https://doi.org/10.1093/nar/gkad189

A red light-triggered chemical tool for sequence-specific alkylation of G-quadruplex and i-motif DNA

Enrico Cadoni*,†, Lessandro De Paepe†, Gertjan Colpaert, Ruben Tack, Dries Waegeman, Alex Manicardi* and Annemieke Madder*

Organic and Biomimetic Chemistry Research Group, Ghent University, Krijgslaan 281 S4, B-9000 Ghent, Belgium

Received October 13, 2022; Revised February 03, 2023; Editorial Decision February 23, 2023; Accepted March 12, 2023

ABSTRACT

The importance of non-canonical DNA structures such as G-quadruplexes (G4) and intercalating-motifs (iMs) in the fine regulation of a variety of cellular processes has been recently demonstrated. As the crucial roles of these structures are being unravelled, it is becoming more and more important to develop tools that allow targeting these structures with the highest possible specificity. While targeting methodologies have been reported for G4s, this is not the case for iMs, as evidenced by the limited number of specific ligands able to bind the latter and the total absence of selective alkylating agents for their covalent targeting. Furthermore, strategies for the sequence-specific covalent targeting of G4s and iMs have not been reported thus far. Herein, we describe a simple methodology to achieve sequence-specific covalent targeting of G4 and iM DNA structures based on the combination of (i) a peptide nucleic acid (PNA) recognizing a specific sequence of interest, (ii) a pro-reactive moiety enabling a controlled alkylation reaction, and (iii) a G4 or iM ligand orienting the alkylating warhead to the reactive residues. This multi-component system allows for the targeting of specific G4 or iM sequences of interest in the presence of competing DNA sequences and under biologically relevant conditions.

INTRODUCTION

In the early 1950s, Watson and Crick defined the classical DNA double-helix model (1). Nevertheless, over the years, it has been shown that DNA can adopt a variety of alternative secondary structures, including G-quadruplexes (G4s) and intercalating-motifs (iMs) (2,3). These structures arise in consecutive G- or C-rich tracts and are found to be prominent in human telomeres and promotor regions of certain genes, such as proto-oncogenes. G4s are composed of G-tetrads, consisting of four Hoogsteen hydrogen-bonded guanines stabilized by cations and are readily formed under physiological conditions. In contrast, iMs often require acidic conditions to guarantee the protonation of a cytosine, needed for the formation of the hemiprotonated cytosine dimer $(C–C)^+$ (4). Therefore, in contrast to G4s, which were intensively studied and characterized in all their aspects, the interest in their C-rich counterparts was put on the back burner after being observed for the first time in 1993 (5). However, more recently, growing evidence of iM formation under physiological conditions mimicking the cellular environment, including neutral pH and molecular crowding conditions was gathered (6), and their formation in a cellular environment has been demonstrated (7). These findings have reinvigorated the interest in such structures. Additionally, the formation of both G4 and iM structures on the same dsDNA sequence has been reported as mutually-exclusive. This is a consequence of the steric hindrance that is generated on the complementary strand when either one or the other is formed (8,9). This interdependent formation can lead to opposite effects when stabilizing one of these structures, hinting towards a functional relationship between G4s and iMs in gene expression regulation.

For a detailed study of their role in cellular processes the development of biochemical tools, able to interfere with their biological mode of action, is needed. This has been intensively explored for G4s, leading to a myriad of small molecules able to bind and stabilize the G4-architecture in a non-covalent manner. Additionally, several covalent methodologies were developed as well, using G4-ligands

*To whom correspondence should be addressed. Tel: +32 92644472; Fax: +32 9244998; Email: Annemieke.Madder@UGent.be
Correspondence may also be addressed to Alex Manicardi. Email: Alex.Manicardi@unipr.it
Correspondence may also be addressed to Enrico Cadoni. Email: Enrico.Cadoni@UGent.be
†First two authors contributed equally to the work.
Present address: Alex Manicardi, Department of Chemistry, Life Sciences and Environmental Sustainability, University of Parma, Viale delle Scienze, 17/A, I-43124 Parma, Italy.

decorated with either reactive (e.g. vinyl derivatives (10), nitrogen mustards (11)) or triggerable (e.g. phenols (12), quinone methides (13,14), benzophenones or arylazides (15) and furan (16)) moieties. These triggerable methods represent a more elegant approach as they offer spatiotemporal control over the alkylation reaction and, at the same time, reduce the risk for off-target interactions. Specifically relevant for their therapeutic applicability are those methodologies that avoid using harmful UV-irradiation (12,16). In particular, the use of red light-responsive systems ensures a deeper tissue penetration, offering an additional therapeutic advantage (17).

Although various reports describing the realization of ligands able to recognize or induce a specific G4-topology have appeared (18), the main limitation of these small-molecule based approaches remains the lack of selectivity for a specific G4 sequence. To date, reported ligands generally display a high promiscuity and are not able to discriminate among G4s exhibiting a similar topology, due to their structural features and binding modes (most of them consist of planar aromatic moieties that tend to stack to the terminal G4-tetrad) (19). To overcome this selectivity problem, several approaches focusing on the use of oligonucleotides, or analogues thereof, have been reported in the past years (20). Richter and co-workers proposed a strategy to induce the formation of a specific G4-isoform and, at the same time, block the other isoforms through the hybridization of a naphthalene diimide-Peptide Nucleic Acid (PNA) conjugate to the proximal DNA region of the G4-target (21). More recently, Di Antonio *et al.* reported a G4-targeting methodology, based on the use of short Locked Nucleic Acid probes complementary to the G-rich tract, to disrupt an individual G4 in living cells (22).

In contrast, the available tools for iM targeting are still rather limited, given the only recently renewed interest in these secondary structures. Indeed, most of the available small molecules are derived from binders originally developed for G4s, and only a small number of them is reported to be truly selective for iM structures (4,23,24). In addition, to the best of our knowledge, no reports focusing on development of iM alkylating agents or sequence-specific iM targeting are available.

Since G4s and iMs are ubiquitous throughout the whole genome, and can exert different cellular functions, it is of utmost importance that one individual G4 or iM sequence can selectively be targeted. Based on what has been discussed above, some of the available methods promise to ensure a high level of selectivity. Nevertheless, none of the methodologies described so far relies upon the formation of a stable covalent linkage. The advantages of a covalent bond over a reversible interaction are numerous, and increasingly exploited in drug discovery. These include the increased potency and duration of action (and, as a consequence, a wider therapeutic window), and, more importantly in this context, a permanent blockade of target activity (25–27). Alternatively, covalent bond formation has been successfully exploited as a chemical tool to enable the pull-down of proteins of interest and to study RNA-protein interactions with a lower background signal and improved efficiency (28–31). However, uncontrolled alkylation of random cellular components, resulting from non-specific interaction with the target molecules, can lead to undesired adverse reactions and even acute, immuno-mediated toxicity (32). It is therefore of crucial importance to ensure high levels of selectivity when designing a covalent drug, to avoid detrimental side effects.

In this study, we report on a simple chemical tool to covalently target DNA secondary structures in a sequence-specific manner (Figure 1). The system relies on the combined use of a PNA probe functionalized with a secondary-structure ligand that is decorated with a pro-reactive moiety, which ensures spatiotemporal control of the alkylation reaction. PNAs, acting as recognition elements able to recognise the flanking region of the targeted secondary structure, were selected over other DNA analogues in view of the exceptional stability of PNA:DNA duplexes and their stability towards nucleases as well as proteases (33–35). Additionally, the ease of functionalization on solid support readily allows the insertion of ligands, spacers and cross-linking warheads with little synthetic effort. The secondary structure ligand, upon recognition of and binding to the desired target, allows positioning of the pro-reactive warhead in a more controlled and reproducible way, thus inducing the necessary proximity for the alkylation reaction to proceed in a selective fashion.

MATERIALS AND METHODS

The synthesis of ligands (Supplementary Schemes S1–S6) and PNA probes (Supplementary Tables S1 and S2) is reported in detail in the ESI, along with the characterization of the purified compounds (Supplementary Figures S20–S55).

Alkylation experiments using PNA-ligand systems

In a typical experiment, a 10 μM working solution of PS (MB or CV, for G4 or iM alkylation experiments respectively) was freshly prepared. In a 1.5 mL Eppendorf tube, 300 μL of a buffered solution (Tris 10 mM, pH 7.4, 100 mM KCl for G4 alkylation experiments or NH_4OAc buffer 20 mM, pH 5.0, 100 mM KCl, pH 6.0; phosphate buffer: 1 mM K_3PO_4, pH 7.0; PBS buffer: Na_2HPO_4 10 mM pH 8.0, 100 mM KCl; for iM alkylation experiments) containing DNA at 5 μM concentration was prepared. All DNA samples were annealed heating at 95°C for 5 min and cooled down to room temperature over a period of 4 h. The PNA probe was added at a final concentration of 5 μM and the solution incubated for 30 min. PS was finally added to the tube at a final concentration of 1 μM (for MB) or 5 μM for CV. The lamp was placed on top of the Eppendorf tube for the entire duration of the experiment (20 min). 50 μL of solution was sampled at different irradiation times and left to react for two h at 37°C, shielded from light. The peaks formed upon irradiation were isolated by HPLC, freeze-dried and analysed via MALDI-TOF to confirm the formation of the alkylation product.

FID assay

In a typical experiment, 400 μL of 0.5 μM DNA solution in ammonium acetate buffer (20 mM NH_4OAc, 100 mM KCl,

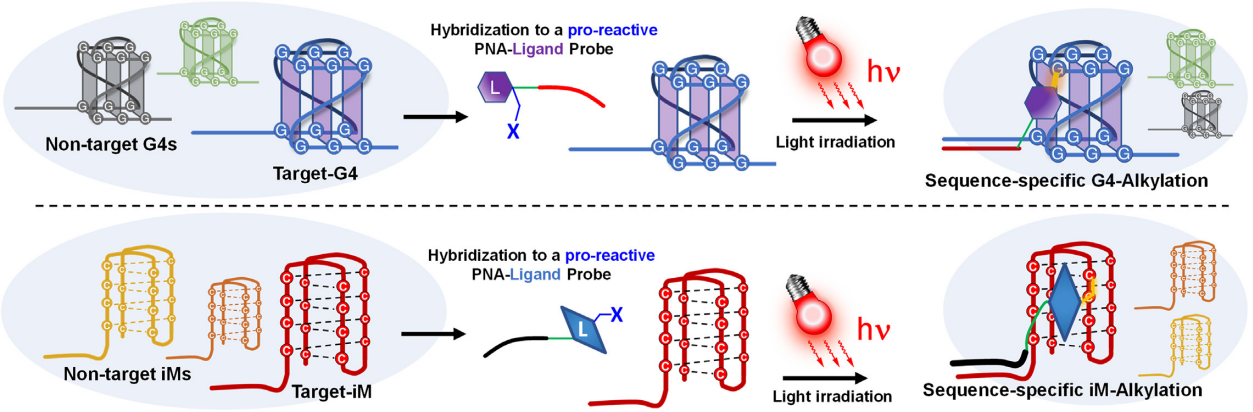

Figure 1. Schematic illustration of the presented work. A PNA probe equipped with a pro-reactive ligand sequence-specifically alkylates the targeted secondary structure (iM or G4). Activation of the moiety is achieved through light irradiation.

pH 5.0) were freshly prepared from a 100 μM stock solution. All DNA samples were annealed by heating at 95°C for 5 min and cooling down to room temperature over a period of 4 h. Thiazole orange (TO) was subsequently added from a 250 μM stock solution to a final concentration of 1 μM and the system equilibrated for 10 min before the collection of the fluorescence signal. For each titration experiment, increasing concentrations of ligand were added to the solution (0–20 μM, 0–40 equivalents). The fluorescence emission was measured using a Cary Eclipse fluorescence spectrophotometer, with a scan rate of 200 nm/min, following the emission from 510 to 700 nm wavelength, at an excitation wavelength of 501 nm (36). The integral of the fluorescence emission was calculated in the interval 515–650 nm and each dataset normalized within respect to the solution containing 0 eq of ligand. The data were plotted following the formula:

$$D_x = 1 - \left(\frac{F_x}{F_0}\right) \quad (1)$$

where D_x is the TO displacement at x equivalents of ligand, F_x is integral of fluorescence emission at x equivalents of ligand, and F_0 is integral of fluorescence emission at 0 equivalents of ligand. The TO displacement and the DC_{50} were calculated using the built-in non-linear fitting (two-phase association) by Graphpad Prism (v 8.4.2).

iM ligand oxidation experiments

A 100 μM working solution of Crystal Violet (CV) was freshly prepared. In a 1.5 mL Eppendorf tube, 300 μL of air saturated buffered solution (NH_4OAc buffer 20 mM, pH 5.0, 100 mM KCl) containing ligand at 50 μM was prepared and placed in the Eppendorf thermomixer at 37°C. When needed, DNA (**k-RAS M**) was added up to 5 μM concentration prior the addition of the ligand. DNA samples were annealed by heating at 95°C for 5 min and cooling down to room temperature over a period of 4 h. CV was added to the tube at desired final concentration. The lamp was placed on top of the Eppendorf tube and the sample irradiated for the entire duration of the experiment (60 min). 50 μL of solution was sampled at intermediate irradiation times and left to react for 2 h at 37°C, shielded from light. Ligand oxidation was evaluated by HPLC, through integration of the 260 nm absorbance of ligand peaks. Integration was performed using Agilent ChemStation for LC 3D systems software. Data reported are normalized to integrals at 0 min irradiation.

iM alkylation experiments using small ligands

A 10 μM working solution of PS was freshly prepared from a 1 mM stock solution. In a 1.5 mL Eppendorf tube, 300 μL of a buffered solution (NH_4OAc buffer 20 mM, pH 5.0, 100 mM KCl) containing DNA at 5 μM concentration was prepared. All DNA samples were annealed by heating at 95°C for 5 min and cooling down to room temperature over a period of 4 h. The ligand was then added at a final concentration of 25 μM and the solution incubated for 30 min. CV was added to the tube at a final concentration of 5.0 μM. The lamp was placed on top of the Eppendorf tube for the entire duration of the experiment (30 min). 50 μL of solution was sampled at different irradiation times and left to react for 2 h at 37°C, shielded from light. The peaks formed upon irradiation were isolated by HPLC, freeze-dried and analysed via MALDI-TOF to confirm the formation of the alkylation product.

RESULTS AND DISCUSSION

In a first step towards the development of a specific alkylation methodology for tetraplexed DNA structures, we directed our efforts to G4-DNA for the initial optimization. The first step involved the introduction of an alkylating ligand onto a suitable PNA sequence, designed to recognize the correct flanking region of the G4 of interest. First, different photo-activatable cross-linkers were evaluated. Subsequently, we selected the best-performing one to be included in the final ligand-PNA conjugates, and then systematically studied their capacity for sequence-specific alkylation in competition experiments.

Screening of alkylating warheads

We previously developed a G4-alkylating methodology exploiting furan-decorated G4-ligands (16). In an attempt to broaden the methodology towards additional warhead

chemistries, next to the oxidation-triggered furan moiety, alternative light-triggerable alkylating units (arylazide and diazirine) were introduced onto the PNA probe and the orienting G4 ligand. For the ligand, pyridine-dicarboxamide (PDC) 360A was chosen, as it was previously exploited in covalent targeting approaches in combination with different photo-crosslinking moieties (15,16), as well as for other non-covalent targeting applications (37–42). The PNA probes were designed to be complementary to the 5′-flanking region of the target and to contain the G4-ligand at their N-terminus, following a spacer and the amino acid containing the cross-linking moiety. L-Photo-leucine and L-azidophenyl alanine, activated upon UV irradiation, and L-furyl-alanine were conveniently introduced as alkylating units through the commercially available Fmoc-protected amino acid building blocks. The furan moiety within L-furyl-alanine can be oxidized into a reactive keto-enal by singlet oxygen ($^1O_2$) (43), which can be generated upon light irradiation of a co-localized PS, and then react with nucleophilic amines of the G4 DNA-target. The use of a G4-binding PS allows a pinpoint $^1O_2$ production, avoiding collateral oxidative damage.

Additionally, a (2-(2-aminoethoxy)ethoxy)acetyl spacer (AEEA) and a β-alanine residue were introduced at the N-terminus of the PNA probe to separate the crosslinking moiety from the PNA, giving the ligand sufficient freedom to better bind to the target G4. The C-terminus of the PNA was functionalized with two negatively charged residues (L-glutamic acid) to deliver an overall neutral molecule and allow a better chromatographic behaviour under the ion-pairing conditions required for the HPLC analysis of DNA sequences. The general structure of the probes is reported in Figure 2A. *c-KIT-21,* which is able to fold into a stable parallel G4 under physiologically relevant conditions, was used as a model target G4 (**c-KIT-T**) given its well-recognised therapeutic importance and the abundance of structural information available for this G4-forming sequence (44).

The alkylating abilities of the different systems were evaluated using the respective photo-activation conditions: UV-light irradiation (λ_{max} = 350 nm; 10 min) or red light irradiation (λ_{max} = 660 nm, 20 min) in presence of methylene blue, a PS able to bind G4-DNA and allow G4-localized $^1O_2$ production (15,45). No DNA consumption was observed for the PNA probe equipped with L-photo-leucine (**PNA-PhLeu**), while for **PNA-PheN$_3$** (equipped with L-4-azidophenylalanine) and **PNA-FAla** (equipped with L-furyl-alanine) reduction of the peak corresponding to the DNA starting material was observed (Figure 2B). Along with the reduction of the DNA peak, the simultaneous appearance of new peaks at higher retention time was observed, hinting towards the formation of alkylation products. MALDI-TOF analysis of the isolated peaks revealed the formation of species with a molecular weight compatible with cross-linked species (see ESI, Supplementary Figure S5). **PNA-FAla** performed better as compared to **PNA-PheN$_3$**, as demonstrated by a more pronounced reduction of the DNA peak upon light irradiation (30% versus 7% DNA consumption). The higher DNA consumption and a more gentle triggering of the system (via localized generation of $^1O_2$ using red light) justify the selection of the furan-containing ligands as alkylating warheads in the subsequent experiments.

Design of G4-alkylating PNA-probes

In order to allow a detailed study with different G4 sequences, a slightly modified version of the alkylating furan-modified G4-ligand was designed for convenient introduction of the ligand and the furan warhead during solid-phase synthesis (see ESI Supplementary scheme S1). Despite the (minor) modification of its structure, PNA bearing this monomer showed DNA alkylation outcomes in line with those obtained with **PNA-FAla** (*vide infra*). To avoid furan alkylation during PNA cleavage, the aromatic furan moiety was temporarily protected through a reversible Diels-Alder adduct with maleimide (46,47). This allowed for an increase in the overall synthesis yield of the entire PNA-ligand construct from 5% to 22% (see ESI, section 3).

Two families of probes complementary to either the 3′ or the 5′ flanking region of **c-KIT-T** (**PNA-G4-3** and **PNA-G4-5** series) were synthesized, to evaluate the influence of the recognition region on the alkylation outcome. In addition, the number of spacers between the ligand and the alkylating warhead on the one hand and the PNA core on the other hand was systematically evaluated. We synthesized probes bearing 0, 1 or 2 AEEA spacers, to check the effect of an increase or decrease in flexibility. Finally, probes lacking the G4-ligand were synthesized to understand if the insertion of the G4-binding ligand was truly beneficial for the G4-targeting. To simplify the nomenclature used for the PNAs, we adopted the following model: **PNA-G4-XY#**. Herein **X** represents the targeted flanking region (**5** or **3**), **Y** the presence of ligand (**L**) or only alkylating unit (**F**) and **#** indicates the number of spacing units (**0**–**2**). An overview of PNA structures and sequences is provided in Figure 3A.

Towards sequence-specific G4-alkylation

Each of the synthesized probes was then evaluated for its alkylation ability in presence of **c-KIT-T**, in a mixture containing an equimolar concentration of the competing G4-sequence **c-KIT-NT**, displaying an identical G4 conformation but lacking the flanking regions (sequences and PNA recognition sites are shown in Figure 3B). In this set-up, for both the **PNA-G4-5L#** and **PNA-G4-3L#** series, alkylation of the target was achieved with a minimal DNA consumption of 20%. Slightly higher alkylation was obtained when at least one spacing unit was used, indicating the need for a minimal spacer length and torsional freedom to allow docking of the ligand on the targeted structure. In contrast, for the furan-only-containing probes (**PNA-G4-XF#**), both the sequence recognition site and the number of spacing units had a significant impact on the alkylation outcome. In detail, we observed a higher alkylation yield only when a single-spacing unit was employed, and a drastic drop of efficiency when targeting the 3′-flanking regions of **c-KIT-T** (Figure 4A and B and Supplementary Figure S8). Similar reaction outcomes were obtained in experiments performed in absence of **c-KIT-NT (**Supplementary Figure S7 in ESI**)**. Formation of an alkylation

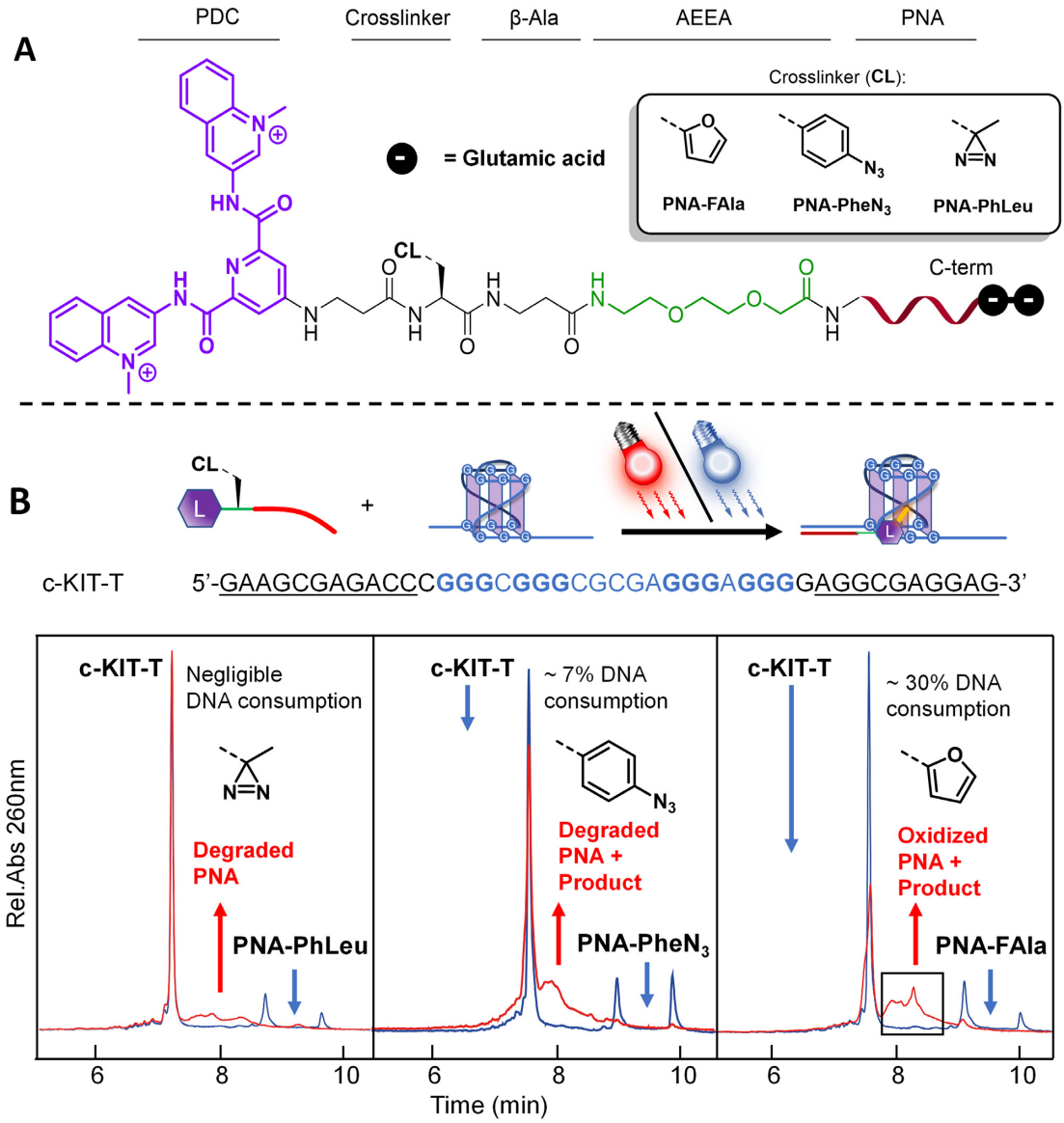

Figure 2. Preliminary screening of three different alkylation moieties (L-photo-leucine, L-azidophenyl alanine, furyl-alanine) incorporated in the PNA probe. (**A**) Design of the PNA probes containing the alkylating units. (**B**) Comparison of the alkylation efficiency (DNA consumption) for the three tested PNA constructs. Experiments were performed in buffered solution (Tris 100 mM, pH 7.4, 10 mM KCl), at 5 μM strand concentration. As a consequence of the adopted chromatographic conditions (suitable for the correct visualization of DNA), the PNA probes appear as double peaks.

product was confirmed by MALDI-TOF analysis, indicating a reaction with guanine as a main product and minor traces of alkylation adducts with adenine and cytosine (Supplementary Figure S12 in ESI). To confirm that the species observed by MALDI-TOF analysis represents a genuine alkylation product and not a stable, but non-covalent, ligand-DNA complex, control experiments with pre-oxidized PNA probes were performed: no reduction of the **c-KIT-T** peak was observed in the chromatogram in that case and MALDI-TOF spectra only showed the presence of the oxidized PNA probe (Supplementary Figure S6).

To further validate the system selectivity, the central G4-forming tract of *c-KIT* was substituted with a mutated, non-G4-forming sequence bearing two G→T substitutions on the central tetrads (**c-KIT-mut**). As expected, the unfolding of the quadruplex led to a dramatic drop in alkylation efficiency (Figure 4C) for the probes bearing the ligand. In contrast, all the PNAs equipped with the furan warhead alone showed the ability to alkylate the mutated target, thus further highlighting the importance of introducing the ligand as a second recognition element to increase the selectivity of the system.

Taken all together, these results demonstrate the importance of the simultaneous presence of the two recognition elements, a PNA for recognition of the flanking region and a G4-ligand for binding the secondary structure, for the selective targeting of the desired sequence and as tool to facilitate the design of the probes. In fact, if the system is flexible enough (1 or more AEEA spacers), after the recognition of the flanking region the ligand can orient the alkylat-

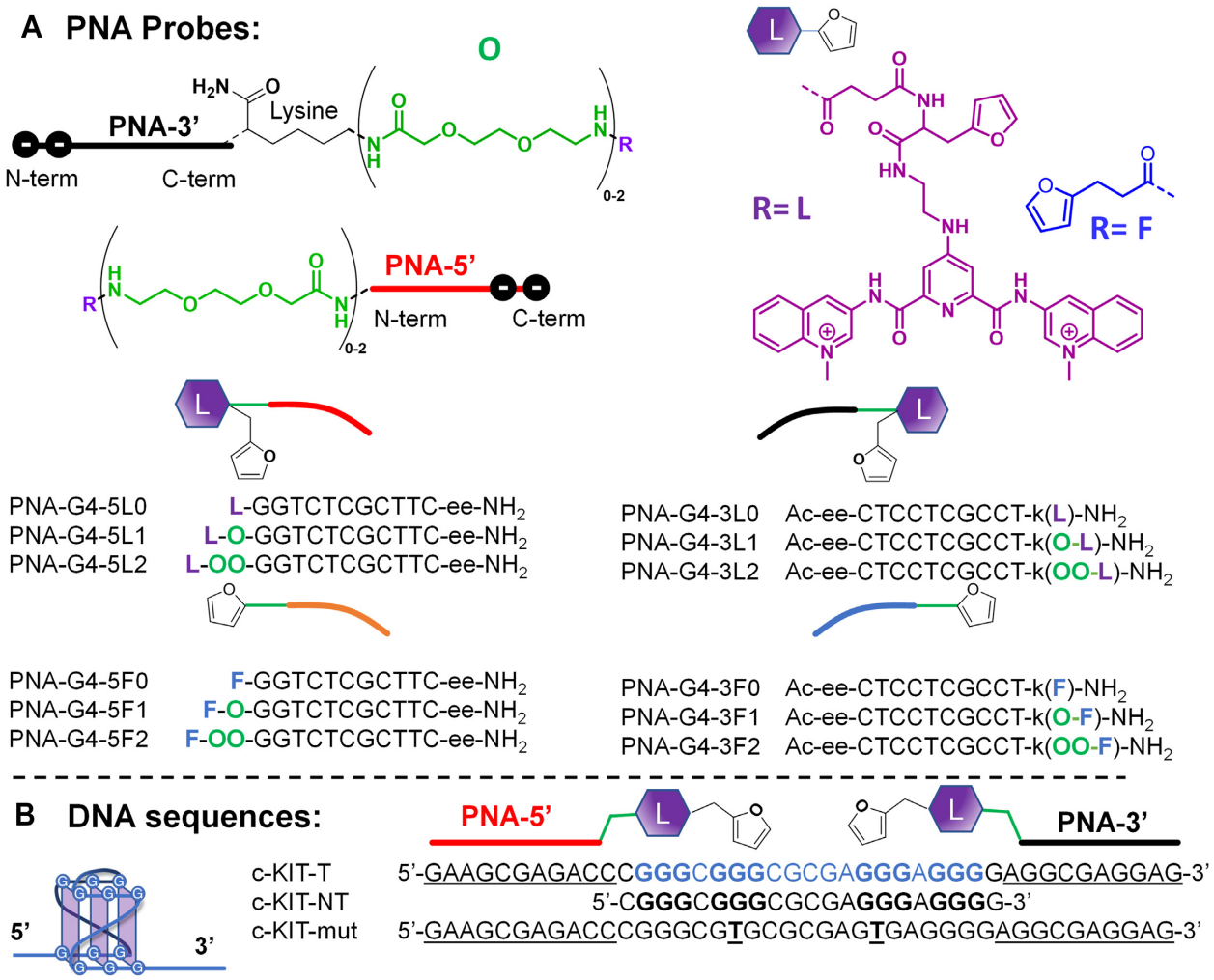

Figure 3. Design and sequences used for the evaluation of the G4-specific alkylation. (**A**) Structure and sequences of the PNA probes employed. (**B**) G4 forming DNA sequences. The G4-forming tract is indicated in blue, the guanines involved in the formation of tetrads are in bold, the mutations are in bold and underlined. O = (2-(2-aminoethoxy)ethoxyacetyl) spacer; L = G4 ligand (derived from PDC-DA-OH, see ESI Supplementary scheme S1), F = furylpropanoyl; e = glutamic acid; Ac = acetyl. c-KIT-T = target; c-KIT-NT = non-target; c-KIT-mut = mutated quadruplex.

ing unit in a suitable position only if the DNA secondary structure is present, ensuring the alkylation of the target. In contrast, probes lacking the ligand can lead to sequence alkylation but require a more careful design (sequence-dependence and flexibility dependence) and cannot discriminate between folded and unfolded sequences.

Expansion to alternative G4-forming sequences and cell lysate experiments

The use of a promiscuous G4-binder, such as the PDC 360A, in principle allows binding and the correct positioning of the alkylating warhead for different G4 structures, resulting in a general methodology for the selective targeting of alternative and widely studied G4s of interest (48–52), with little design and synthetic effort. To test this hypothesis, we designed additional targets bearing a different G4-forming sequence flanked by the same flanking sequences used before. The central G4-forming tracts were selected amongst biologically relevant sequences found in proto-oncogene promoter regions and able to fold into parallel and hybrid G4-structures. The alkylation results (Figure 5A and Supplementary Figure S9) confirm that the designed system is highly tolerant towards the target G4-forming sequence: in fact, moving from *c-KIT* to other structures (*c-MYC, k-RAS and BCL-2*) only marginally affects the alkylation outcome.

To show the possibility for applications in biological contexts, the alkylation methodology was tested in cell lysate. In this setup, mixtures with increasing concentrations of cell-lysate were spiked with the target DNA of interest. Alkylation of the target DNA was confirmed under these experimental conditions (Supplementary Figure S10). Finally, to further validate the selectivity of the approach, we performed a competition experiment at 25% cell lysate content. In this set-up, competing G4s bearing non-matching flanking regions were used as controls to test the specificity of alkylation. To avoid co-elution of the two DNA sequences during the HPLC analysis, the target **c-KIT-T** was modified with a fluorescein (FAM, **c-KIT-FAM**) tag, to decrease its polarity and shift its HPLC retention time. Similarly to what was previously observed in buffered solution, selective

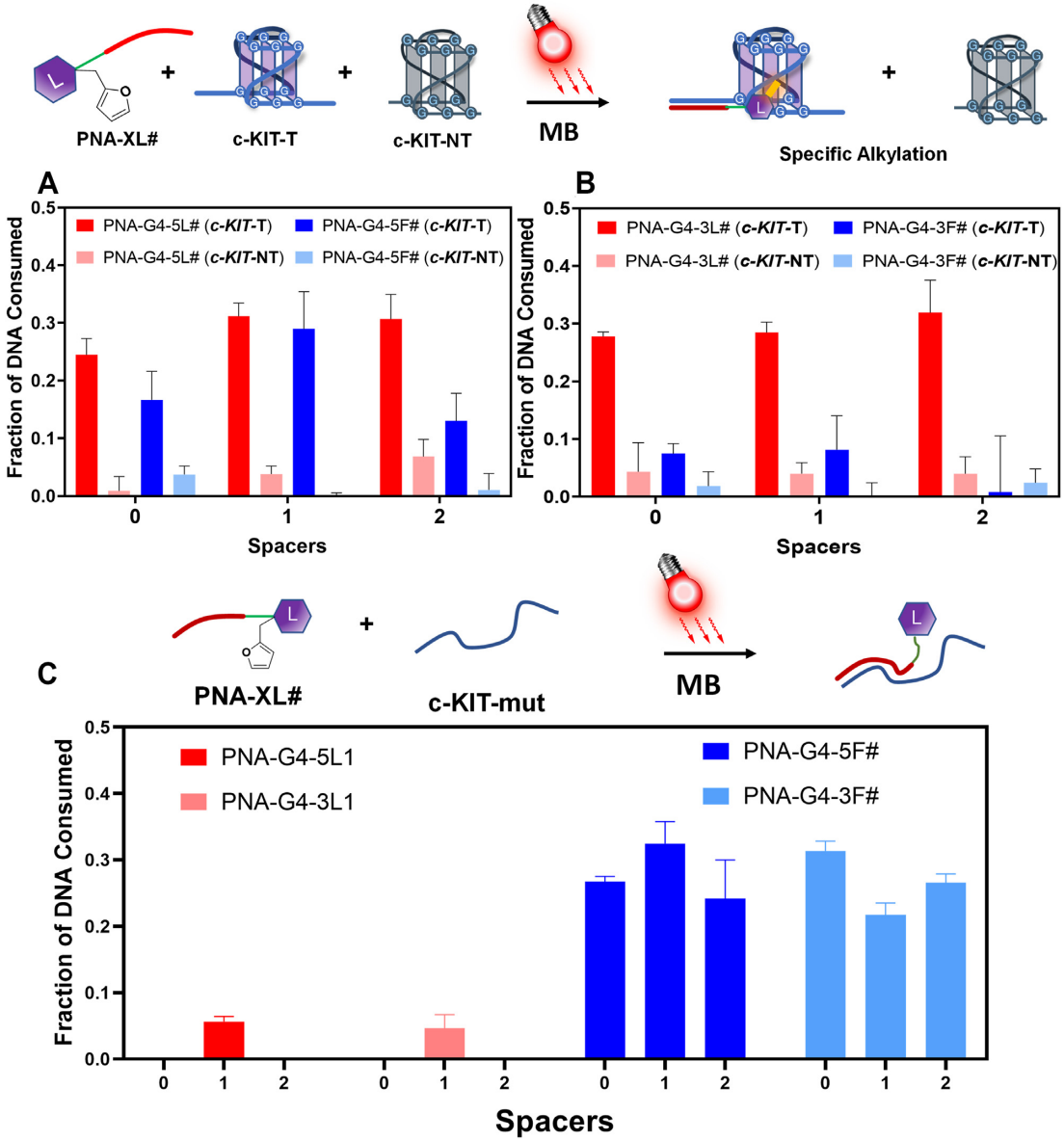

Figure 4. G4-alkylation experiments assessing the sequence-selectivity of the system. (**A**) Competition experiment for PNA-G4-5L# and PNA-G4-5F# in presence of G4-DNA competitor c-KIT-NT. (**B**) Competition experiment for PNA-G4-3L# and PNA-G4-3F# in presence of G4-DNA competitor c-KIT-NT. (**C**) Control experiment in presence of a single-stranded sequence c-KIT-mut. Experiments performed in buffered solution (Tris 10 mM, pH 7.4, 100 mM KCl), at 5 μM strands concentration, and 1 μM MB. # = number of spacers, as indicated on the x-axis.

alkylation of **c-KIT-FAM** over the non-target sequences was achieved (Figure 5B and Supplementary Figure S11).

Moving to the complementary strand: development of an alkylating iM ligand

While alkylating G4-ligands have been developed before by us and other groups (10–15,53), this was not the case for iM-structures. Based on our previous experience with the successful development of a furan-containing G4-targeting ligand, we set out to explore the broadening of the concept and design a suitable alkylating iM-ligand, to be later exploited for the realization of a sequence-selective iM alkylation methodology.

As mentioned in the introduction, most of the ligands exhibiting iM binding ability were originally developed as G4 or general DNA intercalating moieties. We rationally designed a new ligand starting from the structure of a candidate molecule that showed superior binding affinity for iM over other DNA structures (54) and which could be easily functionalized with a furanyl moiety without undermining the binding to the targeted structures: the **PBP1** ligand (Figure 6) (54).

Rational design of iM alkylating agents and evaluation of their iM binding affinity

In line with our previously reported method for G4-alkylation (16), a bimolecular approach based on the

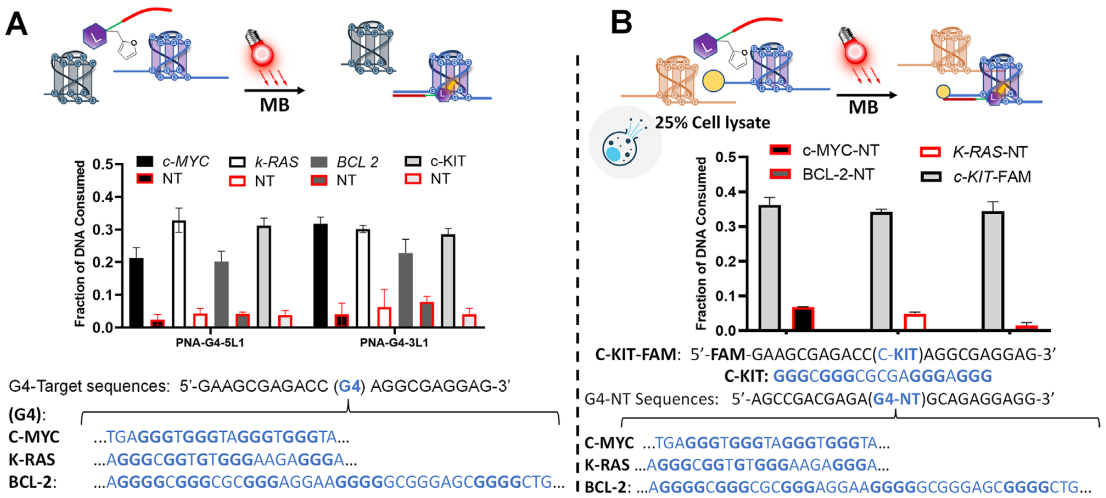

Figure 5. (**A**) Alkylation experiments of c-MYC, k-RAS and BCL-2. Target and non-target (NT) sequences were tested. (**B**) Alkylation of c-KIT-FAM in presence of cell lysate (25%) and competing G4-DNA sequences k-RAS-NT, BCL2-NT or c-MYC-NT. Experiments performed in buffered solution (Tris 10 mM, pH 7.4, 100 mM KCl), at 5 μM strand concentration and 1 μM MB concentration in absence (A) or presence (B) of cell lysate. NT-sequences depicted in ochre yellow, target sequences in blue with the yellow sphere representing the FAM label.

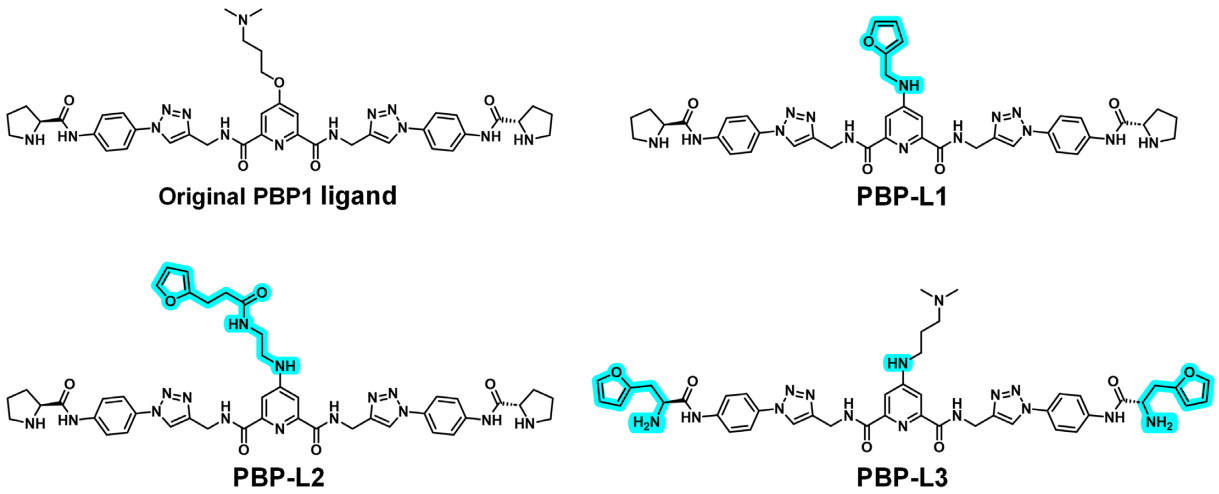

Figure 6. Ligands used for iM alkylation. Modifications of the original ligand (PBP1) are highlighted in blue.

co-localization of two elements to enhance the selectivity for iM over dsDNA was exploited: an iM ligand equipped with a pro-reactive furan warhead and crystal violet (CV), a photosensitizer known to be able to bind iM-DNA (55).

Starting from the selected ligand, the reactive furan-warhead was introduced in three different ways and the effect of the structural modifications on the binding affinity was evaluated. In first instance, the tertiary amine of the central pyridine core was substituted with a furan-containing arm, connected through either a short (**PBP-L1**) or a long (**PBP-L2**) spacing unit, while maintaining the structure of the lateral arms. Alternatively, replacement of the L-proline residues on the lateral chains with L-furylalanine (**PBP-L3**) allowed the simultaneous introduction of two warheads while minimizing perturbation of the original ligand structure. In all cases, the original 4-oxy linkage on the central pyridine core was substituted by its amino version to increase the stability of the system toward acid treatments, which will be required during the synthesis of the envisaged PNA probes. The structures of the furan-containing ligands are shown in Figure 6.

To evaluate if the modifications introduced in the ligand do not interfere with iM binding, a fluorescence-intercalator displacement (FID) assay, based on the displacement of thiazole orange (TO) from the iM-forming sequences (**BCL-2**, **h-TELO**, **k-RAS N/F/M**, **h-RAS**, Table 1) (36) was performed. CD and UV-melting analyses were performed to confirm the folding into iMs under identical buffer conditions (ammonium acetate buffer, pH 5) for all sequences (Supplementary ESI Table S3, Supplementary Figure S1–S3) (56,57).

The reported PBP1 ligand was originally proposed as BCL-2 binder (DC_{50} value of 0.9 μM) (54). After introducing the modifications to include either one or two alkylating moieties, all ligands still displayed good affinity for **BCL-2**. For the two most thermodynamic stable iMs

Table 1. FID assay: %TO displacement with PBP-L1, PBP-L2 and PBP-L3 in presence of iM forming and dsDNA sequences. Experiment performed in NH_4OAc buffer (20 mM, pH 5.0, 100 mM KCl), 0.5 μM DNA and 1.0 μM TO. ND: Not determined (displacement lower than 50%). Ligand concentration range: 0 to 20 μM

	K_d [μM] *(% displacement)*		
Sequence	PBP-L1	PBP-L2	PBP-L3
BCL-2 5′- CAGCCCCGCTCCCGCCCCCTTCCTCCCGCGCCCGCCCCT -3′	3.1 (73.7)	5.4 (66.5)	2.2 (85.0)
k-RAS M 5′-GCCCGGCCCCCGCTCCTCCCCCGCCGGCCCGGCCCGGCCC CCTCCTTCTCCCCG -3′	4.2 (99.8)	4.0 (99.7)	3.0 (98.7)
k-RAS F 5′- GCCCGGCTCGCCACCCTAGACCGCCCCAGCCACCCCT -3′	2.0 (94.0)	2.4 (92.2)	2.0 (84.7)
h-RAS 5′- CGCCCGTGCCCTGCGCCCGCAACCCGA -3′	3.6 (84.3)	4.1 (81.7)	3.5 (83.5)
h-TELO 5′- TAACCCTAACCCTAACCCTAACCC -3′	ND (<50%)	ND (<50%)	ND (<50%)
k-RAS N 5′- TCCCCTCTTCCCTCTTCCCACACCGCCCT -3′	ND (<50%)	ND (<50%)	ND (<50%)
dsLAC 5′- GAATTGTGAGCGCTCACAATTC -3′	ND (<50%)	ND (<50%)	ND (<50%)
T-Loop 5′- TATAGCTATATTTTTTTATAGCTATA -3′	ND (<50%)	ND (<50%)	ND (<50%)

formed in the k-RAS promotor, k-RAS Mid (**k-RAS M**) and k-RAS Far (**k-RAS F**) (58), high affinities were found, with near to complete TO displacement for **k-RAS M** using all three ligands. As for **h-RAS**, all ligands displayed similar affinities and high TO displacement efficiencies. For **h-TELO**, the displacement efficiency was minimal (a total TO displacement lower than 50%, in the ligand concentration range 0–20 μM). A similar behaviour was found for k-RAS Near (**k-RAS N**), the least populated iM of the k-RAS promotor region *in-vivo*. To confirm that the synthesized ligands are iM-selective, their displacement efficiency was also tested in presence of dsDNA sequences (**dsLAC** and **t-loop**) and none of them were able to efficiently displace TO. In general, the ligands showed comparable behaviour in stabilizing the iM-forming sequences, although in all cases **PBP-L3** exhibited a slightly higher affinity for iM-DNA as compared to **PBP-L1** and **PBP-L2**. This can be ascribed to a higher structural similarity to the original ligand PBP1, maintaining not only the lateral amino acid side chains (where L-proline is substituted with L-furylalanine), but also the tertiary N,N-dimethyl propylamine functionality.

iM alkylation using small molecule ligands

Furan-activation kinetics and the benefit of ligand co-localization were first tested for **PBP-L3** in presence of **k-RAS M**, by red-light irradiation of a solution containing CV. From the data reported in Figure 7A, the presence of the iM DNA allows for a reduction of the irradiation time needed for complete furan oxidation, hinting towards an increased local concentration of singlet oxygen in the surroundings of the furan warhead. Similar behaviour was observed with ligand **PBP-L1** (Supplementary Figure S4, ESI). $^1O_2$ quantum yield studies did not show any significant $^1O_2$ production variation for CV in presence or absence of iM-DNA (data not shown).

Next, the alkylation efficiency of each ligand was tested for different iM forming sequences (**k-RAS N/M/F**, **h-RAS**, **BCL-2**, **dsLAC** and **t-Loop**) and analysed via HPLC where the disappearance of the iM-DNA was followed, as an indication of alkylation efficiency. The DNA consumption of all target sequences after CV (1μM) irradiation (1h) in the presence of each of the three ligands (10 eq) was plotted (Figure 7B and Supplementary Figures S13-15). It can be noted that each ligand exhibits different alkylation efficiencies, which can be explained by the different positioning of the furan-warheads on the ligands. **PBP-L1**, despite showing good binding affinities for most of them (*vide supra,* Table 1), was not able to alkylate any of the tested sequences, while **PBP-L2** showed moderate alkylation with **BCL-2** (19%) and **k-RAS M** (43%). This can point towards the need for a longer spacer to ensure the required proximity between the activated furan and the nucleophilic residues. For the bis-furyl ligand **PBP-L3**, higher alkylation yields were found for all iM sequences (65% for **k-RAS M**, 36% for **k-RAS F**, 27% for **BCL-2**, and 17% for **h-RAS**), for which it showed a good affinity in the FID assay (*vide supra*). Under identical experimental conditions, none of the ligands was able to alkylate dsDNA sequences (dsLAC and t-Loop, example shown in Figure 7D) as well as the iM sequences for which they displayed a low affinity in the FID assay (k-RAS N). In summary, **PBP-L3** proves to be the best-performing ligand as a general and high-yielding alkylating tool for iMs.

The newly formed products observed in the HPLC chromatogram (example shown for **k-RAS M** in Figure 7C) at higher retention times (Rt = 8.1 min, **P3**) were purified and characterized by MALDI-TOF to confirm the formation of an alkylation product (ESI Supplementary Figure S16).

Sequence-specific iM-alkylation

In line with the developed G4-strategy, we then aimed to selectively target one specific iM, based on the recognition of its flanks by ligand-PNA conjugates. To test whether the developed methodology can be extended to the sequence-selective targeting of iM DNA secondary structures in

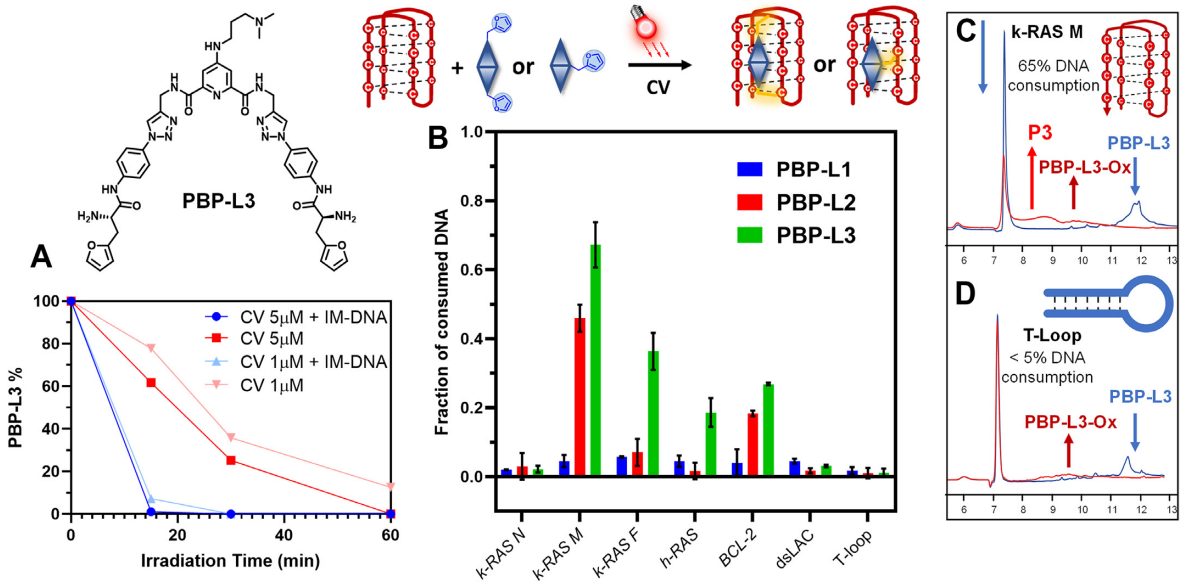

Figure 7. Development of an iM alkylation methodology. (**A**) Oxidation kinetics of PBP-L3 with irradiation of a solution containing CV (1 or 5 μM), in presence (blue lines) or absence (red lines) of iM DNA (k-RAS *M*). (**B**) Alkylation efficiency of PBP-L1 (blue bars), PBP-L2 (red bars) and PBP-L3 (green bars) in presence of different DNA sequences. (**C**) Example of chromatograms resulting from the alkylation of k-RAS-M. (**D**) Control alkylation experiment in presence of dsDNA T-Loop. Alkylation experiments performed in NH_4OAc buffer 20 mM, pH 5.0, 100 mM KCl at 50 μM ligand, 1 μM CV, and 5 μM DNA. PBP-L3-OX: oxidized PBP-L3; P3 = alkylation product with PBP-L3.

presence of competing iM sequences, the structure of **PBP-L3** was adapted to allow its insertion into a PNA strand (Figure 8A). Based on the promising results obtained for **k-RAS-M** with the small molecule binder approach, we used this iM sequence as a model target and included its natural 5′ and 3′ flanking sequences (now referred to as iM-target sequence **k-RAS-T,** depicted in Figure 8B).

Once more, PNA probes were designed to be complementary to either k-RAS 3′- or 5′- natural flanking regions. We focussed on the ligand-containing probes equipped with either 1 or 2 spacing units. In analogy with the previous probes, a similar nomenclature was maintained for the iM-targeting PNAs (**PNA-iM-XY#**). Due to the complexity of the structure and the lack of available information to correctly design a mutated sequence that is not able to fold into iM, we opted for unfolding the iM by changing the pH of the experiment rather than inserting mutations in the iM-forming core. Therefore, we tested the alkylation of the system at different pH values (5.0, 6.0, 7.0 and 8.0) displaying progressive unfolding of the targeted iM structure (refer to CD data in section 5 of the supporting information).

The alkylation efficiency of the PNA-probes bearing the ligand was tested in presence of the two competing sequences **k-RAS-T** and **k-RAS-NT**, at equimolar concentrations (Figure 9). As described above, the DNA consumption data were obtained by integration of the HPLC chromatogram (with the appearance of the alkylation adduct as a broad bump, further confirmed by MALDI-TOF, ESI Supplementary Figures S17–S19). In analogy to the G4 case, sequence specificity was observed, with a significantly lower consumption of the non-target sequence. Similarly, the influence of the spacing unit (number of AEEA spacers) was also not important for the alkylation outcome. In contrast to what was observed for G4, a different efficiency when targeting from the 5′ or the 3′ side of the iM-forming core is observed. Moreover, upon increasing the pH value from 5.0 to 8.0, the alkylation efficiency drops and at pH 8.0 the difference in DNA consumption between **k-RAS-T** and **k-RAS-NT** became less pronounced.

The (significant) difference in terms of DNA consumption when moving from the 5′ to the 3′ flank can be rationalized by considering the binding mode originally proposed for this family of ligands, which are suggested to be iM-groove binders. The formation of the PNA:DNA duplex on the 5′-side or 3′-side could in fact favour the ligand positioning on two different grooves and thus, create a different environment, explaining the different reactivity on the two sides. Although the lack of NMR or XRD data of ligand-iM complexes makes these observations speculative, the data obtained can support the binding mode proposed by Dash and colleagues for PBP-1 hypothesized from CD and fluorescence data (54).

This phenomenon was not observed when targeting the G4 with the PDC ligand: due to the end-stacking binding mode described for the molecule (and, by definition, its docking onto the upper tetrad), its orientation would not change when targeting different flanking regions (**PNA-G4-5L#** versus **PNA-G4-3L#** series).

The drop in alkylation efficiency with the pH increase can be explained by the progressive unfolding of the iM structure. Indeed, the folding of an iM is maximal at pH 5. In line with what was previously reported, the increase of pH leads to the formation of a iM/hairpin hybrid at pH 6 and complete unfolding at pH 8 (58). This was confirmed by CD analysis (Supplementary Figure S3, ESI). This confirms that the structure needs to be folded to allow the ligand to effectively target the secondary structure and correctly orient the alkylating warhead in a suitable position.

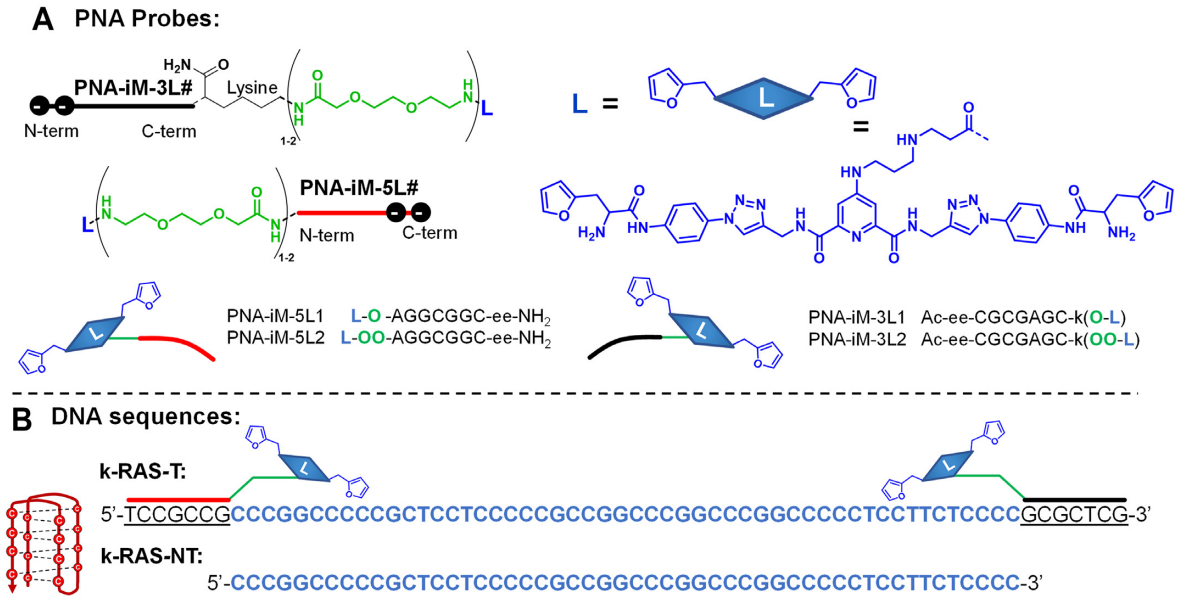

Figure 8. Design of the sequences used for the iM specific alkylation. (**A**) Structure and sequences of the PNA probes employed to target 5′- (red) or 3′-flanking (black) region of c-KIT. (**B**) Sequence of the *k-RAS* derivatives employed in this section. The iM-forming tract is highlighted in blue. O = (2-(2-aminoethoxy)ethoxyacetyl) spacer; L = iM ligand (derived from PBP-L3*, see ESI Supplementary scheme S4), # = number of spacers; F = Furanyl-propanoyl; e = glutamic acid; Ac = acetyl; k-RAS-T = Target; k-RAS-NT = Non-target.

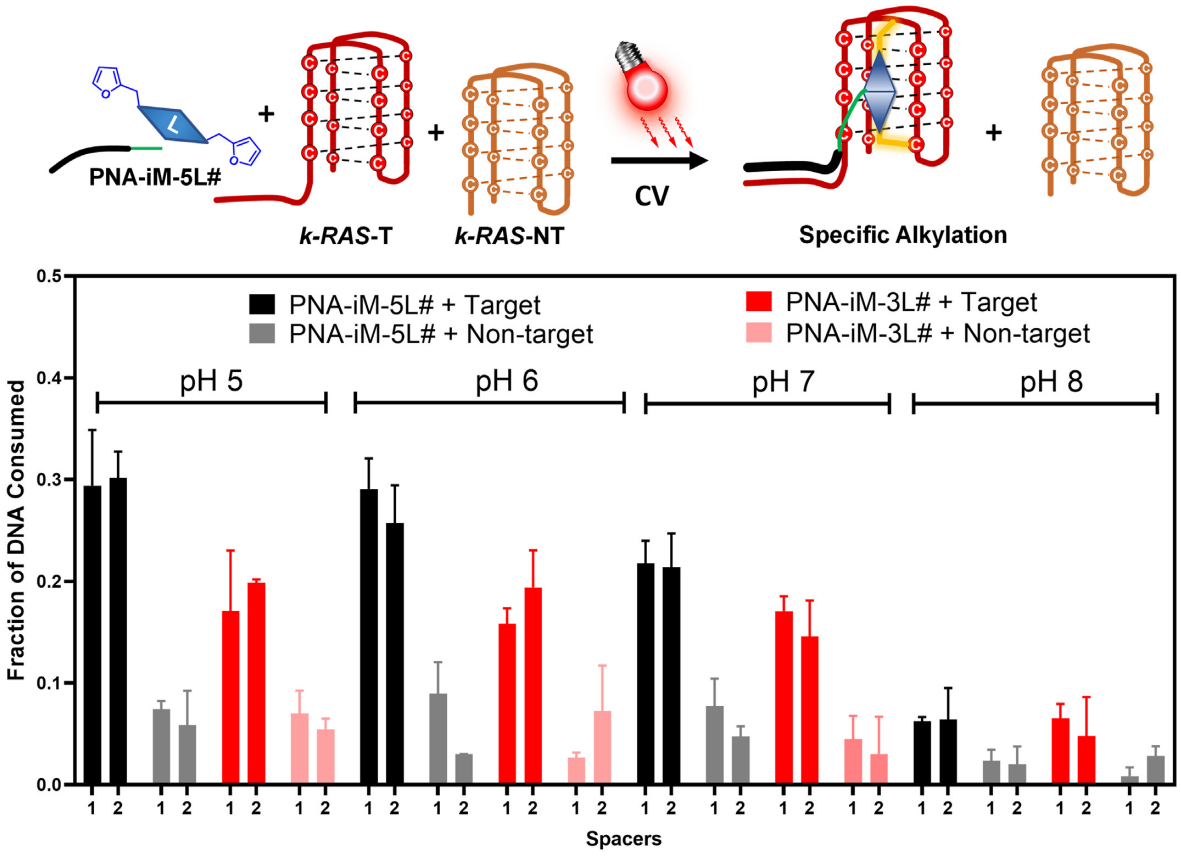

Figure 9. IM competition experiment for PNA-iM-5L# and PNA-iM-3L# in presence of competitor k-RAS-NT tested at different pH values (5, 6 and 7). T = target, NT = non-target, CV = crystal Violet, # = number of spacers, as indicated on the x-axis. Experiments performed in buffered solution (ammonium acetate pH 5, phosphate buffer pH 6 or 7, and PBS pH 8.0, all supplemented with 100 mM KCl), at 5 μM strand concentration.

Following a similar workflow as for the G4s, the alkylation efficiency towards other iM forming sequences (**k-RAS F**, **h-RAS** and **BCL-2**) was investigated. This was done using the same flanking region sequences as recognition elements. The data in Supplementary Figure S18 demonstrate that expanding the methodology to other sequences is possible, with moderate alkylation for each targeted structure. These results clearly indicate that, providing that the PNA can recognize the target flanking region, the ligand can anchor itself in the iM docking site and orient the furan in a suitable position.

CONCLUSIONS

In conclusion, we herein describe a sequence-specific tetraplex-alkylation methodology, which takes advantage of the recognition of G4 or iM flanking regions to target one unique structure in a mixture of different sequences. By incorporating selective G4 or iM binders in a suitable PNA probe, the alkylating moiety can be efficiently directed towards the envisaged target, further enhancing the selectivity of the methodology over non-tetraplex structures. To achieve this, in addition to the previously developed G4-alkylating ligands, we here developed the first example of a selective iM-alkylating ligand.

While there is a wealth of research data available on the development and use of small molecule ligands on the one hand and oligonucleotide-based probes on the other hand, the developed method rests on the rationally designed connection of both modalities, with further strengthening of the alkylation selectivity thanks to the inclusion of a triggerable warhead. The use of a pro-reactive furan moiety, activated by localized $^1O_2$ production, ensures this additional element of selectivity. In this work, we primarily focused on two families of ligands, but further extension of our work to other binders, as well as to alternative alkylating warheads, can be foreseen for future applications. Moreover, further fine-tuning of the methodology is possible in terms of the chosen co-localizing photosensitizers and the corresponding irradiation wavelength used for their activation (cfr. Red light for deeper tissue penetration). In view of the capacity of PNA to bind both DNA and RNA strands, this methodology can be conveniently adapted to target RNA G4 structures. In addition, the methodology can be further exploited for 'on-bench' applications, such as for the pull-down of specific G4 and iM-forming sequences, exploiting the reversibility of this type of alkylation at high temperatures (59).

Concerning future cellular applications, it is no secret that the delivery of oligonucleotides and their analogues has been challenging. In the past few years, however, many applications featuring PNAs have been successfully demonstrated in cells, thanks to the development of delivery strategies able to increase the cellular uptake of these probes (60–63). Also *in vivo* applications in the context of gene therapy have been described, e.g. for the correction of cystic fibrosis in murine models (64–66). Different PNA-based strategies have been developed in the field of G4-targeting, *in cellulo* and in pre-clinical murine models. An overview of these approaches is reported in our recent literature review (20). A further challenge for the success of the approach would be the tailored generation of reactive $^1O_2$ in a cellular environment. In this work, we performed the alkylation using photosensitizers that are not the best G4 or iM-binders. Even if some of these have already reached clinical applications (e.g. MB, which has a well-established clinical application in the context of photodynamic therapy (67)), fine-tuning the localized activation of the pro-reactive moiety will be needed to achieve high alkylation specificity without collateral oxidative damage and adverse toxicity effects. A potential solution could reside in adopting or developing more specific PS-binders, readily available for G4 (68), but not for iM. Furthermore, PS-PNA conjugates targeting the opposite flanking region could be advantageously used in this context. Indeed, limiting the production of $^1O_2$ to the nearby surroundings of the furan activation site has proven useful to prevent collateral oxidative damage, lowering the concentration of PS needed and ensuring, at the same time, the efficient activation of the furan moiety (47).

Given the high level of selectivity, off-target G4 or iM alkylation can be avoided. In view of the continuous discoveries of new G4 and iM structures and the interest in unravelling their biological function, we anticipate that the described methodology could find applications for the development of highly specific molecular tools contributing in the understanding of the role of specific G4 or iM structures. The presented approach expands the repertoire of available tools not only for specific G4-targeting, but also for iM targeting, which we believe to be necessary in the quest for elucidation of the biological role of these important, yet relatively underexplored, secondary structures.

DATA AVAILABILITY

The data underlying this article are available in the article and in is online supplementary material. Raw HPLC data will be shared on reasonable request to the corresponding author.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

The authors would like to thank Prof. Marleen Van Troys for the usage of the CD equipment, Stijn Tanghe and Jos Van De Begin for the technical support and Jack Barr for proofreading the manuscript.

FUNDING

L.D.P. and E.C. have received funding from the Research Foundation Flanders (FWO) [1S32122N to L.D.P., 12B1923N to E.C.]; This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska Curie grant agreement No 956070. European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie [721613 to E.C., 665501 to Al.M]; Al.M. has received funding from the UGent Industrieel OnderzoeksFonds [F2019/IOF-ConcepTT/188, F2020/IOF-ConcepTT/111]. Funding for open access charge: ITN

Oligomed; European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie [956070] and FWO.
Conflict of interest statement. None declared.

REFERENCES

1. Watson,J.D. and Crick,F.H.C. (1953) Molecular structure of nucleic acids: a structure for deoxyribose nucleic acid. *Nature*, **171**, 737–738.
2. Bochman,M.L., Paeschke,K. and Zakian,V.A. (2012) DNA secondary structures: stability and function of G-quadruplex structures. *Nat. Rev. Genet.*, **13**, 770–780.
3. Assi,H.A., Garavís,M., González,C. and Damha,M.J. (2018) I-motif DNA: structural features and significance to cell biology. *Nucleic Acids Res.*, **46**, 8038–8056.
4. Masoud,S.S. and Nagasawa,K. (2018) I-motif-binding ligands and their effects on the structure and biological functions of I-motif. *Chem. Pharm. Bull.*, **66**, 1091–1103.
5. Gehring,K., Leroy,J.L. and Guéron,M. (1993) A tetrameric DNA structure with protonated cytosine-cytosine base pairs. *Nature*, **363**, 561–565.
6. Wright,E.P., Huppert,J.L. and Waller,Z.A.E. (2017) Identification of multiple genomic DNA sequences which form i-motif structures at neutral pH. *Nucleic Acids Res.*, **45**, 2951–2959.
7. Zeraati,M., Langley,D.B., Schofield,P., Moye,A.L., Rouet,R., Hughes,W.E., Bryan,T.M., Dinger,M.E. and Christ,D. (2018) I-motif DNA structures are formed in the nuclei of human cells. *Nat. Chem.*, **10**, 631–637.
8. King,J.J., Irving,K.L., Evans,C.W., Chikhale,R.V., Becker,R., Morris,C.J., Peña Martinez,C.D., Schofield,P., Christ,D., Hurley,L.H. *et al.* (2020) DNA G-quadruplex and i-motif structure formation is interdependent in human cells. *J. Am. Chem. Soc.*, **142**, 20600–20604.
9. Cui,Y., Kong,D., Ghimire,C., Xu,C. and Mao,H. (2016) Mutually exclusive formation of G-quadruplex and i-motif is a general phenomenon governed by steric hindrance in duplex DNA. *Biochemistry*, **55**, 2291–2299.
10. Onizuka,K., Hazemi,M.E., Sato,N., Tsuji,G.I., Ishikawa,S., Ozawa,M., Tanno,K., Yamada,K. and Nagatsugi,F. (2019) Reactive OFF-ON type alkylating agents for higher-ordered structures of nucleic acids. *Nucleic Acids Res.*, **47**, 6578–6589.
11. Di Antonio,M., McLuckie,K.I.E. and Balasubramanian,S. (2014) Reprogramming the mechanism of action of chlorambucil by coupling to a G-quadruplex ligand. *J. Am. Chem. Soc.*, **136**, 5860–5863.
12. Nadai,M., Doria,F., Germani,L., Richter,S.N. and Freccero,M. (2015) A photoreactive G-quadruplex ligand triggered by green light. *Chemistry*, **21**, 2330–2334.
13. Di Antonio,M., Doria,F., Richter,S.N., Bertipaglia,C., Mella,M., Sissi,C., Palumbo,M. and Freccero,M. (2009) Quinone methides tethered to naphthalene diimides as selective G-quadruplex alkylating agents. *J. Am. Chem. Soc.*, **131**, 13132–13141.
14. Lena,A., Benassi,A., Stasi,M., Saint-Pierre,C., Freccero,M., Gasparutto,D., Bombard,S., Doria,F. and Verga,D. (2022) Photoactivatable V-shaped bifunctional quinone methide precursors as a new class of selective G-quadruplex alkylating agents. *Chem. – A Eur. J.*, **28**, e202200734.
15. Verga,D., Hamon,F., Poyer,F., Bombard,S. and Teulade-Fichou,M.P. (2014) Photo-cross-linking probes for trapping G-quadruplex DNA. *Angew. Chemie - Int. Ed.*, **53**, 994–998.
16. Cadoni,E., Manicardi,A., Fossépré,M., Heirwegh,K., Surin,M. and Madder,A. (2021) Teaching photosensitizers a new trick: red light-triggered G-quadruplex alkylation by ligand co-localization. *Chem. Commun.*, **57**, 1010–1013.
17. Juarranz,Á., Jaén,P., Sanz-Rodríguez,F., Cuevas,J. and González,S. (2008) Photodynamic therapy of cancer. Basic principles and applications. *Clin. Transl. Oncol.*, **10**, 148–154.
18. Asamitsu,S., Bando,T. and Sugiyama,H. (2019) Ligand design to acquire specificity to intended G-quadruplex structures. *Chem. - A Eur. J.*, **25**, 417–430.
19. Miglietta,G., Russo,M. and Capranico,G. (2020) G-quadruplex-R-loop interactions and the mechanism of anticancer G-quadruplex binders. *Nucleic Acids Res.*, **48**, 11942–11957.
20. Cadoni,E., De Paepe,L., Manicardi,A. and Madder,A. (2021) Beyond small molecules: targeting G-quadruplex structures with oligonucleotides and their analogues. *Nucleic Acids Res.*, **49**, 6638–6659.
21. Tassinari,M., Zuffo,M., Nadai,M., Pirota,V., Sevilla Montalvo,A.C., Doria,F., Freccero,M. and Richter,S.N. (2020) Selective targeting of mutually exclusive DNA G-quadruplexes: HIV-1 LTR as paradigmatic model. *Nucleic Acids Res.*, **48**, 4627–4642.
22. Chowdhury,S., Wang,J., Nuccio,S.P., Mao,H. and Di Antonio,M. (2022) Short LNA-modified oligonucleotide probes as efficient disruptors of DNA G-quadruplexes. *Nucleic Acids Res.*, **50**, 7247–7259.
23. Pagano,A., Iaccarino,N., Abdelhamid,M.A.S., Brancaccio,D., Garzarella,E.U., Di Porzio,A., Novellino,E., Waller,Z.A.E., Pagano,B., Amato,J. *et al.* (2018) Common G-quadruplex binding agents found to interact with i-motif-forming DNA: unexpected multi-target-directed compounds. *Front. Chem.*, **6**, 281.
24. Day,H.A., Pavlou,P. and Waller,Z.A.E.E. (2014) I-motif DNA: structure, stability and targeting with ligands. *Bioorg. Med. Chem.*, **22**, 4407–4418.
25. Singh,J., Petter,R.C., Baillie,T.A. and Whitty,A. (2011) The resurgence of covalent drugs. *Nat. Rev. Drug Discov.*, **10**, 307–317.
26. Boike,L., Henning,N.J. and Nomura,D.K. (2022) Advances in covalent drug discovery. *Nat. Rev. Drug Discov.*, **21**, 881–898.
27. Ghosh,A.K., Samanta,I., Mondal,A. and Liu,W.R. (2019) Covalent inhibition in drug discovery. *ChemMedChem*, **14**, 889–906.
28. Theil,K., Imami,K. and Rajewsky,N. (2019) Identification of proteins and miRNAs that specifically bind an mRNA in vivo. *Nat. Commun.*, **10**, 4205.
29. McHugh,C.A., Chen,C.K., Chow,A., Surka,C.F., Tran,C., McDonel,P., Pandya-Jones,A., Blanco,M., Burghard,C., Moradian,A. *et al.* (2015) The xist lncRNA interacts directly with SHARP to silence transcription through HDAC3. *Nature*, **521**, 232–236.
30. Chu,C., Zhang,Q.C., Da Rocha,S.T., Flynn,R.A., Bharadwaj,M., Calabrese,J.M., Magnuson,T., Heard,E. and Chang,H.Y. (2015) Systematic discovery of Xist RNA binding proteins. *Cell*, **161**, 404–416.
31. Minajigi,A., Froberg,J.E., Wei,C., Sunwoo,H., Kesner,B., Colognori,D., Lessing,D., Payer,B., Boukhali,M., Haas,W. *et al.* (2015) A comprehensive Xist interactome reveals cohesin repulsion and an RNA-directed chromosome conformation. *Science*, **349**, aab227.
32. Uetrecht,J. (2009) Immune-mediated adverse drug reactions. *Chem. Res. Toxicol.*, **22**, 24–34.
33. Sen,A. and Nielsen,P.E. (2007) On the stability of peptide nucleic acid duplexes in the presence of organic solvents. *Nucleic Acids Res.*, **35**, 3367–3374.
34. Tomac,S., Sarkar,M., Ratilainen,T., Wittung,P., Nielsen,P.E., Norde,B., Gra,A., Norden,B. and Graslund,A. (1996) Ionic effects on the stability and conformation of peptide nucleic acid complexes. *J. Am. Chem. Soc.*, **118**, 5544–5552.
35. Ratilainen,T., Holmén,A., Tuite,E., Haaima,G., Christensen,L., Nielsen,P.E. and Nordén,B. (1998) Hybridization of peptide nucleic acid †. *Biochemistry*, **37**, 12331–12342.
36. Sheng,Q., Neaverson,J.C., Mahmoud,T., Stevenson,C.E.M., Matthews,S.E. and Waller,Z.A.E. (2017) Identification of new DNA i-motif binding ligands through a fluorescent intercalator displacement assay. *Org. Biomol. Chem.*, **15**, 5669–5673.
37. Granotier,C., Pennarun,G., Riou,L., Hoffschir,F., Gauthier,L.R., De Cian,A., Gomez,D., Mandine,E., Riou,J.F., Mergny,J.L. *et al.* (2005) Preferential binding of a G-quadruplex ligand to human chromosome ends. *Nucleic Acids Res.*, **33**, 4182–4190.
38. Monchaud,D., Yang,P., Lacroix,L., Teulade-Fichou,M.P. and Mergny,J.L. (2008) A metal-mediated conformational switch controls G-quadruplex binding affinity. *Angew. Chemie - Int. Ed.*, **47**, 4858–4861.
39. Saintomé,C., Alberti,P., Guinot,N., Lejault,P., Chatain,J., Mailliet,P., Riou,J.F. and Bugaut,A. (2018) Binding properties of mono- and dimeric pyridine dicarboxamide ligands to human telomeric higher-order G-quadruplex structures. *Chem. Commun.*, **54**, 1897–1900.
40. Harikrishna,S., Kotaru,S. and Pradeepkumar,P.I. (2017) Ligand-induced conformational preorganization of loops of c-MYC

G-quadruplex DNA and its implications in structure-specific drug design. *Mol. Biosyst.*, **13**, 1458–1468.
41. Hwang,I.P., Mailliet,P., Hossard,V., Riou,J.F., Bugaut,A. and Roger,L. (2019) Investigating the effect of mono- and dimeric 360A G-quadruplex ligands on telomere stability by single telomere length analysis (STELA). *Molecules*, **24**, 577.
42. Cadoni,E., Magalhães,P.R., Emídio,R.M., Mendes,E., Vítor,J., Carvalho,J., Cruz,C., Victor,B.L. and Paulo,A. (2021) New (Iso)quinolinyl-pyridine-2,6-dicarboxamide G-quadruplex stabilizers. A structure-activity relationship study. *Pharmaceuticals*, **14**, 669.
43. Op De Beeck,M. and Madder,A. (2012) Sequence specific DNA cross-linking triggered by visible light. *J. Am. Chem. Soc.*, **134**, 10737–10740.
44. Hsu,S.-T.T.D., Varnai,P., Bugaut,A., Reszka,A.P., Neidle,S., Balasubramanian,S., Hsu,S.-T.T.D., Reszka,A.P., Varnai,P., Balasubramanian,S. *et al.* (2009) A G-rich sequence within the c-kit oncogene promoter forms a parallel G-quadruplex having asymmetric G-tetrad dynamics. *J. Am. Chem. Soc.*, **131**, 13399–13409.
45. Suchanek,M., Radzikowska,A. and Thiele,C. (2005) Photo-leucine and photo-methionine allow identification of protein-protein interactions in living cells. *Nat. Methods*, **2**, 261–267.
46. Cadoni,E., Rosa-Gastaldo,D., Manicardi,A., Mancin,F. and Madder,A. (2020) Exploiting double exchange diels-alder cycloadditions for immobilization of peptide nucleic acids on gold nanoparticles. *Front. Chem.*, **8**, 4.
47. Manicardi,A., Cadoni,E. and Madder,A. (2020) Visible-light triggered templated ligation on surface using furan-modified PNAs. *Chem. Sci.*, **11**, 11729–11739.
48. Seenisamy,J., Rezler,E.M., Powell,T.J., Tye,D., Gokhale,V., Joshi,C.S., Siddiqui-Jain,A. and Hurley,L.H. (2004) The dynamic character of the G-quadruplex element in the c-MYC promoter and modification by TMPyP4. *J. Am. Chem. Soc.*, **126**, 8702–8709.
49. Ambrus,A., Chen,D., Dai,J., Jones,R.A. and Yang,D. (2005) Solution structure of the biologically relevant G-quadruplex element in the human c-MYC promoter. Implications for G-quadruplex stabilization. *Biochemistry*, **44**, 2048–2058.
50. Cogoi,S. and Xodo,L.E. (2006) G-quadruplex formation within the promoter of the KRAS proto-oncogene and its effect on transcription. *Nucleic Acids Res.*, **34**, 2536–2549.
51. Agrawal,P., Lin,C., Mathad,R.I., Carver,M. and Yang,D. (2014) The major G-quadruplex formed in the human BCL-2 proximal promoter adopts a parallel structure with a 13-nt loop in k+ solution. *J. Am. Chem. Soc.*, **136**, 1750–1753.
52. Cheng,Y., Tang,Q., Li,Y., Zhang,Y., Zhao,C., Yan,J. and You,H. (2019) Folding/unfolding kinetics of G-quadruplexes upstream of the P1 promoter of the human BCL-2 oncogene. *J. Biol. Chem.*, **294**, 5890–5895.
53. Hazemi,M.E., Onizuka,K., Kobayashi,T., Usami,A., Sato,N. and Nagatsugi,F. (2018) Vinyldiaminotriazine-acridine conjugate as G-quadruplex alkylating agent. *Bioorg. Med. Chem.*, **26**, 3551–3558.
54. Debnath,M., Ghosh,S., Chauhan,A., Paul,R., Bhattacharyya,K. and Dash,J. (2017) Preferential targeting of i-motifs and G-quadruplexes by small molecules. *Chem. Sci.*, **8**, 7448–7456.
55. Ma,D.L., Kwan,M.H.T., Chan,D.S.H., Lee,P., Yang,H., Ma,V.P.Y., Bai,L.P., Jiang,Z.H. and Leung,C.H. (2011) Crystal violet as a fluorescent switch-on probe for i-motif: label-free DNA-based logic gate. *Analyst*, **136**, 2692–2696.
56. Simonsson,T., Pribylova,M. and Vorlickova,M. (2000) A nuclease hypersensitive element in the human c-myc promoter adopts several distinct i-tetraplex structures. *Biochem. Biophys. Res. Commun.*, **278**, 158–166.
57. Mergny,J.L., Lacroix,L., Hélène,C., Han,X. and Leroy,J.L. (1995) Intramolecular folding of pyrimidine oligodeoxynucleotides into an i-DNA motif. *J. Am. Chem. Soc.*, **117**, 8887–8898.
58. Kaiser,C.E., Van Ert,N.A., Agrawal,P., Chawla,R., Yang,D. and Hurley,L.H. (2017) Insight into the complexity of the i-motif and G-quadruplex DNA structures formed in the KRAS promoter and subsequent drug-induced gene repression. *J. Am. Chem. Soc.*, **139**, 8522–8536.
59. Cadoni,E., Pennati,F., Muangkaew,P., Elskens,J., Madder,A. and Manicardi,A. (2022) Synthesis and structure–activity relationship of peptide nucleic acid probes with improved interstrand-crosslinking abilities: application to biotin-mediated RNA-pulldown. *RSC Chem. Biol.*, **3**, 1129–1143.
60. Volpi,S., Cancelli,U., Neri,M. and Corradini,R. (2021) Multifunctional delivery systems for peptide nucleic acids. *Pharmaceuticals*, **14**, 14.
61. Roberts,T.C., Langer,R. and Wood,M.J.A. (2020) Advances in oligonucleotide drug delivery. *Nat. Rev. Drug Discov.*, **19**, 673–694.
62. Gait,M.J., Arzumanov,A.A., McClorey,G., Godfrey,C., Betts,C., Hammond,S. and Wood,M.J.A. (2019) Cell-penetrating peptide conjugates of steric blocking oligonucleotides as therapeutics for neuromuscular diseases from a historical perspective to current prospects of treatment. *Nucleic Acid Ther*, **29**, 1–12.
63. Perera,J.D.R., Carufe,K.E.W. and Glazer,P.M. (2021) Peptide nucleic acids and their role in gene regulation and editing. *Biopolymers*, **112**, e23460.
64. Piotrowski-Daspit,A.S., Barone,C., Lin,C.Y., Deng,Y., Wu,D., Binns,T.C., Xu,E., Ricciardi,A.S., Putman,R., Garrison,A. *et al.* (2022) In vivo correction of cystic fibrosis mediated by PNA nanoparticles. *Sci. Adv.*, **8**, 19–22.
65. Bahal,R., Ali McNeer,N., Quijano,E., Liu,Y., Sulkowski,P., Turchick,A., Lu,Y.-C., Bhunia,D.C., Manna,A., Greiner,D.L. *et al.* (2016) In vivo correction of anaemia in β-thalassemic mice by γpna-mediated gene editing with nanoparticle delivery. *Nat. Commun.*, **7**, 13304.
66. McNeer,N.A., Anandalingam,K., Fields,R.J., Caputo,C., Kopic,S., Gupta,A., Quijano,E., Polikoff,L., Kong,Y., Bahal,R. *et al.* (2015) Nanoparticles that deliver triplex-forming peptide nucleic acid molecules correct F508del CFTR in airway epithelium. *Nat. Commun.*, **6**, 6952.
67. Tardivo,J.P., Del Giglio,A., De Oliveira,C.S., Gabrielli,D.S., Junqueira,H.C., Tada,D.B., Severino,D., De Fátima Turchiello,R. and Baptista,M.S. (2005) Methylene blue in photodynamic therapy: from basic mechanisms to clinical applications. *Photodiagn. Photodyn. Ther.*, **2**, 175–191.
68. Kawauchi,K., Urano,R., Kinoshita,N., Kuwamoto,S., Torii,T., Hashimoto,Y., Taniguchi,S., Tsuruta,M. and Miyoshi,D. (2020) Photosensitizers based on G-quadruplex ligand for cancer photodynamic therapy. *Genes (Basel)*, **11**, 1340.