Redefining the Essential Molecular Aspects that Drive Interactions Between Small Molecules and G-Quadruplex DNA

Måns S. Andreasson

Dissertation for PhD
ISBN: 978-91-7855-969-5 (print)
ISBN: 978-91-7855-970-1 (pdf)
Electronic version available at: http://umu.diva-portal.org/
Printed by: Cityprint i Norr AB
Umeå, Sweden 2022

Table of Contents

Abstract

G-Quadruplex (G4) structures are secondary nucleic acid structures located in guanine-rich regions of DNA and RNA sequences, involved in gene regulation and cellular maintenance. Efforts to target G4s in a therapeutic setting are scarce, mainly due to vague details about the binding interactions between the ligands and the G4 structure combined with the lack of emphasis on drug-like properties early in the ligand development process. Furthermore, the ability to target specific G4 structures with small drug-like molecules remains a big challenge to overcome in the field. In this thesis, extensive organic synthesis developments coupled with computational-aided design and orthogonal *in vitro* assays has been used in tandem to reveal in-depth knowledge about ligand-to-G4 interactions. First, a macrocyclic approach was applied to design and discover novel G4 ligands which showed that macrocycles offer a solid foundation for ligand design. Next, computational tools to optimise the macrocyclic molecular conformation were used based on the macrocycles' abilities to stack on the G4 surface. In addition, macrocyclic, and non-macrocyclic ligands that bound G4 with high potency were shown to correlate with electron-deficient electrostatic potential (ESP) maps. The frequent inclusion of cationic residues in G4 ligands and their enhancement on ligand-to-G4 binding was, thereof, ascribed to their impact on the electrostatic character of the ligands' arene-arene interactions with the G4 surface, and not through direct electrostatic ionic interactions. In addition, the dispersion energetic component in the arene-arene interactions between the G4 ligand and the G4 was discovered to be paramount for ligand-to-G4 binding. The implementation of these descriptors in practice resulted in the discovery of potent G4 binders with adequate pharmacokinetic (PK) properties, accentuating the significance of understanding the molecular interactions between ligands and G4s in rational ligand design. Finally, a G4 ligand conjugated to an oligonucleotide was demonstrated as a modular approach to achieve selective binding of a ligand to a specific G4 structure.

List of Publications

This thesis is based on the following papers and manuscripts, referred to in the text by their Roman numerals.

I. **Macrocyclization of bis-indole quinolines for selective stabilization of G-quadruplex DNA structures**
Rabindra Nath Das, Måns Andreasson, Rajendra Kumar, Erik Chorell
Chem. Sci. **2020**, *11*, 10529-10537.

II. **Using Macrocyclic G-Quadruplex Ligands to Decipher the Interactions Between Small Molecules and G-Quadruplex DNA**
Måns Andreasson, Naresh Bhuma, Nils Pemberton, Erik Chorell
Chem. Eur. J. **2022**, *28* (e202202020), 1-14.

III. **The effect of side chain variations on quinazoline-pyrimidine G-quadruplex DNA ligands**
Naresh Bhuma, Karam Chand, Måns Andreasson, James Mason, Rabindra Nath Das, Ankit Kumar Patel, Daniel Öhlund, Erik Chorell
Revised Manuscript, Eur. J. Med. Chem. **2022**.

IV. **"The Synergism of the Dispersion and Electrostatic Components in the Arene-Arene Interactions Between Ligands and G4 DNA"**
Måns Andreasson, Maxime Donzel, Alva Abrahamsson, Andreas Berner, Mara Doimo, Anna Quiroga, Nils Pemberton, Sjoerd Wanrooij, Erik Chorell
Manuscript

V. **"Development of a G4 Ligand-Conjugated Oligonucleotide Modality that Selectively Targets Individual G4 DNA Structures"**
Rabindra Nath Das[#], Andreas Berner[#], Naresh Bhuma, Justyna Golebiewska, Alva Abrahamsson, Måns Andréasson, Namrata Chaudhari, Mara Doimo, Sjoerd Wanrooij*, Erik Chorell*
[#]Equally contributing first authors, *Equally contributing corresponding authors
Manuscript

Publications not included in the thesis:

A Minimalistic Coumarin Turn-On Probe for Selective Recognition of Parallel G-Quadruplex DNA Structures
Marco Deiana, Ikenna Obi, Måns Andreasson, Shanmugam Tamilselvi, Karam Chand, Erik Chorell, Nasim Sabouri
ACS Chem. Biol. **2021**, *16*, 1365−1376.

A complementary chemical probe approach towards customized studies of G-quadruplex DNA structures in live cells
Prasad Bagineni, Mara Doimo, Måns Andreasson, Valentin L'Hote, Erik Chorell Sjoerd Wanrooij
Chem. Sci. **2022**, *13*, 2347–2354.

Contribution to papers I-V

I. Performed half of the synthesis and did parts of the assay evaluation. Interpreted parts of the results and wrote the manuscript draft.

II. Designed the project, performed majority of the synthesis and the assay evaluation. Interpreted the results and wrote the manuscript draft.

III. Did the computational part and illustrations for the NMR figures. Wrote minor parts of the manuscript draft.

IV. Designed the project, performed half of the synthesis. Performed or supervised the assay evaluations. Interpreted the results and wrote the manuscript draft.

V. Performed part of the ligand synthesis and was involved in the design of the project, interpretation of results, and manuscript preparation.

List of Abbreviations

Ac_2O	**Acetic Anhydride**
AcCl	**Acetyl Chloride**
AIBN	**Azobisisobutyronitrile**
Amb. Temp.	**Ambient Temperature**
BCL2	**B-cell Lymphoma 2**
Boc	***Tert*-butoxycarbonyl**
Caco-2	**Cancer Coli**
CD	**Circular Dichroism**
COMU	**(1-Cyano-2-ethoxy-2-oxoethylidenaminooxy)dimethylamino-morpholino-carbenium**
Cy5	**Cyanines or Tetramethylindo(di)-carbocyanines**
DABCO	**1,4-diazabicyclo[2.2.2]octane**
DBU	**1,8-Diazabicyclo[5.4.0]undec-7-ene**
DFT	**Density functional Theory**
DIBAL-H	**Diisobutylaluminium hydride**
DIPEA	***N*,*N*-Diisopropylethylamine or Hünig's base**
DMAP	**4-Dimethylaminopyridine**
DMF	**Dimethylformamide**
DNA	**Deoxyribonucleic acid**
EDC•HCl	***N*-(3-Dimethylaminopropyl)-N'-ethylcarbodiimide hydrochloride**
ESP	**Electrostatic potential**
Et	**Ethyl**
FAM	**Fluorescein amidites**
FID	**Fluorescence intercalator displacement**
FRET	**Fluorescence (Förster) Resonance Energy Transfer**
G4	**G-quadruplex**

G-rich	**Guanine-rich**
GIST	**Gastrointestinal Stromal Tumours**
HATU	**(1-[Bis(dimethylamino)methylene]-1H-1,2,3-triazolo[4,5-b]pyridinium 3-oxide**
IR	**Infra-red**
K_d	**Dissociation constant or Binding affinity**
KIT	**Cytokine receptor**
LCMS	**Liquid chromatography-mass spectrometry**
Me	**Methyl**
MM	**Molecular mechanic**
MST	**Microscale thermophoresis**
MRFO	**Matching random flanking oligo**
mRNA	**messenger-RNA**
NBS	***N*-bromosuccinimide**
NCS	***N*-chlorosuccinimide**
NIS	***N*-iodosuccinimide**
NMR	**Nuclear magnetic resonance**
PNA	**Peptide nucleic acid**
PK	**Pharmacokinetic**
PyAOP	**(7-Azabenzotriazol-1-yloxy) tripyrrolidinophosphonium**
QM	**Quantum mechanic**
Ras	**Rat sarcoma virus**
RFO	**Random flanking oligo**
RNA	**Ribonucleic acid**
S_EAr	**Electrophilic aromatic substitution**
S_NAr	**Nucleophilic aromatic substitution**
T3P	**1-Propanephosphonic anhydride**
TAMRA	**Carboxytetramethylrhodamine**
TBAF	**Tetra-*n*-butylammonium fluoride**

TBHP	***Tert*-butyl hydroperoxide**
TCFH	**Chloro-*N*,*N*,*N*′,*N*′-tetramethylformamidinium**
TFA	**Trifluoro acetic acid**
THF	**Tetrahydrofuran**
TIPS	**Triisopropylsilyl**
T_m	**Melting Temperature**
TO	**Thiazole Orange**
VEGF	**Vascular Endothelial Growth Factor**

Sammanfattning på Svenska

Guanin-rika regioner av DNA- och RNA-sekvenser har förmågan att bilda sekundära strukturer som kallas ”G-Quadruplexes”, även kallade G4. Dessa substrukturer har etablerats som viktiga element vid genreglering och andra viktiga cellulära processer. Försök att utveckla små molekyler som interagerar med G4 för terapeutiska ändamål har i dagsläget inte resulterat i några läkemedel trots deras viktiga roll vid bland annat utveckling av cancer. Begränsningar i både kunskaperna gällande de exakta interaktionerna mellan molekyl och G4, samt bristen av fokus på läkemedelsliknande egenskaper tidigt i utvecklingsprocessen av nya molekyler är troligtvis bidragande till att sådana projekt stannar i ett tidigt stadie. En annan viktig faktor som troligtvis är den mest bidragande, är utmaningen i att kunna utveckla molekyler som endast interagerar med en specifik G4-struktur. Detta anses viktigt eftersom det finns hundratusentals positioner i det mänskliga genomet där G4 skulle kunna bildas. I denna avhandling har omfattande organisk syntesutveckling i kombination med beräknings-baserad design och ortogonala *in vitro*-analyser använts för att i detalj förstå interaktionerna mellan molekyl och G4. Denna information har sedan applicerats för design och utveckling av nya molekyler med bättre egenskaper. Först prövades en makrocyklisk design-strategi för att utveckla nya G4-ligander. Detta visade sig vara en bra grund för molekyl-design som även gjorde molekylerna selektiva för att binda till G4-DNA över dubbelsträngat DNA. För att ytterligare förbättra makrocyklernas förmåga att interagera med de exponerade ytorna på G4-strukturen så undersöktes makrocyklernas molekylära konformation med hjälp av beräkningsmetoder. Fördjupade beräkningsstudier med både makrocykliska och icke-makrocykliska molekyler visade sedan att de molekyler som band G4 med hög affinitet uppvisade elektronfattiga egenskaper genom så kallade ESP-kartor. Den frekventa användningen av kat-joniska funktionella grupper i G4-ligander och dess positiva inverkan på interaktionen mellan molekyl och G4 föreslogs därför vara kopplad till deras inverkan på den elektrostatiska karaktären av molekylernas aren-aren-interaktioner med G4-ytan, och inte genom en direkt jon-jon-interaktion. Utöver detta visade studierna att dispersions-komponenten i arena-aren-interaktionerna mellan molekyl och G4 är avgörande för en stark bindning. Tillämpningen av dessa molekylära egenskaper i praktiken resulterade i upptäckten av potenta G4-ligander med läkemedelsliknande egenskaper, vilket betonar vikten av att förstå de molekylära interaktionerna mellan molekyler och G4. Slutligen utvecklades en ny strategi, baserat på en G4-ligand konjugerad till en oligonukleotid, som visade stark potential för att uppnå selektiv bindning till specifika G4-strukturer.

1. G-Quadruplex DNA

The ability of guanine nucleotides to form aggregates in an aqueous solution was noted as early as 1910.[1] It was later disclosed that such aggregates were composed of arranged self-stacking guanine nucleobases.[2,3] The ordered form of these aggregates in guanine-rich sequences of DNA and RNA have since become well-known as G-quadruplex structures (G4s).[4] The G4s are assembled through the stacking of the G-quartets (plane-assemblies of four guanines), which in turn form a framework of internal Hoogsteen hydrogen bonds (Figure 1A and 1B). The structures are stabilised by the stacking (arene-arene) interactions between each G-quartet and cation coordination through the carbonyl oxygen of each guanine, preferably to K^+.[5, 6] An illustration of the G-quartet and a crystal structure of the *c-MYC* G4 DNA (Pu22), along with a view of how the guanines on the G4 surface presents the accessible area for interactions with molecules, are shown in Figure 1A-C. G4 structures typically show rapid folding kinetics,[7] (dependent on ion concentration) and have thermal stabilities that can match or be even higher than the double-stranded DNA helix.[8]

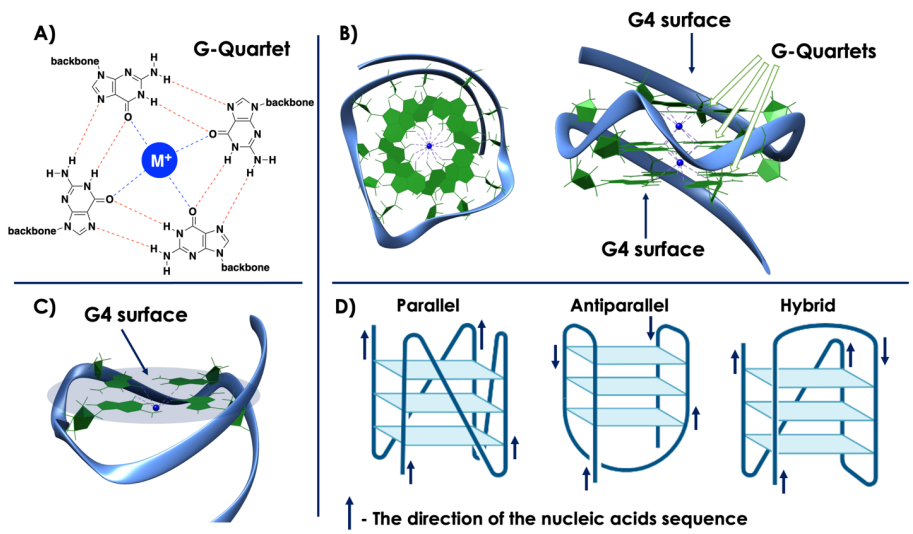


Figure 1. A) Illustration and structure of a single G-quartet. **B)** Crystal structure of G4 DNA Pu22 (5W77) shown from the top (left) and the side (right) edited in the Chimera software. **C)** Illustration of how the G4 surface with its available guanines offer a binding surface where molecules can stack and bind. **D)** Illustration of how intramolecular G4 structures with either a parallel, antiparallel, or hybrid topology can be viewed.

The G4 structures are distinct and can adopt different secondary arrangements, commonly referred to as different topologies. The type of topology is assigned according to if the strands connecting the guanines are aligned parallel,

antiparallel, or hybrid (Figure 1D). Furthermore, the G4 structure can also be inter- or intramolecular with respect to the strands of guanines. What determines the topology of a G4 structure has been attributed to the composition of the other nucleotides in the strand, the length of the loops between the guanines, and the cation composition.[5,9–11] The area of G4 topologies is a small science of its own[10,11] and will not be covered in much more depth. However, for the context of this thesis, it is important to note that the type of topology can affect the steric accessibility of the G4 surface, e.g. if the connecting loops cross the diagonal of the G4 surface. The formation of G4 structures requires single-stranded nucleic acid sequences. Consequently, any local event that disrupts the Watson-Crick base-pairing, and subsequently the DNA helix, can favour G4 folding. Hence, G4 formation is likely to occur naturally during events which result in the formation of single-stranded nucleic acid sequences: During negative supercoiling, DNA damage repair, transcription, or DNA replication.[12,13]

1.1 G4s and Biology

Guanine-rich (G-rich) nucleotide regions capable of forming G4 structures are evolutionary conserved.[14,15] Such G-rich sequences with the possibility to fold into G4s in the human genome have been estimated to be several hundreds of thousands through computational algorithms and whole-genome sequencing.[16,17] The *in vivo* formation of G4 structures has been reinforced by experimental data and the existence of proteins, such as helicases, that can facilitate their folding or unfolding.[4,12,18,19] G4s are not arbitrarily distributed but abundant in transcriptional regulator sequences, highlighting their roles in gene expression regulation, notably in proto-oncogenic sequences.[16,20–25] In addition, G4 forming sequences are localised in the DNA of the mitochondrial genome,[26] evocative of their involvements in managing mitochondrial DNA replication and transcription.[27] Furthermore, G4 involvements in the regulation of mRNA translation has also been explored.[21,22,28] Recently, it was also shown how G4s are enriched at active and poised enhancer regions (distal factors that control the regulation of gene expressions that are essential for development), underscoring their roles in more complex transcriptional mechanisms.[29]

The single-stranded overhangs of telomeric DNA contain G-rich sequences, which have been shown to fold into G4s *in vivo*.[30] Telomeric G4s have been ascribed to be involved in the telomeric maintenance machinery through telomere end-capping and telomerase engagement (nucleotide extension of the telomeric ends).[22,31] In addition, G4s are proposed to play a role in the regulation of the origin of replication by binding protein factors which initiates DNA

replication.[32,33] G4s in the context of neurodegenerative disorders like Alzheimer's and Amyotrophic Lateral Sclerosis (ALS) have been identified. The role of G4 structures in the disease development process have been attributed to their effects on the dysregulation of DNA transcriptional and RNA translational events.[34,35] Finally, G4s has also been connected to important steps in the life cycle of several viruses. [36]

Taken together, it becomes apparent that G4s are involved in central biological processes and manipulation of their presence and stability can have direct effect on central biological processes and human disease development.[37,38]

1.1.1 *c-MYC*

A well-known proto-oncogene where a G4 structure is a part of the intricate gene regulatory mechanism is the *c-MYC*.[13,39] The *c-MYC* gene is associated with cell proliferation and is dysregulated in 70% of all known cancers.[40] The MYC protein is considered hard to drug with small organic molecules due to its undefined binding sites and short half-life.[41] Hence, it has been proposed that a way to inhibit cancer cell growth would be to instead downregulate the *c-MYC* gene expression.[40] The *c-MYC* gene transcription is, in part, regulated by the transcriptional regulatory nuclease hypersensitivity element III_1 (a location in DNA gene sequences that can bind to transcription-regulatory proteins) abbreviated $NHEIII_1$, which consist a guanine-rich nucleotide sequence.[40] This guanine-rich sequence, called Pu27, can form a G4 structure upon the disruption of the double-stranded DNA through negative supercoiling during transcription.[13] Pu27 has been frequently studied *in vitro* by its mutated sequences Pu22 and Pu24T, each representing a single intramolecular parallel G4 structure.[42–44] When the endogenous *c-MYC* G4 sequence (Pu27) is folded, transcription is repressed and the c-MYC protein expression is reduced.[45] Several studies have shown that ligand-induced stabilisation of the G4 structure in this sequence can reduce *c-MYC* expression and cancer growth.[46] Thus, the use of ligands to bind and stabilise the G4 structure in the *c-MYC* promotor region offers an attractive approach to treat *c-MYC*-dependent cancers, which avoid the pitfalls of targeting the c-MYC protein directly.[38,41,47]

1.1.2 *c-KIT* and *Ras*

The *c-KIT* gene is another proto-oncogene elevated in cancers, notably in gastrointestinal stromal tumours (GIST).[48,49] In patients with GIST, clinical resistance to conventional treatment with kinase inhibitors is known to occur due to mutations in the KIT protein, posing challenges in treating GIST through KIT targeting with conventional drug-protein strategies.[50] In the native promotor

region of the *c-KIT* gene, there is a guanine-rich sequence with the ability to form a G4 structure under physiological conditions.[51] This G4 structure is, *in vitro*, represented by the mutated structures cKIT1[51] and cKIT2.[52] Studies indicate that ligand binding to the G4 in the *c-KIT* promotor region could reduce gene transcription through G4 stabilisation.[47,53]

The *Ras* (*KRas, NRas, HRas*) proto-oncogene is the most recurrently dysregulated (by mutations) gene in pancreatic cancers.[54] The Ras promotor sequence has a guanine-rich sequence which, when folded into a G4, silences the transcription of the Ras proteins.[55–57] Along this line, targeting the G4 in the *Ras* sequence is considered a promising way to complement the current strategies for treating *Ras*-associated cancers, especially pancreatic cancer that is associated with a poor patient prognosis.[47,54,55]

In addition to these examples, several other relevant proto-oncogenes have been disclosed where G4 formation in the gene promotor region correlates with a reduction in expression. Such examples are the *BCL2*[58] and *VEGF*[59] genes, among others.[47]

1.2 Therapeutic Potential

There are substantial connections between G4s and the development of cancers. Numerous studies have shown how ligands can bind and stabilise G4 structures, resulting in a reduction in cancer growth.[38,46,47,60,61] In the literature[38], three main mechanisms have been discussed to account for the antiproliferative effects of ligand-induced formation and stabilisation of G4 structures (Figure 2).

The first: Telomeric G4 stabilisation blocks telomerase activity, whose function is crucial for the cancer cells to continuously replicate,[62] and this effect has been shown *in vivo*.[30] The antiproliferative effect of ligands binding to telomeric G4s has also been attributed to the influence of different telomeric maintenance factors.[63,64]

The second: Regulation of proto-oncogene transcription by ligand-induced stabilisation of G4s in gene promotor sequences,[25,47] as discussed in the previous section.

The third: G4 stabilisation during for example replication, can result in genomic instability and apoptosis,[12,18] and G4 ligands have been shown to induce such DNA damage responses.[65]

In addition, there are also several slight variations of the above-mentioned mechanisms that account for the activity of G4 ligands.[38]

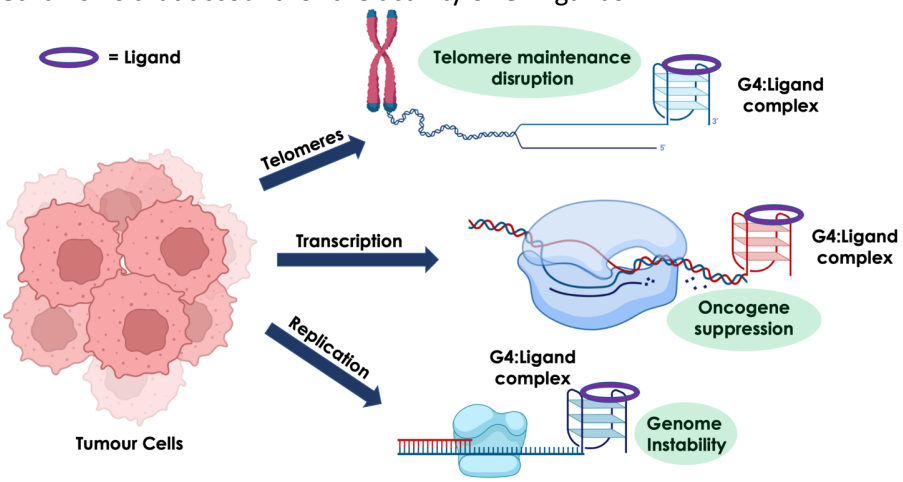


Figure 2. Illustration of how different mechanisms of ligand induced formation and stabilisation of G4 structures in distinct genomic sequences could account for the antiproliferative effects of G4 ligands. This illustration was inspired by graphics in the literature.[38] It is important to emphasise that this picture only represents a visual illustration of the events and not a detailed picture which is accurate for the actual biological events.

Considering the mechanisms of action of G4 ligands (Figure 2), one challenge discussed in the field of targeting G4s in a therapeutic setting is to design ligands that bind only G4s with high potency, typically through the G4 surface (Figure 1C), while not binding other DNA or RNA motifs. The second hurdle to overcome, perhaps the more challenging, is to target specific G4s in the genome. This is considered difficult due to the conserved G4 binding surfaces and overall similarities across the G4 genome. The reason why it is desired to get selectivity for specific G4s is that targeting multiple G4s will likely have adverse effects since several key cellular processes can be affected simultaneously (Figure 2). However, pan-oncogene repressor G4 ligands have been suggested as a possible therapeutic option for certain types of aggressive cancers.[4]

Ways to combat the problems of G4 selectivity have been discussed in terms of exploiting the topological diversity of G4s (figure 1D), and ligands designed to target such specific differences in various G4s could be a way of gaining selectivity.[47] However, this still remains as a big challenge in the field.[12,38] Still, several promising studies show the scientific relevance of targeting G4s.[4,38,46,47,60,66–70] Notably, G4 interacting compounds **5a** and **5b** (Figure 3) have advanced into clinical trials. Compound **5a**[70] was, however, withdrawn. **5b**[68,69] is still undergoing clinical testing (see further below). Such studies further underscore the significance of the development of novel ligands that target G4s in a therapeutic setting.

2. G-Quadruplex Ligands

The extensive efforts within the field to design and develop ligands that bind and stabilise G4 structures have resulted in more than 1000 known G4 ligands.[71] Some notable examples (Figure 3) are: BRACO-19 (**1**),[72] MM41 (**2**),[73] RHPS4 (**8**),[74] and Pyridostatin (**6**)[64] which bind telomeric G4s *in vivo*, reducing cancer cell growth through DNA damage at the telomeres. Another well-known potent G4 ligand is Phen-DC3 (**3**),[75] and it was recently shown how Phen-DC3 could intercalate and stabilise telomeric G4 DNA effectively.[76] However, this compound suffers from poor PK properties, such as poor cell nucleus localisation.[77] The macrocyclic natural product Telomestatin[78] (**4**) binds G4 DNA and prevents telomerase activity, similarly to the synthesised macrocycle TmPyP4 (**7**).[79] Compound CX-3543 (**5a**)[70] is a G4 ligand that entered clinical trials for treating endocrine cancers but was withdrawn due to concerns with bioavailability, highlighting the need for accenting drug-like properties in novel G4 ligand design.[47] An analogue of CX-3543 (**5a**), called CX-5461 (**5b**),[68,69] is another known G4 ligand that has reached phase I/II clinical trials for its use in treating BRCA1/2 deficient tumours. This compound has been disclosed to selectively inhibit RNA polymerase I, in addition, to binding and stabilising G4 DNA.[68] Two other notable examples are compounds **9**[80] and **10**[67]. Compound **9** has been shown to suppress *c-MYC* expression and reduce cancer cell growth through G4 binding. Compound **10** exhibited good binding and stabilisation to G4s and was selective for killing cancer cells over healthy cells.

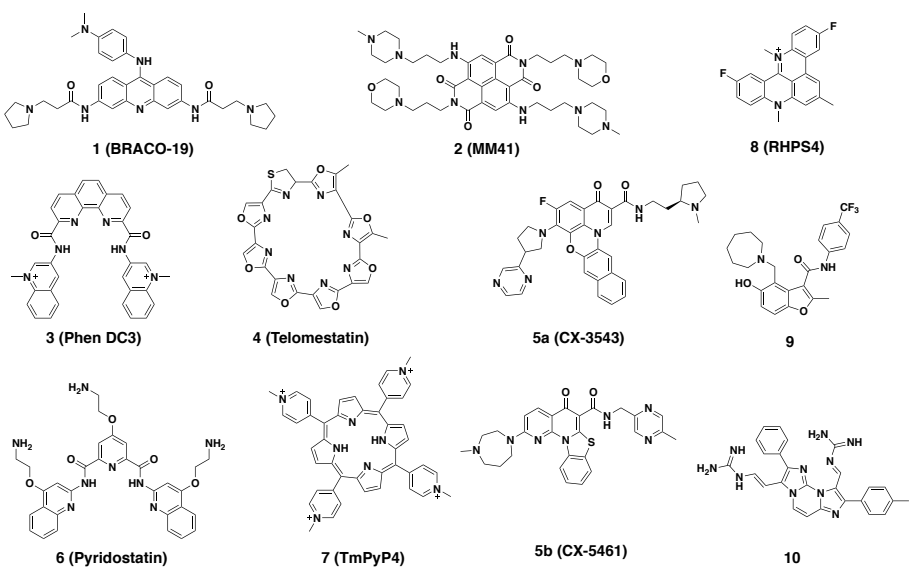


Figure 3. Summary of selected known G4 ligands, highlighting their structural and chemical similarities.

The tell-tale structural features of G4 ligands (Figure 3) are often structurally rigid polyaromatic, typically heterocyclic, compounds that can stack on the G4 surface and engage in arene-arene interactions with the terminal G quartets.[46,71,81] Moreover, such compounds often contain one or more cationic groups, which in the literature are ascribed to engage in strong electrostatic interactions with the nucleic acid phosphate backbone. It has also been stated that electron-withdrawing groups on the compounds can improve ligand-to-G4 binding.[46,81] However, the developments of compounds that target G4 structures in therapeutic applications are scant,[12,38,60] partially due to the poor PK properties often associated with known G4 ligands.[47,4]

A thorough understanding of what types of interactions that are essential for ligand-to-G4 binding and how to carefully tune and modulate them are lacking within the field. The fact that polyaromatic systems can engage in arene-arene interactions with the G4-surface, that electron-withdrawing groups can contribute to binding, and that cations can engage in electrostatic interactions with the phosphate backbone are important pieces of the puzzle, but more in-depth knowledge of the exact interactions are needed to see the big picture. Hence, a better understanding of the ligand-to-G4 interactions is essential to spearhead G4 ligands in therapeutic applications.

3. Molecular Interactions

Understanding the molecular interactions governing how and why a molecule bind to its target is imperative for every medicinal chemistry project focused on developing novel drugs. Non-covalent molecular interactions, and three-dimensional molecular conformations, govern every biological process. The interactions between a small molecule and its target are typically hydrophobic dispersion interactions, arene-arene interactions, hydrogen bonding, and electrostatic interactions. Complementary, understanding the aspects of structural features and the bioactive conformation is as important as understanding the drug-target interactions.[82] Knowledge about the components and conformational preferences that drive interactions in a particular system can thus be used to understand why a molecule bind to its target. This information can subsequently guide the design of novel and improved compounds for the target under study. Hence, both aspects should be carefully considered at the vanguard of drug development projects.

3.1 Macrocycles

The molecular conformation vs bioactive conformation is a fundamental aspect to consider in drug development. Molecules that have several low-energy molecular conformations different from the bioactive conformation will have a higher entropic penalty upon binding to their target. A strategy to lock the compound closer to its bioactive conformation and thereby reduce the entopic penalty upon binding is to use macrocycles.[83] Macrocycles are cyclic compounds, typically 12 atoms or more.[84] They are known to exhibit chemical and structural properties that allow for molecules that violate Lipinski's rule of five (beyond Rule of Five or bRo5),[85] that is, having a mass >500 Da, number of rotatable bonds >10, without sacrificing PK properties like solubility and cell permeability. Furthermore, macrocycles have been shown to tolerate polar surface area (PSA) >140 $Å^2$ which is typically a problem concerning PK properties for small organic molecules.[86–89] In addition, macrocycles can bind flat and featureless binding sites, such as sites for protein-protein interactions, with high potency and selectivity.[87,90] These attributes thus render a macrocyclic design strategy a viable option in the discovery of novel G4 ligands.

3.2 Arene-Arene Interactions

An essential component of ligand-to-G4 binding is the arene-arene interactions between the aromatic systems of the ligands and the guanines on the G4 surface.[46,81] Arene-arene interactions, usually abbreviated π-π-stacking, are the interactions between two arenes which is central for the binding and recognition of organic molecules and biological systems, especially with DNA.[91–94] There are, in principle, two energetic components in arene-arene interactions that contribute to the favourable binding event: Dispersion and Electrostatic interactions.[93,94,94–96] The dispersion (Van Der Waals) energetic component is usually ascribed as the major contributor to the favourable binding.[94,95] The beneficial factor in increasing the dispersion interactions is substituents on the arenes. Hence, the dispersion interactions typically become more favourable with the increasing number of substituents on the arenes, regardless of their electronic character.[93,95,96] This fact is attributed to direct dispersion interactions between the substituents on one arene with the other arene system. It is worth pointing out that the same rationale of direct interactions as for dispersion has been ascribed for rationalising cation/anion-arene interactions (direct interactions between the ion and the substituents on the arene). Substituents can also have negative or no effects on the binding if the substituent, because of positioning, cannot interact or has direct repulsive interactions with the other arene system.[93,96,97] The electrostatic energetic component in arene-arene interactions is based on the electrostatic potential, frequently depicted by electrostatic potential (ESP) maps, of the interacting arenes. The ESP maps show how chemical groups in the molecule, through space, affect the electrostatics of the whole molecule, and therefore also the arene.[93,96] Electrostatic interactions between the arenes can usually be rationalised by comparing their electrostatic character, and electronically matched pairs (an electron-deficient arene and an electron-rich arene) are known to be beneficial for the binding.[92,93,97]

3.3 Ionic Interactions

The interactions between ions, also called salt bridges, are widely known to occur in biological systems between charged groups, for example, aspartic acid and arginine. Such interactions can result in strongly cooperative binding events, especially when a charged residue is liberated from a hydrophobic environment.[82,98,99] As outlined earlier, the inclusion of cationic groups in G4 ligands is a characteristic element that unambiguously contributes to the ligand-to-G4 binding. In the literature, the improved binding is attributed foremost to the electrostatic interactions between the cationic groups and the anionic phosphate backbone.[46,81] Hence, for this statement to be true, the interactions

between the cations and the charged phosphates, both solvated in an aqueous environment, must result in a strongly favourable energetic event (Figure 4). On the contrary, ionic interactions in aqueous environments generally contribute little or nothing to the binding energies due to the high desolvation costs of the charged species.[82,98,99] The rationale that such interactions would contribute considerably to the ligand-to-G4 binding is thus unlikely, and a different explanation for this observation is merited.

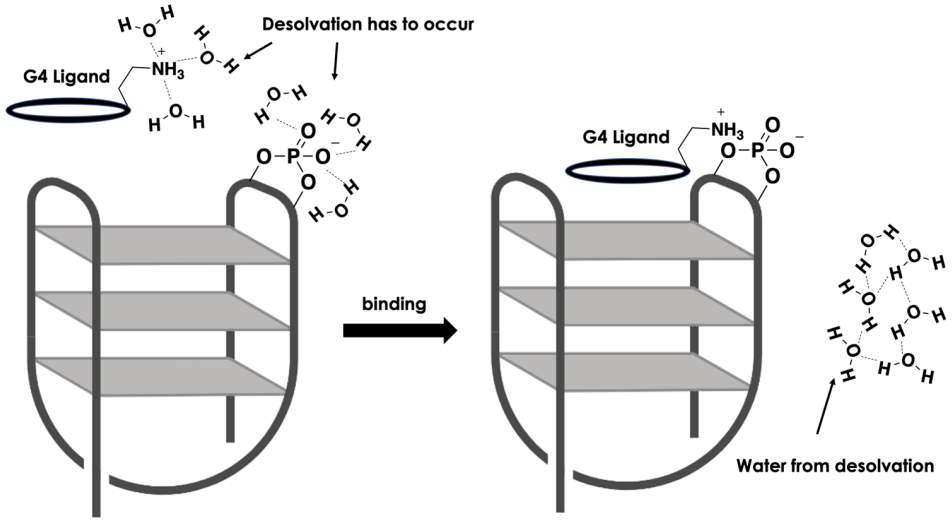


Figure 4. Visual illustration of a ligand binding to a G4 and why ion-pairing in solvent exposed areas generally do not contribute to the overall binding event due to desolvation penalties of the charged species.

4. Compound Evaluations

In this section, some key methods to evaluate how the compounds discussed in the thesis interact with G4 DNA will be listed.

4.1 Fluorescence Resonance Energy Transfer

To study how G4 ligands bind and stabilise G4 DNA, fluorescence resonance energy transfer (FRET) assays are frequently used.[100] In this assay, G4-forming oligonucleotide sequences labelled with two fluorophores, one donor (FAM) and one acceptor (TAMRA), at each end of the DNA sequence is used. In the G4 folded state, the two fluorophores are in proximity to each other, and the donor emission is quenched by the acceptor upon light excitation (496 nm for FAM). If a temperature gradient is applied, the G4 will melt at a specific temperature (T_m), which will cause the donor and acceptor to be far apart, enabling donor emission without quenching (519 nm for FAM) (Figure 5). In this assay, it is important that the acceptor (TAMRA) does not have the same excitation (559 nm for TAMRA) or emission (583 nm for TAMRA) as the donor. The thermal stability of the G4 can then be measured in the absence and presence of different G4 ligands to record how each ligand affects the thermal stability of the G4 (ΔT_m).[46,100,101] Thus, if a ligand affects the stability of the G4, the ΔT_m will be greater in the presence of the ligand than in the G4 alone.

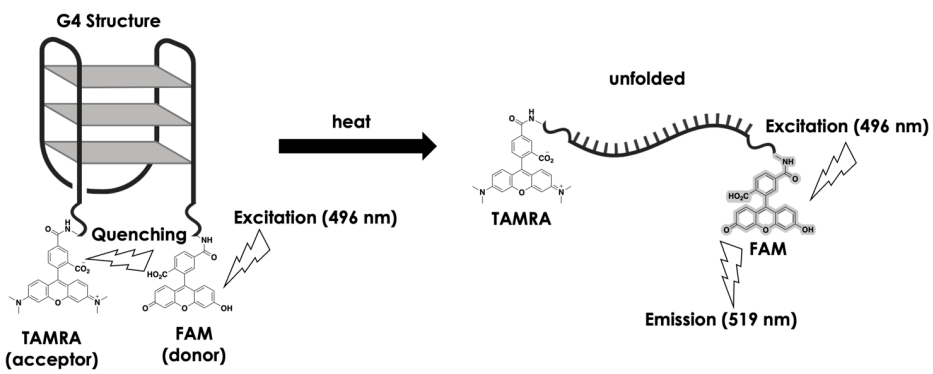


Figure 5. Visual illustration of a FRET melting assay concept.

4.2 Microscale Thermophoresis

Microscale thermophoresis (MST) is a powerful technique to measure the binding affinities (K_d) between a compound and G4 DNA. MST measures the movement of fluorescently labelled molecules (in this case fluorescently labelled DNA) in a microscopic temperature gradient applied by an IR laser (Figure 6A and 6B). Once the IR laser is turned on, the labelled molecules will move out of the heated area (thermophoresis) until complete migration (or simply equilibrium). Once the laser is turned off, the molecules will return to the initial distribution (Figure 6B). Unbound and bound labelled targets will move differently depending on size, hydration shell, and charge. Therefore, the titration of a ligand to the labelled target will allow for quantification of the bound target since a difference in thermophoresis will occur dependent on how much of the target that is bound at different concentrations of added ligand (Figure 6C). Subsequently, a K_d value can be obtained for each ligand by plotting the equilibrium fluorescence from the MST curves against ligand concentration (Figure 6C). In this thesis, the method requires a G4 sequence labelled with a fluorescent dye (typically Cy5) at either end of the sequence. In addition, the ligand should not affect the fluorescence of the dye. If the ligand, however, does induce a fluorescence quenching upon binding to the G4, this quenching can be plotted against concentration to afford the K_d, as an alternative to MST.

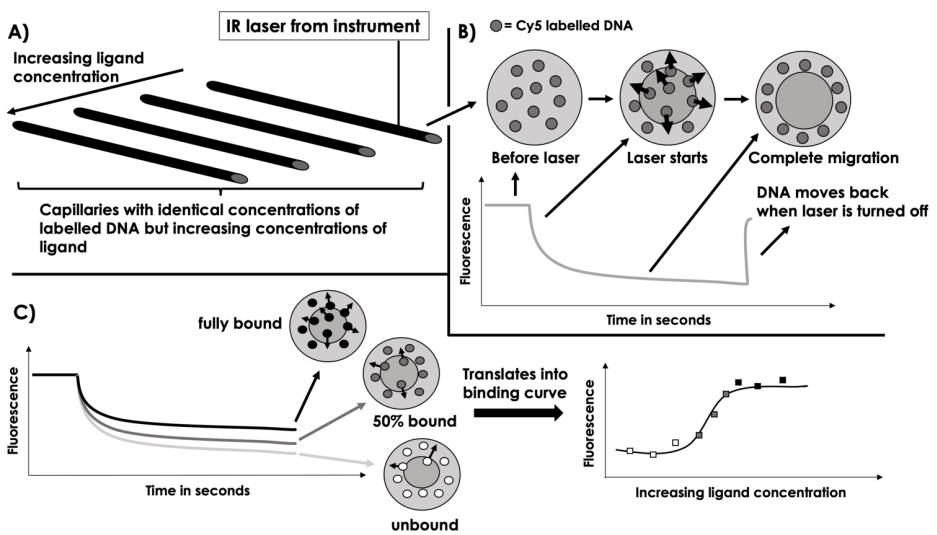


Figure 6. A) In the MST assay, capillaries with identical concentrations of a fluorescently labelled DNA are prepared and varying amounts of ligand are present in each capillary. The IR laser heats up one capillary at a time and measures the thermophoresis of the labelled molecules. **B)** Illustration of how the thermophoresis event in the capillary translates to an MST diagram. **C)** How different concentrations of bound target molecule will result in different degrees of thermophoresis, and consequently different MST curves. The thermophoresis equilibrium fluorescence can then be plotted against the ligand concentration to generate a binding curve.

4.3 Compound Calculations

Complementary to the various *in vitro* experiments used in this thesis, computational techniques were applied. Molecular mechanics (MM) simulations have been used to compare similarities and differences in the compounds' conformational preferences. In addition, quantum mechanics (QM) calculations were used for structural optimisations and the generation of ESP maps. All calculations were performed in the Maestro software.[102] The MM simulations were performed using the OPLS3e[103] force field, as implemented in macromodel.[104] Reported procedures were applied for macrocyclic compound conformation sampling.[105] Geometry optimisations and ESP maps were generated using density functional theory (DFT) calculations with the B3LYP-D3[106–108] basis set as implemented in Jaguar.[109]

4.4 NMR Titration Assay

The tightly bound protons involved in the hydrogen bonding network within each G-quartet, also called imino protons, in G4 structures are not exchanged with the solvent. Therefore, these protons can be observed by ^{1}H NMR recordings as well-defined proton signals, shown by the NMR spectrum of *c-MYC* G4 DNA Pu22 (Figure 7). Each signal corresponds to an imino proton associated with a specific guanine in the G4 structure. Thus, titration with a G4 ligand can upon binding change the chemical environment where it interacts, and consequently, change the imino-proton signals. From this, information regarding if and where the compound bind to the G4 can be acquired.

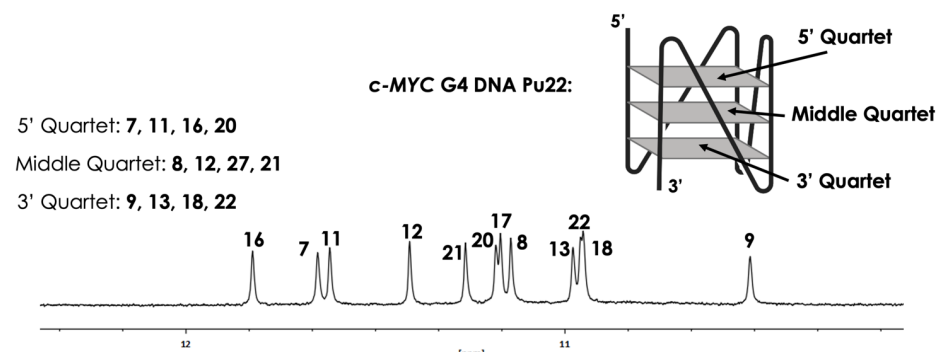


Figure 7. ^{1}H NMR spectrum of c-MYC G4 DNA Pu22 measured at 850 MHz, only showing the region (10-12 ppm) containing the signals of the imino-protons. Which signal that corresponds to which guanine and in which quartet (layer) of the G4 structure is shown to the left. The numbers refer to which number the guanine has in the DNA sequence ($TGAG_3TG_3TAG_3TG_3TA_2$).

5. Aims of the Thesis

The work summarised in this thesis aimed to a) discover and develop new ligands with ability to strongly bind and stabilise G4 DNA structures, b) increase the understanding of which interactions that are critical for ligand-to-G4 binding and selectivity, and c) decipher how such knowledge can be used in the advancement of G4 ligands in therapeutic applications and to further unveil their biological functions.

Specifically, the thesis aims to investigate:

- If macrocycles can be used as baseline design of novel G4 ligands.
- The roles of cationic residues in G4 ligands.
- The interplay of dispersion and electrostatic energetic components in the arene-arene interactions between G4 DNA and ligands.
- If detailed descriptors can be identified that describe interactions between G4 ligands and G4s, and if those can aid in the development of G4 ligands with strong potency.
- How the challenge of selectivity for specific G4 structures can be overcome.

6. Investigations of Macrocycles as an element in G4 Ligand Design (Paper I)

6.1 Introduction: G4 Ligand Starting Point

Studies developed before 2018 in the research group disclosed compounds **11**-**13** (Figure 8) as potent G4 binders with a strong impact on G4 stabilisation.[110,111] Compound **11** suffered from a too-rigid framework that extended the aromatic systems beyond the G4 surface, prohibiting binding. Hence, the greater flexibility generated in **12** by disconnecting the bis-indole C-2 bond afforded a superior binder. The scaffold was further optimised by introducing the bis-indole C-3 methyl-linkage (**13**), allowing for satisfactory placement of the ligand atop the G4 surface. Despite this, compound **13** still suffered from choice of a linear conformation, deviating from its presumed crescent bioactive conformation. Promising studies have previously reported on macrocyclic G4 ligands,[112–117] However, the expansion of such studies was still limited, leaving much to explore in terms of using macrocycles as a design feature in G4 ligand development.[46] Thus, a macrocyclic design was hypothesised as a promising strategy to improve this (**13**) compound class.[83] Furthermore, the macrocyclic design was attractive since the G4 ligands must bind to the flat and featureless G4 surface.[87,90]

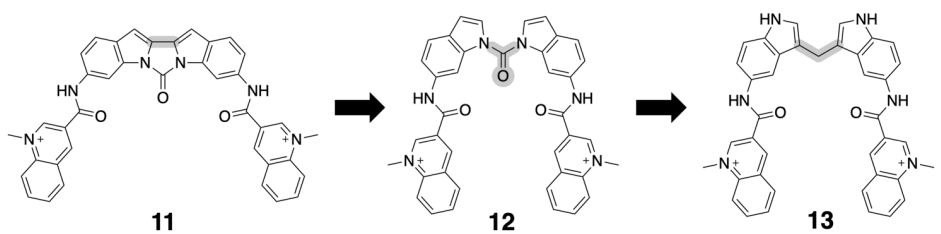


Figure 8. Structures of compound developments **11**-**13** with highlighted structural changes.

6.2 Results & Discussion

6.2.1 The Macrocyclic Design

The novel macrocyclic design was envisioned through a di-amide linkage between the quinolines either in the 7- or 6-position as shown in Figure 9. These

features would allow us to consider different macrocyclic conformations and sizes by varying the linker length. Furthermore, we wanted to investigate the impact of an aliphatic amine incorporated in the linker. The two macrocyclic analogues are summarised in Figure 9A. The retrosynthetic analysis (Figure 9B) first lead to the disconnection of the macrocycle into the di-indole amine and the di-quinoline acid. We anticipated that the macrocyclization event could be favoured by dilute reaction conditions and that the second intramolecular amide cyclisation would be faster than another intermolecular coupling. The di-indole scaffold was planned using a S_NAr condensation and subsequent hydrogenation starting from 5-nitroindole.[118] The plan for the di-quinoline synthesis was based on a di-amide coupling to connect the linker to the quinolines, which in turn could be synthesised from appropriate starting materials.

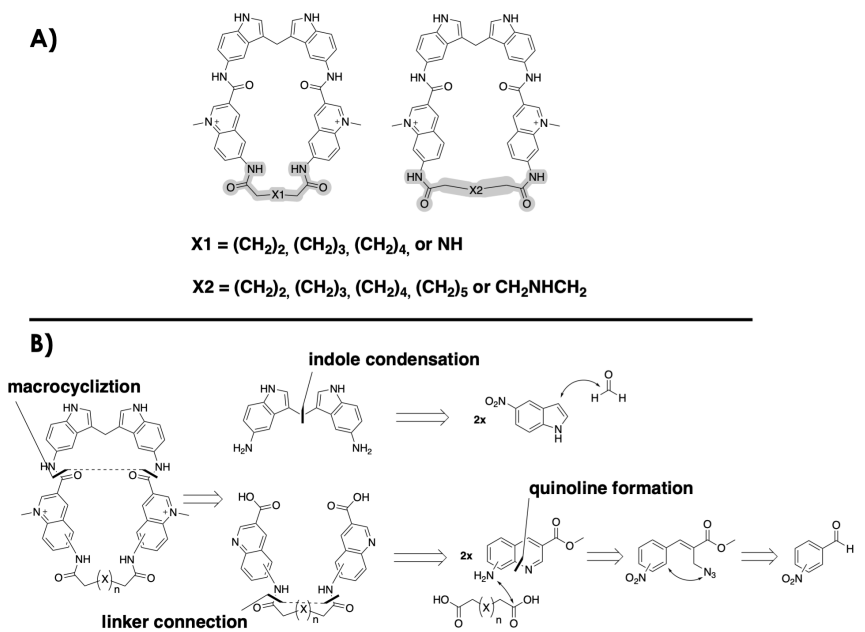
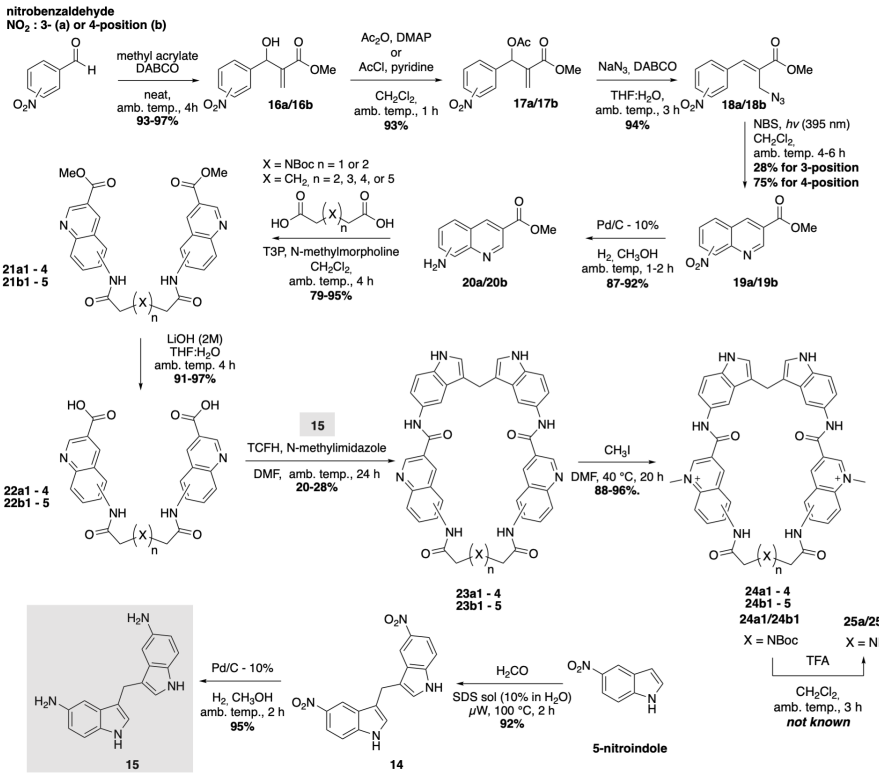


Figure 9. A) Summary of the macrocyclic analogues and their various linkers. **B)** Retrosynthetic analysis of the macrocycles.

6.2.2 Organic Synthesis

The synthesis of the di-indole part (**15**) was achieved from 5-nitroindole through condensation with formaldehyde in water with 10% sodium dodecyl sulphate (SDS) as an emulsifier to afford **14** in 92% yield after purification. Catalytic hydrogenation of 14 afforded **15** in 90% overall yields. The bis-quinoline synthesis was commenced from either 3- or 4-nitrobenzaldehyde to deliver the azide intermediates (**18a**/**18b**) in overall yields >80%. Next, an NBS radical-

mediated cyclisation[119] was applied to afford the quinolines **19a**/**19b**. Starting azide **18b** is symmetric and can, therefore, cyclise to give a single regioisomer. Azide **18a**, however, can cyclise to give either the 8- or 6-nitro product, 6-nitro (**18a**) being the desired product. Unfortunately, the major product of the cyclisation was the 8-nitro. In the reported literature,[119] the 6-nitro product was reported as the major product. The reason for the observed difference might be that the reaction scale in the reported procedure was 0.5 mmol while these reactions were carried out on a 5.3 mmol scale. Running the reaction at a reported 0.5 mmol (130 mg of starting compound **18a**) did not make sense from a perspective of time and to get enough material through to the next steps. Despite the screening of different radical initiators (AIBN, TBHP, benzoyl peroxide), and replacement of NBS with NCS or NIS, the initial conditions gave the best results. The Baran lab has made extensive studies of the radical additions on heteroarenes[120] where π-electron-withdrawing groups strongly direct the radical additions on the ortho positions. This could provide some basis for the observed regioselectivity. Since those studies were conducted on heteroarenes however, the same rules might not be applicable to the system in **18a**. It is worth mentioning that this reaction requires dilute conditions (20 mM) to avoid cross-reactions instead of cyclisation. Reduction to the anilines (**20a**/**20b**) was achieved through catalytic hydrogenation. If the reaction times in this step were extended, partial reduction to the dihydroquinolines was observed. For **20a**, both regioisomers of the nitroquinolines from the previous step were reduced to the aniline and then separated on silica. The reason for this is that the two regioisomers of **20a** (6- and 8-aniline) have dramatically different polarities on silica, likely attributed to the internal hydrogen bonding between the 8-NH_2 and the quinoline nitrogen. Synthesis of the different di-quinoline acids (**22a1-4**, **22b1-5**) was attempted next, with the amide coupling of either aniline (**20a**/**20b**) with their respective diacid linkers. The synthesis is shown in Scheme 1.



Scheme 1. Total synthetic strategy for the macrocyclic compounds.

A few typical amide coupling reagents were investigated: EDC•HCl, PyAOP, and HATU. However, T3P[121] gave superior results, providing the diester quinolines (**21a1-4**, **21b1-5**) in 79-95% yields without the need for column purification. The di-acids (**22a1-4**, **22b1-5**) were obtained by aqueous LiOH-mediated ester hydrolysis without purification in yields between 91-97%. With the diacids (**22a1-4**, **22b1-5**) in hand, we perused the construction of the macrocycles (**23a1-4**, **23b1-5**). Analogous to the amide coupling screen for the synthesis of compounds **21a1-4**/**21b1-5**, we tried different amide coupling reagents (EDC•HCl, PyAOP, HATU, and T3P) at a dilution of 2 mM.[122] None, however, proved successful, and only trace amounts of macrocycle could be observed, whilst polymeric products or unreacted starting materials instead were obtained. We then moved on to try TCFH and N-methylimidazole as these conditions had been reported to be successful for challenging amide couplings.[123] These conditions gave cleaner reaction outcomes and full consumption of the starting materials. Still, mostly polymers were obtained and after extensive optimisation, we ran the reaction at a 2 mM concentration and added the di-indole slowly with a syringe pump which afforded the macrocycles (**23a1-4**, **23b1-5**) in yields between 20-28%. These yields were deemed sufficient

due to the fact the two amide couplings were performed in the same reaction. Subsequent methylation with CH_3I afforded the final macrocyclic products (**24a1-4**, **24b1-5**). Finally, the Boc deprotection of **14a1**/**14b1** to afford the deprotected **25a**/**25b** was conducted. We could, however, not obtain clear NMR spectra for these compounds to characterize them completely. We did however find the correct mass for both compounds with LC-MS analysis and choose to include them in the initial FRET assay anyway. The total synthesis is shown in Scheme 1 and the obtained macrocycles are summarised in Figure 10.

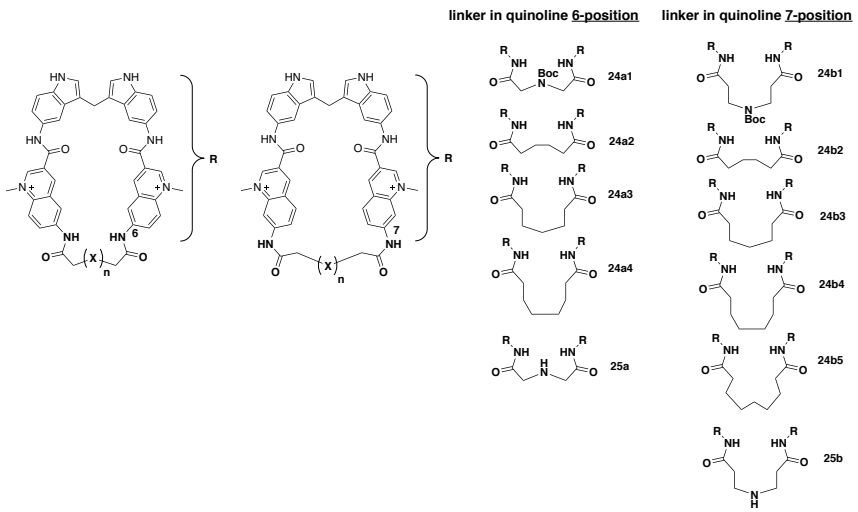


Figure 10. Summary of the synthesised macrocycles.

6.2.3 FRET Melting Assay

To evaluate the macrocycles' abilities to bind and stabilise G4 DNA, a FRET melting assay was applied (see details in section 4.1). The effect on the G4 structure Pu22 (*c-MYC*) showed that all compounds could stabilise the G4 structure (Figure 11A). The macrocycles connected in the quinoline 6-position with a shorter linker were superior (**24a1-2**). This is likely connected to the macrocyclic fit on the G4 surface in combination with less flexibility and a better preorganisation in the macrocyclic structure, e.g. compare macrocycles **24a1** and **24a2** to macrocycles **24a4** and **24b2-5**. Interestingly, non-macrocyclic analogue **13** had a similar ability to bind and stabilise the G4 structure as **24a2**, showing that the macrocyclic design did not make a big difference for this compound. However, the macrocycles did not have an impact on the thermal stability of double-stranded DNA, which was observed for **13** and more pronounced at the higher concentrations (Figure 11B). The amine macrocycles **25a**/**25b** showed a moderate impact on thermal stability, which was surprising due to their smaller ring size like the best macrocycles (**24a2**/**24b2**). The low

stabilising ability of **25a**/**25b** could have been connected to a solvation penalty of the amine upon binding or linked to other factors as **25a**/**25b** were never properly characterised. The Boc-protected macrocycles (**24a1** and **24b1**) also had a significant impact on the thermal stability of the G4 DNA, likely attributed to their smaller macrocyclic size, similar to **24a2**. Macrocycle **24a1** displayed the strongest effect on G4 stabilisation but did also show a significant effect on double-stranded DNA. The compounds were also tested on other G4 DNA structures Pu24T (*c-MYC*) and cKIT2 (*c-KIT*), which revealed the same trends for all compounds.

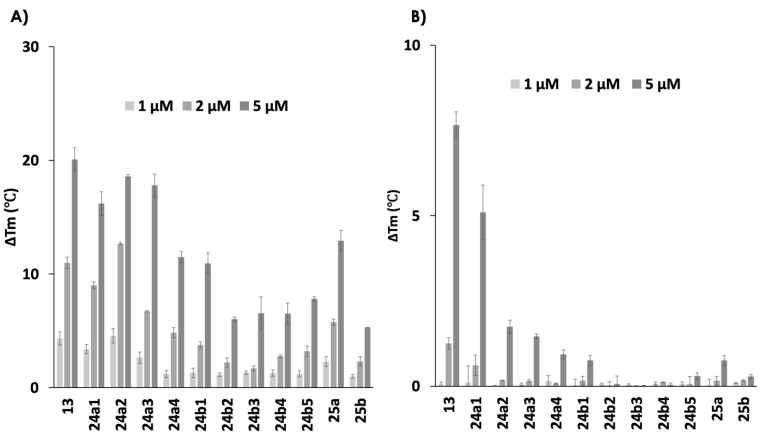


Figure 11. FRET melting assay at 1, 2, and 5 µM of the compound for **A)** Pu22 *c-MYC* G4 DNA, **B)** double stranded DNA, showing the induced thermal stability (ΔT_m) for each compound. DNA concentration is 0.2 µM. T_m in absence of ligands for Pu22 *c-MYC* G4 DNA is 64.3 ± 0.2 °C and ds DNA is 68.0 ± 0.2 °C. Error bars correspond to SD of at least three independent experiments.

The modest difference between compound **13** and macrocycle **24a2** in terms of G4 binding could be rationalised by looking at the conformational preferences of the two compounds. Compound **13** (Figure 12A) has a limited conformational space, where the lowest energy-populated states resided either in the U-shaped form or the opened form. There was an energetic preference for the U-shape, but less than a 4 kcal/mol difference. For the macrocyclic analogue **24a2** (Figure 12B), the compound has a very similar conformational preference to **13**, except that it cannot exist in an open form because of the macrocyclic structure. This similarity in conformational preference can be the reason why there is not much gain in terms of binding for the introduction of the macrocyclic design in this compound class. Furthermore, the di-indole fragment imposes a bent V-shaped structure (Figure 12A and 12B), which constrain the compounds in a twisted shape which likely also reduces ability to bind the flat G4 surface. This will be discussed more in paper II.

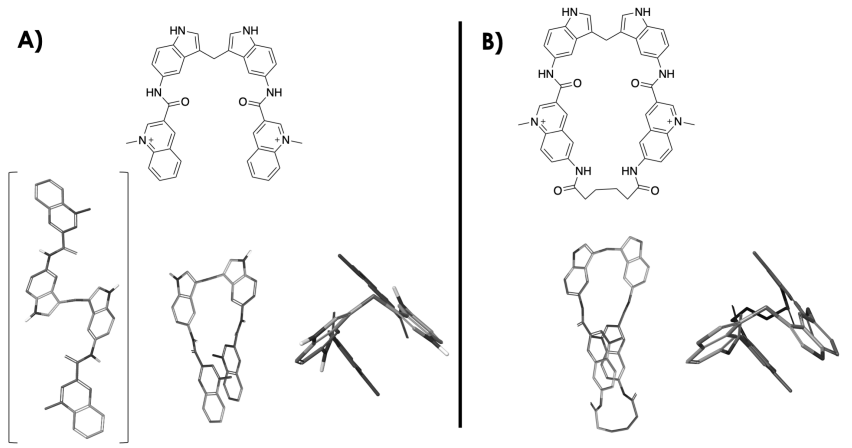


Figure 12. A) The calculated conformational preference for compound **13**. **B)** The calculated conformational preference for macrocycle **24a2**.

6.2.4 Binding Affinities

To assess crude binding affinities for selected macrocycles (**24a2**, **24a3**, **24b3**, and **24b5**) with different G4s (Pu22, Pu24T, and cKIT2), we used a fluorescence intercalator displacement (FID) assay.[124] In this assay, the dye thiazole orange (TO) is bound to the G4 and emits fluorescence. If TO is displaced by another ligand, the fluorescence emission is lost and based on the K_d of TO, a rough K_d value for each ligand can be measured. Overall, the smaller macrocycle **24a2** was the best binder, with K_d values for the different G4s below 1 µM. Compound **13** did show a better K_d (0.49 µM) compared to **24a2** in the TO assay,[111] which likely can be linked to the macrocyclic design for this compound class that did not match the G4 surface. However, the FID assay showed that **13** was better at displacing TO from double-stranded DNA compared to **24a2**, underscoring that the macrocyclic design showed a better selectivity for G4 DNA over double-stranded DNA compared to the non-macrocyclic design.

To record more exact K_d values for the same macrocycles, we employed an MST-based fluorescence quenching assay with Cy5-labelled G4 DNAs (Pu22, Pu24T, and cKIT2). The results from this assay supported macrocycle **24a2** as the best binder. The K_d values from the FID and MST-based fluorescence quenching assay are summarised in Table 1. Complementary, ^{1}H NMR studies recorded at 850 MHz for the same G4s (Pu22, Pu24T, cKIT2) showed a change in the chemical environment for the G4-quartet imino-protons upon titration with the macrocycles (**24a2**, **24a3**, **24b3**, and **24b5**), corroborating the binding event (see paper I for more details). The titrations did, however, not provide much information about which specific guanines that were affected the most, since all

proton signals corresponding to the guanines in the G4 structure displayed considerable broadening in the presence of increasing amount of added ligand.

Macrocycle	K_d (μM) for Pu24T[a]	K_d (μM) for Pu22[a]	K_d (μM) for c-KIT2[a]	K_d (μM) for Pu24T[b]	K_d (μM) for Pu22[b]	K_d (μM) for c-KIT2[b]
24a2	0.99 ± 0.02	0.23 ± 0.01	0.44 ± 0.02	0.77 ± 0.09	0.90 ± 0.11	0.31 ± 0.03
24a3	1.25 ± 0.04	0.33 ± 0.02	0.54 ± 0.06	1.32 ± 0.24	1.09 ± 0.10	0.35 ± 0.04
24b3	2.31 ± 0.06	0.60 ± 0.04	0.93 ± 0.04	1.54 ± 0.21	0.74 ± 0.09	0.62 ± 0.08
24b5	1.80 ± 0.03	0.68 ± 0.05	1.22 ± 0.08	1.39 ± 0.20	3.56 ± 0.45	n.d.

Table 1. [a] K_d values from the FID assay. [b] K_d values from the MST-based fluorescence quenching assay. n.d. = not determined. Errors correspond to SD of at least two independent experiments.

6.2.5 Selectivity and the Macrocyclic Design

To challenge the macrocyclic design and its selectivity for G4 DNA over double-stranded DNA, we conducted another FRET melting assay. In this setup, a constant concentration (2 μM) of the macrocycle (**24a2**) and labelled G4 DNA (0.2 μM) was used, and the melting temperature of the G4 DNA was recorded in the absence and presence of different equivalents (15, 50, and 100 equivalents) of double-stranded DNA. Satisfyingly, the thermal stability of the G4 DNA (Pu22, Pu24T, cKIT2) was unaffected by the addition of double-stranded DNA, underscoring the power of macrocyclization for selectively targeting G4 DNA in the presence of excess double-stranded DNA. Furthermore, **24a2** showed complete discrimination for parallel and hybrid over antiparallel G4 structures, indicating that macrocycles to some extent can be used to gain selectivity for G4s with different topologies (see Figure 1D for description of topologies). Finally, to challenge the concept of macrocyclization and its impact on G4 binding, a more flexible molecule (**26**) was synthesised together with its macrocyclic analogue **27** (Figure 13A). The results from FRET melting assay analysis of these derivatives with Pu24T G4 DNA (Figure 13B) and Pu22 (Figure 13C) clearly showed that macrocyclization can be a powerful tool to improve the G4 binding of a non-macrocyclic compound.

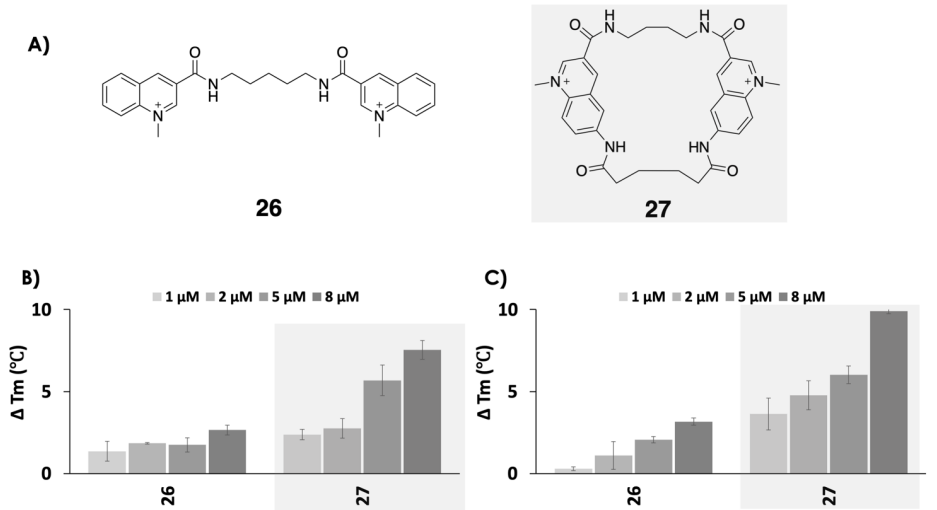


Figure 13. A) Compound **26** and a macrocyclic analogue **27**. FRET melting assay at 1, 2, 5, and 8 µM concentrations for **B)** Pu24T G4 DNA, **C)** Pu22 G4 DNA. DNA concentration is 0.2 µM. T_m in absence of ligands of Pu24T G4 DNA is 62.5 ± 0.3 °C and Pu22 G4 DNA is 64.3 ± 0.2 °C. Error bars correspond to SD of at least three independent experiments.

6.3 Summary Paper I

Synthetic procedures to several novel G4-binding macrocyclic compounds were developed based on a non-macrocyclic G4 ligand **13**. Macrocycle **24a2** showed the best ability to bind and stabilise several G4 DNA structures. The G4 binding and stabilisation proved to be sensitive to small variations in the ring size of the macrocycles and the smaller size of **24a2** was crucial for optimal binding to G4. Overall, the macrocyclic design of this compound class resulted in similar G4 binding compared to non-macrocycle **13**, likely caused by comparable inherent structural preferences in **13** and macrocycle **24a2**. However, macrocycle **24a2** could in contrast to **13** discriminate G4 DNA over double-stranded DNA and display selectivity for parallel and hybrid over antiparallel G4 structures. Finally, the macrocyclic strategy proved very valuable also in terms of improving G4 stabilisation in more flexible compounds, as showcased with the non-macrocyclic compound **26** and macrocycle **27**.

7. Using a Macrocyclic Design to Decipher the Interactions of G4 Ligands (Paper II)

7.1 Optimisation of the Macrocyclic Structure and Novel G4 Ligand Design

Macrocycle **24a2** had strong abilities to bind and stabilise G4 DNA. However, the bent v-shape conformational preference imposed by the di-indole scaffold rendered the structure non-optimal for binding to the flat G4 surface (Figure 12B and 14A). Hence, the replacement of the di-indole scaffold was envisioned to optimise the macrocyclic conformational preference, and consequently, improve the binding to G4. Along this rationale, the replacement of the di-indole (**24a2**) with a tryptamine scaffold (**28a**/**28b**) was proposed. With the tryptamine, the macrocyclic framework could potentially adopt a more planar arrangement and provide a better overall fit on the G4 surface. Two different linker lengths, four (**28a**) and five (**28b**), were used to investigate if the macrocyclic size was important. The indole system was also proposed to be exchanged for a simple phenyl (**29**) to understand if the indole was important for interactions or if it was mainly the conformational aspects of the indole scaffold that was important for binding.

To expand this study further, the possibility to replace the permanently charged methylated quinoline aromatic system with a neutral aromatic system was explored with the aim to improve physicochemical and PK properties. As outlined in the introduction (section 3.2), arene-arene interactions between ligands and G4s are essential for the binding event. However, any description about the interactions between the aromatic systems in G4 ligands and G4s other than that ligands stack on the G4 surface are explained or discussed in the literature.[46,81] Therefore, we wanted to explore how changing the electrostatic component in the arene-arene interactions between G4 ligands and G4s would affect the binding.[93,95,96] Hence, we sought to changing the electrostatic nature of the arenes in the macrocycles. To do this, the benzofuran scaffold was chosen as a replacement for the methylated quinoline. The reason for choosing the benzofuran was that it represents an electrostatically electron-rich aromatic system. This was in sharp contrast to the very electron-deficient nature of the methylated quinoline system, clearly displayed by the ESP maps in Figure 14B.

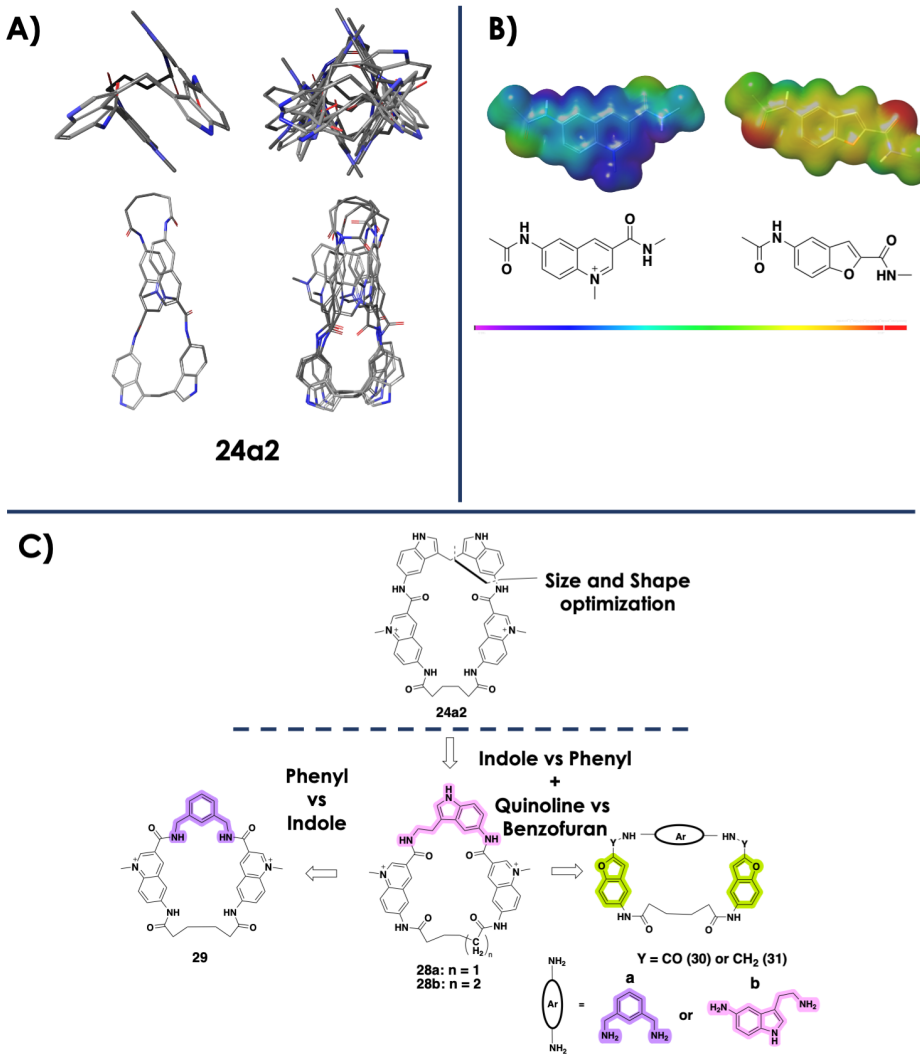


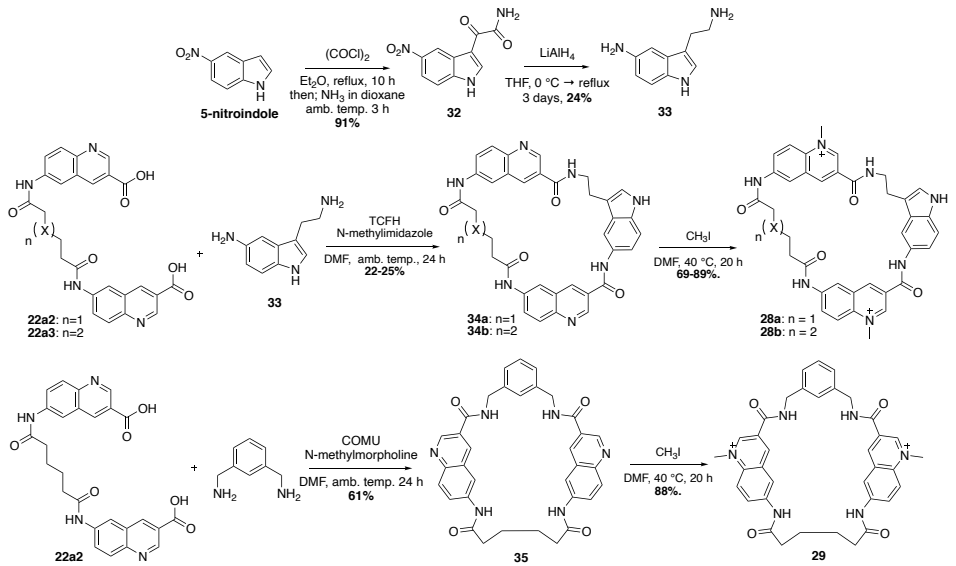
Figure 14. A) Left: The V-shaped di-indole scaffold present in **24a2** (side and top view) forces the entire macrocyclic structure into a distorted twisted shape, not optimal for interacting with the planar G4-surface (represented by the populated lowest energy conformation of **24a2**). **A) Right:** Superimposed image representative of the lowest energy populated states of **24a2**. **B)** Calculated ESP maps for the two fragments (methylated quinoline and benzofuran). The ESP maps are shown with an ISO-value of 0.005 and an energy span of -40-160 kcal/mol. The colour span represents different energy levels going from red (lowest negative, -40 kcal/mol) to blue (highest positive, 160 kcal/mol). **C) top** Exchange of the di-indole scaffold to a tryptamine scaffold could optimise the macrocyclic size and geometry for efficient binding to the G4-surface. **C) bottom** Summary of the novel macrocyclic compounds and their respective structural features included in this paper II.

In addition, benzofurans have been reported as G4 binders.[80] Thus, benzofuran macrocycles with the phenyl (**30a** and **31a**) and tryptamine (**30b** and **31b**) scaffold were proposed. The reason to include an aliphatic amine linkage (**31a** and **31b**) in addition to the amide linkage (**30a** and **30b**) for the benzofuran

macrocycles was to see if cations in G4 ligands contribute to the binding. And in more detail, if this contribution is linked to cation-π interactions[93,96] with the guanines on the G4 surface or if it is mainly an electrostatic effect on the arene electrostatics'. All the planned macrocyclic modifications and their comparisons are shown and highlighted in Figure 14C. In addition, all the non-macrocyclic analogues were also desired to compare the value of macrocyclization.

7.2 Organic Synthesis

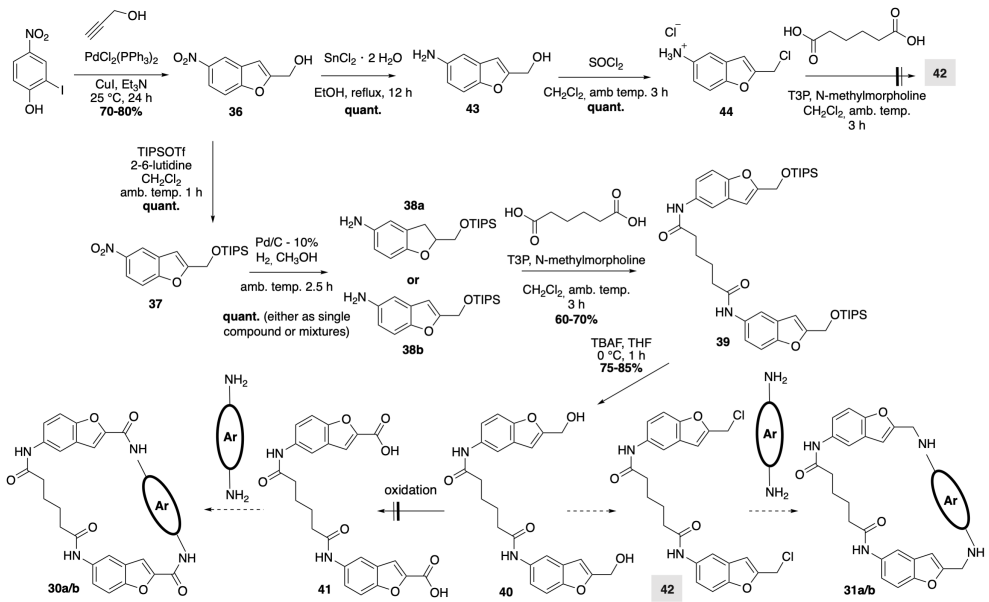
The synthesis of the quinoline macrocycles (**28a**, **28b**, and **29**) was perused first (Scheme 2). Thus, **30** was first assembled through a S_EAr reaction between 5-nitroindole and oxalyl chloride followed by ammonia quenching. Subsequent global reduction of **30** with $LiAlH_4$ in refluxing THF afforded **31** in 24% yield. In the reported literature procedure of **31**,[125] the yield was considerably higher. Despite the higher yields reported in the literature and repeated attempts, tryptamine **31** was never obtained in higher yields, independent of running the reaction in refluxing THF or dioxane. Alternatively, the synthesis of **31** can be accomplished by first reacting 5-nitroindole with the venerable Vilsmeier reagent[126], followed by condensation with nitromethane and subsequent reduction. However, this alternative route did not improve the yield (<20%). Next, the macrocyclization between **22a2**/**22a3** and **31** was performed similarly to paper I. The novel tryptamine macrocycles (**31a** and **31b**) were obtained in yields of 22-25% and subsequent methylation furnished the final macrocyclic compounds **28a** and **28b**. For the macrocyclization between **22a2** and xylene di-amine, different coupling reagents were tested. Ultimately, COMU®[127] was discovered to be the ideal choice, providing satisfying yields of 61% with a similar reaction setup as for the other macrocyclizations. To note, since macrocycles **28a** and **28b** were already obtained, these reaction conditions were never explored for their macrocyclization reactions.



Scheme 2. Synthetic strategy for quinoline macrocycles **28a**, **28b**, and **29**.

To construct the benzofuran macrocycles (**30a**/**b** and **31a**/**b**), it was hypothesised that benzofuran diol (**40**) could be oxidised into the di-acid (**41**) and subsequently converted into the benzofuran amide macrocycles (**30a** and **30b**). Correspondingly, diol **40** could also be converted into dichloride (**42**) and reacted with either diamine to furnish the benzofuran amine macrocycles (**31a** and **31b**). Hence, 2-iodo-4-nitrophenol and propargyl alcohol was subjected to Larock-type conditions to form benzofuran **36** in 70-80% yield, followed by subsequent TIPS protecting of the alcohol to generate **37**. A seemingly simple reduction of the nitro group with catalytic hydrogenation proved challenging. The reaction gave mixtures of desired product **38b** and dihydrofuran **38a**. On a few occasions, **38b** was obtained as a single product without any obvious rationale, and the reaction outcome remained highly irreproducible. Alternative conditions with catalytic Pt/C and cyclohexadiene[128] were attempted, but this only worked in a reaction scale <50 mg, non-optimal for progression. Despite this, progress with the available material of **38b** was attempted, and T3P-mediated di-amide coupling afforded **39** in 60-70% yield. After TIPS deprotection with TBAF, diol **40** was subjected to oxidation. Oxidation of diol **40** proved challenging, and a one-step oxidation (Swern, Dess-Martin, or MnO_2) to the aldehyde was not successful. The inability to generate much material of **40** and the early setbacks indicative of problems with the oxidation warranted a change in strategy. However, one last attempt to utilise this strategy for the synthesis of **31a**/**b** was attempted, circumventing going through problematic intermediate **38b**. Thereof, **36** was reduced directly into the amino alcohol **43** in the presence of $SnCl_2$. Then, deoxychlorination afforded HCl salt **44**. However,

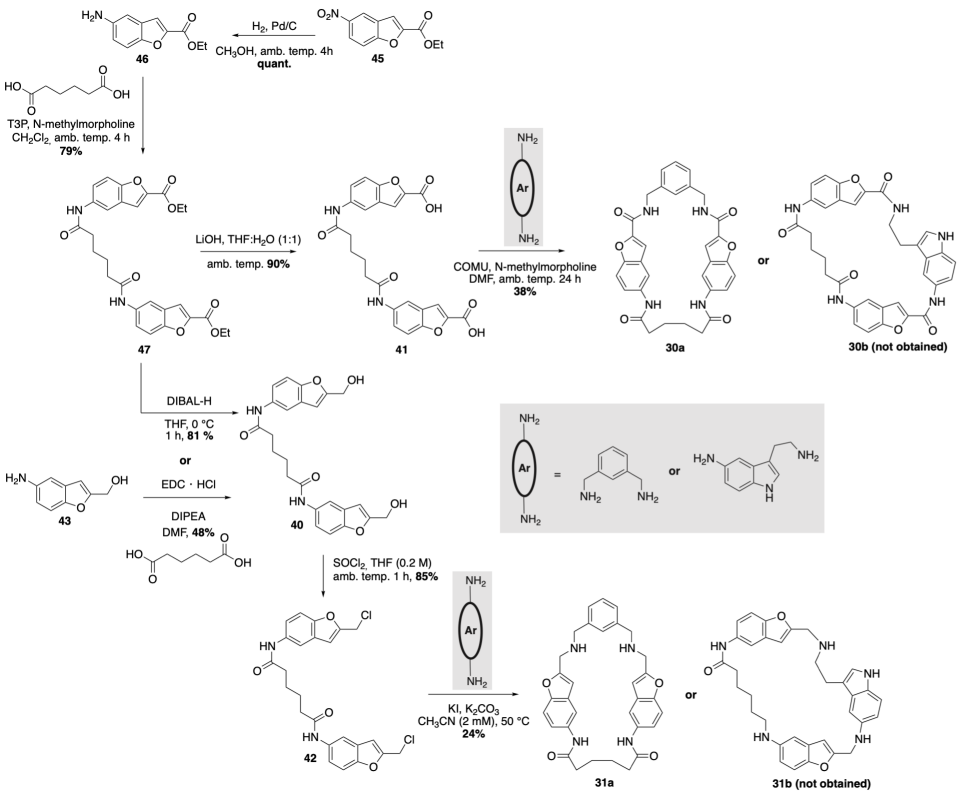
attempts to react **44** into amide **42** were unfruitful, resulting in an unwanted side product resembling a polymer on crude analysis. The problem with this strategy was perhaps, not a surprise considering the daring approach of an intermediate like **44**, having both an electrophile and nucleophile in the same molecule. The failed synthetic strategy is shown in Scheme 3.



Scheme 3. Unsuccessful synthetic strategy for benzofuran macrocycles **30a**/**b** and **31a**/**b**.

After careful reconsideration and revisiting the first strategy, a second approach was devised. It was surmised that starting from the benzofuran ester (**45**) could be a convenient way of accessing either di-acid **41** via hydrolysis or diol **40** through a reduction. In addition, with the electron-withdrawing ester, the furan ring should be less prone to react during the nitro group reduction. Thus, the commercially available benzofuran **45** was first reduced to aniline **46** in quantitative yields. A subsequent amide coupling in the presence of T3P afforded di-ester **47** in 79% yield, and hydrolysis of di-ester **47** afforded intermediate di-acid **41**. Correspondingly, di-ester **47** was reduced with DIBAL-H to afford diol **40** in 81% yield. Slightly baffling, **40** could be obtained directly from benzofuran **43** in the presence of EDC and DIPEA, although in lower yields. Diol **40** could next be chlorinated in the presence of $SOCl_2$ to furnish dichloride **42** in an 85% yield. This reaction was concentration sensitive and if more dilute conditions were used, uncharacterised side-products were obtained, perhaps owing to intramolecular reactions between the monochlorinated product and the remaining alcohol. With each key intermediate (**41** and **42**) obtained, the focus was next turned to the macrocyclizations. First, synthesis of the amide

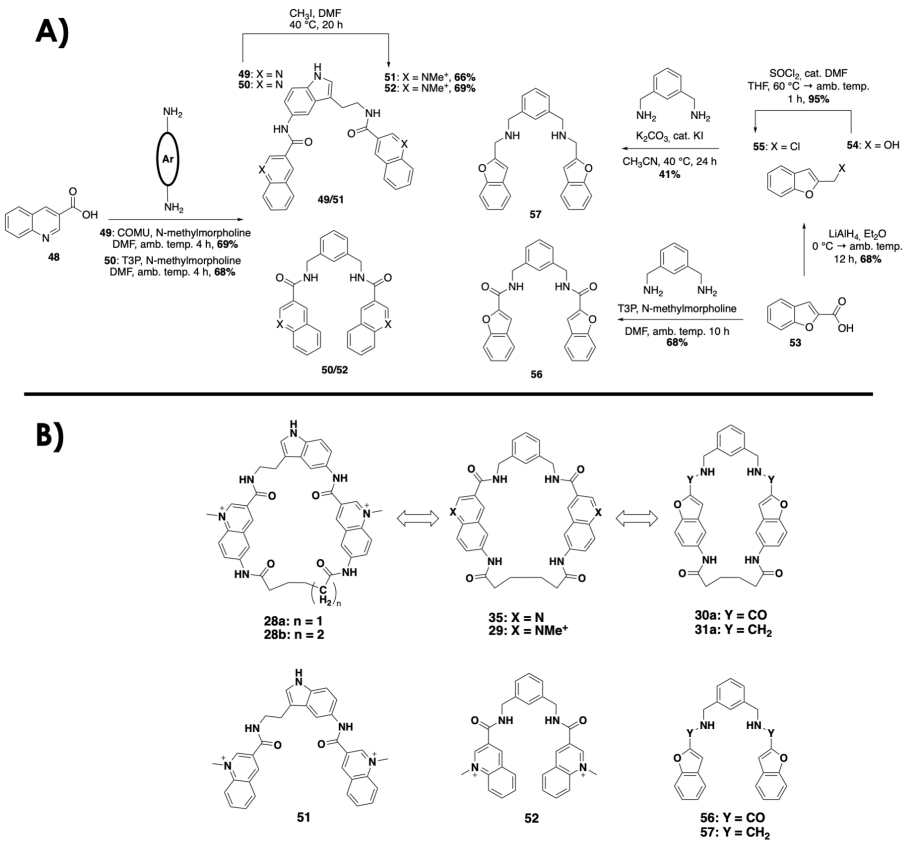
macrocycles (**30a**/**b**) were attempted starting with the xylene di-amine scaffold. Hence, the COMU® reagent was again applied for the macrocyclization, which afforded benzofuran macrocycle **30a** in 38% yield. Despite an extensive screen with different amide coupling reagents like; COMU®, PyAOP, T3P, EDC, HATU, and TCFH, benzofuran amide **30b** could not be obtained nor observed on LC-MS.



Scheme 4. Synthetic strategy for benzofuran macrocycles **30a** and **31a**.

Again, the synthesis of xylene di-amine macrocycle (**31a**) was first attempted. Thus, dichloride (**42**) and xylene di-amine were reacted in THF (2 mM) in the presence of KI and K_2CO_3, but only starting material was obtained. THF was replaced with CH_3CN, and macrocycle **31a** could be observed upon LC-MS analysis, although at a slow rate. The reaction was therefore heated to 50 °C to combat the slow reaction progress. With these conditions, macrocycle **31a** was obtained in 24% yield. The synthesis of tryptamine macrocycle **31b**, like the amide counterpart **30b**, was unfortunately not observed using the same reaction conditions as for the synthesis of **31a**. Synthesis of the benzofuran macrocycles (**30a** and **31a**) are summarised in Scheme 4.

Non-macrocyclic tryptamine compound **51** was synthesised through a double amide coupling between **33** and two equivalences of quinoline-3-carboxylic acid in the presence of COMU which afforded **49** in 69% yield. Subsequent methylation of the quinolines afforded **51** in 66% yield. In a similar fashion, compound **52** was synthesised starting from the amide coupling between xylene diamine and two equivalences of quinoline-3-carboxylic acid in the presence of T3P to afford **50** in 68% yield. Methylation of the quinolines afforded **52** in 69% yield. Benzofuran analogue **56** was synthesised by reacting xylene diamine with 2 equivalences of benzofuran acid (**53**) in the presence of T3P to afford **56** in 68% yield. To obtain compound **57**, benzofuran acid **53** was first reduced to the benzylic alcohol in the presence of $LiAlH_4$ to afford benzofuran alcohol **54** in 68% yield. Subsequent conversion of the alcohol (**54**) to chloride (**55**) was achieved in the presence of the Vilsmeier reagent to afford **55** in 95% yield. Xylene diamine was then reacted with 2 equivalences of **55** in the presence of potassium iodide and K_2CO_3 in CH_3CN at 40 °C to afford **57** in 41% yield. The synthetic strategy for the non-macrocyclic analogues (**51**, **52**, **56**, and **57**) are shown in Scheme 5A.



Scheme 5. A) Synthetic strategy for the non-macrocyclic analogues. **B)** Summary of all the synthesised compounds.

Despite not obtaining either benzofuran macrocycles **30b** or **31b**, the final set of compounds (**28a**, **28b**, **35**, **30a**, **30b**) were still considered adequate to answer the question outlined in the introduction. Macrocycles **28a**/**b** would provide information about the optimisation from starting compound **24a2**. Macrocycle **35** would indicate what effects the indole vs phenyl had on G4 binding. The non-methylated version of **35** (**29**) was also included in the final set. The synthesised compounds are shown in Scheme 5B.

7.3 Interactions with G4

7.3.1 FRET Screen

All the compounds, including starting macrocycle **24a2**, were screened in a FRET melting assay with G4 DNA Pu24T (*c-MYC*) (Figure 15A). Macrocycle **28a**, where the di-indole in **24a2** has been replaced with the tryptamine scaffold, showed a clear improvement in stabilising the G4 structure compared to **24a2**. The improvement is even more evident at the lower concentrations. This supports the notion that the macrocyclic conformation and size in **28a**, where the di-indole in **24a2** has been replaced with the tryptamine scaffold (**28a**), improved binding to the G4 surface. The extension of the carbon linker in **28a** with one carbon (**28b**) results in a drop in G4 stabilisation, rendering it similar to the phenyl analogue (**29**). This data suggests that the indole was more likely to be important for conformational rigidification of the macrocycle and, therefore, favourable for the binding event rather than partaking in specific binding interactions with the G4 surface. Macrocyclization is clearly beneficial for this compound series as shown by comparing **28a**, **28b**, and **29** with their non-macrocyclic analogues (**51** and **52**). The benzofuran amide macrocycle (**30a**) showed poor abilities to stabilise the G4 structure, indicating that the electrostatically electron-rich benzofuran core is not beneficial for binding to the G4 surface. This could be solely due to the benzofuran vs the quinoline core. However, the non-methylated quinoline analogue of **29** (**35**) displays the same inability to stabilise the G4 as **30a**. The benzofuran amine macrocycle (**31a**) did not show any ability to salvage the poor binding seen for **30a**, suggesting that the classical cation-π interactions[93,96] with the G4 surface are not beneficial. This data suggests that an electrostatically electron-deficient arene is essential for the compound to engage in strong arene-arene interactions with the G4 surface.

To conclude that this trend was not simply an artefact for a single G4, different G4s were screened (Figure 15B) for selected compounds (**28a**, **29**, **35**, **30a**, and **31a**). All macrocycles showed some degree of stability towards G4 cKIT1 (*c-KIT*). The same trend was observed for these macrocycles as described in paper I, thus

preferring parallel and hybrid G4 structures over antiparallel. In addition, also these macrocyclic compounds showed selectivity for G4 DNA over double-stranded DNA.

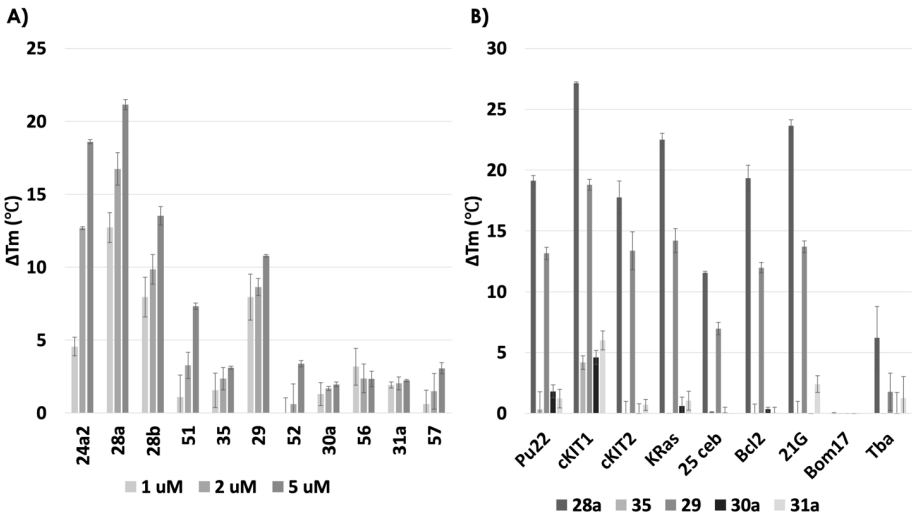


Figure 15. A) The FRET melting assay with all the compounds at 1, 2, and 5 µM concentrations for Pu24T (*c-MYC* promoter) (0.2 µM). **B)** The FRET melting assay with **28a**, **35**, **29**, **30a**, and **31a** at 2 µM concentrations for several G4 DNA structures (0.2 µM). Pu22 (*c-MYC* promoter); cKIT1, cKIT2 (*c-KIT* promoter); KRas (*K-RAS* gene) and 25 ceb (human minisatellite) are parallel G4 forming sequences. Bcl2 (*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. Error bars correspond to SD of at least three independent experiments.

7.3.2 Binding

The new macrocycle **28a** showed satisfactory binding affinities to *c-MYC* G4s (Pu22 and Pu24T) and *c-KIT* G4 (cKIT2) in the sub-micromolar ranges. This data shows considerable improvement compared to macrocycle **24a2**. Macrocycle **29**, with the phenyl scaffold instead of tryptamine, displayed higher binding affinities to the G4s compared to **28a**. This is the same trend as observed in the FRET assay although the difference is even more clear when comparing binding affinity compared to stabilisation, as measured in the FRET assay. Benzofuran macrocycles **30a** and **31a** showed no ability to bind *c-MYC* G4 DNA Pu24T, even at 100 µM, displayed by the fluorescence plots in Figure 16B and 16C. Since no fluorescence quenching of the labelled G4 DNA was observed for **30a** (Figure 16B) and **31a** (Figure 16C), MST measurements were also attempted. However, no binding affinities for **30a** or **31a** were obtained with MST either. All the binding affinities are summarised in Figure 16A.

A)

Macrocycle	K_d (μM) for Pu24T	K_d (μM) for Pu22	K_d (μM) for cKIT2
24a2	0.77 ± 0.09	0.90 ± 0.11	0.31 ± 0.03
28a	0.13 ± 0.01	0.09 ± 0.01	0.13 ± 0.01
29	3.27 ± 0.13	3.45 ± 0.38	5.11 ± 0.96
30a	N. D.	-	-
31a	N. D.	-	-

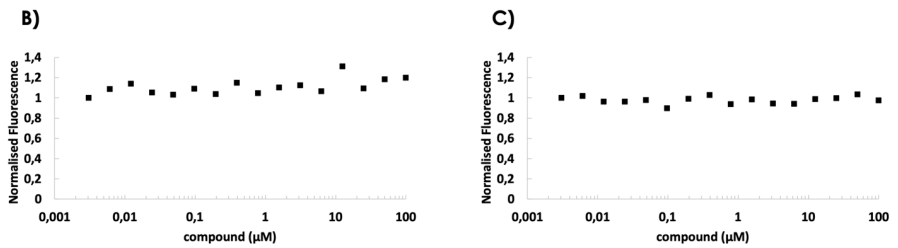


Figure 16 A) Measured K_d-values from the fluorescence quenching assay of different G4 DNA (Pu24T, Pu22, and cKIT2) at 0.25 nM for **24a2**, **28a**, **29**, **30a**, and **31a**. *N. D.* = Not possible to determine. **B)** Fluorescence quenching of 5'Cy5Pu24T at 25 nM for different concentrations for **30a**. **C)** Fluorescence quenching of 5'Cy5Pu24T at 25 nM for different concentrations for **31a**.

7.4 Calculations and PK Properties

To complement the assay data, compound calculations were performed to obtain the lowest energy populated conformational states for selected compounds (**28a**, **28b**, **29**, **30a**, and **31a**). The conformations shown were taken from the 30 lowest energy-populated states (roughly within 5 kcal/mol) that were found more than 15 times in the calculations. More than two conformations, albeit quite similar, were obtained for each compound after this filtering. However, for illustrative purposes, all are not shown (Figure 17A and 17B). All the selected conformations can be found in the supporting information of paper II. Macrocycle **28a**, where the di-indole has been replaced with the tryptamine, possess a more planar conformation (Figure 17A and 17B) compared with the twisted nature of **24a2** (Figure 14A). This is in full agreement with the assay data and underscores the value of computational-aided tools in rational G4 ligand design. For macrocycle **28b**, where the carbon linker has been extended with one carbon, thus making the macrocycle more flexible, the conformational preference is exclusively in a folded rod-like structure (Figure 17A and 17B) that prohibit an efficient fit on the G4 surface. This preference would rationalise why macrocycle **28b** is a worse G4 binder, compared to **28a**, even though the difference between them is only one carbon. Next, macrocycle **29**, which showed similar abilities to stabilise G4 DNA in the FRET assay as **28b**, also has a flexible non-flat conformation (Figure 17A and 17B). The

conformational preference of **29** compared to **28a** (based on a xylene diamine or tryptamine, respectively) suggest that the more rigid indole in **28a** had a constructive effect on the compound geometry that allow better stacking on the G4 surface compared to the flexible xylene diamine in **29**. The benzofuran macrocycles **30a** and **31a** prefer a twisted, more rod-like conformation like **28b** (Figure 17A and 17B). This fact would, however, not solely explain their poor binding to G4.

As a comparison, conformational preferences of the known G4 binder Phen DC3 (**3**), among others (shown in paper II), were conducted. This shows that Phen DC3 likewise prefers a flat and rigid conformation with all the aromatic systems more or less in the same plane (Figure 17A and 17B).

For the same set of compounds (**28a**, **28b**, **29**, **30a**, **31a**, and **3**) geometry optimisations and ESP map calculations were performed. For the compounds with methylated quinolines, the ESP maps have a considerably electron-deficient character, as shown in Figure 17C. in contrast, Benzofuran amide macrocycle (**30a**) displays an electron-rich ESP map, supporting the theory that electron-deficient ESPs in the arenes that bind are essential for strong binding interactions. Benzofuran macrocycle (**31a**) displays a more electron-deficient character due to the protonated nitrogen. However, the G4 binding, and stabilisation data suggest that this does not seem to result in favourable interactions, likely because the cation is located too close to the guanine surface upon binding. However, the effect of the cation in **31a** on the ESP map can partly rationalise why aliphatic amine tails in known G4 ligands,[46,81] that are protonated at physiological pH can improve the binding to G4; by tuning the ESP of the interacting arenes.

Fundamental PK properties of the synthesised compounds revealed that the aqueous solubility of the compounds was low, except for the non-macrocyclic benzofurans (**56** and **57**). Caco-2 measured cell permeability revealed a modest uptake for macrocycle **28a**. The overall poor PK properties highlighted the problematic nature of these types of large polyaromatic systems with permanent charges if applications beyond *in vitro* studies are the aim, despite such compounds being good G4 binders.

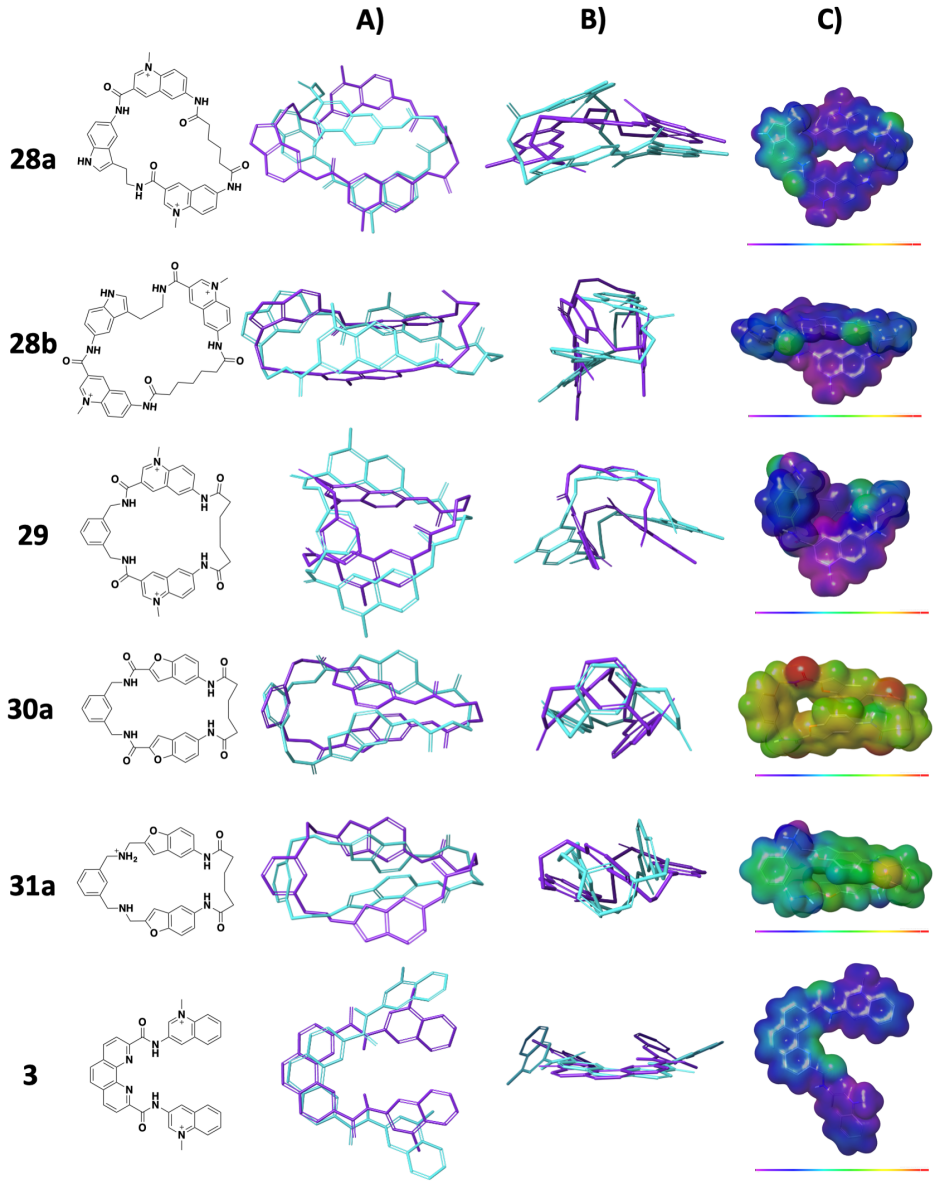


Figure 17. A) Superimposed image representative of the lowest energy populated states for each compound (**28a**, **28b**, **29**, **30a**, **31a**, and **3**) viewed from the top. **B)** Superimposed image representative of the lowest energy populated states for each compound (**28a**, **28b**, **29**, **30a**, **31a**, and **3**) viewed from the side. **C)** Optimised geometry of the compounds (**28a**, **28b**, **29**, **30a**, **31a**, and **3**) showing the ESP map with an ISO-value of 0.005 and an energy span of -40-160 kcal/mol**.** The colour span represents different energy levels going from red (lowest negative, -49 kcal/mol) to purple (highest positive, 160 kcal/mol).

7.5 Summary Paper II

This work disclosed macrocycle **28a** as a potent G4 binder and showed that macrocyclization can be a powerful tool in G4 ligand development and optimisation (**28a** vs **51**). The improved G4 binding upon macrocyclization was attributed to the molecular pre-organisation being closer to the bioactive conformation. Furthermore, the macrocyclic structure made the compounds selective for binding G4 DNA over double-stranded DNA. The comparison between macrocycles **24a2** and **28a** emphasised the value of computational tools in the rational design and optimisation of potent G4 binders. This was further validated in the structural comparison between macrocycles **28b** and **29**. This comparison indicated that the indole was likely not critical for the interactions between the compound and G4 but rather for a better compound conformation and fit on the G4 surface. Hence, the impact that the conformational preference of the macrocycles had on their abilities to bind to the G4 surface underscored the value of computational tools early in G4 ligand developments or optimisations. Furthermore, the results suggested that electron-deficient species, shown by ESP maps, are essential for the compounds to engage in potent arene-arene interactions with the G4 surface. Protonated cations seemed to have a negative effect on the binding interactions when placed directly on the G4 surface. However, protonated aliphatic amines can tune the ESP character of the compounds arene system into a more electron-deficient species (**31a**). Since the ionic interactions between cationic species in G4 ligands and the anionic DNA phosphate backbone should not contribute much to the binding event, due to the high desolvation penalties. Hence, the effect of cations on the G4 ligands ESP could explain why cationic groups in G4 ligands improve the binding to G4. Finally, fundamental PK properties were measured for the compounds. These results underscored the problematic nature of the permanently charged quinolines (poor solubility and low cell permeability) if G4 ligands are to be developed towards therapeutic inventions.

8. The Roles of Aliphatic Amines and the Interplay Between Dispersion and Electrostatics in Ligand to G4 binding (Papers III and IV)

8.1 Introduction Paper III

Paper II suggested that G4 ligands can benefit from cationic groups such as aliphatic protonated amines, which change the ESP map of the compound into a more electron-deficient system. This can be exemplified by looking at the ESP maps of known G4 ligands BRACO-19 (**1**) and Pyridostatin (**6**) (Figure 18A), that both contain aliphatic amines. The general hypothesis for inclusion of aliphatic amines in G4 ligands is that it improves binding through electrostatic interactions between the cationic groups and the anionic phosphate backbone.[46,81] However, ion-pairing in solvent-exposed areas is known to often contribute little or nothing to the total binding energy.[82,98,99] Thus, the rationale that cationic groups strongly improve G4 binding through direct electrostatic interactions with the phosphate backbone can be questioned. With this argument, we proposed that cationic groups instead improve G4 binding by tuning the electrostatic nature of the compound, and consequently boost the arene-arene interactions.

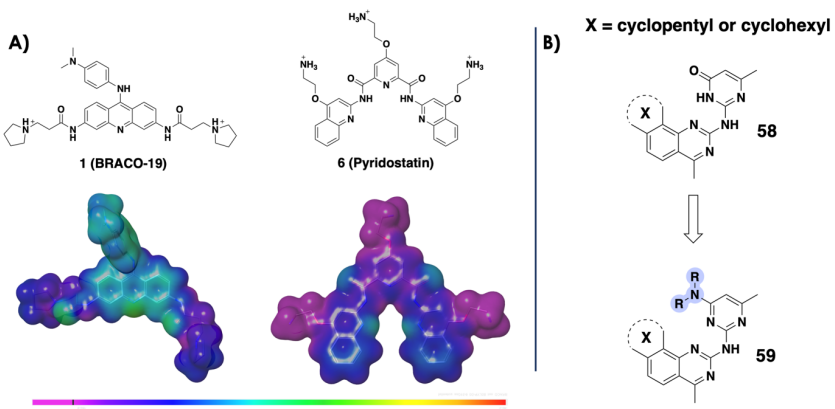
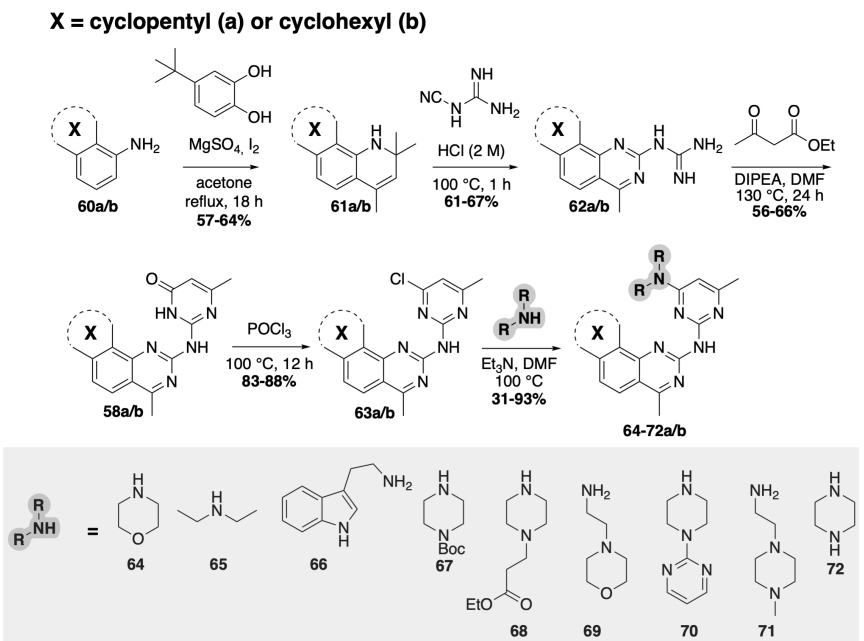


Figure 18. A) ESP maps of the two known G4 ligands BRACO-19 (**1**) and Pyridostatin (**6**), with their aliphatic amine groups protonated. The ESP map is shown with an ISO-value of 0.005 and an energy span of -40-160 kcal/mol. **B)** Known G4 scaffold **58** that would allow introduction of various amine side chain analogues (**59**). The colour span represents different energy levels going from red (lowest negative, -40 kcal/mol) to purple (highest positive, 160 kcal/mol).

To test this hypothesis and gain more in-depth knowledge about the ligand-to-G4 interactions, we decided to use a scaffold known to bind G4[129,130] (**58**, Figure 18B), and append different side chains (**59**) with varying composition (cationic and neutral) and chain length. Scaffold **58** was chosen because this, in contrast to methylated quinolines, is a more attractive central fragment in terms of therapeutic potential; low molecular weight, good solubility, and without permanent charges.

8.2 Results and Discussion Paper III

A set of compounds (**64**-**72**) with various side chains were synthesised for both the cyclopentyl (**a**) and cyclohexyl (**b**) fused quinazoline system. Thus, aniline (**60a** or **60b**) was first reacted with mesityl oxide (formed *in situ*) to form the dihydroquinolines (**61a**/**b**) in yields around 60%. A subsequent ring-opening ring-closing reaction with cyanoguanidine afforded the quinazoline guanidine's (**62a**/**b**) in 61-67%. Condensation between **62a**/**b** and ethyl acetoacetate afforded intermediates **58a**/**b** and deoxychlorination with $POCl_3$, followed by S_NAr, provided the final compounds (**64**-**72a**/**b**) in yields between 31-93%. The synthesis is shown in Scheme 6.



Scheme 6. Synthetic strategy for quinazolines **64**-**72a**/**b**.

The synthesised quinazolines (**64**-**72a**/**b**) were next evaluated for their abilities to stabilise the *c-MYC* G4 DNA Pu24T (Figure 19). The fused cyclopentyl (**a**) or

cyclohexyl (**b**) displayed similar results. The neutral side chains (**64**, **65**, **66**, **67**, and **70**) generally displayed lower abilities to stabilise the G4 structure, with the boc-protected piperazine side chain in **67** sticking out as the weakest stabiliser. However, the tryptamine (**66a**/**b**) did display good stabilisation at higher concentrations, especially **66a**. However, the cationic sidechains are overall superior, foremost **71** and **72**, reaching a high degree of induced thermal stabilisations also at lower concentrations. The G4 stabilisation seem to increase when the cation is positioned further away from the G4 surface, demonstrated by the piperazine chain in **71a**/**b** in comparison with **72a**/**b**. Morpholine (**69**) and ester (**68**), though cationic, are less potent stabilisers. This is likely owed to the larger degree of hydrophobic character in the solvent-exposed side chain, which should cause less favourable interactions with water. Identical trends were observed for the FRET melting assay of the *KRAS* DNA G4 structure.

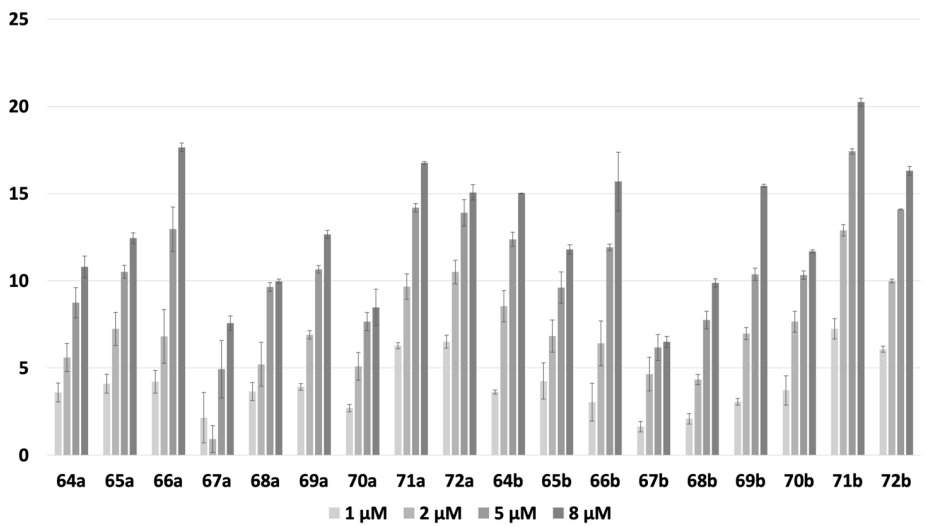


Figure 19. The FRET melting assay with all the compounds at 1, 2, 5, and 8 µM concentrations for Pu24T (0.2 µM).

MM calculations were performed for selected compounds (**66b**, **67b**, **68b**, **71b**, and **72b**), which revealed a concomitant conformational preference (Figure 20A). The low-energy conformations that were found the most times (exclusive for **67b**) were crescent-shaped (Figure 20A), except for **68b** (Figure 20A), where the linear shape was preferred. This could, in part, explain the lower stabilisation seen for **68a**/**b**. Shifting attention to the ESP maps (Figure 20C), the neutral compounds **66b** and **67b** have, as expected, considerably more electron-rich ESPs compared to the cationic compounds (**68b**, **71b**, **72b**). This fact could also explain why **67b** is a less effective stabiliser. The stronger stabilisation observed for the tryptamine (**66b**) is more difficult to rationalise, but due to its flexibility,

the indole can potentially reach additional interactions on the G4 surface to improve the binding. This is, however, purely speculative. Finally, compounds **71b** and **72b** differ in having the piperazine connected directly to the aromatic core (**72b**) or with a two carbon-chain linker in-between (**71b**). The latter is more efficient in stabilising G4s, which is in line with the discussion for the benzofuran macrocyclic analogues in paper II that suggested protonated cations close to the G4 surface to be unfavourable for binding. In addition, the piperazine moiety on a linker can likely partake in solvation interactions during the binding event above the centroid of the solvent-exposed side of the compound, as seen for **71b** in Figure 20A.

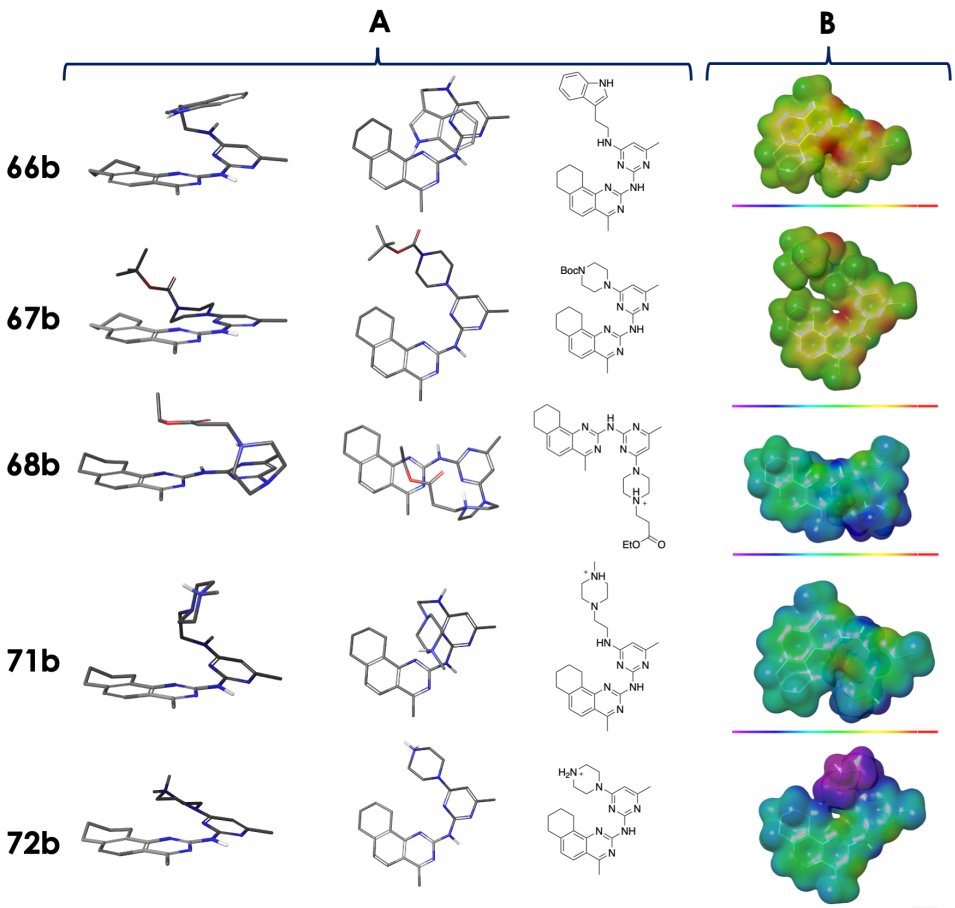


Figure 20. A) MM simulations of selected compounds (**66b**, **67b**, **68b**, **71b**, and **72b**) showing the major populated state of the low energy conformations along with a ChemDraw picture for illustration. **B)** ESP maps of each compound viewed from the side showing the surface of the aromatic core system. The ESP map is shown with an ISO-value of 0.005 and an energy span of -50-140 kcal/mol. Note, both amines in the piperazine ring of **71b** were protonated and calculated for but resulted in more or less identical outcomes. The colour span represents different energy levels going from red (lowest negative, -50 kcal/mol) to purple (highest positive, 140 kcal/mol).

The best compound (**71b**) was next studied together with Pu22 *c-MYC* G4 DNA by using NMR to probe if information about where the compound interacts with the G4 could be obtained. Upon titration of **71b** to Pu22 G4 DNA, the shifting of specific imino protons and consequently also specific guanines could be observed. The affected guanines in the G4 structure are shown in Figure 21A (top view) and 21B (bottom view), illustrated with the Pu22 crystal structure (5W77). On the 5' side (Figure 21A), the two guanines situated in the more accessible hydrophobic space in the crystal structure were the ones affected the most in the NMR spectrum. The two guanines covered by steric bulk (Figure 21A) were also affected less. Similarly, the 3' guanines (Figure 21B) that were more accessible were also the ones more affected, displayed by the orange colour of the three most accessible guanines (Figure 21B). Only one guanine (G12) in the middle quartet displayed a change in the chemical shift, shown in both Figure 21A and 21B. The reason why G12 was the only guanine in the middle quartet that was noticeably affected could be due to its positioning directly below the two strongly affected guanines in the 5' quartet (Figure 21A). Thus, the NMR titration suggests that this compound class bind this G4 at the accessible hydrophobic areas of the G-quartets likely through arene-arene interactions.

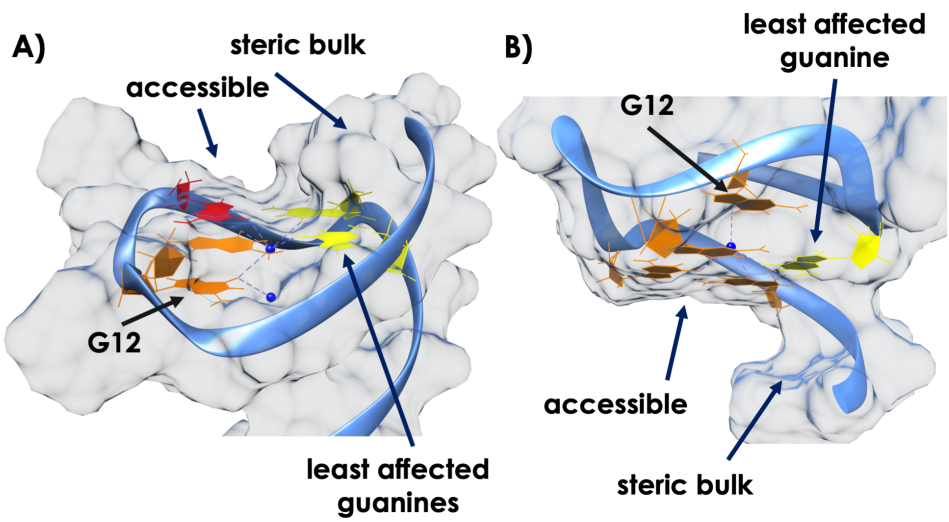


Figure 21. A) 5´, and **B)** 3' views of the affected guanines in the *c-MYC* Pu22 (5W77) G4 DNA structure upon binding of **71b**. Observed chemical shift changes in the NMR spectrum are shown in red (large shifts), orange (moderate shifts), yellow (small shifts), and guanines with no shifts are not shown.

To complement the first set of compounds (**64**-**72a/b**), a new set was synthesised (**73**-**83b**) using the same synthetic strategy as outlined in scheme 6, with foremost different aliphatic amines with varying degrees of chain length (Figure 22A). The FRET melting studies of these compounds (**73**-**83b**) showed that almost any amine that could be protonated at physiological pH was good

for G4 binding and stabilisation. The side chain length is not very sensitive to variations; however, a 3-carbon linker seemed to be the best. In line with previous observations, a lower stabilisation was observed when the cation was placed closer to the arene core of the compound and consequently closer to the G4 surface (like in **73**). A more flexible side chain, e.g., 3 carbons, likely resulted in the protonated amine maintaining solvation interactions during the binding event whilst also affecting the electrostatic character of the arene. A longer side chain (**83b**) should result in more hydrophobic contacts between the additional carbons in the chain and the water, and this could explain why the longer chain is slightly worse for binding. Next, binding affinities for selected compounds (**79**-**82b**) with *c-MYC* G4 DNA Pu24T were recorded with MST. All compounds showed good binding to Pu24T with K_d values between 0.14-0.32 µM, confirming that these compounds are potent G4 binders. Finally, compounds **79b** and **80b** showed promising abilities to target cancer cells over healthy cells, underscoring the value of these compounds as potent G4 binders balanced with good cellular uptake.

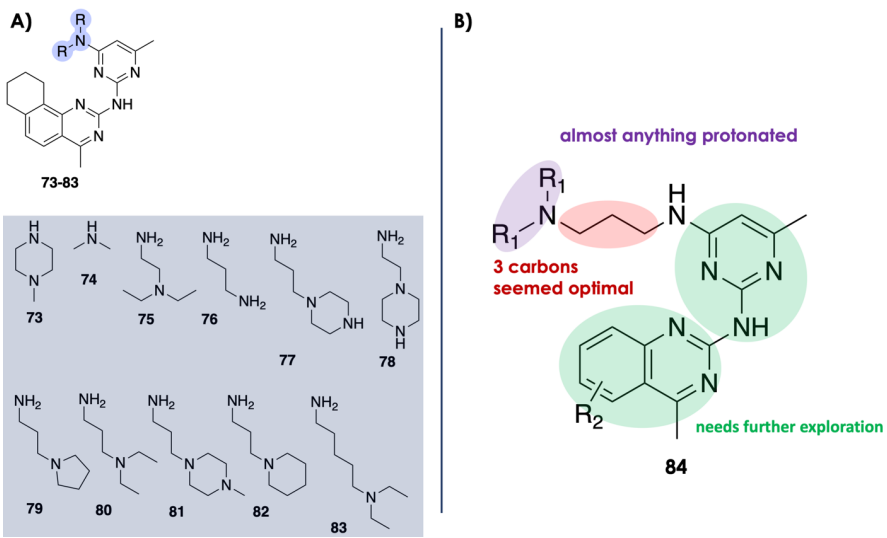


Figure 22. A) Summary of the second series of synthesised compounds (**73**-**83b**). **B)** A general display of scaffold **84**, highlighting the aspects that we proposed to be good for binding and what still needs to be discovered.

8.3 Summary Paper III

This study showed that **84** (Figure 22B) represented a suitable scaffold for the study and development of potent G4 binders with adequate cellular uptake (suggested by the compounds' potent effect on cancer cells). Aliphatic amine side chains with a 3-carbon linker appended to **84** appeared good for binding to

G4 DNA. The protonated amines affected the ESP of the compounds, making them more electron-deficient, clearly shown by the ESP maps. We, therefore, proposed that the inclusion of cationic groups in G4 ligands improves the binding to G4 DNA indirectly by tuning the ESP of the molecule and consequently the electrostatic component in the arene-arene interactions. Further exploration of the central arene fragment is warranted to determine the synergism between the dispersion and electrostatic energetic components of the arene-arene interactions between G4 ligands and G4 DNA structures.

8.4 Introduction Paper IV

In this paper, the focus was to design and synthesise a set of compounds that would provide insight into the dispersion component in the arene-arene interactions between ligands and G4. Dispersion is typically the dominating favourable energetic component in arene-arene interactions (section 3.2). Substituents on one arene increase the favourable dispersion interaction through direct interactions with the other arene partner, and the interactions typically become more favourable with an increasing number of substituents, regardless of electronic character. It is important to remember that steric factors and the positioning of the substituents on the arene also can have an effect.

8.5 Results and Discussion Paper IV

The quinazoline scaffold **84** served as a natural starting point to continue the exploration of G4-ligand interactions with G4 DNA structures. We thus used this central fragment and designed a small series of novel compounds to investigate the dispersion component in the arene-arene interaction between G4 ligands and G4 DNA. Quinazoline compound **85,** carrying no substituents on the central quinazoline core, should have the lowest degree of dispersion interactions with G4s. Sequential addition of methyl substituents on the quinazoline core (**86**-**88**) should increase the interactions with the G4 if dispersion is critical. Correspondingly, the methoxy (**89**) and chlorine (**90**) quinazolines should retain the dispersion interactions, despite their different electronic characteristics compared to the methyl (**87**). In addition to the quinazoline core (**85**), the quinoxaline core (**91**) was chosen. This was to discern if small changes in the aromatic system would result in distinct differences in the arene-arene interactions between the ligand and G4. Same as for the quinazoline core, different substitutions with variable electronic characteristics for the substituents on the quinoxaline core were proposed (**92**-**94**). The amine side

chain for all compounds (**85**-**94**) was kept to a 3-carbon chain length, as disclosed in paper III. Since the nature of substituents on the amine was not paramount, dimethyl was selected. The compounds are summarised in Figure 23A. Moreover, to assure that the discussed modifications are the only factors affecting the binding to G4, conformational and ESP calculations were performed for the compounds (**85**-**94**). Gratifyingly, the compound series shared a uniform conformational preference with the cationic amine side chain extending above one face of the compound, depicted by compounds **85** and **91** (Figure 23B). Likewise, the ESP maps in the presence of the protonated aliphatic dimethylamine were analogous (Figure 23B).

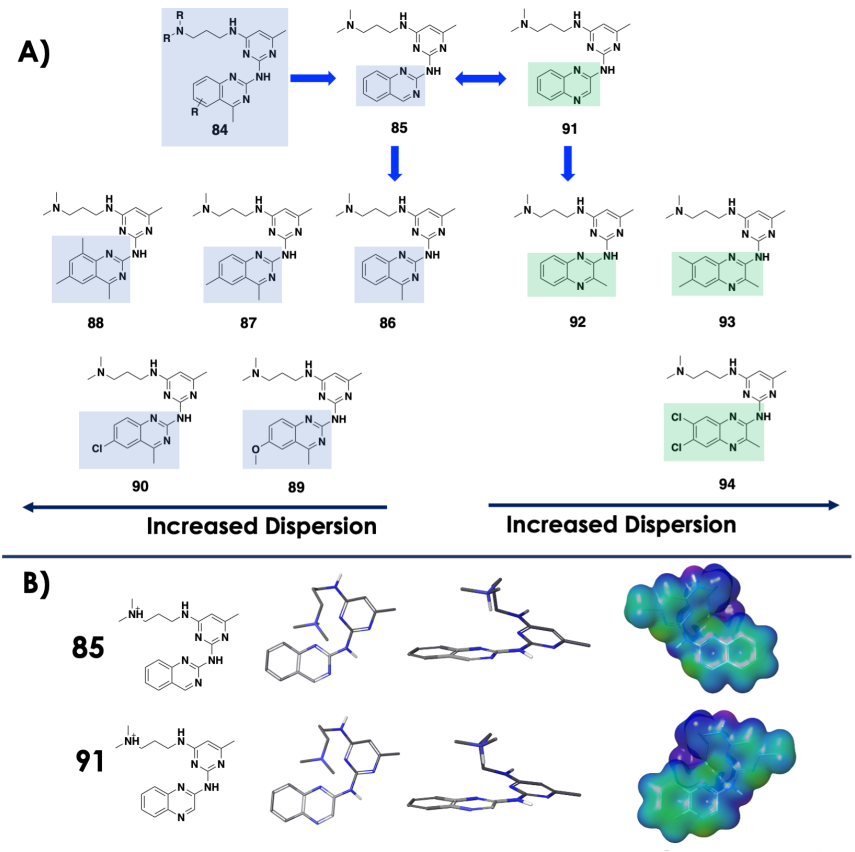
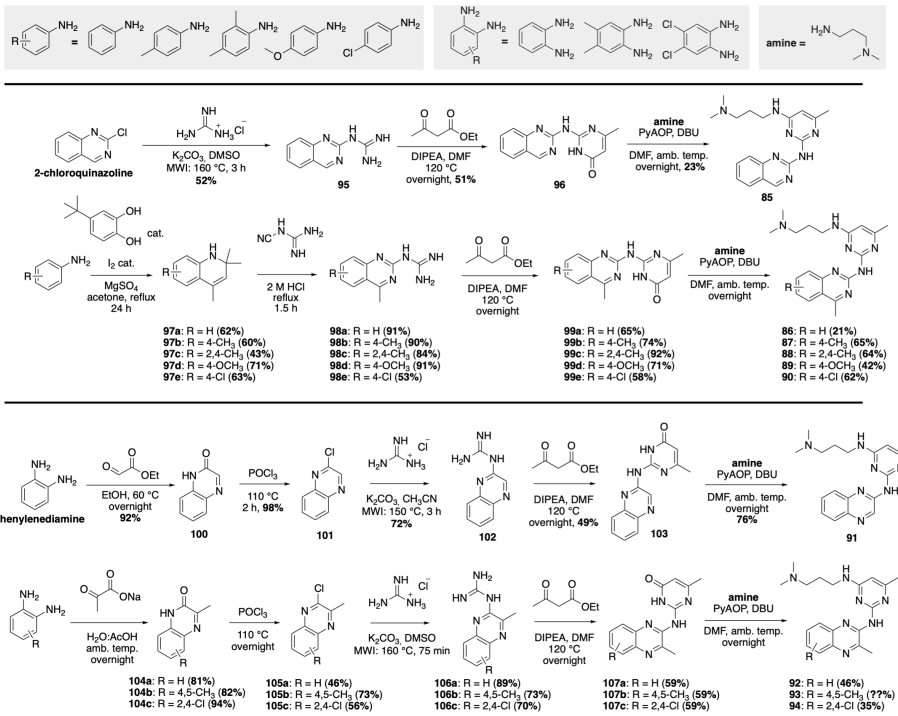


Figure 23. A) Summary of the target compounds either with a quinazoline (**85**-**90**) or quinoxaline (**91**-**94**) core with varying substitution patterns. **B)** Conformational preference of compounds **85** and **91**, showing that the two scaffolds share the same conformational preference as well as electrostatic potential. The ESP map is shown with an ISO-value of 0.005 and an energy span of -40-140 kcal/mol. The colour span represents different energy levels going from yellow/green (lowest negative, -40 kcal/mol) to purple (highest positive, 140 kcal/mol).

8.5.1 Organic Synthesis

Construction of the quinazoline compounds was undertaken first. Hence, 2-chloroquinazoline was reacted with guanidine through a microwave-assisted S_NAr reaction to afford **95** in 52% yield. Subsequent condensation with ethyl acetoacetate followed by a PyAOP-assisted[131] S_NAr reaction to install the amine side chain delivered quinazoline **85** in 12% overall yields. Substituted quinazolines (**86**-**90**) were constructed starting from suitable anilines. Each aniline was reacted with mesityl oxide (formed *in situ*) to afford substituted dihydroquinolines (**97a**-**97e**) in yields between 43-71%. Quinazoline guanidine's (**98a**-**98e**) were obtained in 53-91% and subsequent condensation with ethyl acetoacetate afforded intermediates (**99a**-**99e**) in 58-92% yields. PyAOP-assisted S_NAr finally afforded the quinazolines (**86**-**90**) in yields between 21-65%. The quinoxaline compounds (**91**-**94**) were pursued next. Employing a similar strategy as for the quinazolines, **91** was constructed starting from *o*-phenylenediamine. Thus, condensation with ethyl glyoxylate followed by deoxy chlorination in the presence of $POCl_3$ afforded quinoxaline **101** in 90% overall yields. Then, a microwave-assisted S_NAr followed by condensation, and finally, PyAOP-assisted S_NAr, afforded quinoxaline **91** in 27% overall yield. It is worth pointing out that for the synthesis of **103** from **102**, a side product was formed if the reaction was run for a longer time. This was only observed for compound **102**, likely owed to the unsubstituted C-2 position on the quinoxaline. To secure the substituted quinoxalines (**92**-**94**), the same synthetic strategy as for **91** was applied, starting from the appropriate diamine, and performing the first condensation with sodium pyruvate. The synthesis of all compounds (**85**-**94**) is shown in Scheme 7.



Scheme 7. Total synthetic scheme for the target quinazoline (**85**-**90**) and quinoxaline (**91**-**94**) compounds.

8.5.2 The synergism Between Dispersion and Electrostatic Components

All the compounds were first tested in a FRET melting assay with *c-MYC* G4 DNA Pu22 (Figure 24), and the results were quite striking. First, quinazoline **85**, with no substituents on the quinazoline core, had a clear but moderate effect on the thermal stability of the G4. The addition of one methyl on the quinazoline (**86**) resulted in an increase, almost doubling the effect the compound had on the G4 stability. Adding another methyl (**87**) had a modest increase of around 0.5 °C on thermal stability compared to **86**. The reason for the small difference might be that substituents in that position cannot partake in many additional interactions with the G4 surface. Adding another methyl group (**88**) resulted in an even better compound, with a maximum induced thermal stability of 20 °C. This points in the direction that dispersion is central for the arene-arene interactions between G4 ligands and G4 DNA structures. This was emphasised further by replacing the methyl in **87** with either a methoxy (**89**) or chlorine (**90**) since the compounds retain their ability to stabilise the G4, despite very different electronic characters (methyl vs methoxy vs chlorine).

Changing the location of the nitrogen in quinazoline **85** one step, affording quinoxaline **91**, resulted in a complete loss in thermal stability. The drastic loss in thermal stabilisation as a result of small changes in the arene system underscores that the arene-arene interactions between the ligand and the G4 surface are paramount for ligands binding to G4. It is also interesting that the amine side chain in **91** did not seem to be able to compensate for the loss in activity. In agreement with the quinazoline series, the quinoxalines did advance into modest G4 stabilisers when substituents were added to the quinoxaline core (**92**-**94**), irrespective of their electronic character.

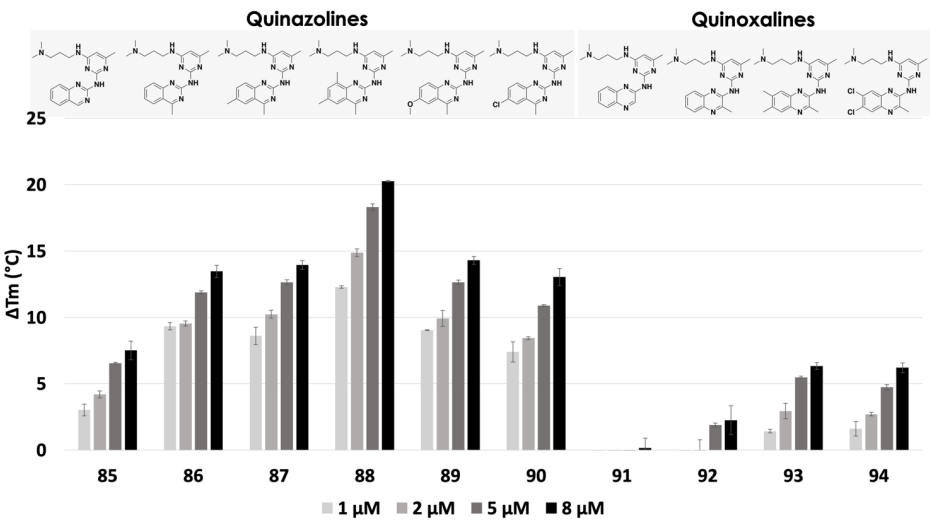


Figure 24. The FRET melting assay with the synthesised compounds (**85**-**94**) at 1, 2, 5, and 8 μM of added compound for *c-MYC* G4 DNA Pu22 (0.2 μM), showing the ability of the compounds to affect the thermal stability of the G4 structure. Error bars correspond to the SD of six independent experiments.

Next, binding affinities for each compound with *c-MYC* G4 DNA Pu22 were recorded using MST (Figure 25). The binding affinities reflected the results from the FRET assay and disclosed potent novel G4 binders. Quinazoline **88**, bearing the most substituents were the strongest binder with a K_d value of 0.05 μM. The two di-substituted quinazolines **87** (methyl) and **89** (methoxy) closely followed, with K_d values of 0.1 μM and 0.09 μM, respectively. Chloro-quinazoline (**90**) displayed slightly worse binding (0.17 μM) and was similar to mono-substituted quinazoline **86** (0.15 μM). Unsubstituted quinazoline **85** displayed a K_d value of 1.4 μM, again underscoring the importance of dispersion interactions for potent G4 binding. In agreement with the FRET data, the quinoxalines showed a worse binding affinity, with the most substituted derivatives (**93** and **94**) showing better binding.

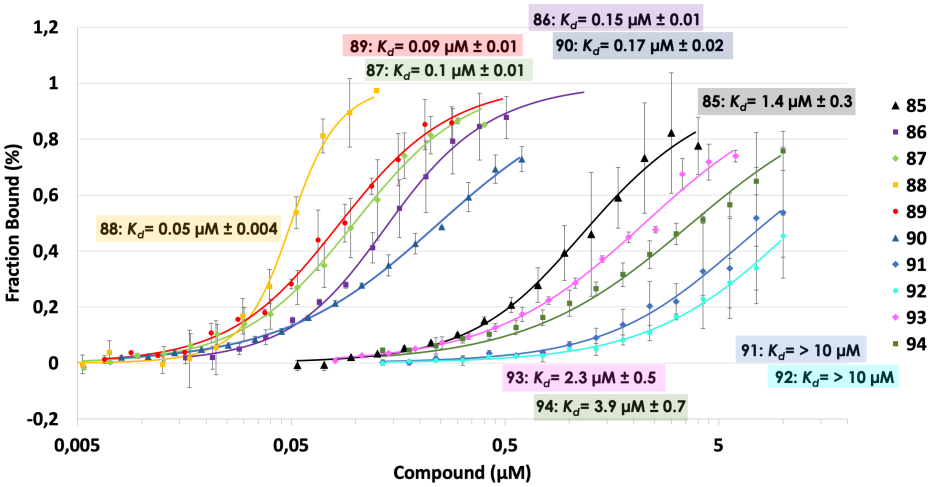


Figure 25. MST binding curves of each compound (**85**-**94**) at varying concentrations (dilution factor: 1:3) with *c-MYC* G4 DNA Pu22 (0.025 µM). The error bars correspond to the SD of two independent experiments. K_d for each compound is written out in the graph and colour coded accordingly.

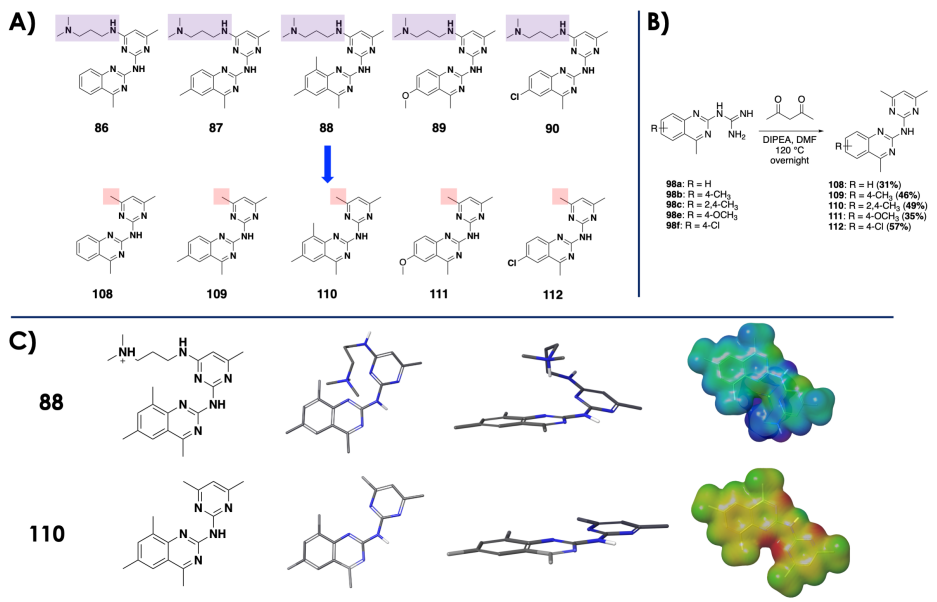


Figure 26. A) Adjusting the amine side chain for the best quinazoline compounds (**86**-**90**) into their corresponding methylated derivatives (**108**-**112**). **B)** Synthesis of the methylated quinazolines (**108**-**112**). **C)** Comparison of the amine side chain quinazolines and their methylated analogues, illustrated by **88** and **110**, highlighting their electrostatic differences. The ESP map is shown with an ISO-value of 0.005 and an energy span of -40-140 kcal/mol. The colour span represents different energy levels going from yellow/green (lowest negative) to purple (highest positive).

As stated in the first section of the results, the compounds (**85**-**94**) shared an analogous electrostatically electron-deficient character (Figure 23B), which

likely considerably impacts the binding interactions to G4. Hence, we next investigated if the dispersion interactions between ligand and G4 followed the same trend if the strong electron-deficient electrostatic component was removed. To probe this, a set of new ligands (**108**-**112**, Figure 26A) with a methyl instead of the aliphatic amine side chain was proposed based on the best quinazoline scaffolds (**86**-**90**, Figure 26A). The synthesis of this novel set of compounds was achieved by condensing appropriate quinazoline guanidine (**98a**-**f**) with acetylacetone to furnish each methyl quinazoline (**108**-**112**) in yields between 31-57% (Figure 26B). The difference in the ESP is shown by comparing **88** and **110**, where the removal of the cationic aliphatic amine drastically changes the ESP map of the compound into a far more electron-rich species (Figure 26C).

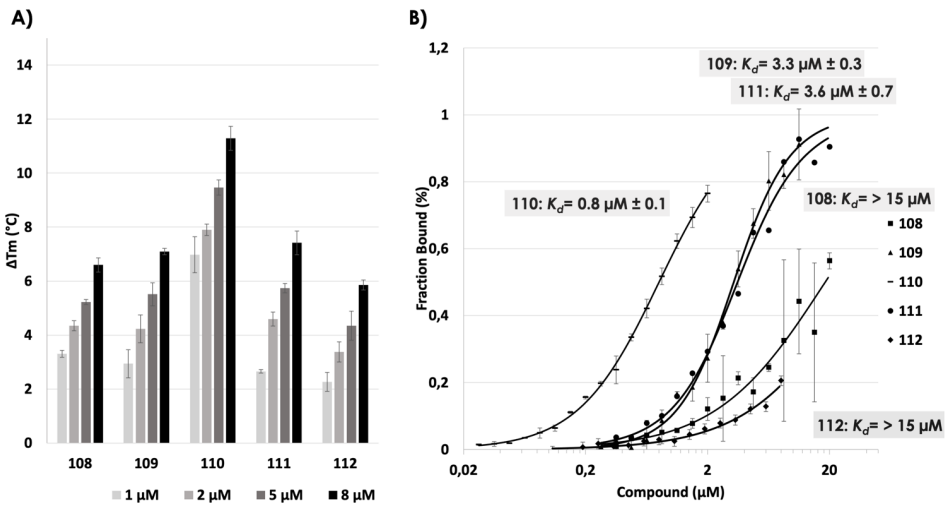


Figure 27. A) The FRET melting assay with the synthesised compounds (**108**-**112**) at 1, 2, 5, and 8 µM of added compound for c-MYC G4 DNA Pu22 (0.2 µM). Error bars correspond to the SD of six independent experiments. **B)** MST binding curves of each compound (**108**-**112**) at varying concentrations (dilution factor: 1:3) with *c-MYC* G4 DNA Pu22 (25 nM). The error bars correspond to the SD of two independent experiments. K_d for each compound is written out in the graph and colour coded accordingly.

The new quinazoline compounds (**108**-**112**) were tested for their abilities to stabilise the *c-MYC* G4 DNA Pu22 (Figure 27A). The trends for the FRET assay were identical to the previous analogues (**86**-**90**), with the most substituted quinazoline (**110**) having the highest impact on thermal stability. This trend was also true for the MST-measured binding affinities for each compound with *c-MYC* G4 DNA Pu22 (Figure 27B), **110** having a measured K_d of 0.8 µM. It is apparent that the dispersion component in the arene-arene interactions is a critical factor in ligand-to-G4 binding that can be used to develop potent G4 binders. However, by comparing the first quinazoline compounds (**86**-**90**) with

the second (**108**-**112**), there is no ambiguity that the aliphatic amines could improve the binding to G4 DNA markedly. We propose that the amines tune the electrostatic potential of the compounds, thus beneficially affecting the electrostatic component in the arene-arene interactions. With this rationale, the arene-arene interactions between ligand and G4 DNA benefit the most from synergism between a strong dispersion component together with an electron-deficient electrostatic component.

To further challenge the *in vitro* assay data, all the compounds (**85**-**94** and **108**-**112**) were tested in a polymerase extension assay (Figure 28). In this assay, a DNA template sequence containing a G4 structure is extended by Taq polymerase. In the absence of ligand, the polymerase is partly at the G4 structure upon advancing along the template sequence but can still advance to synthesise a certain amount of full-length oligo product. However, if a compound can bind and stabilise the G4 structure, the polymerase is halted, resulting in a reduced amount of full-length oligo product and a higher amount of halted oligo product (product up until the G4 structure). The results from this assay were in good agreement with the other assays. For each additional substituent on the quinazoline core (**85**-**90**), the better stabilisation and thus lower amount of full-length product is observed (best viewed at 0.24 µM, showed by the arrows in Figure 28). Compound **88** with the most substituents was the best stabiliser with a 50% inhibition of full-length product at 0.048 µM. Furthermore, either di-methyl (**87**), methoxy (**89**), or chlorine (**90**) (all having two substituents on the quinazoline core) were better than **86**, that only have one methyl (one substituent on the quinazoline core). The quinoxalines (**91**-**94**) were worse binders as expected. Removal of the aliphatic amine to the methyl quinazolines (**108**-**112**) were also not performing as well in this assay. The most substituted methyl quinazoline **110** were, however, good at the higher concentrations. Importantly, none of the compounds displayed any significant effect on the polymerase function when a non-G4 template was used suggesting that the compounds are very selective for G4 DNA.

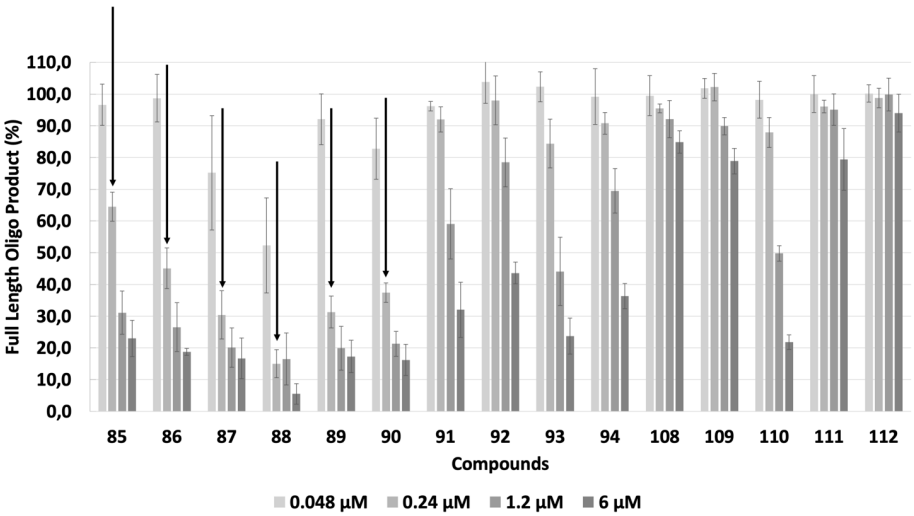


Figure 28. Polymerase extension assay with all the synthesised compounds (**85**-**94** and **108**-**112**) at 0.048, 0.24, 1.2, and 6 µM. The concentration of DNA in the assay is 0.048 µM. Error bars correspond to the SD of at least two independent experiments

8.5.3 Further Evaluations of Compounds 88 and 110

The most promising compound **88** and its methyl analogue (**110**) were tested in a FRET melting assay with different G4 structures. Compound **88** had a clear impact on the thermal stability of cKIT2 (*c-KIT* promoter), Bcl2 (*BCL-2* promoter), 25ceb (human minisatellite), 21G (telomeric), Bom17 (telomeric). Compound **110** did have a similar effect, although weaker. However, **110** did not stabilise Bom17 and only had a slight effect on the Bcl2 G4. Neither of the compounds showed any effect on Tba (thrombin-binding aptamer). Since most structures were affected, this compound class does not seem to discriminate between G4 topologies, unlike the macrocyclic compounds (papers I and II). Both compounds (**88** and **110**) did however show good capacities to selectively stabilise G4 DNA in a FRET competition assay with *c-MYC* G4 DNA Pu22 and Pu24T in the presence of large excess of double-stranded DNA.

8.6 Summary Paper IV

Novel quinazoline (**84**-**90** and **108**-**112**) and quinoxaline (**91**-**94**) compounds were developed with an emphasis on the dispersion and electrostatic components of the arene-arene interactions between ligand and G4. We showed that the dispersion component was paramount for the arene-arene interactions between ligand and G4. Exchanging the quinazoline core to the quinoxaline core (a nitrogen atom shifted one position) resulted in the

compounds that did not bind the G4 DNA well. This underscored that the arene-arene interactions between G4 ligands and G4s are essential for binding, since such small changes in the core aromatic system was detrimental for the binding event. A strong dispersion component in synergism with an electron-deficient ESP of the ligands was optimal for binding to G4. We also showed that these concepts were applicable in more complex assays, such as a polymerase extension assay. This underscored the value of understanding and applying detailed knowledge regarding the binding interactions between ligands and G4 in the discovery and optimisation of novel G4 ligands.

9. Selectivity for Specific G4s (Paper V)

The possibility to selectively bind and stabilise a specific G4 DNA structure is often raised as a key challenge in the field. If successful, this would make it possible to explore the biology of G4 DNA and the potential of G4 DNA as a therapeutic target. Many discussions have been directed towards trying to achieve selectivity based on the topology of the desired G4.[133] A different early approach to specific G4 targeting was to utilise peptide-conjugated compounds that, through the additional interactions between the peptide and G4, gain specificity.[134] Other approaches included the linkage of a G4 ligand with a DNA double-stranded binding ligand, and thereby gain specificity based on the double-stranded sequence adjacent to the desired G4.[135,136] Building on this concept, approaches using PNA-[137] or oligo-conjugated[138] G4 ligands to gain specificity for a desired G4 structure through the complementary interactions with the nucleic acid sequence flanking the selected G4 have shown promising results. Despite this, endeavours in driving the concepts of oligo-conjugated G4 ligands for G4 specificity are so far expanded upon slowly, and this area of research is still in its infancy.

The oligonucleotide research area has over the past 40 years resulted in novel therapeutics for the treatment of several difficult-to-treat diseases. As of 2020, there were 10 FDA-approved oligonucleotide drugs on the market, such as *Spinraza* and *Onpattro*. Oligonucleotides with between 8-40 residues have been reported in the antisense research field, with good selectivity for binding the intended single nucleic acid sequence.[139] Along these lines, an oligonucleotide-G4 ligand conjugate approach was considered a natural starting point to push the apex of targeting specific G4 structures.

9.1 Results and Discussion

9.1.1 Oligonucleotide Design and Compound Conjugation

The suggested approach to gain selectivity for a specific G4 structure was based on the conjugation between a G4 ligand and an oligonucleotide sequence, complementing the nucleotide sequence flanking the desired G4 structure (figure 30A). In this way, the oligonucleotide sequence should guide the G4-ligand to only the target G4 structure. This strategy could work with short (10-

20 bp) oligonucleotide guide sequences because a) this will only match the nucleic acid sequence flanking the G4 structure, b) the conjugated G4 binding compounds are selective for binding G4 DNA, and c) the binding of the compound to a G4 with a non-matching flanking sequence should repel contacts between the compound-conjugated oligo and the flanking sequence (Figure 30B).

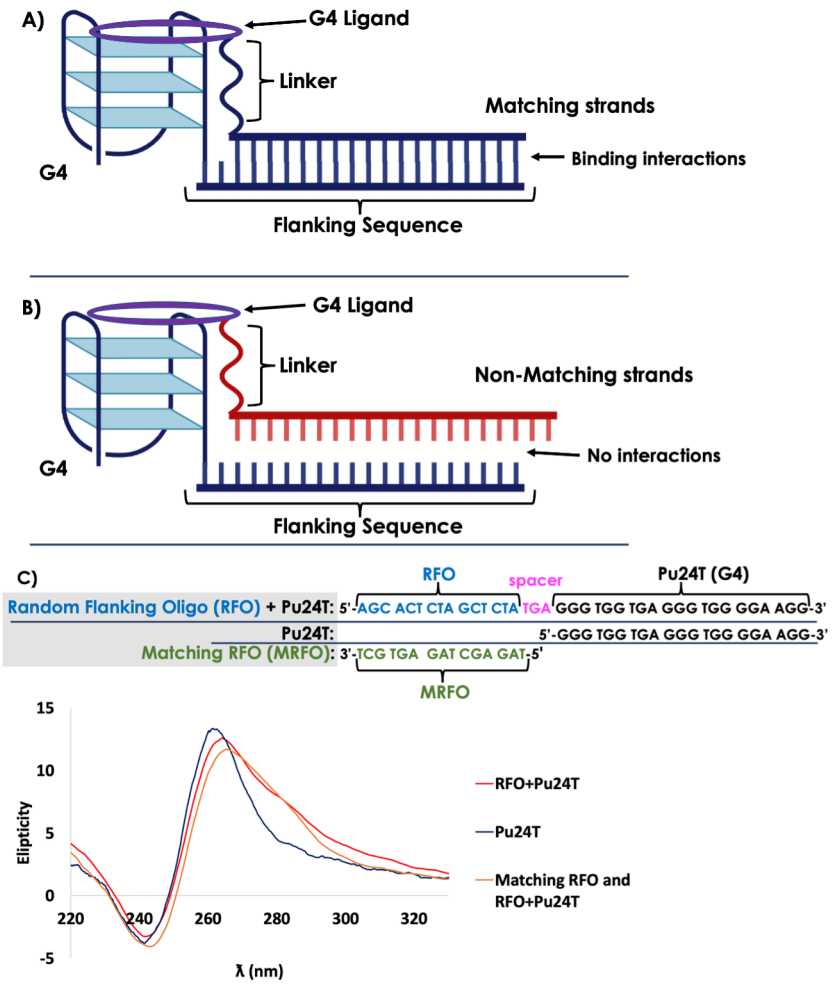


Figure 30. A) Illustration (not representative of details in the binding event) of how a compound + oligonucleotide conjugate (matching sequence) would bind only the G4 structure with a flanking nucleic acid sequence complementing the matching sequence. **B)** Illustration of how a compound + oligonucleotide conjugate (non-matching sequence) should prevent binding to the G4 structure since the flanking nucleic acid sequence will not complement the ligand conjugated oligo sequence. **C)** Written out sequences of *c-MYC* G4 DNA Pu24T, RFO+Pu24T, and MRFO. The CD spectrum for Pu24T, RFO+Pu24T, and MRFO and RFO+Pu24T are shown, colour coded accordingly.

To investigate the concept of this approach, the folding of the G4 structure in the presence of a flanking oligonucleotide sequence *in vitro* was first confirmed. Thus, Pu24T *c-MYC* G4 DNA connected to a random flanking oligonucleotide sequence, alias random flanking oligonucleotide (RFO), was characterised by CD (circular dichroism) spectroscopy (Figure 30C). With CD, we measure the ability of the biomolecule to absorb circularly polarised light, and from this absorption, characterisation of the CD curve of the sample compared to a reference CD signature can be done. CD analysis showed that the Pu24T G4 structure remained folded upon the addition of the RFO (Figure 30C). Furthermore, the addition of the matching oligonucleotide sequence (Matching RFO or MRFO) to the RFO+Pu24T did not affect the G4 structure (Figure 30C).

In addition, the ^{1}H NMR (850 MHz) spectra of MRFO and RFO+Pu24T were recorded. This revealed that the G4 structure remained intact, as shown by the presence of the imino protons. For the combined (MRFO and RFO+Pu24T) recording, the hydrogen-bonded double-stranded protons between the two matched random sequences (MRFO and RFO) could also be observed, showing that the complementary strands can form dsDNA simultaneously as the G4 structure remained intact.

Next, the goal was to conjugate a compound to the MRFO that could bind and stabilise the G4 simultaneously as the double-stranded formation between MRFO and RFO to guide the conjugate to the target sequence. For this, quinazoline compound **113**, a known potent G4 binder of the type disclosed in papers III and IV, with an azide at the end of the amine side chain (Figure 31), was chosen. This ligand (**113**) was chosen before much of the work from papers III and IV, and a different ligand might have been more optimal. The click (Huisgen) reaction for the conjugation reaction was chosen, based on its wide applications in bioconjugation approaches.[140] Hence, three different MRFO to **113** conjugates were proposed (Figure 31), one with simply the MRFO ($MRFO_1$), one with an added nucleotide complementary to the spacer in RFO+Pu24T ($MRFO_2$), and one with two complementary spacer nucleotides ($MRFO_3$). Compound **113** was consequently conjugated with each MRFO to generate three different MRFO+**113** conjugates, namely, $MRFO_1$+**113**, $MRFO_2$+**113**, and $MRFO_3$+**113** (Figure 31). The conjugation was performed on the 5' side of MRFO since that should be adjacent to the G4 structure.

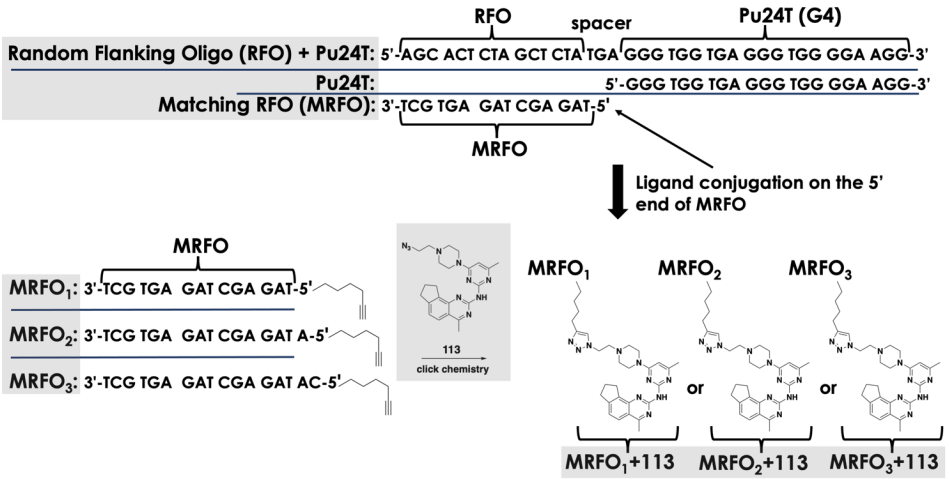
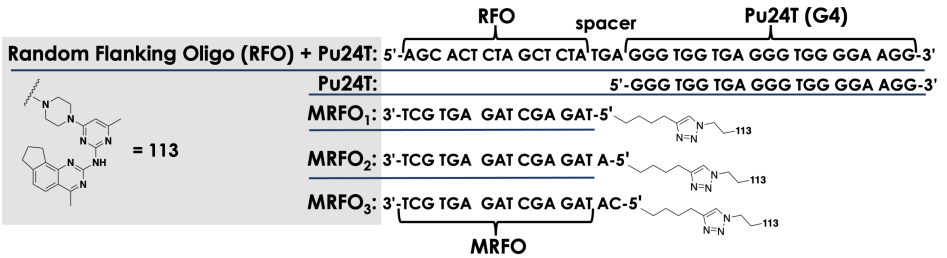


Figure 31. Written out sequences of *c-MYC* G4 DNA Pu24T, RFO+Pu24T, and MRFO. The conjugation of compound **113** to the 5' side of MRFO via the click reaction between either $MRFO_1$, $MRFO_2$, or $MRFO_3$. The MRFO+**113** conjugates ($MRFO_1$+**113**, $MRFO_2$+**113**, and $MRFO_3$+**113**) are finally shown.

9.1.2 Binding to of Stabilisation of RFO+Pu24T

With the compound conjugated MRFOs in hand, the binding was recorded using MST with 5'Cy-labelled RFO+Pu24T. Thus, each $MRFO_{1\text{-}3}$+**113**, compound **113** alone, and a non-matching MRFO (non-MRFO), were tested for their abilities to bind RFO+Pu24T (Figure 32). Compound **113** alone had a K_d value above 200 nM, as expected for this compound (paper III). The MRFO alone was a potent binder to RFO+Pu24T with a K_d value of 60 nM. In contrast, when the non-MRFO was used, there was no observable binding to RFO+Pu24T since no matched double-stranded can assemble. When the G4 ligand conjugated MRFOs ($MRFO_{1\text{-}3}$+**113**) were analysed, the binding was notably improved compared to **133** and MRFO alone, suggesting that this concept is a viable strategy for specific G4 targeting. The best was $MRFO_1$+**113**, which has the shortest length between the compound and MRFO and was the best binder with a K_d value of 12 nM. Extending the length by the introduction of either one ($MRFO_2$+**113**) or two ($MRFO_3$+**113**) nucleotides showed a modest increase in binding, however, both still being very potent binders. When the RFO was removed from the G4, resulting in G4 Pu24T structure alone, no observable binding of the compound conjugated oligos could be observed either. Hence, this strategy should be viable for targeting specific G4 structures.

The $MRFO_1$+**113** was challenged further with MST competition experiments. In this setup, the binding affinity of $MRFO_1$+**113** to RFO+Pu24T was measured, in the presence of Pu24T. In both scenarios, the $MRFO_1$+**113** retained its potent binding to RFO+Pu24T. This underscored the ability of this concept to discriminate binding to other G4s.



Oligo and Ligand	Binding constant K_d (nM)
RFO+Pu24T and **113**	> 200
RFO+Pu24T and MRFO	60 ± 5
RFO+Pu24T and non-MRFO	No binding
RFO+Pu24T and $MRFO_1$+**113**	12 ± 3
RFO+Pu24T and $MRFO_2$+**113**	18 ± 4
RFO+Pu24T and $MRFO_2$+**113**	22 ± 6

Figure 32. Illustration of the RFO+Pu24T, Pu24T, $MRFO_1$+**113**, $MRFO_2$+**113**, and $MRFO_3$+**113**, highlighting their differences and similarities. The K_d values for the different species (**113**, MRFO, non-MRFO, and $MRFO_{1\text{-}3}$+**113**) for 5'Cy5-RFO+Pu24T.

Variable temperature ^{1}H NMR (850 MHz) studies were conducted to investigate the difference in thermal stability of the RFO+Pu24T in the presence of either simply MRFO or $MRFO_1$+**113**. In this experiment, the temperature was increased in increments of 5 °C, and the proton signals for the double-stranded formation and G4 were observed separately. This is possible since the double-stranded signals appear at 12-14 ppm, while the protons from the G4 appear between 10-12 ppm. However, at 50 °C, the double-stranded signals were almost completely gone, while the signals for the G4 remained, showing that the G4 structure is more stable than the double-stranded DNA. When ^{1}H NMR of the RFO+Pu24T and $MRFO_1$+**113** were recorded, the double-stranded signals remained intact until 60 °C. Hence, the double-stranded DNA and G4 structure had a higher degree of thermal stability in the presence of the G4 ligand conjugated to the recognition sequence ($MRFO_1$+**113**), compared to only the MRFO. These results were validated in a CD melting experiment.

9.1.3 Specificity and Synergism of the Conjugated Oligonucleotide Approach

Next, the strategy was challenged in a polymerase extension assay. In this setup, the DNA template was comprised of either Pu24T with the non-matching RFO (non-RFO+Pu24T) or the RFO (RFO+Pu24T). The ability of the Taq polymerase to extend beyond the G4 structure in both templates was measured in the presence of either **113** alone or $MRFO_1$+**113** (Figure 33A and 33B). Compound **113** did, as expected, reduce the full-length DNA product for both DNA template strands since it should not differentiate between different G4s.

Strikingly, when the $MRFO_1$+**113** and non-RFO+Pu24T were combined, the DNA synthesis was virtually unaffected until above 1 µM concentrations (Figure 33A and 33B). In contrast, for the complementary RFO+Pu24T and $MRFO_1$+**113**, the DNA end-product was significantly reduced, already at 0.2 µM (Figure 33A and 33B). This data shows that this approach offers a powerful way towards specificity for a desired G4 with the matching oligo sequence and improves the G4 stabilisation compared to the G4 ligand (**113**) alone.

As another control experiment, the Taq polymerase extension assay was carried out with the DNA template containing RFO+Pu24T, but with either just the MRFO sequence, compound **113**, MRFO and **113** together (but not covalently linked), or with $MRFO_1$+**113** (Figure 34). This setup was designed to provide information on how well the conjugation strategy synergised in comparison to simply having a G4 stabilising compound and a separate oligonucleotide that can hybridise with the flanking sequence. The MRFO alone did not show a considerable effect on the G4 stability, as expected, based on the observed low thermal stability of the double-stranded in the variable temperature NMR measurements. Compound **113** alone could reduce the amount of DNA end-product, analogous to its ability to stabilise the G4 structure in the FRET assay (paper III). When compound **113** and the MRFO were not conjugated but added as separate units in combination, a reduction in the DNA end-product was observed, which was slightly more prominent compared to either **113** or MRFO alone. However, when $MRFO_1$+**113** was used, the results showed an even higher reduction in DNA end-product. This showed that the conjugated compound to oligo approach results in an even better G4 stabiliser compared to the two components alone without conjugation. Thus, this approach does provide a way to gain G4 specificity, but also a synergistic effect which results in improved G4 binding and stabilisation of the oligo+compound, not achieved by the separate components.

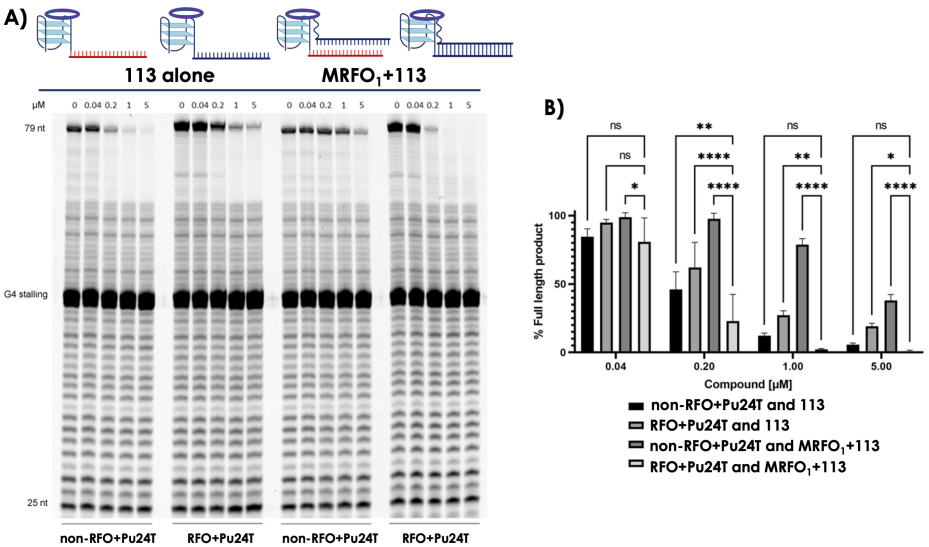


Figure 33. A) Taq polymerase extension assays in the presence of control compound **113** (lanes 2-5 and 7-10) without oligonucleotide addition or with $MRFO_1$+**113** (lanes 12-15 and 17-20) on two different templates that includes a 15 nts DNA sequence that was either the RFO+Pu24T (lanes 6-10 and 16-20, DNA indicated in blue) or non-RFO+Pu24T (lanes 1-5 and 11-15, DNA indicated in red) to the oligo portion (oligonucleotide addition indicated in orange) of $MRFO_1$+**113**. **B)** Full-length (79 nts DNA run-off product) quantification of the Taq polymerase extension assays. The full-length product is expressed in % of the full-length band intensity that was obtained in the control reaction that did not contain compound. Mean and standard deviation was calculated (n=4).

Finally, the two other compound-conjugated oligonucleotides ($MRFO_{2\text{-}3}$+**113**) were tested for their abilities to stabilise the G4 structure in the Taq polymerase extension assay (for more information, see paper V). The different MRFOs+**113** had similar binding affinities to the RFO+Pu24T on the MST measurements (Figure 30). Despite this, $MRFO_2$+**113**, having one extra nucleotide in the spacer (Figure 30), showed to be superior (see paper V for more information) in stabilising the G4 structure in the Taq polymerase extension assay, indicating that the best ability for the compound and MRFO to bind simultaneously is when the MRFO began two nucleotides away from the start of the G4 structure (with a strong effect already at 0.04 μM). This showed that the distance between the oligo and ligand was very important, and the linker strategy will thus be crucial in future developments.

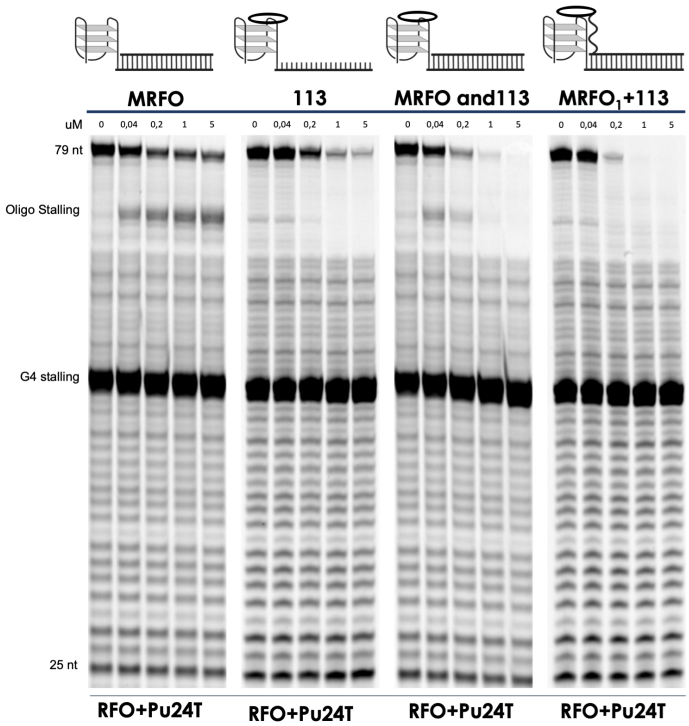


Figure 34. Taq-polymerase extension assays in the presence of MRFO alone (lanes 1-5), **113** alone (lanes 6-10), MRFO and **113** (lanes 11-15), or $MRFO_1$+**113** (lanes 16-20) on a DNA template that includes a 15 nts DNA sequence complementary to the oligo (in orange).

9.2 Summary Paper V

The concept of a compound conjugated to an oligonucleotide sequence resulted in a very potent G4 binder, exemplified by the interactions between RFO+Pu24T and $MRFO_2$+**113**. Furthermore, the data show that no binding between the MRFO+**113** and G4 structure occurs when the RFO is not matching the MRFO conjugated to the compound. In summary, both G4 specificity and improved G4 binding were achieved using this strategy. We surmise that this strategy could be applied in a highly modular way to access G4 specificity through the conjugation between short oligonucleotides and known G4 ligands. The ideality of this approach can be emphasised in the vast access to synthetic strategies for short and modified oligos, as well as conjugation strategies. In addition, the rapid expansion of the oligonucleotide drug delivery field could serve as a bridgehead for specified deliveries and modulation of unwanted PK properties in G4 ligand-conjugated oligos. The hope is that this work can inspire other researchers to apply this strategy to expand the knowledge of G4s in biology and to strive towards possible therapeutic applications.

10. Concluding Remarks

This thesis describes the discovery and synthetic development of novel G4 ligands and investigations of the essential components in the molecular interactions between ligands and the G4 surface.

Major synthetic efforts resulted in the development of several novel macrocyclic G4 ligands. The work shows that macrocyclic compounds can serve as a basis for novel G4 ligand design. Careful consideration and optimisation of the macrocyclic ring size were critical for binding to the G4 surface. The macrocycles presented discriminate between G4 DNA and double-stranded DNA and preferred parallel and hybrid G4 topologies, and this strategy can thus improve G4 selectivity. Computational tools were successfully applied in optimising the macrocyclic design. This optimisation resulted in macrocycles with a conformational preference similar

ESP map calculations of different G4 ligands revealed that the compounds that efficiently bound and stabilised G4s displayed an electrostatically electron-deficient character. This electron-deficient character could be attributed to cationic charged side chains such as protonated aliphatic amines. However, protonated cations in the ligands did not seem beneficial for the binding interactions when located close to the G4 surface upon binding. Hence, the inclusion of cationic groups in G4 ligands should be placed outside of the G4 surface and was suggested to improve the binding to G4s indirectly through the electrostatic component of the arene-arene interactions.

A small set of heterocyclic compounds were designed, aided by computational methods, to disclose the role of the dispersion energetic component in the arene-arene interactions between G4 ligands and G4 DNA structures. This showed that the dispersion component is paramount for binding. The relocating of a nitrogen atom from the quinazoline aromatic core to afford the quinoxaline core resulted in no binding. This underscored the importance of the arene-arene interactions between G4 ligands and G4 DNA. Furthermore, a strong dispersion component in synergism with an electron-deficient ESP was imperative for optimal binding to G4. In addition, this revealed small organic compounds without permanent charges demonstrating a balance between good PK

properties and potent G4 binding. Thus, the discoveries revealed in this thesis, especially regarding ligand-to-G4 interactions, is ample to assist in novel G4 ligand developments and can be valuable to catalyse the advance of G4 ligands in therapeutic settings.

Finally, an oligonucleotide conjugated to a G4 ligand approach was demonstrated as a tool to generate a very potent G4 binder. The compound conjugated oligo did not bind the G4 structure when the flanking oligonucleotide sequence did not complement the conjugated oligonucleotide. Hence, the data shows that considerable G4 specificity was achieved. Thus, both increased G4 binding and G4 specificity were accomplished by this strategy. This approach must now be evaluated to study if specific targeting of G4s will affect the transcription of the target gene and not others. Furthermore, cell permeability and the effect on cell viability should be considered. Perhaps, the methods currently being developed for the delivery of oligonucleotide therapies could be applied to the delivery of oligonucleotide-conjugated G4 ligands. Thereof, a Phosphorothioate (PS) backbone is known to improve oligonucleotide stability. Locked nucleic sugars (LNAs) are a popular strategy for generating more specificity in binding to the correct complementary sequence, thereby reducing off-target effects.

To conclude, most of the aims of this thesis have been reached. Although, a lot of future work remains to fully understand the interactions between G4 DNA and G4 ligands, as well as to explore approaches to achieve specificity for individual G4 structures.

11. Acknowledgements

I have learnt a lot during my PhD at Umeå University, both personal and professional. I want to thank everyone who contributed to my time during these 4 and a half years. But I especially wish to thank the following people.

First, I want to thank my supervisor **Erik Chorell** who took me on as his first PhD student, despite me not even having obtained my master's degree. I am thankful for all of the understanding you have showed, and how much freedom and responsibility I have been given, which I think has been very important for my learning outcome. I think that we have had a lot of great discussions that has resulted in that both of us has learned a great deal.

My co-supervisor, **Anna Linusson Jonsson**, thank you for being such an involved second supervisor and excellent role-model. I have learned so much from our discussions, and your support has been very meaningful to me throughout the PhD.

Everyone in "the best lab", **Maxime Donzel**, **Jessica Giacoboni**, **Manesh Nautiyal**, **Justyna Gołębiewska**, and **Alva Abrahamsson**. Thank you for creating such a great work environment, the problem-solving sessions, and all the fun stuff outside of work. It has been really important for me.

Previous members in the Chorell group. **Rabindra Nath Das**, **Naresh Bhuma**, **Karam Chand**, **Prasad Bagineni**, and **Rajendra Kumar**, thank you for all the help and good advice. It has been helpful for me in my learning process during the PhD.

The **Fredrik Almqvist research group**. Thank you for all the nice group meeting discussions and continuous feedback on yours and my research projects.

The **Anna Linusson Jonsson research group**. Thank you for letting me join your group meetings and journal clubs. The medicinal chemistry discussions have been valuable, and I have gotten a lot of great feedback.

The rest of the Organic Chemistry corridor at KBC, thank you for always being friendly and helpful. I have had a lot of great advice and discussions around synthetic chemistry.

The “PhD (and others) gang”. **Ioana**, **Mareike**, **Norman**, **Sofie**, **Fredrik**, **Maria**, **James**, **Mason**, and others. Thank you for all the great beer evenings and dinners. It has been great to have a solid group of people to do things with. A special thanks to **Tilde** who has been a great friend, both at and outside of work.

I would like to thank my collaborators for contributing to some great publications and good teamwork. Especially **Nils Pemberton** as AstraZeneca and the research group of **Sjoerd Wanrooij**.

I would also like to thank my friends and family in Gothenburg for helping me get through this long period, both by visiting me in Umeå and spending time with me from a distance.

I also would like to thank the department of chemistry at Umeå University and Vetenskapsrådet for the financial support, facilities, and the use of available research instruments.

Finally, but for sure not least, I would like to thank **Olivia Spain** for helping me through the PhD. I think it would have been very difficult to manage without your help, both at work and outside, and I am very grateful for that support.

12. References

(1) Bang, I. Untersuchungen über die Guanylsäre. *Biochem. Z.* **1910**, *26*, 293–311.

(2) Gellert, M.; Lipsett, M. N.; Davies, D. R. Helix Formation by Guanylic Acid. *Proceedings of the National Academy of Sciences* **1962**, *48* (12), 2013–2018. https://doi.org/10.1073/pnas.48.12.2013.

(3) Zimmerman, S. B.; Cohen, G. H.; Davies, D. R. X-Ray Fiber Diffraction and Model-Building Study of Polyguanylic Acid and Polyinosinic Acid. *Journal of Molecular Biology* **1975**, *92* (2), 181–192. https://doi.org/10.1016/0022-2836(75)90222-3.

(4) Spiegel, J.; Adhikari, S.; Balasubramanian, S. The Structure and Function of DNA G-Quadruplexes. *Trends in Chemistry* **2020**, *2* (2), 123–136. https://doi.org/10.1016/j.trechm.2019.07.002.

(5) Dolinnaya, N. G.; Ogloblina, A. M.; Yakubovskaya, M. G. Structure, Properties, and Biological Relevance of the DNA and RNA G-Quadruplexes: Overview 50 Years after Their Discovery. *Biochemistry (Mosc)* **2016**, *81* (13), 1602–1649. https://doi.org/10.1134/S0006297916130034.

(6) Sen, D.; Gilbert, W. A Sodium-Potassium Switch in the Formation of Four-Stranded G4-DNA. *Nature* **1990**, *344* (6265), 410–414. https://doi.org/10.1038/344410a0.

(7) Lane, A. N.; Chaires, J. B.; Gray, R. D.; Trent, J. O. Stability and Kinetics of G-Quadruplex Structures. *Nucleic Acids Res* **2008**, *36* (17), 5482–5515. https://doi.org/10.1093/nar/gkn517.

(8) Dingley, A. J.; Peterson, R. D.; Grzesiek, S.; Feigon, J. Characterization of the Cation and Temperature Dependence of DNA Quadruplex Hydrogen Bond Properties Using High-Resolution NMR. *J. Am. Chem. Soc.* **2005**, *127* (41), 14466–14472. https://doi.org/10.1021/ja0540369.

(9) Burge, S.; Parkinson, G. N.; Hazel, P.; Todd, A. K.; Neidle, S. Quadruplex DNA: Sequence, Topology and Structure. *Nucleic Acids Research* **2006**, *34* (19), 5402–5415. https://doi.org/10.1093/nar/gkl655.

(10) Zhou, J.; Bourdoncle, A.; Rosu, F.; Gabelica, V.; Mergny, J.-L. Tri-G-Quadruplex: Controlled Assembly of a G-Quadruplex Structure from Three G-Rich Strands. *Angewandte Chemie International Edition* **2012**, *51* (44), 11002–11005. https://doi.org/10.1002/anie.201205390.

(11) *Topologies of G-quadruplex: Biological functions and regulation by ligands | Elsevier Enhanced Reader*. https://doi.org/10.1016/j.bbrc.2019.12.103.

(12) Varshney, D.; Spiegel, J.; Zyner, K.; Tannahill, D.; Balasubramanian, S. The Regulation and Functions of DNA and RNA G-Quadruplexes. *Nat Rev Mol Cell Biol* **2020**, *21* (8), 459–474. https://doi.org/10.1038/s41580-020-0236-x.

(13) Sun, D.; Hurley, L. H. The Importance of Negative Superhelicity in Inducing the Formation of G-Quadruplex and i-Motif Structures in the c-Myc Promoter: Implications for Drug Targeting and Control of Gene Expression. *J. Med. Chem.* **2009**, *52* (9), 2863–2874. https://doi.org/10.1021/jm900055s.

(14) Capra, J. A.; Paeschke, K.; Singh, M.; Zakian, V. A. G-Quadruplex DNA Sequences Are Evolutionarily Conserved and Associated with Distinct Genomic Features in Saccharomyces Cerevisiae. *PLoS Comput Biol* **2010**, *6* (7), e1000861. https://doi.org/10.1371/journal.pcbi.1000861.

(15) Nakken, S.; Rognes, T.; Hovig, E. The Disruptive Positions in Human G-Quadruplex Motifs Are Less Polymorphic and More Conserved than Their Neutral Counterparts. *Nucleic Acids Res* **2009**, *37* (17), 5749–5756. https://doi.org/10.1093/nar/gkp590.

(16) Chambers, V. S.; Marsico, G.; Boutell, J. M.; Di Antonio, M.; Smith, G. P.; Balasubramanian, S. High-Throughput Sequencing of DNA G-Quadruplex Structures in the Human Genome. *Nat Biotechnol* **2015**, *33* (8), 877–881. https://doi.org/10.1038/nbt.3295.

(17) Marsico, G.; Chambers, V. S.; Sahakyan, A. B.; McCauley, P.; Boutell, J. M.; Antonio, M. D.; Balasubramanian, S. Whole Genome Experimental Maps of DNA G-Quadruplexes in Multiple Species. *Nucleic Acids Res* **2019**, *47* (8), 3862–3874. https://doi.org/10.1093/nar/gkz179.

(18) Bochman, M. L.; Paeschke, K.; Zakian, V. A. DNA Secondary Structures: Stability and Function of G-Quadruplex Structures. *Nat Rev Genet* **2012**, *13* (11), 770–780. https://doi.org/10.1038/nrg3296.

(19) Brázda, V.; Hároníková, L.; Liao, J. C. C.; Fojta, M. DNA and RNA Quadruplex-Binding Proteins. *Int J Mol Sci* **2014**, *15* (10), 17493–17517. https://doi.org/10.3390/ijms151017493.

(20) Eddy, J.; Maizels, N. Gene Function Correlates with Potential for G4 DNA Formation in the Human Genome. *Nucleic Acids Res* **2006**, *34* (14), 3887–3896. https://doi.org/10.1093/nar/gkl529.

(21) Maizels, N.; Gray, L. T. The G4 Genome. *PLOS Genetics* **2013**, *9* (4), e1003468. https://doi.org/10.1371/journal.pgen.1003468.

(22) Rhodes, D.; Lipps, H. J. G-Quadruplexes and Their Regulatory Roles in Biology. *Nucleic Acids Res* **2015**, *43* (18), 8627–8637. https://doi.org/10.1093/nar/gkv862.

(23) Huppert, J. L.; Balasubramanian, S. G-Quadruplexes in Promoters throughout the Human Genome. *Nucleic Acids Res* **2007**, *35* (2), 406–413. https://doi.org/10.1093/nar/gkl1057.

(24) Robinson, J.; Raguseo, F.; Nuccio, S. P.; Liano, D.; Di Antonio, M. DNA G-Quadruplex Structures: More than Simple Roadblocks to Transcription? *Nucleic Acids Research* **2021**, *49* (15), 8419–8431. https://doi.org/10.1093/nar/gkab609.

(25) Tian, T.; Chen, Y.-Q.; Wang, S.-R.; Zhou, X. G-Quadruplex: A Regulator of Gene Expression and Its Chemical Targeting. *Chem* **2018**, *4* (6), 1314–1344. https://doi.org/10.1016/j.chempr.2018.02.014.

(26) Bharti, S. K.; Sommers, J. A.; Zhou, J.; Kaplan, D. L.; Spelbrink, J. N.; Mergny, J.-L.; Brosh, R. M. DNA Sequences Proximal to Human Mitochondrial DNA Deletion Breakpoints Prevalent in Human Disease Form G-Quadruplexes, a Class of DNA Structures Inefficiently Unwound by the Mitochondrial Replicative Twinkle Helicase. *J Biol Chem* **2014**, *289* (43), 29975–29993. https://doi.org/10.1074/jbc.M114.567073.

(27) Wanrooij, P. H.; Uhler, J. P.; Shi, Y.; Westerlund, F.; Falkenberg, M.; Gustafsson, C. M. A Hybrid G-Quadruplex Structure Formed between RNA and DNA Explains the Extraordinary Stability of the Mitochondrial R-Loop. *Nucleic Acids Res* **2012**, *40* (20), 10334–10344. https://doi.org/10.1093/nar/gks802.

(28) Wolfe, A. L.; Singh, K.; Zhong, Y.; Drewe, P.; Rajasekhar, V. K.; Sanghvi, V. R.; Mavrakis, K. J.; Jiang, M.; Roderick, J. E.; Van der Meulen, J.; Schatz, J. H.; Rodrigo, C. M.; Zhao, C.; Rondou, P.; de Stanchina, E.; Teruya-Feldstein, J.; Kelliher, M. A.; Speleman, F.; Porco, J. A.; Pelletier, J.; Rätsch, G.; Wendel, H.-G. RNA G-Quadruplexes Cause EIF4A-Dependent Oncogene Translation in Cancer. *Nature* **2014**, *513* (7516), 65–70. https://doi.org/10.1038/nature13485.

(29) Lyu, J.; Shao, R.; Kwong Yung, P. Y.; Elsässer, S. J. Genome-Wide Mapping of G-Quadruplex Structures with CUT&Tag. *Nucleic Acids Res* **2021**, *50* (3), e13. https://doi.org/10.1093/nar/gkab1073.

(30) *Telomeric G-quadruplexes are a substrate and site of localization for human telomerase | Nature Communications*. https://www.nature.com/articles/ncomms8643 (accessed 2022-10-07).

(31) Paeschke, K.; Simonsson, T.; Postberg, J.; Rhodes, D.; Lipps, H. J. Telomere End-Binding Proteins Control the Formation of G-Quadruplex DNA Structures in Vivo. *Nat Struct Mol Biol* **2005**, *12* (10), 847–854. https://doi.org/10.1038/nsmb982.

(32) Besnard, E.; Babled, A.; Lapasset, L.; Milhavet, O.; Parrinello, H.; Dantec, C.; Marin, J.-M.; Lemaitre, J.-M. Unraveling Cell Type–Specific and Reprogrammable Human Replication Origin Signatures Associated with G-Quadruplex Consensus Motifs. *Nat Struct Mol Biol* **2012**, *19* (8), 837–844. https://doi.org/10.1038/nsmb.2339.

(33) Prorok, P.; Artufel, M.; Aze, A.; Coulombe, P.; Peiffer, I.; Lacroix, L.; Guédin, A.; Mergny, J.-L.; Damaschke, J.; Schepers, A.; Cayrou, C.; Teulade-Fichou, M.-P.; Ballester, B.; Méchali, M. Involvement of G-Quadruplex Regions in Mammalian Replication Origin Activity. *Nat Commun* **2019**, *10* (1), 3274. https://doi.org/10.1038/s41467-019-11104-0.

(34) Mirkin, S. M. Expandable DNA Repeats and Human Disease. *Nature* **2007**, *447* (7147), 932–940. https://doi.org/10.1038/nature05977.
(35) Wang, E.; Thombre, R.; Shah, Y.; Latanich, R.; Wang, J. G-Quadruplexes as Pathogenic Drivers in Neurodegenerative Disorders. *Nucleic Acids Research* **2021**, *49* (9), 4816–4830. https://doi.org/10.1093/nar/gkab164.
(36) Ruggiero, E.; Richter, S. N. G-Quadruplexes and G-Quadruplex Ligands: Targets and Tools in Antiviral Therapy. *Nucleic Acids Res* **2018**, *46* (7), 3270–3283. https://doi.org/10.1093/nar/gky187.
(37) G4-Associated Human Diseases. *EMBO reports* **2015**, *16* (8), 910–922. https://doi.org/10.15252/embr.201540607.
(38) Kosiol, N.; Juranek, S.; Brossart, P.; Heine, A.; Paeschke, K. G-Quadruplexes: A Promising Target for Cancer Therapy. *Mol Cancer* **2021**, *20* (1), 40. https://doi.org/10.1186/s12943-021-01328-4.
(39) Siddiqui-Jain, A.; Grand, C. L.; Bearss, D. J.; Hurley, L. H. Direct Evidence for a G-Quadruplex in a Promoter Region and Its Targeting with a Small Molecule to Repress c-MYC Transcription. *Proc Natl Acad Sci U S A* **2002**, *99* (18), 11593–11598. https://doi.org/10.1073/pnas.182256799.
(40) Dang, C. V. MYC on the Path to Cancer. *Cell* **2012**, *149* (1), 22–35. https://doi.org/10.1016/j.cell.2012.03.003.
(41) Gregory, M. A.; Hann, S. R. C-Myc Proteolysis by the Ubiquitin-Proteasome Pathway: Stabilization of c-Myc in Burkitt's Lymphoma Cells. *Mol Cell Biol* **2000**, *20* (7), 2423–2435. https://doi.org/10.1128/MCB.20.7.2423-2435.2000.
(42) Deng, N.; Wickstrom, L.; Cieplak, P.; Lin, C.; Yang, D. Resolving the Ligand-Binding Specificity in c-MYC G-Quadruplex DNA: Absolute Binding Free Energy Calculations and SPR Experiment. *J Phys Chem B* **2017**, *121* (46), 10484–10497. https://doi.org/10.1021/acs.jpcb.7b09406.
(43) Ambrus, A.; Chen, D.; Dai, J.; Jones, R. A.; Yang, D. Solution Structure of the Biologically Relevant G-Quadruplex Element in the Human c-MYC Promoter. Implications for G-Quadruplex Stabilization. *Biochemistry* **2005**, *44* (6), 2048–2058. https://doi.org/10.1021/bi048242p.
(44) Phan, A. T.; Kuryavyi, V.; Gaw, H. Y.; Patel, D. J. Small-Molecule Interaction with a Five-Guanine-Tract G-Quadruplex Structure from the Human MYC Promoter. *Nat Chem Biol* **2005**, *1* (3), 167–173. https://doi.org/10.1038/nchembio723.
(45) Ohnmacht, S. A.; Neidle, S. Small-Molecule Quadruplex-Targeted Drug Discovery. *Bioorganic & Medicinal Chemistry Letters* **2014**, *24* (12), 2602–2612. https://doi.org/10.1016/j.bmcl.2014.04.029.
(46) Chaudhuri, R.; Bhattacharya, S.; Dash, J.; Bhattacharya, S. Recent Update on Targeting C-MYC G-Quadruplexes by Small Molecules for Anticancer Therapeutics. *J. Med. Chem.* **2021**, *64* (1), 42–70. https://doi.org/10.1021/acs.jmedchem.0c01145.
(47) Balasubramanian, S.; Hurley, L. H.; Neidle, S. Targeting G-Quadruplexes in Gene Promoters: A Novel Anticancer Strategy? *Nat Rev Drug Discov* **2011**, *10* (4), 261–275. https://doi.org/10.1038/nrd3428.
(48) Nannini, M.; Biasco, G.; Astolfi, A.; Pantaleo, M. A. An Overview on Molecular Biology of KIT/PDGFRA Wild Type (WT) Gastrointestinal Stromal Tumours (GIST). *J Med Genet* **2013**, *50* (10), 653–661. https://doi.org/10.1136/jmedgenet-2013-101695.
(49) Sommer, G.; Agosti, V.; Ehlers, I.; Rossi, F.; Corbacioglu, S.; Farkas, J.; Moore, M.; Manova, K.; Antonescu, C. R.; Besmer, P. Gastrointestinal Stromal Tumors in a Mouse Model by Targeted Mutation of the Kit Receptor Tyrosine Kinase. *Proc Natl Acad Sci U S A* **2003**, *100* (11), 6706–6711. https://doi.org/10.1073/pnas.1037763100.
(50) Fletcher, J. A.; Rubin, B. P. KIT Mutations in GIST. *Curr Opin Genet Dev* **2007**, *17* (1), 3–7. https://doi.org/10.1016/j.gde.2006.12.010.

(51) Rankin, S.; Reszka, A. P.; Huppert, J.; Zloh, M.; Parkinson, G. N.; Todd, A. K.; Ladame, S.; Balasubramanian, S.; Neidle, S. Putative DNA Quadruplex Formation within the Human C-Kit Oncogene. *J Am Chem Soc* **2005**, *127* (30), 10584–10589. https://doi.org/10.1021/ja050823u.

(52) Fernando, H.; Reszka, A. P.; Huppert, J.; Ladame, S.; Rankin, S.; Venkitaraman, A. R.; Neidle, S.; Balasubramanian, S. A Conserved Quadruplex Motif Located in a Transcription Activation Site of the Human C-Kit Oncogene. *Biochemistry* **2006**, *45* (25), 7854–7860. https://doi.org/10.1021/bi0601510.

(53) Wei, D.; Husby, J.; Neidle, S. Flexibility and Structural Conservation in a C-KIT G-Quadruplex. *Nucleic Acids Research* **2015**, *43* (1), 629–644. https://doi.org/10.1093/nar/gku1282.

(54) Waters, A. M.; Der, C. J. KRAS: The Critical Driver and Therapeutic Target for Pancreatic Cancer. *Cold Spring Harb Perspect Med* **2018**, *8* (9), a031435. https://doi.org/10.1101/cshperspect.a031435.

(55) Cogoi, S.; Xodo, L. E. G-Quadruplex Formation within the Promoter of the KRAS Proto-Oncogene and Its Effect on Transcription. *Nucleic Acids Res* **2006**, *34* (9), 2536–2549. https://doi.org/10.1093/nar/gkl286.

(56) Kumari, S.; Bugaut, A.; Huppert, J. L.; Balasubramanian, S. An RNA G-Quadruplex in the 5′ UTR of the NRAS Proto-Oncogene Modulates Translation. *Nat Chem Biol* **2007**, *3* (4), 218–221. https://doi.org/10.1038/nchembio864.

(57) Cogoi, S.; Shchekotikhin, A. E.; Xodo, L. E. HRAS Is Silenced by Two Neighboring G-Quadruplexes and Activated by MAZ, a Zinc-Finger Transcription Factor with DNA Unfolding Property. *Nucleic Acids Res* **2014**, *42* (13), 8379–8388. https://doi.org/10.1093/nar/gku574.

(58) Dai, J.; Dexheimer, T. S.; Chen, D.; Carver, M.; Ambrus, A.; Jones, R. A.; Yang, D. An Intramolecular G-Quadruplex Structure with Mixed Parallel/Antiparallel G-Strands Formed in the Human BCL-2 Promoter Region in Solution. *J. Am. Chem. Soc.* **2006**, *128* (4), 1096–1098. https://doi.org/10.1021/ja055636a.

(59) Sun, D.; Guo, K.; Rusche, J. J.; Hurley, L. H. Facilitation of a Structural Transition in the Polypurine/Polypyrimidine Tract within the Proximal Promoter Region of the Human VEGF Gene by the Presence of Potassium and G-Quadruplex-Interactive Agents. *Nucleic Acids Res* **2005**, *33* (18), 6070–6080. https://doi.org/10.1093/nar/gki917.

(60) Carvalho, J.; Mergny, J.-L.; Salgado, G. F.; Queiroz, J. A.; Cruz, C. G-Quadruplex, Friend or Foe: The Role of the G-Quartet in Anticancer Strategies. *Trends Mol Med* **2020**, *26* (9), 848–861. https://doi.org/10.1016/j.molmed.2020.05.002.

(61) Cimino-Reale, G.; Zaffaroni, N.; Folini, M. Emerging Role of G-Quadruplex DNA as Target in Anticancer Therapy. *Curr Pharm Des* **2016**, *22* (44), 6612–6624. https://doi.org/10.2174/1381612822666160831101031.

(62) *Specific Association of Human Telomerase Activity with Immortal Cells and Cancer*. https://doi.org/10.1126/science.7605428.

(63) Tan, J.; Lan, L. The DNA Secondary Structures at Telomeres and Genome Instability. *Cell & Bioscience* **2020**, *10* (1), 47. https://doi.org/10.1186/s13578-020-00409-z.

(64) Rodriguez, R.; Müller, S.; Yeoman, J. A.; Trentesaux, C.; Riou, J.-F.; Balasubramanian, S. A Novel Small Molecule That Alters Shelterin Integrity and Triggers a DNA-Damage Response at Telomeres. *J Am Chem Soc* **2008**, *130* (47), 15758–15759. https://doi.org/10.1021/ja805615w.

(65) De Magis, A.; Manzo, S. G.; Russo, M.; Marinello, J.; Morigi, R.; Sordet, O.; Capranico, G. DNA Damage and Genome Instability by G-Quadruplex Ligands Are Mediated by R Loops in Human Cancer Cells. *Proc Natl Acad Sci U S A* **2019**, *116* (3), 816–825. https://doi.org/10.1073/pnas.1810409116.

(66) Marchetti, C.; Zyner, K. G.; Ohnmacht, S. A.; Robson, M.; Haider, S. M.; Morton, J. P.; Marsico, G.; Vo, T.; Laughlin-Toth, S.; Ahmed, A. A.; Di Vita, G.; Pazitna, I.; Gunaratnam, M.; Besser, R. J.; Andrade, A. C. G.; Diocou, S.; Pike, J. A.; Tannahill, D.; Pedley, R. B.; Evans, T. R. J.; Wilson, W. D.; Balasubramanian, S.;

Neidle, S. Targeting Multiple Effector Pathways in Pancreatic Ductal Adenocarcinoma with a G-Quadruplex-Binding Small Molecule. *J. Med. Chem.* **2018**, *61* (6), 2500–2517. https://doi.org/10.1021/acs.jmedchem.7b01781.

(67) Marzano, S.; Miglietta, G.; Morigi, R.; Marinello, J.; Arleo, A.; Procacci, M.; Locatelli, A.; Leoni, A.; Pagano, B.; Randazzo, A.; Amato, J.; Capranico, G. Balancing Affinity, Selectivity, and Cytotoxicity of Hydrazone-Based G-Quadruplex Ligands for Activation of Interferon β Genes in Cancer Cells. *J. Med. Chem.* **2022**, acs.jmedchem.2c00772. https://doi.org/10.1021/acs.jmedchem.2c00772.

(68) Sanij, E.; Hannan, K. M.; Xuan, J.; Yan, S.; Ahern, J. E.; Trigos, A. S.; Brajanovski, N.; Son, J.; Chan, K. T.; Kondrashova, O.; Lieschke, E.; Wakefield, M. J.; Frank, D.; Ellis, S.; Cullinane, C.; Kang, J.; Poortinga, G.; Nag, P.; Deans, A. J.; Khanna, K. K.; Mileshkin, L.; McArthur, G. A.; Soong, J.; Berns, E. M. J. J.; Hannan, R. D.; Scott, C. L.; Sheppard, K. E.; Pearson, R. B. CX-5461 Activates the DNA Damage Response and Demonstrates Therapeutic Efficacy in High-Grade Serous Ovarian Cancer. *Nat Commun* **2020**, *11* (1), 2641. https://doi.org/10.1038/s41467-020-16393-4.

(69) *CX-5461 is a DNA G-quadruplex stabilizer with selective lethality in BRCA1/2 deficient tumours | Nature Communications*. https://www.nature.com/articles/ncomms14432 (accessed 2022-10-10).

(70) Drygin, D.; Siddiqui-Jain, A.; O'Brien, S.; Schwaebe, M.; Lin, A.; Bliesath, J.; Ho, C. B.; Proffitt, C.; Trent, K.; Whitten, J. P.; Lim, J. K. C.; Von Hoff, D.; Anderes, K.; Rice, W. G. Anticancer Activity of CX-3543: A Direct Inhibitor of RRNA Biogenesis. *Cancer Research* **2009**, *69* (19), 7653–7661. https://doi.org/10.1158/0008-5472.CAN-09-1304.

(71) Li, Q.; Xiang, J.-F.; Yang, Q.-F.; Sun, H.-X.; Guan, A.-J.; Tang, Y.-L. G4LDB: A Database for Discovering and Studying G-Quadruplex Ligands. *Nucleic Acids Res* **2013**, *41* (Database issue), D1115-1123. https://doi.org/10.1093/nar/gks1101.

(72) Burger, A. M.; Dai, F.; Schultes, C. M.; Reszka, A. P.; Moore, M. J.; Double, J. A.; Neidle, S. The G-Quadruplex-Interactive Molecule BRACO-19 Inhibits Tumor Growth, Consistent with Telomere Targeting and Interference with Telomerase Function. *Cancer Research* **2005**, *65* (4), 1489–1496. https://doi.org/10.1158/0008-5472.CAN-04-2910.

(73) Micco, M.; Collie, G. W.; Dale, A. G.; Ohnmacht, S. A.; Pazitna, I.; Gunaratnam, M.; Reszka, A. P.; Neidle, S. Structure-Based Design and Evaluation of Naphthalene Diimide G-Quadruplex Ligands As Telomere Targeting Agents in Pancreatic Cancer Cells. *J. Med. Chem.* **2013**, *56* (7), 2959–2974. https://doi.org/10.1021/jm301899y.

(74) Gowan, S. M.; Heald, R.; Stevens, M. F. G.; Kelland, L. R. Potent Inhibition of Telomerase by Small-Molecule Pentacyclic Acridines Capable of Interacting with G-Quadruplexes. *Mol Pharmacol* **2001**, *60* (5), 981–988. https://doi.org/10.1124/mol.60.5.981.

(75) Piazza, A.; Boulé, J.-B.; Lopes, J.; Mingo, K.; Largy, E.; Teulade-Fichou, M.-P.; Nicolas, A. Genetic Instability Triggered by G-Quadruplex Interacting Phen-DC Compounds in Saccharomyces Cerevisiae. *Nucleic Acids Res* **2010**, *38* (13), 4337–4348. https://doi.org/10.1093/nar/gkq136.

(76) Ghosh, A.; Trajkovski, M.; Teulade-Fichou, M.-P.; Gabelica, V.; Plavec, J. Phen-DC3 Induces Refolding of Human Telomeric DNA into a Chair-Type Antiparallel G-Quadruplex through Ligand Intercalation. *Angew Chem Int Ed Engl* **2022**, *61* (40), e202207384. https://doi.org/10.1002/anie.202207384.

(77) Prasad, B.; Doimo, M.; Andréasson, M.; L'Hôte, V.; Chorell, E.; Wanrooij, S. A Complementary Chemical Probe Approach towards Customized Studies of G-Quadruplex DNA Structures in Live Cells. *Chem. Sci.* **2022**, *13* (8), 2347–2354. https://doi.org/10.1039/D1SC05816A.

(78) Kim, M.-Y.; Vankayalapati, H.; Shin-Ya, K.; Wierzba, K.; Hurley, L. H. Telomestatin, a Potent Telomerase Inhibitor That Interacts Quite Specifically with the Human Telomeric Intramolecular g-Quadruplex. *J Am Chem Soc* **2002**, *124* (10), 2098–2099. https://doi.org/10.1021/ja017308q.

(79) Izbicka, E.; Wheelhouse, R. T.; Raymond, E.; Davidson, K. K.; Lawrence, R. A.; Sun, D.; Windle, B. E.; Hurley, L. H.; Von Hoff, D. D. Effects of Cationic Porphyrins as G-Quadruplex Interactive Agents in Human Tumor Cells. *Cancer Res* **1999**, *59* (3), 639–644.

(80) Calabrese, D. R.; Chen, X.; Leon, E. C.; Gaikwad, S. M.; Phyo, Z.; Hewitt, W. M.; Alden, S.; Hilimire, T. A.; He, F.; Michalowski, A. M.; Simmons, J. K.; Saunders, L. B.; Zhang, S.; Connors, D.; Walters, K. J.; Mock, B. A.; Schneekloth, J. S. Chemical and Structural Studies Provide a Mechanistic Basis for Recognition of the MYC G-Quadruplex. *Nat Commun* **2018**, *9* (1), 4229. https://doi.org/10.1038/s41467-018-06315-w.

(81) Duarte, A. R.; Cadoni, E.; Ressurreição, A. S.; Moreira, R.; Paulo, A. Design of Modular G-Quadruplex Ligands. *ChemMedChem* **2018**, *13* (9), 869–893. https://doi.org/10.1002/cmdc.201700747.

(82) Bissantz, C.; Kuhn, B.; Stahl, M. A Medicinal Chemist's Guide to Molecular Interactions. *J. Med. Chem.* **2010**, *53* (14), 5061–5084. https://doi.org/10.1021/jm100112j.

(83) Fang, Z.; Song, Y.; Zhan, P.; Zhang, Q.; Liu, X. Conformational Restriction: An Effective Tactic in 'Follow-on'-Based Drug Discovery. *Future Medicinal Chemistry* **2014**, *6* (8), 885–901. https://doi.org/10.4155/fmc.14.50.

(84) Yudin, A. K. Macrocycles: Lessons from the Distant Past, Recent Developments, and Future Directions. *Chem. Sci.* **2014**, *6* (1), 30–49. https://doi.org/10.1039/C4SC03089C.

(85) Lipinski, C. A.; Lombardo, F.; Dominy, B. W.; Feeney, P. J. Experimental and Computational Approaches to Estimate Solubility and Permeability in Drug Discovery and Development Settings. *Advanced Drug Delivery Reviews* **1997**, *23* (1), 3–25. https://doi.org/10.1016/S0169-409X(96)00423-1.

(86) Doak, B. C.; Kihlberg, J. Drug Discovery beyond the Rule of 5 - Opportunities and Challenges. *Expert Opin Drug Discov* **2017**, *12* (2), 115–119. https://doi.org/10.1080/17460441.2017.1264385.

(87) Doak, B. C.; Zheng, J.; Dobritzsch, D.; Kihlberg, J. How Beyond Rule of 5 Drugs and Clinical Candidates Bind to Their Targets. *J Med Chem* **2016**, *59* (6), 2312–2327. https://doi.org/10.1021/acs.jmedchem.5b01286.

(88) Over, B.; Matsson, P.; Tyrchan, C.; Artursson, P.; Doak, B. C.; Foley, M. A.; Hilgendorf, C.; Johnston, S. E.; Iv, M. D. L.; Lewis, R. J.; McCarren, P.; Muncipinto, G.; Norinder, U.; Perry, M. W. D.; Duvall, J. R.; Kihlberg, J. Structural and Conformational Determinants of Macrocycle Cell Permeability. *Nature Chemical Biology* **2016**, *12* (12), 1065–1074. https://doi.org/10.1038/nchembio.2203.

(89) Kelly, C. N.; Townsend, C. E.; Jain, A. N.; Naylor, M. R.; Pye, C. R.; Schwochert, J.; Lokey, R. S. Geometrically Diverse Lariat Peptide Scaffolds Reveal an Untapped Chemical Space of High Membrane Permeability. *J. Am. Chem. Soc.* **2020**, jacs.0c06115. https://doi.org/10.1021/jacs.0c06115.

(90) Villar, E. A.; Beglov, D.; Chennamadhavuni, S.; Porco, J. A.; Kozakov, D.; Vajda, S.; Whitty, A. How Proteins Bind Macrocycles. *Nat Chem Biol* **2014**, *10* (9), 723–731. https://doi.org/10.1038/nchembio.1584.

(91) Meyer, E. A.; Castellano, R. K.; Diederich, F. Interactions with Aromatic Rings in Chemical and Biological Recognition. *Angew Chem Int Ed Engl* **2003**, *42* (11), 1210–1250. https://doi.org/10.1002/anie.200390319.

(92) Salonen, L. M.; Ellermann, M.; Diederich, F. Aromatic Rings in Chemical and Biological Recognition: Energetics and Structures. *Angewandte Chemie International Edition* **2011**, *50* (21), 4808–4842. https://doi.org/10.1002/anie.201007560.

(93) Wheeler, S. E.; Bloom, J. W. G. Toward a More Complete Understanding of Noncovalent Interactions Involving Aromatic Rings. *J. Phys. Chem. A* **2014**, *118* (32), 6133–6147. https://doi.org/10.1021/jp504415p.

(94) Andersson, C. D.; Mishra, B. K.; Forsgren, N.; Ekström, F.; Linusson, A. Physical Mechanisms Governing Substituent Effects on Arene–Arene Interactions in a Protein Milieu. *J. Phys. Chem. B* **2020**, *124* (30), 6529–6539. https://doi.org/10.1021/acs.jpcb.0c03778.

(95) Watt, M.; Hardebeck, L. K. E.; Kirkpatrick, C. C.; Lewis, M. Face-to-Face Arene−Arene Binding Energies: Dominated by Dispersion but Predicted by Electrostatic and Dispersion/Polarizability Substituent Constants. *J. Am. Chem. Soc.* **2011**, *133* (11), 3854–3862. https://doi.org/10.1021/ja105975a.

(96) Wheeler, S. E. Understanding Substituent Effects in Noncovalent Interactions Involving Aromatic Rings. *Acc. Chem. Res.* **2013**, *46* (4), 1029–1038. https://doi.org/10.1021/ar300109n.

(97) Wheeler, S. E.; Houk, K. N. Substituent Effects in the Benzene Dimer Are Due to Direct Interactions of the Substituents with the Unsubstituted Benzene. *J. Am. Chem. Soc.* **2008**, *130* (33), 10854–10855. https://doi.org/10.1021/ja802849j.

(98) Fischer, A.; Smieško, M.; Sellner, M.; Lill, M. A. Decision Making in Structure-Based Drug Discovery: Visual Inspection of Docking Results. *J. Med. Chem.* **2021**, *64* (5), 2489–2500. https://doi.org/10.1021/acs.jmedchem.0c02227.

(99) Ferreira de Freitas, R.; Schapira, M. A Systematic Analysis of Atomic Protein–Ligand Interactions in the PDB. *Med. Chem. Commun.* **2017**, *8* (10), 1970–1981. https://doi.org/10.1039/C7MD00381A.

(100) Mergny, J. L.; Lacroix, L.; Teulade-Fichou, M. P.; Hounsou, C.; Guittat, L.; Hoarau, M.; Arimondo, P. B.; Vigneron, J. P.; Lehn, J. M.; Riou, J. F.; Garestier, T.; Hélène, C. Telomerase Inhibitors Based on Quadruplex Ligands Selected by a Fluorescence Assay. *Proc Natl Acad Sci U S A* **2001**, *98* (6), 3062–3067. https://doi.org/10.1073/pnas.051620698.

(101) Murat, P.; Singh, Y.; Defrancq, E. Methods for Investigating G-Quadruplex DNA/Ligand Interactions. *Chem. Soc. Rev.* **2011**, *40* (11), 5293. https://doi.org/10.1039/c1cs15117g.

(102) Maestro, Schrödinger, LLC, New York, NY, 2021.

(103) Roos, K.; Wu, C.; Damm, W.; Reboul, M.; Stevenson, J. M.; Lu, C.; Dahlgren, M. K.; Mondal, S.; Chen, W.; Wang, L.; Abel, R.; Friesner, R. A.; Harder, E. D. OPLS3e: Extending Force Field Coverage for Drug-Like Small Molecules. *J. Chem. Theory Comput.* **2019**, *15* (3), 1863–1874. https://doi.org/10.1021/acs.jctc.8b01026.

(104) *MacroModel, Schrödinger, LLC, New York, NY, 2021.*

(105) Olanders, G.; Alogheli, H.; Brandt, P.; Karlén, A. Conformational Analysis of Macrocycles: Comparing General and Specialized Methods. *J Comput Aided Mol Des* **2020**, *34* (3), 231–252. https://doi.org/10.1007/s10822-020-00277-2.

(106) Becke, A. D. Density-Functional Exchange-Energy Approximation with Correct Asymptotic Behavior. *Phys. Rev. A* **1988**, *38* (6), 3098–3100. https://doi.org/10.1103/PhysRevA.38.3098.

(107) Lee, C.; Yang, W.; Parr, R. G. Development of the Colle-Salvetti Correlation-Energy Formula into a Functional of the Electron Density. *Phys. Rev. B* **1988**, *37* (2), 785–789. https://doi.org/10.1103/PhysRevB.37.785.

(108) Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. A Consistent and Accurate Ab Initio Parametrization of Density Functional Dispersion Correction (DFT-D) for the 94 Elements H-Pu. *J. Chem. Phys.* **2010**, *132* (15), 154104. https://doi.org/10.1063/1.3382344.

(109) Bochevarov, A. D.; Harder, E.; Hughes, T. F.; Greenwood, J. R.; Braden, D. A.; Philipp, D. M.; Rinaldo, D.; Halls, M. D.; Zhang, J.; Friesner, R. A. Jaguar: A High-Performance Quantum Chemistry Software Program with Strengths in Life and Materials Sciences. *International Journal of Quantum Chemistry* **2013**, *113* (18), 2110–2142. https://doi.org/10.1002/qua.24481.

(110) Prasad, B.; Jamroskovic, J.; Bhowmik, S.; Kumar, R.; Romell, T.; Sabouri, N.; Chorell, E. Flexible Versus Rigid G-Quadruplex DNA Ligands: Synthesis of Two Series of Bis-Indole Derivatives and Comparison of Their Interactions with G-Quadruplex DNA. *Chemistry – A European Journal* **2018**, *24* (31), 7926–7938. https://doi.org/10.1002/chem.201800078.

(111) Prasad, B.; Das, R. N.; Jamroskovic, J.; Kumar, R.; Hedenström, M.; Sabouri, N.; Chorell, E. The Relation Between Position and Chemical Composition of Bis-Indole Substituents Determines Their Interactions

with G-Quadruplex DNA. *Chem. Eur. J.* **2020**, *26* (43), 9561–9572. https://doi.org/10.1002/chem.202000579.

(112) Carvalho, J.; Quintela, T.; Gueddouda, N. M.; Bourdoncle, A.; Mergny, J.-L.; Salgado, G. F.; Queiroz, J. A.; Cruz, C. Phenanthroline Polyazamacrocycles as G-Quadruplex DNA Binders. *Org. Biomol. Chem.* **2018**, *16* (15), 2776–2786. https://doi.org/10.1039/C8OB00247A.

(113) Ma, Y.; Iida, K.; Sasaki, S.; Hirokawa, T.; Heddi, B.; Phan, A. T.; Nagasawa, K. Synthesis and Telomeric G-Quadruplex-Stabilizing Ability of Macrocyclic Hexaoxazoles Bearing Three Side Chains. *Molecules* **2019**, *24* (2), 263. https://doi.org/10.3390/molecules24020263.

(114) Agarwal, T.; Roy, S.; Chakraborty, T. K.; Maiti, S. Selective Targeting of G-Quadruplex Using Furan-Based Cyclic Homooligopeptides: Effect on c-MYC Expression. *Biochemistry* **2010**, *49* (38), 8388–8397. https://doi.org/10.1021/bi1005927.

(115) Shirude, P. S.; Gillies, E. R.; Ladame, S.; Godde, F.; Shin-ya, K.; Huc, I.; Balasubramanian, S. Macrocyclic and Helical Oligoamides as a New Class of G-Quadruplex Ligands. *J. Am. Chem. Soc.* **2007**, *129* (39), 11890–11891. https://doi.org/10.1021/ja073775h.

(116) Monchaud, D.; Granzhan, A.; Saettel, N.; Guédin, A.; Mergny, J.-L.; Teulade-Fichou, M.-P. " *One Ring to Bind Them All"* —Part I: The Efficiency of the Macrocyclic Scaffold for G-Quadruplex DNA Recognition. *Journal of Nucleic Acids* **2010**, *2010*, 1–19. https://doi.org/10.4061/2010/525862.

(117) Granzhan, A.; Monchaud, D.; Saettel, N.; Guédin, A.; Mergny, J.-L.; Teulade-Fichou, M.-P. " *One Ring to Bind Them All"* —Part II: Identification of Promising G-Quadruplex Ligands by Screening of Cyclophane-Type Macrocycles. *Journal of Nucleic Acids* **2010**, *2010*, 1–11. https://doi.org/10.4061/2010/460561.

(118) Pillaiyar, T.; Köse, M.; Sylvester, K.; Weighardt, H.; Thimm, D.; Borges, G.; Förster, I.; von Kügelgen, I.; Müller, C. E. Diindolylmethane Derivatives: Potent Agonists of the Immunostimulatory Orphan G Protein-Coupled Receptor GPR84. *J. Med. Chem.* **2017**, *60* (9), 3636–3655. https://doi.org/10.1021/acs.jmedchem.6b01593.

(119) Wang, W.-X.; Zhang, Q.-Z.; Zhang, T.-Q.; Li, Z.-S.; Zhang, W.; Yu, W. N-Bromosuccinimide-Mediated Radical Cyclization of 3-Arylallyl Azides: Synthesis of 3-Substituted Quinolines. *Advanced Synthesis & Catalysis* **2015**, *357* (1), 221–226. https://doi.org/10.1002/adsc.201400637.

(120) O'Hara, F.; Blackmond, D. G.; Baran, P. S. Radical-Based Regioselective C–H Functionalization of Electron-Deficient Heteroarenes: Scope, Tunability, and Predictability. *J. Am. Chem. Soc.* **2013**, *135* (32), 12122–12134. https://doi.org/10.1021/ja406223k.

(121) Dunetz, J. R.; Xiang, Y.; Baldwin, A.; Ringling, J. General and Scalable Amide Bond Formation with Epimerization-Prone Substrates Using T3P and Pyridine. *Org. Lett.* **2011**, *13* (19), 5048–5051. https://doi.org/10.1021/ol201875q.

(122) Martí-Centelles, V.; Pandey, M. D.; Burguete, M. I.; Luis, S. V. Macrocyclization Reactions: The Importance of Conformational, Configurational, and Template-Induced Preorganization. *Chem. Rev.* **2015**, *115* (16), 8736–8834. https://doi.org/10.1021/acs.chemrev.5b00056.

(123) Beutner, G. L.; Young, I. S.; Davies, M. L.; Hickey, M. R.; Park, H.; Stevens, J. M.; Ye, Q. TCFH–NMI: Direct Access to *N* -Acyl Imidazoliums for Challenging Amide Bond Formations. *Org. Lett.* **2018**, *20* (14), 4218–4222. https://doi.org/10.1021/acs.orglett.8b01591.

(124) *G4-FID: A Fluorescent DNA Probe Displacement Assay for Rapid Evaluation of Quadruplex Ligands | SpringerLink*. https://link.springer.com/protocol/10.1007/978-1-59745-363-9_15 (accessed 2022-10-13).

(125) *Synthesis of 4-, 5-, 6-, and 7-azidotryptamines | Elsevier Enhanced Reader*. https://doi.org/10.1016/j.tetlet.2008.10.091.

(126) Giles, P. R.; Marson, C. M. Dimethylchloromethyleneammonium Chloride. In *Encyclopedia of Reagents for Organic Synthesis*; John Wiley & Sons, Ltd, 2001. https://doi.org/10.1002/047084289X.rd319m.

(127) El-Faham, A.; Albericio, F. COMU: A Third Generation of Uronium-Type Coupling Reagents. *Journal of Peptide Science* **2010**, *16* (1), 6–9. https://doi.org/10.1002/psc.1204.

(128) Quinn, J. F.; Bryant, C. E.; Golden, K. C.; Gregg, B. T. Reduction of Nitroarenes with Hydrogen Transfer Catalysis. *Synfacts* **2010**, *2010* (4), 492–492. https://doi.org/10.1055/s-0029-1219464.

(129) Jamroskovic, J.; Livendahl, M.; Eriksson, J.; Chorell, E.; Sabouri, N. Identification of Compounds That Selectively Stabilize Specific G-Quadruplex Structures by Using a Thioflavin T-Displacement Assay as a Tool. *Chemistry – A European Journal* **2016**, *22* (52), 18932–18943. https://doi.org/10.1002/chem.201603463.

(130) Jamroskovic, J.; Doimo, M.; Chand, K.; Obi, I.; Kumar, R.; Brännström, K.; Hedenström, M.; Nath Das, R.; Akhunzianov, A.; Deiana, M.; Kasho, K.; Sulis Sato, S.; Pourbozorgi, P. L.; Mason, J. E.; Medini, P.; Öhlund, D.; Wanrooij, S.; Chorell, E.; Sabouri, N. Quinazoline Ligands Induce Cancer Cell Death through Selective STAT3 Inhibition and G-Quadruplex Stabilization. *J. Am. Chem. Soc.* **2020**, *142* (6), 2876–2888. https://doi.org/10.1021/jacs.9b11232.

(131) Wan, Z.-K.; Wacharasindhu, S.; Binnun, E.; Mansour, T. An Efficient Direct Amination of Cyclic Amides and Cyclic Ureas †. *Organic letters* **2006**, *8*, 2425–2428. https://doi.org/10.1021/ol060815y.

(132) Felsenstein, K. M.; Saunders, L. B.; Simmons, J. K.; Leon, E.; Calabrese, D. R.; Zhang, S.; Michalowski, A.; Gareiss, P.; Mock, B. A.; Schneekloth, J. S. Small Molecule Microarrays Enable the Identification of a Selective, Quadruplex-Binding Inhibitor of MYC Expression. *ACS Chem. Biol.* **2016**, *11* (1), 139–148. https://doi.org/10.1021/acschembio.5b00577.

(133) Asamitsu, S.; Bando, T.; Sugiyama, H. Ligand Design to Acquire Specificity to Intended G-Quadruplex Structures. *Chemistry – A European Journal* **2019**, *25* (2), 417–430. https://doi.org/10.1002/chem.201802691.

(134) Redman, J. E.; Granadino-Roldán, J. M.; Schouten, J. A.; Ladame, S.; Reszka, A. P.; Neidle, S.; Balasubramanian, S. Recognition and Discrimination of DNA Quadruplexes by Acridine-Peptide Conjugates. *Org. Biomol. Chem.* **2009**, *7* (1), 76–84. https://doi.org/10.1039/B814682A.

(135) Asamitsu, S.; Obata, S.; Phan, A. T.; Hashiya, K.; Bando, T.; Sugiyama, H. Simultaneous Binding of Hybrid Molecules Constructed with Dual DNA-Binding Components to a G-Quadruplex and Its Proximal Duplex. *Chemistry – A European Journal* **2018**, *24* (17), 4428–4435. https://doi.org/10.1002/chem.201705945.

(136) Nguyen, T. Q. N.; Lim, K. W.; Phan, A. T. A Dual-Specific Targeting Approach Based on the Simultaneous Recognition of Duplex and Quadruplex Motifs. *Sci Rep* **2017**, *7* (1), 11969. https://doi.org/10.1038/s41598-017-10583-9.

(137) Tassinari, M.; Zuffo, M.; Nadai, M.; Pirota, V.; Sevilla Montalvo, A. C.; Doria, F.; Freccero, M.; Richter, S. N. Selective Targeting of Mutually Exclusive DNA G-Quadruplexes: HIV-1 LTR as Paradigmatic Model. *Nucleic Acids Res* **2020**, *48* (9), 4627–4642. https://doi.org/10.1093/nar/gkaa186.

(138) Chen, S.-B.; Hu, M.-H.; Liu, G.-C.; Wang, J.; Ou, T.-M.; Gu, L.-Q.; Huang, Z.-S.; Tan, J.-H. Visualization of *NRAS* RNA G-Quadruplex Structures in Cells with an Engineered Fluorogenic Hybridization Probe. *J. Am. Chem. Soc.* **2016**, *138* (33), 10382–10385. https://doi.org/10.1021/jacs.6b04799.

(139) Roberts, T. C.; Langer, R.; Wood, M. J. A. Advances in Oligonucleotide Drug Delivery. *Nat Rev Drug Discov* **2020**, *19* (10), 673–694. https://doi.org/10.1038/s41573-020-0075-7.

(140) Stump, B. Click Bioconjugation: Modifying Proteins Using Click-Like Chemistry. *ChemBioChem* **2022**, *23* (16), e202200016. https://doi.org/10.1002/cbic.202200016.