary antibodies (1:10,0000) goat anti-rabbit/mouse HRP- Advansta, R-05071-500, R-05072-500) for 1h at room temperature. After three washing steps of the membrane in TBS-T, 0.5 ml of HRP substrate (Merck, WBLUC0100) was added to the top of the membrane, and the excess was removed. The signal was developed using Image Quant LAS 4000 (Cytiva) with the following setting: chemiluminescence for the WB signal and Cy5 for visualizing the protein marker.

CAPA assay

MCF7 cells stably expressing dCas9-Halo were plated the day before at 30x10^3 cells/ well in 96-well plates pre-coated with Poly-D-Lysine (Thermo Fisher Scientific, A3890401). Cells were incubated for 2h in the presence of serial dilutions of Cl-PDS$_n$ or Cl-PhenDC3$_n$ probes ranging from 0.2µM to 10 µM (5% CO2, 37 °C). Cells were then washed two times with media, and every washing step included 10 minutes of incubation (5% CO2, 37 °C). Cells were then incubated with HaloTag® Oregon Green® Ligand (Promega, G2801) for 45 minutes, followed by two washing steps (10 minutes,5% CO2, 37 °C). Cells were washed with DPBS, treated with 0.25% trypsin, and resuspended in FACS buffer (5%BSA in DPBS) before flow cytometry analysis (Attune Nxt).

Transfection and incubation with G4 ligands

MCF7 cells stably expressing dCas9-Halo were plated the day before at 25x10^3 cells/ well in 96-well plates pre-coated with Poly-D-Lysine (Thermo Fisher Scientific, A3890401). According to manufacturer protocol, 50 ng of guide expressing plasmid was transfected using Lipofectamine 3000 (Thermo Fisher Scientific, L3000001). The day after, cells were incubated with either DMSO (0.5%) or Cl-PDS$_n$ or Cl-PhenDC3$_n$ probes and incubated for 48h.

RNA isolation

RNA was harvested 72 hours post-transfection. Cells were washed with 100 µL of 1X DPBS (Gibco, 14190144) and incubated with RLT buffer from RNeasy mini kits (Qiagen, 74104) according to the manufacturer's instructions. Following the manufacturer's instructions, the eluted RNA was reverse transcribed to cDNA using the RevertAid cDNA prep kit (K1621, Thermo FisherScientific).

Quantitative Real-Time PCR (RT-qPCR)

RT-qPCR reactions were prepared as follows: cDNA was mixed with 500 nM of forward and reverse primer (Table.S5) and 5µL of Fast SYBR Green Master Mix (Applied Biosystems: 4385612) in a final volume of 10 µL. RT-qPCR was performed on a 96-well plate format using an Agilent Stratagene Mx3000P machine with the following program: 95 °C for 20 seconds, 40× (95 °C for 3 s, 60 °C for 30 s), 95 °C for 15 s, 60 °C for 1 min, and 95 °C for 15 s (melting curve). The C_t values obtained were used to calculate the relative fold-change in gene expression using the $\Delta\Delta C_t$ method and normalized to the housekeeping gene *GAPDH* in both treated and untreated samples.

Cut&Tag

Cells were plated in a 6-well plate format one day before transfection following the previously described transfection protocol. 24h after transfection, cells were incubated with Cl-PDS_n probes for 6h and then harvested, counted, and inspected for viability. To prepare nuclei, 200k cells per reaction (plus 10% excess) were centrifuged for 3 min at 600×*g* at room temperature and then resuspended in nuclei extraction buffer (20 mM HEPES–KOH, pH 7.5, 10 mM KCl, 0.1% Triton X-100, 20 mM Glycerol, 0.5 mM Spermidine, 1X Roche cOmplete Mini EDTA-free Protease Inhibitor (Roche, 11836170001) and incubated for 10 minutes on ice. Nuclei were spined at 600 x g for 3 min at 4 °C and then gently resuspended in 200 µL/per reaction cold Nuclei Extraction Buffer. 2 µL aliquot was taken mixed with 2 µL trypan blue to examine nuclei integrity under the microscope. In the meantime, ConA beads (BioMag®Plus Concanavalin A, BP531) were prewashed and equilibrated in binding buffer (20 mM HEPES–KOH, pH 7.5, 10 mM KCl, 1 mM $CaCl_2$, 1 mM $MnCl_2$). Nuclei for each condition that needed to be tested were mixed with 10 µL of beads and incubated for 10 minutes at RT while shaking. After checking the binding of the beads to the ConA beads under the microscope (2 µL sample+ 2 µL trypan blue), the samples were washed two times with 300 µL of 1%BSA antibody buffer (20mM HEPES–KOH, pH 7.5, 150 mM KCl,0.5 mM Spermidine, 1X Roche cOmplete Mini EDTA-free Protease Inhibitor, 5% Digitonin, 2mM EDTA and 1%BSA). The tubes were placed on the magnetic rack, and the liquid was withdrawn. Samples were then resuspended in 50µL/ reaction of 1%BSA antibody buffer and incubated for 1h at RT while shaking. The samples (3 technical replicates for each condition) were incubated with 216 nM in-house prepared BG4 antibody and 0.5 µL of H3K27me3 (Cell Signaling Technology,9733S) antibody for the positive control (1 replicate for each condition) and left overnight at 4°C while shaking.

The day after, samples were checked to ensure that they were bound to the beads and then washed two times with Dig-wash buffer (20mM HEPES–KOH, pH 7.5, 150 mM KCl,0.5 mM Spermidine, 1X Roche cOmplete Mini EDTA-free Protease Inhibitor, 5% Digitonin). Samples containing BG4 and samples for the negative control were resuspended in 50 µL of Dig-wash buffer containing 2 µL of anti-flag antibody (DYKDDDDK Tag Antibody, 2368S) and incubated for 1h at RT while shaking. Then, the BG4 and negative samples were washed three times in 100 µL Dig wash buffer, resuspend in 50 µL Dig wash buffer supplemented with 0.5 µL of tertiary antibody (anti-rabbit, Abcam, ABIN101961) and incubated for 1h at RT while shaking. Samples were then washed three times with 100µL Dig-wash buffer, resuspend in 50 µL Dig-300 buffer (20 mM HEPES–KOH, pH 7.5, 300 mM KCl, 0.5 mM Spermidine, 0.01% Digitonin, 1X Roche cOmplete Mini EDTA-free Protease Inhibitor) containing pA-Tn5 (CUTANA™ pAG-Tn5 for CUT&Tag, 15-1017) adapter complex (1:20) and incubated for 1h at RT while shaking. Samples were then, washed three times in Dig-300 buffer, resuspend in 100 µL of Tagmentation buffer (Dig-300 buffer supplemented with 10 mM $MgCl_2$) and incubated for 1h at 37°C. Samples were then, washed twice with TAPS buffer (10 mM TAPS buffer-J63268.AE, Thermo Scientific Chemicals, 0.2 mM EDTA), mixed by vortexing in 100 µL of Protenaise

K buffer (0.5mg/mL Proteinase K, Thermo Fisher Scientific EO0492/EO0491, 0.5% SDS in 10mM Tris-HCl buffer pH 8.0) and incubate 1 hr at 55 ºC. Samples were then purified with Zymo DNA Clean & Concentrator-5 (D4013) and samples were then, amplified using NEBNext HiFi 2x PCR Master mix (NEB, M0541) with the following set up (Initial 72 °C for 5 min, 20X cycle (98 °C for 30 sec, 98 °C for 10 sec, 63 °C for 10 sec) and 72°C for 1 min). Samples were then, purified using Ampure XP beads (A63880) following manufacturer's instruction. Quality control of the sample was perfomed using TapeStation ScreenTape HSD1000 (5067-5584) and TapeStation ScreenTape HSD1000 reagents (5067-5585) according to manufacturer instruction in a TapeStation (TapeStation Agilent 4150).

CUT&Tag-qPCR

CUT&Tag libraries were diluted 1:10, and 2 µL of the diluted CUT&Tag library was mixed with 1 µM of primer mix (Table S6) and 5 µL of SYBR Green PCR Master Mix (Applied Biosystems: 4385612) according to the following protocol 20 seconds at 95 °C, 39 cycles of 10 seconds at 95 °C, 30 seconds at 60 °C, and 10 seconds at 72 °C, and finally heating to 90 °C. The C_t values obtained were used to calculate the relative fold-change in gene expression using the $\Delta\Delta C_t$ method and assess the relative fold change at G4 target sites when compared against G4-positive regions (MAZ, RBBP4, RPA3).

RNAseq sample preparation and bioinformatic analysis

For transcriptome-wide gene expression changes in transfected samples, cells were plated the day before in 6-well plates, and 24h after transfection, they were incubated with either (0.5%) DMSO or 2.5 µM Cl-PDS_2 for 6 hours. For DC-34/PyPDS-treated samples, cells were plated in a 6-well plate the day before treatment with respectively 7.5 µM and 2.5 µM of compound for 6h. Cells were harvested and washed with 1x DPBS (300xg, 5 minutes). RNA was then extracted using RNeasy mini kits (Qiagen, 74104) according to the manufacturer's instructions. The quality of the extracted RNA was assessed by TapeStation (High Sensitivity RNA Screen Tape, Agilent) and sent to Novogene service for sequencing. RNA libraries were prepared using Novogene kit for library preparation and then sequenced with Illumina NovaSeq X-plus. CASAVA (version 1.8) was used to perform base calling and Phred score. The quality of the reads was assessed with FastQC (Fastp v.0.23.1). Sequence alignment to the reference genome (GRCh38-NCBI:GCA_000001405.27) was performed using HISAT2. Sorted bam files were generated with (Hisat2 v.2.0.5), and gene counts using FPKM. Differential gene expression analysis was performed with the DESeq2, and data were plotted using iDEP.96.

Data availability

All new datasets have been deposited in the gene expression omnibus (GEO): GSEXXXXXX

Acknowledgments

M.D.A is supported by a Biotechnology and Biological Sciences Research Council (BBSRC) David Phillips Fellowship (BB/R011605/1) and is a Lister Institute Research Prize holder (2022). S.P.N. acknowledges support from the Engineering and Physical Sciences Research Council (EPSRC) (EP/S023518/1). T.E.M acknowledges support from the Engineering and Physical Sciences Research Council (EPSRC) (EP/S023518/1). J.S.S. is supported by the Intramural Research Programs of the National Institutes of Health, National Cancer Institute (NCI), Center for Cancer Research, Project BC011585. This work was supported by a Leverhulme Trust Research Grant (RPG-2022-184). We are thankful to Prof Luca Magnani for providing MCF7 cells and for helpful discussions. The authors also acknowledge Prof Ramon Vilar, Dr Michele Stasi, and Dr Koustav Pal for insightful discussions. The authors also acknowledge Dr Denise Liano, Dr Aisling Minard, Dr Souroprobho Chowdhury and Dr Anna Di Porzio for initial experimental efforts on this project.

Authors Contribution

S.P.N. designed research with M.D.A. and performed all the experiments unless otherwise specified. E.C, L.S.L., R.N., A.J.L., E.F, T.E.M. have contributed to synthesizing the G4-ligands. E.C. and A.J.L. perform FRET melting assays. S.G. performed with S.P.N. CUT&Tag experiments. J.S.S. conceived DC-34 and helped edit the manuscript. C.R.F synthesized DC-34 used in this study. M.D.A. wrote the manuscript supported by S.P.N and with input from all authors. All authors contributed to critical discussion and data interpretation. M.D.A. conceived and supervised the research project.

References

1. Davis, J.T. G-quartets 40 years later: from 5'-GMP to molecular biology and supramolecular chemistry. *Angew. Chem. Int. Ed. Engl.* **43**, 668-698 (2004).
2. Gellert, M., Lipsett, M.N. & Davies, D.R. Helix formation by guanylic acid. *Proc. Natl. Acad. Sci. U. S. A.* **48**, 2013-2018 (1962).
3. Parkinson, G.N., Lee, M.P. & Neidle, S. Crystal structure of parallel quadruplexes from human telomeric DNA. *Nature* **417**, 876-880 (2002).
4. Biffi, G., Tannahill, D., McCafferty, J. & Balasubramanian, S. Quantitative visualization of DNA G-quadruplex structures in human cells. *Nat. Chem.* **5**, 182-186 (2013).
5. Di Antonio, M. *et al.* Single-molecule visualization of DNA G-quadruplex formation in live cells. *Nat. Chem.* **12**, 832-837 (2020).
6. Summers, P.A. *et al.* Visualising G-quadruplex DNA dynamics in live cells by fluorescence lifetime imaging microscopy. *Nat. Commun.* **12**, 162 (2021).
7. Lyu, J., Shao, R., Kwong Yung, P.Y. & Elsasser, S.J. Genome-wide mapping of G-quadruplex structures with CUT&Tag. *Nucleic Acids Res.* **50**, e13 (2022).
8. Esnault, C. *et al.* G4access identifies G-quadruplexes and their associations with open chromatin and imprinting control regions. *Nat. Genet.* **55**, 1359-1369 (2023).
9. Hansel-Hertsch, R. *et al.* G-quadruplex structures mark human regulatory chromatin. *Nat. Genet.* **48**, 1267-1272 (2016).
10. Hansel-Hertsch, R. *et al.* Landscape of G-quadruplex DNA structural regions in breast cancer. *Nat. Genet.* **52**, 878-883 (2020).
11. Chambers, V.S. *et al.* High-throughput sequencing of DNA G-quadruplex structures in the human genome. *Nat. Biotechnol.* **33**, 877-881 (2015).
12. Robinson, J., Raguseo, F., Nuccio, S.P., Liano, D. & Di Antonio, M. DNA G-quadruplex structures: more than simple roadblocks to transcription? *Nucleic Acids Res.* **49**, 8419-8431 (2021).
13. Balasubramanian, S., Hurley, L.H. & Neidle, S. Targeting G-quadruplexes in gene promoters: a novel anticancer strategy? *Nat. Rev. Drug Discov.* **10**, 261-275 (2011).
14. Kosiol, N., Juranek, S., Brossart, P., Heine, A. & Paeschke, K. G-quadruplexes: a promising target for cancer therapy. *Mol. Cancer* **20**, 40 (2021).
15. Marsico, G. *et al.* Whole genome experimental maps of DNA G-quadruplexes in multiple species. *Nucleic Acids Res.* **47**, 3862-3874 (2019).
16. Figueiredo, J., Mergny, J.L. & Cruz, C. G-quadruplex ligands in cancer therapy: Progress, challenges, and clinical perspectives. *Life Sci.* **340**, 122481 (2024).
17. Boddupally, P.V. *et al.* Anticancer activity and cellular repression of c-MYC by the G-quadruplex-stabilizing 11-piperazinylquindoline is not dependent on direct targeting of the G-quadruplex in the c-MYC promoter. *J. Med. Chem.* **55**, 6076-6086 (2012).
18. Galli, S., Flint, G., Ruzickova, L. & Di Antonio, M. Genome-wide mapping of G-quadruplex DNA: a step-by-step guide to select the most effective method. *RSC Chem. Biol.* **5**, 426-438 (2024).
19. Spiegel, J. *et al.* G-quadruplexes are transcription factor binding hubs in human chromatin. *Genome Biol.* **22**, 117 (2021).
20. Calabrese, D.R. *et al.* Chemical and structural studies provide a mechanistic basis for recognition of the MYC G-quadruplex. *Nat. Commun.* **9**, 4229 (2018).

21. Cadoni, E., De Paepe, L., Manicardi, A. & Madder, A. Beyond small molecules: targeting G-quadruplex structures with oligonucleotides and their analogues. *Nucleic Acids Res.* **49**, 6638-6659 (2021).
22. Minard, A. *et al.* A short peptide that preferentially binds c-MYC G-quadruplex DNA. *Chem. Commun.* **56**, 8940-8943 (2020).
23. Chowdhury, S., Wang, J., Nuccio, S.P., Mao, H. & Di Antonio, M. Short LNA-modified oligonucleotide probes as efficient disruptors of DNA G-quadruplexes. *Nucleic Acids Res.* **50**, 7247-7259 (2022).
24. Heddi, B., Cheong, V.V., Martadinata, H. & Phan, A.T. Insights into G-quadruplex specific recognition by the DEAH-box helicase RHAU: Solution structure of a peptide-quadruplex complex. *Proc. Natl. Acad. Sci. U. S. A.* **112**, 9608-9613 (2015).
25. Keppler, A. *et al.* A general method for the covalent labeling of fusion proteins with small molecules in vivo. *Nat. Biotechnol.* **21**, 86-89 (2003).
26. Rodriguez, R. *et al.* A novel small molecule that alters shelterin integrity and triggers a DNA-damage response at telomeres. *J. Am. Chem. Soc.* **130**, 15758-15759 (2008).
27. De Cian, A., Delemos, E., Mergny, J.L., Teulade-Fichou, M.P. & Monchaud, D. Highly efficient G-quadruplex recognition by bisquinolinium compounds. *J. Am. Chem. Soc.* **129**, 1856-1857 (2007).
28. Deng, W., Shi, X., Tjian, R., Lionnet, T. & Singer, R.H. CASFISH: CRISPR/Cas9-mediated in situ labeling of genomic loci in fixed cells. *Proc. Natl. Acad. Sci. U. S. A.* **112**, 11870-11875 (2015).
29. Larson, M.H. *et al.* CRISPR interference (CRISPRi) for sequence-specific control of gene expression. *Nat. Protoc.* **8**, 2180-2196 (2013).
30. McCarty, N.S., Graham, A.E., Studena, L. & Ledesma-Amaro, R. Multiplexed CRISPR technologies for gene editing and transcriptional regulation. *Nat. Commun.* **11**, 1281 (2020).
31. Prasad, B. *et al.* A complementary chemical probe approach towards customized studies of G-quadruplex DNA structures in live cells. *Chem. Sci.* **13**, 2347-2354 (2022).
32. Anders, C., Niewoehner, O., Duerst, A. & Jinek, M. Structural basis of PAM-dependent target DNA recognition by the Cas9 endonuclease. *Nature* **513**, 569-573 (2014).
33. Tomasello, G., Armenia, I. & Molla, G. The Protein Imager: a full-featured online molecular viewer interface with server-side HQ-rendering capabilities. *Bioinformatics* **36**, 2909-2911 (2020).
34. Lam, E.Y., Beraldi, D., Tannahill, D. & Balasubramanian, S. G-quadruplex structures are stable and detectable in human genomic DNA. *Nat. Commun.* **4**, 1796 (2013).
35. Peraro, L. *et al.* Cell Penetration Profiling Using the Chloroalkane Penetration Assay. *J. Am. Chem. Soc.* **140**, 11360-11369 (2018).
36. Ahmadi, S.E., Rahimi, S., Zarandi, B., Chegeni, R. & Safa, M. MYC: a multipurpose oncogene with prognostic and therapeutic implications in blood malignancies. *J. Hematol. Oncol.* **14**, 121 (2021).
37. Albert, T. *et al.* The chromatin structure of the dual c-myc promoter P1/P2 is regulated by separate elements. *J. Biol. Chem.* **276**, 20482-20490 (2001).
38. Esain-Garcia, I. *et al.* G-quadruplex DNA structure is a positive regulator of MYC transcription. *Proc. Natl. Acad. Sci. U. S. A.* **121**, e2320240121 (2024).
39. Zhang, X., Spiegel, J., Martinez Cuesta, S., Adhikari, S. & Balasubramanian, S. Chemical profiling of DNA G-quadruplex-interacting proteins in live cells. *Nat. Chem.* **13**, 626-633 (2021).
40. Hui, W.W.I., Simeone, A., Zyner, K.G., Tannahill, D. & Balasubramanian, S. Single-cell mapping of DNA G-quadruplex structures in human cancer cells. *Sci. Rep.* **11**, 23641 (2021).
41. Pennarun, G. *et al.* Apoptosis related to telomere instability and cell cycle alterations in human glioma cells treated by new highly selective G-quadruplex ligands. *Oncogene* **24**, 2917-2928 (2005).
42. Yu, Z. *et al.* Chem-map profiles drug binding to chromatin in cells. *Nat. Biotechnol.* **41**, 1265-1271 (2023).
43. Cuddapah, S. *et al.* Genomic profiling of HMGN1 reveals an association with chromatin at regulatory regions. *Mol. Cell Biol.* **31**, 700-709 (2011).

44. Siddiqui-Jain, A., Grand, C.L., Bearss, D.J. & Hurley, L.H. Direct evidence for a G-quadruplex in a promoter region and its targeting with a small molecule to repress c-MYC transcription. *Proc. Natl. Acad. Sci. U. S. A.* **99**, 11593-11598 (2002).
45. Qin, G. *et al.* Targeting specific DNA G-quadruplexes with CRISPR-guided G-quadruplex-binding proteins and ligands. *Nat. Cell. Biol.* **26**, 1212-1224 (2024).
46. Felsenstein, K.M. *et al.* Small Molecule Microarrays Enable the Identification of a Selective, Quadruplex-Binding Inhibitor of MYC Expression. *ACS Chem. Biol.* **11**, 139-148 (2016).
47. Flusberg, D.A. *et al.* Identification of G-Quadruplex-Binding Inhibitors of Myc Expression through Affinity Selection-Mass Spectrometry. *SLAS Discov.* **24**, 142-157 (2019).
48. Robinson, J. *et al.* G-quadruplex structures regulate long-range transcriptional reprogramming to promote drug resistance in ovarian cancer. *bioRxiv*, 2024.2006.2024.600010 (2024).