

■ Nitrogen Heterocycles

The Relation Between Position and Chemical Composition of Bis-Indole Substituents Determines Their Interactions with G-Quadruplex DNA

Bagineni Prasad,^[a] Rabindra Nath Das,^[a] Jan Jamroskovic,^[b] Rajendra Kumar,^[a] Mattias Hedenström,^[a] Nasim Sabouri,^{*[b]} and Erik Chorell^{*[a]}

Abstract: G-quadruplex (G4) DNA structures are linked to fundamental biological processes and human diseases, which has triggered the development of compounds that affect these DNA structures. However, more knowledge is needed about how small molecules interact with G4 DNA structures. This study describes the development of a new class of bis-indoles (3,3-diindolyl-methyl derivatives) and detailed studies of how they interact with G4 DNA using orthogonal assays, biophysical techniques, and computational studies. This revealed compounds that strongly bind and

stabilize G4 DNA structures, and detailed binding interactions which for example, show that charge variance can play a key role in G4 DNA binding. Furthermore, the structure–activity relationships generated opened the possibilities to replace or introduce new substituents on the core structure, which is of key importance to optimize compound properties or introduce probes to further expand the possibilities of these compounds as tailored research tools to study G4 biology.

Introduction

Single-stranded DNA sequences containing four repeats of guanines (normally two-three), bridged by random nucleotides, can fold into G4 DNA structures. In the G4 DNA structure, the guanines bind to each other by Hoogsteen hydrogen bonds and form a quartet of four guanines, one from each guanine stretch, called G-tetrads. The G-tetrads stack on top of each other to form strong π – π interactions and are further stabilized by central monovalent cations, such as potassium or sodium.^[1,2] G4 DNA structures can be more stable than the classical double-stranded helical DNA.^[3] However, the number of G-tetrads and the length and base composition of the nucleotides that bridge the guanine stretches, called loops, are highly important for the stability of the G4 DNA structure.^[4–9]


G4 DNA structures can form when DNA is single-stranded, for instance during cellular processes, such as, DNA replication, transcription, and DNA repair. Furthermore, superhelical stress,^[10] molecular crowding,^[11] and G4 DNA binding proteins can also favor their formation.^[12]


There are about 700 000 predicted G4 DNA structures in the human genome and many of these are evolutionary conserved,^[13,14] suggesting that they have important functions. One of the first reported biological implications of G4 DNA structures was their presence at the telomeres, the chromosomal ends.^[15] Since then, it has also been shown that G4 DNA structures can inhibit or enhance gene expression and G4 DNA motifs are present in every second human promoter, and most frequently at promoters of oncogenes and regulatory genes.^[16,17] Additionally, G4 DNA motifs are also present in the promoters of other organisms.^[18–20] G4 DNA structures have thus been linked to fundamental cellular events such as maintenance of the telomeres and transcription, but they also affect DNA replication progression and many proteins and enzymes are reported to bind and/or resolve G4 DNA structures.^[15,21–23]

G4 DNA structures are also connected to human diseases such as neurodegenerative diseases and different types of cancers.^[24–27] One of the most extensively studied G4 DNA structures in this regard is located in the *c-MYC* promoter, which is a transcriptional factor affecting cell proliferation that is over-expressed in many human cancers.^[28] Mutations that prevent formation of the G4 DNA structure in the *c-MYC* promoter downregulates transcription of this gene,^[29] and selective stabilization of this G4 DNA structure is therefore a potential che-

[a] Dr. B. Prasad, Dr. R. N. Das, Dr. R. Kumar, Dr. M. Hedenström, Dr. E. Chorell
Department of Chemistry, Umeå University
90187 Umeå (Sweden)
E-mail: erik.chorell@umu.se

[b] Dr. J. Jamroskovic, Dr. N. Sabouri
Department of Medical Biochemistry and Biophysics
Umeå University, 90187 Umeå (Sweden)
E-mail: nasim.sabouri@umu.se

 Supporting information and the ORCID identification number(s) for the author(s) of this article can be found under:
<https://doi.org/10.1002/chem.202000579>.

 © 2020 The Authors. Published by Wiley-VCH Verlag GmbH & Co. KGaA. This is an open access article under the terms of Creative Commons Attribution NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

motherapeutic strategy. The G4-forming region of the *c-MYC* promoter is a 27 nucleotide sequence, Pu27, that is important for regulating transcription of the *c-MYC* gene.^[30,31] The Pu27 sequence form parallel polymorphic G4 DNA structures but the predominant G4 DNA structure involved in *c-MYC* transcriptional regulation is captured in the modified and truncated versions of the Pu27 sequence, Pu24T (G13T) and Pu22 (G14T/G23T) *c-MYC* G4 DNA structures.^[32,33]

The prevalence and biological relevance of G4 DNA structures as well as their links to human diseases has resulted in a significant interest in the development of compounds that stabilize G4 DNA structures. However, although there are many examples of G4 DNA stabilizing compounds, the reports that fully explore the structure–activity relationships and correlates this to the compound’s interactions with G4 DNA structures are scarce. This type of information is of key importance to develop this field of research both towards tailored research tools and to efficiently explore G4 DNA structures as therapeutic targets.

Indoles are frequently occurring in drugs and biologically active compounds and are therefore considered to be privileged structures. Bis-indoles have been shown to efficiently bind and stabilize G4 structures with effects similar or even improved over the most well-known and efficient G4 stabilizing compound Phen-DC3.^[34–37] However, these reported bis-indoles are either linked at the indole nitrogen or in position 2, leaving bis-indoles linked at position 3 unexplored. Furthermore, these studies also show that flexibility is important in the design of G4 ligands and that diindolyl-methyl derivatives can be efficient G4 binders even though their most energetically favorable conformation is non-planar.^[35,36] In this work, we have thus designed and synthesized 3,3-diindolyl-methyl analogues with different side chains to achieve structure–activity relationships (SARs). Additionally, we have developed synthetic methods that allowed us to vary the side chains on the 3,3-bisindolyl analogues in all open positions for all side chains and for a combination of side chains to enhance the SAR study (Figure 1).

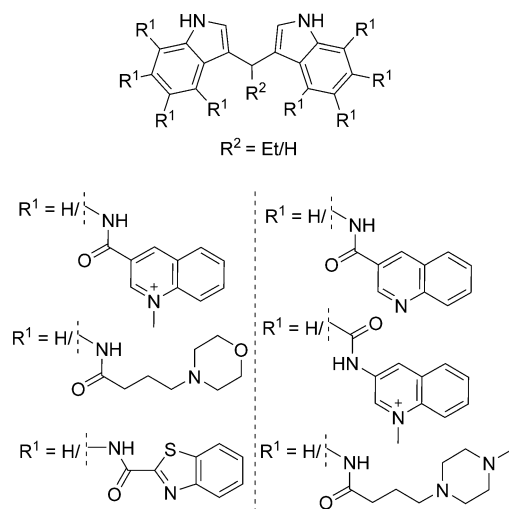


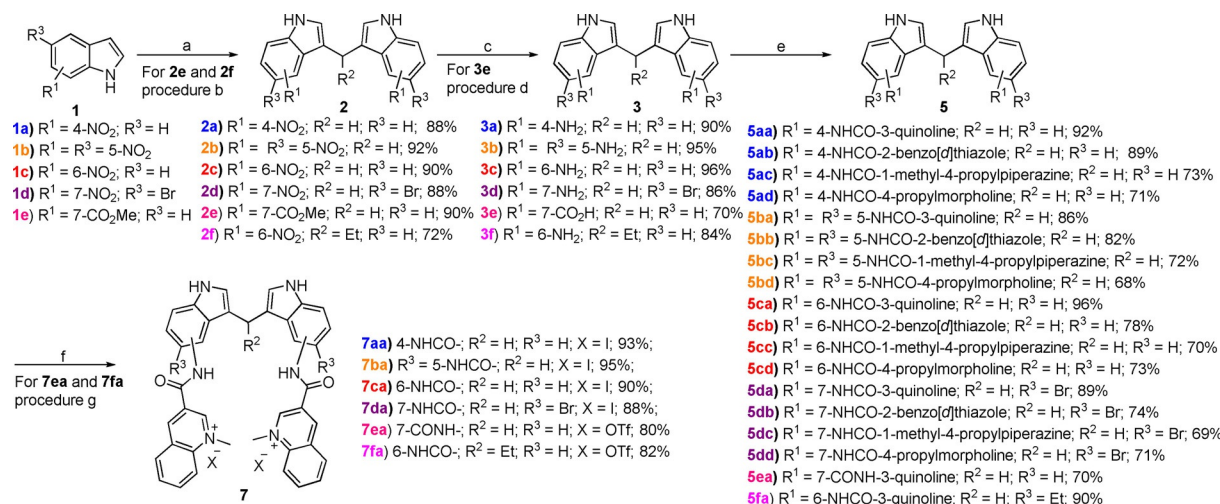
Figure 1. Summary of the 3,3-diindolyl-methyl central fragment and the variations of the substituents and their position.

To analyze the compounds’ abilities to bind and stabilize G4 DNA structures, we used a wide range of methods including fluorescence resonance energy transfer (FRET) melting assay, circular dichroism (CD) spectroscopy, microscale thermophoresis (MST), fluorescence intercalator displacement (FID) assay, nuclear magnetic resonance (NMR), and computational techniques such as molecular dynamics (MD) simulations to understand the molecular details of the results. Overall, this revealed compounds that strongly bind and stabilize G4 DNA structures and key components in the design of these types of compounds and their interactions with G4 DNA structures.

Results and Discussion

Synthesis

To determine how the position and composition of different side chains of the 3,3-diindolyl-methyl affect G4 binding and stabilization, we synthesized different analogues. The nitro-indoles (**1 a–d**; Scheme 1) were the key to investigate the effect of the position and composition of the side chains of the 3,3-diindolyl-methyl derivatives. The 4- and 5-nitro-indoles are commercially available whereas the 6- and 7-nitro-indoles were synthesized according to published procedures in two and five steps, respectively (supporting information). The 7-nitro-indole synthesis also resulted in a 5-bromo substituent for this set of derivatives (which depending on the data generated from these derivatives potentially can be used for late stage functionalization or be dehalogenated). From these four nitro-indoles, attempts to synthesize the di-nitro-3,3-diindolyl-methyl intermediates were performed by reacting the corresponding nitro-indoles with formaldehyde in water and 3 equivalents of acetic acid under refluxing conditions for 12 h. This gave the desired di-nitro-3,3-diindolyl-methyl intermediates (**2 a–c**) in 65–75% yields. Unfortunately, the 5-Bromo-7-nitro indole (**1 d**) did not react under these conditions. However, when the nitro indoles were reacted with formaldehyde using 10% sodium dodecyl sulfate (SDS) in water under microwave irradiation at 100 °C for 2 h, all the desired di-nitro-3,3-diindolyl-methyl intermediates (**2 a–d**) were obtained in 88–92% yields. Hydrogenation of the di-nitro-3,3-diindolyl-methyl analogues under Pd/C conditions in methanol at room temperature afforded di-amino-3,3-diindolyl-methyl analogues (**3 a–d**) in 86–96% yields. Next, the four di-amino-3,3-diindolyl-methyl analogues were each coupled with 4 different carboxylic acid substituted side chains using *N*-[(Dimethylamino)-1*H*-1,2,3-triazolo-[4,5-*b*]pyridin-1-ylmethylene]-*N*-methylmethanaminium hexafluorophosphate *N*-oxide (HATU) as coupling reagent and diisopropylethylamine (DIPEA) as base to provide the 16 desired analogues (**5 aa–ad**, **5 ba–bd**, **5 ca–cd**, and **5 da–dd**) in high yields. The choice of side chains was based on both aromatic and aliphatic systems with possibilities to form stacking interactions and/or positive charges under physiological pH or through permanent methylations. Hence, the analogues having quinolines as side chains were also quaternized using methyl iodide (MeI) in dimethylformamide (DMF) at 40 °C for 24 h to give the four methylated analogues (**7 aa–da**) in 88–96% yields. Unfortu-



Scheme 1. Synthesis of 3,3-diindolyl-methyl analogues. Reaction conditions: a) HCHO, 10% SDS in H₂O, MW, 100 °C, 2 h; b) EtCHO/HCHO, AcOH, H₂O, reflux, 12 h; c) Pd/C, H₂, Methanol, rt, 3 h; d) LiOH, THF, 2 h; e) R-COOH (quinoline-3-carboxylic acid (**4a**)/benzo[d]thiazole-2-carboxylic acid (**4b**)/4-(4-methylpiperazin-1-yl)butanoic acid (**4c**)/4-morpholinobutanoic acid (**4d**), HATU, DIPEA, DMF, rt or 50 °C, 12–24 h; f) MeI (**6a**), DMF, 40 °C, 24 h; g) MeOTf (**6b**), DCE, 85 °C, 3 h.

nately, methylations of the side chain containing benzo[d]thiazoles (**5ab**, **5bb**, **5cb**, and **5db**) were not successful even though various methods were applied (summarized in Scheme S1). To further expand the scope of the study, an analogue with the reversed amide (**5ea** and **7ea**) as well as an example with an ethyl substituent on the central carbon (**5fa** and **7fa**) was also synthesized (Scheme 1). In total 24 derivatives (**5aa–5fa** and **7aa–7fa**) with different side chain composition and position on the 3,3-diindolyl-methyl central fragment was prepared.

FRET Melting assay

We next analyzed how the variations of the side chains, both composition and position, affect the compounds ability to stabilize G4 DNA structures by the FRET melting assay. This assay measures the ligand-induced change in melting temperature (ΔT_m) of the labeled parallel Pu24T *c-MYC* G4 DNA structure (see Experimental Section for details, the oligonucleotides used in this study are summarized in Table S1). This showed that both the position and the composition of the side chains on the 3,3-diindolyl-methyl is crucial for an efficient G4 stabilization (Figure 2A and Figure S1). The derivatives with methylated quaternized quinoline side chains in position 5 and 6 (**7ba** and **7ca**) clearly gave the most efficient stabilization with a dramatic difference compared to analogues with substituents in position 4 and 7 (**7aa** and **7da**). Furthermore, the methylated quinoline side chain in the 5- and 6-substituted derivatives (**7ba** and **7ca**) proved to be of utmost importance for an efficient G4 stabilization when compared to their homologues with non-methylated quinoline side chains, suggesting that the charge is important for the quinoline side chains. This was supported by the derivatives with neutral benzo[d]thiazoles side chains (**5ab–5db**) that showed weak stabilization. However, the charge alone could not be responsible for the observed ef-

fects as the piperazine and morpholine derivatives (**5ac–5dc** and **5ad–5dd**, respectively), which are positively charged in the physiological pH used in the experiments, only showed weak stabilization. The piperazine and morpholine side chains lack the ability to form π -stacking interactions which suggest that this is important and cannot in full be compensated for by the potential electrostatic interactions of these derivatives. Reversing the amides (**7ea** compared to **7da**) did not have any major impact on the compounds ability to stabilize G4 DNA in this assay. Importantly, all compounds showed good selectivity for G4 DNA structures over double-stranded DNA (Figure S2). A potential problem with the FRET assay is the risk that compounds can interfere with the FRET process and thereby give false positive/negative results. However, these effects should appear immediately and not be temperature-dependent. None of the tested compounds displayed a temperature-independent effect on the FRET process thus suggesting that they do not affect the fluorophores or disrupt the G4 DNA structure (except **7fa** with a central ethyl group which affected the FRET process likely by interacting with the carboxytetramethylrhodamine (TAMRA) fluorophore, which can be the reason for the decreased stabilization ability observed for this compound compared to **7ca**).

Taq polymerase stop assay

The synthesized analogues were next evaluated for their ability to affect the progression of the Taq DNA polymerase on a DNA template with the Pu24T *c-MYC* G4 DNA structure and compared to a non-G4 control template. Due to their intrinsic stability, G4 DNA structures can form obstacles for DNA replication resulting in DNA synthesis pausing just before the G4 DNA structure on the template strand. Compounds that stabilize G4 DNA structures can thus be detected by an increased pausing before the G4 DNA structure and hence a reduced

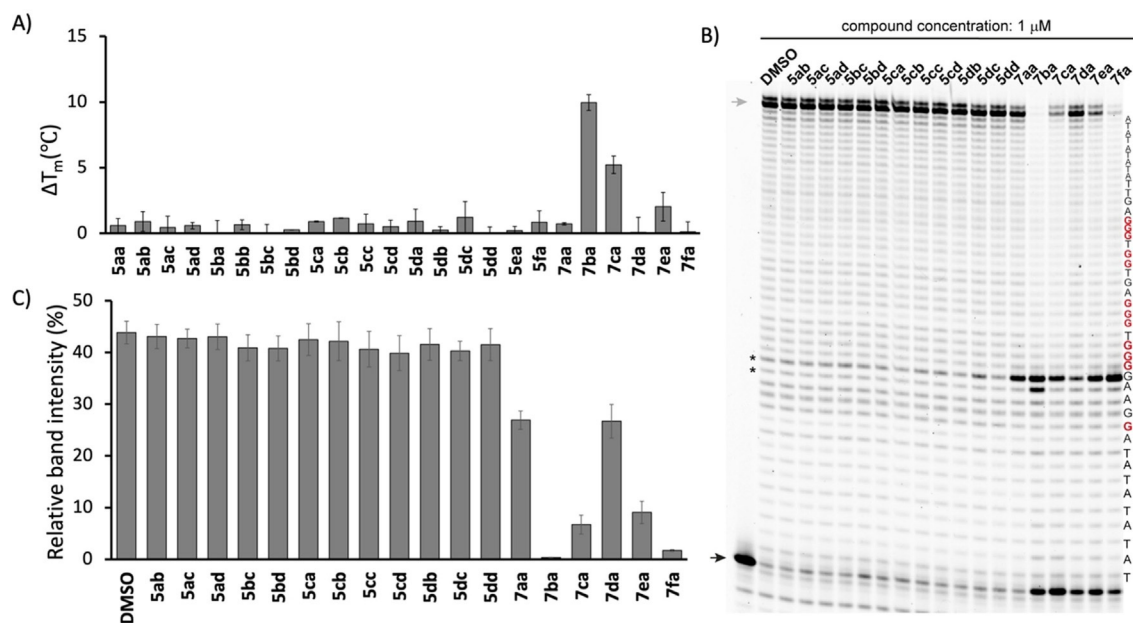


Figure 2. (A) Evaluation of the compounds ability to stabilize the Pu24T *c*-MYC G4 DNA structure using the FRET assay (at 2 μM ligand concentration). T_m in absence of ligands of Pu24T is 62.7 ± 0.3 °C. (B) Primary screening of the compounds (at 1 μM ligand concentration) by Taq-polymerase stop assay using Pu24T *c*-MYC G4 DNA template. *indicates pausing sites. Grey arrow represents full-length product, black arrow represent primer—start of the reaction. (C) Relative quantification of full-length product in each reaction. Error bars correspond to SD of at least three independent experiments.

amount of full-length products. Importantly, compounds that bind DNA non-selectively and therefore cause DNA replication arrest, can also be detected when running the experiments in parallel with a non-G4 control template. We first performed a primary screen that showed a clear G4 stabilization by **7ba**–**7fa** as indicated by a strong replication pausing one or two nucleotides before the first G-tract in the G4 structure (Figure 2B, denoted with asterisks; Figure S3B). **7ba**–**7fa** also displayed a clear selectivity for G4 DNA compared to the non-G4 control template (Figure 2 versus S4) although they also slightly suppressed primer utilization for both the G4 and non-G4 DNA (Figure 2B, Figures S3A, S4A, and S4C). In line with the FRET melting studies, none of the tested neutral 3,3-diindolylmethyl derivatives (**5**) showed G4 stabilization in this assay (Figure 2). Both the FRET melting studies and this assay clearly showed that **7ba** induced the strongest stabilization (Figure 2). The 4- and 7-substituted derivatives **7aa** and **7da** both have low effects, which likely can be attributed to steric hindrance to adopt a conformation that match the G-tetrad and also sug-

gest that the 5-bromo substituent likely is not of key importance for the effect of **7da**. In the Taq polymerase stop assay, we also detected strong pausing and reduced amount of full-length products by **7fa**, however this was not found with the FRET melting assay (Figure 2). Furthermore, both the central ethyl substituent and reversing the amide had only mild impact on the stabilization ability (**7fa** versus **7ca** and **7ea** versus **7da**, respectively) in this assay.

Next, we performed dose-response Taq polymerase stop assays to determine the efficiency and selectivity of **7aa**, **7ba**, and **7ca** by calculating the IC_{50} values (Table 1, Figures S5 and S6). All three compounds showed a clear dose response, and the IC_{50} for the G4 DNA template was 1.5, 0.17, and 0.24 μM for **7aa**, **7ba**, and **7ca**, respectively (Table 1, Figure S5). For the non-G4 DNA template, the IC_{50} values were > 5 , 2, and 1.7 μM , respectively (Table 1, Figure S6). These data demonstrate the lowest IC_{50} values for **7ba** and **7ca**, and a strong selectivity by both compounds for the G4 DNA template compared to the non-G4 DNA.

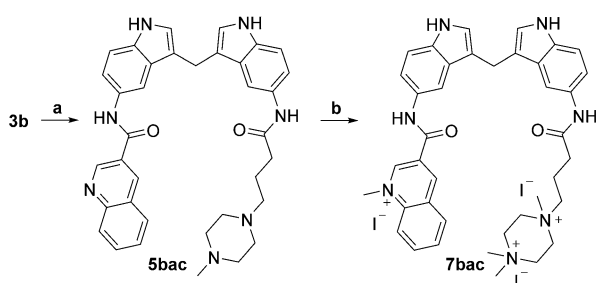
Table 1. Summary of binding and stabilization of the most promising 3,3-diindolyl-methyl derivatives.

Ligands	ΔT_m [°C] ^[a]			ΔT_m [°C] ^[b]	IC_{50} [μM] ^[c]		K_d [μM] ^[d]	Apparent K_d [μM] ^[e]
	1 μM	2 μM	5 μM		G4 DNA	dsDNA		
7aa	0.7 ± 0.2	0.9 ± 0.1	1.8 ± 0.3	5.3	1.55 ± 0.13	> 5	no saturation	no displacement
7ba	4.2 ± 0.8	10.9 ± 0.6	17.5 ± 1.5	4.0	0.17 ± 0.06	2.01 ± 0.3	0.49 ± 0.04	1.42 ± 0.06
7ca	0.8 ± 0.1	5.2 ± 0.6	10.4 ± 1.0	3	0.24 ± 0.06	1.72 ± 0.07	n.d.	n.d.
5bac	0	0.2 ± 0.1	0.6 ± 0.3	n.d.	n.d.	n.d.	no binding	no displacement
7bac	1.5 ± 0.8	8.8 ± 1.3	12.4 ± 0.9	> 12	0.19 ± 0.015	4.02 ± 0.45	3.05 ± 0.3	1.67 ± 0.05

[a] Using the FRET melting assay at the tabulated compound concentrations (see also Figure S1 and S10). [b] Using CD melting (see also Figure S8). [c] Using the Taq polymerase stop assay (see also Figure S5 and S6). [d] Using MST (see also Figure S11). [e] Using FID assay (see also Figures S12 and S13). n.d.-not determined.

Synthesis of asymmetrical compounds

Based on the data above, the positioning of the side chains on the 3,3-diindolyl-methyl central fragment proved to be highly important with position 5 giving the most efficient G4 DNA stabilizing compounds in both assays. Derivative **7ba** with methylated quinoline side chains in position 5 was the most promising derivative and other side chains in position 5 did not result in compounds with any G4 stabilizing capacity at the tested concentrations. However, we hypothesized that not both side chains are required for an efficient binding and stabilization, which thus would open for the possibility to use one of the side chain positions for example, improving compound properties or reaching additional electrostatic interactions with the phosphate backbone. Hence, a combined derivative with one quinoline and one piperazine side chain in position 5 of the 3,3-diindolyl-methyl central fragment was synthesized (Scheme 2).



Scheme 2. Synthesis of asymmetrical 3,3-diindolyl-methyl analogues. Reaction conditions: a) 4-(4-methylpiperazin-1-yl)butanoic acid (**4c**), quinoline-3-carboxylic acid (**4a**), HATU, DIPEA, DMF, rt, 12 h, 70% yield; b) MeI (**6a**), DMF, 40 °C, 24 h, 88% yield.

Many different strategies were tested to achieve these derivatives (Scheme S2). However, the successful synthetic method (Scheme 2) started with a coupling reaction between one amino group of 3,3'-methylenebis(1*H*-indol-5-amine) **3b** with 4-(4-methylpiperazin-1-yl)butanoic acid directly followed by the coupling of the other amino group with quinoline-3-carboxylic acid using HATU and DIPEA conditions to afford the unsymmetrical **5bac** in 70% yield. Subsequent methylation gave the tri-methylated compound **7bac** in 88% yield.

G4 stabilization by the asymmetrical analogues

Next, the 3,3-diindolyl-methyl derivatives with one quinoline and one piperazine side chain (**5bac** and **7bac**) were tested for their ability to affect the progression of the Taq DNA polymerase with Pu24T *c-MYC* G4 DNA and non-G4 DNA (Figure S7). Similar to the other **5** and **7** analogues, the charged **7bac** stabilized the G4 structure whereas the neutral **5bac** did not (Figure S7). To test the efficiency of **7bac** to selectively stabilize G4 structures, we also performed the dose-response Taq polymerase stop assay (Figures S5 and S6). **7bac** displayed a clear dose response and efficiently blocked the Taq polymerase

1–2 nucleotides before the G4 DNA structure. In fact, the 3,3-diindolyl-methyl derivative with one quinoline and one piperazine side chain, **7bac**, displayed an IC_{50} of 0.2 μM , showing that it is equally active as the derivative with two methylated quinolines (**7ba**) (Table 1). Furthermore, this asymmetric analogue demonstrates an impressive selectivity for the G4 template over the non-G4 template (20-fold).

To further confirm that these derivatives stabilize the Pu24T *c-MYC* G4 DNA structure, we used thermal CD spectroscopy. Thermal CD melting experiments again showed that the compounds with methylated side chains are able to stabilize the G4 DNA structure from thermal melting, whereas the derivative with only piperazine side chains (**5ac**) did not have any effect (Figure S8). Importantly, the CD studies also showed that the compounds did not affect the structural folding of the Pu24T G4 DNA structure (Figure S9). However, this assay did not give the same internal order of efficiency as the Taq DNA polymerase stop assay and showed that **7bac** (with one methylated quinoline and one piperazine side chain) stabilized the G4 DNA structure much more efficient than **7ba** (with two methylated quinolines). This can likely be explained by the difference in the compound's interactions with the G4 DNA (stacking versus electrostatic interactions) which may affect these assays differently. Furthermore, the FRET melting assay showed similar results as the Taq polymerase stop assay with high melting temperatures for both **7ba** and **7bac** (Table 1, Figure S10). The FRET assay further showed intermediate stabilization for **7ca** whereas both **5bac** (the non-methylated version of **7bac**) and **7aa** were unable to stabilize the Pu24T *c-MYC* G4 DNA structure from thermal denaturation.

Binding affinity studies

To investigate the compounds (**7aa**, **7ba**, **5bac**, and **7bac**) binding affinity to the Pu24T *c-MYC* G4 DNA structure, we used MST (Table 1, Figure S11). The 3,3-diindolyl-methyl derivative **7ba** with two methylated quinoline side chains in position 5 was the strongest binder in this assay with a K_d of 0.49 μM , whereas **7aa** with two methylated quinoline side chains in position 4 did not reach saturation at the highest concentration (10 μM). This again emphasize the importance of the correct positioning of the side chains on the central fragment for an efficient binding to the G4 DNA structure. As expected, the methylated **7bac** (with one quinoline and one piperazine side chain) displayed a much stronger binding affinity (3.05 μM) compared to the non-methylated homolog **5bac** which did not bind to the Pu24T *c-MYC* G4 DNA in this assay. However, the difference in binding affinity between **7ba** and **7bac** was surprisingly high. Therefore, the FID assay with thiazole orange (TO) was used to confirm these binding affinities (Table 1, Figures S12 and 13). This assay ranked the compounds in the same internal order, but showed a more similar binding affinity (1.4 μM and 1.7 μM for **7ba** and **7bac**, respectively).

NMR studies

To further probe the compounds' abilities to bind the Pu24T *c-MYC* G4 DNA structure, we performed NMR titration studies with **7aa**, **7ba**, **5bac**, and **7bac**, ranging from 0.1–5 equivalents and using the peaks in the imino-region of the spectra as probes for G4 interactions (Figure 3). This showed some clear differences between the compounds; **7bac** induced the biggest changes on the G4 structure with a strong effect even at a 0.1:1 compound to DNA ratio, and an almost complete disappearance of NMR signals at the higher concentrations was found. Based on the other assays with **7bac**, it is unlikely that this compound induced unfolding of the G4 DNA. Instead, the most likely explanation for the observed effect is exchange between free and bound form of the G4 DNA on an intermediate time-scale, resulting in line-broadening. Compound **7aa** also induced strong effects on the G4 DNA structure at substoichiometric concentrations whereas **7ba** required almost equimolar concentrations of compound to DNA before a clear effect could be observed. However, at these concentrations of **7ba**, a new set of peaks starts to appear that was not observed with the other compounds, suggesting that **7ba** have a stronger binding, resulting in a shift from intermediate to slow exchange on the NMR time-scale where peaks from both free and bound form of the DNA are observed. As expected, **5bac** did not induce any big changes. However, at equimolar and higher amounts of **5bac** compared to DNA, a clear effect on

the guanines at the 5'-G-tetrad was clearly visible indicating that this compound indeed binds G4 DNA structures although it does not stabilize them to the same extent as the other compounds. All compounds primarily affected the guanines on the 5'-G-tetrad although almost all guanines were affected at the higher concentrations of **7aa**, **7ba**, and **7bac**.

Molecular dynamics simulations

We next used molecular dynamics (MD) simulations to understand the molecular details of the observed G4 DNA binding and stabilization. This will also give valuable information about the binding mode of the 3,3-diindolyl-methyl derivatives with G4 DNA. Based on the NMR results, we modeled compounds **7aa**, **7ba**, **5bac**, and **7bac** at the 5'-G-tetrad of the Pu24T *c-MYC* G4 DNA structure and performed 1 μ s MD simulations of these complexes. The sampled binding modes for these compounds were subsequently clustered using principal component analysis (PCA).

Five clusters were obtained for **7aa** and contrary to the expected binding mode, all these clusters showed a partially self-stacked conformation of **7aa** with one indole stacking on a quinoline (Figure 4, Figure S14, and Table S2). This conformation reduced the stacking interaction of **7aa** with the top G-tetrad due to the loss of the planar compound conformation. The self-stacked conformation of **7aa** was also present in its unbound form thus highlighting the rigidity of this conforma-

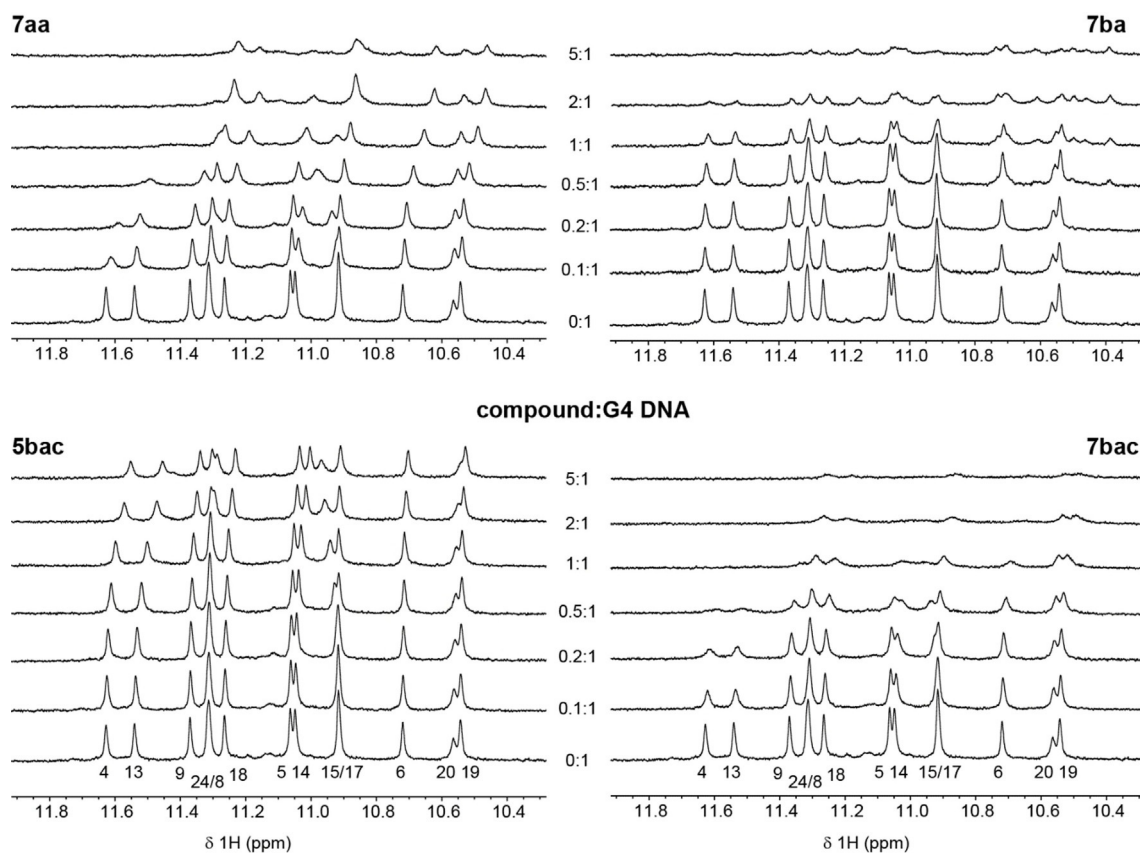


Figure 3. NMR spectra of Pu24T *c-MYC* G4 DNA in presence of different amount of **7aa**, **7ba**, **5bac**, and **7bac** (0.0:1, 0.1:1, 0.2:1, 0.5:1, 1:1, 2:1, 5:1 compound:G4 DNA). Numbers below peaks in the lower panel refers to the nucleotide position in the DNA sequence.

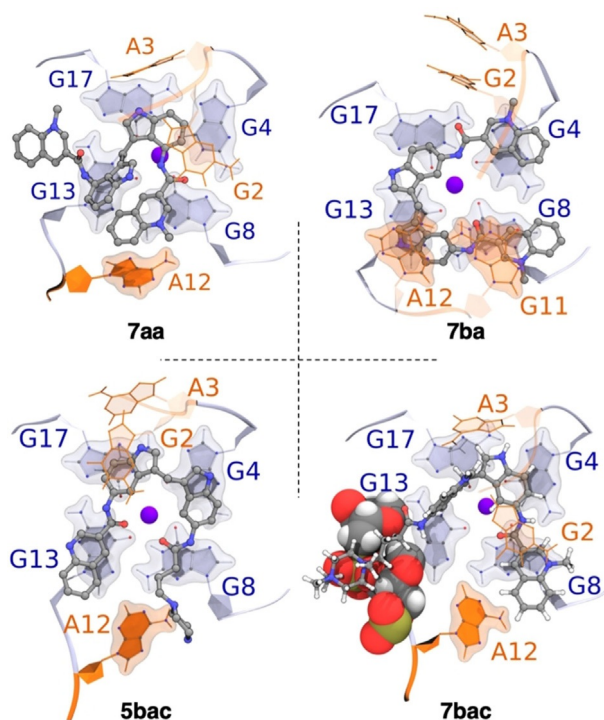


Figure 4. Representative binding poses of compounds during MD simulations. The 5'-G-tetrad (light blue), nucleotides flanking the G4 DNA structure (orange), and compounds (ball-stick model) from the largest cluster central structure **7aa**, **7ba**, **5bac**, and **7bac**. The interactions of **7bac** with the DNA backbone is highlighted with a space filling model. The potassium ion (central purple sphere) is also shown. The binding energies are tabulated in Table S6.

tion, which also can explain the high selectivity for **7aa** to G4 DNA over other types of DNA. The indole in **7aa** that is not involved in self-stacking and one of the quinolines were involved in stacking interactions with the top G-tetrad while the other free quinoline either made strong van der Waals interactions with the phosphate and sugar moiety or stacks with A-12 (first and second vs. third and fourth clusters, respectively).

The MD simulations with **7ba** could also be clustered into five main binding modes on the top G-tetrad (Figure 4, Figure S15, and Table S3). During 50% of the simulations (first and fourth clusters by occurrence), all four aromatic rings of **7ba** were stacked on top of the G-tetrad. **7ba** was also sandwiched by A-12 and G-11 from above in the first binding mode, which made this the most stable binding mode. When these additional stacking interactions were temporarily lost, **7ba** was able to rotate by 90 degrees to attain the fourth binding mode (fourth by occurrence, weakest binding affinity, Figure S15). Although **7ba** was stacked by A-3 from above, this fourth binding mode was flexible and could shift towards either the third or the fifth binding modes. This shift pushed both the quinoline rings outside of the G-tetrad and these rings either stacked with A-12 or interacted with phosphate and sugar moieties. The second binding mode was stabilized by strong stacking interactions with one quinoline and one indole group while the other quinoline ring remained outside of the G-tetrad. Overall, **7ba** could stack very efficiently with

all the four top guanines and was also further stabilized by stacking interactions with A-3, A-12, and G-11, however it remained flexible and was able to rotate on the top of the G-tetrad.

Because of the asymmetry in the 3-3'-bisindolylmethanes with two different side chains, **5bac** and **7bac**, we started with two different binding modes in these simulations, which gave seven and six binding modes from the MD simulations, respectively (Figure 4, Figures S16 and S17). One of the starting binding modes of **7bac**, where the piperazine is close to G-13, was surprisingly stable and this binding mode was populated at 42% of the simulation time which was the highest occurrence of all compounds (Figure S17 and Table S5). This binding mode also included interactions between the methylated piperazine and the DNA backbone and gave the highest binding affinity of all binding modes for all compounds. A similar starting binding mode for **5bac** did not give the same stable binding mode and yielded three different clusters (second, third and fourth clusters, Figure S16 and Table S4). The other starting binding mode in both **5bac** and **7bac** yielded four clusters. The obtained single largest cluster in **7bac** suggest that both the piperazine and the quinoline groups were interacting much stronger with the G4 DNA than that of **5bac**.

Overall, the obtained binding energy suggests that **7bac** is the strongest binder while **5bac** is a very weak binder (Table S6). However, the computed binding energy does not include entropic contribution and therefore, these cannot be directly compared with experimental values.

Additionally, when the quinoline side chain was methylated, the partial charge variance significantly decreased (Figure S18). This reduction in partial charge variance strengthen the stacking interaction with the guanines in the G-tetrad because guanine also possess large charge variance and two groups with large charge variance will thus repulse each other which will destabilize the stacking interaction. Therefore, a low partial charge variance is preferable for a strong stacking interaction with the guanines in the G-tetrad, as observed for the methylated quinoline in **7bac** compared to the non-methylated quinoline in **5bac**.

Selectivity studies

Finally, we investigated if **7ba** and **7bac** were able to bind and stabilize other types of G4 DNA structures in addition to the Pu24T *c-MYC* G4 DNA structure. A selection of 11 different G4 DNA structures were screened using the FRET assay and the results showed that **7ba** and **7bac** efficiently bind to many types of G4 structures (Figure 5A and Figure S19). The *c-MYC* Pu22 and Pu24T G4 DNA structures are very similar and it is thus not surprising that both **7ba** and **7bac** effectively stabilize both these structures. Furthermore, the compounds were also able to stabilize several other G4 DNA structures correlated to oncogenes, such as the *c-Kit* and *Bcl-2* G4 DNA structures.^[38,39] This can be a valuable attribute as it has been proposed that it in fact can be more effective to target several G4 DNA structures simultaneously in cancer therapy.^[40]

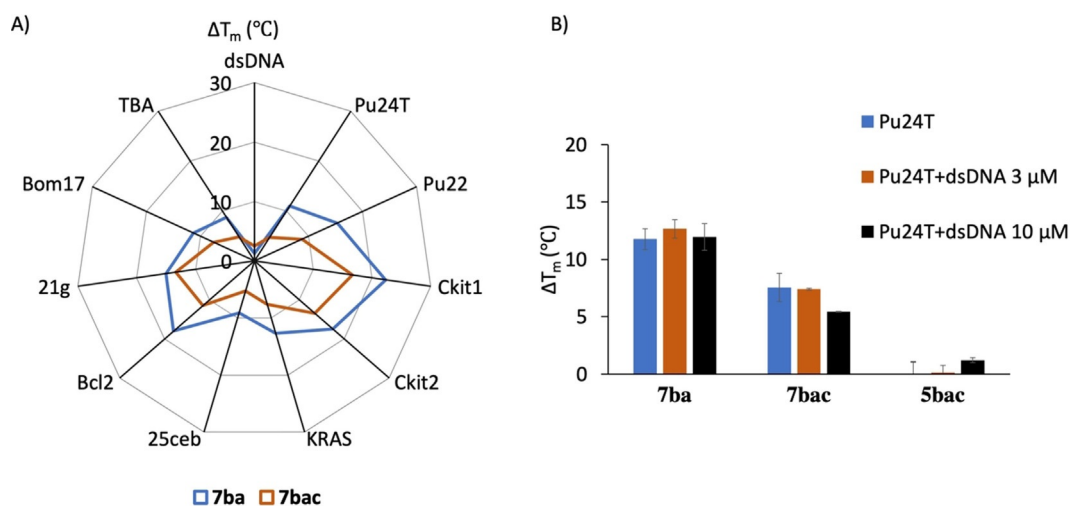


Figure 5. A) Selectivity study of **7ba** and **7bac** ($2 \mu\text{M}$) for their ability to stabilize different G4 DNA ($0.2 \mu\text{M}$) and dsDNA ($0.2 \mu\text{M}$) (see also Figure S19). Pu24T, Pu22 (*c-MYC* promoter); *c-kit1*, *c-kit2* (*c-Kit* promoter); *K-RAS* (*K-RAS* gene) and 25ceb (human minisatellite) are parallel G4 forming sequences. *Bcl-2* (*Bcl-2* promoter) and 21g (human telomere) are hybrid G4 forming sequences. Bom17 (*Bombyx* telomere) and TBA (thrombin binding aptamer) are antiparallel G4 forming sequences. ΔT_m could not be calculated for dsDNA. B) Stabilization of Pu24T *c-MYC* G4 DNA ($0.2 \mu\text{M}$) by the ligands **7ba**, **7bac**, and **5bac** ($2 \mu\text{M}$) in the presence of various concentrations of a double-stranded competitor dsDNA (ds26). Comparison of ΔT_m in the absence of competitor (blue) vs. ΔT_m in the presence of ds26 (brown and black bars).

To further study the selectivity of **7ba** and **7bac** to G4 DNA, we investigated if double-stranded DNA can compete out the binding of **7ba** and **7bac** to the G4 DNA structure. We therefore performed the FRET melting assay with G4 DNA, compound, and increasing concentrations of double-stranded DNA. This again showed the compound's selectivity for G4 DNA, as even 50-fold excess of double-stranded DNA hardly affected ΔT_m in the FRET melting studies (Figure 5B).

Conclusions

Bis-indoles are known to have the potential to efficiently bind and stabilize G4 DNA. In this work, we have developed a new bis-indole core structure and evaluated the effect of both the position and chemical composition of the side chains attached to this scaffold. Multistep synthetic method developments generated convenient methods to synthesize the target compounds, starting from readily available starting materials. Using these methods, a set of 26 target compounds were synthesized with varying substituents both in terms of their position on the core structure and their chemical composition. The compounds were evaluated using an array of different biochemical and biophysical techniques (FRET melting assay, CD, NMR, Taq polymerase stop assay, MST, FID) as well as computational techniques (MD simulations and charge variance calculations). This revealed compounds that bind and stabilize G4 DNA structures with well-defined dose-response curves and similar or even improved selectivity and activity compared to previously reported bis-indoles. Furthermore, this work also resulted in important structure-activity relationships that were correlated to the compound's interactions with the G4 DNA to understand why certain structural motifs are so important for the interactions with G4 DNA structures. For example, the methylated quinoline side chains generated the most efficient bis-

indoles in terms of G4 binding and stabilization, and the methylation of the quinoline proved crucial for this activity. The MD simulations and charge variance calculations show that the reason for this does not lie in the possibility for electrostatic interactions but rather the formation of a substituent with low charge variance that can stack efficiently on the top 5' G-tetrad. However, even though this charged quinoline is required for binding, it did not prove to be crucial to have two of these substituents as one could be replaced, while still retaining most of the binding and stabilization abilities. Moreover, the position of the substituents proved crucial for both the selectivity and stabilization ability with substitutions in position 5 being preferred as this has the ability to adopt a conformation that place the substituents in the right position to bind the G-tetrad. This type of detailed structure-activity relationships connected to a deep understanding of the compounds G4 DNA interactions open the possibilities to replace or introduce new substituents. This is of key importance to improve pharmacokinetic properties or to broaden the use of these G4 binding compounds through for example, the introduction of chemical handles for pull-down experiments or fluorophores and, ultimately, to better understand G4 biology and its therapeutic potential.

Experimental Section

Selected synthesis

Procedure for the preparation of *N,N'*-(3,3'-methylene-bis(1*H*-indole-4,3-diyl))bis(quinoline-3-carboxamide) (ECH-108) (5aa**):** HATU (322 mg, 0.85 mmol) was added to the solution of quinoline-3-carboxylic acid (**4a**) (140 mg, 0.80 mmol) in a dry DMF (3 mL) and the reaction mixture was stirred at room temperature until the solids were dissolved completely. DIPEA (155 mg, 1.2 mmol) was thereafter slowly added to the reaction mixture that was stirred for

10 min followed by slow addition of di-amine **3a** (110 mg, 0.40 mmol) in DMF (3 mL). The mixture was stirred at room temperature under nitrogen atmosphere for 12 h. The progress of the reaction was monitored by TLC. Upon completion, the reaction mixture was poured into ice and filtered through sintered funnel. The solid was dried under vacuum. The pure compound **5aa** was isolated in 92% yield as a light-yellow solid. ¹H NMR (600 MHz, [D₆]DMSO): δ = 10.93 (d, *J* = 2.5 Hz, 2H), 10.06 (s, 2H), 9.31 (s, 2H), 8.44 (d, *J* = 2.3 Hz, 2H), 8.04 (d, *J* = 8.4 Hz, 2H), 7.85 (t, *J* = 8.4 Hz, 2H), 7.69 (d, *J* = 8.1 Hz, 2H), 7.61 (t, *J* = 7.3 Hz, 2H), 7.25 (d, *J* = 7.7 Hz, 2H), 7.18 (d, *J* = 8.1 Hz, 2H), 7.03 (t, *J* = 7.8 Hz, 2H), 6.89 (d, *J* = 2.3 Hz, 2H), 4.62 ppm (s, 2H); ¹³C NMR (150 MHz, [D₆]DMSO): δ = 164.84, 149.22, 138.60, 136.11, 131.85, 130.28, 129.95, 128.47, 127.67, 127.57, 126.80, 124.54, 122.41, 121.47, 121.21, 116.37, 114.09, 110.23, 23.70 ppm; MS (ES mass): *m/z* 587.3 (*M*+1); HRMS: *m/z* calcd for C₃₇H₂₆N₆O₂[*M*+H]⁺ 587.2190, found 587.2180; Purity: 96.8% (according to HPLC).

Procedure for the preparation of *N,N'*-(3,3'-methylenebis(1*H*-indole-4,3-diyl))bis(benzo[d]thiazole-2-carboxamide) (ECH-109) (5ab): HATU (322 mg, 0.85 mmol) was added to the solution of benzo[d]thiazole-2-carboxylic acid (**4b**) (143 mg, 0.80 mmol) in a dry DMF (3 mL) and the reaction mixture was stirred at room temperature until the solids were dissolved completely. DIPEA (155 mg, 1.2 mmol) was thereafter slowly added to the reaction mixture that was stirred for 10 min followed by slow addition of di-amine **3a** (110 mg, 0.40 mmol) in DMF (3 mL). The mixture was stirred at room temperature under nitrogen atmosphere for 12 h. The progress of the reaction was monitored by TLC. Upon completion, the reaction mixture was poured into ice and filtered through sintered funnel. The solid was dried under vacuum. The pure compound **5ab** was isolated in 89% yield as a light-yellow solid. ¹H NMR (600 MHz, [D₆]DMSO): δ = 11.02 (s, 2H), 10.23 (s, 2H), 8.12 (d, *J* = 7.7 Hz, 2H), 7.55 (dd, *J* = 8.2, 4.5 Hz, 4H), 7.52–7.47 (m, 4H), 7.28 (d, *J* = 8.1 Hz, 2H), 7.12 (t, *J* = 7.9 Hz, 2H), 6.92 (d, *J* = 2.3 Hz, 2H), 4.74 ppm (s, 2H); ¹³C NMR (150 MHz, [D₆]DMSO): δ = 164.52, 158.14, 152.70, 138.72, 136.66, 129.74, 127.33, 127.29, 124.67, 123.21, 121.61, 120.61, 113.61, 113.50, 110.11, 24.52 ppm; MS (ES mass): *m/z* 599.2 (*M*+1); HRMS: *m/z* calcd for C₃₃H₂₂N₆O₂S₂[*M*+H]⁺ 599.1318, found 599.1298; Purity: 99.1% (according to HPLC).

Procedure for the preparation of *N,N'*-(3,3'-methylenebis(1*H*-indole-4,3-diyl))bis(4-(4-methylpiperazin-1-yl)butanamide) (ECH-110) (5ac): HATU (456 mg, 1.2 mmol) was added to the solution of 4-(4-methylpiperazin-1-yl)butanoic acid (**4c**) (149 mg, 0.80 mmol) in a dry DMF (3 mL) and the reaction mixture was stirred at room temperature until the solids were dissolved completely. DIPEA (155 mg, 1.2 mmol) was thereafter slowly added to the reaction mixture that was stirred for 10 min followed by slow addition of di-amine **3a** (110 mg, 0.40 mmol) in DMF (3 mL). The mixture was stirred at room temperature under nitrogen atmosphere for 12 h. The progress of the reaction was monitored by TLC. Upon completion, the reaction mixture was diluted with (15 mL) cold water and extracted with isopropanol and dichloromethane (3:1) (3 × 20 mL). The combined organic phases were washed with brine (50 mL), dried over anhydrous Na₂SO₄, filtered, and concentrated under a reduced pressure. The residue was purified by column chromatography over basic alumina using MeOH-DCM (0.1:9.9) to give *N,N'*-(3,3'-methylenebis(1*H*-indole-4,3-diyl))bis(4-(4-methylpiperazin-1-yl)butanamide) **5ac** in 73% yield as white solid. ¹H NMR (600 MHz, [D₆]DMSO): δ = 10.90 (s, 2H), 8.85 (s, 2H), 7.17 (t, *J* = 8.0 Hz, 4H), 7.02 (t, *J* = 7.9 Hz, 2H), 6.81 (s, 2H), 4.39 (s, 2H), 2.40–2.21 (m, 16H), 2.11 (s, 6H), 2.03 (d, *J* = 7.5 Hz, 4H), 1.90 (d, *J* = 8.1 Hz, 4H), 1.45 ppm (q, *J* = 7.5 Hz, 4H); ¹³C NMR (150 MHz, [D₆]DMSO): δ = 171.73, 138.64, 130.99, 124.30, 121.55, 120.92, 114.61, 113.78,

109.19, 57.60, 55.15, 52.99, 46.18, 34.35, 23.73, 22.61 ppm; MS (ES mass): *m/z* 613.2 (*M*+1); HRMS: *m/z* calcd for C₃₅H₄₈N₈O₂[*M*+H]⁺ 613.3973, found 613.3998; Purity: 99.1% (according to HPLC).

Procedure for the preparation of *N,N'*-(3,3'-methylenebis(1*H*-indole-4,3-diyl))bis(4-mor-pholinobutanamide) (ECH-111) (5ad): HATU (456 mg, 1.2 mmol) was added to the solution of 4-morpholinobutanoic acid (**4d**) (138 mg, 0.80 mmol) in a dry DMF (3 mL) and the reaction mixture was stirred at room temperature until the solids were dissolved completely. DIPEA (155 mg, 1.2 mmol) was thereafter slowly added to the reaction mixture that was stirred for 10 min followed by slow addition of di-amine **3a** (110 mg, 0.40 mmol) in DMF (3 mL). The mixture was stirred at room temperature under nitrogen atmosphere for 12 h. The progress of the reaction was monitored by TLC. Upon completion, the reaction mixture was diluted with (15 mL) cold water and extracted with isopropanol and dichloromethane (3:1) (3 × 20 mL). The combined organic phases were washed with brine (50 mL), dried over anhydrous Na₂SO₄, filtered, and concentrated under a reduced pressure. The residue was purified by column chromatography over basic alumina using MeOH-DCM (0.1:9.9) to give *N,N'*-(3,3'-methylenebis(1*H*-indole-4,3-diyl))bis(4-(4-methylpiperazin-1-yl)butanamide) **5ad** in 71% yield as white solid. ¹H NMR (600 MHz, [D₆]DMSO): δ = 10.90 (s, 2H), 8.87 (s, 2H), 7.19–7.16 (m, 4H), 7.02 (t, *J* = 7.8 Hz, 2H), 6.82 (s, 2H), 4.39 (s, 2H), 3.50 (t, *J* = 4.7 Hz, 8H), 2.33–2.21 (m, 8H), 2.07–2.04 (m, 4H), 1.92 (t, *J* = 7.4 Hz, 4H), 1.47 ppm (t, *J* = 7.4 Hz, 4H); ¹³C NMR (150 MHz, [D₆]DMSO): δ = 171.66, 138.64, 130.96, 124.35, 121.57, 120.95, 114.64, 113.78, 109.24, 66.46, 57.92, 53.55, 34.18, 23.72, 22.05 ppm; MS (ES mass): *m/z* 587.2 (*M*+1); HRMS: *m/z* calcd for C₃₃H₄₂N₆O₄[*M*+H]⁺ 587.3340, found 587.3329; Purity: 98.2% (according to HPLC).

Procedure for the preparation of 3,3'-(((3,3'-methylenebis(1*H*-indole-4,3-diyl))bis(azanedi-yl))bis(carbonyl))bis(1-methylquinolin-1-ium) iodide (ECH-112) (7aa): To the compound **5aa** (30 mg, 0.05 mmol) dissolved in DMF (2 mL) was added methyl iodide (**6a**) (212 mg, 1.5 mmol) drop wise at room temperature under nitrogen atmosphere. The reaction mixture was stirred at room temperature for 24 h. The progress of the reaction was monitored by LC-MS. Upon completion, the reaction mixture was precipitated by slow addition of acetone. The precipitate was filtered through a sintered funnel and washed with diethyl ether and acetone. The solid was dried under vacuum to obtain pure compound **7aa** as a light brown solid in 93% yield. ¹H NMR (600 MHz, [D₆]DMSO): δ = 10.97 (s, 2H), 10.37 (s, 2H), 9.72 (s, 2H), 9.26 (s, 2H), 8.53 (d, *J* = 8.9 Hz, 2H), 8.37 (m, 2H), 8.29 (d, *J* = 8.9 Hz, 2H), 8.08 (t, *J* = 7.6 Hz, 2H), 7.13 (d, *J* = 8.0 Hz, 2H), 6.85–6.82 (m, 4H), 6.78 (d, *J* = 7.4 Hz, 2H), 4.67 (s, 6H), 4.40 ppm (s, 2H); ¹³C NMR (150 MHz, [D₆]DMSO): δ = 161.37, 149.43, 144.46, 138.65, 138.02, 136.86, 131.61, 130.34, 128.39, 127.82, 127.08, 124.67, 122.53, 120.80, 119.11, 116.64, 113.53, 110.29, 45.73, 34.39 ppm; MS (ES mass): *m/z* 615.2 (*M*–1); HRMS: *m/z* calcd for C₃₉H₃₂N₆O₂[*M*–H]⁺ 616.2581, found 616.2517; Purity: 98.7% (according to HPLC).

Procedure for the preparation of 3,3'-(((3,3'-methylenebis(1*H*-indole-5,3-diyl))bis(azanedi-yl))bis(carbonyl))bis(1-methylquinolin-1-ium) iodide (ECH-113) (7ba): To the compound **5ba** (30 mg, 0.05 mmol) dissolved in DMF (2 mL) was added methyl iodide (212 mg, 1.5 mmol) drop wise at room temperature under nitrogen atmosphere. The reaction mixture was stirred at room temperature for 24 h. The progress of the reaction was monitored by LC-MS. Upon completion, the reaction mixture was precipitated by slow addition of acetone. The precipitate was filtered through a sintered funnel and washed with diethyl ether and acetone. The solid was dried under vacuum to obtain pure compound **7ba** as a light brown solid in 95% yield. ¹H NMR (600 MHz, [D₆]DMSO): δ = 10.89

(s, 2H), 10.79 (s, 2H), 10.01 (s, 2H), 9.80 (s, 2H), 8.59–8.56 (m, 4H), 8.41–8.38 (m, 2H), 8.15 (t, $J=7.6$ Hz, 2H), 8.01 (d, $J=1.9$ Hz, 2H), 7.49 (dd, $J=8.7, 2.0$ Hz, 2H), 7.42 (d, $J=8.6$ Hz, 2H), 7.15 (d, $J=2.2$ Hz, 2H), 4.73 (s, 6H), 4.17 ppm (s, 1H); ^{13}C NMR (150 MHz, $[\text{D}_6]\text{DMSO}$): $\delta=160.60, 150.53, 145.66, 139.09, 137.23, 134.53, 131.80, 131.12, 130.01, 129.04, 128.56, 127.40, 124.51, 119.84, 116.08, 114.41, 111.96, 111.37, 46.07, 21.61$ ppm; MS (ES mass): m/z 615.2 ($M-1$); HRMS: m/z calcd for $\text{C}_{39}\text{H}_{32}\text{N}_6\text{O}_2$ [$M-H$] $^+$ 615.2503, found 615.2517; Purity: 99.3% (according to HPLC).

Procedure for the preparation of *N*-(3-((4-(4-methylpiperazin-1-yl)butanamido)-1*H*-indol-3-yl)methyl)-1*H*-indol-4-yl)quinoline-3-carboxamide (ECH-114) (5 bac): DIPEA (155 mg, 1.2 mmol) was added to a solution of di-amine **3b** (110 mg, 0.40 mmol) in a dry DMF (2 mL) and then 4-(4-methylpiperazin-1-yl)butanoic acid (**4c**) (59 mg, 0.32 mmol) and HATU (125 mg, 0.33 mmol) in a dry DMF (2 mL) was slowly added drop wise over 6 h under nitrogen atmosphere at room temperature. Next, quinoline-3-carboxylic acid (**4a**) (69 mg, 0.4 mmol) and HATU (152 mg, 0.4 mmol) in a dry DMF (2 mL) was added drop wise to the reaction mixture. The mixture was stirred at room temperature under nitrogen atmosphere for 2 h. The progress of the reaction was monitored by LCMS. Upon completion, the reaction mixture was diluted with (15 mL) cold water and extracted with isopropanol and dichloromethane (3:1) (3×20 mL). The combined organic phases were washed with brine (50 mL), dried over anhydrous Na_2SO_4 , filtered, and concentrated under reduced pressure. The residue was purified by column chromatography over basic alumina using MeOH-DCM (0.1:9.9) to give *N,N'*-(3,3'-methylenebis(1*H*-indole-4,3-diy))bis(4-(4-methylpiperazin-1-yl)butanamide) **5 bac** in 70% yield as a white solid. ^1H NMR (600 MHz, $[\text{D}_6]\text{DMSO}$): $\delta=10.78$ (s, 1H), 10.67 (s, 1H), 10.45 (s, 1H), 9.62 (s, 1H), 9.39 (s, 1H), 8.97 (s, 1H), 8.15 (d, $J=8.25$ Hz, 1H), 8.12 (d, $J=8.32$ Hz, 1H), 8.01 (s, 1H), 7.90 (t, $J=7.8$ Hz, 1H), 7.85 (s, 1H), 7.73 (t, $J=7.6$ Hz, 1H), 7.50 (d, $J=8.8$ Hz, 1H), 7.35 (d, $J=8.6$ Hz, 1H), 7.23 (s, 2H), 7.06 (d, $J=18.1$ Hz, 2H), 4.08 (s, 2H), 2.34–2.22 (m, 8H), 2.14 (s, 3H), 1.71 (t, $J=7.3$ Hz, 2H), 1.03 ppm (s, 4H); ^{13}C NMR (150 MHz, $[\text{D}_6]\text{DMSO}$): $\delta=170.85, 163.96, 149.67, 148.87, 136.12, 134.17, 133.60, 131.60, 131.32, 130.67, 129.60, 129.25, 128.54, 127.86, 127.38, 127.33, 127.00, 124.13, 123.89, 118.79, 116.31, 115.25, 114.58, 114.36, 111.57, 111.45, 111.37, 109.75, 57.78, 55.14, 53.01, 46.13, 46.09, 34.76, 22.94, 21.44$ ppm; MS (ES mass): m/z 600.4 ($M+1$); HRMS: m/z calcd for $\text{C}_{36}\text{H}_{37}\text{N}_7\text{O}_2$ [$M+H$] $^+$ 600.3081, found 600.3088; Purity: 97.1% (according to HPLC).

Procedure for the preparation of 1,1,4-trimethyl-4-((3-((4-(1-methylquinolin-1-ium-3-carboxamido)-1*H*-indol-3-yl)methyl)-1*H*-indol-4-yl)amino)-4-oxobutyl)piperazine-1,4-dium iodide (ECH-115) (7 bac): To the compound **5 bac** (30 mg, 0.05 mmol) dissolved in DMF (2 mL) was added methyl iodide (212 mg, 1.5 mmol) drop wise at room temperature under nitrogen atmosphere. The reaction mixture was stirred at room temperature for 24 h. The progress of the reaction was monitored by LC-MS. Upon completion, the reaction mixture was precipitated by slow addition of acetone. The precipitate was filtered through a sintered funnel and washed with diethyl ether and acetone. The solid was dried under vacuum to obtain pure compound **7 bac** as a light brown solid in 88% yield. ^1H NMR (600 MHz, $[\text{D}_6]\text{DMSO}$): $\delta=10.87$ (s, 1H), 10.78 (s, 1H), 10.72 (s, 1H), 10.01 (s, 1H), 9.81 (d, $J=4.3$ Hz, 2H), 8.61–8.57 (m, 2H), 8.41 (t, $J=8.0$ Hz, 1H), 8.16 (t, $J=7.7$ Hz, 1H), 7.98 (s, 1H), 7.84 (s, 1H), 7.48 (d, $J=8.7$ Hz, 1H), 7.42 (d, $J=8.7$ Hz, 1H), 7.29–7.25 (m, 2H), 7.10 (s, 1H), 7.04 (s, 1H), 4.74 (s, 3H), 4.10 (s, 2H), 3.85 (s, 6H), 3.64 (dd, $J=10.8, 5.8$ Hz, 2H), 3.38 (s, 8H), 3.30 (s, 2H), 3.26 (s, 3H), 2.04 ppm (s, 2H); ^{13}C NMR (150 MHz, $[\text{D}_6]\text{DMSO}$): $\delta=169.17, 160.58, 150.51, 145.66, 139.12, 137.25, 134.50, 133.78, 131.82, 131.14, 131.00, 130.01, 129.03, 128.57, 127.40, 127.34, 124.45,$

124.06, 119.87, 116.05, 115.29, 114.44, 114.06, 111.96, 111.63, 111.32, 109.78, 54.87, 53.75, 46.09, 34.87, 32.52, 26.67, 21.58, 17.65 ppm; MS (ES mass): m/z 642.2 ($M-2$); HRMS: m/z calcd for $\text{C}_{39}\text{H}_{44}\text{N}_7\text{O}_2$ [$M-2H$] $^+$ 642.3551, found 642.3536; Purity: 96.9% (according to HPLC).

Methods

Folding of G4 structures for FRET study: Synthetic oligonucleotides for FRET study were purchased from Eurofins Genomics. Stock solutions were prepared in water at 100 μM concentration. The sequences used are listed in supplementary Table S1. All the oligonucleotides except Pu22 were prefolded in 10 mM lithium cacodylate buffer (pH 7.4), with 10 mM KCl and 90 mM LiCl by heating for 5 min at 95 $^\circ\text{C}$ and then quick cooling on ice. Pu22 was folded in 10 mM lithium cacodylate buffer (pH 7.4), with 2 mM KCl and 98 mM LiCl.

FRET melting assay: The fluorescence resonance energy transfer (FRET) occurs between two dyes (5'-FAM as donor and 3'-TAMRA as acceptor) linked at both extremities of a DNA oligonucleotide. When the oligonucleotides are folded into G4 structures, the donor and acceptor are in close proximity, which results in an energy transfer from the donor to the acceptor. This process can be detected by a reduction in the fluorescence emission of the donor. Fluorescence emission of the donor is recovered when the temperature increment triggers the thermal denaturation of the G4 structure. The experiments were performed in a Bio-rad CFX96 real-time PCR device at temperatures from 10 to 95 $^\circ\text{C}$ with a heating rate of 1.5 $^\circ\text{C min}^{-1}$ using a 492-nm excitation wavelength and a 516-nm detection wavelength in 96-well plates. Each condition was tested in duplicate and analysis of the data was carried out by using Excel and Origin 8 software. In each well, 200 nM of labelled oligonucleotide was heated in the presence or absence of the ligand (and with or without the competitor dsDNA) at the specified concentrations. Emission of 5'-FAM was normalized between 0 and 1, and the melting temperature (T_m) is defined as the temperature at which 50% of the G4 structures are denatured (the temperature when the normalized emission was 0.5). The stabilization (ΔT_m) is calculated from comparison of T_m of the fluorescently labeled oligonucleotide in the presence or absence of the ligand.

Taq-polymerase stop assay: 1 μM TET-labeled primer was annealed to 1.25 μM template DNA in 50 mM KCl by heating the reaction at 95 $^\circ\text{C}$ for 5 min and slowly cool down overnight to room temperature. Folding of the G4 structure and primer annealing proceeded simultaneously. For each reaction in the screening procedure, 40 nM annealed DNA was incubated with 1 μM compounds or 1% (v/v) DMSO (control reaction) for at least 30 min in 50 mM KCl, 10 mM Tris-HCl (pH 7.5), 1.5 mM MgCl_2 , and 200 mM dNTPs at room temperature. DNA synthesis was started by the addition of 0.625 U of recombinant Taq DNA-polymerase (Thermo Scientific) into the reaction mixture (10 μL), and incubated 30 min at 50 $^\circ\text{C}$. The reaction was stopped by addition of one volume of stop solution (95% formamide, 20 mM ethylenediaminetetraacetic acid (EDTA), and 0.1% bromophenol blue) into the reaction. A total of 8 μL of the mixture was loaded into a 10% polyacrylamide gel containing 8 M urea, 25% formamide, and 1 \times tris/borate/ethylenediaminetetraacetic acid (TBE). The gel was visualized with a Typhoon Scanner 9400 (GE Healthcare) at $\lambda=532$ nm and quantified with the Image Quant TL 8.1 software (GE Healthcare). Quantifications are displayed as an average of two experiments \pm absolute error. For the determination of IC_{50} values, the DNA was incubated with different concentrations (5, 2, 0.8, 0.32, 0.13, 0.05, 0.025, and 0.008 μM) of the tested compounds or 5% (v/v) DMSO. IC_{50} values

were calculated by fitting the full-length products into the dose response function in the OriginPro 2016 software. All reactions were independently repeated three times and standard deviation was calculated.

CD melting experiment: 50 μM folded oligonucleotides diluted to 5 μM concentration into 10 mM Tris (pH 7.5), 10 mM KCl and 10 μM compounds or 1.25% DMSO were used for CD melting studies. A quartz cuvette with a path length of 1 mm was used for the measurements by JASCO-720 spectropolarimeter (Jasco Internatiol Co. Ltd.). First, CD spectra were recorded at 25 $^{\circ}\text{C}$ over $\lambda = 230\text{--}350$ nm with an interval of 0.2 nm and a scan rate of 100 nmmin $^{-1}$. The corresponding buffer was used alone for baseline correction of each measurement. Melting curves were recorded at 264 nm between 25–95 $^{\circ}\text{C}$ at a speed of 1 $^{\circ}\text{Cmin}^{-1}$. Melting temperature (T_m) is defined as the temperature at which 50% of the G4 structures are unfolded. T_m values were estimated by fitting the melting curves into a dose response function using the OriginPro 2016 software.

Fluorescence intercalator displacement experiments: The experiments were performed at 25 $^{\circ}\text{C}$ on a Jasco FP-6500 spectrofluorometer equipped with a temperature controller. 0.25 μM of pre-folded G4-DNA in 10 mM potassium phosphate (pH 7.4), 100 mM KCl buffer was mixed with 0.50 μM Thiazole Orange (TO) and incubated for 2 minutes before the fluorescence spectrum was recorded ($\lambda_{\text{ex}} = 501$ nm; $\lambda_{\text{em}} = 510\text{--}650$ nm). Then ligands were added to the mixture stepwise with a 2 min equilibration period, and the fluorescence spectra were recorded. The percentage of TO displacement was calculated from the fluorescence intensity (F) at the emission maxima, using Equation (1):

$$\text{Percentage of TO displacement} = 100 - \left(\frac{F}{F_0} \times 100 \right) \quad (1)$$

where, F_0 is the initial fluorescence intensity of TO bound to G4-DNA.

The percentage of TO displacement was plotted as a function of the concentration of added ligands and DC_{50} is determined. The association binding constant (K_a) of the ligands were calculated from Equation (2) using K_a^{TO} as $5.01 \times 10^6 \text{ M}^{-1}$:

$$K_a^{\text{ligand}} = \frac{K_a^{TO} \times [TO]}{[\text{ligand}]_{50}} \quad (2)$$

Microscale thermophoresis (MST): The Pu24T G4 DNA sequence with a 5' CY5 label was folded in KCl buffer (10 mM phosphate, 100 mM KCl, pH 7.4) by heating at 95 $^{\circ}\text{C}$ for 5 min and then cooling to room temperature. All the experiments were performed in 10 mM phosphate pH 7.4, 100 mM KCl, 0.05% Tween20 and 4% BSA. The labelled DNA concentration is held constant at 25 nM and ligand concentration is varied from 0.15 nM to 10 μM (fourteen 1:1 dilutions). The samples were loaded into standard MST graded glass capillaries and MST experiment is performed using Monolith NT.115 (Nano Temper, Germany) with 40% LED power. Data was analyzed using the Nano Temper analysis software and plotted in OriginPro 8.

Nuclear magnetic resonance: The G4 DNA stock solutions was prepared by folding 100 μM *c-MYC* Pu24T in 10 mM potassium phosphate buffer (pH 7.4) and 35 mM KCl by heating to 95 $^{\circ}\text{C}$ and slowly cooling to room temperature overnight. 10% D_2O and 10% [D_6]DMSO was added to the DNA stock solutions, yielding a final DNA concentration of 82 μM . NMR samples were prepared by sequential addition of **7aa**, **7ba**, **5bac**, and **7bac** from 20 mM DMSO stock solutions to 220 μL of the DNA solution which was then

transferred to 3 mm NMR tubes. Control samples with Pu24T *c-MYC* G4 DNA with and without 10% DMSO was also performed to verify that DMSO did not have a significant effect on the DNA structure. All spectra were recorded at 298 K on a Bruker 850 MHz Avance III HD spectrometer equipped with a 5 mm TCI cryoprobe. Excitation sculpting was used in the 1D ^1H experiments, and 256 scans were recorded. Processing was performed in Topspin 3.5 (Bruker Biospin, Germany).

G4 complex modeling: For Pu24T *c-MYC* G4 DNA, structure coordinates were downloaded from PDB with PDB-ID 2MGN.^[41] From the Pu24T structure, the bound ligand was removed. In the next step, three-dimensional structural coordinates of **7aa**, **7ba**, **5bac**, and **7bac** were generated using Avogadro package.^[42] To model the Pu24T *c-MYC* G4 DNA complexes, docking was performed at the 5'-terminal G-tetrad using Autodock Vina.^[43] Two potassium ions were preserved inside the G4 channel to maintain its stability during the simulations.

Molecular dynamics simulations: All G4-compound complexes were prepared for molecular dynamics simulations using GROMACS^[44] by placing these at the center of a periodic dodecahedron box and solvating with water molecules. Subsequently, the system was neutralized by adding an excess of 0.100 M KCl using the GROMACS tools. For the DNA, Amber99SB^[45] with PARMBSC1^[46] improvements were used as force-field parameters. For water, the TIP3P model^[47] was considered while ion parameters were taken from the following reference [48]. For **7aa**, **7ba**, **5bac**, and **7bac**, at first, their geometry was optimized and ESP was calculated with HF/6-31G* basis set using Gaussian-16^[49] and subsequently the partial atomic charges were calculated with the RESP method using the AmberTools package.^[50] All four compounds' force-field parameters were generated from GAFF using the AmberTools package and converted to GROMACS format using acpype script.^[51] Subsequently, MD simulations were performed using GROMACS package.^[44] Parameter settings for all these stages were previously described in the following reference^[36].## Length of the MD simulations were 5*200 ns for each compound. All trajectories were merged for respective complexes, processed, and further used for the analysis.

Conformational clustering for each compound bound to their corresponding G4 DNA structure was performed with the *gmx clusterByFeatures* tool using PCA based conformational clustering (https://github.com/rjdkmr/gmx_clusterByFeatures). Subsequently, the first 50 frames of each cluster were considered for binding energy calculation using the *g mmpbsa* tool.^[52,53] The obtained MD trajectories were visualized and images were rendered using VMD.^[54]

Acknowledgements

Work in the Chorell lab was supported by the Kempe foundations (SMK-1632) and the Swedish Research Council (2017-05235). Work in the Sabouri lab was supported by Knut and Alice Wallenberg Foundation (KAW2015-0189), the Swedish Research Council (VR-MH 2018-2651), the Swedish Cancer Society (CAN2019/126), and the Medical Faculty of Umeå University. The MD simulations were performed on resources provided by the Swedish National Infrastructure for Computing (SNIC) at HPC2N Umeå, Sweden. We also thank the Knut and Alice Wallenberg foundation program NMR for Life (www.nmrforlife.se) for NMR spectroscopy support.

Conflict of interest

The authors declare no conflict of interest.

Keywords: bis-indole · DNA structures · drug design · G-Quadruplexes · nitrogen heterocycles

- [1] D. Sen, W. Gilbert, *Nature* **1990**, *344*, 410–414.
- [2] J. R. Williamson, M. K. Raghuraman, T. R. Cech, *Cell* **1989**, *59*, 871–880.
- [3] A. N. Lane, J. B. Chaires, R. D. Gray, J. O. Trent, *Nucleic Acids Res.* **2008**, *36*, 5482–5515.
- [4] M. P. Cheng, Y. Cheng, J. Y. Hao, G. Q. Jia, J. Zhou, J. L. Mergny, C. Li, *Nucleic Acids Res.* **2018**, *46*, 9264–9275.
- [5] A. Risitano, K. R. Fox, *Nucleic Acids Res.* **2004**, *32*, 2598–2606.
- [6] A. Guédin, J. Gros, P. Alberti, J. L. Mergny, *Nucleic Acids Res.* **2010**, *38*, 7858–7868.
- [7] R. Tippiana, W. K. Xiao, S. Myong, *Nucleic Acids Res.* **2014**, *42*, 8106–8114.
- [8] P. A. Rachwal, T. Brown, K. R. Fox, *FEBS Lett.* **2007**, *581*, 1657–1660.
- [9] G. Sattin, A. Artese, M. Nadai, G. Costa, L. Parrotta, S. Alcaro, M. Palumbo, S. N. Richter, *PLoS One* **2013**, *8*, e84113.
- [10] D. Sun, L. H. Hurley, *J. Med. Chem.* **2009**, *52*, 2863–2874.
- [11] D. Miyoshi, H. Karimata, N. Sugimoto, *J. Am. Chem. Soc.* **2006**, *128*, 7957–7963.
- [12] E. N. Nikolova, E. Kim, A. A. Wise, P. J. O'Brien, I. Andricioaei, H. M. Al-Hashimi, *Nature* **2011**, *470*, 498–484.
- [13] V. S. Chambers, G. Marsico, J. M. Boutell, M. Di Antonio, G. P. Smith, S. Balasubramanian, *Nat. biotechnol.* **2015**, *33*, 877–881.
- [14] G. Marsico, V. S. Chambers, A. B. Sahakyan, P. McCauley, J. M. Boutell, M. Di Antonio, S. Balasubramanian, *Nucleic Acids Res.* **2019**, *47*, 3862–3874.
- [15] D. Rhodes, H. J. Lipps, *Nucleic Acids Res.* **2015**, *43*, 8627–8637.
- [16] J. Eddy, N. Maizels, *Nucleic Acids Res.* **2006**, *34*, 3887–3896.
- [17] J. L. Huppert, S. Balasubramanian, *Nucleic Acids Res.* **2007**, *35*, 406–413.
- [18] J. A. Capra, K. Paeschke, M. Singh, V. A. Zakian, *Plos Comput. Biol.* **2010**, *6*, e1000861.
- [19] P. Rawal, V. B. R. Kumarasetti, J. Ravindran, N. Kumar, K. Halder, R. Sharma, M. Mukerji, S. K. Das, S. Chowdhury, *Genome Res.* **2006**, *16*, 644–655.
- [20] V. K. Yadav, J. K. Abraham, P. Mani, R. Kulshrestha, S. Chowdhury, *Nucleic Acids Res.* **2007**, *36*, D381–D385.
- [21] M. L. Bochman, K. Paeschke, V. A. Zakian, *Nature Reviews: Genetics* **2012**, *13*, 770–780.
- [22] N. Maizels, L. T. Gray, *Plos Genet.* **2013**, *9*, e1003468.
- [23] O. Mendoza, A. Bourdoncle, J. B. Boule, R. M. Brosh, J. L. Mergny, *Nucleic Acids Res.* **2016**, *44*, 1989–2006.
- [24] N. Maizels, *Embo Rep.* **2015**, *16*, 910–922.
- [25] S. Balasubramanian, L. H. Hurley, S. Neidle, *Nat. Rev. Drug Discovery* **2011**, *10*, 261–275.
- [26] H. Xu, M. Di Antonio, S. McKinney, V. Mathew, B. Ho, N. J. O'Neil, N. Dos Santos, J. Silvester, V. Wei, J. Garcia, F. Kabeer, D. Lai, P. Soriano, J. Banath, D. S. Chiu, D. Yap, D. D. Le, F. B. Ye, A. N. Zhang, K. Thu, J. Soong, S. C. Lin, A. H. C. Tsai, T. Osako, T. Algara, D. N. Saunders, J. Wong, J. Xian, M. B. Bally, J. D. Brenton, G. W. Brown, S. P. Shah, D. Cescon, T. W. Mak, C. Caldas, P. C. Stirling, P. Hieter, S. Balasubramanian, S. Aparicio, *Nat. Commun.* **2017**, *8*, 14432.
- [27] R. Simone, P. Fratta, S. Neidle, G. N. Parkinson, A. M. Isaacs, *FEBS Lett.* **2015**, *589*, 1653–1668.
- [28] S. Pelengaris, M. Khan, G. Evan, *Nat. Rev. Cancer* **2002**, *2*, 764–776.
- [29] T. Simonsson, P. Pecinka, M. Kubista, *Nucleic Acids Res.* **1998**, *26*, 1167–1172.
- [30] U. Siebenlist, L. Hennighausen, J. Battey, P. Leder, *Cell* **1984**, *37*, 381–391.
- [31] D. Z. Yang, L. H. Hurley, *Nucleos. Nucleot. Nucl.* **2006**, *25*, 951–968.
- [32] A. Ambrus, D. Chen, J. X. Dai, R. A. Jones, D. Z. Yang, *Biochemistry* **2005**, *44*, 2048–2058.
- [33] A. T. Phan, V. Kuryavyi, H. Y. Gaw, D. J. Patel, *Nat. Chem. Biol.* **2005**, *1*, 167–173.
- [34] A. De Cian, E. DeLemos, J. L. Mergny, M. P. Teulade-Fichou, D. Monchaud, *J. Am. Chem. Soc.* **2007**, *129*, 1856.
- [35] M. Livendahl, J. Jamroskovic, S. Ivanova, P. Demirel, N. Sabouri, E. Chorrell, *Chem. Eur. J.* **2016**, *22*, 13004–13009.
- [36] B. Prasad, J. Jamroskovic, S. Bhowmik, R. Kumar, T. Romell, N. Sabouri, E. Chorrell, *Chem. Eur. J.* **2018**, *24*, 7926–7938.
- [37] J. Dash, R. Nath Das, N. Hegde, G. D. Pantoş, P. S. Shirude, S. Balasubramanian, *Chem. Eur. J.* **2012**, *18*, 554–564.
- [38] S. Rankin, A. P. Reszka, J. Huppert, M. Zloh, G. N. Parkinson, A. K. Todd, S. Ladame, S. Balasubramanian, S. Neidle, *J. Am. Chem. Soc.* **2005**, *127*, 10584–10589.
- [39] P. Agrawal, C. Lin, R. I. Mathad, M. Carver, D. Z. Yang, *J. Am. Chem. Soc.* **2014**, *136*, 1750–1753.
- [40] S. Neidle, *J. Med. Chem.* **2016**, *59*, 5987–6011.
- [41] W. J. Chung, B. Heddi, F. Hamon, M.-P. Teulade-Fichou, P. Anh Tuan, *Angew. Chem. Int. Ed.* **2014**, *53*, 999–1002; *Angew. Chem.* **2014**, *126*, 1017–1020.
- [42] M. D. Hanwell, D. E. Curtis, D. C. Lonie, T. Vandermeersch, E. Zurek, G. R. Hutchison, *J. Cheminf.* **2012**, *4*, 17.
- [43] O. Trott, A. J. Olson, *J. Comput. Chem.* **2010**, *31*, 455–461.
- [44] M. J. Abraham, T. Murtola, R. Schulz, S. Páll, J. C. Smith, B. Hess, E. Lindahl, *SoftwareX* **2015**, *1–2*, 19–25.
- [45] V. Hornak, R. Abel, A. Okur, B. Stockbake, A. Roitberg, C. Simmerling, *Proteins Struct. Funct. Bioinf.* **2006**, *65*, 712–725.
- [46] I. Ivani, P. D. Dans, A. Noy, A. Perez, I. Faustino, A. Hospital, J. Walther, P. Andrio, R. Goni, A. Balaceanu, G. Portella, F. Battistini, J. L. Gelpi, C. Gonzalez, M. Vendruscolo, C. A. Loughton, S. A. Harris, D. A. Case, M. Orozco, *Nat. Methods* **2016**, *13*, 55–58.
- [47] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, M. L. Klein, *J. Chem. Phys.* **1983**, *79*, 926–935.
- [48] L. X. Dang, *J. Am. Chem. Soc.* **1995**, *117*, 6954–6960.
- [49] Gaussian 16, Revision A03, M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, G. A. Petersson, H. Nakatsuji, X. Li, M. Caricato, A. V. Marenich, J. Bloino, B. G. Janesko, R. Gomperts, B. Mennucci, H. P. Hratchian, J. V. Ortiz, A. F. Izmaylov, J. L. Sonnenberg, Williams, F. Ding, F. Lipparini, F. Egidi, J. Goings, B. Peng, A. Petrone, T. Henderson, D. Ranasinghe, V. G. Zakrzewski, J. Gao, N. Rega, G. Zheng, W. Liang, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, K. Throssell, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. J. Bearpark, J. J. Heyd, E. N. Brothers, K. N. Kudin, V. N. Staroverov, T. A. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. P. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, J. M. Millam, M. Klene, C. Adamo, R. Cammi, J. W. Ochterski, R. L. Martin, K. Morokuma, O. Farkas, J. B. Foresman, D. J. Fox, Wallingford, CT, **2016**.
- [50] D. A. Case, D. S. Cerutti, I. T. E. Cheatham, T. A. Darden, R. E. Duke, T. J. Giese, H. Gohlke, A. W. Goetz, D. Greene, N. Homeyer, S. Izadi, A. Kovalenko, S. T. S. Lee, LeGrand, P. Li, C. Lin, J. Liu, T. Luchko, R. Luo, D. Mermelstein, K. M. Merz, G. Monard, H. Nguyen, I. Omelyan, A. Onufriev, F. Pan, R. Qi, D. R. Roe, A. Roitberg, C. Sagui, C. L. Simmerling, W. M. Botello-Smith, J. Swails, R. C. Walker, J. Wang, R. M. Wolf, X. Wu, L. Xiao, D. M. York, P. A. Kollman, AMBER, University of California, San Francisco, **2017**.
- [51] A. W. Sousa da Silva, W. F. Vranken, *BMC Res. Notes* **2012**, *5*, 367.
- [52] N. A. Baker, D. Sept, S. Joseph, M. J. Holst, J. A. McCammon, *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 10037–10041.
- [53] R. Kumari, R. Kumar, C. A. Lynn, *J. Chem. Inf. Model.* **2014**, *54*, 1951–1962.
- [54] W. Humphrey, A. Dalke, K. Schulten, *J. Mol. Graph.* **1996**, *14*, 33–38.

Manuscript received: February 3, 2020

Revised manuscript received: March 16, 2020

Accepted manuscript online: March 18, 2020

Version of record online: July 3, 2020