



**HAL**  
open science

# Combinatorial chemistry approaches for the development of G-quadruplex DNA and RNA ligands

Oksana Reznichenko

► **To cite this version:**

Oksana Reznichenko. Combinatorial chemistry approaches for the development of G-quadruplex DNA and RNA ligands. Organic chemistry. Université Paris-Saclay, 2021. English. NNT : 2021UPASF014 . tel-03620871

**HAL Id: tel-03620871**

**<https://theses.hal.science/tel-03620871>**

Submitted on 27 Mar 2022

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Les approches de chimie combinatoire  
pour le développement des ligands de G-  
quadruplexes de l'ADN et de l'ARN  
*Combinatorial chemistry approaches for the  
development of G-quadruplex DNA and RNA ligands*

**Thèse de doctorat de l'université Paris-Saclay**

École doctorale n°571 Sciences chimiques : molécules, matériaux,  
instrumentation et biosystèmes (2MIB)  
Spécialité de doctorat: chimie  
Unité de recherche : Université Paris-Saclay, CNRS, Inserm, Chimie et modélisation  
pour la biologie du cancer, 91405, Orsay, France  
Réfèrent : Faculté des sciences d'Orsay

**Thèse présentée et soutenue à Paris-Saclay,  
le 26/03/2021, par**

**Oksana REZNICHENKO**

**Composition du Jury**

**Dominique GUIANVARC'H**

Professeure, Université Paris-Sud (UMR 8182)

Présidente

**Thomas LAVERGNE**

Chargé de recherche, HDR, Université Grenoble Alpes

Rapporteur & Examineur

**Liliya YATSUNYK**

Professeur agrégé, equiv. HDR, Swarthmore College

Rapporteur & Examineur

**Marc BLONDEL**

Professeur, Université de Brest

Examineur

**Jean-Louis MERGNY**

Directeur de recherche, École Polytechnique

Examineur

**Direction de la thèse**

**Anton GRANZHAN**

Chargé de recherche, CNRS, Institut Curie (UMR 9187 –  
U 1196)

Directeur de thèse

**Anthony BUGAUT**

Chargé de recherche, Université de Bordeaux

Membre invité



**Titre :** Les approches de chimie combinatoire pour le développement des ligands de G-quadruplexes de l'ADN et de l'ARN

**Mots clés :** chimie dynamique combinatoire, chimie combinatoire, ligands de l'ADN, G-quadruplexes, chimie organique, chimie supramoléculaire.

**Résumé :** Les G-quadruplexes (G4s) sont des structures non-canoniques d'acides nucléiques (ADN et ARN) constituées d'au moins deux quartets de guanines. L'une des propriétés importantes des G4s est leur capacité à former des complexes avec de petites molécules exogènes (« ligands ») et d'influencer ainsi les processus biologiques dans lesquels ils sont impliqués. Ainsi, l'interaction de petites molécules avec certaines structures G4s permettrait de diminuer l'expression de certains oncogènes, d'inhiber la télomérase ou encore d'induire des dommages à l'ADN.

Ce travail vise à développer des méthodologies rapides et simples pour la synthèse et le criblage des molécules afin d'identifier des ligands sélectifs et affins de structures non-canoniques d'acides nucléiques, en particulier des G4s. Plus précisément, ce travail explore la synthèse réversible d'acylhydrazones, jusqu'ici peu appliquée pour le développement de ligands de l'ADN et de l'ARN. Dans un premier temps, une série de 20 bis(acylhydrazones), analogues des ligands PDC (360A) et PhenDC3, a été obtenue par la synthèse préparative. Les expériences de dénaturation thermique suivie par fluorescence ont démontré que certains de ces composés avaient une bonne affinité pour l'ADN G4. Ces expériences ont permis de valider le potentiel du motif acylhydrazone pour le développement de ligands des G4s. Ensuite, une méthode de chimie dynamique combinatoire (CDC) a été développée. Cette dernière consiste en génération de bibliothèques combinatoires comportant jusqu'à 20 composés, suivie par l'isolement des ligands les plus affins par la précipitation avec la cible, immobilisée sur des billes magnétiques.

Ainsi, un bis(acylhydrazone) non-symétrique a été identifié comme un ligand prometteur du G4 parallèle Pu24T. Cependant, les expériences avec ses proches analogues n'ont pas confirmé son affinité aux G4 augmentée par rapport aux dérivés symétriques. Il a été supposé que les résultats d'expériences de CDC pouvaient être biaisés par des interactions non-spécifiques entre les ligands et les billes magnétiques. Pour améliorer l'analyse des

bibliothèques combinatoires, une nouvelle méthode basée sur l'extraction en phase solide des ligands a été développée et appliquée à deux bibliothèques d'acylhydrazones non-symétriques. Huit hits ont été obtenus à partir de 70 composés générés *in situ*. Trois d'entre eux ont été sélectionnés pour la synthèse préparative et une étude de l'interaction avec l'ADN G4.

En parallèle, une approche classique de chimie combinatoire a été élaborée, ce qui a conduit à la génération d'une bibliothèque combinatoire de 90 dérivés bis(acylhydrazone) sous forme de solutions 2 mM dans DMSO prêtes à l'emploi, avec une pureté moyenne de 87%. Ces échantillons ont été utilisés directement dans le criblage biophysique contre quatre G4s de l'ADN de trois topologies différentes. Les composés les plus actifs ont été synthétisés d'une manière préparative et leur interaction avec les G4s a été étudiée en détail par des méthodes biophysiques, y compris la spectrométrie de masse native. Ainsi, au moins un dérivé avec une affinité pour les G4s supérieure à celle de PhenDC3 et une sélectivité inédite pour le G4 antiparallèle a été identifié.

Enfin, dans le cadre d'un projet collaboratif (M. Blondel, Université de Bretagne Occidentale), des ligands synthétisés au cours de ce travail ont été étudiés vis-à-vis de leur capacité à moduler d'évasion immune du virus d'Epstein-Barr (EBV). Il a été démontré que certains bis(acylhydrazones) interagissent *in vitro* avec la séquence riche en guanines de l'ARNm codante pour le domaine riche en glycine-alanine (GAR) de la protéine virale EBNA1. Deux de ces dérivés déplacent le facteur de la cellule hôte (nucléoline) de l'ARNm d'EBNA1, conduisant ainsi à la surexpression de la protéine et à la présence exacerbée de peptides antigéniques sur les cellules infectées. Cet effet représente une opportunité thérapeutique pour le traitement des cancers associés à l'EBV.

**Title :** Combinatorial chemistry approaches for the development of G-quadruplex DNA and RNA ligands

**Keywords :** dynamic combinatorial chemistry, combinatorial chemistry, DNA ligands, G-quadruplexes, organic chemistry, supramolecular chemistry.

**Abstract :** G-quadruplexes (G4s) are four-stranded structures of nucleic acids (DNA or RNA) that consist of at least two coplanar guanine quartets. An important feature of G4s is their ability to form stable complexes with exogenous small molecules (ligands) and thus influence biological processes in which they are involved. G4 targeting is often associated with oncology, where G4 ligands may suppress the expression of oncogenes, inhibit telomerase, or induce DNA damage in cancer cells.

This work aims to develop methodologies for rapid and simple synthesis and screening of compounds, in order to identify selective and highly affine ligands of given non-canonical structures of nucleic acids, in particular G4s. Specifically, this work exploits the chemistry of reversible synthesis of acylhydrazones, which has been barely applied for the development of DNA or RNA ligands before. First, a small library of 20 cationic bis(acylhydrazones), analogues of the previously reported G4-ligands PDC (360A) and PhenDC3, was obtained by preparative synthesis. Through fluorescence melting experiments it is demonstrated that some of compounds indeed have high affinity to G4-DNA, validating the suitability of the acylhydrazone motif as a scaffold for the development of G4 ligands. Next, a method of dynamic combinatorial chemistry (DCC), which consists in simultaneous one-pot generation of libraries of up to 20 compounds with consecutive pull-down of most affine ligands by bead-immobilized targets (i.e., G4-DNA), was developed. By using this method, a non-symmetrical bis(acylhydrazone) was identified as a promising ligand of a parallel G4-DNA Pu24T. However, biophysical experiments with its close structural analogues did not confirm their preferential binding in comparison with the symmetrically substituted compound. It is proposed that the outcome of DCC experiments may be biased by non-specific interactions of ligands with magnetic beads, leading to false-positive results. In order to improve the analysis of dynamic combinatorial libraries, a novel method based on solid-phase extraction of the G4-

ligand complex was developed and applied to two libraries of non-symmetric acylhydrazones. In a few rounds of selection, 13 hits were obtained out of 70 *in situ* generated compounds. Three of them were selected for preparative synthesis and detailed study of interaction with G4-DNA.

In parallel, a classical combinatorial chemistry approach was developed, resulting in generation of a combinatorial library of 90 individual bis(acylhydrazone) derivatives in the form of ready-to-use 2 mM solutions in DMSO, with an average purity of 87%. These samples were directly used for biophysical screening experiments towards four G4-DNA targets of three different topologies. Three most active compounds were obtained in preparative manner and their interaction with the mentioned biological targets was studied in detail by several biophysical methods, including native mass spectrometry experiments. This way, at least one derivative with a G4-DNA affinity superior to that of PhenDC3 and unprecedented selectivity towards anti-parallel G4-DNA could be identified.

Finally, in the framework of a collaborative project (M. Blondel, University of Western Brittany) the ligands synthesized in this work were studied with respect to their capacity to act as modulators of the immune evasion of Epstein-Barr virus (EBV). Specifically, it was shown that several bis(acylhydrazones) bind *in vitro* to G4-RNA structures formed by the guanine-rich repeat sequence of mRNA encoding for the glycine-alanine rich (GAR) domain of viral genome maintenance protein EBNA1. Moreover, two derivatives were found to displace the host cell factor nucleolin from EBNA1 mRNA, leading to overexpression of EBNA1 protein and a concomitant increase of antigen presentation in EBV-infected cell cultures. This effect represents an interesting therapeutic opportunity for treatment of EBV-related cancers.

## ***Acknowledgments***

First of all, I would like to thank Marie-Paule Teulade-Fichou and Florence Mahuteau-Betzer for allowing me to complete my thesis within their unit. I am also grateful for their interesting remarks and good advises.

I would like to thank my thesis director Dr. Anton Granzhan. I am very happy that my research was supervised by such a kind, delicate, supportive, interesting person of outstanding mind. I am truly grateful for everything he taught me, including not only methods and techniques connected to my research, but also for helping me to become more critical, attentive to details, consistent and organized.

I would like to thank Thomas Lavergne for being part of my thesis monitoring committee, for giving me good advises about my research and career and for accepting, along with Jean-Louis Mergny to evaluate my thesis work as a referee. Thank you both, I am looking forward to our discussion. I would also like to thank Marc Blondel, Antony Bugaut and Dominique Guianvarc'h for the time they will devote to examine the results of these three years of my research.

I would also like to thank Dr. Valérie Gabelica for welcoming me to her laboratory at IECB, Pessac, France. Within two short stays, I was given an exciting opportunity to perform native mass spectrometry experiments. I would like to thank Dr. Eric Largy, Dr. Jorge Gonzalez Garcia and Dr. Frederic Rosu and all members of Dr. Valérie Gabelica's team for their assistance and warm reception.

I would also like to thank Dr. Lukáš Trantírek for welcoming me to his laboratory at CEITEC in Brno, Czech Republic, and giving me a great opportunity to test some hypothesis and learn new techniques. I am deeply grateful to all member of his team, especially to Pavlína Víšková and Eva Ištvánková, for sharing with me their passion to science, for teaching me basics of cell manipulating techniques and cell imaging, for conducting with me native PAGE experiments and helping me with analysis of NMR spectra of G-quadruplexes and *i*-motifs. I have a very warm memory of this so short, but so exciting stay in Brno.

I would like to thank all people that helped me with my experiments. I thank Dr. Guillot Régis for his help with crystallographic analysis, Dr. Carlos Kikuti (Institut Curie), as well as Dr. Patrick England and Dr. Bruno Baron (Platform Biophysics of Macromolecules, Institut Pasteur) for their assistance for MST experiments.

Special thanks is to Anne Cucchiarini, our “HPLC Master” and “social manager” who was helping me with HPLC analyses and made a huge contribution in the development of dynamic combinatorial methods, discussed here.

I would like to thank Dr. Liliane Mouawad (Institut Curie) and Denis Leclerc for performing docking experiments.

I would like to thank Sylvain Trepout (Institut Curie) for TEM images of nanoparticles.

I would like to thank all members of our friendly team. I would like to thank Jaime Pinto Franco for being so helpful, attentive and encouraging and also for our interesting talks. I would like to thank all ancient members of our team, including Dr. Michela Zuffo and Dr. Coralie Caron and Anne Cucchiarini for their helpful advices and suggestions. I am keeping warm memories of our adventures in Cabourg, Montpellier and our volleyball sessions.

I would also like to thank all members of our unit, especially Dr. Daniela Verga and Thibaut Masson for their openness to discussions, helpful advices and assistance for experiments. I would like to thank Nathalie, Claire, Delphine, Pauline, Romain, Corinne, Sandrine, Chloé, Julie, Sounderya, Yu, Rahima and Pedro for their assistance and nice company.

I am grateful to my family, particularly to my mother, Olena Reznichenko, for unconditional support and assistance, wise insights and eternal believe in her daughter. I am also grateful to my boyfriend, Johannes Salomonsson, for his support and all joy that he brings in my life.

# *Curriculum vitae*

## **Education**

2017-2021 – DOCTORAL STUDIES – Paris-Saclay University (France) – Chemical Sciences: Molecules, Materials, Instrumentation and Biosystems, Organic and Biomolecular Chemistry

2015-2017 – MASTER DEGREE – Erasmus+ double-diploma

– University of Maine (France) - Functional Materials and Nanoscience, Fine Chemistry

– Taras Shevchenko National University of Kyiv (Ukraine), Chemistry, Organic chemistry

2011-2015 – BACHELOR DEGREE – Taras Shevchenko National University of Kyiv (Ukraine) – Faculty of Chemistry, Organic chemistry

## **Professional experience**

2017-2020 – Institut Curie, Orsay – Doctoral project “Combinatorial chemistry approaches for the development of G-quadruplex DNA and RNA ligands”

2016-2017 – The Institute of Molecules and Materials of Le Mans – Master project «Synthesis of novel NHC-ligands for their application in asymmetric metathesis»

2012 - 2016 – Enamine Ltd., Kyiv – Internship at the chemical company: synthesis of fine chemicals

2012 - 2015 – Enamine Ltd., Kyiv – Bachelor’s thesis «Synthesis and exploration of products of Horner-Wadsworth-Emmons reaction for perfluoralkylhetarylketones».

## **Publications**

Xie, X.; Reznichenko, O.; Chaput, L.; Martin, P.; Teulade-Fichou, M.-P.; Granzhan, A. Topology-Selective, Fluorescent “Light-Up” Probes for G-Quadruplex DNA Based on Photoinduced Electron Transfer. *Chem. Eur. J.* **2018**, *24*, 12638-12651. <https://doi.org/10.1002/chem.201801701>

Reznichenko, O.; Quillévéré, A.; Martins, R. P.; Loaëc, N; Kang, H.; Lista, M.J.; Beauvineau, C.; González-García, J.; Guillot, R.; Voisset, C; Daskalogianni, C.; Fåhraeus, R.; Teulade-Fichou, M.P.; Blondel, M.; Granzhan, A. Novel cationic bis(acylhydrazones) as modulators of Epstein–Barr virus immune evasion acting through disruption of interaction between nucleolin and G-quadruplexes of EBNA1 mRNA. *Eur. J. Med. Chem.* **2019**, *178*, 13-29. <https://doi.org/10.1016/j.ejmech.2019.05.042>

Reznichenko, O.; Cucchiaroni, A.; Granzhan, A. Quadruplex DNA-guided ligand selection from dynamic combinatorial libraries of acylhydrazones. *Org. Biomol. Chem.*, 2021, **19**, 379-386.

<https://doi.org/10.1039/D0OB01908A>



## Patents

Granzhan, A.; Reznichenko, O.; Beauvineau, C.; Teulade-Fichou, M.P.; Daskalohgianni, C.; Prado-Martins, R.; Farhaeus, R.; Lista, M.J.; Voisset, C.; Quillévéré, A.; Blondel, M. Novel hydrazone derivatives for preventing or treating ebv-related cancer. EP18305392 (2018)

## Conferences

First French industrial chemistry symposium, (Paris, France, 2020) – poster *“Combinatorial chemistry of acylhydrazones for G-quadruplex targeting”*;

PhD Day of the 2MIB Doctorale School of Université Paris-Saclay (Evry, France, 2019) – poster presentation *“Novel combinatorial chemistry approach for rapid synthesis and evaluation of bis(acylhydrazone)-type G-quadruplex binders”*

PhD Day of the 2MIB Doctorale School of Université Paris-Saclay (Orsay, France, 2018) – poster presentation *“Development of bis(acylhydrazone) libraries of G-quadruplex DNA binders: synthesis, pull-down assay and virtual screening”*

International Symposium Biology of non-canonical nucleic acids: from humans to pathogens (Padova, Italy, 2018) – small talk presentation and poster presentation *“Topology-Selective Fluorescent “Light-Up” Probes for G-Quadruplex DNA Based on Photoinduced Electron Transfer”*

## Languages

Ukrainian – native

French – advanced

Russian – native

Swedish – beginner

English – advanced

## Hobbies

Violin, dance, hiking, reading.

# Contents

Acknowledgments.....	- 1 -
Curriculum vitae .....	- 3 -
List of abbreviations .....	- 9 -
INTRODUCTION .....	- 11 -
1. General overview of G-quadruplexes .....	- 12 -
1.1 G-quadruplex structure: general principles .....	- 12 -
1.2 Biological relevance of G-quadruplexes.....	- 14 -
1.3 G-quadruplexes at telomeres.....	- 15 -
1.4 G-quadruplexes as transcription regulators .....	- 18 -
1.5 G-quadruplexes as translation regulators: the case of viral genome maintenance protein EBNA1.....	- 23 -
2. G4-ligands.....	- 25 -
3. Combinatorial chemistry approaches to discover G4 ligands.....	- 36 -
3.1. Classical combinatorial chemistry approaches .....	- 37 -
3.2. DNA-encoded libraries (DECL).....	- 42 -
3.3. Phage display.....	- 44 -
3.4. Fragment-based drug discovery.....	- 46 -
3.5. Kinetic target-directed synthesis .....	- 48 -
3.6 Dynamic combinatorial chemistry .....	- 52 -
4. Biophysical techniques to study DNA–ligand interactions .....	- 59 -
4.1 Circular dichroism (CD).....	- 59 -
4.2 FRET-melting .....	- 61 -
4.3 FID assay.....	- 64 -
4.4 Fluorescence quenching assay (fluorimetric titrations).....	- 65 -

4.5 Native mass spectrometry (MS).....	- 67 -
AIMS OF THE WORK .....	- 71 -
RESULTS AND DISCUSSION .....	- 77 -
1. Synthesis and studies of model ligands .....	- 78 -
1.1 Synthesis of building blocks .....	- 78 -
1.2 Preparative synthesis of model bis(acylhydrazones).....	- 81 -
1.3 Structural characterization of bis(acylhydrazone) compounds .....	- 83 -
1.4 Evaluation of affinity of bis(acylhydrazones) to G-quadruplexes by FRET-melting-	85 -
1.5 Molecular modeling (in collaboration with Dr Liliane Mouawad, M.Sc. 2 project of Denis Leclerc) .....	- 88 -
1.6 Conclusions.....	- 96 -
2. Dynamic combinatorial chemistry .....	- 98 -
2.1 Preliminary experiments and optimization of conditions for DCC generation.....	- 98 -
2.2 Pull-down experiment with streptavidin-coated magnetic beads .....	- 104 -
2.3 Hits identification & synthesis.....	- 110 -
2.4 Synthesis of building blocks and selected compounds.....	- 112 -
2.5 Evaluation of G4-binding properties .....	- 116 -
2.6 Conclusions.....	- 128 -
3. Identification of Optimal G-quadruplex Ligands by Combining Parallel Synthesis and Profiling of Cationic Bis(acylhydrazones) .....	- 131 -
3.1 Library design and synthesis. ....	- 131 -
3.2 Screening of interaction with G4-DNA by FRET-melting experiments. ....	- 139 -
3.3 Hit validation. ....	- 142 -
3.4 Conclusions.....	- 163 -
4. Ligand pull-down beyond streptavidin-coated magnetic beads.....	- 165 -
4.1 Experiments with gold-coated magnetic nanoparticles .....	- 165 -

4.2 Solid phase extraction .....	- 170 -
5. Cationic bis(acylhydrazones) as potential drugs for treating EBV-related cancers ...	- 183 -
5.1. <i>In silico</i> evaluation .....	- 183 -
5.2 Biophysical studies of interaction of novel ligands with G4 of EBNA1 mRNA and G4-DNA.....	- 187 -
5.3. Biological tests.....	- 189 -
5.4 Pharmacological properties of selected drug candidates.....	- 194 -
5.5 Conclusions.....	- 196 -
CONCLUSIONS AND PERSPECTIVES .....	- 199 -
EXPERIMENTAL SECTION .....	- 205 -
Synthesis.....	- 206 -
RESULTS AND DISCUSSION. Part 1. ....	- 206 -
RESULTS AND DISCUSSION. Part 2. ....	- 231 -
RESULTS AND DISCUSSION. Part 3. ....	- 239 -
RESULTS AND DISCUSSION. Part 4. ....	- 248 -
Single-crystal X-ray diffraction analysis.....	- 251 -
Biophysical assays .....	- 252 -
FRET-melting assay.....	- 252 -
TO displacement assay.....	- 252 -
Fluorimetric titrations .....	- 253 -
CD spectroscopy.....	- 254 -
Electrospray mass spectrometry.....	- 254 -
Protocols of generation of dynamic combinatorial libraries. ....	- 255 -
RESULTS AND DISCUSSION. Part 2. ....	- 255 -
RESULTS AND DISCUSSION. Part 4. ....	- 255 -
REFERENCES .....	- 259 -

Annexes .....	- 277 -
<i>Table A</i> .....	- 278 -
Resumé en français .....	- 282 -

## *List of abbreviations*

5' UTR: 5' untranslated region

ALT: alternative lengthening of telomeres

BLI: bio-layer interferometry

Bn: benzyl

DCC: dynamic combinatorial chemistry

DCL: dynamic combinatorial library

DCM: dichloromethane

DMA: dimethylacetamide

DMF: dimethylformamide

DMSO: dimethyl Sulfoxide

DNA: Deoxyribonucleic acid

Ds: double-strand

EBNA1: Epstein-Barr nuclear antigen 1

EBV: Epstein-Barr virus

G4: G-quadruplex

HPLC: high performance liquid chromatography

ITC: isothermal titration calorimetry

MeCN: acetonitrile

MST: microscale thermophoresis

NHE: nuclease hypersensitive element

NMR: nuclear magnetic resonance

NP: nanoparticle

POT1: protection of telomeres protein 1

PQS: putative G4 sequences

RNA: ribonucleic acid

SPR: surface plasmon resonance.

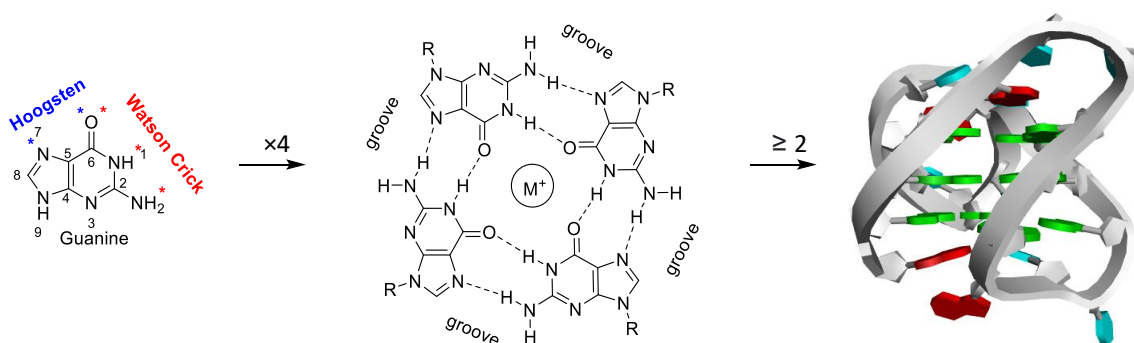
# ***INTRODUCTION***



# 1. General overview of G-quadruplexes

## 1.1 G-quadruplex structure: general principles

G-quadruplex (G4) is a four-stranded structure of nucleic acids that contains at least two quartets of guanines stacked one upon another (*Figure 1*). Guanine bases in the quartet are connected with Hoogsteen hydrogen bonds. Every guanine in a quartet has two hydrogen-bond donors (N1H and N2H) and two hydrogen-bond acceptors (O6 and N7). Cations of appropriate size, such as  $\text{Na}^+$ ,  $\text{K}^+$  or  $\text{NH}_4^+$ , stabilize the G-quadruplex by coordinating its O6 guanine atoms.<sup>1</sup> While  $\text{Na}^+$  (ionic radius 0.95 Å) coordinates four O6 atoms and stays in the plane of a G-quartet,  $\text{K}^+$  and  $\text{NH}_4^+$  (ionic radii 1.33 and 1.48 Å, respectively) due to their size cannot be placed between guanines in the plane and are situated between two G-quartets. Other metals, such as  $\text{Rb}^+$ ,  $\text{Cs}^+$ ,  $\text{Sr}^{2+}$ ,  $\text{Ba}^{2+}$ ,  $\text{Mg}^{2+}$ ,  $\text{Ca}^{2+}$  and  $\text{Pb}^{2+}$ , were also found to stabilize G-quadruplexes. On the contrary, in the absence or low concentration of monovalent cations, ions of  $\text{Cd}^{2+}$ ,  $\text{Co}^{2+}$ ,  $\text{Mn}^{2+}$ ,  $\text{Zn}^{2+}$  and  $\text{Ni}^{2+}$  can destabilize G-quadruplexes favoring their dissociation.<sup>2</sup>  $\text{Li}^+$  has no effect on G4-folding.



*Figure 1. From guanines to G-quadruplexes. From left to right: guanine residue (blue and red asterisks: atoms involved in Hoogsteen and Watson Crick hydrogen bonding, respectively), G-quartet and G-quadruplex (guanine bases are green, adenine – red, thymine – blue).*

Depending on the strand orientation, G-quadruplexes can be classified as parallel (all strands are of the same direction), antiparallel (two strands have one direction and other two – the opposite) or hybrid (three strands have the same direction, one the opposite) (*Figure 2*). Generally, the stability of these structures decreases in a row parallel > antiparallel > hybrid,

in agreement with different conformation of glycosidic bonds that the bases in G-quartets adopt (*syn* or *anti*, the latter being more stable). While DNA can form G4s of all three topologies, RNA adopts only the parallel G-quadruplex conformation that can be explained by sterical defavorization of *syn* conformation of glycosidic bond angle in the RNA sugar.

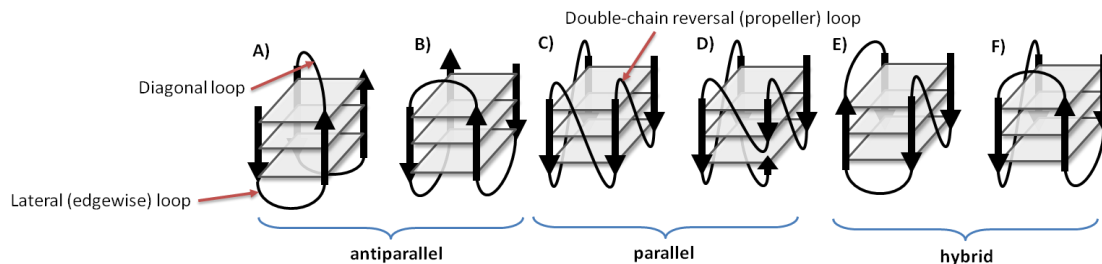


Figure 2. G-quadruplex topologies and loops: A) basket-type antiparallel; B) simple chair-type antiparallel; C) simple parallel; D) snap-back parallel; E,F) hybrid. Loops: lateral (or edgewise), diagonal and double-chain reversal (propeller).

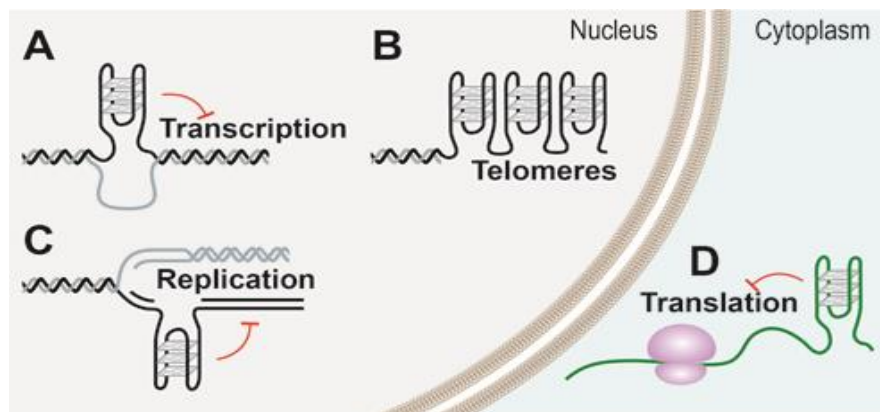
G-quadruplexes can be intramolecular (formed within one strand) and intermolecular (consist of two, three, or four strands). The majority of biologically relevant G-quadruplexes are intramolecular. Loops that connect G-tracks in G4s are, generally, of three kinds: lateral (or edgewise), double-chain reversal (or propeller) of 1-3 bases each and diagonal that usually contains 3-6 bases. There is also short V-loop, connecting guanines of different conformation. In the presence of  $K^+$  the stability of G-quadruplexes is inversely proportional to loop length, while in  $Na^+$  conditions this tendency is less prominent.<sup>3</sup>

A lot of efforts are put into the development of methods of theoretical prediction of G4-folding.<sup>4</sup> Different bioinformatics tools, such as Quadparser, Quadruplexes, QGRS Mapper, G4Hunter, AllQuads etc. were designed. These prediction methods started with a canonical sequence  $G_{3-5}N_{1-7}G_{3-5}N_{1-7}G_{3-5}N_{1-7}G_{3-5}$  for identification of G4-forming sequences.<sup>5</sup> However, they often give numbers of putative G4 sequences (PQS) inferior than those, obtained empirically. The calculation is complicated by the possibility of formation of multimeric G-quadruplexes or bulged structures, when a G-track is interrupted by another nucleotide. Taking all this into account new tools, like G4-iM Grinder<sup>6</sup> or QPARSE,<sup>7</sup> are emerging.

Although double-stranded DNA is very stable, some of its parts occur to be unwinded during various biological processes (replication etc.) and alternative secondary structures with Hoogsteen base-pairing, such as G-quadruplexes, can be formed within these single stranded parts. Their formation is facilitated or even can be induced by the environment. Formation of G4s, for example, is promoted by high potassium concentration in nucleus (145 mM), molecular crowding, negative supercoiling and specific G4-binding proteins.

## 1.2 Biological relevance of G-quadruplexes

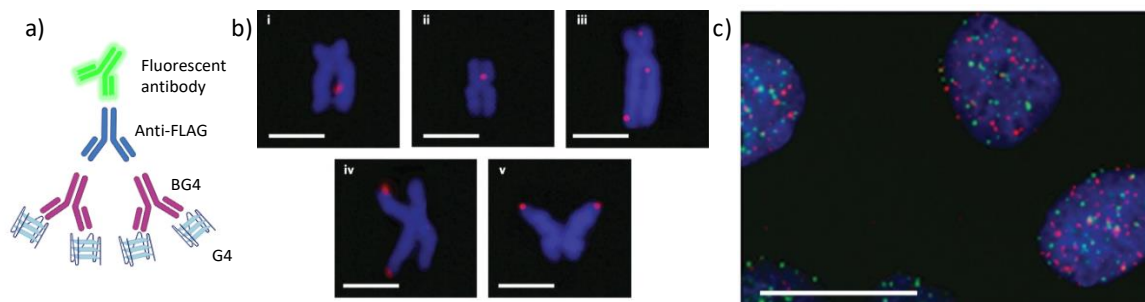
Over 700 000 G-quadruplex-forming sequences were detected by high-throughput sequencing in human genome.<sup>8</sup> Within the genome, the majority of G-quadruplex structures are formed in telomeres, in promoters of oncogenes and in 5'- and 3'- untranslated regions (UTRs) of messenger RNA (*Figure 3*).<sup>9</sup> G4 forming sequences are also found in viral species<sup>10</sup> or in bacteria,<sup>11</sup> not to mention other mammalian species.



*Figure 3. Possible locations of G-quadruplex structures in cells. (Adapted from <sup>9</sup>)*

The existence of G-quadruplexes *in vivo* was proved with the help of specific antibodies developed to these structures. For the first time the formation of G4 was detected in telomeric region of macronucei of the ciliate *Stylonychia lemnae* using single-chain antibody fragment (scFv) probes.<sup>12</sup> In human cells the G4-structures were detected using the G4-specific antibody

BG4, which is highly affine both to intramolecular ( $K_d = 0.5 - 1.6 \text{ nM}$ ) and intermolecular ( $K_d = 2 \text{ nM}$ ) G-quadruplexes.<sup>13</sup> In order to visualize G4s in human cells, fixed cells were incubated with BG4 and treated with secondary and then with tertiary fluorochrome-labelled antibodies (*Figure 4A*). Experiments with G4- and single strand-pretreated BG4 antibody, additional oligonucleotide transfection of cells and treatment of cells with DNase confirmed the specificity of BG4 to G-quadruplexes. Interestingly, the majority of G4s was detected outside the telomeres, contrary to what would be expected (*Figure 4B-C*). Authors explained this fact by low accessibility of telomeric G4s due to their binding to specific proteins (shelterins). The hypothesis, that the majority of detected G-quadruplexes are formed during transcription, was confirmed in experiments with populations within synchronized cell cycles. The highest density of G-quadruplexes was observed during S-phase, when the replication of DNA occurs. Pretreatment of cells with a G4-binding ligand (PDS) led to almost 3-fold increase in nuclear staining.



*Figure 4. Detection of G-quadruplex structures with BG4 antibody. A) Scheme of G4-imaging with BG4 (Created with BioRender.com); B) Discrete BG4 foci (red) were observed both within the non-telomeric regions i–iii and at the telomeres iv and v. Chromosomes are counterstained with DAPI (blue). Scale bars, 2.5  $\mu\text{m}$ ; C) Absence of large co-localization between telomeric TRF2 proteins (green) and G-quadruplexes (red) in U2OS cells. Nuclei are counterstained with DAPI (blue). Scale bar, 20  $\mu\text{m}$ . (adapted from <sup>13</sup>).*

### 1.3 G-quadruplexes at telomeres

Telomeres are extreme parts of DNA whose main function is to protect ends of chromosomes from fusion. They consist of repetitive nucleotide sequence  $(\text{TTAGGG})_n$  and comprise a

double-strand region of 2-15 kilobases and a 3'-single strand overhang of around 200 bases. Telomeres form loops by displacement of the part of ds-DNA (double stranded DNA) by 3'-overhang. Their integrity and functioning is regulated by shelterin, a protein complex of six subunits that allows the cells to distinguish the natural ends of chromosomes from the sites of DNA-damage. As telomeric sequences are guanine-rich, they can form G-quadruplexes that was demonstrated using specific antibodies. The stabilization of G4s has a negative impact on the replication and can cause telomere instability. Because of this reason, G4-unwinding helicases and single-strand binding proteins exist.

With every cell division telomeres become shorter due to the end-replication problem and when they bypass certain limit, cells stop to grow or the process of apoptosis is induced. However, in certain types of rapidly growing cells, such as germ cells, and some types of stem cells, telomeres are constantly elongated by the reverse transcriptase enzyme telomerase. Interestingly, telomerase is also activated in 85% of cancer cells. The stabilization of G4-quadruplexes in the 3'-overhang with small molecules inhibits telomerase activity by making an obstacle to enzyme binding and leads to senescence and apoptosis of cancer cells.<sup>14</sup> In the rest of cancer cells the telomere length is maintained by alternative lengthening of telomeres (ALT) pathway which also requires the 3'-overhang in the form of single strand. As both telomerase and ALT pathway of telomere elongation in cancer cells require terminal 8 or 12 bases respectively in the form of single strand,<sup>15</sup> stable complex of G4 with ligand prevents the telomere from being both a substrate for telomerase and a primer for ALT.

Telomeric G-quadruplexes are highly polymorphic. Sequences that contain one, two, three or four repeats form G-quadruplexes of different topologies depending on conditions. Thus, the same human telomeric sequence forms G-quadruplex of antiparallel topology in Na<sup>+</sup> conditions,<sup>16</sup> several hybrid structures in K<sup>+</sup> solutions<sup>17</sup> and adapts parallel topology upon molecular crowding<sup>18</sup> and in the crystal form.<sup>19</sup> By varying the flanking nucleotides or/and by modifying/changing bases, it is possible to favor the formation of one major conformation.

Several well-studied variants of the human telomeric sequence are summarized in *Table 1*. Among them, sequence **24TTG** is a variant of a human telomeric sequence that contains two modifications of terminal residues: A→T in second position and T→A in the 24<sup>th</sup> (*Table 1*). In K<sup>+</sup> solution it forms a G-quadruplex of major (about 95%) hybrid-1 (3+1) topology in which the

three strands run in one direction and one strand in the opposite and its three G-tetrads are *anti•syn•syn•syn* or *syn•anti•anti•anti* (Figure 5A).<sup>20</sup> Two base modifications stabilize the hybrid-1 structure by additional interactions of flanking bases with G-tetrads, namely, Watson-Crick base pairing T1•A20 and reversed Watson-Crick T13•A21 stacked on the tetrad core. This structure has one narrow, one wide and two medium grooves, one of which is occupied by double-chain-reversal loop.

Table 1. Telomeric G4 sequences used in this work.

Acronym	Sequence (5' → 3')	PDB	Structure
24TTG	d[TT GGG (TTA GGG) <sub>3</sub> A]	2GKU	[G4 hybrid-1] in K <sup>+</sup> (solution)
25TAG	d[TA GGG (TTA GGG) <sub>3</sub> TT]	2JSL	[G4 hybrid-2] in K <sup>+</sup> (solution)
22CTA	d[ A GGG (CTA GGG) <sub>3</sub> ]	2KM3	G4 antiparallel in K <sup>+</sup> (solution)

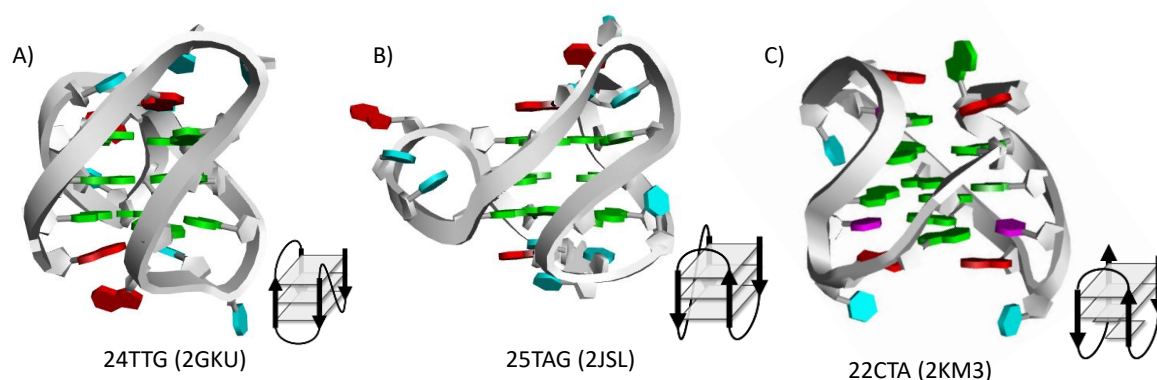


Figure 5. Structures and schematic views of telomeric G-quadruplexes: A) 24TTG G4 of Hybrid-1 topology in K<sup>+</sup> solution (replotted from PDB 2GKU);<sup>20</sup> B) 25TAG sequence of Hybrid-2 topology in K<sup>+</sup> solution (replotted from PDB 2JSL);<sup>17</sup> C) antiparallel 22CTA (replotted from PDB 2KM3).<sup>21</sup> Base colors: G green, A red, T blue.

The **25TAG** sequence (Table 1) adopts major (3+1) hybrid-2 structure in K<sup>+</sup>-rich solution (Figure 5B).<sup>17</sup> Similarly to 24TTG, it has three G-tetrad core and contains two lateral loops that span wide and narrow grooves and double-chain-reversal loop in one of two medium grooves. It was suggested that the residues T1, T12 and T14 interact with possible coordination of K<sup>+</sup>-ion.

As for additional interactions, on the bottom of G-quadruplex the residues T7, A8 and T24 interact with the formation of a (T7–A8)•T24 triad, below which T25 is stacked.

Sequence **22CTA**: The (CTAGGG) repeat is associated with high mutation rate in male germline and somatic cells.<sup>21</sup> It is suggested that about 7% of all human telomeres contain (CTAGGG)<sub>n</sub> repeats. Sequence 22CTA (Table 1) adapts a chair-type antiparallel conformation in the presence of K<sup>+</sup> ions (Figure 5C).<sup>22</sup> It contains two G-quartets with glycosidic conformations *syn•anti•syn•anti* and has two narrow and two wide grooves. All three loops are lateral; the first and the third loops span narrow grooves, while the second spans a wide groove. Watson-Crick pairing is observed for G8•C17 and G20•C5 on the top and G14•C11 at the bottom of G-tetrad core. In addition, two GC pairs on the top (G8•C17 and G20•C5) form an unusual G•C•G•C quartet over which adenines A7 and A19 are further stacked.

#### 1.4 G-quadruplexes as transcription regulators

According to different computational studies, above 40% of human genes contain PQSs in proximity to their promoters, especially near the transcription starting sites (TSS).<sup>23</sup> Interestingly, PQS are particularly represented in oncogenes and regulatory genes. The role of G4 in the regulation of transcription is ambiguous: they can act both as suppressors and as activators of gene expression. A very prominent example of the ambiguous role of G4 in transcription regulation is its role in *c-myc* promoter. On the one hand, G-quadruplex acts as an activator of *c-myc* expression: G4 is a binding site of some proteins, such as CNBP25 and NM23-H that activate the gene transcription. On the other hand, overrepresentation of G4 structures (achieved by overexpression of nucleolin, a G4-binding protein) can inhibit transcription (Figure 6).<sup>24</sup> The repression of *c-Myc* expression occurs both in the case of inducing of G4 formation by treatment with G4-binding ligand TMPyP4<sup>25</sup> and in the case of blocking of G4 folding with locked nucleic acid complementary strand.<sup>26</sup>

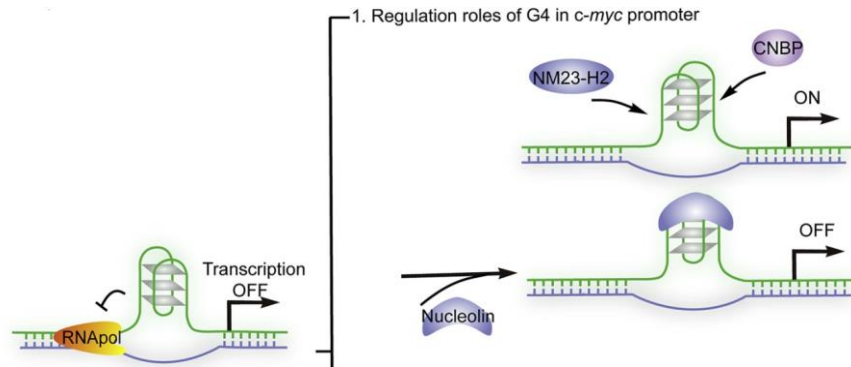


Figure 6. Regulation role of G-quadruplex in *c-myc* promoter (adapted from <sup>24</sup>).

#### 1.4.1 G-quadruplexes from the MYC promoter: Pu24T and *c-myc*

*c-Myc* protein is overexpressed in around 80% of cancers and is involved in multiple hallmarks of cancer.<sup>27</sup> The upregulation of *c-Myc* gene is usually connected with the nuclease hypersensitivity element NHE III<sub>1</sub>, a region of 27 nucleobases that is situated 142 to 115 base pairs upstream from the P1 promoter and controls 85-90% of *c-Myc* transcription. The abundance of cytosines in the coding strand and guanines in the template strand of NHE III<sub>1</sub> implies a possible formation of non-canonical structures in this region (i-motifs and G-quadruplexes, respectively)<sup>25</sup> upon different conditions. This guanine-rich region is also called *Pu27* (from purine-rich sequence of 27 nucleotides) and contains six guanine stretches (five of which are of three or more guanines) that can assemble in numerous G-quadruplex structures.<sup>28</sup> These G-quadruplexes are believed to exist in equilibrium with the double-strand and contribute to regulation of *c-Myc* expression in different extent.<sup>29</sup> In our studies, we decided to work with two truncated sequences of *Pu27* in which three guanines of first G-track were removed. They contain mutations that favor formation of G-quadruplexes of definite topologies (Table 2): *Pu24T* (snap-back parallel) and *myc22* (simple parallel).

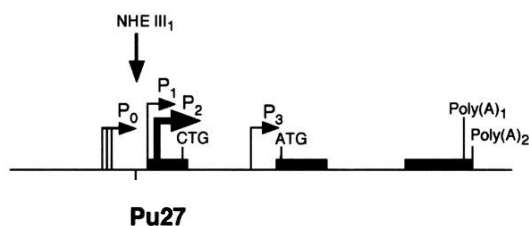


Figure 7. Location of NHE III1 element in *c-Myc*-gene.



Table 2. Sequences from *c-Myc* gene promoter.

Acron ym	Sequence (5' → 3')	PDB	Structure
	1 2 3      4 5 6 7      8 9 10 11 12      13 14 15 16      17 18 19 20 21 22      23 24		
<i>Pu27</i>	TGGGGA GGGT GGGGA GGGT GGGGAA GG		polymorphic
<i>Pu24T</i>	TGA GGGT GG <b>T</b> GA GGGT GGGGAA GG	2A5P	G4 s4nap-back parallel
<i>Myc22</i>	TGA GGGT GGG <b>T</b> A GGGT GGG <b>T</b> AA	1XAV	G4 simple parallel

The ***Myc22*** sequence is a *Pu27* variant that besides truncation of the first G-track contains two G→T modifications in positions 14 and 23 (*Table 2, Figure 8A*).<sup>30</sup> In K<sup>+</sup> solution it forms a G-quadruplex structure of three G-tetrads in which four parallel strands are connected with three double-chain-reversal loops. The first and third loops are of one nucleotide and showed to be very stable, second loop is of two nucleotides. All four grooves of the G-quadruplex are of the same width. Glycosidic conformation of all core guanines is *anti*. The 3'-flanking TAA region forms a stable fold-back conformation stacking with bottom tetrad: T23 is stacked with G22, A25 is stacked with G9 and G22, and A24 is stacked with T23 residue. Interestingly, T23 and A25 are paired not through Watson-Crick, but Hoogsteen type of hydrogen bonding. In the 5'-flanking region (TGA) the position of A6 is fixed by stacking with the top tetrad. Capping of G-core by 5'- and 3'-flanking regions increases the stability of the G-quadruplex.

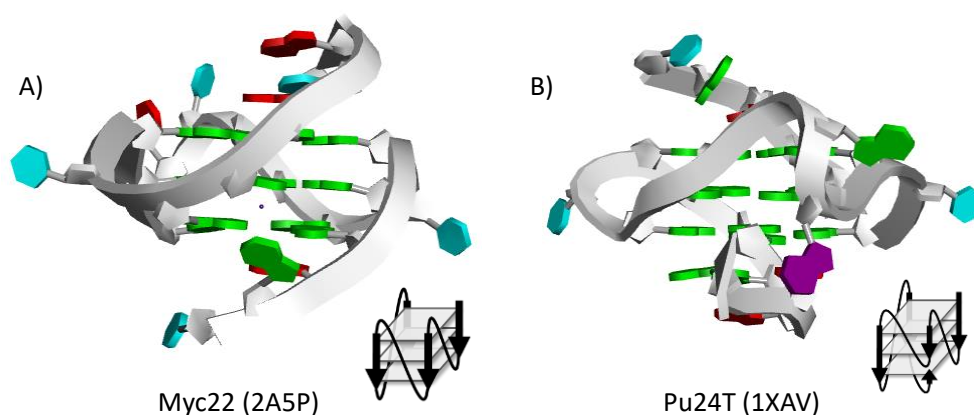


Figure 8. PDB structures and schematic views of G-quadruplexes from *c-Myc* promotor: A) simple parallel *myc22* (replotted from BDB 1XAV);<sup>30</sup> B) snap-back parallel *Pu24T* (replotted from PDB 2A5P).<sup>31</sup>

In the ***Pu24T*** sequence the third guanine of second G-track is replaced by thymine (*Table 2, Figure 8B*). The structure of this G-quadruplex was solved by NMR and in order to obtain a clearer spectrum T10 was replaced with inosine (*Pu24T*).<sup>31</sup> G-quadruplex structure, formed by this sequence in K<sup>+</sup> solution, is also parallel; it contains three tetrads, one of which contains the foldback guanine G24. The core contains three G-tetrads that are formed by five G-tracks and four loops: three double-chain-reversal and one diagonal. The first and third of double-chain-reversal loops are of one nucleotide (thymine) and bridge three tetrad layers, while the second one contains three nucleotides (GGA or IGA) and bridges two layers. The fourth, diagonal loop is of four nucleotides (GAAG); it connects the opposite bases of the bottom G-tetrad. The glycosidic configuration of G24 is *syn* and all other guanines of G-core have *anti* conformation. The bases in the diagonal loop are fixed by formation of G20●(A22-G23) triad, on the bottom of which A21 is stacked. On the top of the core pair A3●A12 is stacked. It was suggested that the stability of this triad leads to such particular structure of this G-quadruplex. The structure is additionally stabilized by triad G10●(G11-A12) that caps G-core on the top. When two last guanines are removed from the sequence, the formation of a simple parallel G-quadruplex is observed (*myc22*).

#### ***1.4.2 G-quadruplex from RAS promoter***

Ras proteins are low-weight GTPase proteins that are involved in cellular signal transduction. They exist in non-active state (affine to GDP) or active state (affine to GTP). In the active (or “switched on”) state they indirectly activate the genes involved in cell-growth, differentiation and survival. Single-point mutations that can convert these proteins to uncontrolled activated state were found in many types of cancer (about 20% of human cancers, in 90% of pancreas tumor).<sup>32</sup>

The family of RAS proteins contains HRAS, NRAS and KRAS proteins, the latter being essential. Due to their extremely high affinity to GTP/GDP (in the picomolar range) targeting of these proteins is not a trivial task and only recently the first covalent KRAS inhibitor reached clinical trials.<sup>33</sup> So important attention is now being paid to alternatives strategies, including the

regulation of the protein expression. KRAS promoter contains purine-rich nucleic acid hypersensitive element that plays an important role in transcription regulation. Sequence 32R (Table 3) in NHE is extremely rich in guanines and is able to form G-quadruplexes of different conformations.<sup>34</sup> Sequence 21R (Table 3) was shown to form a G-quadruplex of a single conformation. A single 16G→T mutation and an additional 3'-terminal adenine (22RT) led to additional stabilization of the structure. A G-quadruplex formed by the 22RT sequence contains three G-tetrads and three double-chain reversal loops (Figure 9). Two loops are of one nucleotide; one contains four nucleotides and completely covers one of the grooves. All four grooves are approximately of the same size ( $12 \pm 2$  Å). Glycosidic conformations of all guanines involved in G-tetrads are *anti* suggesting parallel topology of the sequence. Adenine residues at the extremities cap the G-core. One thymine bulge is found between G8 and G9.<sup>35</sup>

Table 3. KRAS sequences.

Acronym	Sequence (5' → 3')																Topology														
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16		17	18	19	20	21	22								
32R	G	G	A	G	G	G	G	G	A	G	A	G	G	G	C	G	G	T	G	T	G	G	G	A	G	A	G	G	G	A	
21R						A	G	G	G	C	G	G	T	G	T	G	G	G	A	A	G	A	G	G	G	A					
22RT						A	G	G	G	C	G	G	T	G	T	G	G	G	A	A	T	A	G	G	G	A	A				

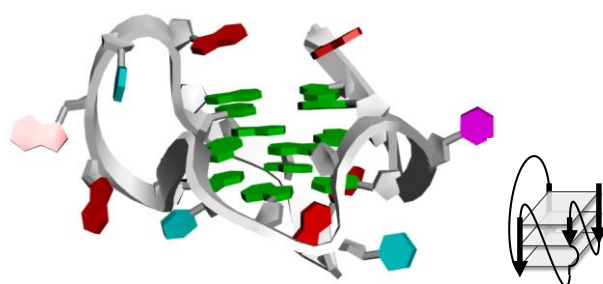


Figure 9. PDB structure and schematic views of G-quadruplexes from RAS promoter KRAS-22RT (replotted from PDB 5I2V).<sup>35</sup>

## 1.5 G-quadruplexes as translation regulators: the case of viral genome maintenance protein EBNA1

G-quadruplex motifs were found in abundance in 5'-untranslated regions (5' UTR of mRNA).<sup>36</sup> The formation of G-quadruplexes in 5'-UTR impedes the initiation of translation so helicases, such as DHX36, are employed to resolve them.<sup>37</sup>

As a matter of example, G-quadruplexes play a very important role in regulation of translation of EBNA1 protein associated with Epstein-Barr virus (EBV), a herpesvirus that affects more than 95% of human population worldwide. Like other herpesviruses, EBV established a latent lifecycle in the nuclei of hosts' cells and has a genome as circular genetic elements not incorporated in host DNA. Under certain conditions that remain unclear it can cause some human cancers such as Burkitt's lymphoma, nasopharyngeal carcinoma etc.<sup>38</sup> EBNA1 protein (Epstein-Barr nuclear antigen 1) is a genome maintenance protein that is required for replication and mitotic segregation of the viral genome during host cell division and is a hallmark of all EBV-associated cancers. This protein contains a central glycine-alanine (GAR) domain, two arginine-glycine (RG) domains (LR1 and LR2), and a C-terminal DNA-binding domain (DBD) (*Figure 10A*). Interestingly, LR1 and LR2 domains of EBNA1 are known to interact with G4 RNA structures (including those formed in its own mRNA). EBNA1 is highly antigenic; however, the host's immune system fails to detect it due to the fine mechanism of regulation of EBNA1-expression evolved by the EBV (called immune evasion) where a crucial role is played by its central GAR domain.

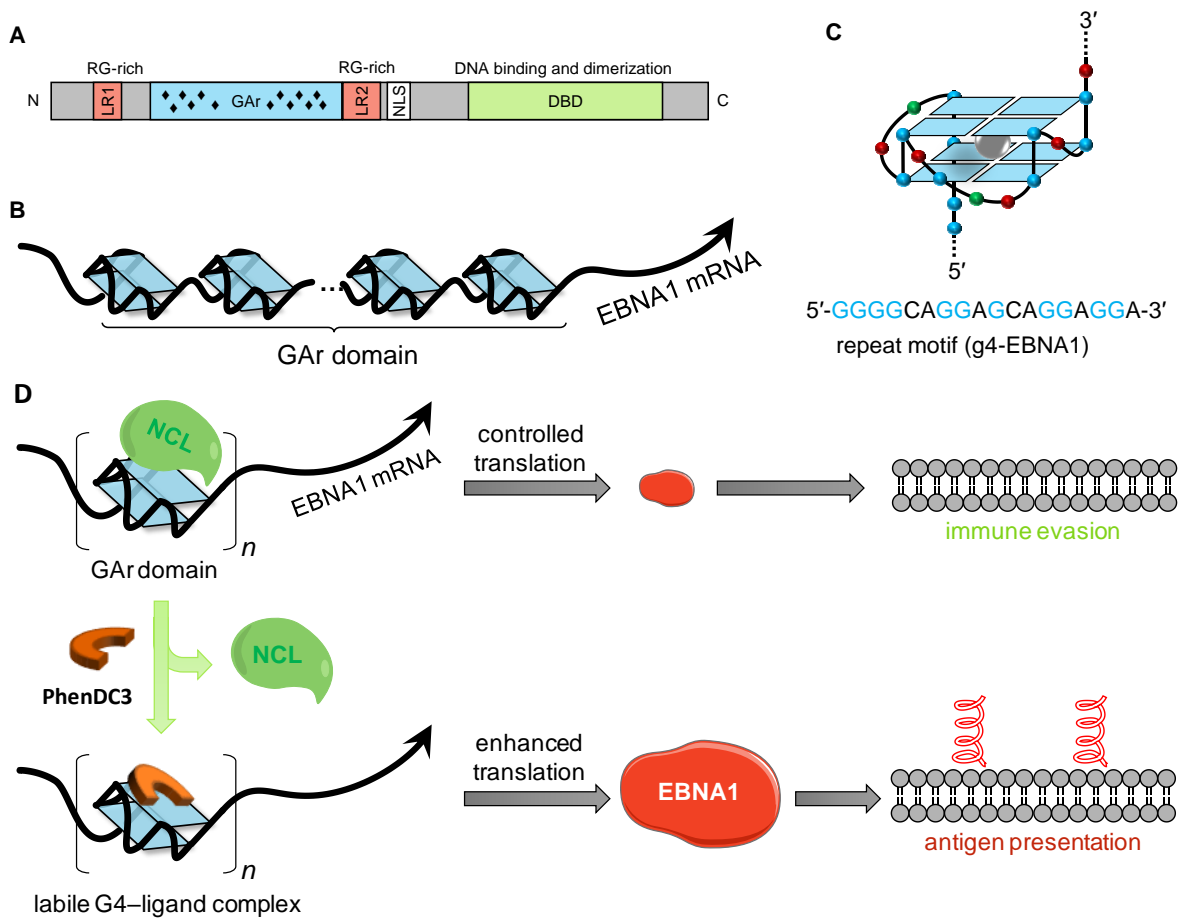


Figure 10. A) EBNA1 sequence showing the central Gly-Ala repeat (GAR) domain surrounded by the N- and C-terminal domains; B) GAR domain of EBNA1 mRNA; C) The proposed model of g4-EBNA1; D) Schematic representation of the mechanism of immune evasion of EBV and of the effect of nucleolin-displacing G4 ligand **PhenDC3** (adapted from <sup>39</sup>).

One of hypothesis explaining the mechanism of immune evasion of EBV includes the regulation of translation of EBNA1 mRNA by G-quadruplexes. As mRNA EBNA1 GAR-coding sequence is of high GC content, the formation of G-quadruplexes was suggested in this region (Figure 10B). Host protein nucleolin binds to G-quadruplexes and inhibits EBNA1 translation through an unknown mechanism. It was demonstrated that the treatment of EBV-infected cells with the benchmark G-quadruplex ligand **PhenDC3** leads to disruption of the complex of EBNA1 mRNA G4 with nucleolin and results in enhanced translation of EBNA1 mRNA (Figure 10D).<sup>40</sup>

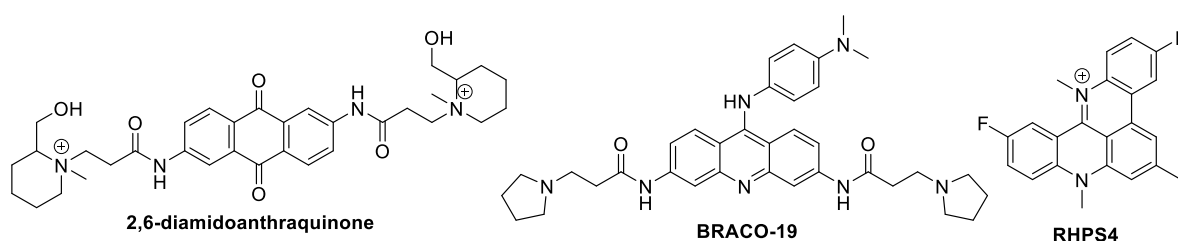
In this work we used the 18-nucleotide sequence 5'-r(GGGGCAGGAGCAGGAGGA)-3' **g4-EBNA1** that is capable to form a G-quadruplex structure and that occurs 13 times through the

GAr-coding sequence.<sup>41</sup> In vitro it adapts simple parallel conformation, although in vivo formation of polymorphic structures with longer loops is possible. G4-EBNA1 presumably contains two G-tetrads (*Figure 10C*) even though the exact structure of this G-quadruplex was not reported yet.

## 2. G4-ligands

Small organic molecules (or metal complexes) that bind to G4 structures with high affinity are termed G4 ligands. Nowadays, hundreds of compounds are reported to be G-quadruplex ligands, as summarized in several reviews.<sup>42,43,44</sup> Their structures in most cases contain an aromatic (often, polyheterocyclic) core, that interacts with external G-quartets by  $\pi$ - $\pi$  stacking, and cationic branches that directly or indirectly interact with loops and grooves. Nevertheless, some of these molecules also intercalate into double-stranded DNA (dsDNA), so the selectivity towards dsDNA is a very important feature for potent G4-ligands. I will briefly introduce the chemical families as well as the most referenced compounds in the field of G-quadruplex ligands with proven biological activity.

The first synthetically developed G-quadruplex ligand, 2,6-diamidoanthraquinone (*Chart 1*) was reported in 1997 as a telomerase inhibitor.<sup>45</sup> Since then, a large number of G4-ligands were described in the same context. Among structurally related acridone and acridine derivatives **RHPS4**<sup>46</sup> and **BRACO-19**<sup>47</sup> (*Chart 1*) stepped out as efficient telomerase inhibitors.



*Chart 1. Structures of 2,6-diamidoanthraquinone, BRACO-19 and RHPS4.*

**RHPS4** not only inhibits telomerase at submicromolar concentrations ( $IC_{50}$  in the TRAP assay of 330 nM), but was also shown to act synergistically with the currently used anticancer drug paclitaxel (Taxol) in MCF-7 breast cancer cells.<sup>48</sup> Despite its low cellular uptake, **BRACO-19** inhibits cancer cells proliferation<sup>49</sup> and has an anti-HIV-1 activity.<sup>50</sup> It binds to *c-myc* sequence with a  $K_d$  of 19 nM (by SPR).<sup>51</sup> The  $K_d$  value for the telomeric sequence was measured in three different studies and determined as 29 nM (SPR),<sup>51</sup> 310 nM (SPR),<sup>52</sup> and 288 nM (BLI).<sup>53</sup> Such a discrepancy (ten fold) in values reveals an interesting observation questioning the reliability of affinity measurements for ligand–G4 interaction. It is generally accepted that dissociation constants determined by different methods may vary but should at least be of the same order of magnitude. In this case, the measurements were performed with different systems of telomeric G4 sequence: the first value was obtained in the experiment with intramolecular hybrid G4 (that changed its topology to antiparallel upon binding of ligand). Two other values were obtained using intermolecular antiparallel G4 assembled on a cyclic peptidic construct. Such a difference in  $K_d$  values can be explained by different affinity of ligand to different topologies and point out that **BRACO-19** is more affine to hybrid than antiparallel G-quadruplex. The crystal structure of the complex of **BRACO-19** with telomeric bimolecular G-quadruplex  $d(TAGGGTTAGGGT)_2$  was also reported (*Figure 11*).<sup>54</sup> The reported biological unit contained two parallel G-quadruplexes stacked via **BRACO-19** molecule. Interestingly, the drug was stacked to 3'-G-tetrad of one G-quadruplex and T●A●T●A base pair capping the 5'-end external tetrad of another G4.  $\pi$ - $\pi$  stacking of **BRACO-19** was reported for two guanines of G-tetrad and T●A of another tetrad. The positively charged ring nitrogen was found to form H-bond with the water molecule that was positioned in line with the  $K^+$  ions coordinated between G-quartets. Besides, one direct H-bond was observed between thymine T24 and side-chain amide as well as two water mediated H-bonds between the same thymine and carbonyl group of another side chain and N9 of acridine of the drug. The positively charged side chains are protruding into the grooves due to electrostatic interactions.

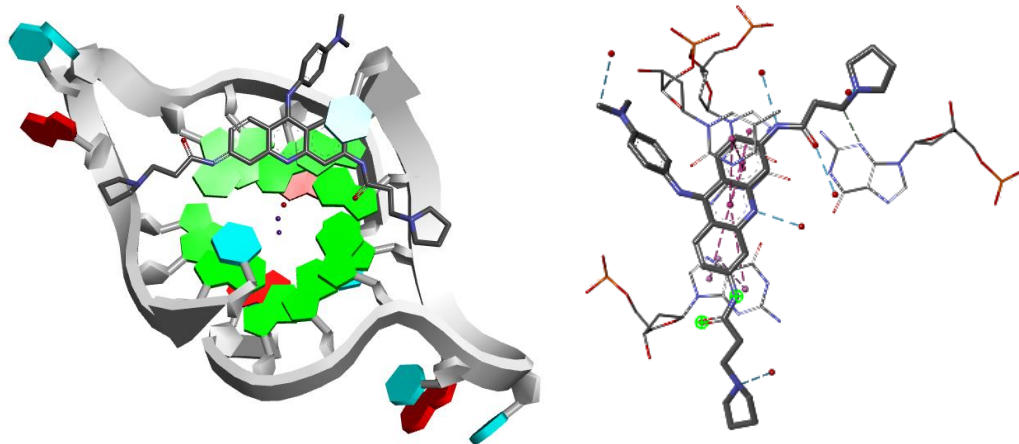


Figure 11. Crystal structure of telomeric bimolecular G-quadruplex with BRACO-19 and ligand-interaction view (replotted from PDB 3CE5).<sup>54</sup>

Quaternary bis-quinolinium dicarboxamide derivatives of pyridine (**PDC** or **360A**) and phenantroline (**PhenDC3**, Chart 2) became benchmark G4 ligands due their high affinity to both promoter and telomeric G4s and good selectivity with respect to duplexes.

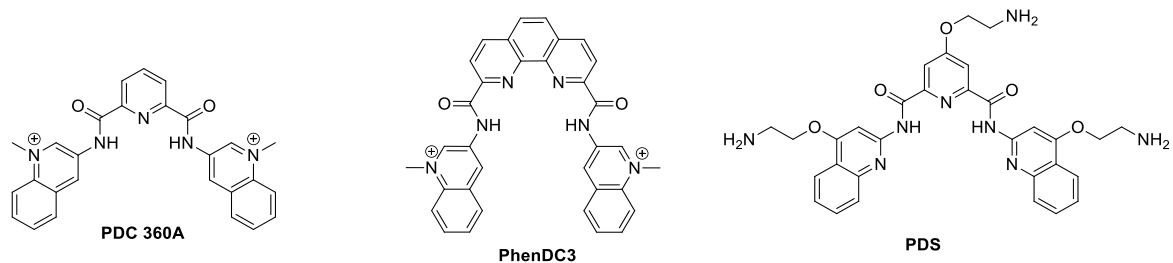


Chart 2. Structures of **PDC (360A)**, **PhenDC3** and **PDS**.

Pyridine-2,6-dicarboxamides inhibit proliferation and induce apoptosis of telomerase-positive and ALT tumor cells by inducing end fusion of chromosomes.<sup>55</sup> Among this series (Figure 12), the tritiated (<sup>3</sup>H) derivative of **PDC** (or **360A**) was found (by radioactive analysis) localized in extremes of chromosomes.<sup>56</sup> With the help of PDC-based photo-cross-linking probes (PDC-XL) it is now possible to selectively alkylate DNA G-quadruplexes;<sup>57</sup> the Pt(II)-NHC derivative of **PDC** perturb telomere integrity and causes significant (up to 50%) delocalization of one of shelterin's proteins, TRF2.<sup>58</sup> **PDC** scaffold was also used for the development of fluorescent probes. For example, the conjugate with thiazole orange **PDC-M-TO** selectively binds G-



quadruplexes over dsDNA with fluorescent enhancement for telomeric G4 almost two-fold bigger than for promoter.<sup>59</sup> In this context, we recently reported a series of fluorescent PDC-coumarin conjugates, among which two derivatives (**PDC-L2-C2** and **PDC-L3-C2**) discriminate between G4 topologies.<sup>60</sup> When the probe binds to antiparallel or hybrid G-quadruplex, the coumarin part is flipped out and fluorescent enhancement is observed, while no fluorescent gain is observed upon their binding to parallel G-quadruplex (*Scheme 1*).

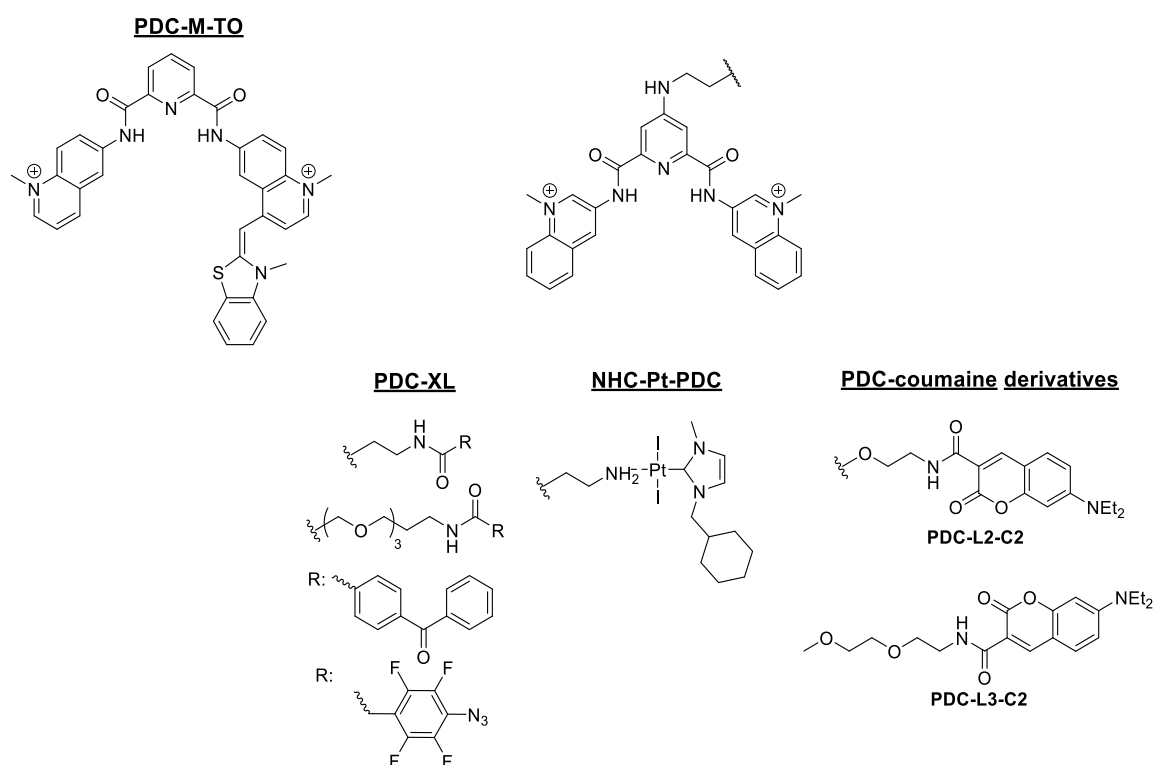
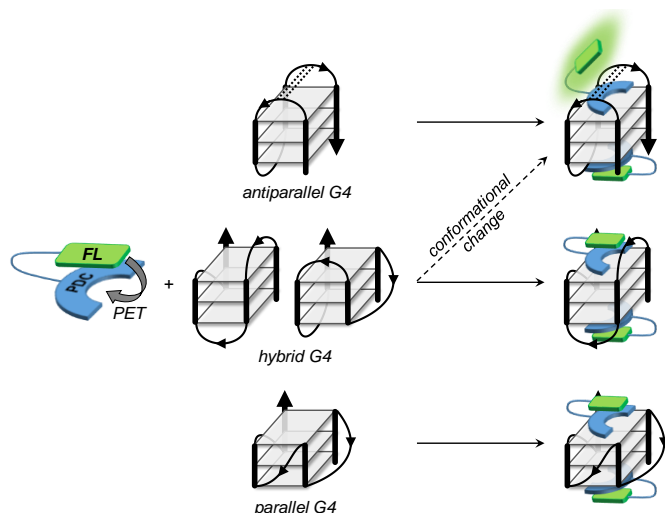


Figure 12. Structures of functional derivatives of PDC and fluorescent probes.



*Scheme 1. Proposed model of fluorimetric response of PDC-coumarin conjugates towards various G4-DNA structures. The dashed lines depict the base–base interactions defining the antiparallel G4 topology.<sup>60</sup>*

Phenantroline dicarboxidiamide **PhenDC3**<sup>61</sup> represents probably the best compromise between affinity and selectivity towards G-quadruplexes. It binds to assembled antiparallel telomeric G4 with  $K_d$  down to 14 nM measured by BLI and 4.4 nM by SPR.<sup>53</sup> It binds to G-quadruplex by  $\pi$ - $\pi$  stacking, as showed by an NMR structure of a complex with *Pu24T* (Figure 13A)<sup>62</sup> and even though it disrupts capping of the top G-tetrad by the A3•A12 base pair, it significantly stabilizes the structure. Concerning its action on cancer cells, **PhenDC3** causes genome instability and inhibition of telomere elongation.<sup>63,64</sup> Similar compound **pyridostatin** (or **PDS**, *Chart 2*)<sup>65</sup> differs from **PDC** by the three branches (positively charged in water) while quinoline nitrogen atoms are not quaternized. Its  $K_d$  to the antiparallel G-quadruplex was estimated as 12, 17 or 96 nM.<sup>53,66</sup> **PDS** uncaps the protein POT1 from the shelterin complex inducing DNA-damage response.<sup>65</sup> A biotinylated analogue of **PDS** was used for the first small-molecule-mediated isolation of G-quadruplexes from human cells.<sup>67</sup> **PDC**, **PhenDC3** and **PDS** are quite easy to synthesize and now are commercially available.

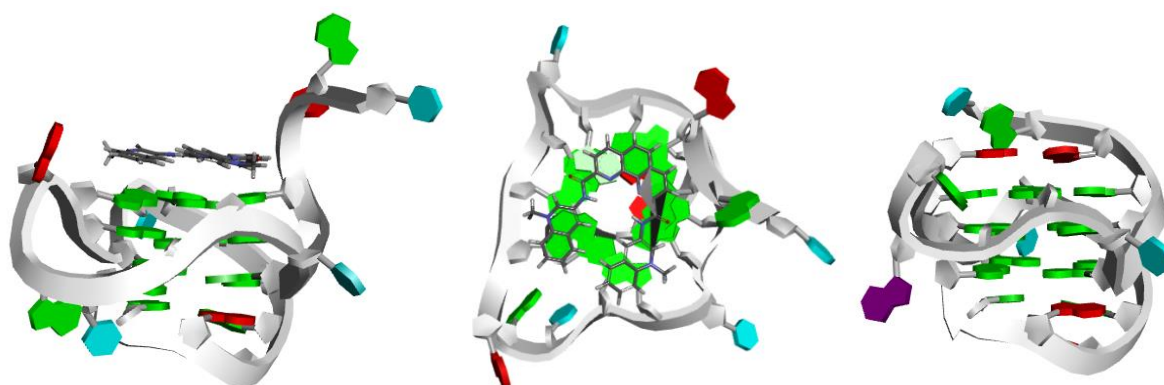


Figure 13. A) Structure of complex of **PhenDC3** with Pu24T (replotted from PDB 2MGN),<sup>62</sup> B) Unbound Pu24I (replotted from PDB 2A5P)<sup>31</sup> for comparison.

The derivatives of naphthyridine-1,8-dicarboxamide, **3AQN** and **6AQN** (Chart 3), were found to have high G4 binding affinity and selectivity to G4 over double-stranded DNA.<sup>68</sup> Besides high thermal stabilization of telomeric and promoter G-quadruplexes in FRET-melting and high efficiency of these ligand in the FID assay, they were found to be very efficient in the polymerase stop assay with telomeric G-rich primer sequence and even outperformed **360A** (IC<sub>50</sub> of 0.8 μM, 0.7 μM and 1.2 μM for **3AQN**, **6AQN** and **360A**, respectively). These molecules, similar to **360A**, caused inhibition of growth of human malaria parasite (with IC<sub>50</sub> of 1.8, 1.5 and 0.9 μM for **3AQN**, **6AQN** and **360A**, respectively for 3D7 strain) but had low cytotoxicity to human cells.<sup>69</sup>

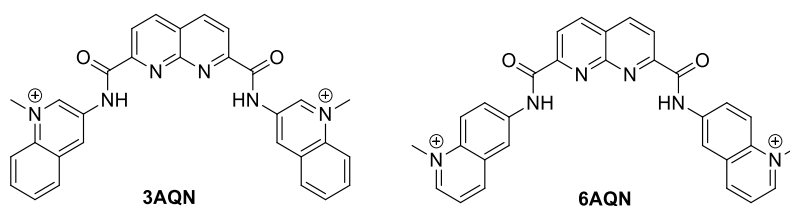


Chart 3. Structures of naphthyridine derivatives **3AQN** and **6AQN**.

Natural compounds and their derivatives were also shown to bind G-quadruplexes,<sup>70</sup> e.g. groove binders **steroid FG**<sup>71</sup> and oligopeptides such as **distamycin A** (Chart 4).<sup>72</sup> According to molecular docking and molecular dynamics study, **steroid FG** mostly interacts with hydrophobic groove/loop regions<sup>73</sup>. **Distamycin A** binds to G-quadruplexes with ligand to DNA stoichiometry of 4:1 in two grooves, by forming antiparallel **distamycin A** pairs.

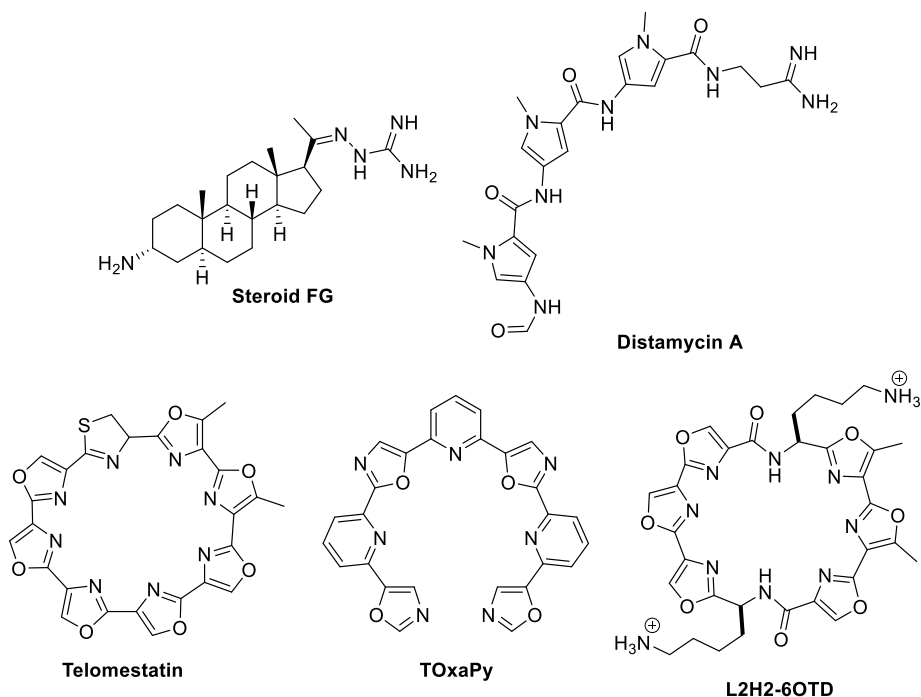
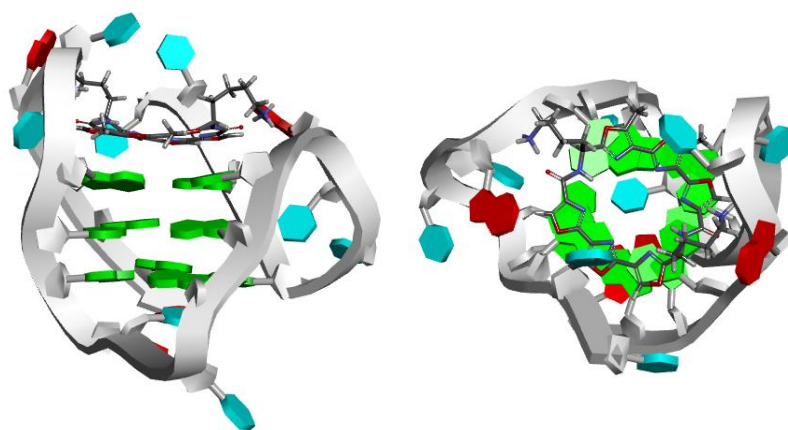


Chart 4. Structures of natural compounds and their derivatives as G4-binders.

The biggest attention among natural compounds was paid to the macrocyclic peptide **telomestatin** (Chart 4),<sup>74,75</sup> which was isolated from actinobacteria species *Streptomyces anulatus*. This compound contains five oxazole, two methyloxazole and one thiazoline ring, and binds to G4 by stacking on external tetrads. It is a very efficient telomerase inhibitor ( $IC_{50}$  in the nanomolar range).<sup>64</sup> Interestingly, unlike the majority of G4-ligands, **telomestatin** is uncharged. It is cytotoxic for many cancer cell lines at concentrations below 5  $\mu$ M and has no effect on normal cells. The synthesis of **telomestatin** is one of the biggest obstacles for the use of this compound: its total synthesis is rather complicated and includes low-yield last-step purification;<sup>76</sup> at the same time, its extraction from bacteria is not cost-effective due to the presence of large amount of other metabolites. A method of heterologous expression of this compound was reported recently.<sup>77</sup> Due to difficult accessibility of this ligand, multiple synthetic analogues of **telomestatin** were developed. For example, interesting properties were reported for the acyclic heteroarylic compound **TOxaPy** (Chart 4) which binds to the antiparallel but not hybrid conformations of human telomeric sequence.<sup>78</sup> Cyclic telomestatin analogue **L2H2-6OTD** (Chart 4), which contains six oxazole rings and two alkyl amine chains binds to telomeric G-quadruplexes by  $\pi$ -stacking and additional electrostatic interaction of side chains with phosphate groups. It inhibits telomerase activity of Namalwa cells ( $IC_{50} = 20$

nM) and proliferation of HeLa cells ( $IC_{50} = 7.4 \mu\text{M}$ ).<sup>79</sup> The structure of the complex of **L2H2-6OTD** with telomeric sequence 24TTG of hybrid topology was resolved by NMR (*Figure 14*).<sup>80</sup>



*Figure 14. Structure of the complex of L2H2-6OTD with telomeric 24TTG hybrid G4 (replotted from PDB 2MB3).*<sup>80</sup>

Porphyrins represent another important class of G4-binders. Along with heme, that binds to telomeric G4 with  $K_d$  of 25 nM,<sup>81</sup> the most studied and widely used compound of this family is **TMPyP4** (*Figure 15*).<sup>82</sup> Cationic branches of **TMPyP4** increase the solubility and, in the same time, affinity to nucleic acids. Upon light activation **TMPyP4** can generate reactive oxygen species that allows its use in photodynamic therapy.<sup>83</sup> Thanks to its photosensitizing properties it has found various applications, from cancer to antimicrobial therapies.<sup>84</sup> Curiously, **TMPyP4** is considered as DNA G4-stabilizer<sup>82</sup> and RNA G4-unfolder.<sup>85</sup> It binds to G4s by stacking on external G-tetrad (*Figure 15*).<sup>31</sup> Although it binds not only to quadruplexes (e.g.,  $K_d$  to antiparallel G4 of 247 nM),<sup>53</sup> but also to duplexes ( $K_d$  to a hairpin of 110 nM),<sup>53</sup> a large number of fundamental studies related to biological roles of G-quadruplexes were performed with this compound.<sup>25,29</sup>

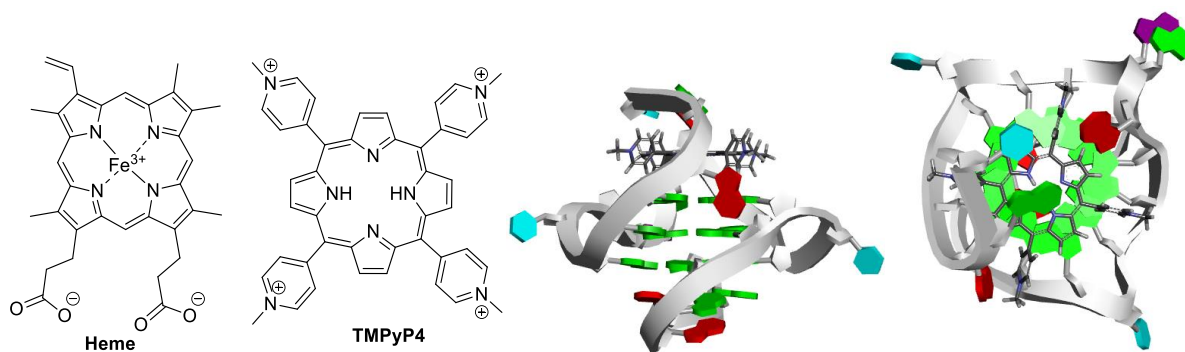


Figure 15. Structure of heme, **TMPyP4** and the complex of **TmPyP4** with parallel G4 Pu24I (replotted from PDB 2A5R).<sup>31</sup>

*N*-Methyl mesoporphyrin IX (or **NMM**) was first reported to bind G-rich DNA sequences in 1996.<sup>86</sup> Since then this molecules became extensively used in the G4 field due to its fluorescent properties,<sup>87,88</sup> good G4 selectivity over double-stranded DNA, and much higher affinity to parallel over antiparallel and mixed (hybrid) G-quadruplex topologies.<sup>89</sup> The latter can be well illustrated by the results of SPR experiments in which the dissociation constants of **NMM** with G4s of different topologies were measured.<sup>90</sup> Five different parallel sequences formed with **NMM** stable complexes with  $K_d$  values varying from 62 to 330 nM. At the same time, hybrid G4s formed complexes with  $K_d$  of micromolar range (1.5-7.2  $\mu$ M) and for **NMM** complexes with antiparallel G4s  $K_d$  was higher than 17  $\mu$ M. In a reported X-ray structure of the complex of **NMM** with telomeric sequence dAGGG(TTAGGG)<sub>3</sub> (*Tel22*) the drug is not completely planar (*Figure 16*).<sup>91</sup> The biological unit of the reported structure contained two parallel G-quadruplexes stacked with their 5'-5' tetrads. Bridging K<sup>+</sup> cation is positioned between them. One **NMM** molecule is stacked on each of the 3'-external tetrads in this dimer. The binding of **NMM** to G-tetrad is stabilized by efficient  $\pi$ - $\pi$  stacking. The *N*-Me group of **NMM** is positioned in the central channel upon K<sup>+</sup> cations most likely because of sterical reason.

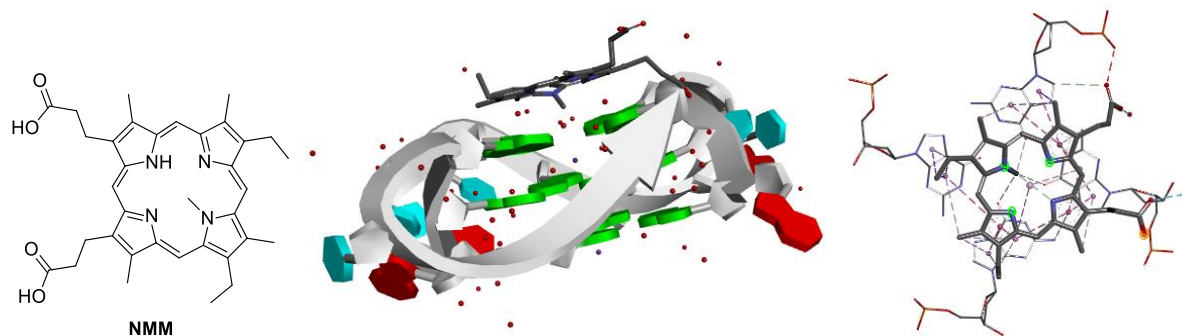


Figure 16. Structure NMM and its complex with telomeric G4 Tel22 (replotted from PDB 4FXM).<sup>91</sup>

To date, only two G-quadruplex ligands reached clinical trials. The first one, **quarfloxin (CX-3543 or intrafloxin)**, was developed by Cylene pharmaceuticals as a result of optimization of compound **QQ58** presented by the Hurley group (*Chart 5*).<sup>92</sup> It entered Phase II of clinical trials in early 2008 ([ClinicalTrials.gov](http://ClinicalTrials.gov) identifier: NCT00780663) as a treatment of carcinoid and neuroendocrine tumors but failed due to bioavailability issues. The optimized derivative **CX-5461** (*Chart 5*) was originally described as ribosomal RNA synthesis inhibitor in cancer cells.<sup>93</sup> Later it was found that CX-5461 is a G-quadruplex stabilizer.<sup>94</sup> It is currently in Phase I clinical studies as an RNA polymerase I inhibitor for treatment of many tumors, particularly breast cancer deficient in BRCA1 and/or BRCA2 ([ClinicalTrials.gov](http://ClinicalTrials.gov) identifier: NCT02719977).

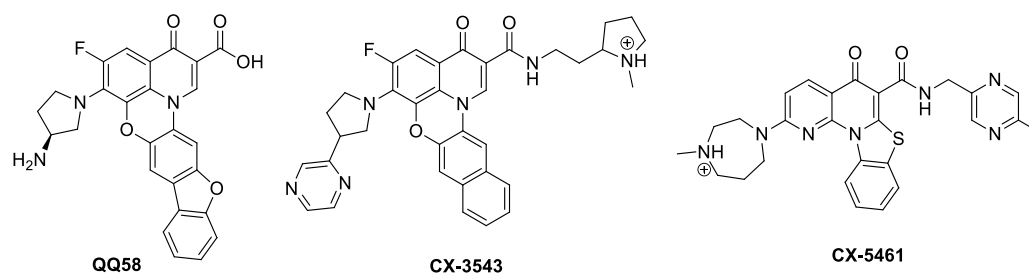
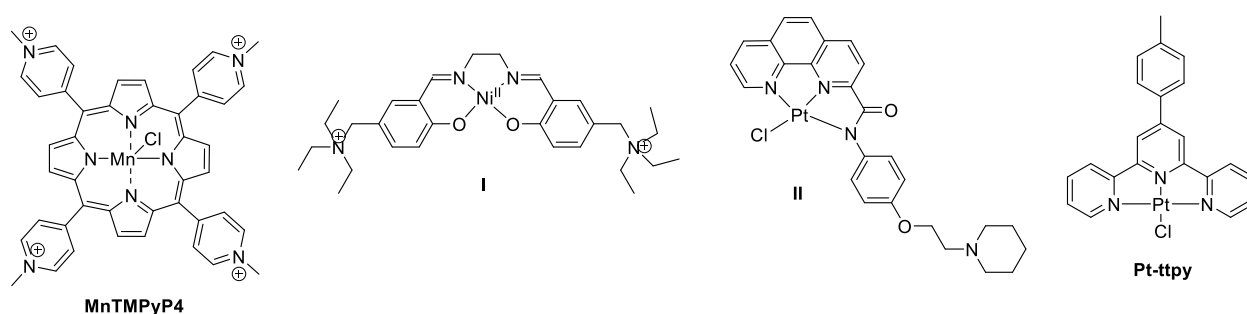


Chart 5. Structures **QQ58** and clinical candidates **CX-3543** and **CX-5461**.

Metal complexes constitute another large family of G-quadruplex ligands. They are rather easy to synthesize and their magnetic, optical and catalytical properties open a wide area for applications. Metal cations not only make  $\pi$ - $\pi$  stacking of the ligand more efficient by diminishing electron density of the aromatic core, but also provide additional coordination bonds with guanines of G-tetrads if placed in the central canal. The development of this class

of ligands started with porphyrine derivatives of Cu(II),<sup>95</sup> Ni(II), Mn(III),<sup>96</sup> and then expanded to other metals, such Zn(II)<sup>97</sup> or Fe(II).<sup>98</sup> Complexation of metal also leads to increased selectivity of ligand with respect to *dsDNA* (up to 1000 fold as for **Mn(III) TmPyP4** derivative, *Chart 6*). Ni(II) salphene derivative **I** (*Chart 6*) showed antiproliferative effect on breast cancer cells.<sup>99</sup> Pt(II) complex of phenantroline **II** (*Chart 6*) inhibits telomerase<sup>100</sup> and **Pt-ttpy** (*Chart 6*) derivative was shown to inhibit the proliferation of cancer cells that are resistant to cisplatin<sup>101</sup> that makes them attractive for cancer therapy.



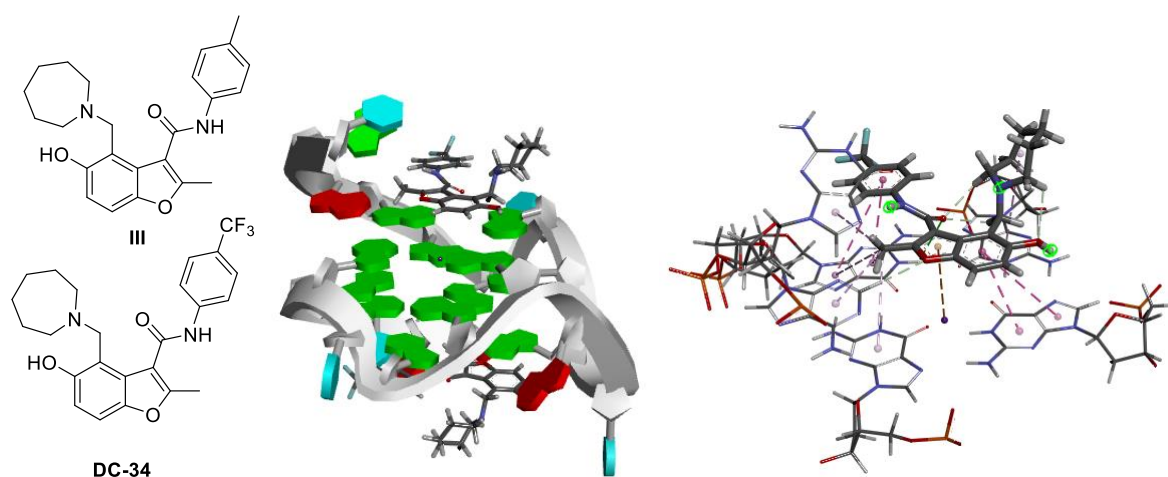
*Chart 6. Structure of G4-binding metal complexes.*

One of major challenges in the search of G-quadruplex ligands is to find compounds that would be able to discriminate between different G4 structures or topologies. To date, there are several families of compounds that selectively bind to G4s of parallel topology (e.g., **NMM**, indole and selenium-containing cyanine-based dyes, distyrylpyridinium dyes etc.).<sup>44</sup> Most of ligands that bind to telomeric hybrid G-quadruplexes induce the change their conformation to antiparallel. What concerns the specific recognition of antiparallel topology, there is still a room for improvement.

In the search of selective G4 ligands, a small-molecule microarray screening of 20 000 compounds was performed using a fluorescently labeled oligonucleotide.<sup>102</sup> As a result, compound **III** was identified as a selective binder of *c-myc* G4. Although thermal stabilization of 2.1 °C (CD-melting, 1 eq. of ligand, *c* = 25 μM) and a *K<sub>d</sub>* value of 4.5 μM (as determined by SPR) are not convincing enough when comparing with the benchmark G4 ligands (for comparison, the value of *K<sub>d</sub>* of **BRACO-19** binding to *c-myc* is 19 nM), analogs of compound **III** were developed. The trifluorinated compound **DC-34** has better affinity to G4 (*K<sub>d</sub>* = 1.4 μM to *Pu27* sequence as per SPR) and represent a rare example of a drug-like G4-ligand that can



discriminate between promoter G-quadruplexes with preferential binding to *c-myc*.<sup>103</sup> The structure of the complex of *c-myc* G4 with two molecules of **DC-34** was solved by NMR (*Figure 17*). It indicates the stabilization of the complex not only through  $\pi$ - $\pi$  stacking, but also by hydrogen bonds between NH<sub>2</sub>-group of A25 and benzofuran's oxygen on 3'-end and NH<sub>2</sub>-groups of G7 and G18 and trifluoromethyl group on 5'-end.



*Figure 17. Structure of the 2:1 complex of DC-34 with c-myc sequence (replotted from PDB 5W77).*<sup>103</sup>

### 3. Combinatorial chemistry approaches to discover G4 ligands

Combinatorial chemistry refers to the synthesis of libraries of small-molecules or peptides in relatively simple reactions, often performed in automated manner. Peptide coupling and click chemistry represent the most significant breakthroughs in the field. In combination with high-throughput screening, it is a very powerful tool in current drug discovery. High-throughput screening is a technique that permits the simultaneous screening of thousands to millions compounds towards various biological targets. Until recently, it was a prerogative of industry considering the high cost of equipment. Nowadays the situation is different: not only many universities have their own screening centers where researchers can develop and perform

their screening assays, the screening can be done even in the researchers' laboratories. Technologies such as microarrays, DNA-encoded libraries and phage display allow the latter. In this part, I will give some examples of combinatorial chemistry applied to the search of G-quadruplex ligands, including classical combinatorial synthesis, as well as more recent developments such as DNA-encoded libraries, phage display and dynamic combinatorial chemistry.

### 3.1. Classical combinatorial chemistry approaches

The screening of synthesized (targeted) libraries of compounds is a well established way of the search for G-quadruplex ligands. Some examples of this approach are given below.

In an attempt to find analogues of the previously reported bis(amidinohydrazone) **105.1** (identified by a HTS campaign)<sup>104</sup> that was capable of high thermal stabilization of G4s but possessed low biological activity, Randazzo et al.<sup>105</sup> synthesized a library of ten amidinohydrazones (*Figure 18A*). CD-melting experiments, performed with G4s of different topologies as well as with dsDNA, showed that compounds **105.6-105.8** were not selective to G4s and had some affinity to dsDNA, and compounds **105.9** and **105.10** showed no significant stabilization of any target. Instead, compounds **105.1-104.3** and **105.5** stabilized G4s and not dsDNA, among which compound **105.3** was the most selective towards parallel *c-myc* G4 as demonstrated by  $\Delta T_m$  values ( $\Delta T_m > 20$  °C for *c-myc*,  $\approx 10$  °C for parallel *c-kit* and telomeric G4s and  $\approx 0$  °C for hybrid telomeric G4 and dsDNA). The authors suggested that two equivalents of **105.3** could bind to *c-myc* sequence and that this interaction was stabilized not only by  $\pi$ - $\pi$  stacking but also by hydrogen bonds between the positively charged amino groups with phosphate backbone and between formyl group of the ligand with amines of the bases A6 and T23 that cap the structure (as suggested by docking, *Figure 18B*). In biological assays compound **105.3** showed significantly higher cytotoxicity for cancerous cells than **105.1** (IC<sub>50</sub> for U2OS cells after 24 h of 0.845  $\mu$ M and 15.95  $\mu$ M, respectively).

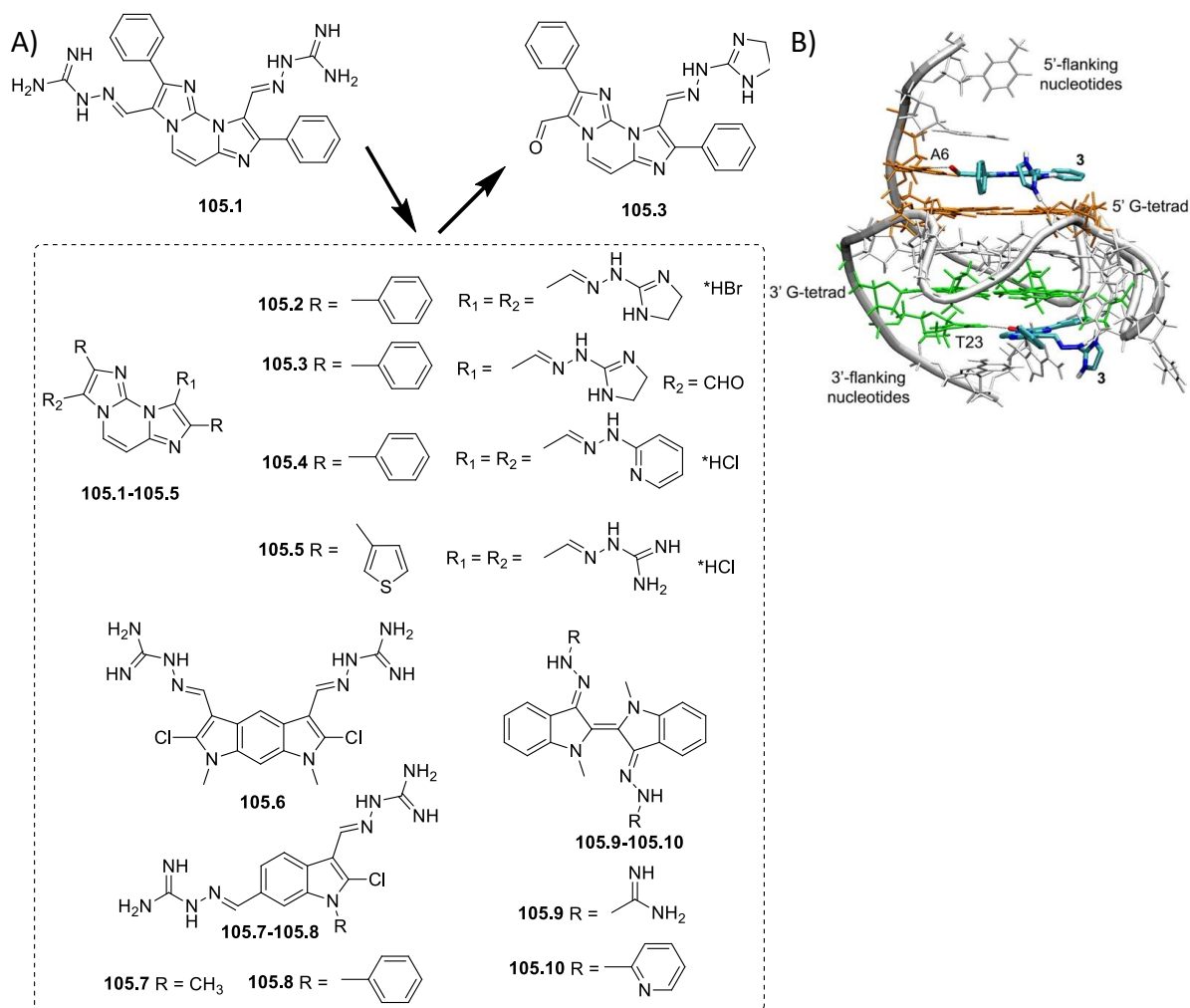
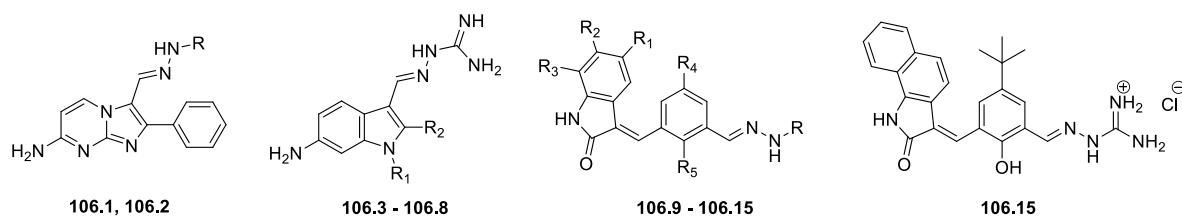


Figure 18. Structure of compounds **105.1**, **105.3** and the combinatorial library; B) structure of the complex of **105.3** with *c-myc* G-quadruplex (suggested by docking).<sup>105</sup>

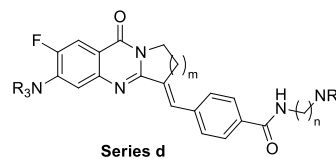
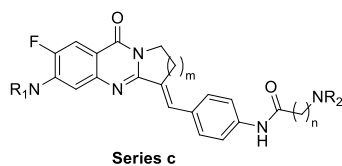
Another amidinohydrazone derivative **106.15** was reported as a promising G4-ligand by the same group as a result of the screening of a library of 15 compounds by means of CD-melting towards telomeric and promoter G-quadruplexes (Figure 19).<sup>106</sup> Compound **106.15** is cytotoxic to cancer U2OS and HeLa cell lines at micromolar concentrations. In CD-melting experiments it selectively stabilizes the telomeric sequence ( $\Delta T_m$  of 13 °C), although in immunofluorescence microscopy experiment with BG4 antibody neither **106.15** nor **BRACO-19** significantly increased the number of telomeric G4 foci. Instead, compound **106.15** was shown to induce DNA damage as seen from the increased quantity of hallmarks of DNA double-strand breaks (53BP1 and  $\gamma$ -H2AX foci) and by formation of micronuclei.



Compd	R	R <sub>1</sub>	R <sub>2</sub>	R <sub>3</sub>	R <sub>4</sub>	R <sub>5</sub>	Compd	R	R <sub>1</sub>	R <sub>2</sub>	R <sub>3</sub>	R <sub>4</sub>	R <sub>5</sub>
106.1		-	-	-	-	-	106.9		H	CBR	H	H	
106.2		-	-	-	-	-	106.10		H	CBR	H	H	
106.3	-	H	Cl	OCH <sub>3</sub>	CH <sub>3</sub>	-	106.11		Cl	H	H	H	H
106.4	-	CH <sub>3</sub>	Cl	OCH <sub>3</sub>	CH <sub>3</sub>	-	106.12		Cl	H	H	H	H
106.5		-	Cl	H	H	-	106.13		OH	CH <sub>3</sub>	H	H	H
106.6	-	H	H	OCH <sub>3</sub>	CH <sub>3</sub>	-	106.14		OH	CH <sub>3</sub>	H	H	H
106.7		-	H	H	H	-	106.15		H	CBR		OH	
106.8		-	H	H	H	-							

Figure 19. Composition of the library of amidinohydrazone derivatives of three heterocyclic scaffolds and the structure of the lead compound **106.15**.<sup>106</sup>

In the search of *c*-Myc transcriptional inhibitors, Huang et al. designed two libraries of isaindigotone derivatives: series **c** (of 19 compounds) and series **d** (of 24 compounds) (Figure 20).<sup>107</sup> First, the ELISA assay was performed with all compounds of series to estimate their capacity to disrupt the complex between *c*-myc G4 and transcriptional factor NM23-H2. Compounds **107.19d** and **107.22d** showed the strongest inhibitory activities. They also provided the highest thermal stabilization of *c*-myc G4 in a FRET-melting assay ( $\Delta T_m$  of 12.1 and 12.9 °C, respectively) and showed rather weak binding to the NM23-H2 protein in MST experiment ( $K_d$  of 17.1  $\mu$ M for  $K_d$  **107.19d** and not determined value for **107.22d**). Both compounds reduced level of *c*-myc mRNA and *c*-Myc protein in dose-dependent level in SiHa cells.



Compd	R <sub>1</sub>	R <sub>2</sub>	R <sub>3</sub>	R <sub>4</sub>	R <sub>3</sub>	R <sub>4</sub>
107.1c m = 1, n = 1						
107.5c m = 1, n = 2						
107.9c m = 2, n = 1						
107.13c m = 2, n = 2						
107.2c m = 1, n = 1						
107.6c m = 1, n = 2						
107.10c m = 2, n = 1						
107.14c m = 2, n = 2						
107.3c m = 1, n = 1						
107.7c m = 1, n = 2						
107.11c m = 2, n = 1						
107.15c m = 2, n = 2						
107.4c m = 1, n = 1						
107.8c m = 1, n = 2						
107.12c m = 2, n = 1						
107.16c m = 2, n = 2						
107.17c m = 1, n = 2						
107.18c m = 1, n = 2						
107.19c m = 1, n = 2						
107.17d m = 1, n = 2						
107.4d m = 1, n = 2						
107.8d m = 1, n = 3						
107.12d m = 2, n = 2						
107.16d m = 2, n = 3						
107.17d m = 1, n = 2						
107.18d m = 1, n = 2						
107.22d m = 1, n = 3						
107.23d m = 1, n = 3						
107.19d m = 1, n = 3						
107.20d m = 1, n = 3						
107.21d m = 1, n = 3						
107.24d m = 1, n = 3						

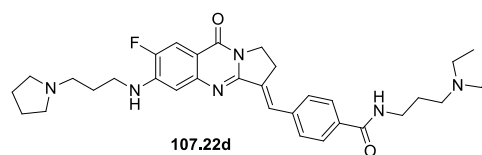
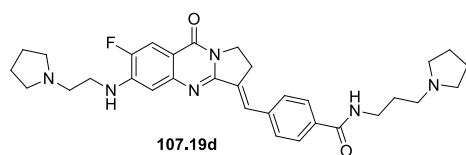
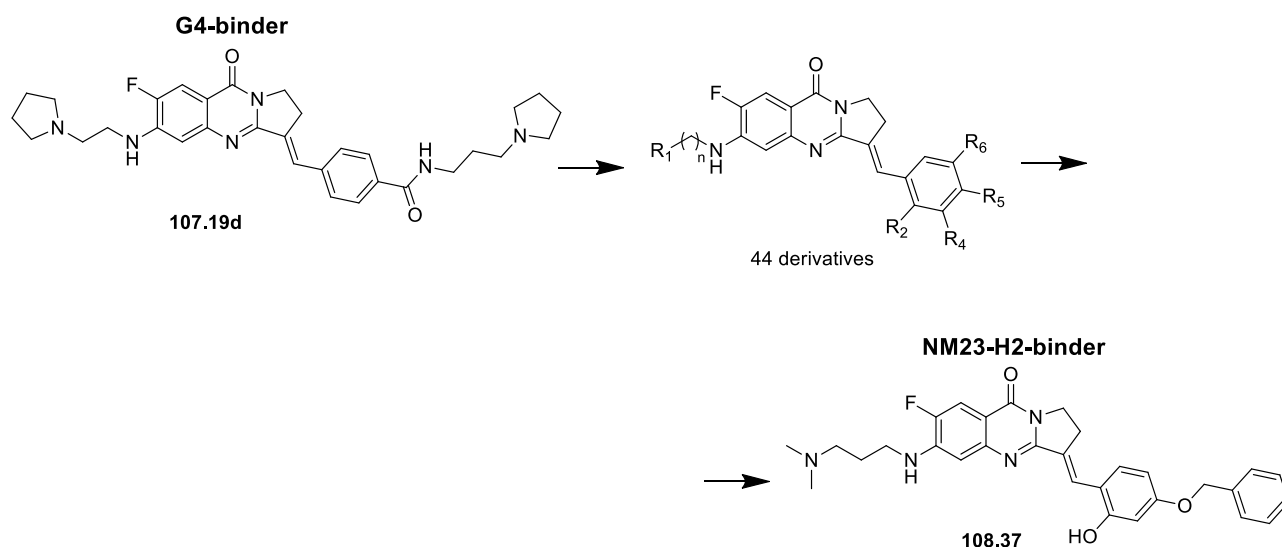


Figure 20. Structure of isaindigotone derivatives series c and d and the best binders, **107.19d** and **107.22d**.<sup>107</sup>

Interestingly, along with stabilization of G-quadruplexes, inhibition of transcription factors can also be used in order to inhibit c-Myc protein expression. The protein NM23-H2 among its other functions is an oncogene transcription factor and was found to inhibit transcription of c-Myc by binding and unfolding G-quadruplexes. The authors supposed, that a potential inhibitor of NM23-H2 can be designed starting from a molecule that binds G-quadruplex, since NM23-H2 and G4 should have complementary shapes of their binding sites in the complex, that is, a molecule that binds to G4 can have moderate affinity to NM23-H2 as well. So the compound **107.19d** initially discovered as a G4 binder was used as a prototype for a protein-targeting of c-Myc transcriptional inhibitor. The library of 44 compounds was designed on the base of **107.19d** and screened for the affinity to NM23-H2 oncogene transcription factor (Scheme 2). The compounds' affinities towards the protein were measured by MST and SPR and **108.37** was selected as the most affine binder ( $K_d$  of 3.1  $\mu\text{M}$  by SPR and 2.0  $\mu\text{M}$  by

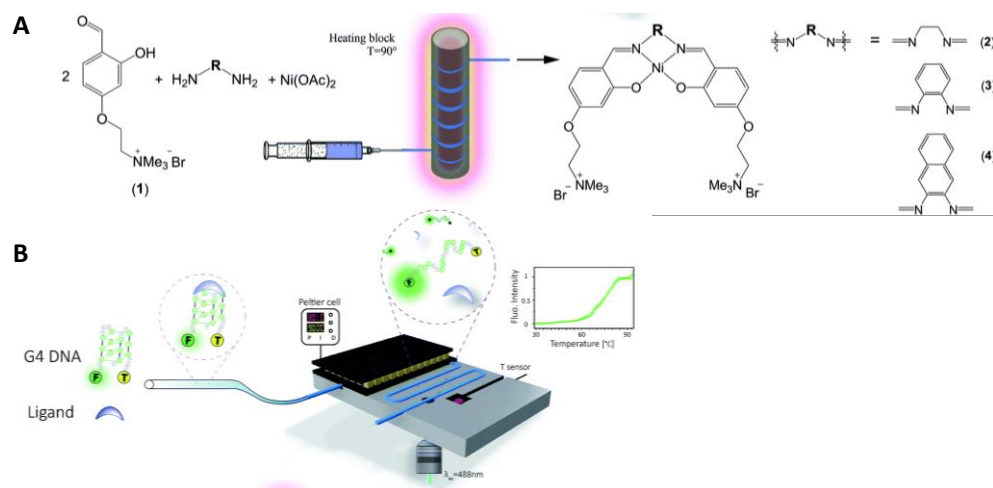
MST). It inhibits transcription of *c-myc* protein in dose-dependent manner and exhibited antitumor activity in mice with a SiHa xenograft.<sup>108</sup>



*Scheme 2. The discovery of NM23-H2 binder 108.37.*<sup>108</sup>

Microfluidics is the science and technology of systems that process or manipulate small ( $10^{-9}$  to  $10^{-18}$  litres) amounts of fluids, using channels with dimensions of tens to hundreds of micrometres.<sup>109</sup> Two major features of microfluidics synthesis that explain its increasing popularity recently are the effective mixing and heat/mass transfer and precise control of every parameter of the reaction.<sup>110</sup> Coupling of on-chip synthesis with on-chip screening represents a very promising route for combinatorial chemistry. Such an example exists also in G4-field where the synthesis of G4 ligands in a microfluidic setup is combined with biophysical evaluation of these compounds in FRET-melting experiment on microfluidic platform.<sup>111</sup> Three nickel(II)-salphen complexes (similar to ligand I, page 35) were prepared in-flow by injecting starting materials into thermostable polytetrafluoroethylene (PTFE) tubing coiled around a thermoblock and heated at 90 °C (*Scheme 3A*). The reaction was monitored by  $^1\text{H}$  NMR and was completed after 9 hours. The system allows the usage of much smaller quantities of reagents as it is performed in small volumes (350  $\mu\text{L}$ ). Interestingly, the G4 affinity of the “as-synthesized” ligands can also be monitored in a microfluidics-based DNA FRET-melting platform (*Scheme 3B*) and revealed to be identical to that of ligands synthesized in a “classical”

preparative way. To summarize, the microfluidics is a very promising method that can be applied for the synthesis and screening of compounds in combinatorial chemistry.



Scheme 3. A) Synthesis of nickel(II)-salphens in microfluidics platform; B) online FRET-melting assay (adapted from <sup>111</sup>).

### 3.2. DNA-encoded libraries (DECL)

The technique of DNA-encoded libraries was developed in the beginning of 1990s and now represents a rapidly evolving technique of drug discovery. It permits to analyze billions of compounds in one test tube, where each library component is linked to a unique barcoding DNA-sequence (tag). The library is mixed with the target and the most affine partners are separated from unbound library members by a pull-down technique. Structures of the best binders are then revealed by amplification and sequencing of DNA-tags. Reactions of synthesis of libraries should be performed in water (sometimes with co-solvents, such as DMSO, MeCN, DMA) with highly-diluted DNA-coupled starting materials and should preserve the genetic information that limits the scope of potential chemical scaffolds.

To date, only one example of the use of DNA-encoded libraries for the search of G-quadruplex ligands was reported (by X-Chem Pharmaceuticals).<sup>112</sup> Although in this study 33 libraries of a total 120 billion compounds were used, the report concerned only two libraries obtained by split-and-pool DNA encoding (Figure 21). Library A was a three-cycle library. The tagged

bromoacetyl derivative was split to 1024 wells and the unique tag was added to each derivative after a reaction of  $S_N2$  nucleophilic substitution with primary amines (step 1, S1). Then all products were mixed (pooled) and split into 85 wells of different formyl carboxylic acids for the reaction of reductive amination (S2). After tagging, all products were pooled and split for the last time to 2500 amines for acylation reaction with consecutive unique tagging (S3). The final pool of library **A** contained 217.6 million of tagged compounds. The synthesis of library **B** started with reaction of DNA-bound 2-chloro-3-nitro-pyridine with 1 051 secondary amines (S1). Tagged derivatives of  $S_NAr$  substitution were pooled and their nitro-group was reduced. The resulting amines were split to 1 213 wells containing aldehydes where reaction of oxidative cyclization took place (S2). The pool of tagged products contained 1 127 863 individual compounds. Two rounds of selection were performed with each library by addition of the library to immobilized G4 from *c-myc* promoter with consequent elution of bound compounds. Tags of selected compounds were then amplified, their structures were decoded, filtered by criteria such as logP etc., and one compound from every library was re-synthesized without the encoding tag. In a reverse SPR experiment (biotin-tagged compound + untagged G4) compound **112.2** showed affinity of  $K_d$  328 nM to *Pu27* G-quadruplex and even higher affinity with  $K_d$  59 nM to *c-myc* G-quadruplex containing only first four G-tracts. Compound **112.1** was shown to be less affine to these G4s ( $K_d$  values of 1.06  $\mu$ M and 328 nM, respectively). Though in cell culture experiments the compound did not show extraordinary properties, the published results concerned only compounds from two of 33 analyzed libraries, demonstrating the first example of the usage of DECL beyond protein targeting.



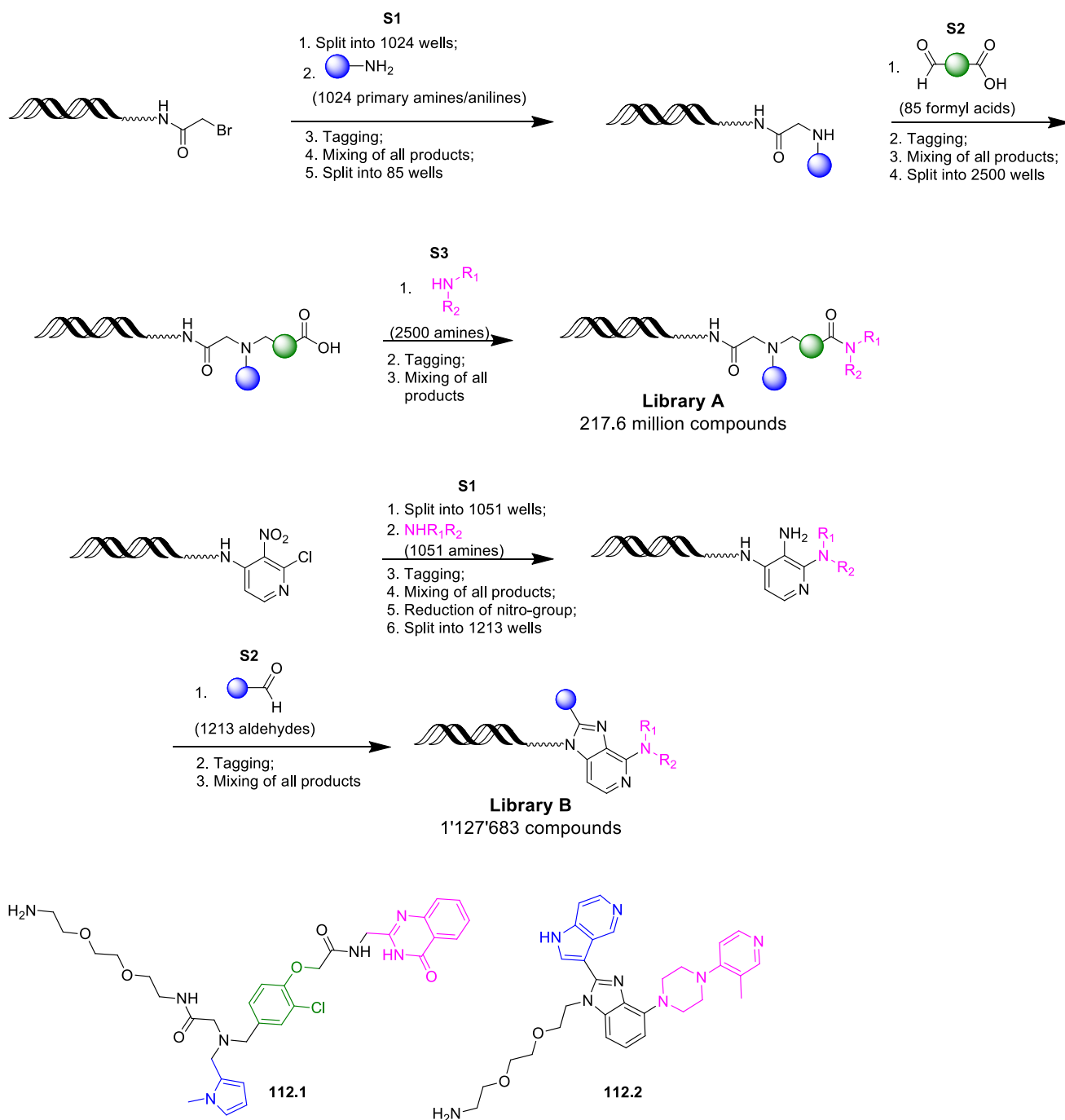


Figure 21. Synthesis of DNA-encoded libraries A and B and their best *c-myc* G4 binders **112.1** and **112.2**.<sup>112</sup>

### 3.3. Phage display

Phage display was discovered in 1980s by George P. Smith and Greg Winter who were awarded a half of the 2018 Nobel Prize in Chemistry for its development. In this assay, the

“ligands” are the peptides expressed on the surface of a phage, and the encoding sequence is incorporated into phage DNA which is situated inside the capsid, preventing any potential tag–target interactions. The phages are incubated with the target and the best binders are separated by pull down. The phages, eluted from the complex are then amplified in *E. coli* cells and the process of selection can be repeated. In this regard, phage display may be considered as a variant of DNA-encoded libraries. The genes from the phage are then sequenced and the searched protein structure is elucidated.

Recently, phage display of bicyclic peptides was applied to the search of G-quadruplex binders.<sup>113</sup> Bicyclic peptides are obtained from linear amino acid sequences containing three cysteines, upon reaction with tris(bromomethyl)benzene (*Figure 22A*). Incubation of biotinylated *c-kit1* promoter G-quadruplex with phages displaying a 3 × 3 bicyclic peptide library (ACX<sub>3</sub>CX<sub>3</sub>CG, X = any of the 20 canonical amino acids) and consecutive sequencing resulted in selection of the peptide ACPPIKIFCG (**G4pep2**, *Figure 22B*) having high affinity to the target ( $\Delta T_m$  of 20 °C in FRET-melting at 5  $\mu$ M and  $K_d$  of 1.0  $\mu$ M in fluorescence quenching experiment). Furthermore, the elucidation of structure of all peptides in the evolved library permitted to generate the structure of **G4pep3** (ACPRLCRRFCG, *Figure 22B*) from amino acids, the most probable for every defined position in the library. **G4pep3** possessed enhanced G4-binding properties:  $\Delta T_m$  of 34 °C in FRET-melting at 5  $\mu$ M and  $K_d$  of 630 nM in fluorescence quenching experiment. To compare, benchmark ligand **PhenDC3** provides thermal stabilization of more than 35 °C at concentration 1  $\mu$ M<sup>61</sup> and has a  $K_d$  value around 50 nM in the fluorescence quenching experiment.<sup>66</sup> Molecular modeling calculations were performed to study the interaction of **G4pep3** with G4 (*Figure 22C*). The peptide binds to 3'-tetrad of the G4. The complex was stabilized by hydrophobic interactions between the G-tetrad and leucine as well as N-terminus of the peptide. Besides, electrostatic interaction takes place between the belt-like architecture of arginines' positive charges of the peptide and the AGGAG loop.

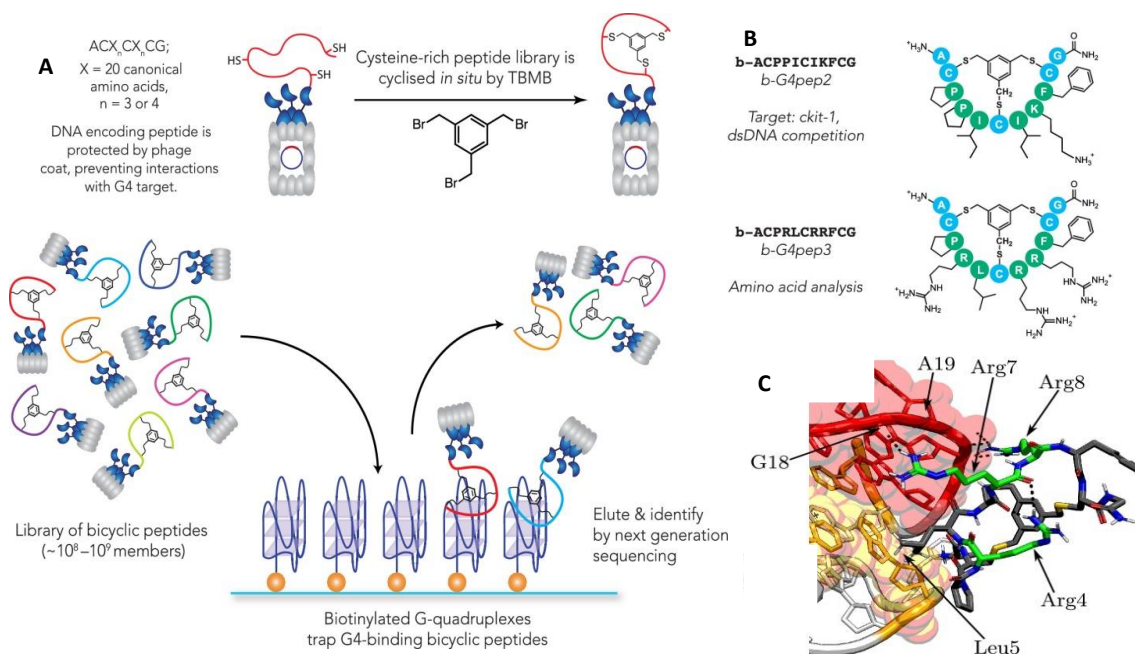


Figure 22. A) Phage display of bicyclic peptides for affinity selection of G4 ligands; B) Structures of **G4pep2** and **G4pep3**; C) Structure of the complex of *c-kit1*:**G4pep3**, proposed by molecular modeling. Interactions between Arg7/8 belt (green) and the phosphate groups of G18/A19 in the AGGAG loop (red), with stacking on 3'-tetrad (yellow) through Leu5 and the N-terminus (behind Leu5). Adapted from 113.

### 3.4. Fragment-based drug discovery

The approach of fragment-based drug discovery (FBDD) consists in a multistep selection of fragments moderately affine to the target with their subsequent assembling in the hit compounds. In the first step, small fragments (of molecular mass generally <300 Da, logP<3 etc.) are screened towards the target and molecules with some affinity (micro- or even millimolar) are identified. This preliminary screening is usually performed by NMR as it was first described by Abbot,<sup>114</sup> by X-ray crystallography, SPR, ITC, MST, or any other biophysical method. The second step consists in assembling the previously identified fragments in one molecule. Information about binding sites of fragments indicates how to better connect the fragments; the final molecule usually has significantly improved affinity characteristics.<sup>115</sup>

Fragment-based drug discovery has found wide application mostly for protein chemistry. The unique example of application of FBDD for the search of G4-binders was reported by Richter *et al.*<sup>116</sup> In a preliminary screening 150 molecules were tested in the FRET-melting assay

towards G-quadruplex from the HIV-1 long terminal repeat (LTR-III), human telomeric G-quadruplex and dsDNA. In the most active group of mono- and bifunctionalised aromatic and heteroaromatic rings, five compounds (**116.5**, **116.11**, **116.12**, **116.14** and **116.15**) provided thermal stabilization of G4s of more than 1 °C in the presence of 4000-fold excess of fragments and three of them (**116.5**, **116.11**, **116.12**), containing an amidoxime moiety, were selected for further studies. In the second round of selection, after incorporation of different heteroaryl units and side chains, 17 compounds were synthesized and two of them, **116.30** and **116.32**, (with  $\Delta T_m$  of 3.5 and 5.3 °C for LTR-III G4 in the presence of 100-fold excess of fragments) were selected for further optimization. Finally, hits **116.35** and **116.36** were selected from the third round for deeper investigation ( $\Delta T_m$  of 9.0 and 14.1 °C for LTR-III G4 in the presence of 100-fold excess of fragments in FRET-melting experiment;  $\Delta T_m$  of 1.7 and 5.1 °C for LTR-III G4 in the presence of 10-fold excess of fragments in CD-melting experiment). In a FRET-melting competition experiment, the  $T_m$  of labeled LTR-III in the presence of 100-fold excess of **116.36** was not affected by addition up to 32 equivalents of unlabelled *hTel* although a decrease of 7 °C was observed upon addition of 32 equivalents of unlabeled LTR-III. Compound **116.36** was shown to inhibit (at concentrations of 50–200 nM) the polymerase progression on a DNA matrix containing NHE-1 LTR forming sequence in Taq polymerase stop assay.

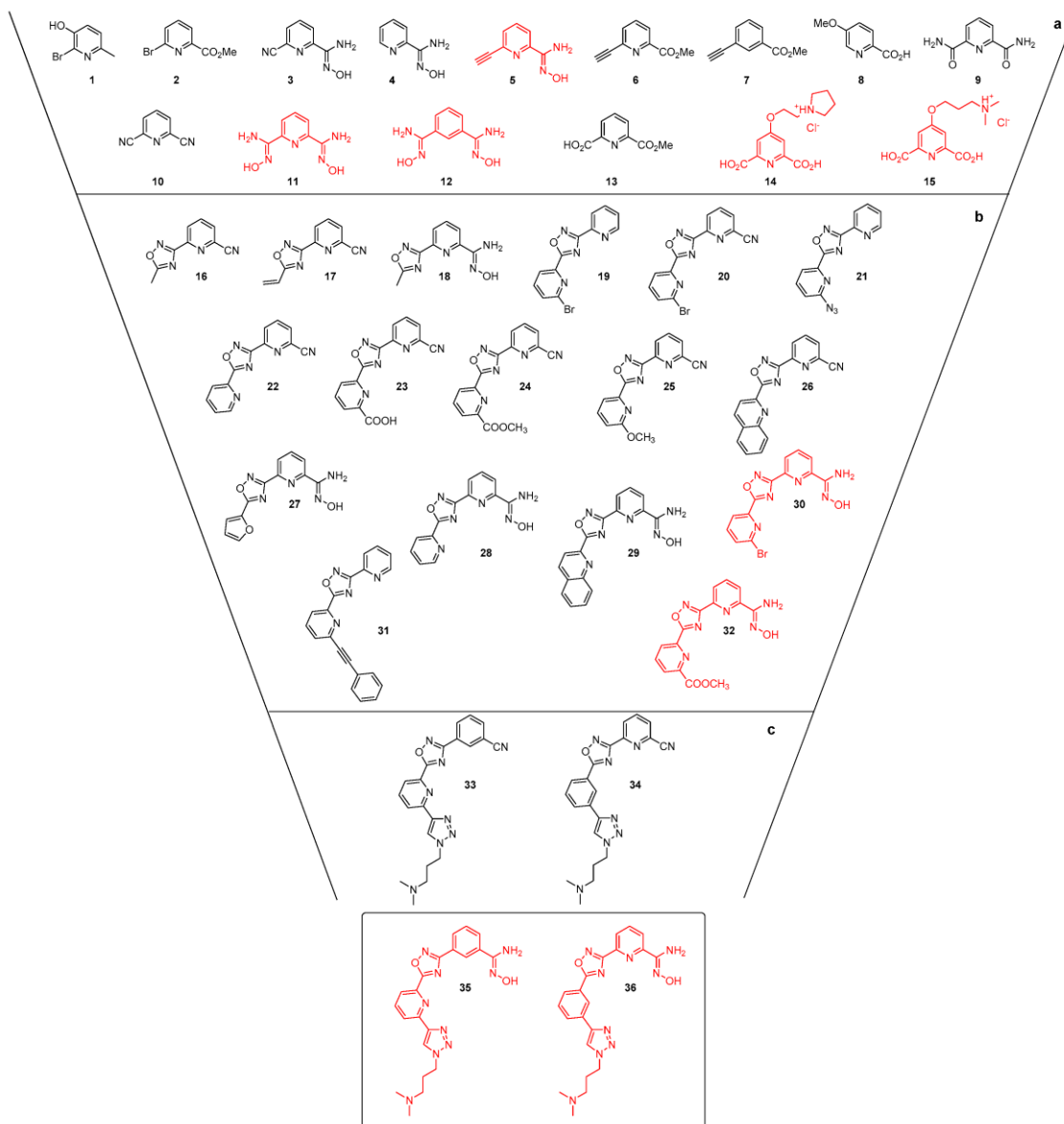


Figure 23. Fragment-based approach starting from (a) mono-aryl fragments, evolving through (b) low level of complexity bi- and tri-aryl derivatives and (c) final tetra-heteroaryl hits. Adapted with permission from <sup>116</sup>

### 3.5. Kinetic target-directed synthesis

Target-directed synthesis concerns the synthesis of small molecules in the presence of a biological target. It can be performed in two ways: upon kinetic control (kinetic target-guided synthesis, or KTGS) or in the conditions of dynamic equilibrium (dynamic combinatorial chemistry or DCC). In KTGS the target acts as a template for an irreversible reaction acting by

approaching the reacting species. The reagents bind to specific pockets (or binding sites) and if they occur to be at a suitable proximity, an *irreversible* formation of the product occurs. The most used KTGS reaction is *in situ* click chemistry of azide-alkyne cycloaddition.<sup>117</sup> In contrast, in the DCC approach the synthesis reaction is *reversible* and occurs regardless the presence of the target. The addition of the target changes the dynamic equilibrium shifting it towards a formation of products that are affine to the target and therefore are more thermodynamically stable in the “new” system. While dynamic combinatorial chemistry has found wide application in the field of G-quadruplexes (see Section 3.6), to date there are only two examples of target-directed synthesis of G-quadruplex binders and both employ the reaction of Huisgen cycloaddition.

The first example of application of KTGS in the G4 field was reported by Balasubramanian *et al.* The reaction of cycloaddition was performed in the system of two alkyne derivatives of **PDS** (**118.1** and **118.2**) and six azides (**118.3-118.8**) (*Figure 24A*).<sup>118</sup> Interestingly, in copper-free conditions the formation of only one product **118.10** (adduct **118.1-118.4**), a sugar derivative, was observed in the presence of telomeric G4 DNA and no product was formed in the absence of the target (*Figure 24B*). Conversely, in the presence of copper catalyst the products of all possible combinations were formed regardless the presence of G-quadruplex. Sugar derivatives **118.10** (**118.1-118.4**) and **118.16** (**118.2-118.4**) as well as some other derivatives of **118.2** were, however, amplified in the presence of telomeric G4. In FRET-melting experiments **118.10** and **118.16** provided the highest thermal stabilization of telomeric G4 (30.0 and 28.2 °C, respectively). In further studies, compound **118.10** was found to reduce the number of TRF1 foci in a dose-dependent manner (IC<sub>50</sub> of 1.3 μM) that supports the idea of its binding to telomere G-quadruplexes. When the reaction was carried out in the presence of copper (I) and a telomeric RNA G-quadruplex (TERRA), compound **118.13** (**118.1-118.7** adduct, also known as carboxypyridostatin or **cPDS**) was amplified in comparison to reactions performed without target, with dsDNA, or telomeric G-quadruplex. Interestingly, **118.13** contains a negatively charged carboxylic group that is rather unusual for G4 DNA ligands. The results of FRET-melting experiments indicated that although **118.13** provided slightly lower temperature of stabilization of RNA G4 than **PDS** ( $\Delta T_m$  of 20.7 and 22.2 °C, respectively),  $\Delta T_m$  for **118.13** was not affected by the presence of up to 100 equivalents of telomeric G4 competitor that was not the case for **PDS**. In order to explain such selectivity of **cPDS** to RNA,

docking and molecular dynamic simulations were performed by Artese *et al.*<sup>119</sup> The selectivity of **cPDS** to RNA G4 was explained by much lesser repulsion of ligand's carboxylic group from the negatively charged phosphate backbone, the last being involved in hydrogen bonds with the 2'-OH group of RNA's sugar constituent.

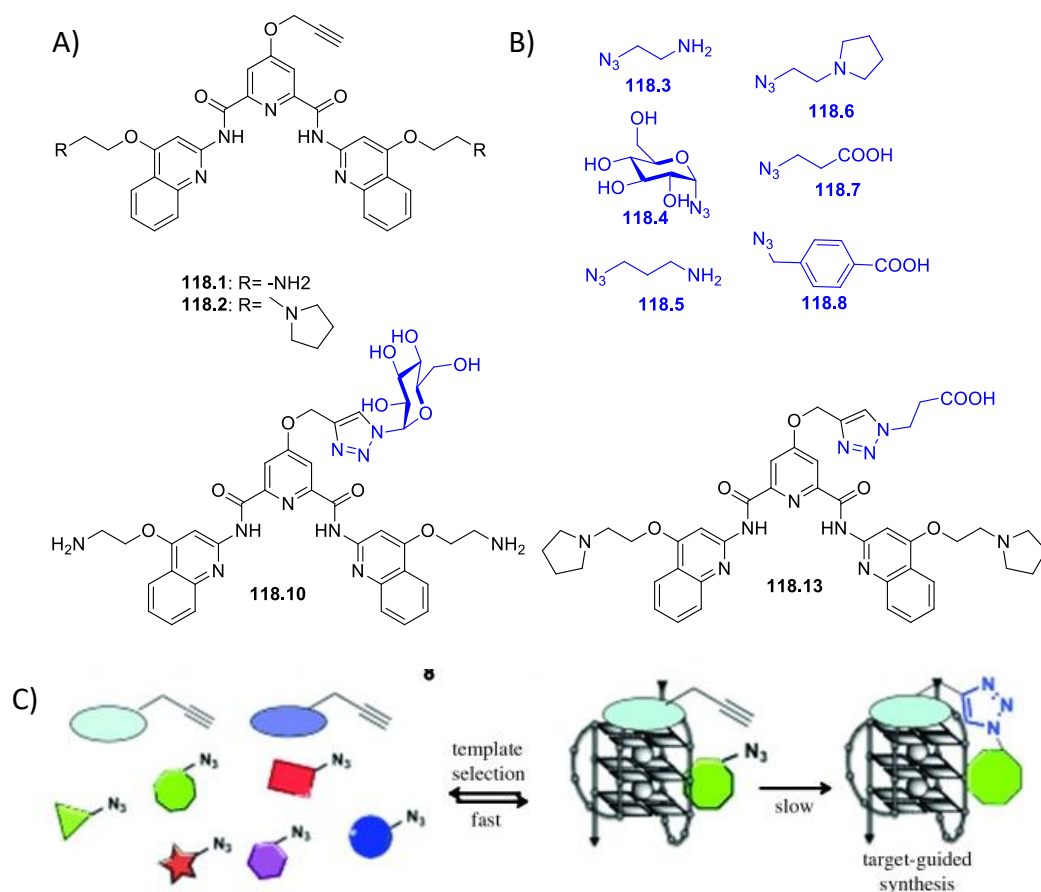


Figure 24. A) Structures of alkyne and azide building blocks; B) structures of amplifiers compounds **118.10** and **118.13** (cPDS); C) in situ synthesis of triazoles catalyzed by H-Telo.<sup>118</sup>

Another example of target-guided synthesis was reported by Dash *et al.*<sup>120</sup> for the system composed of three alkynes (**120.1a-c**) and 11 azides (**120.2-118.12**) (Figure 25). The target in this system was bound to paramagnetic nanoparticles covered with gold (Fe<sub>3</sub>O<sub>4</sub>@Au NP) to enable the pull-down of the DNA-ligand complex. The mixture of reagents and the NP-bound target (*c-myc* G4 or dsDNA) was incubated for 6 days and the complex of target with bound compounds was pulled-down with a magnet. The analysis of released ligands showed that in the presence of *c-myc* G4 three adducts of **120.1a** were formed: **120.1a-11**, **120.1a-3** and

**120.1a-7** in a 62:7:31 ratio, while only one compound (**120.1a-3**) was formed in the presence of dsDNA. Further studies of interaction of **120.1a-120.11** (**120.Tz1**) with *c-myc* G4 revealed a  $K_d$  of the complex of 0.17  $\mu\text{M}$  (measured by fluorescence titration) and thermal stabilization of *c-myc* G4 of 19.3  $^\circ\text{C}$  at 1  $\mu\text{M}$  ligand concentration in FRET-melting experiment (however, the concentration of labeled DNA used in this experiment was 100 nM in contrast to standard 200 nM, used for most reported ligands). Compound **120.Tz1** inhibited the proliferation of colorectal adenocarcinoma cancer cell line HCT116 (with  $\text{IC}_{50}$  of 2.1  $\mu\text{M}$  after 24 hours) and was capable to inhibit *c-myc* protein expression in a dose-dependent manner.

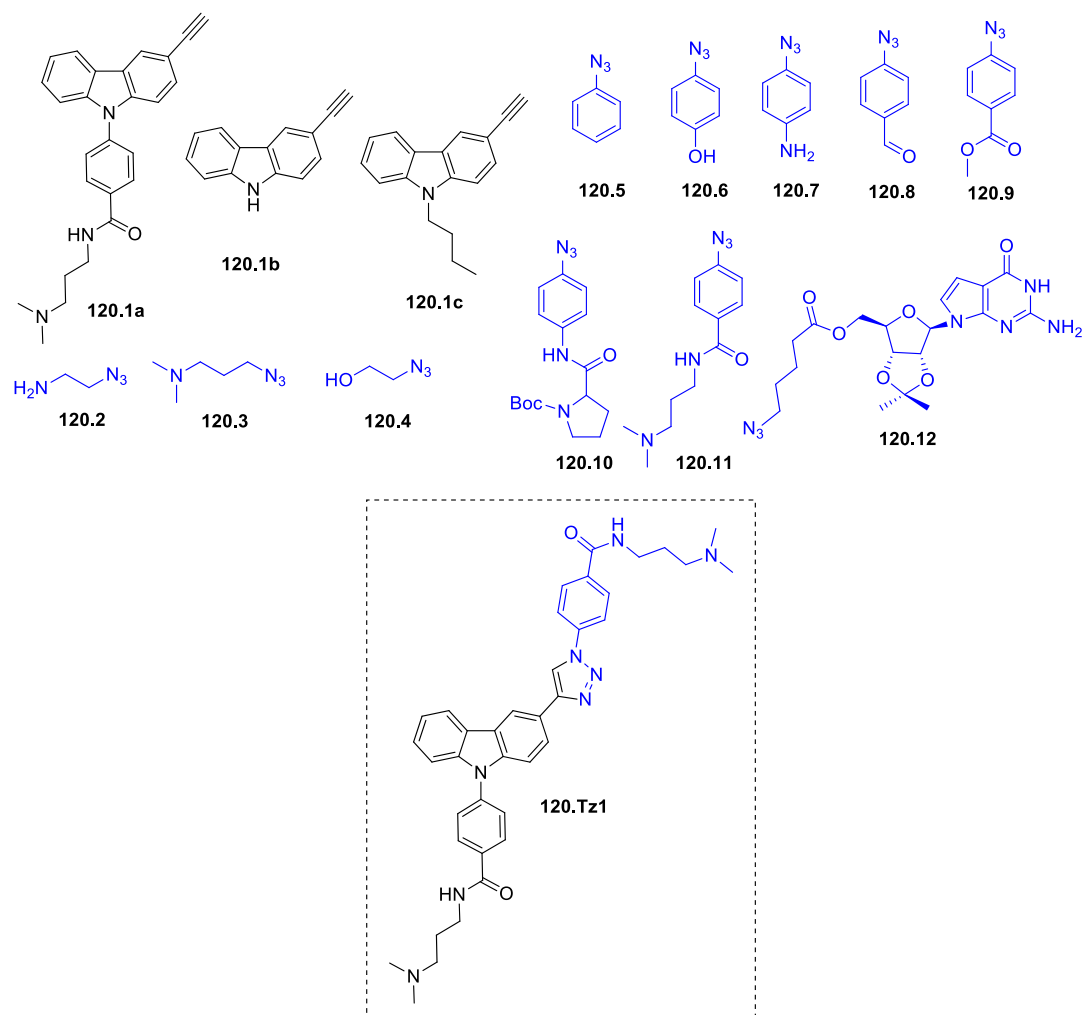


Figure 25. Structure of alkynes **120.1a-c**, azides **120.2-120.12** and the best binder **120.Tz1**.<sup>120</sup>



### 3.6 Dynamic combinatorial chemistry

Dynamic combinatorial chemistry (DCC) is an approach of combinatorial chemistry that permits to generate large libraries of compounds (dynamic combinatorial libraries, or DCL) by means of a reversible reaction between simple building blocks, performed under thermodynamic control.<sup>121</sup> These combinatorial libraries theoretically contain all possible combinations of building blocks (if the reactivity of building blocks is approximately the same). The reaction can be performed in the presence of a target (*adaptive DCC*) so that the composition of obtained DCL is compared to the “blank” library obtained in the absence of the target. Alternatively, the target can be added to the system where the equilibrium has already been established (*pre-equilibrated DCC*) (*Figure 26A*). In the latter case, the addition of the target induces the re-equilibration of the system in favor of formation of compounds that have some affinity to the target. In both cases the composition of the equilibrated library should contain the same distribution of components.<sup>122</sup> In the comparative approach of DCC the composition of libraries obtained in the presence and in the absence of the target is analyzed by HPLC or LC/MS, and an increase of the compound’s peak area (called amplification) normally witnesses its affinity to the target (*Figure 26B*). In the non-comparative approach the ligand-target complex is isolated and the hits are analyzed in the complex with the target or after being released from the target (*Figure 26C*).<sup>123</sup> The reaction used in dynamic combinatorial chemistry with nucleic acid targets should be water-compatible, give access to diverse and stable products and must be performed at near-neutral pH to keep native target’s structure. To date, two types of reactions have been used in G4-field for DCC: disulfide and imine exchange.

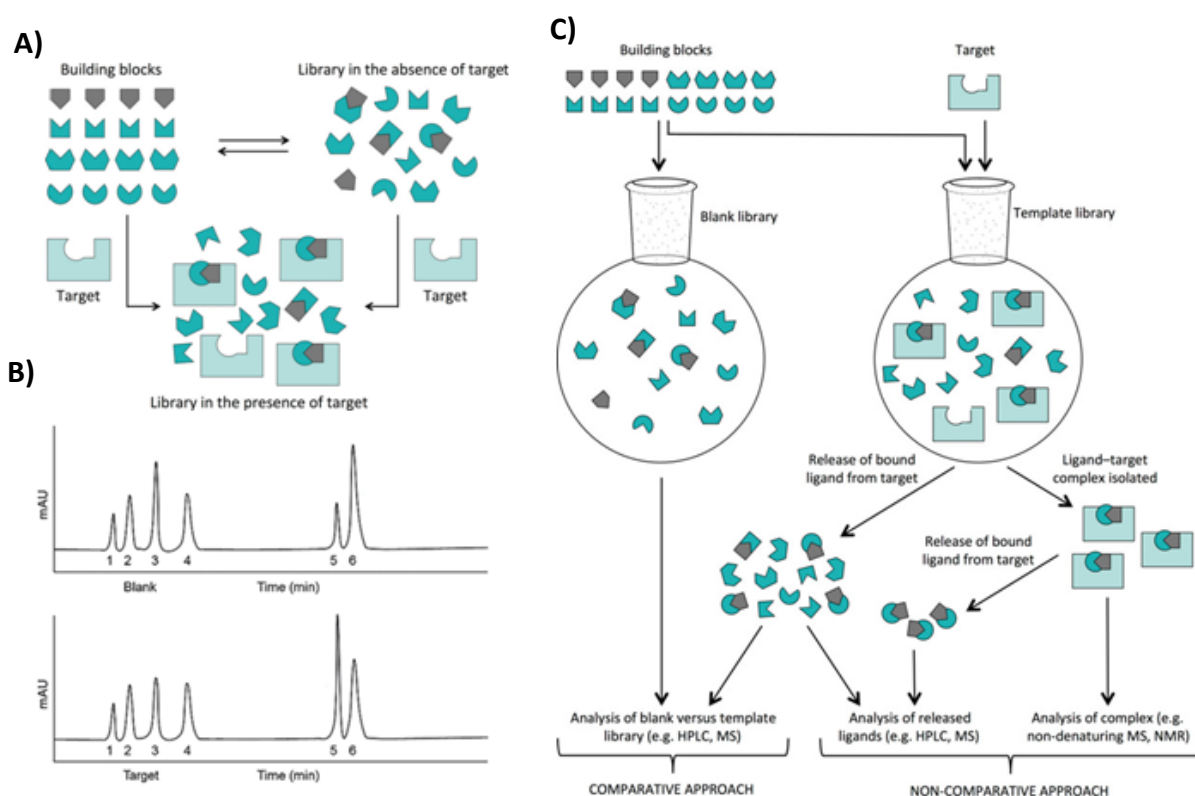
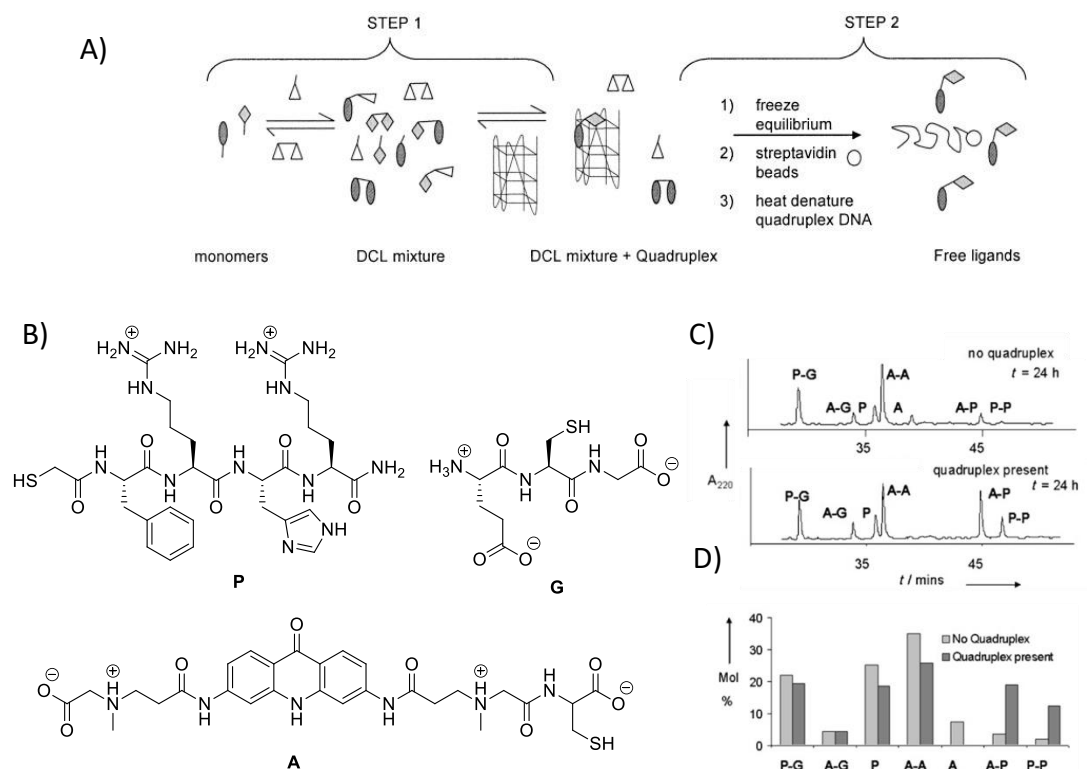


Figure 26. A) The principle of dynamic combinatorial chemistry; B) chromatograms of dynamic combinatorial libraries in the absence and in the presence of the target; C) overview of approaches of DCC (adapted from <sup>122</sup> and <sup>123</sup>).

The first application of DCC to the search of G-quadruplex ligands was described by Balasubramanian et al. and employed the disulfide exchange.<sup>124</sup> A relatively simple system was composed by the acridone derivative **A** (that was supposed to interact with external G-tetrads of G4s by  $\pi$ -stacking) and tetrapeptide FRHR **P** (presumable loop/groove-binder) in a buffer that contained excess of both oxidized (**G**) and reduced (**G-G**) forms of glutathione (**DCL1**, Scheme 4). When reaction reached the equilibrium it was stopped by lowering pH value to 2 and the “frozen” libraries were analyzed by HPLC. In a DNA-free system the peptide was predominantly conjugated with glutathione (**P-G**) whereas the acridone was in the form of homodimer (**A-A**). Upon addition of telomeric G-quadruplex to the system, an increase of peaks of **A-P** and **P-P** was observed, indicating the affinity of **A-P** and **P-P** to the target. In SPR experiments these compounds showed binding to G4 with  $K_d$  values of 30.0 and 22.5  $\mu\text{M}$  for **A-P** and **P-P**, respectively.



*Scheme 4. A) General scheme of the DCC experiment; B) Structures of building blocks; C) Chromatograms of equilibrated system without and with G-quadruplex. D) Histogram showing the change in equilibrium mixture composition on introduction of quadruplex DNA.<sup>124</sup>*

Another example, also reported by the same team, employed the peptidic spacing containing one, two or three *N*-methylpyrrole heterocycles that mimic the previously mentioned distamycin ligand (page 30) that binds DNA duplexes and G-quadruplexes (**DCL2**, Figure 27).<sup>125</sup> The reaction was performed in the presence of large excess of glutathione to avoid self-assembly of **P3** (up to 75% in low glutathione buffer). In the presence of human telomeric G-quadruplex, disulfides **P2-P3** and **P3-P3** were amplified (130 and 140%, respectively). In the presence of A/T-rich duplex DNA amplification observed much higher (660, 295 and 220% for **P2-P3**, **P3-P3** and **P2-P2**, respectively), suggesting that these type of polyheterocyclic peptides have higher affinity to hairpins than to G-quadruplexes.

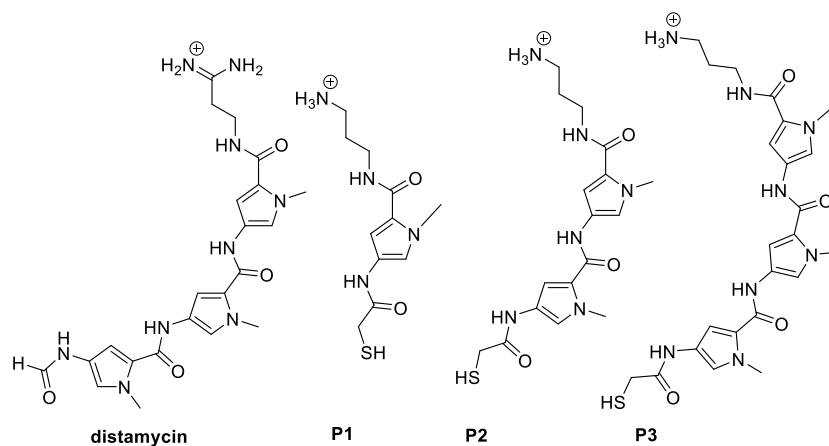


Figure 27. Structure of distamycin and building blocks of **DCL2**.<sup>125</sup>

In an attempt to find ligands that discriminate between two parallel G-quadruplexes, *myc22* and *c-kit21*, DCC of disulfides was applied to the system containing an oxazole-based peptide macrocycle moiety **125.1** (a G4 binder) and two libraries of *para*-substituted benzyl thiols (**DCL3a** and **DCL3b**, Figure 28A).<sup>126</sup> In the library **DCL3a**, two disulfide adducts were amplified upon addition of both G4s: compound **125.1-A** that was found to be more affine to *c-myc* G-quadruplex, and **125.1-E**, more selective to *c-kit*, as demonstrated by the respective  $K_d$  values (Figure 28B). Interesting results were obtained upon analysis of **DCL3b**: the adducts that contained  $\alpha$ - and  $\beta$ -lyxose chains had a large difference in amplification and affinity. With respect to *c-kit* G-quadruplex,  $\beta$ -lyxose derivative **125.1-F** was found to be the best binder with a  $K_d$  of 9.1  $\mu\text{M}$  (SPR), while the  $\alpha$ -isomer **125.1-G** was the worst in the whole library with a  $K_d$  of only 23.6  $\mu\text{M}$  (SPR). In the presence of *c-myc* G-quadruplex the amplification was much lower and all library members had rather low affinity to target (more than of 20  $\mu\text{M}$ , as per SPR).

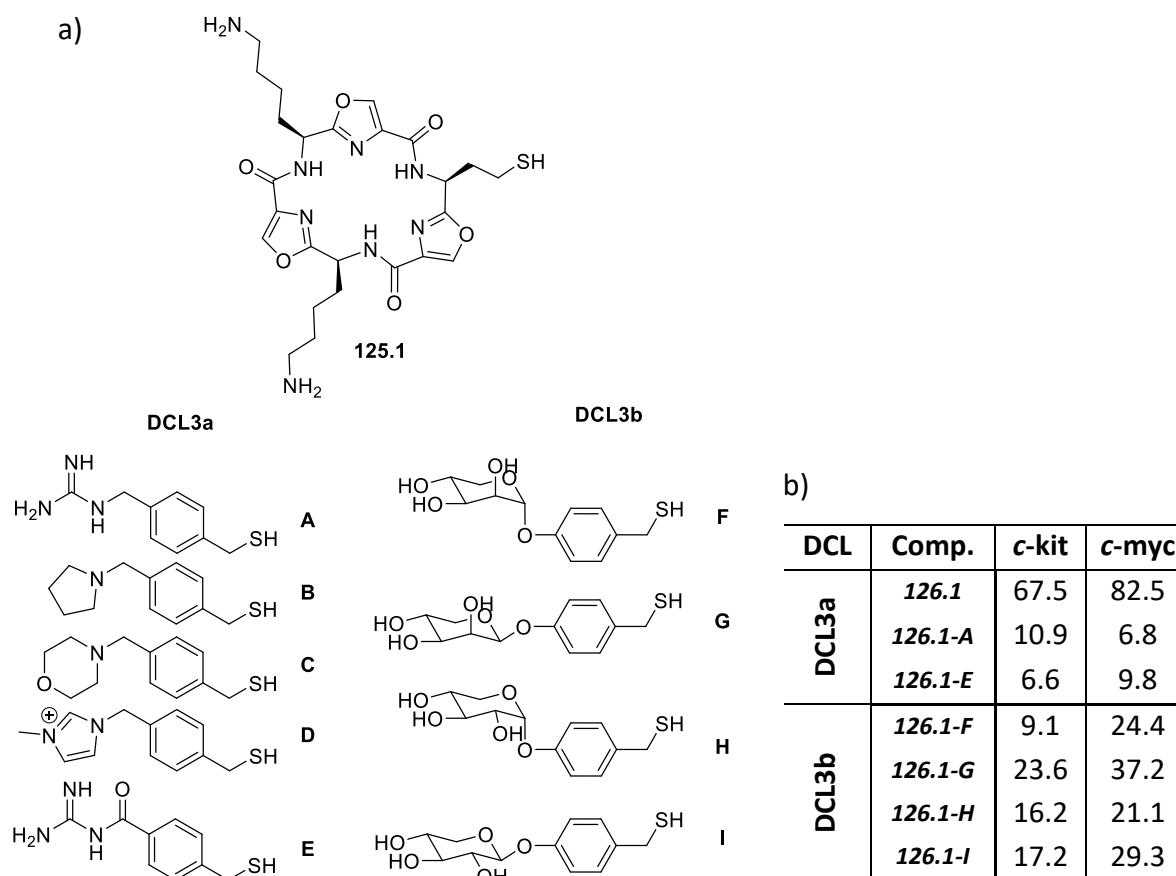


Figure 28. A) **DCL3**;<sup>126</sup> B) Affinity ( $K_d$ ,  $\mu\text{M}$ ) of members of **DCL3a** and **DCL3b** determined by SPR.<sup>126</sup>

A non-comparative DCC analysis protocol was applied by Ulven et al. to the system of disulfides of aromatic scaffolds **Aa**, **Ba<sub>2</sub>**, **Ca<sub>3</sub>** and cationic side chains **a<sub>2</sub>**, **b<sub>2</sub>** and **c<sub>2</sub>** (**DCL4**, Figure 29).<sup>127</sup> In this experiment the binders were identified by the analysis of ligands eluted from the pulled-down complex with a G-quadruplex, and not by comparing the composition of entire DCLs in the presence and absence of the target. The acridone derivative **126.Ac**, identified as the best binder to telomeric G-quadruplex in this experiment, demonstrated a  $K_d$  of 0.78  $\mu\text{M}$  to telomeric G-quadruplex (by SPR) and a  $\Delta T_m$  of 12.8 °C (by CD-melting).

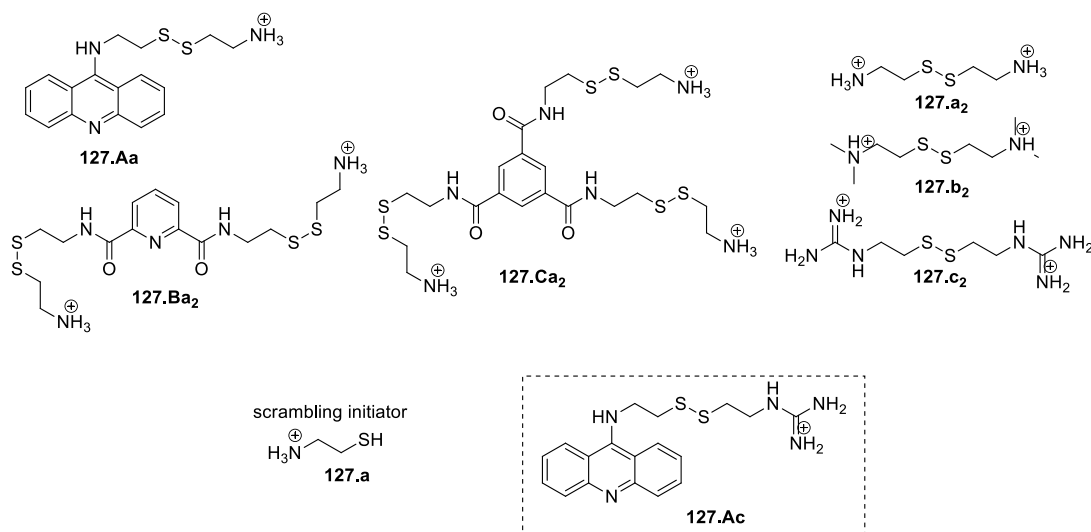


Figure 29. The composition of initial DCL4 and the structure of best telomeric G4 binder **127.Ac**.<sup>127</sup>

Not only disulfide exchange was used for the DCC search of G4-ligands. A library of imines was generated from the aldehyde **128.1** and ten amines **128.2a** - **128.2j** in the presence and in the absence of *c*-myc G-quadruplex by Dash et al. (Figure 30).<sup>128</sup> After 18 hours of equilibration, a reducing agent (NaBH<sub>3</sub>CN) was added and all imines were reduced to stable amines. As in the previous example, the identification of best binders was performed only by analyzing the molecules pulled down from the DCL. Specifically, DNA oligonucleotides (*c*-myc G4 and dsDNA control) were functionalized with thiol groups and attached to magnetic gold-coated nanoparticles. Interestingly, in the presence of G4-target after 4 hours only three products were observed (**128.4e**, **128.4g** and **128.4h**) and after 18 hours the mixture contained only one product **128.4e**. On the contrary, in the presence of dsDNA numerous adducts were present in the mixture after 18 hours. All amines were synthesized in the preparative manner and FRET-melting experiment was performed to rank them according to their capacity to thermally stabilize the *c*-myc G-quadruplex. The three abovementioned compounds **128.4e**, **128.4g** and **128.4h** provided  $\Delta T_m$  of more than 10 °C (23.4, 13.6 and 11.8 °C, respectively). Compound **128.4e** gave a  $K_d$  value of 1.08  $\mu$ M to *c*-myc G4 (according to the fluorescent titration experiment).

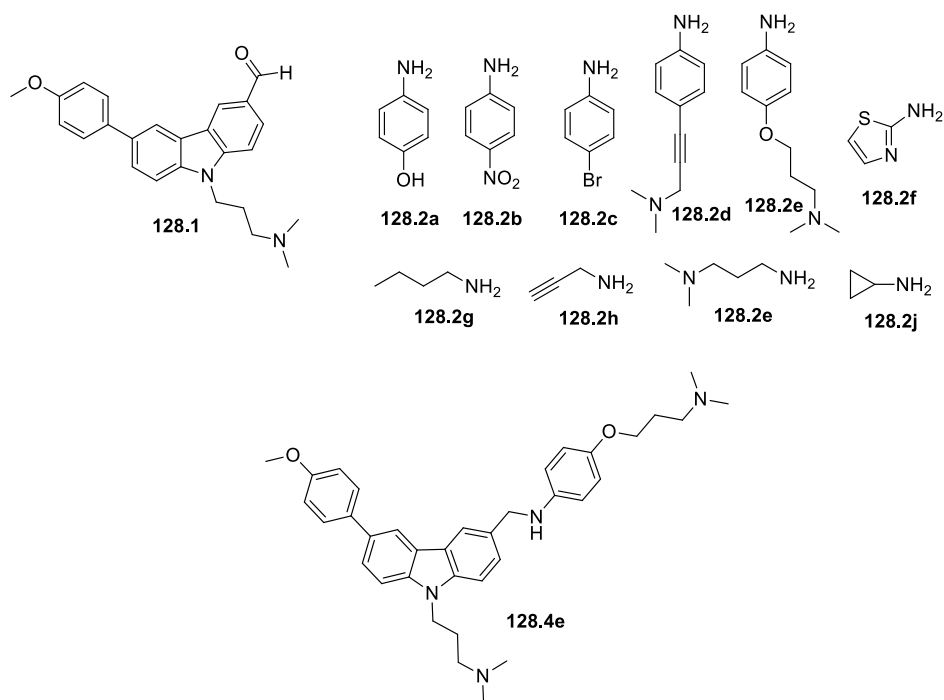


Figure 30. Building blocks of **DCL5** and structure of the lead compound **128.4e**.<sup>128</sup>

In summary, various approaches of combinatorial chemistry have been applied for the search of G-quadruplex ligands. Each approach has its own benefits and drawbacks. Thus, classical combinatorial chemistry coupled with high-throughput screening is a proven efficient approach of drug discovery that can be applied to a wide variety of chemical scaffolds. A more recent method of DNA-encoded libraries permits to significantly reduce the sample consumption but, at the same time, has some limitations regarding the structures of constituents and the reactions used to construct the libraries. Finally, dynamic combinatorial chemistry permits to screen “virtual” libraries but has limitations with respect to the size of libraries (up to 20 compounds) due to existing analysis methods.

## 4. Biophysical techniques to study DNA– ligand interactions

The interactions between G-quadruplexes and ligands can be studied by numerous physico-chemical methods aiming to determine the structure of the complex, or thermodynamic or kinetic parameters of its formation. They include, for example, crystallography, NMR-spectroscopy, CD-spectroscopy, and a variety of fluorescence-based methods such as direct fluorescence quenching methods, FRET-melting, UV-melting, microscale thermophoresis, fluorescence anisotropy etc. Only the methods employed in this work will be discussed in detail below.

### 4.1 Circular dichroism (CD)

Circular dichroism is a method of absorption spectroscopy that allows to get an insight into the conformation of nucleic acids and proteins. The method is based on differential absorption of left- and right-handed circularly polarized light by chiral molecules. Thus, CD spectroscopy may be used for molecules that are chiral and absorb light.

Circularly polarized light is an electric wave of constant magnitude that oscillates in a plane perpendicular to its propagation. It can be polarized in two directions: right (RCP), when the electric field vector rotates in the clockwise direction, and left (LCP), when it rotates counterclockwise. When LCP and RCP are of the same magnitude they are superimposed and result in a linearly polarized wave. While passing through a medium that contains chiral adsorbing compounds, the RCP and LCP light components are adsorbed in different extents and their resulting wave becomes not linearly, but elliptically polarized. The differential absorption of the radiation polarized in two directions as a function of frequency is called dichroism. It can be expressed as molar circular dichroism, a difference in molar adsorption coefficient of components ( $\Delta\varepsilon = \varepsilon_l - \varepsilon_r$ ) [ $M^{-1} \times cm^{-1}$ ] or as ellipticity  $\Theta$  [ $deg \times cm^2 \times dmol^{-1}$ ] that



corresponds to the angle whose tangent is the ratio between the minor and major axes of the resulting elliptically polarized light.<sup>129</sup>

Nucleic acids contain a chiral sugar backbone and nucleobases that adsorb light in the ultraviolet region. Absorption peaks of nucleobases correspond to wavelengths of electron  $\pi \rightarrow \pi^*$  transitions. Though even simple nucleobases are chiral, the main source of nucleic acid chirality comes from their secondary structure. Guanine has two adsorption bands that corresponds to its two electronic ( $\pi \rightarrow \pi^*$ ) transitions at 248 and 279 nm and are coupled to each other. The coupling depends on the positions of neighboring stacked guanines that are rotated by a certain angle. Although folding of phosphate backbone usually determines the topology of G-quadruplex, the mutual orientation of strands and, more importantly, the polarity of stacked guanines contribute most to CD spectrum of G-quadruplexes (*Figure 31A*).

CD spectrum of a G-quadruplex of parallel topology is characterized by a positive peak at 260 nm and a negative one at 240 nm; antiparallel G-quadruplexes show a positive band at 290 nm and a negative one at 260 nm; and hybrid G-quadruplexes have a positive band at 290 nm and a shoulder at 260-270 nm (*Figure 31B*).<sup>130</sup> Upon ligand binding, the CD spectrum of a G-quadruplex can increase, diminish, stay the same or change the pattern. Rather often, upon ligand addition the change of topology from hybrid to antiparallel (or to parallel) is observed, which can be a significant proof of ligand-macromolecule interaction. The increase of the signal often indicates an interaction. However, if the pattern and intensity of the spectrum do not change upon addition of the ligand, it does not necessarily point to the absence of interaction. The decrease of the CD signal sometimes can point to the formation of aggregates and can also therefore indicate the interaction of a ligand with the macromolecule. However, in certain cases, this situation is interpreted as a ligand-induced “unfolding” of a G-quadruplex structure.<sup>131</sup> Usually, it is difficult to obtain any quantitative information about ligand binding from CD titration but sometimes it can give some insights about  $K_d$  or stoichiometry of the complex. CD spectroscopy is a very good tool to monitor the stability of the structure. This feature is used in CD-melting experiments, where thermal stability of G-quadruplex alone is compared with stability of the complex of G4 with a ligand. Concentration of DNA necessary to record its CD spectrum usually varies between 2-50  $\mu\text{M}$ .

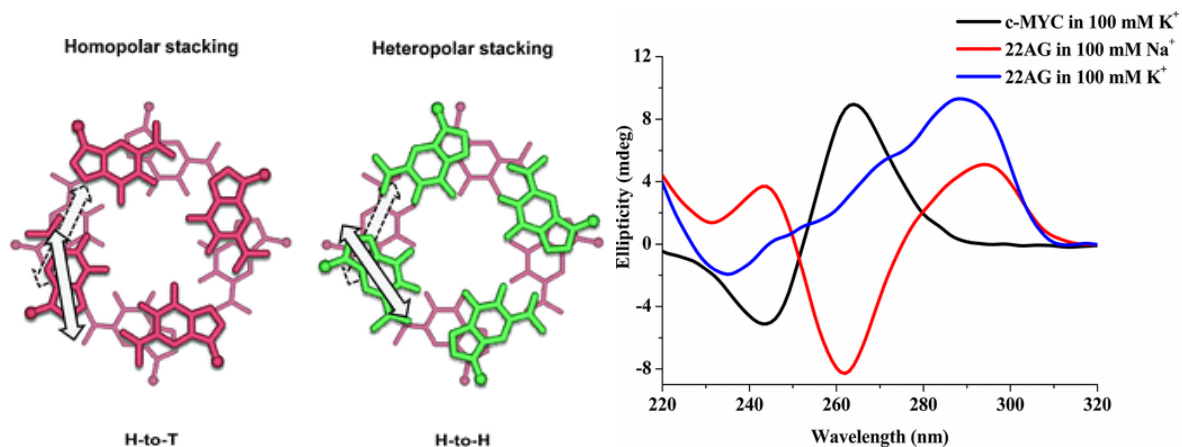


Figure 31. A) Top view of the homopolar (parallel topology) and heteropolar (antiparallel topology) stacking of two G-quartets: the “head” and the “tail” sides of the G-quartets are represented in red and green, respectively (the double-head arrows represent the transition moments corresponding to the absorption band at ca. 250 nm) (adapted from <sup>130</sup>); B) CD spectra of guanine quadruplexes of parallel c-myc (black), antiparallel 22AG in Na<sup>+</sup>-buffer (red) and hybrid 22AG in K<sup>+</sup>-buffer (blue) G-quadruplexes (adapted from <sup>132</sup>).

The binding of ligand can also be observed by the appearance (or change) of the “induced” CD signal from ligand: even if the ligand is a symmetric planar molecule, its two faces are not equivalent any more in a complex with a macromolecule (induced chirality).

The main limitation of CD spectroscopy is long time of the experiment (around 5 min for one spectrum) and the absence of a direct information regarding the DNA affinity of the ligand.

## 4.2 FRET-melting

The fluorescence-melting or FRET (Föster Resonance Energy Transfer)-melting experiment resides in a measurement of a temperature of denaturation of a G-quadruplex (or its complex with a ligand) and therefore permits to evaluate the ligand-induced thermal stabilization. Double-labeled oligonucleotides are used in this experiment. The most common pair of donor-acceptor dyes is FAM (carboxyfluorescein) and TAMRA (Carboxytetramethylrhodamine) (Figure 32A). The donor-acceptor pair is selected in such a way that the donor’s emission spectrum overlaps with the acceptor’s excitation spectrum. Excitation energy is transferred from the excited donor to the quencher so that relaxation to the ground state of donor takes

place and the photon will be emitted by the acceptor. This phenomenon is called Förster Resonance Energy Transfer (FRET) and highly depends on the distance between two dyes ( $\sim 1/R^6$ ). At room temperature, when a G-quadruplex is in a folded state, the donor and the quencher are in proximity so that energy transfer takes place and the fluorescence of the donor is quenched. When G-quadruplex unfolds into a single strand upon heating, the distance between the two dyes increases and therefore the enhancement of donor's fluorescence takes place (*Figure 32B*). The fluorescence of the acceptor changes in the opposite way but, due to technical issues, usually is not recorded in this experiment. In FRET-melting experiment the dependence of donor's fluorescence on the temperature is recorded (the typical curve is shown on the *Figure 32C*). Temperature of melting ( $T_m$  or  $T_{1/2}$ ) of G-quadruplex corresponds to the maximum of the peak of the first derivative of fluorescence vs. temperature curve, or to the temperature at which normalized fluorescence is 0.5. When G4 forms a complex with a ligand that stabilizes its structure, the G4 unfolds at a higher temperature characterized by  $\Delta T_m$  that is a difference between  $T_m$  of a G4-ligand complex and  $T_m$  of a G4 alone.  $\Delta T_m$  may be employed as an indirect indication of a ligand's affinity (more affine ligands usually provide higher  $\Delta T_m$  values). However, possible inconsistencies of FRET-melting with isothermal assays may be due to formation of complexes of different composition at higher temperatures that will affect the  $\Delta T_m$ .<sup>133</sup> A competition experiment can be easily designed by addition of an unlabelled competitor (most often, a self-complementary 26-mer oligonucleotide *ds26*: 5'-d(CAATCGGATCGAATTCGATCCGATTG)-3'). A ligand is considered selective if its  $\Delta T_m$  value is not affected by a presence of an unlabeled competitor. Standard concentration of a double-labeled DNA or RNA substrate in this experiment is 0.2  $\mu\text{M}$  and the ligand is often used in slight excess (0.2-2  $\mu\text{M}$  for "good" and 1-10  $\mu\text{M}$  for "poor" ligands). In "standard conditions" a concentration of the competitor is 3 and 10  $\mu\text{M}$  but is not limited to these values.<sup>134</sup> A FRET-melting experiment can be carried out on fluorescence spectrophotometers in cuvettes, but it allows a simultaneous analysis of only a few samples with volumes of at least 500  $\mu\text{L}$  in this setup. Usually, FRET-melting is performed on a real-time PCR (RT-PCR) machine in a 96-well plate that makes the experiment faster (more samples at the same time) and permits to decrease a sample volume to 25  $\mu\text{L}$ . In addition, new PCR machines have an option of an automatic analysis of melting curves including the plotting of

the first derivative, that significantly facilitates data analysis. The temperature range employed in FRET-melting is normally between 25 and 95 °C.<sup>134</sup>

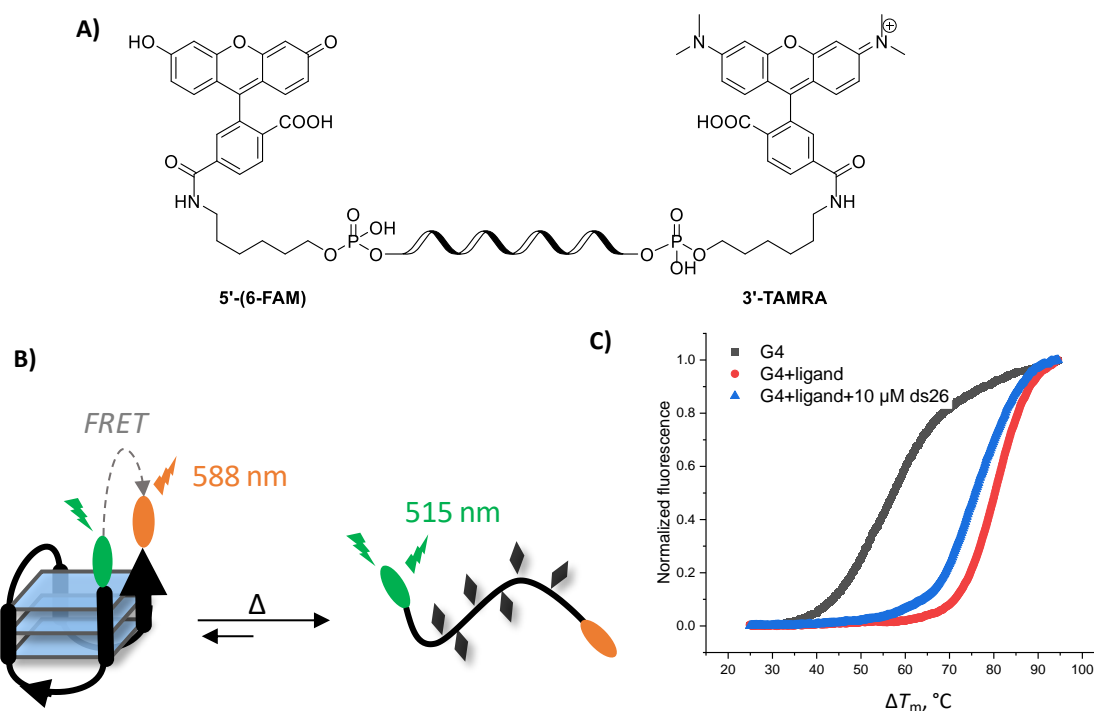


Figure 32. A) Structures of a pair of most used fluorophores in FRET-melting experiment (FAM and TAMRA); B) Principle of a FRET-melting experiment; C) Example of melting curves of F-G4-T without (black) and with (red, blue) G4-stabilising ligand and 10  $\mu\text{M}$  of competitor ds26 (blue).

Several precautions should be taken in order to perform FRET-melting experiment correctly. First, the  $T_m$  of the competitor should be significantly higher than  $T_m$  of the G4 and, ideally, higher than  $T_m$  of a complex G4-ligand. Second, a competitor should not be complementary to G4 (though in a hybridization experiment it is a requirement). Third, ideally, a ligand should not interact with the fluorophores. Finally, the fluorescence of the most widely used fluorophore donor (fluorescein) is pH-dependent and decreases at pH below 7. It should be noted that the efficiency of a ligand in the FRET-melting assay does not guarantee that this compound is going to be active in cell-experiments. To overcome this issue Morgan et al.<sup>135</sup> suggested using a buffer with additives such as 10% glycerol and two annealing cycles to improve the significance of results obtained by FRET-melting experiment.

### 4.3 FID assay

A fluorescent indicator displacement (FID) experiment is a widely used method that assesses a ligand's capacity to displace a fluorescent dye from its complex with G4 and provides an indirect measure of a ligand's affinity to G4 (Figure 33A). The most common dye used in FID is Thiazole Orange (TO) (Figure 33B). TO is highly fluorescent in the DNA-bound state ( $\lambda_{\text{ex}} = 488$  nm,  $\lambda_{\text{em}} = 544$  nm) where its conformation is fixed. In the unbound state, it loses fluorescence due to conformational rotation. A FID experiment consists in a titration of a complex G4-TO with ligand, aiming at a determination of the  $\text{DC}_{50}$  value, that is, the concentration of a ligand at which the fluorescence of TO is decreased by 50% (that is to say, 50% of TO is displaced from the complex). Some examples of titration curves are demonstrated on Figure 33C.

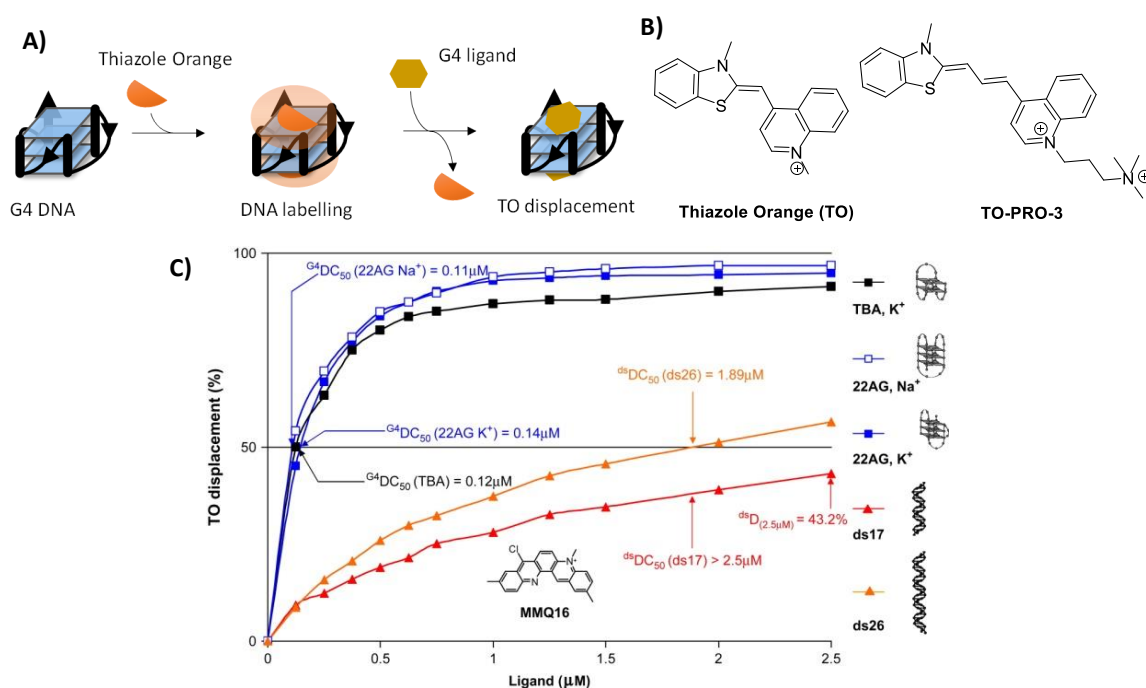


Figure 33. A) Principle of the FID assay; B) Structures of Thiazole Orange and TO-PRO-3; C) Example of FID assay with **MMA16** and different G-quadruplexes and duplexes (adapted from <sup>136</sup>).

TO binds to G4 with  $K_d$  in the micromolar range,<sup>136</sup> so the information about concentration at which ligand is capable to displace TO gives some insights about the ligand's affinity to the given G4. FID is an easy and quick method. A titration can be performed in a cuvette by step-by-step ligand addition, or solutions of the G4-TO complex can be mixed with the different concentrations of the ligand in 96- or 384-well plates. Attention should be paid for the analysis

of fluorescent G4-ligands: the ligand should not be fluorescent in the same spectral range as TO. If this is the case, other fluorescent indicators should be used for the experiment (such as TO-PRO-3). Also, in order to effectively rank the ligands, TO and the scrutinized binder should have the same binding site. However, some false negative results may occur when the ligand binds to the preformed complex of G4 × TO leading to formation of G4–ligand–TO (or G4–ligand<sub>2</sub>–TO) ternary complexes.<sup>137</sup>

#### 4.4 Fluorescence quenching assay (fluorimetric titrations)

The method of fluorescence titration was described by Balasubramanian *et al.*<sup>66</sup> The authors showed that the binding of a ligand to a terminally labelled G-quadruplex induced proximal quenching of fluorescence of the dye (*Figure 34A*) and that the fluorescence was restored by adding an excess of non-labeled G4 competitor. Moreover, by varying a fluorophore's position (5' vs. 3') it was possible to determine  $K_d$  values of a ligand binding to the two G4's external tetrads. The method was cross-validated with respect to literature data by comparing the affinities of **TmPyP4** to telomeric, NRAS, c-kit and c-myc sequences measured by fluorescence quench assay and those reported previously (obtained in UV-visible titration, FRET-quenching, SPR and ITC experiments). They also measured the affinities of **TmPyP4**, **PDS** and **PhenDC3** towards the different tetrads of telomeric and c-kit sequences. Interestingly, while **PDS** seems to do not distinguish between the tetrads of the c-kit structure ( $K_d$  values of 0.64 and 0.44  $\mu$ M to 5'- and to 3'- G-tetrads, respectively), **PhenDC3** shows a 46-fold tighter 5'-tetrad binding ( $K_d$  values of approximately 9 and 330 nM to 5'- and to 3'- G-tetrads, respectively). In the described assay the oligonucleotides were labeled with the Cy5 dye (*Figure 34B*).

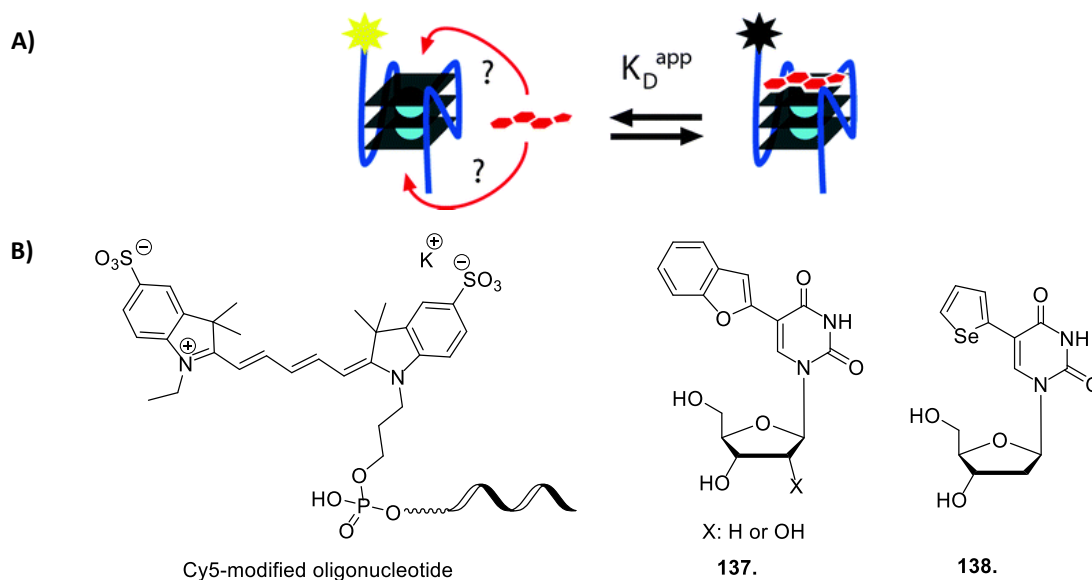


Figure 34. A) Principle of fluorescent quench assay. G4-tetrad ligand's selectivity can be study by labeling different G4-tetrads of G-quadruplex (adapted from <sup>66</sup>); B) Structures of Cy5-modified oligonucleotide, **137**. and **138**.

A similar principle of proximal fluorescence quenching was used by Srivatsan *et al.*<sup>138</sup> for the development of topology-specific fluorescent turn-on probes for DNA and RNA G-quadruplexes. Fluorescent 5-(benzofuran-2-yl)uracil **137**. (Figure 34B) was incorporated into telomeric DNA and RNA G4-forming sequences. The topology selectivity manifested in different fluorescence enhancement of monomolecular G4s comparing to double-strands formed with their complementary strands: for antiparallel G4s, formed in Na<sup>+</sup>-conditions, the enhancement of fluorescence was 9-fold while for G4s of mixed structures, formed in the presence of K<sup>+</sup>, it was only 4-fold. RNA G4 showed fluorescence enhancement of about 5-fold comparing to DNA:RNA double-stranded structure. The addition of **PDS**, a G4-ligand, on the contrary, induced a quenching of fluorescence: 4-fold for RNA G4 and DNA G4 of hybrid and mixed G4-conformations (in K<sup>+</sup>) and 8-fold for antiparallel DNA G4 structures. By monitoring fluorescence quenching upon gradual addition of a ligand, a saturation curve could be obtained and  $K_d$  value could be calculated. For example, it was shown that **PDS** has a 2-fold higher affinity for an antiparallel telomeric G4 than to a G4 of mixed conformations obtained in K<sup>+</sup> ( $K_d$  of 440 and 919 nM, respectively). **BRACO-19** showed an opposite tendency: it binds more tightly to G4 of hybrid topology ( $K_d$  of 476 nM compared to 938 nM for antiparallel G4).

Dissociation constants of both ligands to RNA G4 were higher (244-294 nM) comparing to DNA of any topology.

Recently, the same group developed a 5-selenophene-modified 2'-deoxyuridine nucleoside probe **138**. (*Figure 34B*) that was capable not only to report the formation of a G4 structure and its binding to a ligand, but also had a potential to facilitate the X-ray structural analysis of an oligonucleotide or its complex.<sup>139</sup> In this work, they modified a long terminal G4-forming repeat sequence of the HIV-1 proviral genome (LTR-IV) with the probe. As in the previous example, the fluorescence was much higher in the system containing G4-structure comparing to the duplex. Dose-dependent fluorescence quenching was observed upon addition of G4-ligands such as **BRACO-19** or **TmPyP4**.

The main advantage of probes based on nucleoside-modification is that they essentially do not affect topology of G-quadruplexes, which is not always the case for Cy5-modified G4-oligonucleotide sequences. For example, a conformation change of telomeric sequence from hybrid (unlabeled sequence) to parallel (5'-Cy5 labeled oligonucleotide) topology was reported.<sup>66</sup> At the same time, Cy5-modified oligonucleotides can be easily ordered from a commercial supplier, whereas in order to introduce a more specific nucleoside modification, a solid-phase DNA synthesizer and a corresponding phosphoramidite are needed.

The main advantage of the method of fluorescence quenching is that it gives a direct access to  $K_d$  at a constant temperature (that is not a case in the FRET-melting experiment) that is closer to native conditions of a functioning cell machinery.

#### 4.5 Native mass spectrometry (MS)

Non-covalent interactions of G-quadruplexes with ligands can also be followed by mass-spectrometry. Despite the fact that the analysis takes place in the gas phase, the methods of soft ionization such as matrix-assisted laser desorption/ionization (MALDI) or electrospray ionization (ESI) do not perturb the composition of a complex. The buffer used for ESI-MS experiments is usually slightly different than standard buffers for G4s study: instead of high



concentrations of potassium or sodium chloride (up to 150 mM), their amount is reduced to 1 mM and the ionic strength of solution is achieved by addition of ammonium or trimethylammonium acetate (TMAA). For the analysis of G-quadruplexes in ammonium acetate buffer, Na<sup>+</sup> or K<sup>+</sup> salts can even be omitted, because NH<sub>4</sub><sup>+</sup> coordinates guanines by itself, however, G-quadruplexes are less stable and more polymorphic in this case. On the contrary, the bulky trimethylammonium cation is not coordinated by G-quadruplex so the resulting spectrum is much clearer than in the case of NH<sub>4</sub><sup>+</sup> usage.<sup>140</sup>

A lot of important information about G4-ligand complexes can be extracted from mass spectra. First of all, it is a stoichiometry of complex. The majority of ligands bind to G4 by stacking to external G-tetrads and therefore the stoichiometry of their complexes with G4 is expected to be 1:1 and 2:1 (ligand:G4). For example, 2:1 complexes can be observed for *Pu24T* with **PDC** and **PhenDC3** even in solutions containing G4:ligand in 1:1 ratio. One-site binding might point on a more specific ligand-loop interaction or just a lower binding affinity to G-quadruplex, as in a case of *Pu24T* and **PDS**.<sup>141</sup>

Second, the number of G-tetrads in the G-quadruplex can be determined considering the fact that one K<sup>+</sup> is placed between two G-tetrads. It is particularly interesting in the systems where the binding of a ligand leads to a change of G4 topology. For example, a binding of **PDC** and **PhenDC3** ligands to three-quartet hybrid G-quadruplex 24TTG induces a change of its topology to an antiparallel two-quartet structure. This change is accompanied by one K<sup>+</sup> ejection and can be easily followed by mass-spectrometry (*Figure 35*).<sup>142</sup>

Finally, ligand's affinity to G-quadruplexes can be measured by mass-spectrometry as a result of direct titration of G4 with ligand, considering the fact that peaks areas of the unbound and ligand-bound DNA forms are directly proportional to the concentration of the corresponding species. However, one should be careful while interpreting these results: when the ligand does not change the G4 topology, the apparent  $K_d$  ( $K_{app}$ ) corresponds to the "real"  $K_d$ . But if a change of topology upon ligand binding takes place, the system's equilibrium includes not only ligand's binding to each conformation but also the equilibrium between the two G4 conformations and the unfolded (or partially unfolded) oligonucleotide's state,<sup>141</sup> leading to a biased  $K_{app}$  values.

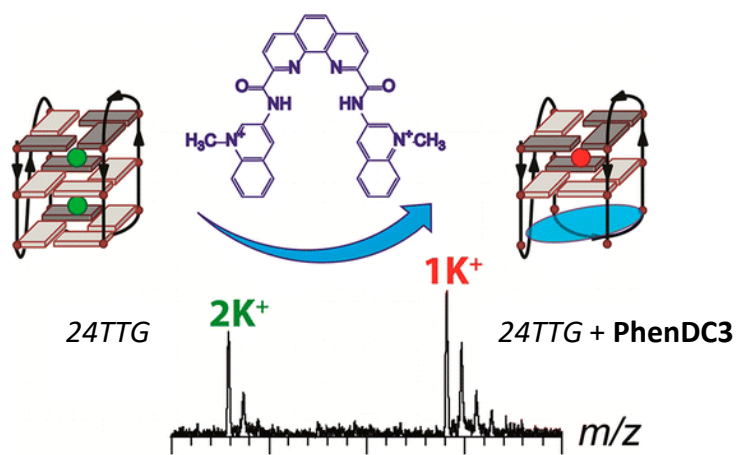


Figure 35. Ligand-induced conformational change of 24TTG G-quadruplex (adapted from <sup>142</sup>).



# ***AIMS OF THE WORK***

Being rather unusual structures, G-quadruplexes have very important regulation functions. Already the fact that they are distributed non-randomly in the genome points to this. G4-targeting is often associated with (but not limited to) oncology, where ligands that stabilize G-quadruplexes have three major applications: to suppress oncogenes, to inhibit telomerase, and to induce DNA damage in cancer cells.

Targeting of G-quadruplexes is challenging taking into account the fact that they have rather similar structures and ligands that are selective to G4 over duplex DNA often do not distinguish between different G-quadruplexes. The capacity to selectively bind only a specific subtype of G-quadruplexes is, however, a very desirable quality of a G-quadruplex ligand since, for example, stabilization of some G4s enhances transcription while stabilization of others leads to transcription suppression. In addition, cellular machinery involves many G4-binding proteins, making G4s not easily accessible targets. Although numerous compounds were reported as G-quadruplex ligands, the search of compounds that could differentiate between distinct G4 structures is still a very challenging task.

The goal of my project is to develop a methodology for rapid and simple synthesis and screening of compounds, that could facilitate the search of structurally diverse, potentially selective, highly affine and biologically active G-quadruplex binders. Towards this end, we put emphasis on dynamic combinatorial chemistry (DCC), as it permits to generate *in situ* and analyze libraries of up to 20 compounds and directly select the best binders to the target among them. Obviously, not every chemical reaction can be used for DCC and there are numerous restrictions for its choice. Firstly, the reaction should be reversible in order to permit the re-equilibration of DCL after target addition. Secondly, the reaction should be performed in conditions that are compatible with native structures of biological targets (i.e., aqueous conditions, around physiological pH and temperature). The last requirement for library analysis is the stability of the products in the conditions of analysis, DNA-binding experiments (e.g. melting experiments) and, last but not least, conditions of biological experiments. Here are several reactions that could be used for DCC, depending on the type of bond formed (*Table 4*):

Table 4. Types of DCC

Reaction	Structure
imine formation	$\text{R}_1\text{-C(=O)-H} + \text{H}_2\text{N-R}_2 \rightleftharpoons \text{R}_1\text{-C=N-R}_2$
hydrazone formation	$\text{R}_1\text{-C(=O)-R}_2 + \text{H}_2\text{N-NH-R}_3 \rightleftharpoons \text{R}_1\text{-C=N-NH-R}_3$
acylhydrazone formation	$\text{R}_1\text{-C(=O)-R}_2 + \text{H}_2\text{N-NH-C(=O)-R}_3 \rightleftharpoons \text{R}_1\text{-C=N-NH-C(=O)-R}_3$
reaction of metathesis	$\text{R}_1\text{-CH=CH-R}_1 + \text{R}_2\text{-CH=CH-R}_2 \rightleftharpoons \text{R}_1\text{-CH=CH-R}_2 + \text{R}_2\text{-CH=CH-R}_1$
reaction of thiol-disulfide exchange	$\text{R}_1\text{-SH} + \text{HS-R}_2 \rightleftharpoons \text{R}_1\text{-S-S-R}_2$
Michael-type addition to enones	$\text{R}_1\text{-C(=O)-CH=CH}_2 + \text{XH-R}_2 \rightleftharpoons \text{R}_1\text{-C(=O)-CH}_2\text{-CH}_2\text{-X-R}_2$ $\text{X} = \text{O, S}$

In the literature, the most exploited reactions for nucleic acid- and protein-targeted DCC are those of N-acylhydrazone formation<sup>143</sup> and disulfide exchange.<sup>126</sup> As discussed in the Introduction (Part 1.3.6), DCC of disulfides and imines have been already exploited in the field of G-quadruplex targeting. However, both disulfides and imines are not stable enough for subsequent studies or therapeutic applications: imines are not stable even in the conditions of analysis as they readily hydrolyze; disulfides are instable in biologically relevant conditions due to the reaction of disulfide exchange with glutathione abundant in living cells. Therefore, the modification of amplified imines or disulfides into more stable analogues is often required.<sup>122</sup> DCC of acylhydrazones, on the contrary, is free of this limitation as it gives direct access to compounds stable in aqueous solutions even at high temperatures. Other advantages of DCC of acylhydrazones include commercial availability or easy access to aldehyde and acylhydrazide building blocks (in particular, heterocyclic derivatives); as well as the presence of both donors and acceptors of hydrogen bonds in the acylhydrazone group, that can contribute to ligand's binding to nucleic acids (Figure 36). The most important is that the reaction of acylhydrazone exchange can occur in aqueous conditions at pH around 6 in the presence of a suitable catalyst, such as aniline or its derivatives (Figure 36).<sup>143</sup>

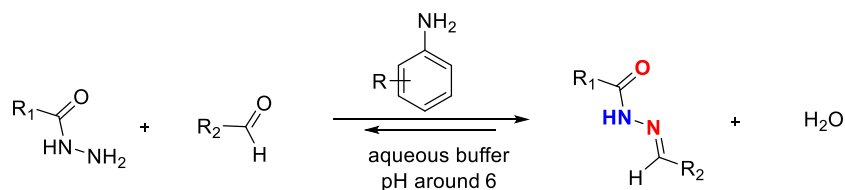


Figure 36. Reaction of acylhydrazone synthesis. H-bond acceptors in N-acylhydrazone motif are highlighted with red color, H-bond donors – with blue.

On the other hand, numerous biologically active hydrazone derivatives were reported. This scaffold is the part of several antitumoral, antimicrobial, antiviral compounds, antidepressants, anticonvulsants and analgetics.<sup>144</sup> Despite the fact that acylhydrazone derivatives are rapidly metabolized in highly acidic media, they are rather stable in plasma. This motif is found in several approved drugs (e.g., nifuroxazide and dantrolene)<sup>145</sup> as well as two promising drug candidates for cancer treatment that are currently in clinical trials: PAC-1 (NCT02355535, completed Phase-1) and aldoxorubicin (NCT02049905, completed Phase-3) (Figure 37). In addition, as mentioned in Introduction (Part 1.3.1), the hydrazone motif was already encountered in several G4-ligands (**105.3** and **106.15**, Figure 37).<sup>105, 106</sup>

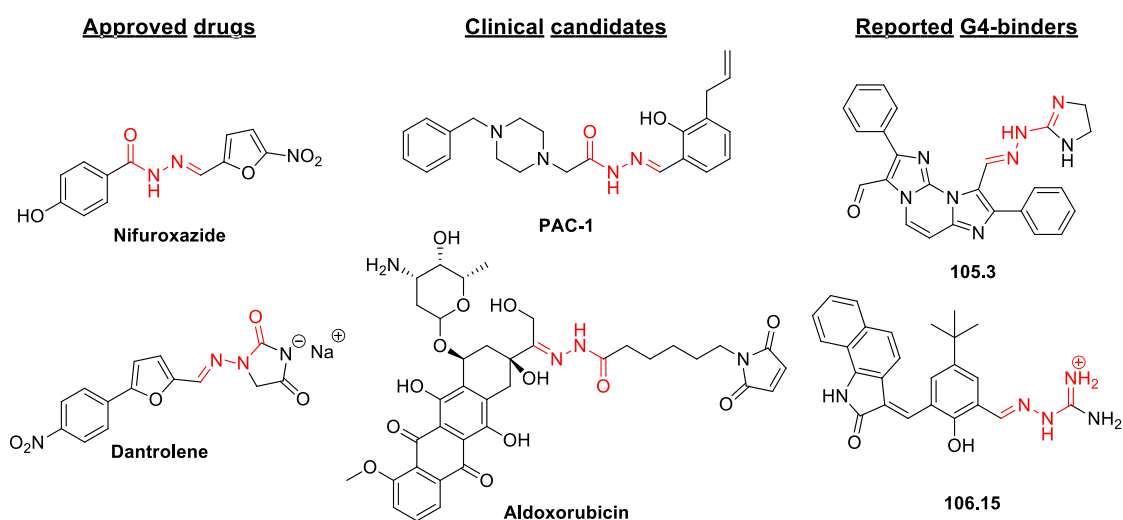
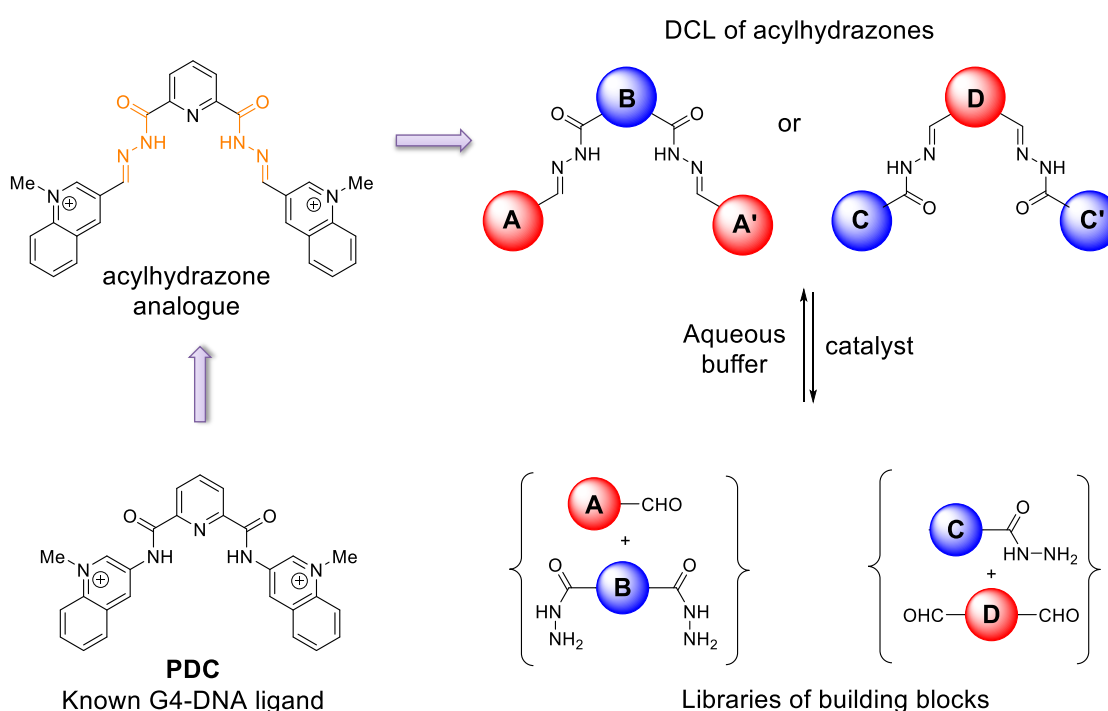


Figure 37. The presence of N-acylhydrazone motif in approved drugs, clinical candidates and reported G4-binders.

Considering the fact that DCC of acylhydrazones had not been studied in the context of G4-targeting and taking into account all aforementioned advantages of this method, we decided to implement DCC of acylhydrazone exchange. In order to adapt the acylhydrazone chemistry

for the search of G4 ligands, we designed a model ligand by replacing the amide bonds in the structure of **PDC**, an already known G4-ligand, by acylhydrazone ones (*Figure 38*). Depending on the mode of replacement, model ligands could be synthesized from a bis(acylhydrazide) and two aldehydes or, on the contrary, from a dialdehyde and two acylhydrazides. For our experiments we chose the first case (A–B–A').



*Figure 38. Design of model acylhydrazones. Structure of the known ligand PDC, its acylhydrazone analogue, two model libraries and building blocks for their synthesis.*

The project starts with the synthesis of building blocks: aldehydes (A) and bis(acylhydrazides) (B in Figure 38). In order to validate the methods that we are going to develop we also need to have a reference library of bis(acylhydrazones) with known G4 affinities. For this purpose, a set of bis(acylhydrazone) ligands will be prepared by classical organic synthesis and their binding to selected G4 targets (those described in the Introduction, Part 1.1.3-1.1.4) will be studied, using classical biophysical techniques.

The main goal of the project is to implement the comparative approach of dynamic combinatorial chemistry of acylhydrazones for G-quadruplex targeting. As mentioned in the Introduction (Part 1.3.6), the classical, i.e., comparative approach consists in generation of



dynamic combinatorial libraries in the presence and in the absence of the target (*Figure 26*). The comparison of the composition of these libraries permits to identify compounds that were amplified by the target, i.e. enriched in the target-templated library. To put this into practice, the reaction of library generation should be optimized. This includes the choice of appropriate buffer, catalyst, time of the reaction and concentrations of the reactants and target. Another important point is the optimization of the method of separation and analysis of libraries. This includes the choice of suitable column for HPLC analysis, eluents and the gradient of elution. The direct analysis of protein-templated libraries is possible because proteins easily denature with releasing bound ligands and can be separated by simple centrifugation or filtration.<sup>143, 146</sup> However, we have not found any example of direct analysis of G-quadruplex templated libraries. Therefore, we took into consideration several fall-back solutions.

The first fall-back solution is to use non-comparative approach of dynamic combinatorial chemistry. In this case, the ligand-target complex is isolated (pulled-down) from the library and bound ligands are released and analyzed afterwards. Among numerous possible ways to do this we consider either the use of streptavidin-coated magnetic beads in combination with biotin-modified oligonucleotides (as used by Balasubramanian et al.)<sup>124</sup> or gold-coated magnetic nanoparticles in combination with thiol-modified oligonucleoties (as described by Dash et al.).<sup>120</sup>

Finally, another fall-back solution is to focus on approach of classical combinatorial chemistry and perform multiple parallel combinatorial synthesis of acylhydrazones, followed by a subsequent study of their binding to nucleic acid targets by high-throughput screening techniques (such as FID or FRET-melting).

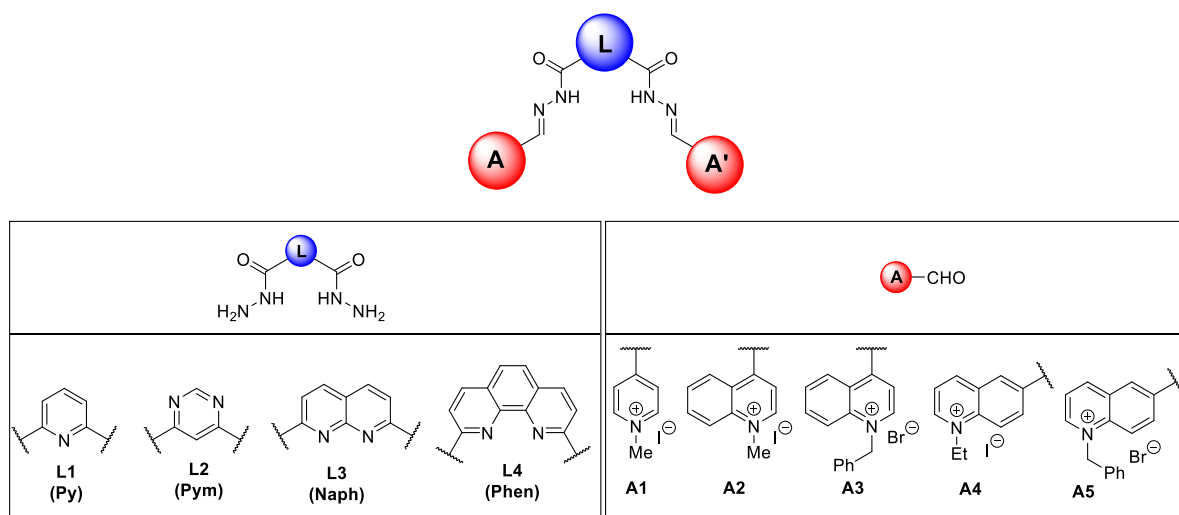
In parallel with our investigations, the team of M. Blondel (Université de Bretagne Occidentale) expressed interest in novel G4 ligands for targeting mRNA EBNA1 of Epstein-Barr virus (Part 1.1.5).<sup>38</sup> Some of ligands synthesized in the course of this work were specifically studied with respect to their binding to G4 EBNA1 and cellular activity. These results will also be presented and discussed.

# ***RESULTS AND DISCUSSION***

# 1. Synthesis and studies of model ligands

## 1.1 Synthesis of building blocks

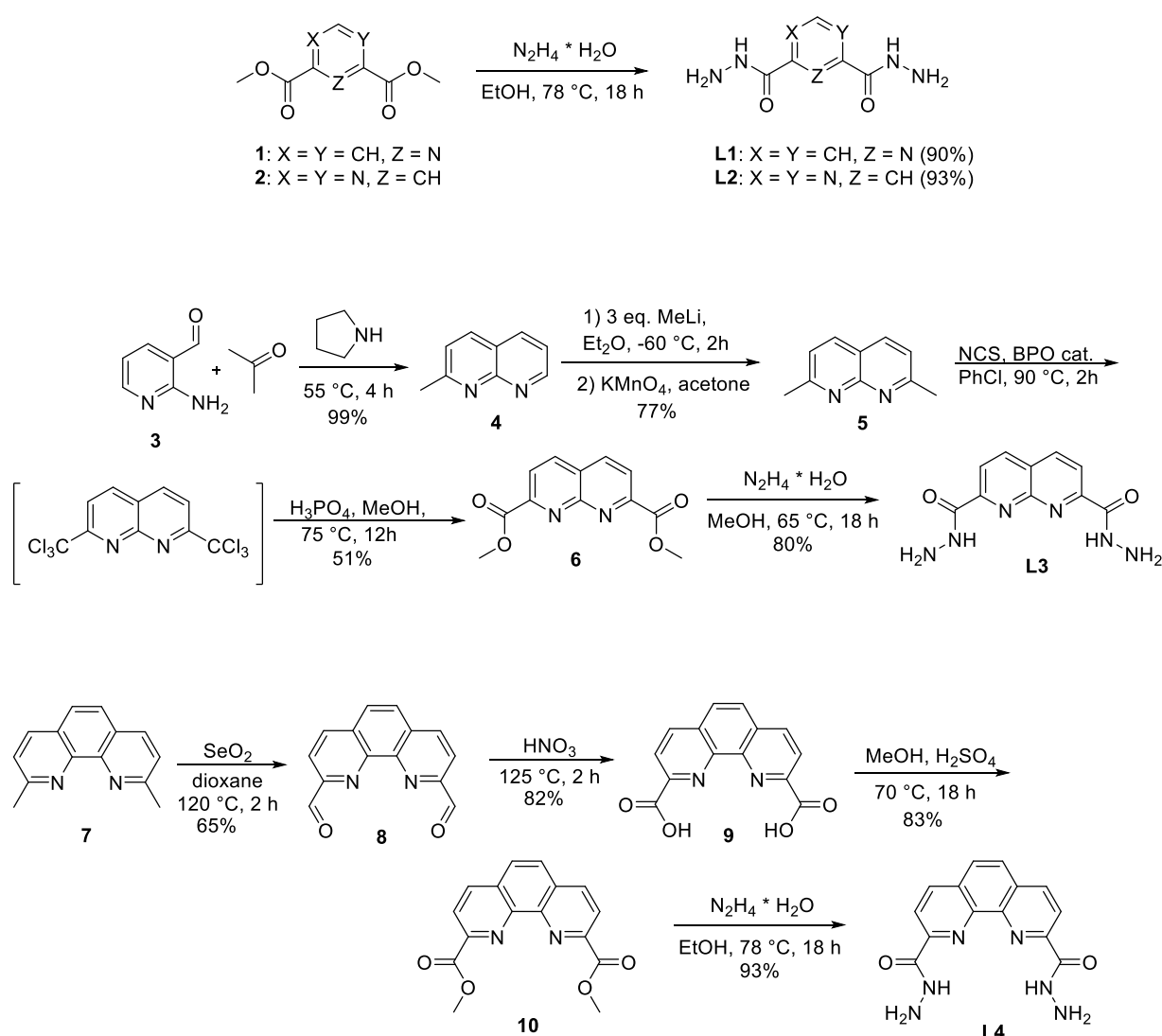
As mentioned in *AIMS OF THE WORK*, the model bis(acylhydrazone) ligand was designed from the **PDC** molecule and can be assembled from two aldehydes and one bis(acylhydrazide). The project starts with the synthesis of the first selection of building blocks which is presented on *Chart 7*. We started with the cationic aldehydes **A1–A5**, as positively charged molecules typically have higher affinity to nucleic acids due to electrostatic interactions. Selected bis(acylhydrazides) **L1–L4** were preferentially of heterocyclic nature to provide ligands with additional hydrogen bonding functions as well as some biological activity.



*Chart 7. First selection of building blocks for bis(acylhydrazone) synthesis; in this part, A=A'.*

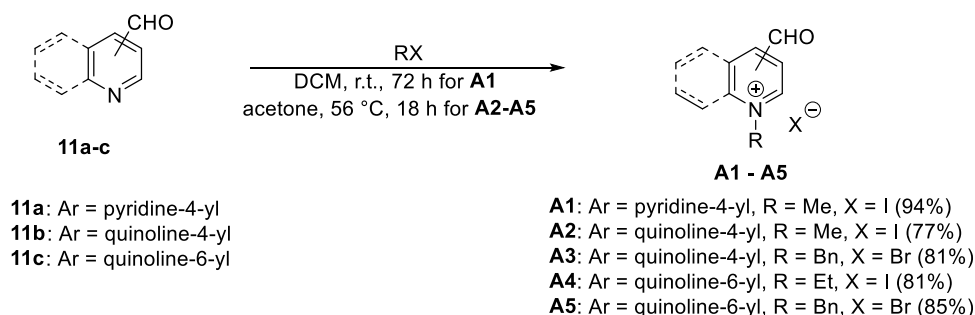
Bis(acylhydrazides) **L1–L4** were prepared from the dimethyl ester precursors by the reaction of hydrazinolysis. Methyl esters of pyridine-2,6- and pyrimidine-4,6-dicarboxylic acids **1** and **2** were purchased from commercial suppliers, while naphthyridine and phenanthroline analogues **6** and **10** were prepared in-house (*Scheme 5*). The synthesis of naphthyridine derivative **6** involves the construction of the heterocycle **4** from 2-aminopyridine-3-carbaldehyde **3** and acetone in excellent yield. Subsequent incorporation of a methyl group by MeLi and oxidation of the obtained dihydroderivative by  $\text{KMnO}_4$  led to 2,7-dimethyl-1,8-naphthyridine **5** in 77% yield. It was halogenated in a radical reaction with *N*-

chlorosuccinimide following a described procedure.<sup>147</sup> Subsequent hydrolysis of hexachloro-2,7-dimethyl-1,8-naphthyridine and esterification of obtained acid gave the desired dimethyl-1,8-naphthyridine-2,7-dicarboxylate **6** in 51% yield. The synthesis of bis(acylhydrazide) **L4** started with an oxidation of 2,9-dimethyl-1,10-phenanthroline to dialdehyde **8** with  $\text{SeO}_2$  in a 65% yield and then to dicarboxylic acid **9** with  $\text{HNO}_3$  in a 82% yield.<sup>148</sup> After the esterification step, the dimethyl ester **4** was transformed into **L4** through a hydrazinolysis reaction in excellent yield.



Scheme 5. Synthesis of bis(acylhydrazide) building blocks **L1–L4**

The cationic aldehydes **A1–A5** were prepared by simple alkylation of the commercially available pyridine-4-, quinoline-4- and quinoline-6-carbaldehydes (*Scheme 6*). 4-Formyl-1-methylpyridin-1-ium iodide **A1** was obtained in 94% yield by stirring the aldehyde **11a** with the excess of iodomethane in DCM at room temperature as described in the literature.<sup>149</sup> Quinoline-4-carbaldehyde **11b** and quinoline-6-carbaldehyde **11c** were alkylated by an excess of iodomethane, iodoethane, or benzyl bromide in boiling acetone, which is the typical conditions for alkylation of quinoline derivatives,<sup>150,151</sup> to give building blocks **A2–A5** in good yields ranging from 77 to 85%.



*Scheme 6. Synthesis of aldehyde building blocks A1–A5.*

Being very electrophilic, these aldehydes can readily form hydrates. In aqueous solutions, aldehydes **A1–A3** exist predominantly in the hydrated form: in <sup>1</sup>H NMR spectra the peak of the aldehyde group is absent; instead, there is a characteristic peak at 6.2 ppm of a proton of the geminal diol group (*Figure 39a*). In contrast, aqueous solutions of aldehydes **A4–A5** contain only about 30% of the hydrated form (*Figure 39b*), due to different positions of aldehyde groups with respect to quaternary nitrogen atom. As we will see later, the hydration does not change the reactivity of the aldehydes, since all aldehydes **A1–A5** were able to form acylhydrazide products, including in aqueous solutions.

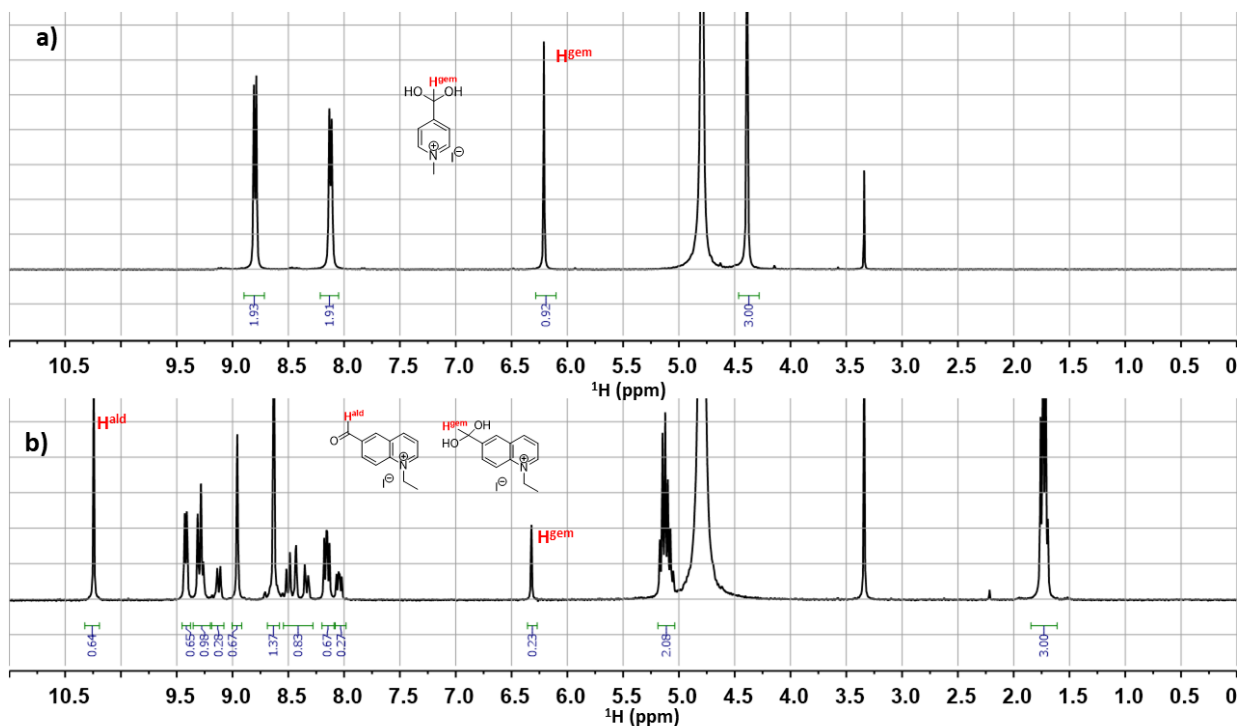
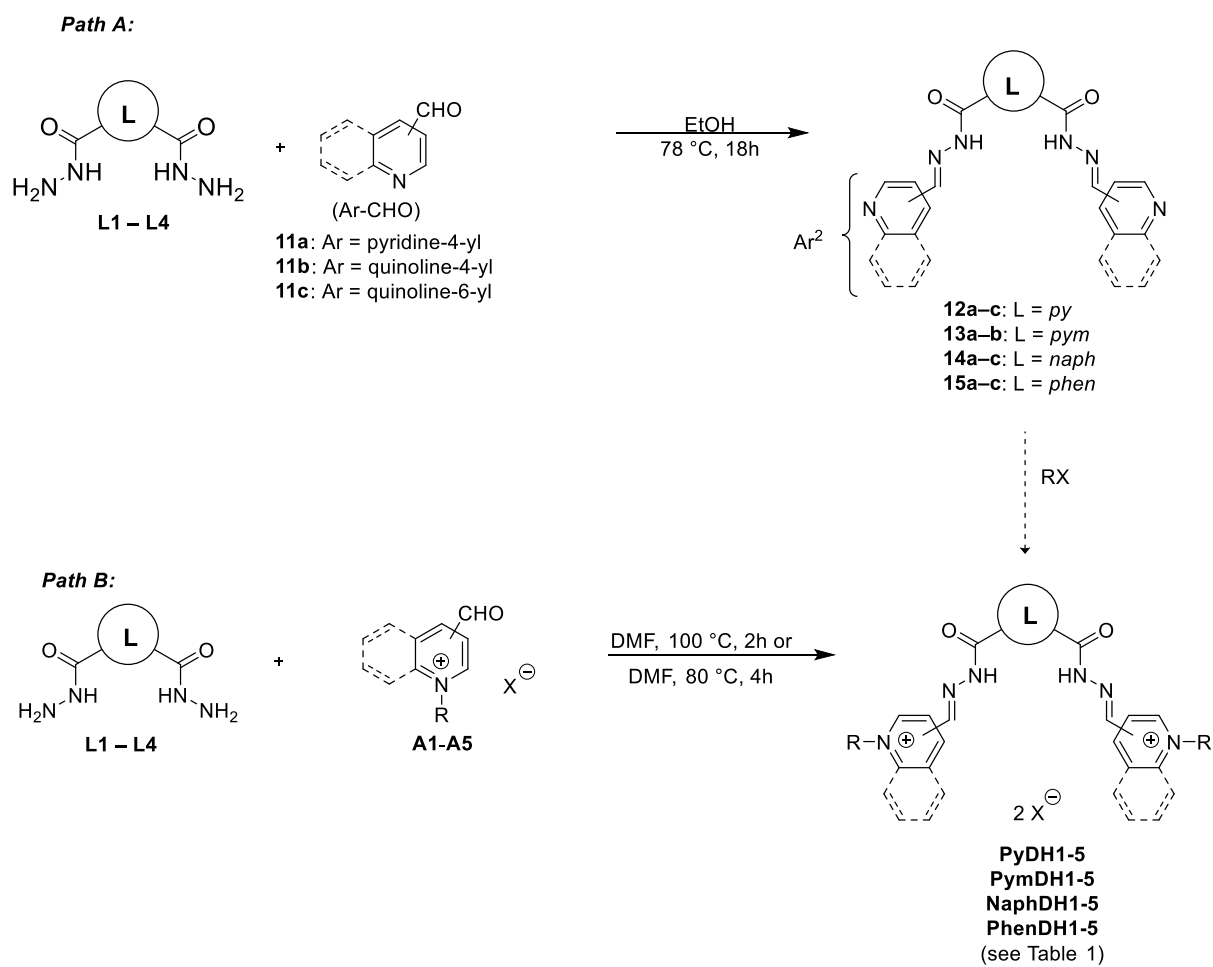


Figure 39. NMR spectra (300 MHz) of aldehydes **A1** (a) and **A4** (b) in  $\text{D}_2\text{O}$  (+MeOH for reference).

## 1.2 Preparative synthesis of model bis(acylhydrazones)

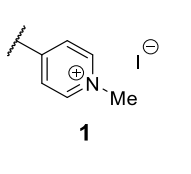
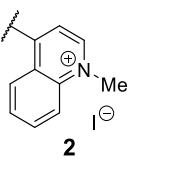
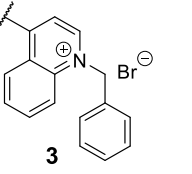
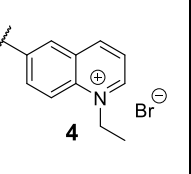
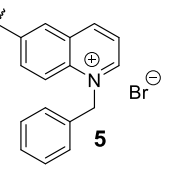
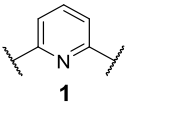
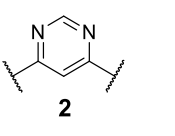
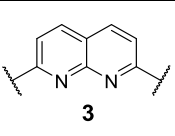
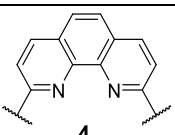
For the synthesis of bis(acylhydrazones) initially a two-step approach was explored, namely the condensation of aldehydes with acylhydrazides, followed by alkylation of the formed acylhydrazone (Path A, *Scheme 7*). This approach is similar to that typically used for the synthesis of **PDC** and **PhenDC3**.<sup>61</sup> The reaction of condensation was successfully carried out in boiling methanol or ethanol with yields of 55-90%. However, the reaction of quaternization performed with an excess (150 eq.) of alkylation agent in some cases was incomplete and the impurity of monoalkylated acylhydrazone (up to 20% as per  $^1\text{H}$  NMR) could not be removed by recrystallization of the product. So we turned our attention to the method of one-step synthesis of model ligands from acylhydrazides and alkylated aldehydes (Path B, *Scheme 7*). This reaction is complete after 2 hours in DMF at 100 °C (for certain cases: 80 °C, 4 hours), without the use of catalyst. The products were then purified by a recrystallization in MeCN/ $\text{H}_2\text{O}$  (1:1). One product (**NaphDH2**) could not be recrystallized because of its poor solubility but revealed sufficiently pure according to  $^1\text{H}$  NMR, LC/MS and combustion analysis

data. This way, twenty symmetrical bis(acylhydrazones) were synthesized in good yields (Table 5).



Scheme 7. Synthesis of model bis(acylhydrazones).

Table 5. Structures and acronyms of the synthesized model ligands, reaction conditions and yields.

A \ L					
	<b>PyDH1</b> , [OR28], A1-L1-A1 90% <sup>a</sup>	<b>PyDH2</b> , [OR19], A2-L1-A2 87% <sup>a</sup>	<b>PyDH3</b> , [OR20], A3-L1-A3 89% <sup>a</sup>	<b>PyDH4</b> , [OR21], A4-L1-A4 92% <sup>a</sup>	<b>PyDH5</b> , [OR40], A5-L1-A5 52% <sup>a</sup>
	<b>PymDH1</b> , [OR31], A1-L2-A1 97% <sup>a</sup>	<b>PymDH2</b> , [OR27], A2-L2-A2 92% <sup>a</sup>	<b>PymDH3</b> , [OR127], A3-L2-A3 93% <sup>a</sup>	<b>PymDH4</b> , [OR129], A4-L2-A4 96% <sup>a</sup>	<b>PymDH5</b> , [OR128], A5-L2-A5 97% <sup>a</sup>
	<b>NaphDH1</b> , [OR29], A1-L3-A1 78% <sup>a</sup>	<b>NaphDH2</b> , [OR30], A2-L3-A2 81% <sup>a</sup>	<b>NaphDH3</b> , [OR22], A3-L3-A3 90% <sup>a</sup>	<b>NaphDH4</b> , [OR23], A4-L3-A4 75% <sup>a</sup>	<b>NaphDH5</b> , [OR42], A5-L3-A5 80% <sup>b</sup>
	<b>PhenDH1</b> , [OR33], A1-L4-A1 76% <sup>a</sup>	<b>PhenDH2</b> , [OR34], A2-L4-A2 78% <sup>a</sup>	<b>PhenDH3</b> , [OR35], A3-L4-A3 87% <sup>a</sup>	<b>PhenDH4</b> , [OR36], A4-L4-A4 72% <sup>a</sup>	<b>PhenDH5</b> , [OR41], A5-L4-A5 91% <sup>b</sup>

Conditions of synthesis: <sup>a</sup> DMF, 100 °C, 2 h; <sup>b</sup> DMF, 80 °C, 4 h.

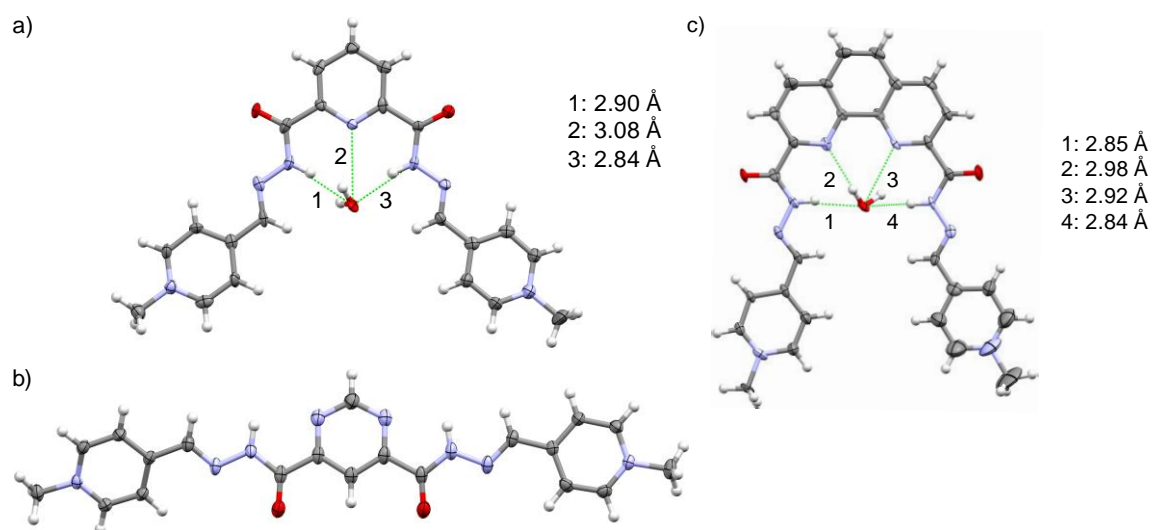
### 1.3 Structural characterization of bis(acylhydrazone) compounds

X-ray quality crystals of **PyDH1** (2I<sup>-</sup>), **PymDH1** (2I<sup>-</sup>), and **PhenDH1** (2I<sup>-</sup>) could be obtained from acetonitrile–water (in the case of **PymDH1** (2I<sup>-</sup>), a simultaneous formation of yellow and orange polymorphs was observed, both belonging to the  $\bar{P}1$  space group but differing by the packing of cations). Single-crystal diffraction analysis revealed that in all three compounds the organic cations were essentially planar, with minimal twisting at the level of terminal heterocyclic rings breaking the two-fold molecular symmetry (Figure 40). In all cases, both *N*-acylhydrazone groups adopted *E*-antiperiplanar conformation. Remarkably, **PyDH1** and **PhenDH1** revealed co-crystallized water molecules located in between the two acylhydrazone groups and forming hydrogen bonds with both amide NH groups (**PyDH1**:  $d(\text{NH}\cdots\text{O}) = 2.84$  and  $2.90$  Å,  $\alpha(\text{N-H}\cdots\text{O}) = 150^\circ$ ; **PhenDH1**:  $d(\text{NH}\cdots\text{O}) = 2.85$  and  $2.84$  Å,  $\alpha(\text{N-H}\cdots\text{O}) \approx 170^\circ$ ). In addition, in **PyDH1** the water molecule formed a weak hydrogen bond with the pyridine nitrogen ( $d(\text{N}\cdots\text{O}) = 3.08$  Å, Figure 40a), whereas in **PhenDH1** it formed a stronger, bifurcated



hydrogen bonds with both phenanthroline nitrogens (mean  $d(\text{N}\cdots\text{O}) = 2.95 \text{ \AA}$ , *Figure 40c*). The coordination of water molecules thus determines the inward-folded, V-shaped conformation of **PyDH1**, similar to what was observed with related, charge-neutral pyridine-2,6-bis(acylhydrazones),<sup>152</sup> as well as the U-shaped conformation of **PhenDH1** which is further stabilized due to the number and favorable arrangement of hydrogen bonds. In contrast, **PymDH1** (*Figure 40b*) was found to adopt a linear conformation, partially stabilized by intramolecular bonds between the amide NH groups and nitrogen atoms of the pyrimidine core ( $d(\text{N}\cdots\text{N}) = 2.66 \text{ \AA}$ ). Although this compound also crystallized along with water molecules, the latter were found to coordinate to amide CO groups, on one hand, and iodide anions, on the other hand (not shown).

It may be hypothesized that the described interactions with water molecules persist in aqueous solutions, governing the molecular shape of bis(acylhydrazone) ligands. In this context, V- or U-shaped derivatives **PyDHn** and **PhenDHn**, presumably, have a more favorable structure for binding to G4-quadruplex structures due to a more complete overlap with G-tetrads, as compared to linear analogues **PymDHn**.



*Figure 40. Solid-state structures of a) **PyDH1** ( $2\Gamma$ )  $\times$  2  $\text{H}_2\text{O}$ , b) **PymDH1** ( $2\Gamma$ )  $\times$  2  $\text{H}_2\text{O}$   $\times$  MeCN, and c) **PhenDH1** ( $2\Gamma$ )  $\times$  2  $\text{H}_2\text{O}$   $\times$  MeCN, from single-crystal X-ray diffraction analysis. Non-bound water, acetonitrile molecules and counter-ions were omitted for clarity. CPK atom colors; green lines and labels indicate intermolecular hydrogen bonds with crystallized water molecules and the corresponding  $\text{N}\cdots\text{O}$  distances*

## 1.4 Evaluation of affinity of bis(acylhydrazones) to G-quadruplexes by FRET-melting

The interaction of 20 model bis(acylhydrazone) derivatives with five representative G-quadruplex-forming sequences and one hairpin duplex was studied in FRET-melting experiment. The chosen G4-forming sequences were the three variants of the human telomeric sequence (*24TTG*, *25TAG* and *22CTA*) and two sequences from promoter of the oncogene *c-myc* (*Pu24T* and *myc22*, Table 6). As mentioned in the *Introduction*, where their structures are discussed in more details, these sequences represent different topologies, namely hybrid-1 (*24TTG*), hybrid-2 (*25TAG*), antiparallel (*22CTA*), simple parallel (*myc22*) and snap-back parallel (*Pu24T*). For the FRET-melting assay we used the sequences, labeled with fluorescein (5') and TAMRA (3'). As a reference we used the benchmark G4-ligand **PhenDC3**. The experiments were performed in three different buffers (K100, K10 and K1), that contain different concentrations of K<sup>+</sup>-ions, which were chosen such as that the temperature of melting of G-quadruplex alone was approximately the same for each oligonucleotide (in our case, around 60 °C, Table 6). Thus, buffer K1 was used for *myc22* and *Pu24T*, K10 for *22CTA*, and K100 for *25TAG*, *24TTG* and *hp2* control.

Table 6. Sequences used for fluorescence melting experiments and their melting temperatures in the absence of ligands.<sup>a, b</sup>

Acronym	Sequence (5' – 3')	Fold	Length	T <sub>m</sub> , °C K100	T <sub>m</sub> , °C K10	T <sub>m</sub> , °C K1
<b>24TTG</b>	TTGGGTTAGGGTTAGGGTTAGGGA	G4 hybrid-1	24	61.3	52.9	57.3
<b>25TAG</b>	TAGGGTTAGGGTTAGGGTTAGGGTT	G4 hybrid-2	25	58.2	43.4	39.1
<b>22CTA</b>	AGGGCTAGGGCTAGGGCTAGGG	G4 antiparallel	22	63.3	56.9	56.0
<b>myc22</b>	TGAGGGTGGGTAGGGTGGGTAA	G4 parallel	22	84.9	65.6	61.0
<b>Pu24T</b>	TGAGGGTGGGTAGGGTGGGGAAGG	G4 parallel	24	88.3	79.6	60.0
<b>hp2</b>	GTTATATCT-18heg-TGATATAAC	hairpin duplex	19	65.7	67.8	68.2
<b>ds26</b>	CAATCGGATCGAATTCGATCCGATTG	Duplex (competitor)	26			

<sup>a</sup> Systems, chosen for subsequent FRET-melting experiment with ligands, are highlighted in orange;

<sup>b</sup> Composition of K100 buffer: 10 mM LiAsO<sub>2</sub>Me<sub>2</sub>, 100 mM KCl, pH 7.2; K10 buffer: 10 mM LiAsO<sub>2</sub>Me<sub>2</sub>, 10 mM KCl, 90 mM LiCl, pH 7.2; K1 buffer: 10 mM LiAsO<sub>2</sub>Me<sub>2</sub>, 1 mM KCl, 99 mM LiCl, pH 7.2.

The experiments with ligands were performed using 0.2 μM of double-labeled oligonucleotide and 1 μM of ligand, standard conditions for FRET-melting assay.<sup>134</sup> Many (but not all) of our

compounds revealed a capacity to bind and stabilize different G4-forming nucleic acid sequences, as revealed by high  $\Delta T_m$  values, i.e., increase of melting temperatures in the presence of ligands (Figure 41).

Among the tested compounds, compounds **PyDH2**, **NaphDH2**, **PhenDH4** showed good stabilization of all tested G4-sequences:  $\Delta T_m > 20$  °C for *myc22*, *25TAG*, *22CTA* and *24TTG* and  $\Delta T_m > 15$  °C for *Pu24T*. Of note, for the sequence *Pu24T* the temperatures of stabilization were not as high compared with others sequences, that could be explained by its initially very high stability. **PhenDH1**, **PhenDH2**, **PhenDH3** and **PhenDH5** also provide significant increase of temperature of melting: **PhenDH1** for *25TAG* and *24TTG*, **PhenDH2** for *myc22*, *25TAG* and *24TTG*, **PhenDH3** for *25TAG* and *24TTG* and **PhenDH5** for *myc22*, *25TAG*, *22CTA* and *24TTG*. At the same time, pyrimidine derivatives **PymDH1–PymDH5** do not significantly affect the temperature of melting of G4-forming nucleic acid sequences. Having in hands both types of ligands, namely those that stabilize G4, and those without such a capacity, within the same library of chemically similar compounds (identical charge, same functional groups) is beneficial for the project as it makes possible a “calibration” of the following DCC experiments and molecular modeling data.

Only a small drop of melting temperatures observed in the presence of double-stranded competitor *ds26* proves that ligands bind selectively to G4. A control experiment with double-strand hairpin DNA sequence (*F-hp2-T*) shows the same: none of our ligands significantly affects the stability of a double-stranded DNA.

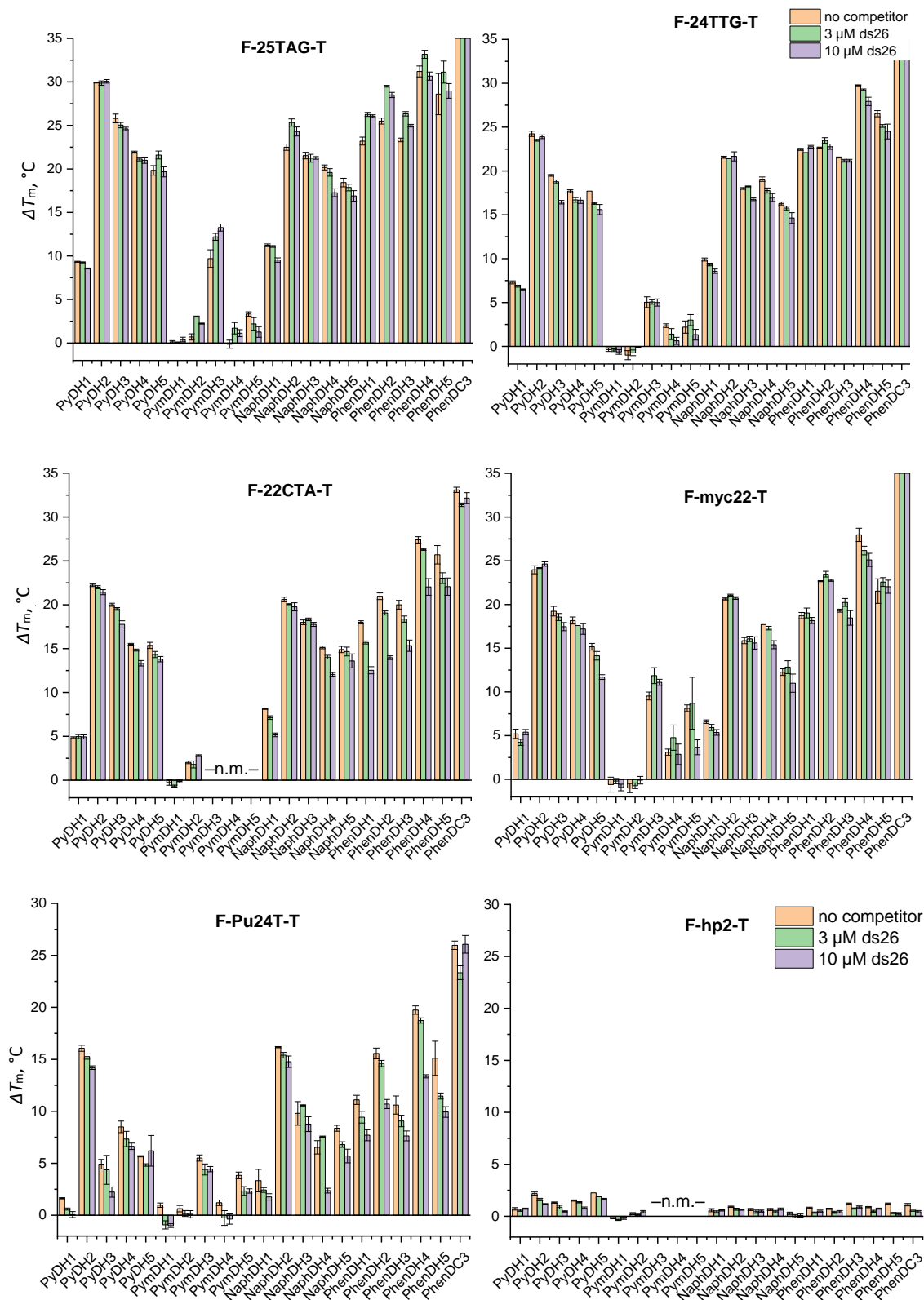


Figure 41. Thermal stabilization of selected DNA G-quadruplexes (0.2  $\mu$ M) by tested compounds (1.0  $\mu$ M), assessed by fluorescence melting experiments in the absence (orange bars) or in the presence of duplex DNA competitor ds26 (green bars: 3  $\mu$ M, violet bars: 10  $\mu$ M). N.m. – not measured.

Taking into account all above mentioned observations, we can make following conclusions:

- The structure of the central heterocycle (**L**) plays a crucial role in determining the binding capacity of bis(acylhydrazone) to G-quadruplexes. All derivatives of phenanthroline, as well as the majority of derivatives of pyridine were found to be “good” ligands. On the contrary, derivatives of pyrimidine-4,6-diacylhydrazide, **PymDH1–PymDH5**, possessed almost no stabilizing activity with respect to all sequences tested, even though the structures of pyridine (Py) and pyrimidine (Pym) heterocycles are very similar. Presumably, it is due to the linear shape of pyrimidine derivatives that is stabilized by intramolecular H-bonds between amide hydrogens and N-atoms of pyrimidine ring. In contrast, the derivatives of other heterocycles had U- or V-shaped form, more favorable for stacking with external G-tetrads (*Figure 40*).
- The structure of the lateral heterocycle also plays an important role. Among U-shaped ligands, the derivatives of quinolinium carbaldehydes **A2–A5** provide higher thermal stabilization than derivatives of pyridinium carbaldehyde **A1**. This is in accordance with a common concept that suggests an increase of affinity of the ligand to a G4-structure with the augmentation of its planar aromatic surface (due to better  $\pi$ - $\pi$  stacking).
- The structure of the alkyl chain at the quaternary atom of nitrogen does not drastically affect the binding affinity. However, in the majority of cases methyl derivatives showed the highest stabilization, revealing no additional stabilization with the incorporation of benzyl group despite additional  $\pi$ -surface area.

## **1.5 Molecular modeling (in collaboration with Dr Liliane Mouawad, M.Sc. 2 project of Denis Leclerc)**

In order to simplify the search of potential ligands for G-quadruplexes, we made an attempt to exploit structure-based virtual screening. Molecular modeling for G4-ligands is popular in the field, with docking being the most used approach.<sup>153</sup> Molecular docking consists in finding the best orientation of two molecules in the complex. For each such an orientation (pose) the docking score, a value that represents free energy of binding, or  $\Delta G$ , it is calculated. However,

a comparison of molecular modeling scores with experimental affinities is seldom performed. With our ligands in hand, we had a unique opportunity to make such a correlation. For this work, we used ligand-docking software Glide (Schrödinger, Inc.), with the default parameters. Several high-resolution structures complexes of G-quadruplexes with ligands were available in PDB (*Figure 42*). The first structure, 5MVB, was the complex of telomeric G-quadruplex *wtTel26* with the ligand **Auoxo6**, that contained two atoms of gold (*Figure 42A*).<sup>154</sup> The sequence *wtTel26* differed from *25TAG* only by the presence of an additional thymine base in 5'. It adopted exclusively hybrid-2 topology in the complex with **Auoxo6**. In comparison to the free G-quadruplex, capping of the top tetrad by nucleotides was absent in the complex; however, adenine base (A3) was stacked upon the ligand. The second studied structure, 2L7V, was the complex of oligonucleotide *myc22* with two molecules of quindoline (*Figure 42B*).<sup>155</sup> The G4 in this complex was of parallel topology. **Quindoline** molecules alone with bases adjacent to external tetrads (A6 and T23) formed a plane capping these tetrads. These interactions were stabilized by additional capping of these planes by bases G5 (on the top) and A24 (in the bottom). 2A5R was a complex *Pu24T* oligonucleotide with a porphyrin ligand **TMPyP4** (*Figure 42C*).<sup>31</sup> The ligand was stacked upon the top tetrad. Its binding did not affect the snap-back parallel topology of the G4, but prevented the stacking of the pair A3•A12. Instead, two bases (T1 and G2) were capping the ligand. The last structure that we used for docking is 2MGN, a complex of *Pu24T* with **PhenDC3** (*Figure 42D*).<sup>62</sup> As in the previous complex, the ligand was stacked on the top tetrad. Due to its large aromatic surface and rather stable U-shape conformation, **PhenDC3** efficiently covered the tetrad, interacting with all guanines of the external quartet. No additional capping was reported.

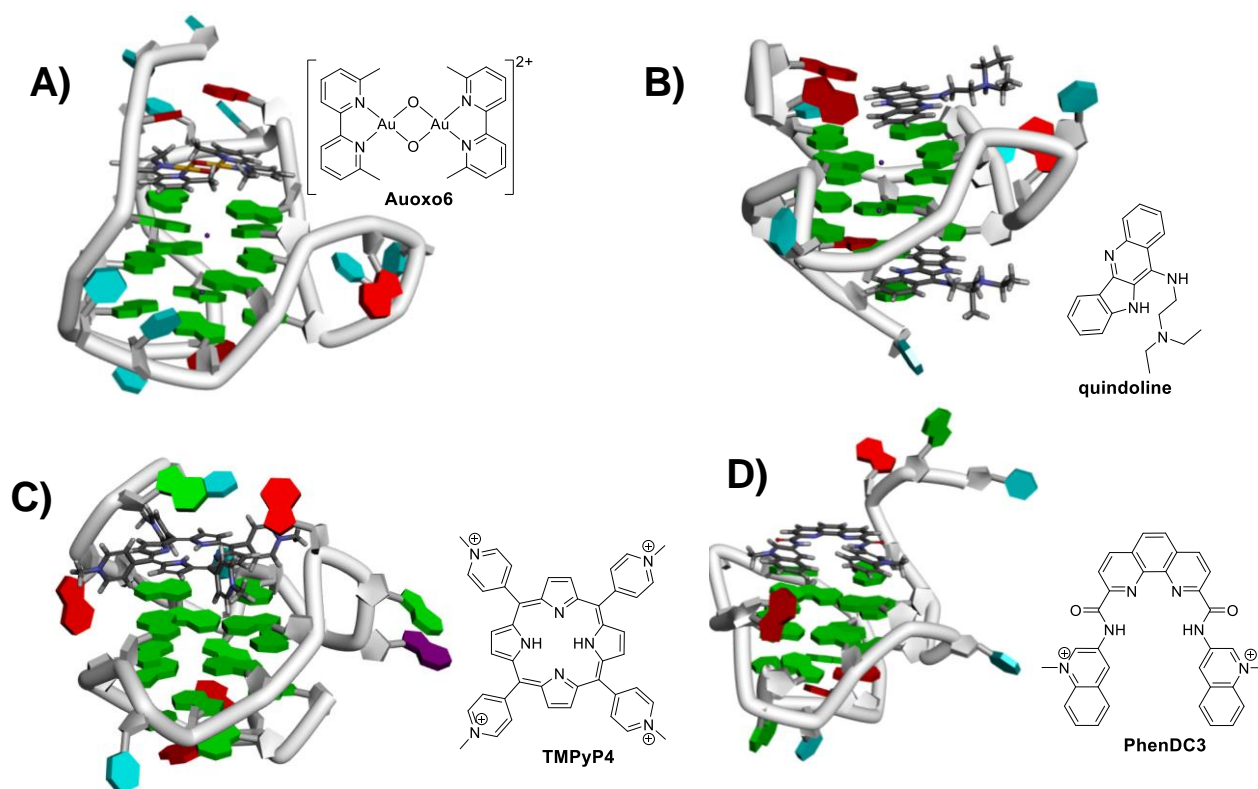


Figure 42. PDB-structures of G4-nucleic sequences with ligands used for docking studies: a) 5MVB – wtTel26 with **Auoxo6**; b) 2L7V – myc22 with **quindoline**; c) 2A5R – Pu24T with **TMPyP4**; d) 2MGN – Pu24T with **PhenDC3**.

However, as these structures were solved by NMR, the positions of the  $K^+$  ions, coordinated by G-tetrads, were not identified. For docking calculations, bound ligands were removed and  $K^+$  ions were placed in the center of mass of the eight oxygen atoms of the two stacked tetrads. We started with re-docking the original ligands of each PDB structure (Figure 42) into their own binding sites. 30 poses were generated for each ligand and compared to the pose of the PDB file. In each G4 system, two poses were considered: the top-rank pose that had the lowest docking score (or gScore, which represents free energy of binding) and the “best” pose that was the one with the smallest RMSD, Root Mean Square Deviation from the original PDB pose (Figure 43).

- **5MVB**: As gold atoms are not recognized by Glide, the re-docking could not be performed for this system.
- **2L7V**: For the top cavity, the top-rank pose had an RMSD of 7.268 Å and a gScore of -9.058 kcal/mol. The best pose had an RMSD of 7.320 Å and a gScore of -7.360 kcal/mol. For the bottom cavity, the top-rank pose had an RMSD of 9.338 Å and a gScore of -

7.552 kcal/mol, and the best pose had an RMSD of 9.035 Å and a gScore of -7.160 kcal/mol. Both RMSDs and scores for bottom cavity were higher than those observed for the top cavity showing that the ligand was positioned even further from the reference pose. In both cavities there was no pose close to the experimental positions of the ligand. However, since experimentally the ligand adopts very different positions in the top and in the bottom cavities and interacts mostly with the closest G-tetrad, we may consider that there is a large number of possibilities in positioning such a small molecule.

- **2A5R:** The top-score docking pose had an RMSD of 8.215 Å and gScore of -13.493 kcal/mol. The best pose had an RMSD of 7.478 Å and gScore of -13.342 kcal/mol. Both top-score and best poses had rather high RMSD, but this is mainly due to the rotation of the methylpyridines. Also, both poses were shifted and not centered on the G-tetrad anymore. This could be explained by the protonation error during the preparation of the ligand by ligprep software: the docked porphyrin had four protonated nitrogen atoms at its center, when experimentally only two of four were protonated.
- **2MGN:** The top-score pose obtained by docking had an RMSD of 8.098 Å and gScore of -7.708 kcal/mol and the best pose had an RMSD of 6.681 Å and Score -5.407 kcal/mol. Interestingly, these two docked poses of **PhenDC3** were not planar: one of ligand's arms was raised to establish a hydrogen bond with the oxygen of the sugar of one guanine of the tetrad. This pose seems to be less constrained than that of the PDB structure. To verify the reason why this ligand adopts experimentally a rather constrained structure, we checked the NMR constraints available from the PDB. It appeared that, since the molecule is symmetrical, the constraints were applied indifferently on both arms, impeding the formation of the H-bond observed in Glide poses. To verify if these poses are nevertheless plausible, the distances corresponding to the NOE (Nuclear Overhauser effect) constraints have been calculated for each pose given by Glide and the number of constraints violation calculated. One pose was found to respect the NMR constraints as satisfactory as the PDB pose while establishing the H-bond with the guanine, thus being more plausible.



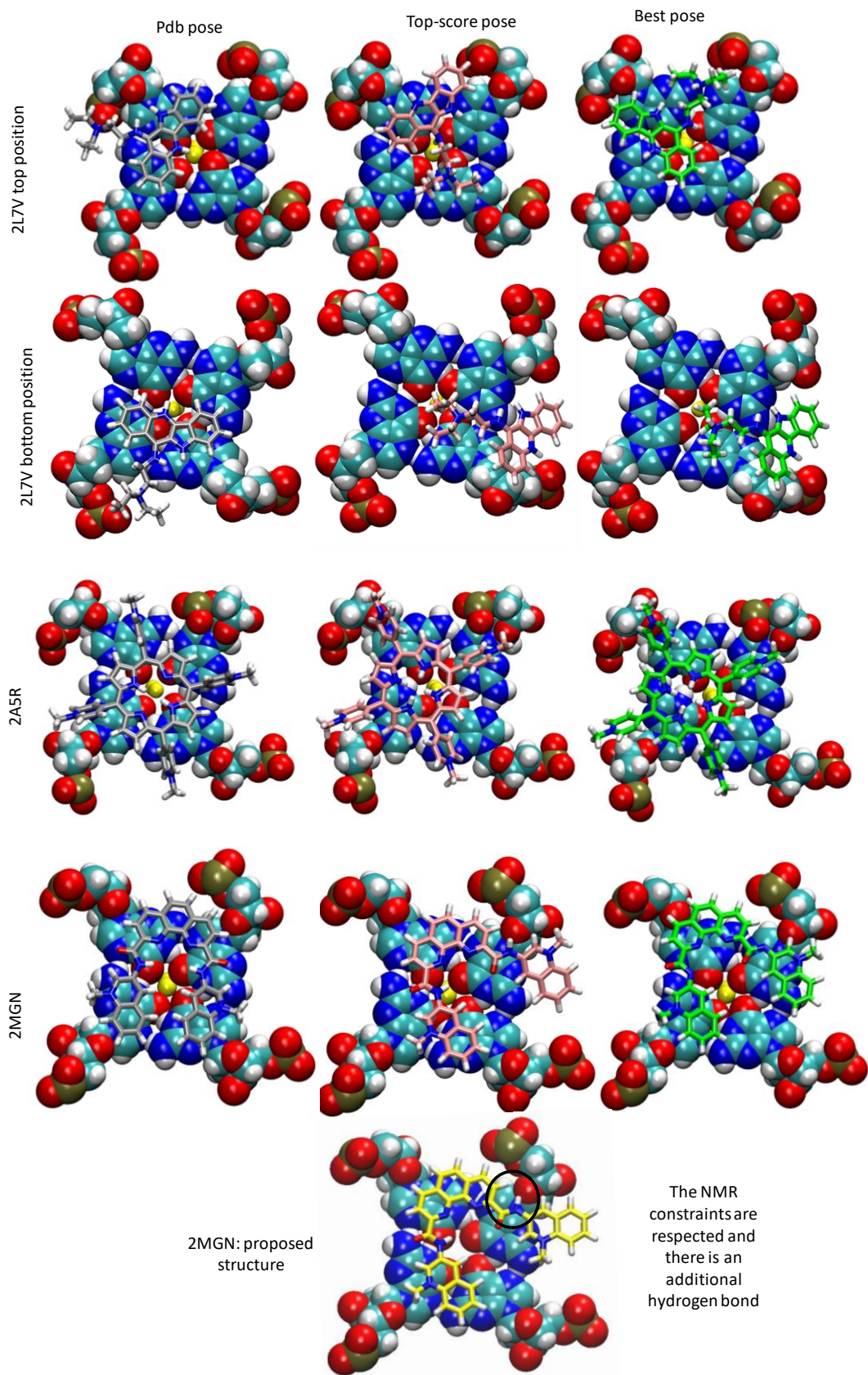


Figure 43. PDB, top-score and best positions for ligands from corresponding PDB structures (2L7V, 2A5R and 2MGN) obtained by Glide as well as proposed structure for 2MGN.

Next, 17 compounds of the prepared library, derivatives of pyridine (**PyDH1–PyDH5**), pyrimidine (**PymDH1–PymDH2**), naphthyridine (**NaphDH1–NaphDH5**) and phenanthroline (**PhenDH1–PhenDH5**), as well as two benchmark G4 ligands **PDC** and **PhenDC3** were docked to the above mentioned G4 structures using the same settings of Glide software, and gScore values were collected. Then we tried to find a correlation between gScore and experimental affinity values for our 19 molecules to the above-mentioned structures (*Figure 44*). As a proxy of the latter, we considered the temperature of stabilization of G4 in fluorescence melting experiments ( $\Delta T_m$ ) since those data were readily available. Unfortunately, although the general trend, where compounds with higher  $\Delta T_m$  values tend to have lower gScores was respected, the correlation found between experimental and calculated affinities was very poor ( $R^2$  between 0.115 and 0.330, *Figure 44*). This observation questions the significance of published report on G4 ligand identification by virtual screening / docking programs.<sup>153</sup>

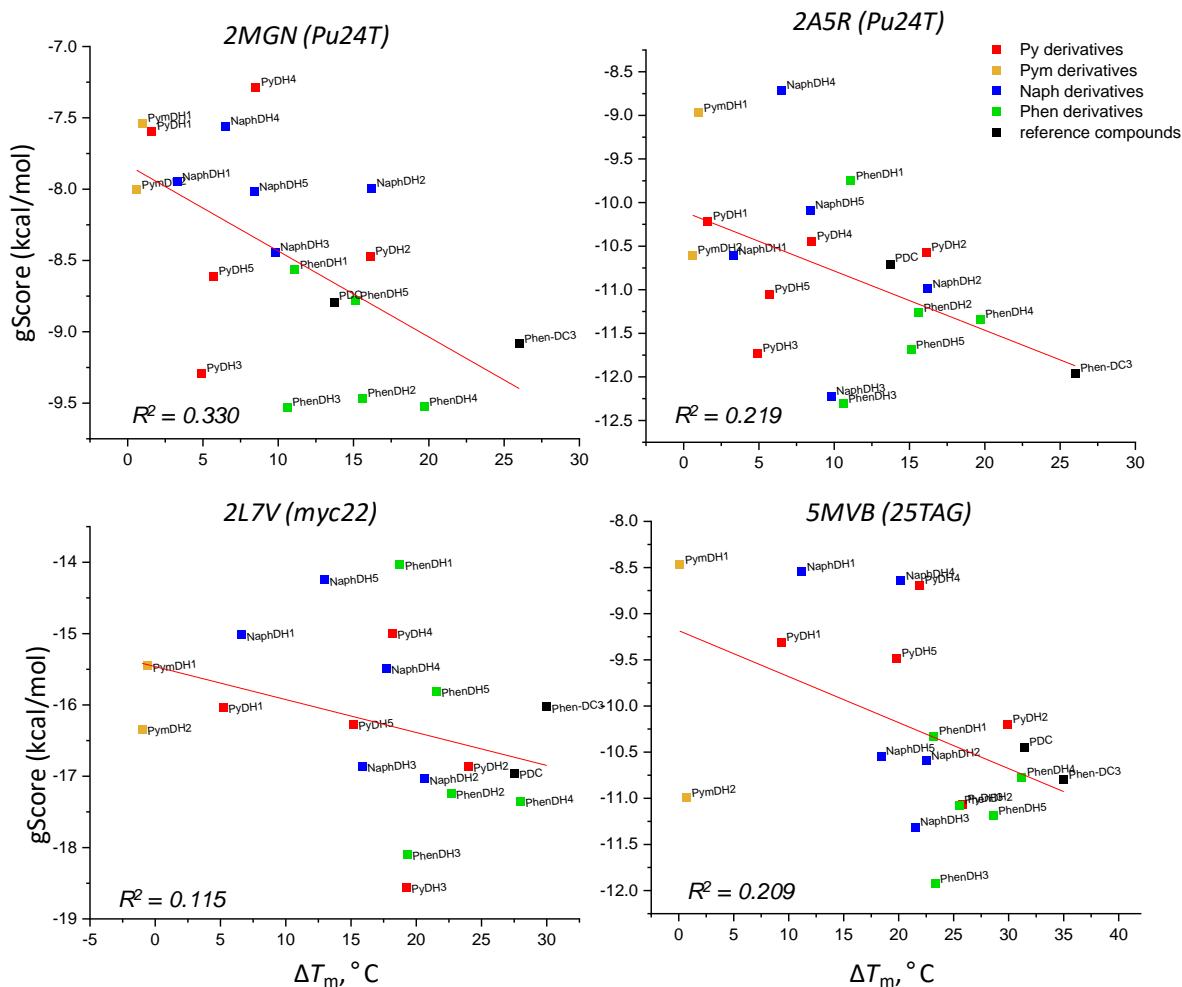


Figure 44. Comparison of the calculated affinities with the experimental ones. For each G4, the Glide score (gScore) is plotted versus the experimental ligand stabilization ( $\Delta T_m$ , °C) of the tested ligands (red – derivatives of pyridine, orange – pyrimidine, blue – naphthyridine, green – phenantroline, black – reference compounds **PDC** and **PhenDC3**).

Among the possible reasons of poor correlation between gScore and  $\Delta T_m$  we may assume:

- 1) Docking is a rather coarse method of molecular modeling. The use of force field to represent the G-quadruplex is not completely accurate. While it can be well suited to simulate binding pockets of defined structure in proteins, the G-quadruplexes are highly dynamic structures that can readily reorganize upon ligand binding. A structure of G4 from the complex in PDB makes sense only for the ligand, for which it was resolved. The molecule of another structure can induce a rearrangement of the loops and flanking regions and in some cases may even lead to a complete change of the topology of the G4. A good illustration of above statement is that we found the best

correlation ( $R^2 = 0.33$ ) between gScore and  $\Delta T_m$  values for PDB 2MGN, the complex of *Pu24T* with **PhenDC3**, a ligand structurally very similar to our set.

- 2) Ligprep, the procedure used for the optimization of the conformation of the ligand prior to docking, does not take into account the solvent molecules. The majority of our ligands are of U-shape that is stabilized by the H-bonds through the molecule of water (*Figure 40a,c*). This important feature may be omitted when the conformation of the ligand is optimized.
- 3) Even though  $\Delta T_m$  obtained in FRET-melting experiment is sufficient for ranking of the set of ligands according to their affinity to G4 structures, it may not correctly reflect the value of affinity of the molecule to the target. Multiple factors are not taken into account in this assay (stoichiometry, interaction with unfolded structure,  $\Delta H$  and  $\Delta S$  contributions to binding). Finally, in FRET-melting experiment we measure temperature of thermal stabilization of the structure while in biological systems these conditions (60-90 °C) will never be the case.
- 4) In addition, we noticed a systematic bias due to overestimation of some type of interactions. It mostly concerns derivatives of **A3**, in particular **PyDH3** and **PhenDH3**. Low gScores systematically have been assigned to these compounds, however, in FRET-melting they usually had only moderate  $\Delta T_m$ . To better understand the origin of this bias we looked at the best pose, generated for **PyDH3** with *Pu24T* G-quadruplex from 2MGN PDB (*Figure 45*). As we can see, the ligand is very well positioned on the top tetrad of the G4 and its lateral benzyl moieties are protruding in the grooves. We may hypothesize that even though geometrically it is an optimal position, in reality benzyl groups are too hydrophobic to interact with negatively charged grooves. If we exclude these two compounds from the set, the  $R^2$  value will increase up to 0.526 for 2MGN.

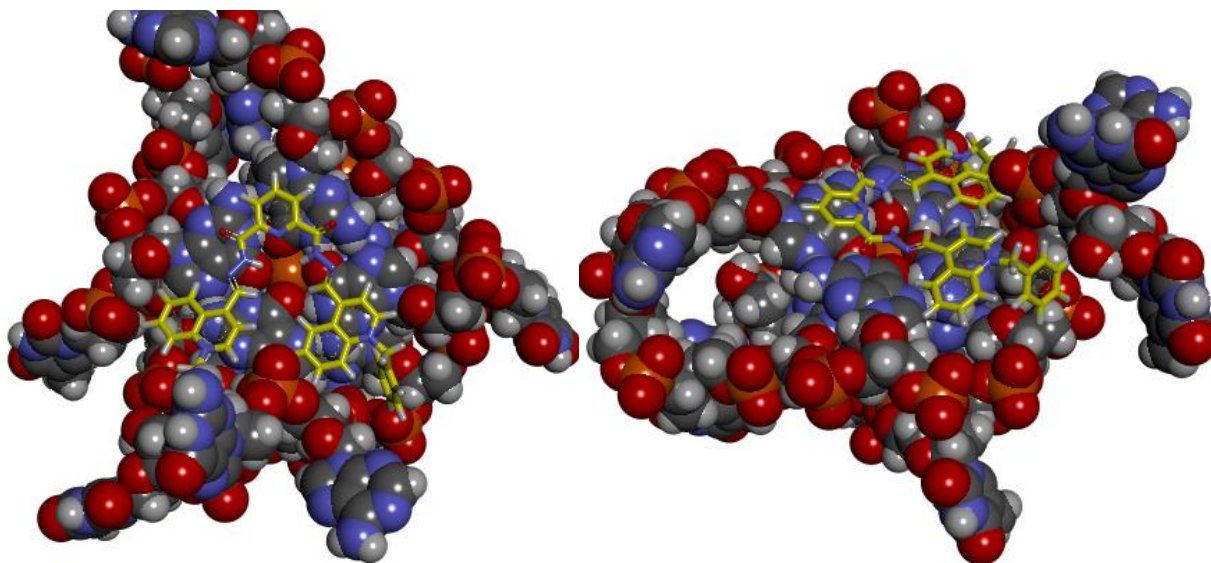


Figure 45. The best pose of **PyDH3** with Pu24T from 2MGN PDB file.

Taking into account all above mentioned drawbacks, we decided not to use molecular docking for the follow-up work.

## 1.6 Conclusions

We started this work with the synthesis of the model library of acylhydrazone compounds. For this purpose, the building blocks, namely bis(acylhydrazides) **L1–L4** and cationic aldehydes **A1–A5** were synthesized from available precursors. Two synthetic routes of acylhydrazone synthesis were tested and the most efficient approach was selected for the synthesis of the model library of 20 acylhydrazones.

The capacity of model acylhydrazones to thermally stabilize G-quadruplexes of different topologies was evaluated using the FRET-melting assay. We found that most derivatives with central pyridine (**Py**), naphthyridine (**Naph**) or phenanthroline (**Phen**) units were efficient G4 binders, in contrast to pyrimidine (**Pym**) counterparts. Structural characterization of compounds using X-ray confirmed a significantly different molecular geometry of pyrimidine derivatives (linear), in contrast to U- or V-shape of analogues with other central heterocycles.

We also attempted to use the docking method of molecular modelling for our project. To validate this method, firstly, we analyzed three reported complexes of G4s with small molecules from the Protein Data Bank and re-docked the ligands from these complexes into their own binding sites. Secondly, we docked structures of our model acylhydrazones to four G4 structures from PDB. We then attempted to find the correlation between melting temperatures and gScore obtained from docking. Unfortunately, due to numerous reasons such a correlation was rather poor so we have not continued docking experiments in our project.

Of note, other methods of molecular modeling (such as molecular dynamics or flexible docking) may be better suited for simulation of G4-ligand interaction. However, these methods are computationally expensive (MD) or not available with the software in our disposition (Glide).

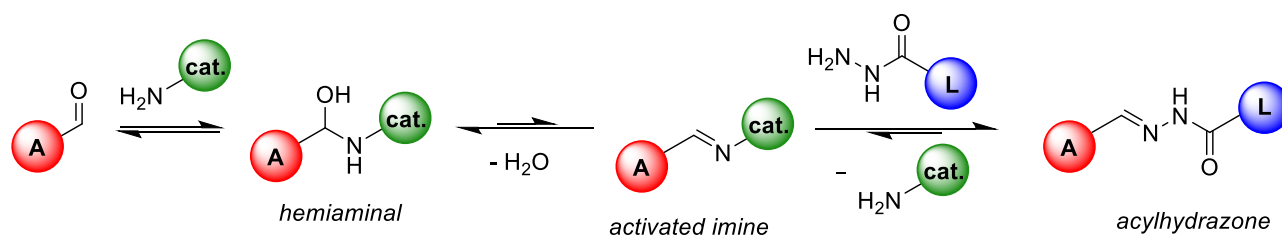
## 2. Dynamic combinatorial chemistry

Dynamic combinatorial chemistry (DCC) is a supramolecular approach that exploits reversible chemical reactions to generate dynamic combinatorial libraries (DCL) of products under thermodynamic control.<sup>122,156</sup> This method is particularly well-suited for the discovery of ligands and receptors, since the addition of an external species (a “target”, or a “template”) shifts the dynamic equilibrium towards the formation of the product(s) having the highest affinity to the target.<sup>157</sup> While DCC has now been well established for protein targeting,<sup>123,158</sup> its application in the field of nucleic acids is much more restricted.<sup>159,160</sup> Along these lines, Balasubramanian<sup>124–126</sup> and Ulven<sup>161</sup> labs exploited DCC for the discovery (or optimization) of G4-DNA ligands; of note, all these approaches relied on the reaction of disulfide exchange and involved only small number of components, typically, a central scaffold possessing a G4 affinity and a set of side chains. More recently, Dash et al. utilized a DCC approach based on the reversible formation of imines followed by pull-down and reduction, to identify novel carbazole-based G4-DNA ligands.<sup>128</sup> However, as discussed in *AIMS OF THE WORK*, a serious drawback of imine-based DCC is that it requires the conversion of imines, identified from DCL analysis, into stable amine analogues whose affinity and selectivity to the target may be significantly different. Therefore, there is a need for the development of alternative chemistries suitable for nucleic acid-targeted DCC.

### 2.1 Preliminary experiments and optimization of conditions for DCC generation

The advantages of the use of reaction of acylhydrazone formation in DCC were discussed in *AIMS OF THE WORK*. It was also mentioned that this reaction is possible to perform in water in near physiological pH with the help of suitable catalysts. This possibility was first reported back in 1962 when the catalysis of semicarbazone formation by anilines was described.<sup>162</sup> In this reaction aniline was found to work as a nucleophilic catalyst *via* the rate-limiting formation of a Schiff base. The latter, being highly reactive, forms adducts with semicarbazide.

Since then, numerous nucleophilic catalysts with enhanced reactivity, such as anthranilic acids,<sup>163</sup> *p*-phenylenediamine<sup>164</sup> or 2-(aminomethyl)benzimidazoles,<sup>165</sup> were reported. The general mechanism of acylhydrazone formation, catalyzed by nucleophilic catalysts, is presented on *Scheme 8*.



*Scheme 8. The mechanism of nucleophilic catalyst assisted acylhydrazone formation.*

To assess the feasibility of synthesis of bis(acylhydrazones) in the conditions compatible with native structures of G4-DNA, we started with the study of the reaction of equimolar amounts of the cationic aldehyde **A1** with pyridine-2,6-dihydrazide **L1** in the aqueous DCC buffer (100 mM NH<sub>4</sub>OAc, 1.5 mM KCl, pH 6.4) in the presence of four anthranilic acid derivatives as nucleophilic catalysts **cat.a–d** (*Figure 46a*).<sup>163</sup> HPLC analysis of reaction mixtures (*Figure 46b*) demonstrated that in presence of 5-methoxyanthranilic acid (**cat.b**) the reaction almost reached equilibrium after 24 hours, whereas **cat.a** and **cat.c** were less efficient, and only traces of products were detected when the reaction was performed in the presence of **cat.d**. In view of this result, **cat.b** was employed as a catalyst in the following DCC experiments.



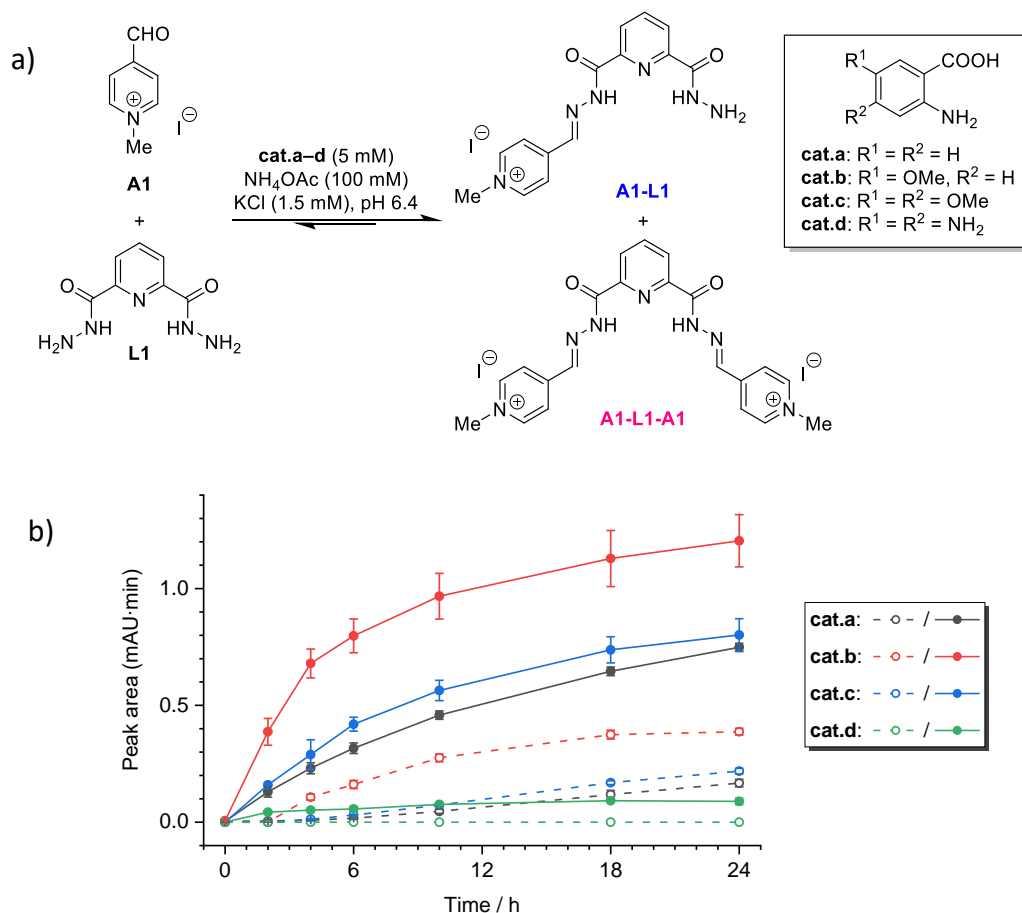


Figure 46. Optimization of the catalyst for the synthesis of cationic bis(acylhydrazones). a) Reaction scheme; b) time course of formation of **A1-L1** (empty circles, dashed lines) and **A1-L1-A1** (filled circles, solid lines) upon the reaction of **A1** and **L1** (40  $\mu\text{M}$  each) in DCC buffer in the presence of catalysts **cat.a-d** (5 mM). Data from three independent experiments.

Next, we generated a first combinatorial library (DCL1) through the reaction of two aldehydes (**A1** and **A2**) and three dihydrazides (**L1-L3**). The employed conditions (10 mM of **cat.b**, incubation time of 24 hours) led to the formation of 15 products (Figure 47a), including six mono-acylhydrazones and nine bis(acylhydrazones), all of which could be successfully separated by HPLC (Figure 47b). The peaks were assigned by using pairwise combinations of reagents. We also investigated the conversion of the reagents (Table 7) to make sure that the reactivity of building blocks is approximately equal and there will no bias in DCC experiments because of substrates with low reactivity.

All six symmetric bis(acylhydrazones) present in this library (pink in Figure 47) were synthesized in the preparative fashion in the previous part of this work and their G4-binding properties are already known (RESULTS AND DISCUSSION, Part 1), enabling a subsequent

validation of the G4-templated DCC approach. This library contained both “good” **A2-L1-A2** (**PyDH2**), **A2-L3-A2** (**NaphDH2**) and “poor” **A1-L1-A1** (**PyDH1**), **A1-L2-A1** (**PymDH1**) ligands.

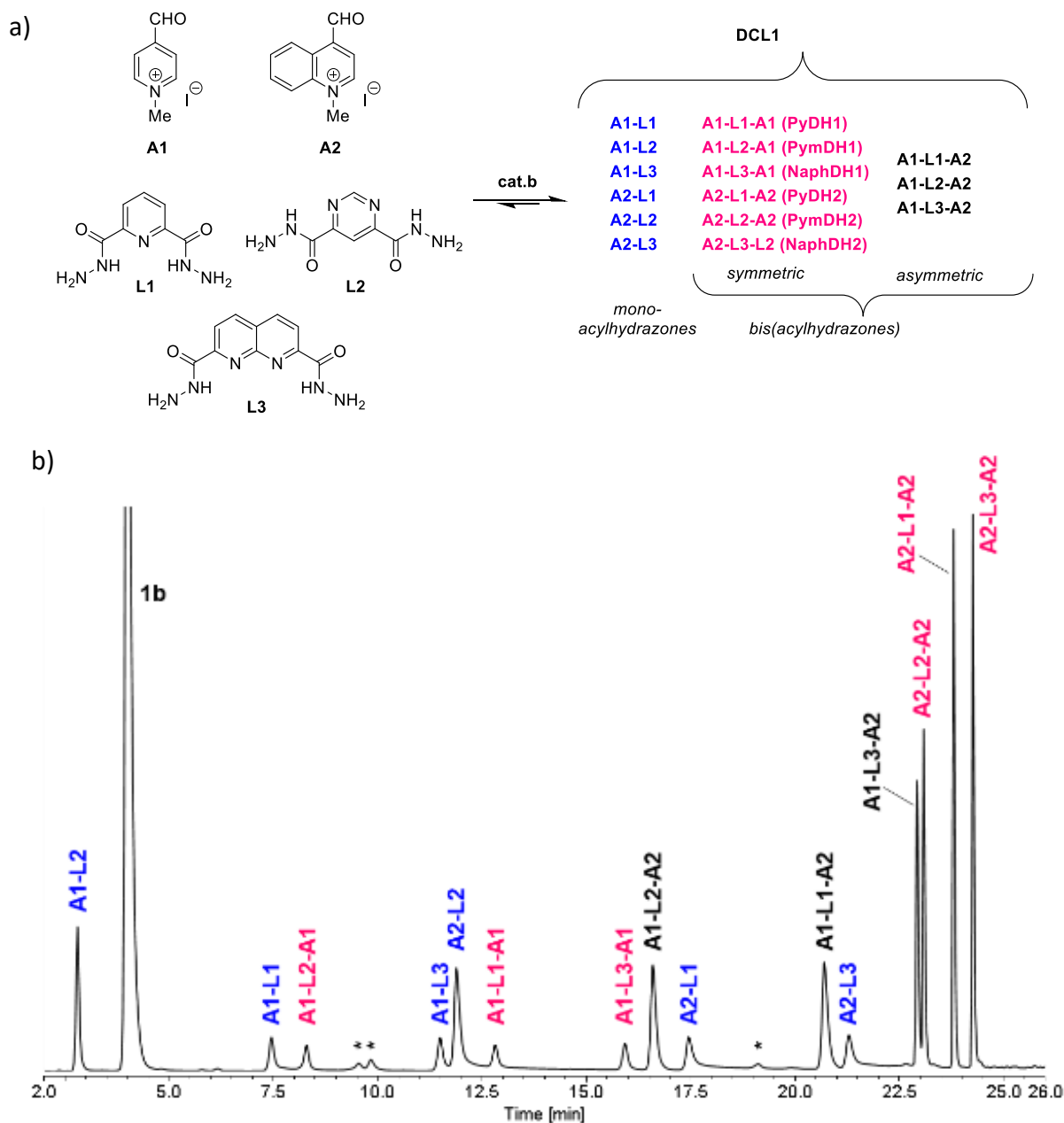


Figure 47. a) Schematic representation and b) HPLC profile of a 15-membered DCL1 obtained upon 24-h reaction of two cationic aldehydes with three dihydrazides.  $c(\mathbf{A1}) = c(\mathbf{A2}) = 120 \mu\text{M}$ ,  $c(\mathbf{L1}) = c(\mathbf{L2}) = c(\mathbf{L3}) = 40 \mu\text{M}$ ,  $c(\text{cat.b}) = 10 \text{ mM}$  in  $100 \text{ mM NH}_4\text{OAc}$ ,  $1 \text{ mM KCl}$  buffer,  $\text{pH } 6.0$ . Peaks labelled with asterisks in b) could not be assigned. Conditions: column Atlantis T3,  $3 \mu\text{m}$ ,  $3 \times 100 \text{ mm}$  (Waters), eluent A:  $\text{H}_2\text{O} + 0.01\% \text{ TFA}$ ; eluent B:  $\text{ACN} + 0.01\% \text{ TFA}$ ; flow rate:  $0.8 \text{ mL min}^{-1}$ ; detection wavelength:  $310 \text{ nm}$ . Gradient: from 0 to 80% B.

Table 7. Initial composition of DCL1 and conversion rate of the corresponding reactants.

Reactant	c (μM)	conv. (%) <sup>a</sup>
<b>A1</b>	120	81
<b>A2</b>	120	61
<b>L1</b>	40	97
<b>L2</b>	40	88
<b>L3</b>	40	≈100
<b>cat.b</b>	5 mM	
<b>DNA</b>	5 <sup>b</sup>	

<sup>a</sup> Conversion, calculated as per relative peak area of the reactants before and after addition of the catalyst and incubation for 24 h (in the absence of DNA). <sup>b</sup> In templated libraries.

One of advantages of acylhydrazone-based DCC is that it gives access to stable products that can be directly analyzed by HPLC, opposing to imine-based DCC where “freezing” step borohydride reduction is usually needed.<sup>128</sup> Thus, protein-templated DCLs of acylhydrazones are typically analyzed by direct HPLC injection; in HPLC conditions, the protein template denatures and bound ligands are instantly released, allowing their quantification and comparison with the “blank” library obtained in the absence of the template.<sup>143,146,166</sup> To verify the applicability of this method to G4-templated DCLs, we performed HPLC injections of equimolar mixtures (25 μM each) of pure, extemporarily prepared **A2-L1-A2** (i.e., **PyDH2**, a strong G4-DNA binder, *RESULTS AND DISCUSSION, Part 1*) and **A1-L2-A1** (a very poor G4-DNA binder)<sup>167</sup> in the absence and in the presence of 50 μM of a G4-DNA oligonucleotide (*Pu24T, Table 8*). The presence of G4-DNA had no effect on **A1-L2-A1** that was eluted in a sharp peak. However, the peak of **A2-L1-A2** was shifted and strongly distorted in the presence of G4-DNA (*Figure 48: (1) and (2)*). A possible explanation of this phenomenon is that the complex of G4-DNA with strong binders is not instantly dissociated, but retained in the column head after HPLC injection, leading to a gradual dissociation occurring during the HPLC run and manifested as peak tailing. Indeed, G4-DNA structures were shown to be extremely stable in a range of non-physiological conditions, including up to 50% acetonitrile;<sup>168,169</sup> moreover, the dissociation constants of best G4-DNA binders (including **A2-L1-A2**) are typically found in the low-nanomolar range, i.e., one to three orders of magnitude lower than that of typical protein inhibitors discovered by protein-templated DCC,<sup>146,166</sup> and this strong binding can lead to additional stabilization of the folded form of G4-DNA.

Table 8. DNA oligonucleotides used in this work.

Acronym	Sequence (5' → 3')	Structure	Use			
			DCC <sup>a</sup>	FM <sup>b</sup>	FT <sup>c</sup>	MS <sup>d</sup>
Pu24T	TGAGGGTGGTGGAGGGTGGGGAAGG	G4 (snap-back parallel)	X	X	X	X
myc22	TGAGGGTGGGTAGGGTGGGTAA	G4 (parallel)		X	X	
25TAG	TAGGGTTAGGGTTAGGGTTAGGGTT	G4 (hybrid, Type 2)		X	X	X
22CTA	AGGGCTAGGGCTAGGGCTAGGG	G4 (anti-parallel)		X	X	
hp2	GTTATATCT-HEG-TGATATAAC <sup>e</sup>	hairpin	X	X	X	
dT <sub>22</sub>	TTTTTTTTTTTTTTTTTTTTTTTT	single-stranded	X		X	
dT <sub>6</sub>	TTTTTT	single-stranded				X <sup>f</sup>
ds26	CAATCGGATCGAATTCGATCCGATTG	self-complementary duplex		X <sup>g</sup>		

<sup>a</sup> DCC pull-down experiments: the oligonucleotides were 5'-modified with biotin. <sup>b</sup> Fluorescence melting experiments: oligonucleotides were labelled with 5'-FAM and 3'-TAMRA. <sup>c</sup> Isothermal fluorimetric titrations: oligonucleotides were labeled with 5'-Cy5. <sup>d</sup> Native ESI mass spectrometry. <sup>e</sup> HEG = hexa(ethylene glycol) linker. <sup>f</sup> Internal reference. <sup>g</sup> Unlabeled ds26 was used as a competitor in fluorescence melting experiments.

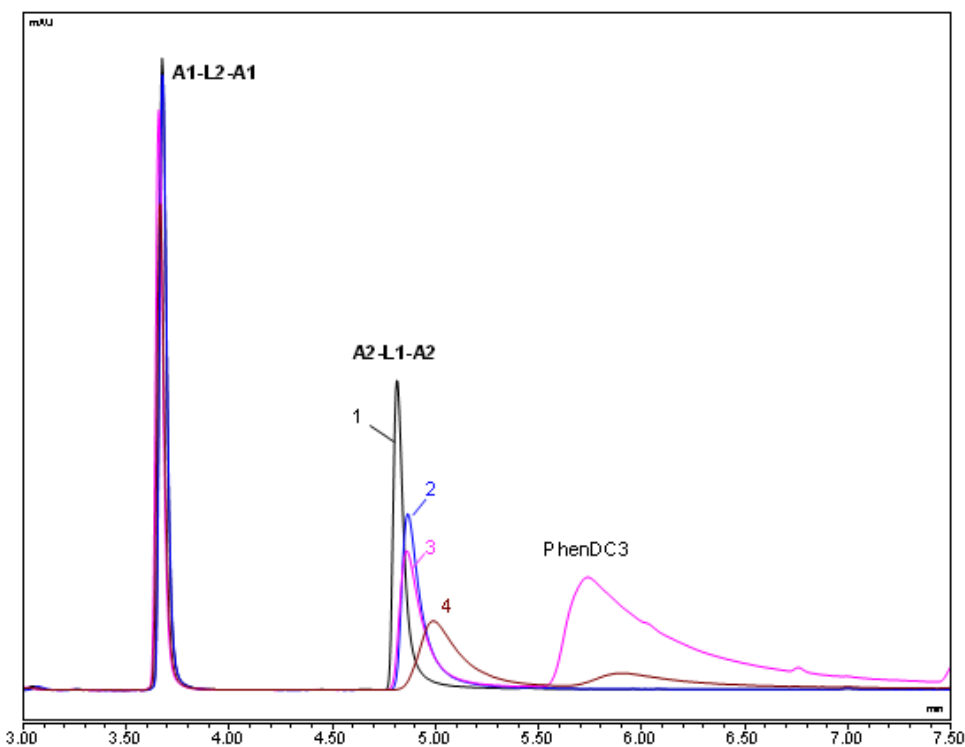


Figure 48. HPLC profiles of mixtures of **A1-L2-A1** and **A2-L1-A2** (25  $\mu$ M each in 100 mM NH<sub>4</sub>OAc, 1.5 mM KCl buffer, pH 6): (1) in the absence of G4-DNA; (2) in the presence of Pu24T (50  $\mu$ M); (3) same as (2) followed the addition of **PhenDC3** (125  $\mu$ M); (4) same as (2) followed addition of equal volume of formamide prior to injection. Samples were incubated for 1 h at room temperature prior to analysis. The injection volume was 5  $\mu$ L in (1–3) and 10  $\mu$ L in (4), to compensate for the addition of formamide.

We attempted to release bound ligands by adding an excess (5 molar equivalents) of **PhenDC3** as a competitor prior to HPLC analysis, as was described by Ernst et al. (in this work, the authors were blocking binding sites of a protein with excess of highly affine ligand to show that the shift in composition of DCC was not due to non-specific interactions)<sup>170</sup> or by supplementing the samples with equal volume of formamide. However, in either case the original peak shape and area of **A2-L1-A2** could not be restored (*Figure 48: (3) and (4)*), making peak integration unreliable. These observations implied that the comparative approach using the direct HPLC analysis could not be employed for reliable assessment of G4-DNA-templated libraries.

## **2.2 Pull-down experiment with streptavidin-coated magnetic beads**

We hypothesized that the issues discussed above could be overcome by implementation of “pull-down” (or “capturing”) protocol when DNA is biotinylated and can be removed from the solution with the help of streptavidin-coated paramagnetic beads. Instead of comparing templated and non-templated libraries, this approach seeks to identify the best binders pulled-down by DNA-covered beads from “frozen” libraries, following ligand release in more harsh conditions.<sup>122</sup> The principle of this method is shown on *Figure 49*.

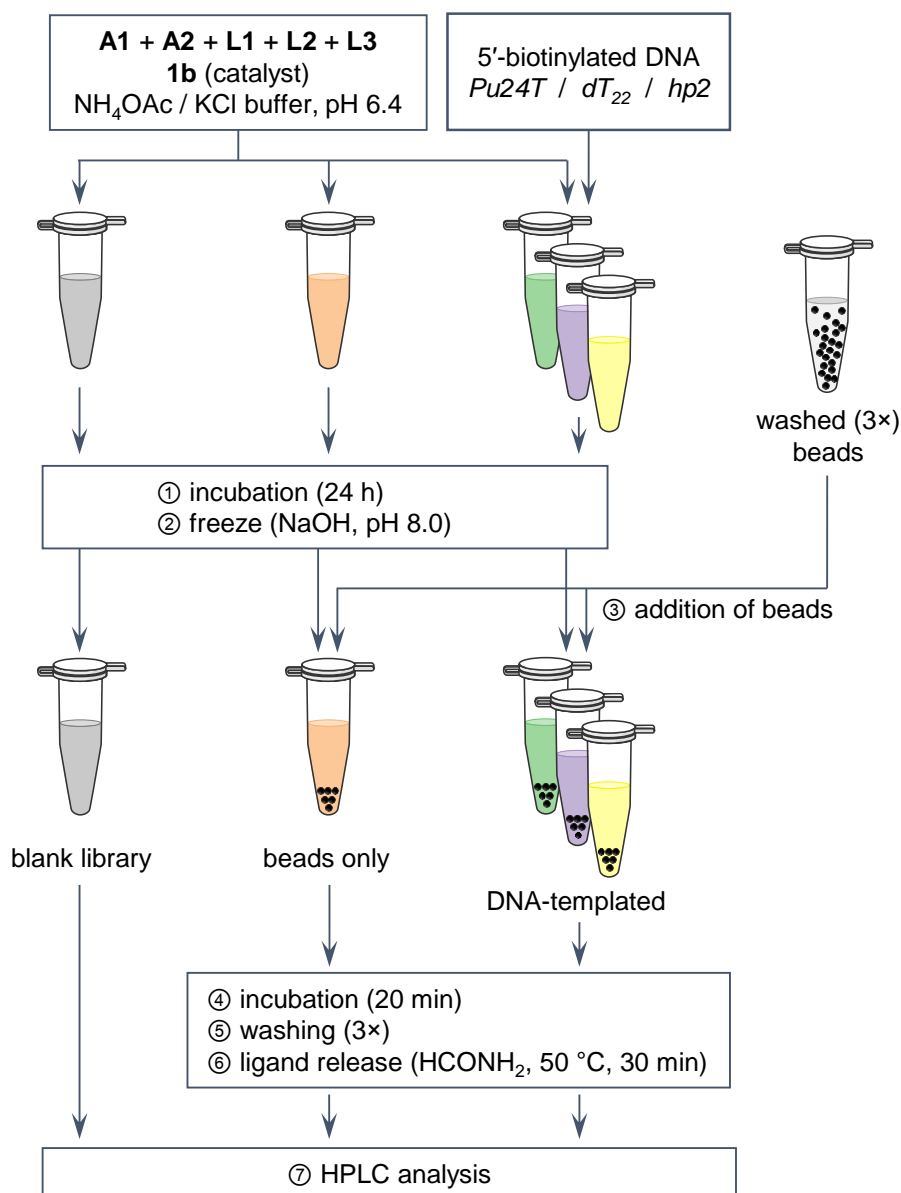


Figure 49. Experimental protocol for DNA-targeted DCC (here: DCL1) using biotinylated DNA oligonucleotides and streptavidin-coated magnetic beads.

To validate this approach, DCL1 was incubated in the presence of 5'-biotinylated *Pu24T*. In parallel, we employed two 5'-biotinylated, non-G4 controls of comparable molecular size, namely a single-stranded oligonucleotide (*dT<sub>22</sub>*) and an 18-mer hairpin (*hp2*) containing a non-nucleoside (i.e., hexaethyleneglycol) loop, introduced to avoid the possible binding of ligands in the loop region (sequences: Table 8). Following a 24-hour incubation, the reaction was arrested by increase of pH up to 8 by addition of  $\text{NaOH}$  and pre-washed magnetic beads (Dynabeads MyOne T1, Thermo Fisher) were added, resulting in pull-down of DNA–ligand complexes. After a thorough washing, the pulled-down ligands were released from the beads

by addition of neat formamide and heating to 50 °C, and analyzed by HPLC; the resulting peak areas were normalized with respect to control (blank, *Figure 49*) library to account for the differences in initial composition and the differences in UV absorption by ligands, and compared with a sample treated with magnetic beads in the absence of biotinylated DNA, to account for non-specific ligand absorption on the beads.

The results (*Figure 50a*) demonstrated that five ligands were selectively pulled down from the library by *Pu24T*-covered beads, in the following relative order: **A2-L3-A2** > **A2-L1-A2** > **A1-L3-A2**  $\approx$  **A2-L2-A2** > **A1-L1-A2**. Satisfyingly, this ranking was in a good agreement with the ligand-induced stabilization of *Pu24T* observed in fluorescence-melting experiments performed with pure samples of symmetric ligands that demonstrated that **A2-L3-A2 (NaphDH2)** and **A2-L1-A2 (PyDH2)**, but not **A2-L2-A2 (PymDH2)** were good G4 binders (*Figure 50b*). Of note, non-specific binding of ligands to magnetic beads was not negligible in these experiments, as evidenced by the amount of ligands pulled-down in the absence of biotinylated DNA (“beads only”); at the same time, no ligand was selectively pulled-down in the presence of single-stranded or hairpin DNA, demonstrating the absence of binding to non-G4 DNA targets. This fact also agrees with the results of fluorescence-melting experiments that demonstrate minimal effect of all ligands on thermal stability of the hairpin *hp2* (*Figure 50b*).

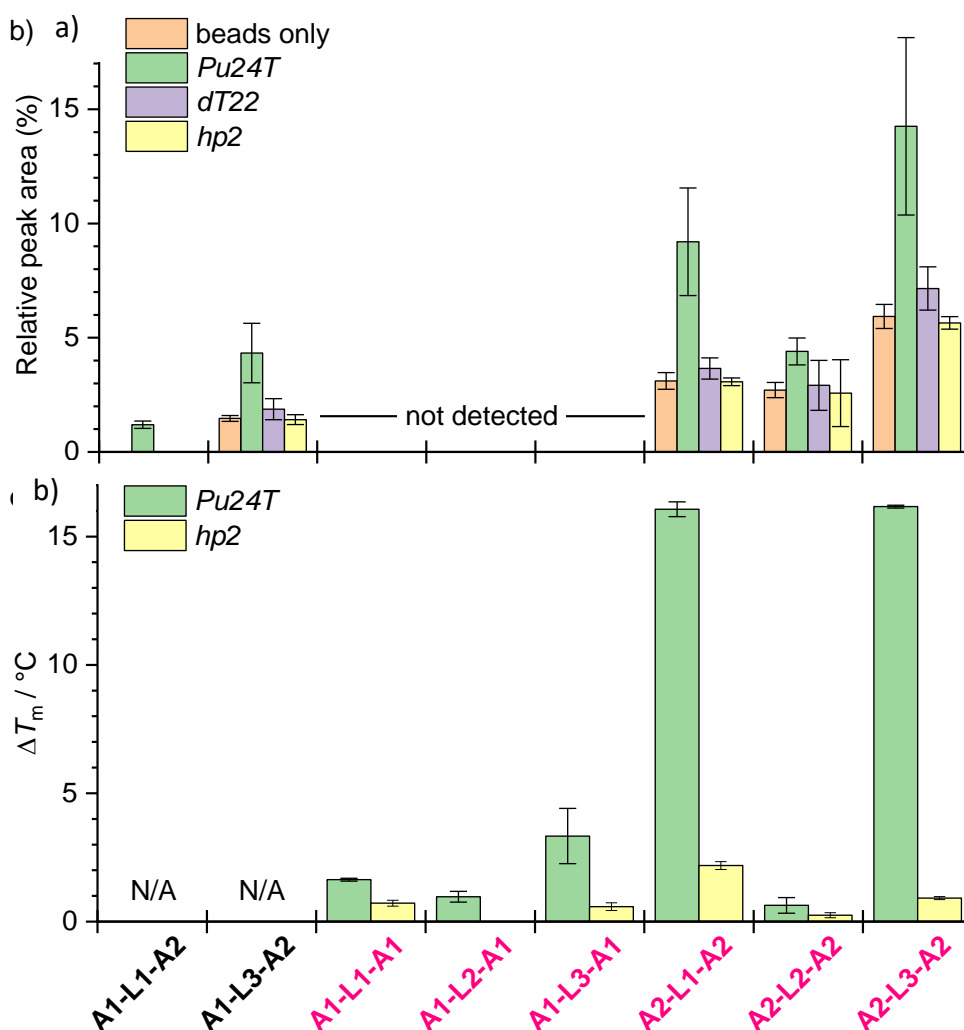


Figure 50. a) Relative (with respect to the “blank library”) amounts of components of DCL1 released after pull-down with naked beads or biotinylated oligonucleotides; data are mean  $\pm$  s.d. from three independent experiments. b) Ligand-induced stabilization of G4-DNA Pu24T and hairpin hp2 (0.2  $\mu$ M) in fluorescence-melting experiments performed with pure, extemporarily synthesized symmetrical components of DCL1 (1.0  $\mu$ M). Data are mean  $\pm$  s.d. from three measurements.

Inspired by these results, we designed two other libraries: a 14-membered DCL2, prepared from the dihydrazide **L1**, cationic aldehyde **A2** and its isomer **A6**, as well as two novel aldehydes (**A7** and **A8**) (Figure 51a), and an 18-membered DCL3 was constructed from **L1**, its 4-(2-morpholinoethoxy) derivative **L5**, and aldehydes **A2**, **A6** and **A8** (Figure 52a). In both cases, we introduced the fragments that were uncharged (**A7**) or partially charged in the assay conditions (**A8**) in a hope to obtain ligands displaying a G4 affinity comparable to that of the prototype bis(acylhydrazones) (**A2-L1-A2**, **PyDH2**) but a lower permanent cationic charge.



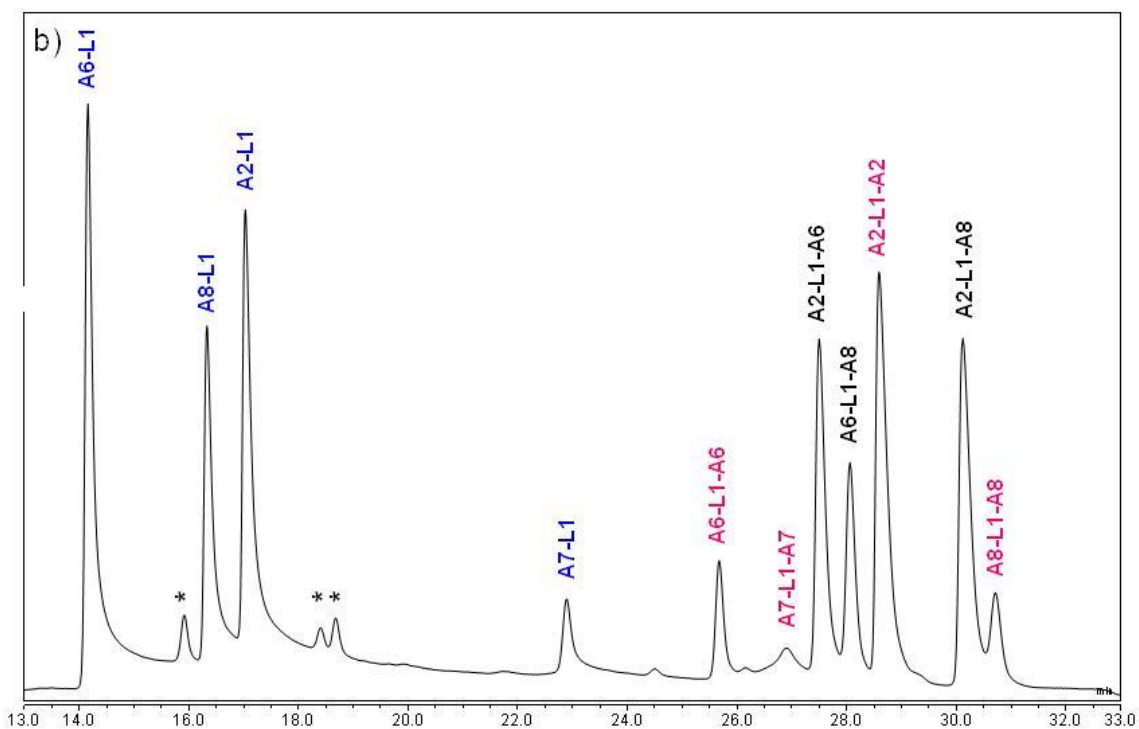
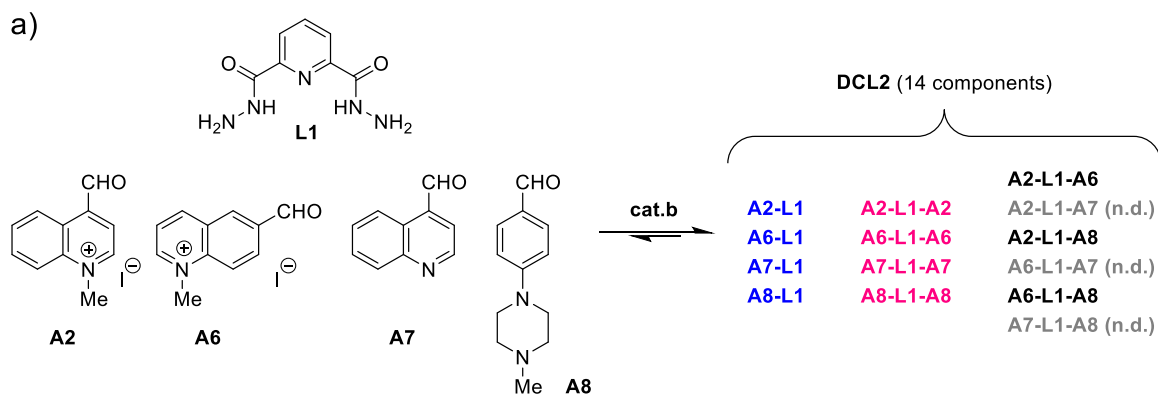


Figure 51. Design (a) and corresponding chromatogram (b) of 14-component DCL2. N.d. = components not detected in the blank library. *Italicized numbers indicate the conversion of the reagents in "blank" libraries.*

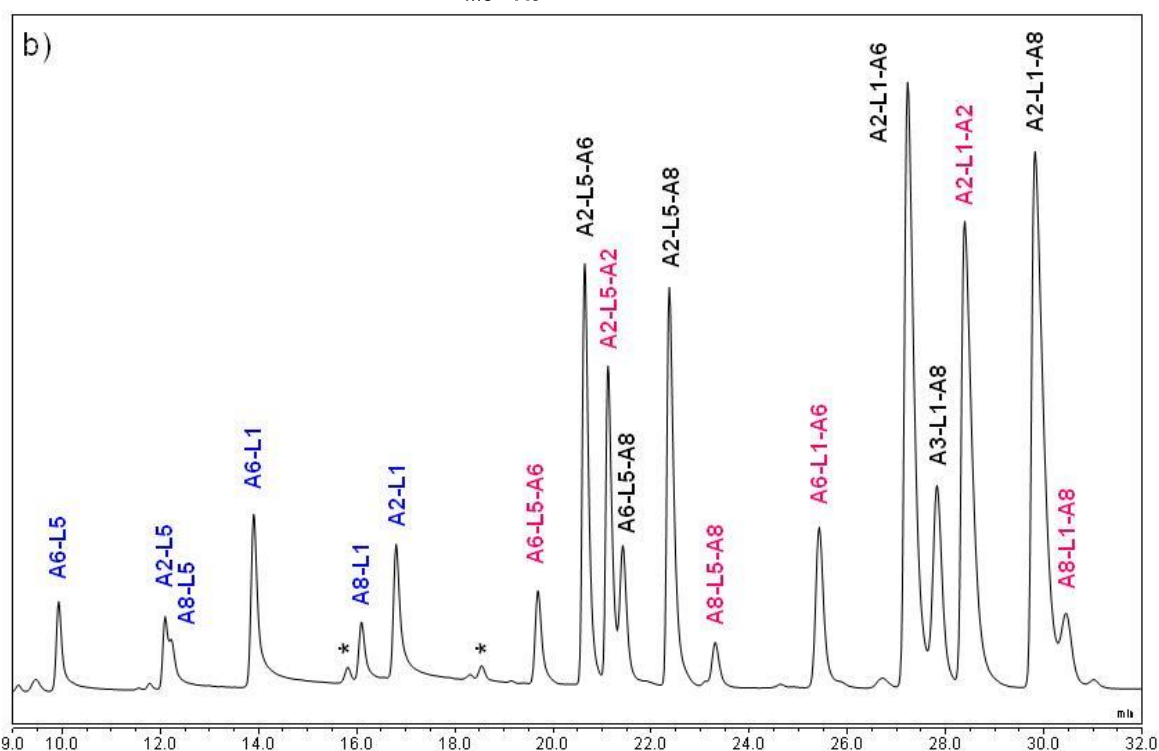
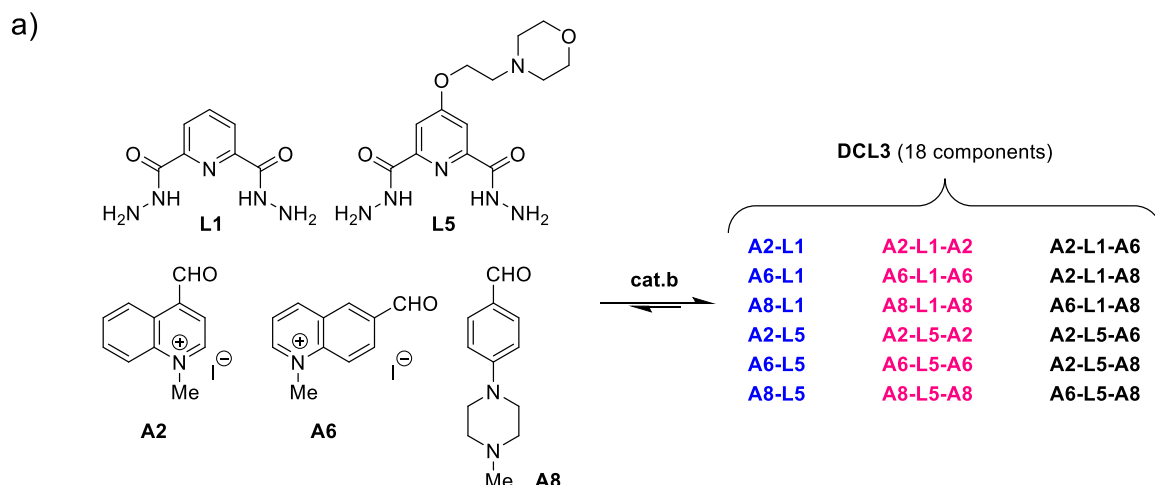


Figure 52. Design (a) and corresponding chromatogram (b) of 18-component DCL3. N.d. = components not detected in the blank library. *Italicized numbers indicate the conversion of the reagents in "blank" libraries.*

Of note, the conversion of aldehyde **A8** in the conditions of DCL synthesis was significantly lower (34–41%) comparing with heterocyclic aldehydes **A2** and **A6–A7** (80–100%) (Table 9), presumably due to its lower reactivity owing to the presence of an electron-donating amine substituent in the benzene ring. Nevertheless, acylhydrazone products containing the **A8** fragment could be clearly detected in the resulting libraries (Figure 51b, Figure 52b).

Table 9. Initial compositions of DCLs 2–3 and conversion rate of the corresponding reactants.

Reactant	DCL2		DCL3	
	c (μM)	conv. (%)	c (μM)	conv. (%)
<b>A2</b>	60	87	80	83
<b>A6</b>	60	94	80	97
<b>A7</b>	60	≈100		
<b>A8</b>	60	34	80	41
<b>L1</b>	120	91	60	95
<b>L5</b>			60	97
<b>cat.b</b>	5 mM		5 mM	
<b>DNA<sup>a</sup></b>	5		5	

<sup>a</sup> In templated libraries.

### 2.3 Hits identification & synthesis

The pull-down experiments, performed with magnetic beads in the conditions identical to those presented in *Figure 50*, demonstrated that, in the case of DCL2, **A2-L1-A8** was the only compound massively and selectively extracted in the presence of *Pu24T*; even in this case, significant amounts of this ligand were detected in samples treated with *dT<sub>22</sub>* and *hp2* oligonucleotides (*Figure 53a*). Interestingly, **A2-L1-A2**, a good G4-DNA binder, was not selected from this library. In contrast, in the case of DCL3, a number of compounds were selectively extracted in the presence of *Pu24T*, in the following order: **A2-L1-A8** ≈ **A2-L5-A8** > **A2-L1-A2** ≈ **A2-L5-A2** > **A2-L1-A6** ≈ **A2-L5-A6** (*Figure 53b*). These compounds (in addition to **A2-L1-A2** already described above) could be considered as promising G4-DNA binders, and we next sought to synthesize some of them in preparative fashion in order to confirm their G4-binding properties.

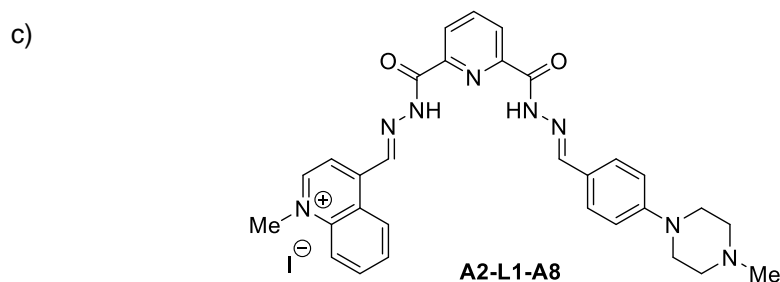
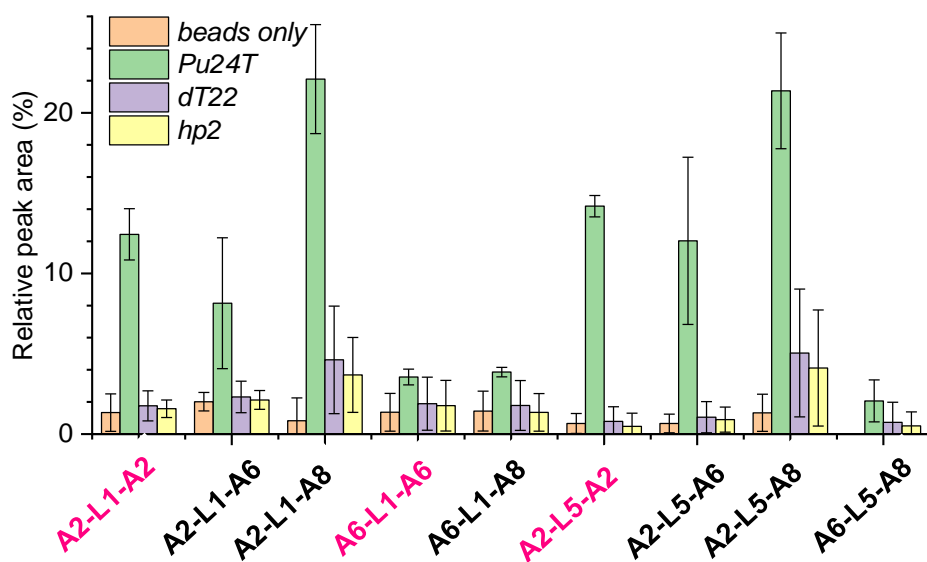
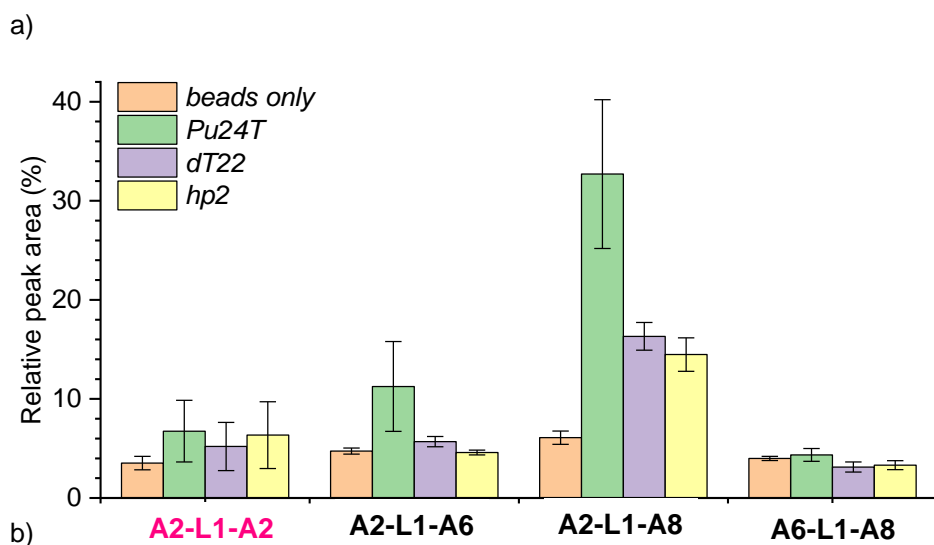
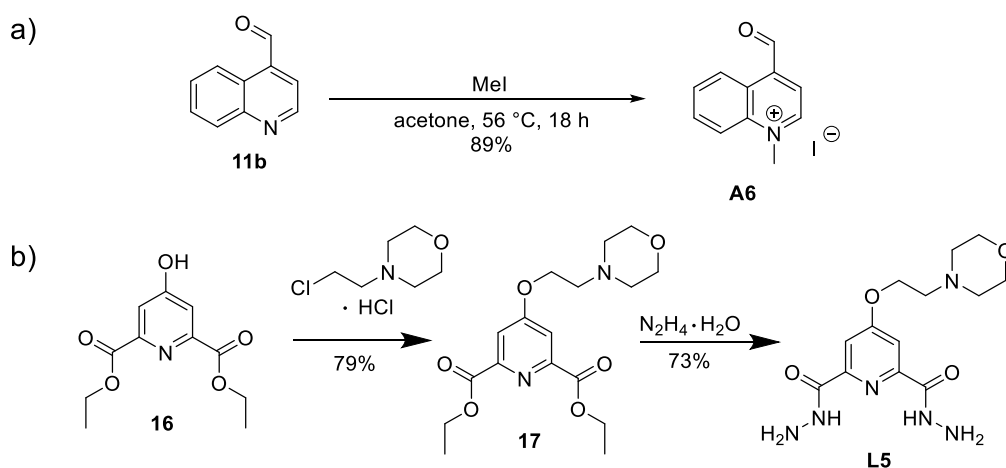


Figure 53. Relative (with respect to the “blank library”) amounts of components of a) DCL2 and b) DCL3, released after pull-down with naked beads or biotinylated oligonucleotides; data are mean  $\pm$  s.d. from three independent experiments. c) Structure of compound A2-L1-A8, pulled down from both DCL2 and DL3 libraries.

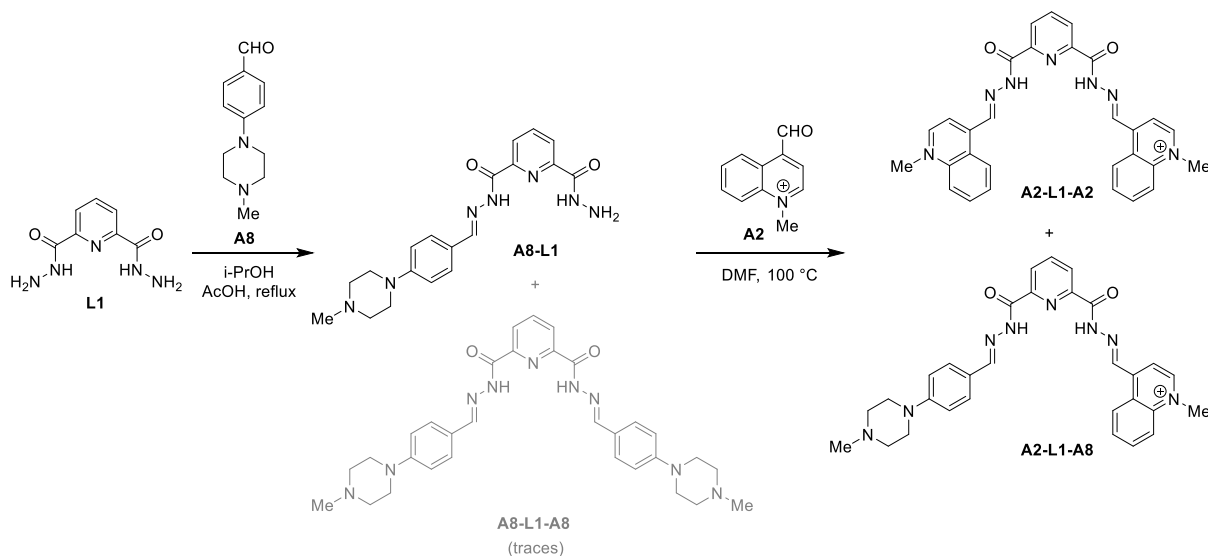
## 2.4 Synthesis of building blocks and selected compounds

Aldehydes **A1–A2** and acylhydrazides **L1–L3** were synthesized previously in this work (*RESULTS AND DISCUSSION, Part 1*). Aldehyde **A6** was obtained similarly to **A2** (*Scheme 9a*). Acylhydrazide **L5** was obtained in two steps starting from diethyl chelidamate **16** via reaction of alkylation with 4-(2-chloroethyl)morpholine hydrochloride, followed by hydrazinolysis of diethyl 4-(2-morpholinoethoxy)pyridine-2,6-dicarboxylate **17** (*Scheme 9b*). Aldehydes **A7** and **A8** were purchased from commercial suppliers.



*Scheme 9. Synthesis of building blocks **A6** (a) and **L5** (b).*

The non-symmetric bis(acylhydrazone) **A2-L1-A8**, selected from both DCL2 and DCL3, was considered as a priority target. An initial attempt to synthesize this compound was undertaken through a “one-pot” reaction performed by heating of **L1** with equimolar amounts of **A2** and **A8** in DMF (*Scheme 10*), followed by chromatographic separation of the products. However, while mass-spectrometric analysis showed three compounds (**A2-L1-A2**, **A2-L1-A8** and **A8-L1-A8**) formed in this reaction, they co-eluted on a preparative RP-HPLC column, rendering their separation impossible on a preparative scale (*Figure 54*).



Scheme 10. The attempt to synthesize non-symmetric compound **A2-L1-A8**.

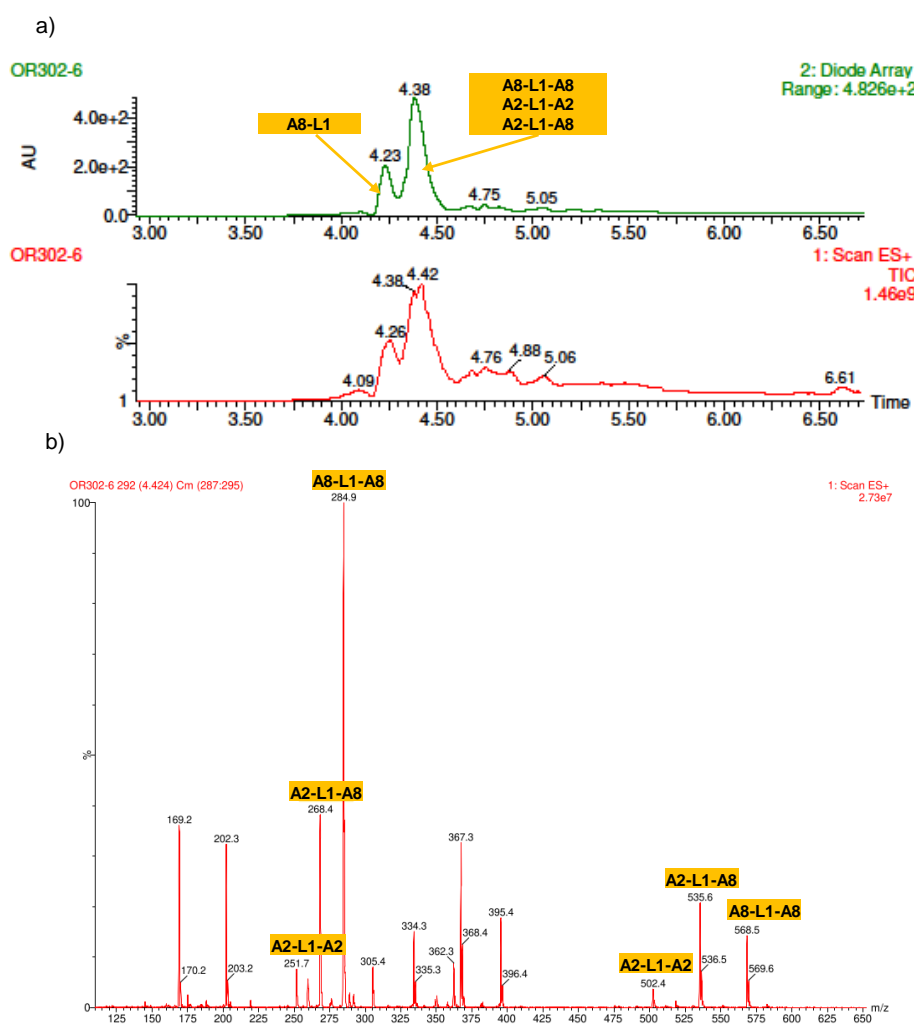
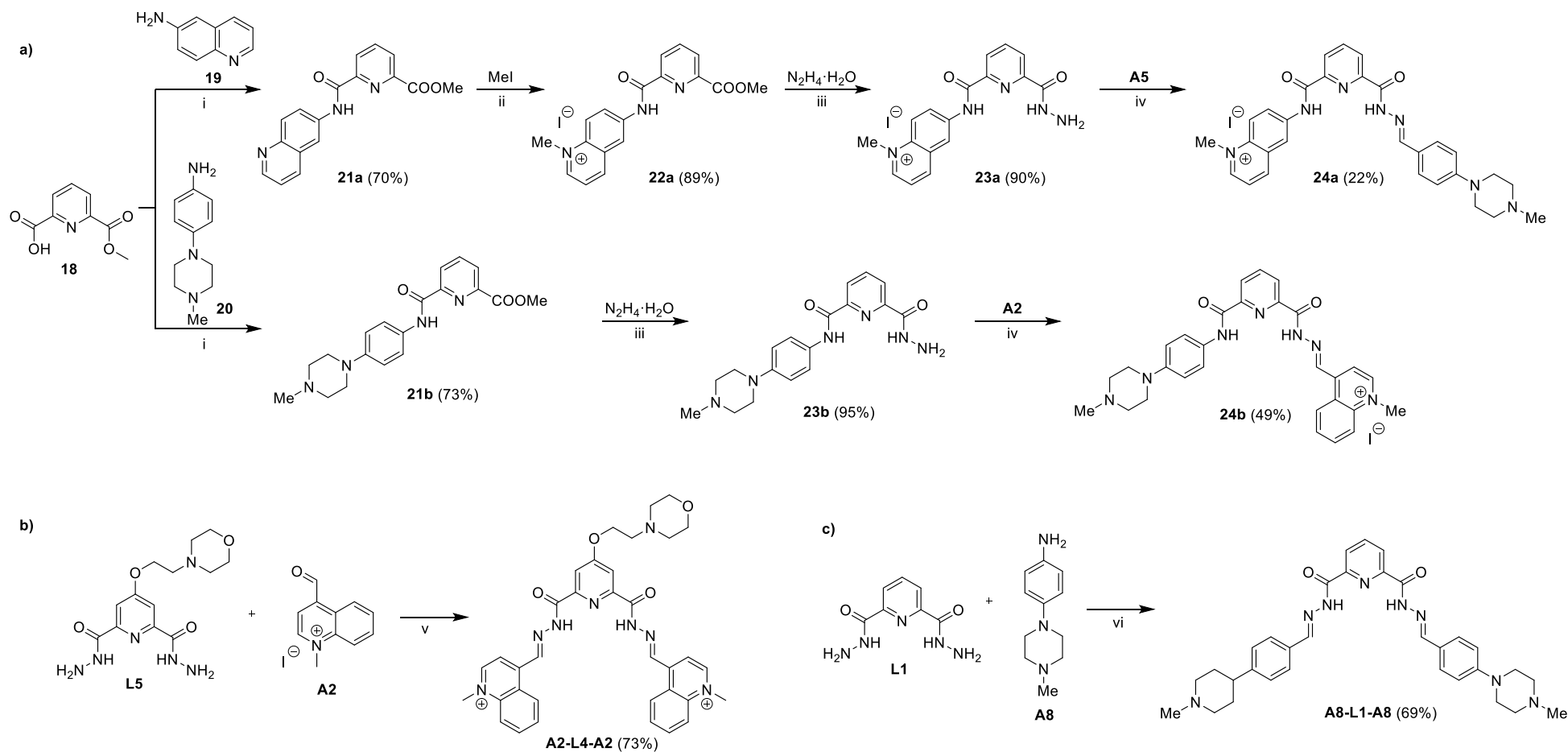


Figure 54. a) The chromatogram of the reaction mixture after addition of **A2** to the solution of **A8-L1** with traces of **A8-L1-A8**; b) mass spectra corresponding to peak at 4.38 minutes.

Therefore, we designed hybrid ligands **24a** and **24b**, in which *N*-methylquinolinium and 4-(4-methylpiperazino)phenyl moieties were linked to the central 2,6-pyridyl unit through two different combinations of carboxamide and acylhydrazone bonds.

The synthesis of **A2-L1-A8** analogues **24a** and **24b** is presented on *Scheme 11a*. Both isomers were prepared from the commercially available mono-methyl pyridine-2,6-dicarboxylate **18** through amidation with the corresponding aromatic amines **19** and **20** to give the amides **21a-b**, followed by methylation of the quinoline substituent yielding the salt **22a**, hydrazinolysis of methyl esters to give the acylhydrazides **23a** and **23b**, and the reaction of the latter with aldehydes **A8** or **A2**, respectively.

In addition, symmetric bis(acylhydrazones) **A2-L5-A2** (selected from DCL3) and **A8-L1-A8** (not selected either from DCL2 or from DCL3, and chosen as a negative control) were prepared by straightforward condensation of the corresponding building blocks (*Scheme 11b-c*).





## 2.5 Evaluation of G4-binding properties

### 2.5.1 Fluorimetric titrations

G4-binding properties of novel derivatives were initially studied by isothermal fluorimetric titrations that harness the ligand-induced fluorescence quenching of 5'-Cy5-labelled oligonucleotides (cf. INTRODUCTION, part 1.4.4). Performed at a constant temperature, this method is suitable to characterize strong binders with  $K_d$  values down to few nM, and is gaining increasing popularity in characterizing ligand–G4 interactions. In addition, contrary to fluorescence-melting experiment, this method directly provides the values of dissociation constants ( $K_d$ ) of G4–ligand complexes.

Prior to perform fluorimetric titrations, we recorded CD-spectra of 5'-Cy5-labelled telomeric sequences *5'-Cy5-22CTA* and *5'-Cy5-25TAG* (Figure 55), because in the original article introducing this essay<sup>66</sup> the change of topology of telomeric sequence to parallel was reported. This phenomenon was explained by the influence of the fluorophore. It appeared to us that sequence *5'-Cy5-22CTA* had mixed hybrid topology in the conditions, suggested for this assay (buffer K100 with additives CHAPS and Triton-X), regardless cooling mode (fast or slow) (Figure 55A, black and red curves). However, when the annealing of this sequence was performed in the buffer without additives, the sequence formed a G4 structure of antiparallel topology (Figure 55B, black curve). We added CHAPS and Triton-X to this prefolded G4 to see if the additives can affect the topology of prefolded sequences after annealing. Even after 18 hours, antiparallel topology of *5'-Cy5-22CTA* sequence remained unchanged (Figure 55B, red curve). The topology of *5'-Cy5-25TAG*, annealed in the buffer with additives was parallel (Figure 55C, black curve), as reported by Balasubramanian et al.<sup>66</sup> However, if the same sequence was annealed in K100 buffer without additives it formed G4 of hybrid topology (Figure 55C, red curve). These experiments indicate that the change of topology of labelled sequences is mostly due to the presence of additives during the annealing step and that the influence of fluorophore (Cy5) on G4 topology is less prominent in these cases. For our experiments, we used 5'-Cy5-labelled DNA annealed in K100 buffer (at 1  $\mu$ M) and slowly cooled to room temperature.

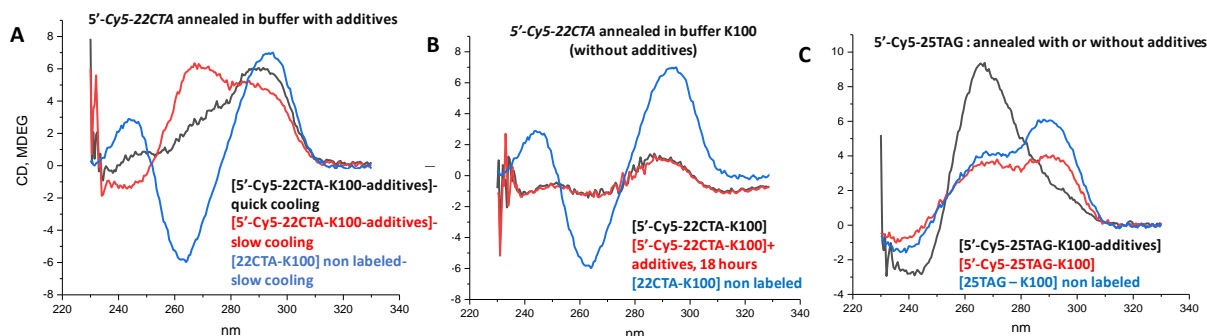


Figure 55. CD spectra of telomeric 5'-Cy5-22CTA (A,B) and 5'-Cy5-25TAG (C) prefolded in different conditions. A) 5'-Cy5-22CTA annealed in K100 with additives with consecutive quick (black) or slow (red) cooling. B) 5'-Cy5-22CTA annealed in K100 buffer without additives, slow cooling before (black) and 18 hours after (red) addition of additives to the solution of prefolded sequence. C) 5'-Cy5-25TAG prefolded in K100 with (black) or without (red) additives. Spectra of corresponding unlabelled sequences prefolded in K100 buffer without additives with slow cooling are showed in blue in all panels.<sup>a</sup> Conditions **K100 with additives**: 10 mM LiAsO<sub>2</sub>Me<sub>2</sub>, 100 mM KCl, 0.5 w/v % CHAPS, 0.05 v/v % Triton X-100, pH 7.2. **K100**: 10 mM LiAsO<sub>2</sub>Me<sub>2</sub>, 100 mM KCl, pH 7.2. Annealing: 5 minutes at 95°C; fast cooling: the sample is put on the ice for 15 minutes; slow cooling: the samples was cooling in thermoblock to room temperature for 4-12 hours. <sup>a</sup> Spectra of 5'-Cy5-22CTA and unlabelled 22CTA are recorded in different cuvettes.

Fluorimetric titrations were performed by mixing the 5'-Cy5-labelled DNA ( $c(\text{final}) = 2 \text{ nM}$ ) with the range 0 to 10  $\mu\text{M}$  of ligand concentrations and measuring the spectra of fluorescence after 2 hours of incubation. The fluorescence intensity was normalized by dividing the raw value by the mean intensity of no-ligand wells. The obtained titration curves were fitted to a 1:1 binding model using Equation (1):

$$F = F_0 - (F_0 - F_b) \times \frac{A + \frac{x}{c} + 1 - \sqrt{(A + \frac{x}{c} + 1)^2 - 4 \times \frac{x}{c}}}{2}, \quad (1)$$

where  $F_0$  and  $F_b$  are fluorescence intensity in the absence and in the presence of saturating concentration of the ligand, respectively;  $x$  is a concentration of the ligand and  $A = \frac{K_d}{c}$ , where  $K_d$  is the dissociation constant and  $c$  is the concentration of labelled oligonucleotide (2 nM).

The results of titrations, performed with four fluorophore-labelled G4-DNA oligonucleotides belonging to different folding topology groups (*Pu24T* and *myc22*: parallel, *25TAG*: hybrid, *22CTA*: antiparallel), confirmed the G4-binding properties of novel derivatives (

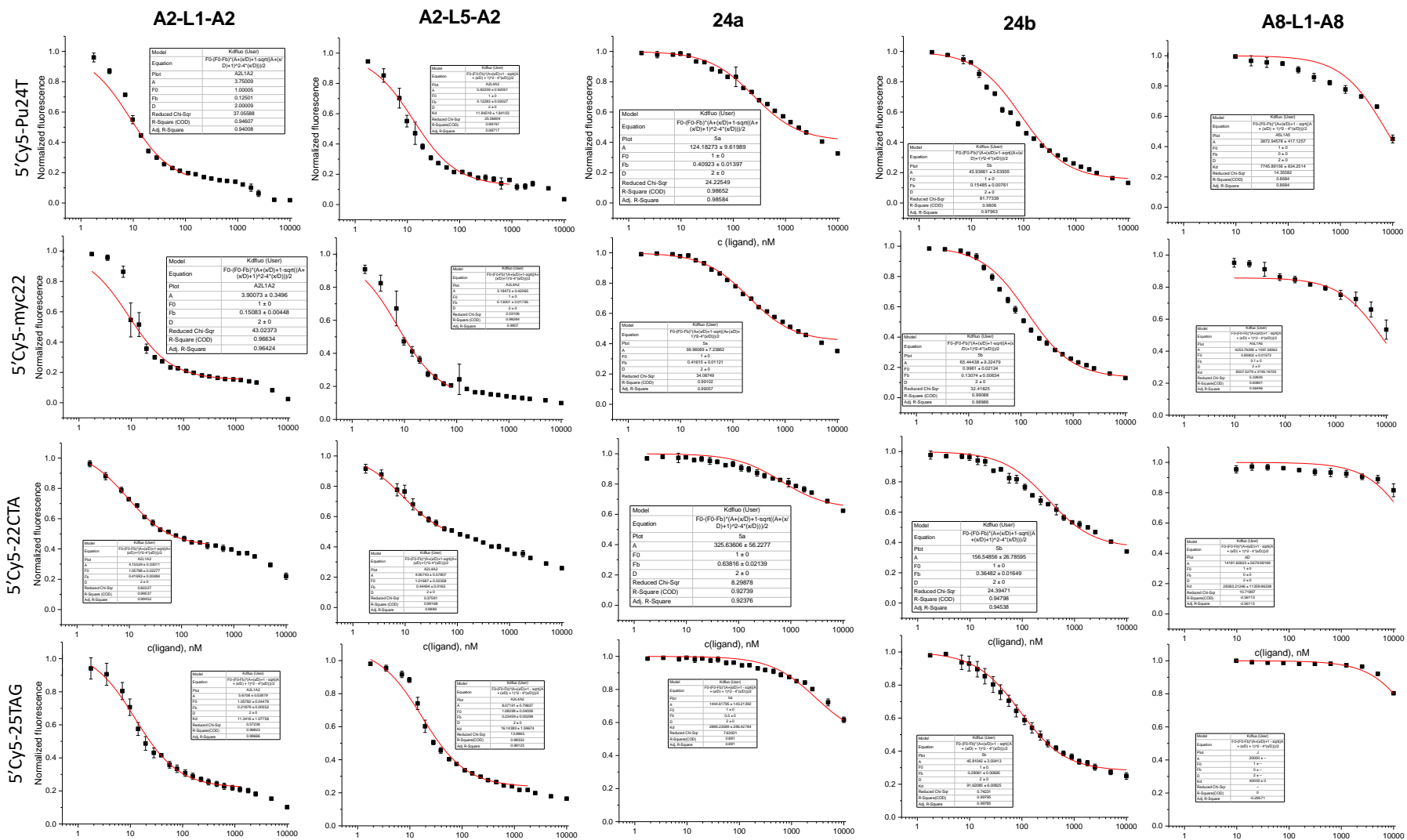
*Table 10, Figure 56*). Compounds **A2-L1-A2** and **A2-L5-A2** had very similar and very high affinity to all four G-quadruplexes, with  $K_d$  values of 6 to 16 nM, which is comparable to that of the benchmark ligand **PhenDC3** ( $K_d$  of 7 to 17,

*Table 10*) and largely supersedes the affinity of several recently reported G4-binding peptides (determined by the same method).<sup>171,172</sup> Hybrids **24a** and **24b** were significantly less active (with  $K_d \geq 90$  nM), with **24a** generally having lower affinity than **24b**. In particular, the interaction of **24a** with telomeric G4-DNA (*25TAG*) was characterized by a  $K_d$  value of 2.9  $\mu$ M, more than ten-fold larger than in case of parallel quadruplexes *Pu24T* and *myc22* ( $K_d \approx 250$  and 200 nM, respectively). Compound **A8-L1-A8** did not quench the fluorescence of Cy5-labelled G4 substrates, a fact that can be interpreted either as poor G4 binding of this compound, or as unfavourable electron transfer with the fluorophore, resulting in the lack of quenching effect upon binding. Finally, none of ligands showed quenching of fluorophore-labelled hairpin *hp2* and single-stranded control *dT22* at concentrations lower than 1  $\mu$ M (*Figure 56*), giving evidence of the absence of non-specific (i.e., unrelated to G4-DNA binding) quenching effect of ligands on the Cy5 probe in the employed concentration range.

*Table 10. Dissociation constants ( $K_d$ /nM) of ligands to G4-DNA, determined from fluorimetric titrations with 5'-Cy5-labelled oligonucleotides.<sup>a</sup>*

Ligand	<i>Pu24T</i>	<i>myc22</i>	<i>25TAG</i>	<i>22CTA</i>
<b>A2-L1-A2</b>	7.6 $\pm$ 0.6	7.8 $\pm$ 0.3	11.3 $\pm$ 0.5	8.3 $\pm$ 0.3
<b>A2-L5-A2</b>	11.8 $\pm$ 1.8	6.4 $\pm$ 0.3	16.1 $\pm$ 0.8	8.1 $\pm$ 0.6
<b>24a</b>	248 $\pm$ 10	199 $\pm$ 7	2890 $\pm$ 140	651 $\pm$ 56
<b>24b</b>	88 $\pm$ 4	131 $\pm$ 9	92 $\pm$ 3	313 $\pm$ 27
<b>A8-L1-A8</b>	> 10 <sup>4</sup>	> 10 <sup>4</sup>	> 10 <sup>4</sup>	> 10 <sup>4</sup>
<b>PhenDC3</b>	6.7 $\pm$ 0.8	8.7 $\pm$ 2.0	n.m. <sup>b</sup>	16.8 $\pm$ 4.7

<sup>a</sup> Conditions:  $c(\text{DNA}) = 2$  nM,  $c(\text{ligand}) = 0$  to 10  $\mu$ M in 10 mM  $\text{LiAsO}_2\text{Me}_2$ , 100 mM KCl, 0.5 w/v% CHAPS, 0.05 v/v% Triton X-100 buffer, pH 7.2;  $\lambda_{\text{ex}} = 590$  nm,  $\lambda_{\text{em}} = 675$  nm.  $K_d$  values are determined by numeric fitting of the data (means from three technical replicates) to a 1:1 binding model, with errors representing the standard error of the fitting parameter; <sup>b</sup>) not measured.



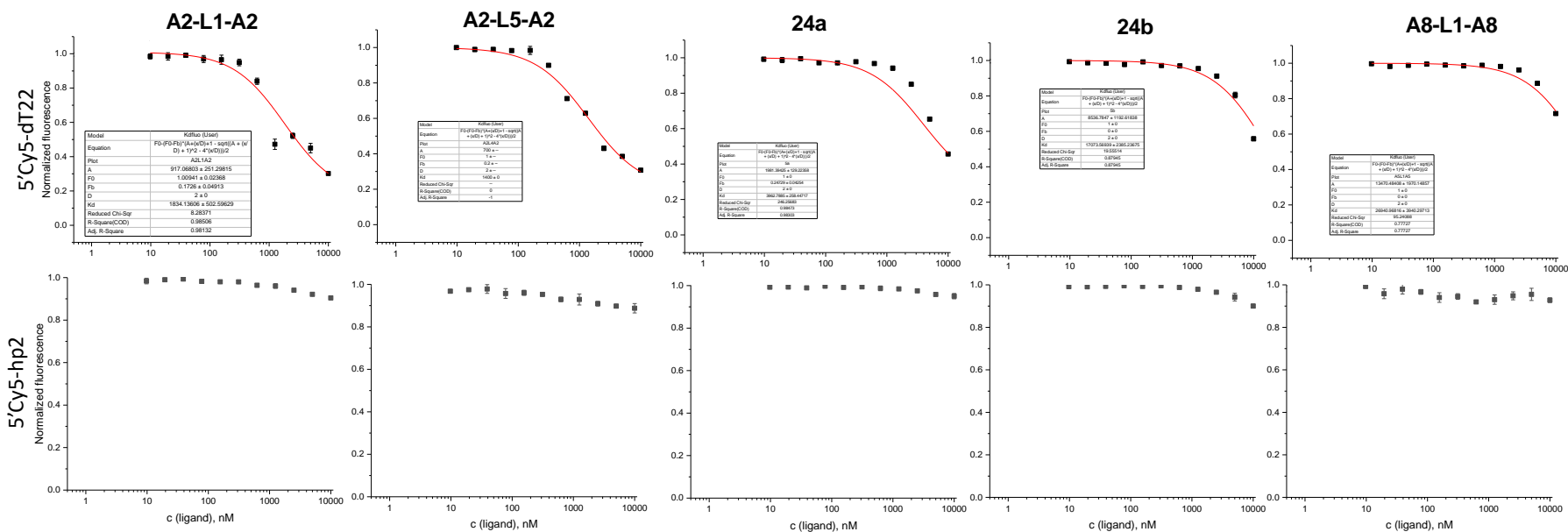


Figure 56. Normalized changes of fluorescence intensity ( $\lambda_{ex} = 590^{BP50} \text{ nm}$ ,  $\lambda_{em} = 675^{BP50} \text{ nm}$ ) upon titrations of ligands (final concentration: 1.8 nM to 10  $\mu\text{M}$ ) to 5'-Cy5-labelled oligonucleotides ( $c = 2 \text{ nM}$  in 10 mM  $\text{LiAsO}_2\text{Me}_2$ , 100 mM KCl, 0.5 w/v % CHAPS, 0.05 v/v % Triton X-100, pH 7.2). Data are means  $\pm$  s.d. from three independent experiments. Red lines represent fits of experimental data to a 1:1 binding model.

### 2.5.2 Native mass spectrometry<sup>1</sup>

Additional information about G4 binding of novel derivatives was obtained from native mass spectra of G4–ligand complexes that were performed with two G4-DNA (*Pu24T* and *25TAG*). In all cases, the peaks of G4–ligand complexes could be clearly observed. Interestingly, mass spectra of *25TAG* in the presence of **A2-L1-A2** and **A2-L5-A2** were largely dominated by the peaks of 1:1 complexes, formed upon ejection of one of the two K<sup>+</sup> ions initially present in this substrate (*Figure 57*). Ejection of K<sup>+</sup> cations was previously observed with other ligands of the bis(quinolinium) family, and attributed to the ligand-induced change from hybrid to anti-parallel conformation of this G4-DNA.<sup>142</sup> In contrast, 1:1 complexes of *25TAG* with hybrids **24a** and **24b** showed the peaks attributed to both hybrid (2K<sup>+</sup>) and anti-parallel (1K<sup>+</sup>) species. This observation, along with the additional signals of 2:1 (ligand : G4) complexes, suggests their multiple binding modes to the *25TAG* quadruplex. In the case of *Pu24T*, no K<sup>+</sup> ejection was observed, and only minor peaks of 2:1 complexes could be detected with all ligands (*Figure 58*), consistent with the presence of a single ligand-accessible binding site (most likely, the 5'-tetrad).<sup>31</sup>

---

<sup>1</sup> Experiments performed in the lab of V. Gabelica, IECB, with assistance of Dr. E. Largy and Dr. F. Rosu.

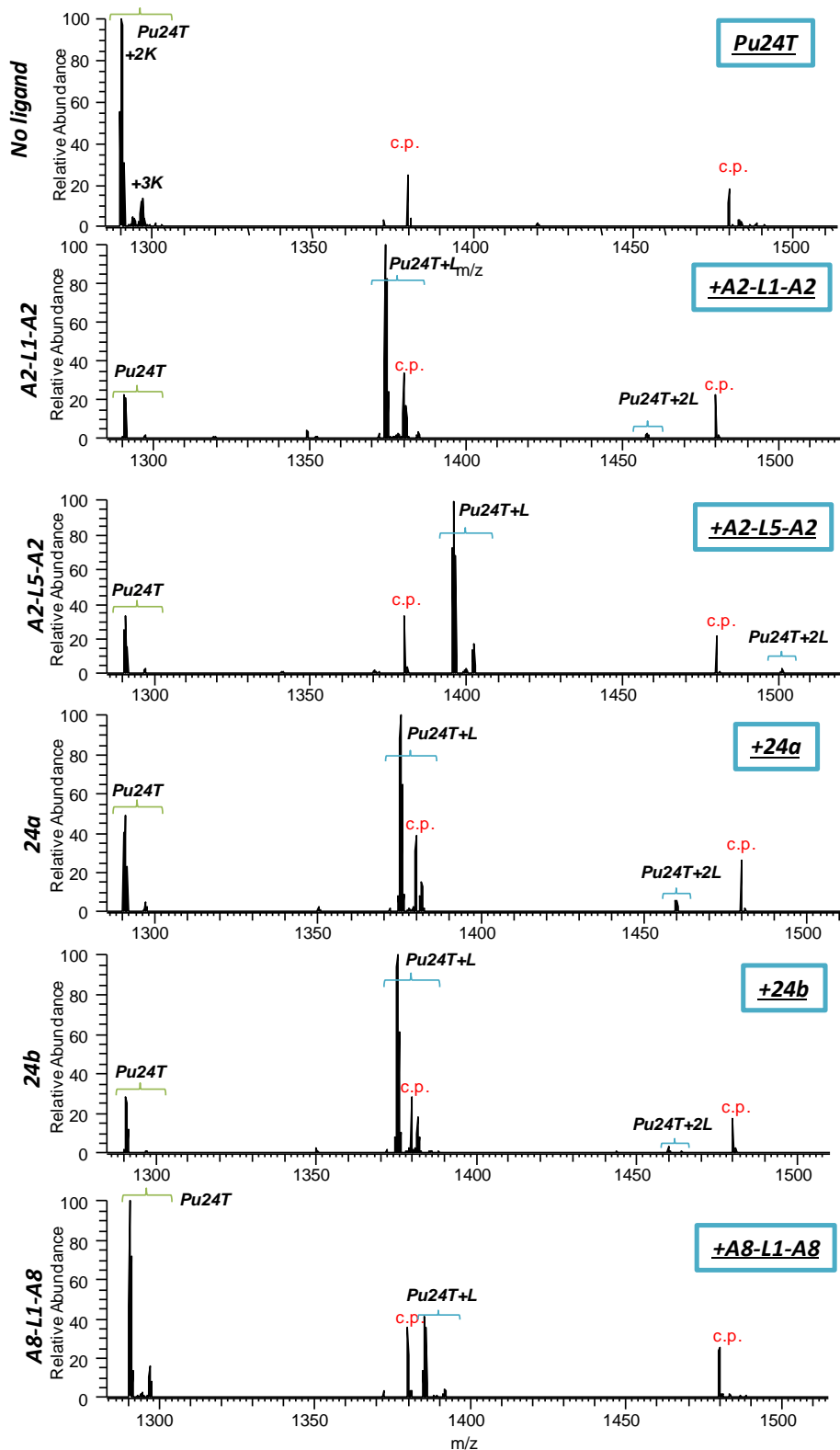


Figure 57. ESI-MS spectra of 25TAG (5  $\mu$ M) in the absence (top) and in the presence of equimolar concentrations of ligands (L) **A2-L1-A2**, **A2-L5-A2**, **24a**, and **24b**, in 100 mM TMAA + 1 mM KCl buffer. *c.p.* = calibration peak ( $dT_6$ ).

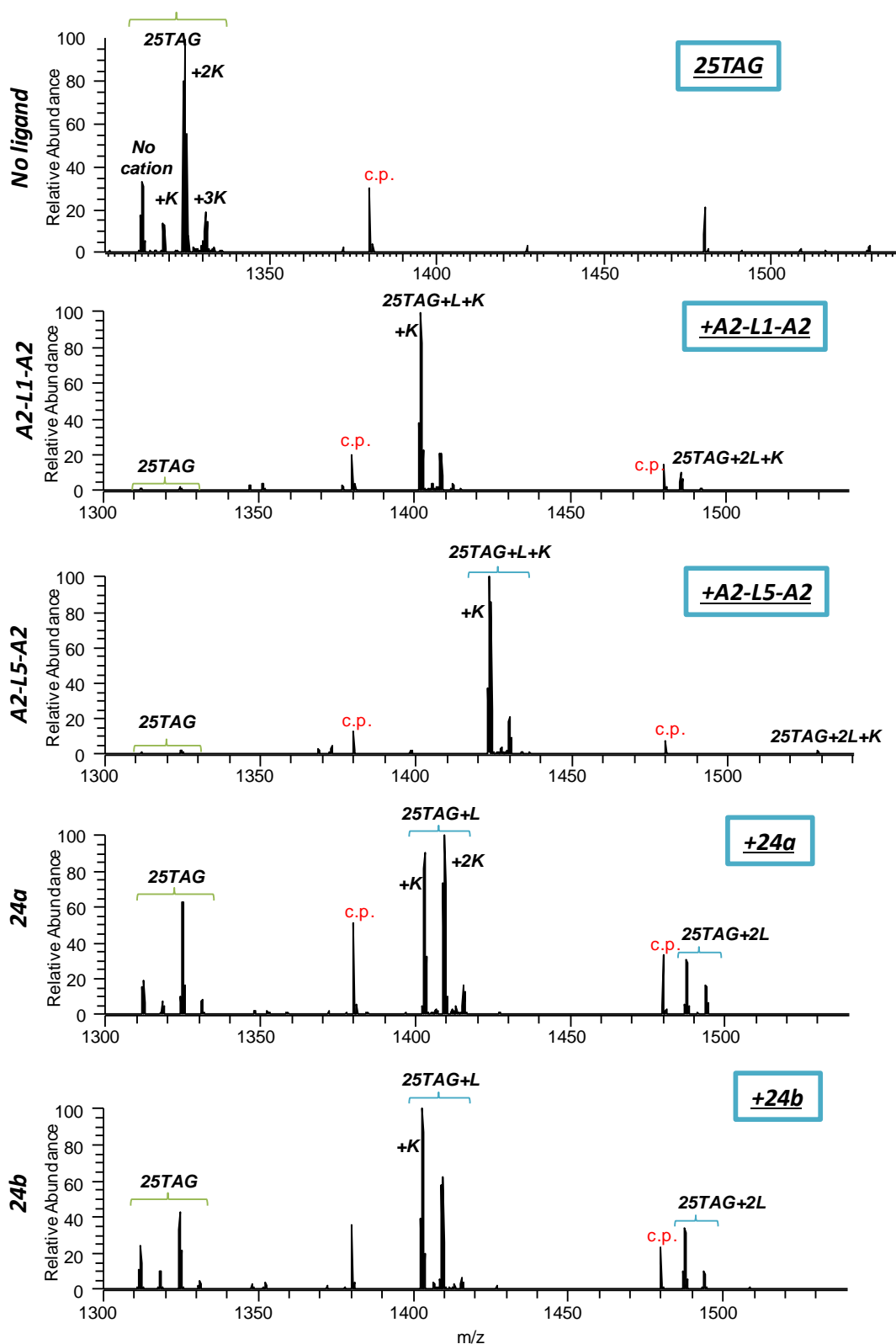


Figure 58. ESI-MS spectra of Pu<sub>24</sub>T (5 μM) in the absence (top) and in the presence of equimolar concentrations of ligands (L) **A2-L1-A2**, **A2-L4-A2**, **5a**, **5b**, and **A5-L1-A5**, in 100 mM TMAA + 1 mM KCl buffer. *c.p.* = calibration peak (*dT<sub>6</sub>*).



The relative intensities of the peaks of the ligand-bound and free G4-DNA were used for the calculation of apparent binding constants (*Table 11*). We examined charge state 6<sup>-</sup> because it contained fewer nonspecific potassium adducts. The concentration of each DNA form was determined from the mass balance equation and from the peak areas:

$$[DNA]_{free} = [DNA]_{tot} \cdot \frac{A(DNA)}{A(DNA)+A(DNA+L)+A(DNA+2L)} \quad (2)$$

$$[DNA + L] = [DNA]_{tot} \cdot \frac{A(DNA+L)}{A(DNA)+A(DNA+L)+A(DNA+2L)} \quad (3)$$

$$[DNA + 2L] = [DNA]_{tot} \cdot \frac{A(DNA+2L)}{A(DNA)+A(DNA+L)+A(DNA+2L)} \quad (4)$$

The concentration of free ligand was determined through the mass balance equation:

$$[L]_{free} = [L]_{tot} - [DNA + L] - 2[DNA + 2L] \quad (5)$$

Apparent consecutive dissociation constants ( $K_{D1}$  and  $K_{D2}$ ) were calculated using Equations (6–7):

$$K_{D1} = \frac{[DNA]_{free} \times [L]_{free}}{[DNA+L]} \quad (6)$$

$$K_{D1} = \frac{[DNA+L] \times [L]_{free}}{[DNA+2L]} \quad (7)$$

Table 11. Apparent dissociation constants ( $K_d$  / nM) determined from mass-spectra of G4-DNA–ligand complexes.<sup>a</sup>

Ligand	<i>Pu24T</i>	<i>25TAG</i>
<b>A2-L1-A2</b>	85	n.d. <sup>b</sup>
<b>A2-L5-A2</b>	220	< 7.0
<b>24a</b>	470	330
<b>24b</b>	240	120
<b>A8-L1-A8</b>	9000	not measured

<sup>a</sup> Conditions: DNA (5  $\mu$ M) in the absence (top) or in the presence of 5  $\mu$ M of ligands. Buffer: 100 mM TMAA + 1 mM KCl buffer Time of incubation: 2 hours. <sup>b</sup> The  $K_d$  value could not be determined since the peak of unbound G4-DNA was undetectable

The values of apparent binding constants were generally found in a relatively good agreement with the results of fluorimetric titrations, except for the interaction of **24a** with telomeric *25TAG* that was characterized by a  $K_d$  value of 0.33  $\mu$ M, almost 10-fold lower than the value obtained by fluorimetry (2.89  $\mu$ M,

Table 10). This discrepancy may be due to a particular binding mode of this ligand, unfavourable for interaction with the fluorophore, such as binding in the vicinity of the unlabelled 3'-tetrad.<sup>66</sup> The fact, that the values of dissociation constants for complexes of ligands **A2-L1-A2** and **A2-L5-A2** with *Pu24T* obtained by fluorimetric titrations method are higher than those obtained by mass spectrometry may be explained by the influence of fluorophore: Cy5 may be positioned above external 5'-tetrad in such a way that it provides an additional stabilization of the complexes with ligands and in this way increases the ligands' affinities to *Pu24T*. It would be very interesting to test this hypothesis by the recording of mass spectra of complexes of ligands with 5'-Cy5-labelled oligonucleotides and comparing the corresponding  $K_d$  values with those obtained with unlabelled sequences.

### 2.5.3 Fluorescence-melting experiments

Finally, DNA binding of ligands was assessed by fluorescence-melting experiments, based on thermal denaturation of double-labelled sequences. The conditions of these experiments, namely the  $K^+$  content of the buffers, were selected so that the four G4-DNA substrates and the hairpin *hp2* denature at comparable temperature in the absence of ligands ( $T_m^0 \approx 60$  °C, cf. *RESULTS AND DISCUSSION, Part 1*) thus allowing direct comparison of ligand-induced thermal shifts ( $\Delta T_m$ ). The results (*Figure 59*) show that, in all cases, no preferential stabilization of *Pu24T* over three other G4-DNA substrates was observed; instead, most compounds preferentially stabilized the *25TAG* quadruplex, with  $\Delta T_m$  values of up to 30 °C (**A2-L1-A2**). Consistently with the results of fluorimetric titrations and mass-spectrometry experiments, hybrids **24a** and **24b** induced lower stabilization of most substrates compared with **A2-L1-A2**, but higher than **A8-L1-A8**. However, in the case of antiparallel quadruplex *22CTA*, compounds **24a** ( $\Delta T_m = 15.3$  °C) and, particularly, **24b** ( $\Delta T_m = 18.8$  °C) were almost as efficient as **A2-L1-A2** ( $\Delta T_m = 22.2$  °C) and **A2-L5-A2** ( $\Delta T_m = 17.7$  °C) despite large differences in binding constants ( *Table 10*). Conversely, compound **A2-L5-A2** systematically induced less efficient stabilization of all G4-DNA substrates compared with the parent ligand **A2-L1-A2**, although both compounds were pulled down from DCL3 at almost identical levels (*Figure 53b*) and displayed very close  $K_d$  values in the fluorimetric assay ( *Table 10*). These differences between  $K_d$  values and thermal stabilization effects may be due to their different binding mode; indeed, ligands with large entropic contributions to the binding yield higher thermal stabilization effects.<sup>133</sup> Of note, fluorescence melting experiments confirmed the excellent level of selectivity of all ligands with respect to duplex DNA, as evidenced by negligible stabilization of the hairpin *hp2* ( $\Delta T_m \leq 2$  °C).

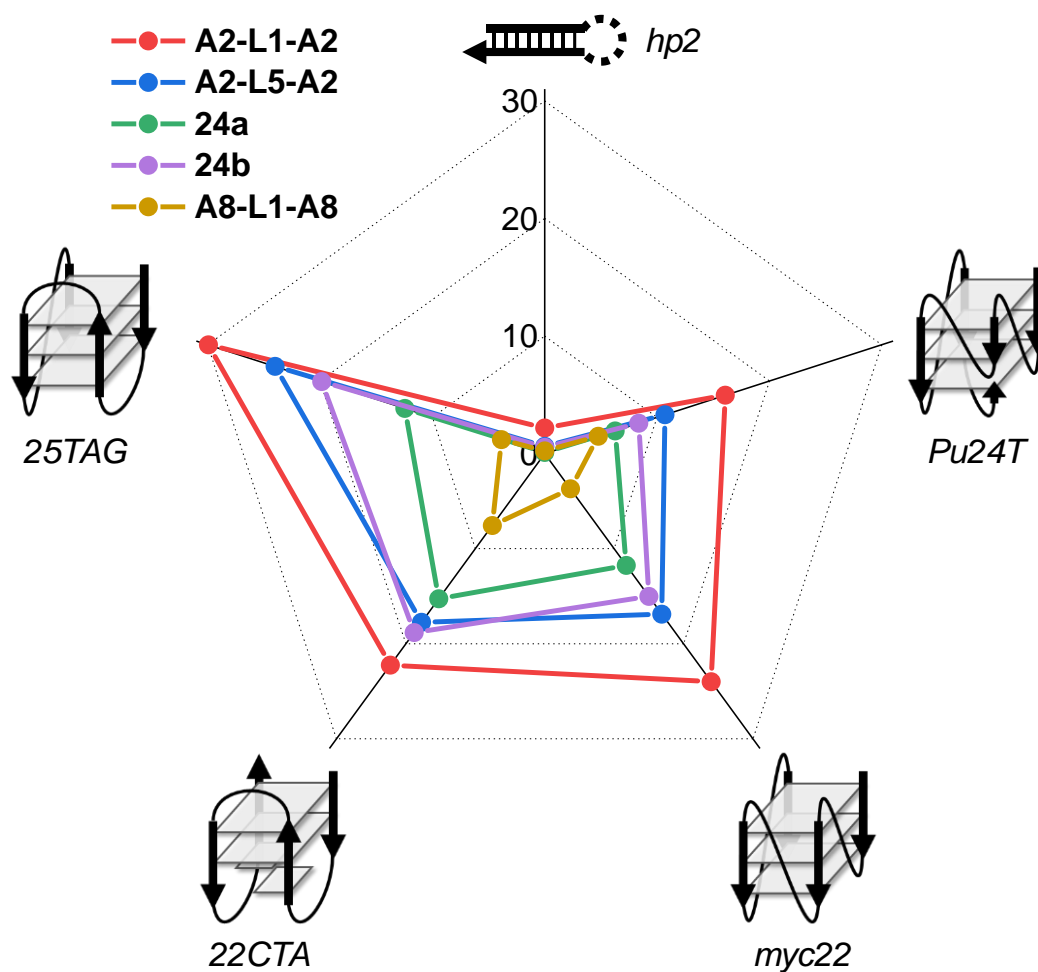


Figure 59. Ligand-induced stabilization ( $\Delta T_m$ , °C) of G4-DNA (Pu24T, myc22, 22CTA, 25TAG) and hairpin control (hp2) observed in fluorescence-melting experiments performed with compounds **A2-L1-A2**, **A2-L5-A2**, **24a**, **24b**, and **A8-L1-A8** [ $c(\text{DNA}) = 0.2 \mu\text{M}$ ,  $c(\text{ligand}) = 1.0 \mu\text{M}$ , buffer composition: the same as in RESULTS AND DISCUSSION, Part 1. Data are mean values from three measurements. The structures of DNA substrates are schematically shown next to each axis and the structures of compounds below the graph.

In addition, almost no drop of ligand-induced stabilization was observed when fluorescence melting experiments were performed in the presence of excess double-stranded competitor (self-complementary 26-mer *ds26*), and only in the case of *Pu24T*, the addition of the duplex resulted in decrease of the stabilization induced by ligands **24a** and **24b** by 2–3 °C (Figure 60). This observation agrees with the results of DCC pull-down experiments, which also indicated moderate selectivity of **A2-L1-A8** for *Pu24T* with respect to duplex DNA, as non-negligible amounts of this ligand were pulled-down with *dT22* and *hp2* (Figure 53).

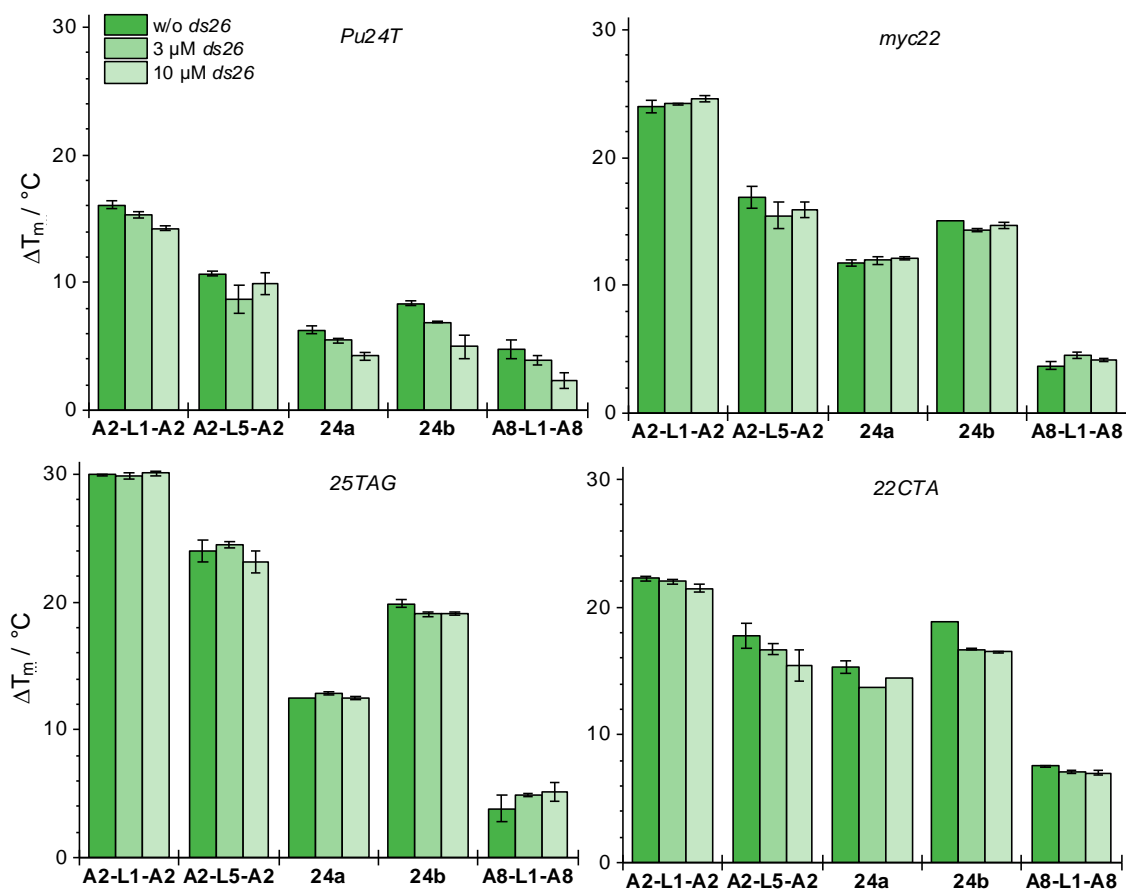


Figure 60. Ligand-induced stabilization of G4-DNA (Pu24T, myc22, 25TAG and 22CTA,  $c = 0.2 \mu\text{M}$ ) observed in fluorescence-melting experiments performed with compounds **A2-L1-A2**, **A2-L5-A2**, **24a**, **24b**, and **A8-L1-A8** ( $c = 1.0 \mu\text{M}$ ) in the absence (dark green bars) and in the presence of duplex competitor ds26 (3 and 10  $\mu\text{M}$ , green and pale green bars, respectively). Data are mean  $\pm$  s.d. from three measurements.

## 2.6 Conclusions

In this part of thesis, we developed a DCC approach based on the formation of DCLs of acylhydrazones in the conditions compatible with native structures of G4-DNA (aqueous media, near-physiological pH), followed by identification of putative best binders through a pull-down with biotinylated oligonucleotides immobilized on streptavidin-coated magnetic beads. This approach was initially validated using DCL1, where the ranking of pulled-down

ligands was found in a good agreement with their G4-binding properties. Subsequently, we constructed two novel libraries, DCL2 and DCL3, which pointed out to several novel bis(acylhydrazones), including **A2-L1-A8**, **A2-L5-A8**, and **A2-L5-A2**, as promising G4-DNA binders. A preparative synthesis of **A2-L1-A8** was attempted but was not successful. Indeed, to the best of our knowledge, no non-symmetric bis(acylhydrazone) was reported in the literature, so far.

Due to the hard synthetic accessibility of non-symmetric bis(acylhydrazones), compound **A2-L1-A8** was emulated through two isomeric carboxamide/acylhydrazone hybrids **24a** and **24b**. Both hybrids, as well as the bis(acylhydrazone) **A2-L5-A2**, indeed demonstrated fair affinity to various G4-DNA substrates (with  $K_d$  values in the 10–650 nM range) and a good level of G4-vs.-duplex selectivity, as evidenced by isothermal fluorimetric titrations, mass-spectrometric and fluorescence-melting experiments, but failed to outperform the prototype ligand **A2-L1-A2**.

At the current stage, we were not able to obtain ligands with significant selectivity towards one or another G4 topology or another structural feature. Development of such ligands would probably require generation of DCLs from building blocks featuring higher structural diversity and/or more complex structural motifs (e.g., peptide derivatives) as well as optimized pull-down protocols. Nonetheless, our approach resulted in hybrid ligands (**24a** and **24b**) whose fair G4 affinity combines with favourable drug-like properties, as illustrated by the values of the corresponding physico-chemical descriptors (calculated using SwissADME).<sup>173</sup>

Table 12.  $F(sp^3)$  and  $clogP$  of compounds **24a**, **24b** and **A2-L1-A2**.

Ligand	$F(sp^3)$	$clogP$
<b>24a</b>	0.21	-1.3
<b>24b</b>	0.21	-1.7
<b>A2-L1-A2</b>	0.07	-7.1

Thus, non-symmetric ligands combining an N-methylquinolinium moiety with another aromatic moiety represent an interesting motif for the development of other biologically active G4-DNA binders.

# 3. Identification of Optimal G-quadruplex Ligands by Combining Parallel Synthesis and Profiling of Cationic Bis(acylhydrazones)

As mentioned in *AIMS OF THE WORK*, one of fall-back solutions envisaged for this project was to develop an approach of classical combinatorial chemistry by performing multiple parallel combinatorial synthesis of acylhydrazones, followed by a subsequent study of their binding to nucleic acid targets by high-throughput screening techniques. Towards this end we designed a combinatorial chemistry strategy based on the straightforward preparation of “ready-to-screen” samples of cationic bis(acylhydrazones), coupled with high-throughput profiling of their binding to a panel of G4-DNA targets. The implementation of this method can lead to an easy, rapid and not-expensive way to analyze numerous compounds for their affinity to the target.

## 3.1 Library design and synthesis.

### 3.1.1 Synthesis of aldehydes

The selection of currently used aldehydes is presented on Chart 8.



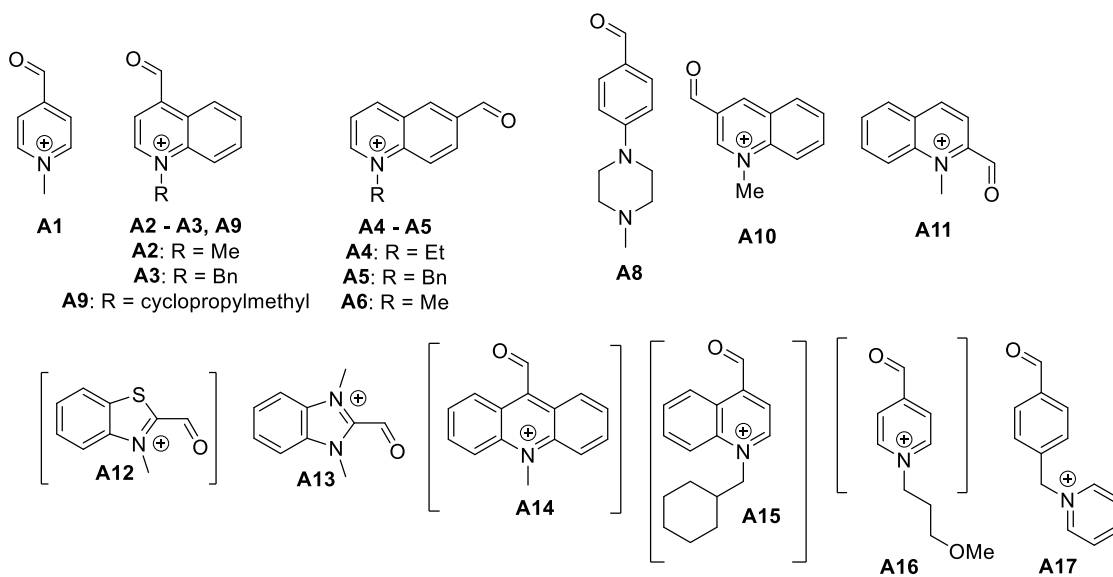
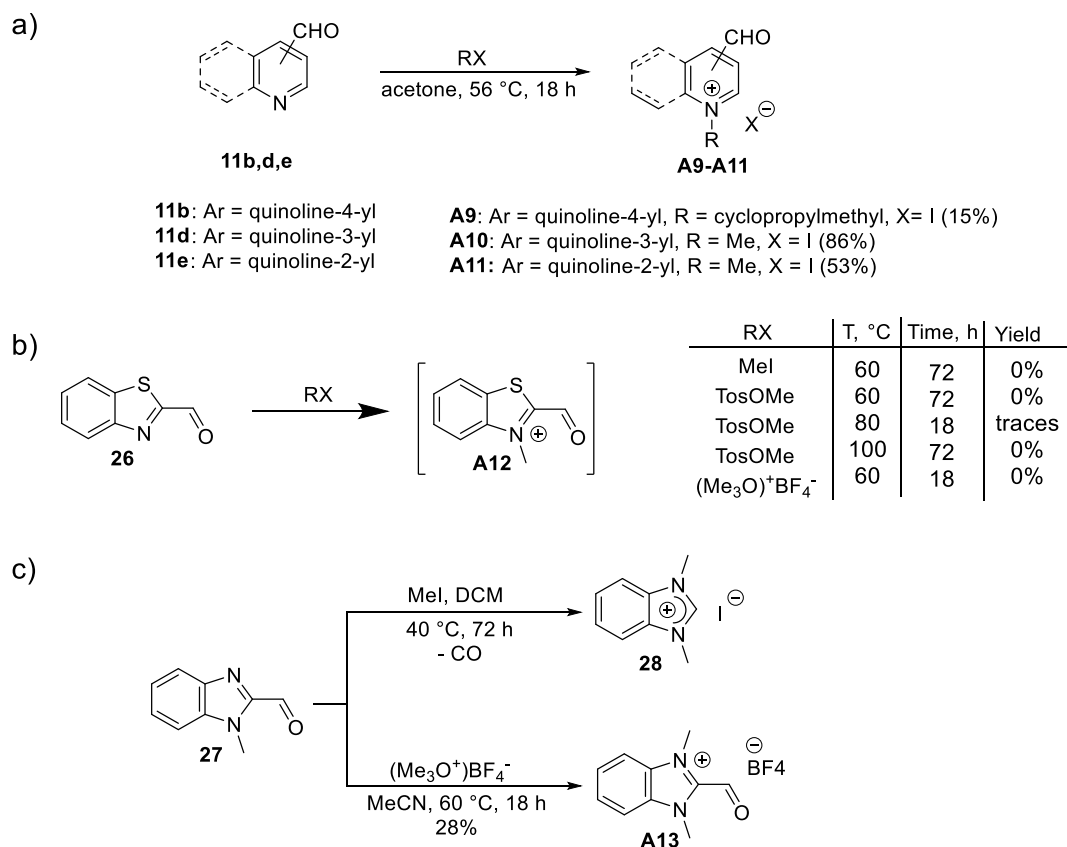


Chart 8. Selection of aldehyde building blocks. The compounds shown in brackets could not be obtained.

Aldehydes **A1–A6** were already synthesized in the course of this work (cf. *RESULTS AND DISCUSSION, Parts 1 and 2*). We prepared aldehydes **A9–A11** using the same conditions as for **A2–A5** (22 equivalents of alkylation agent, boiling acetone, 18 hours, *Scheme 12a*).



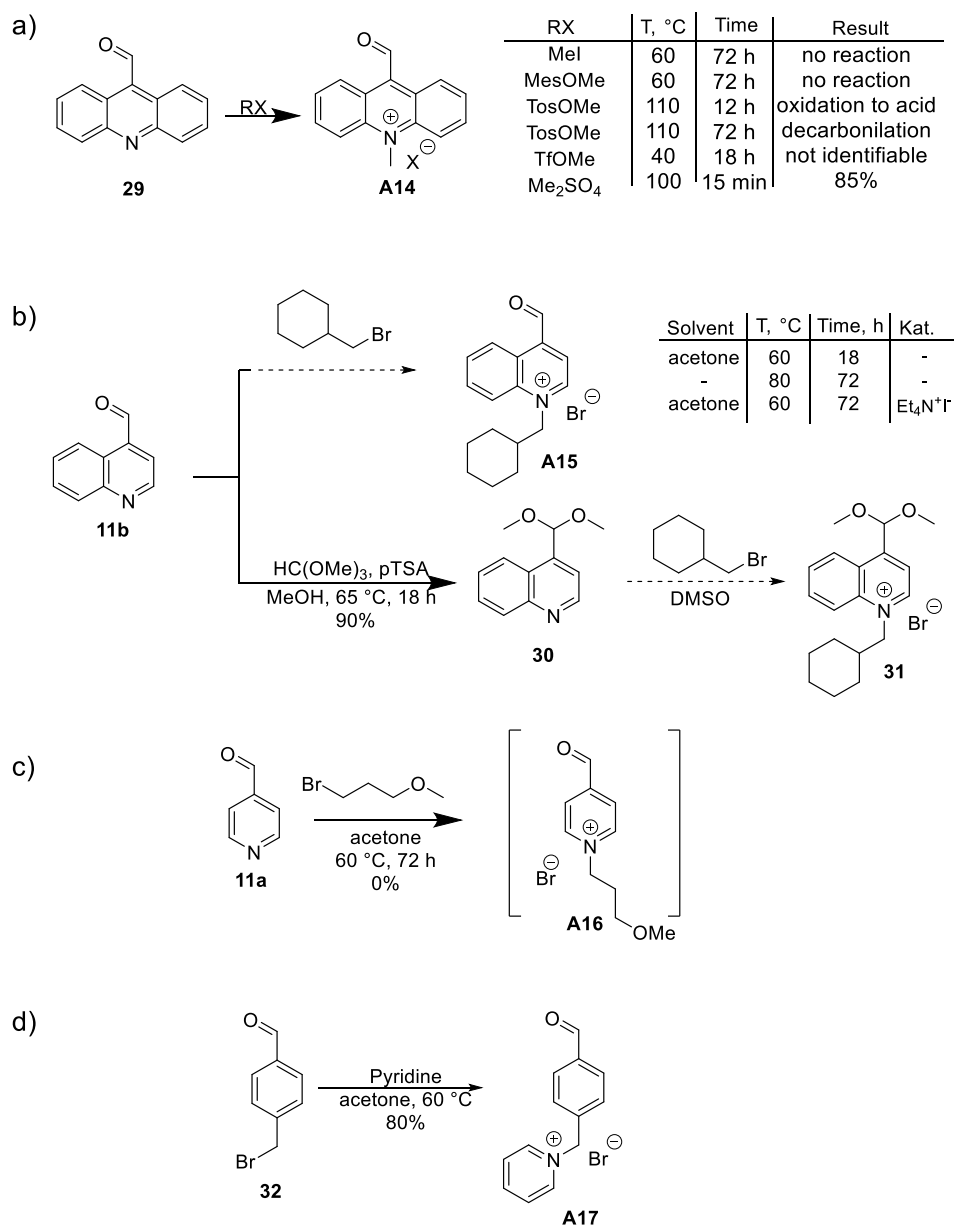
Scheme 12. Synthesis of quaternized aldehydes **A7–A16**.

The attempts to methylate benzothiazole-2-carbaldehyde **26** by iodomethane or methyl *p*-toluene sulfonate at 60 °C during 18 and 72 hours respectively were ineffective (Scheme 12b). On the contrary, heating with methyl *p*-toluenesulfonate without solvent at 60 °C for 18 hours gave traces of product **A12**. Attempt to raise the temperature to 100 °C leads to products of polymerization.

The methylation of 1-methyl-benzoimidazole-2-carbaldehyde **27** with iodomethane in DCM at room temperature during 72 hours results in a precipitation of light-yellow crystals of 1,3-dimethyl-benzoimidazol-3-ium **28** (Scheme 12c) that presumably formed through decarbonylation of **27**. The assay with trimethyloxonium tetrafluoroborate, on the contrary, was successful and the desired product **A13** was isolated in 28% yield. This aldehyde is rather stable in water that was confirmed by <sup>1</sup>H NMR.

The attempt to synthesize 9-formyl-10-methylacridinium **29** was also problematic (Scheme 13a). While no reaction occurs with methyl iodide or methyl methanesulfonate in different

conditions, with methyl trifluoromethanesulfonate the alkylating agent was consumed, but the product could not be identified. A reaction with 12 equivalents of dimethylsulfate without solvent at 100 °C during 15 minutes gave the desired product contaminated with 15% of starting material.



Scheme 13. Synthesis of aldehydes **A14-A18**.

The attempt to quaternize quinoline-4-carbaldehyde **11b** with cyclohexylmethyl bromide in the “standard” conditions (acetone, 60 °C, 18 hours) gave no result (*Scheme 13b*). Stirring the mixture of reagents without solvent at 80 °C for 72 hours, without or with catalyst did not lead to formation of **A15** neither. The attempt to incorporate cyclohexylmethyl group in the compound **30** with a protected aldehyde group (to increase the nucleophilicity of the endocyclic N atom) also failed.

The attempt to quaternize pyridine-4-carbaldehyde **11a** with methoxypropylbromide in “standard” conditions was not successful (*Scheme 13c*).

In contrast, aldehyde **A17** was easily obtained from pyridine and 4-bromomethylbenzaldehyde (*Scheme 13d*). This aldehyde is different from all other aldehydes of the series by that the quaternary N atom and the aldehyde group are located in different ring systems.

### ***3.1.2 Synthesis of bis(acylhydrazides)***

The set of bis(acylhydrazides) (*Chart 9*) was constructed to include heterocyclic derivatives with a varied number of condensed rings (**L1–L5**, **L11–L12**), a benzene counterpart (**L10**), three aliphatic derivatives with varied length of the spacer between the acylhydrazide functions (**L7–L9**), and carbohydrazide (**L6**) as the minimal dihydrazide.

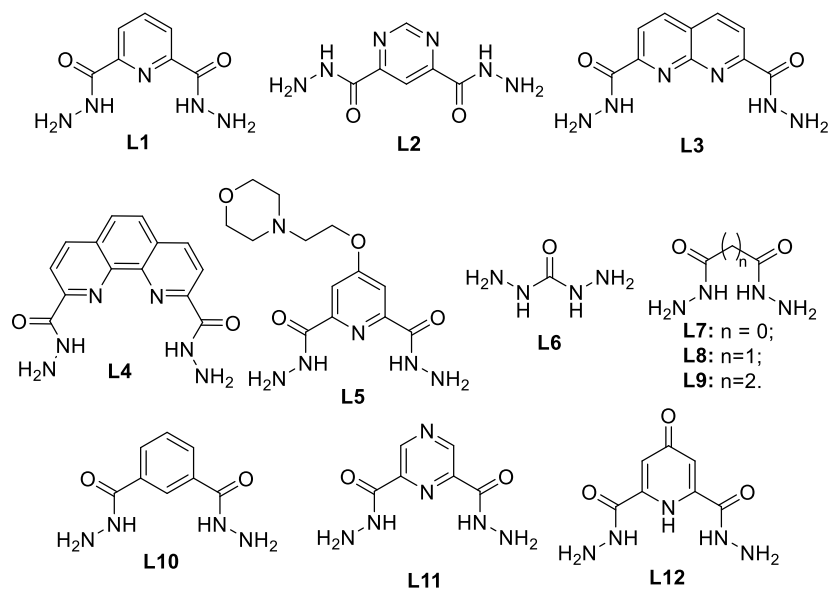
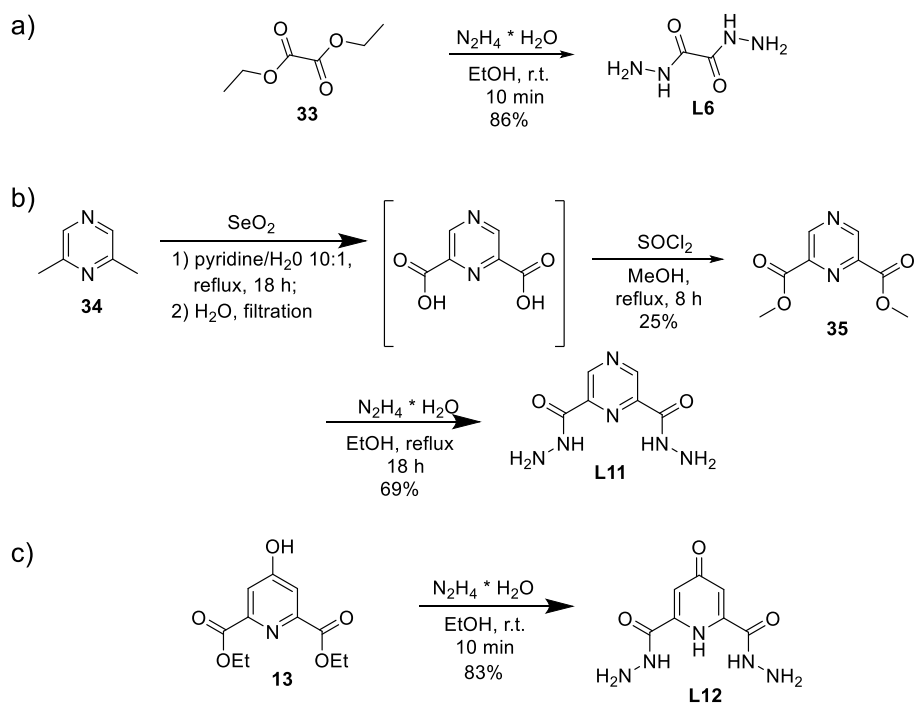


Chart 9. Selection of bis(acylhydrazides)

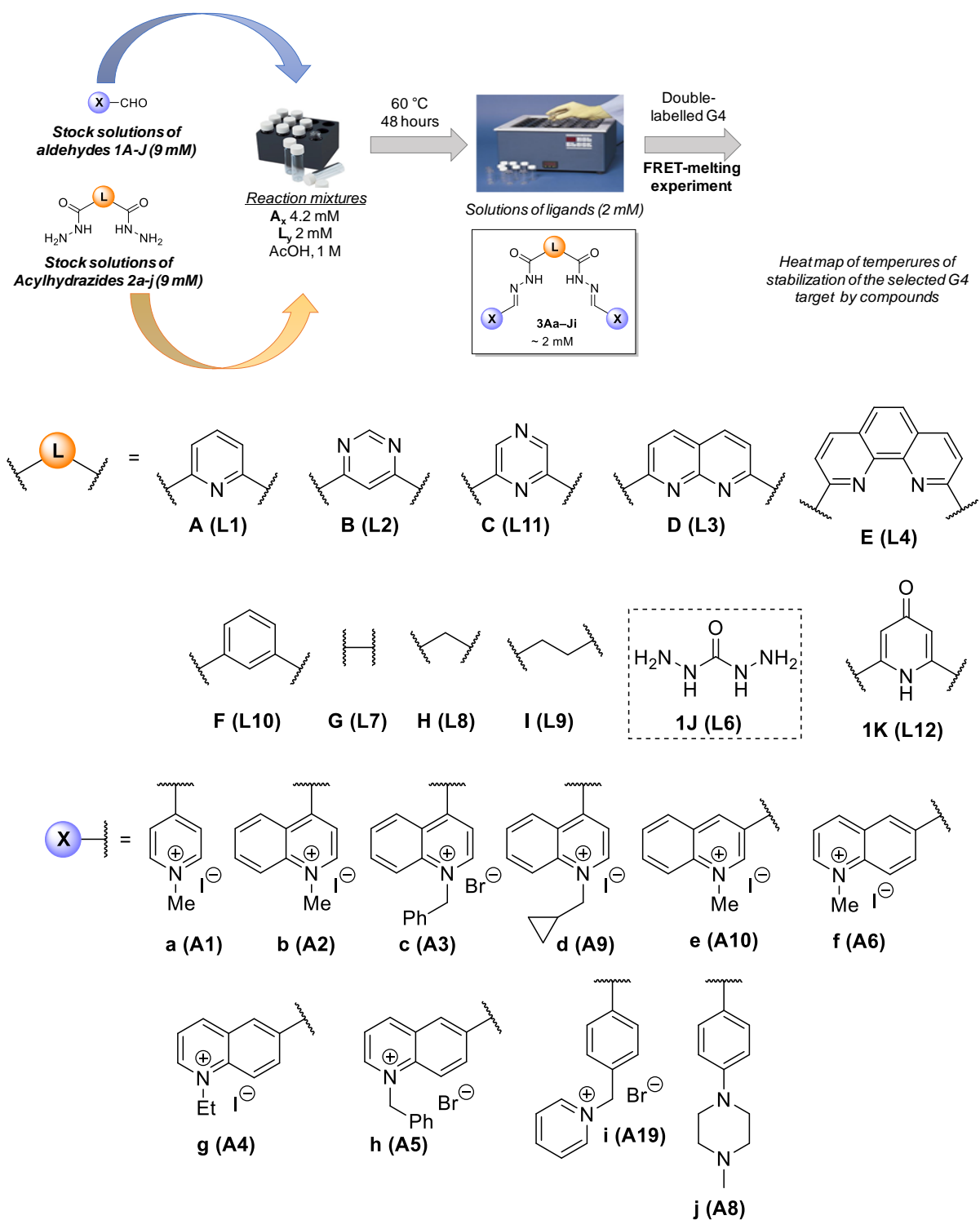
Acylhydrazides **L1–L5** were synthesized in the course of this work (cf. *RESULTS AND DISCUSSION, Parts 1 and 2*). Compounds **L6**, as well as **L8–L10** were purchased. Compounds **L7** and **L12** were prepared from the corresponding esters (*Scheme 14a* and *c*). The pyrazine derivative **L11** was prepared starting from 2,6-dimethylpyrazine **31** (*Scheme 14b*).<sup>174</sup>



Scheme 14. Synthesis of bis(acylhydrazides) **L6**, **L11–L12**.

### 3.1.3 Optimization of conditions of acylhydrazone synthesis

Our previous experiments towards the synthesis of model bis(acylhydrazones) indicated that these products can be obtained by heating the aldehyde and bis(acylhydrazide) reagents in a suitable solvent. Here, we attempted to prepare “ready-to-screen” samples of bis(acylhydrazides) by performing reactions in small volume ( $\approx 1$  ml) of the solvent without isolation of product, considering that the conversion of the starting materials is nearly-quantitative and that minor impurities present in such samples (i.e. unreacted starting material, monoacylhydrazides) have little impact on the outcome of the screening experiment. In order to generate “ready-to-screen” solutions of the products, stock solutions of eleven bis(acylhydrazides) **1A–K** (9 mM in DMSO) were mixed with solutions of ten aldehydes **2a–j** (9 mM in DMSO) in a one-by-one fashion, to give the reaction mixtures containing 2.0 and 4.4 mM (i.e. 2.2 molar equivalents) of dihydrazide and aldehyde components, respectively (*Scheme 15*; the names of aldehydes and bis(acylhydrazides) were changed to follow the results easier). The reaction mixtures were supplemented with AcOH as a catalyst (final concentration: 1 M), and heated at 60 °C for 48 h to ensure complete conversion of the dihydrazide. The use of higher temperature (100 °C) and shorter reaction time was disadvantageous, as it led to the formation of significant amounts of dealkylated products. After several preliminary experiments, chelidamic acid dihydrazide (**1K**) was discarded from the set of bis(acylhydrazides) due to unwanted formation of Michael adducts in the course of reaction. Aldehyde **2j** (**A6**) was also tested, but was discarded due to its low reactivity in the mentioned conditions. After exclusion of these building blocks, the resulting solutions were analyzed by LC/MS and revealed an average purity of 87% and minimal purity of 80%, as per peak area (*Table 13*). It goes in a range with commercial combinatorial libraries the purity of which normally is around 80-90%. The as-synthesized sample solutions were employed for the screening by fluorescence-melting experiments without further treatment (*Scheme 15*).



Scheme 15. Library design and one-step synthesis of “ready-to-screen” solutions of bis(acylhydrazones).

Table 13. Purity of as-synthesized samples of bis(acylhydrazones) **3Aa–Ji** as determined by LC/MS analysis (UV peak area, %).<sup>a</sup>

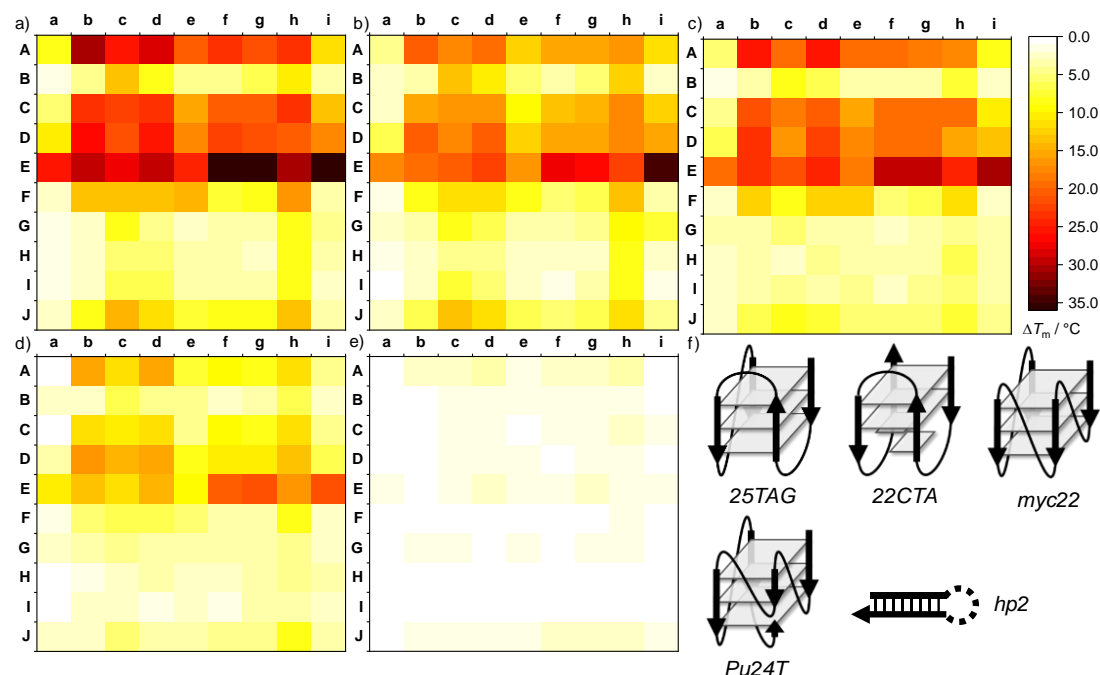
Aldehyde	Bis(acylhydrazide)									
	1A	1B	1C	1D	1E	1F	1G	1H	1J	1I
<b>2a</b>	87	86	84	91	81	89	99	99	99	99
<b>2b</b>	84	86	81	81	84	91	83	80	80	81
<b>2c</b>	85	83	87	84	96	89	90	95	81	88
<b>2d</b>	93	91	90	94	85	90	90	80	83	91
<b>2e</b>	92	93	93	93	93	92	93	93	98	93
<b>2f</b>	85	84	83	86	88	80	82	94	88	87
<b>2g</b>	89	88	90	89	89	92	89	93	94	88
<b>2h</b>	88	87	89	88	92	88	83	84	94	83
<b>2i</b>	80	82	80	80	90	85	84	86	85	86

<sup>a</sup> Conditions: colon Waters Alliance 2695 equipped with a Waters XBridge C<sub>18</sub>-3.5 μm column. Eluent A: H<sub>2</sub>O + 0.1% HCOOH, eluent B : acetonitrile + 0.1% HCOOH, gradient elution with 2 to 100% of eluent B.

### 3.2 Screening of interaction with G4-DNA by FRET-melting experiments.

The capacity of the “as-synthesized” ligands to stabilize G4-DNA was assessed by FRET-melting experiments, which monitor thermal denaturation of G4-forming oligonucleotides labelled with a fluorophore–quencher pair (cf. *Introduction, Part 1.4.2*). Four representative G4-DNA targets belonging to different topology classes were selected for screening: 25TAG (hybrid-2); 22CTA (antiparallel), *myc22* (parallel), and *Pu24T* (snap-back parallel), as schematically presented in *Figure 61f* (for more detailed information about sequences cf. *Introduction, Parts 1.1.3–1.1.4*). A hairpin with 8 base pairs in the stem part (*hp2*) was included as a mimic of double-stranded DNA. The conditions of melting experiments were used the same as described in *RESULTS AND DISCUSSION, Part 1*. The results of FRET-melting experiments, presented as heatmaps of ligand-induced stabilization ( $\Delta T_m$ ), are presented in *Figure 61, a-e* (The corresponding values are given in Table A1, Annex).





**Figure 61.** Results of FRET-melting assay with 90 ‘as-synthesized’ bis(acylhydrazone) ligands **3Aa–Ji**. a)–e) Heatmaps of ligand-induced stabilization ( $\Delta T_m$ ) of G4-DNA: a) 25TAG, b) 22CTA, c) myc22, d) Pu24T, e) hp2; f) Schematic depiction of G4 and hp2 substrates used in the assay; all oligonucleotides were labelled with a fluorophore (5′-FAM) and a quencher (3′-TAMRA). Data are average values from three technical replicates. Conditions:  $c(\text{G4-DNA}) = 0.2 \mu\text{M}$ ,  $c(\text{ligands}) \approx 1 \mu\text{M}$ .

The results of FRET-melting screening present a number of interesting regularities. With respect to molecular structure, the capacity of bis(acylhydrazones) to stabilize, or not, G4-DNA substrates seems to be mostly due to the combination of two major factors: the total number of aryl rings ( $N_{\text{Ar}}$ ) and the nature of the central moiety **L**. The importance of the first factor is demonstrated by the strong positive correlation between  $N_{\text{Ar}}$  and ligand-induced stabilization ( $\Delta T_m$ , Figure 62). With respect to the nature of the central moiety, the derivatives containing pyridine (**A**), pyrazine (**C**), naphthyridine (**D**), and, in particular, phenanthroline (**E**) fragments, generally, demonstrated very high capacity to stabilize G4-DNA, in stark contrast to the analogues containing pyrimidine (**B**) and benzene (**F**) fragments, as well as all aliphatic derivatives (**G–J**). Among the latter, only carbonylhydrazone derivatives **3Jc** and **3Jh** ( $N_{\text{Ar}} = 6$ ) demonstrated significant level of thermal stabilization of G4-DNA ( $\Delta T_m = 12\text{--}14 \text{ }^\circ\text{C}$ , with 25TAG and 22CTA). This fact confirms the assumption that G4-binding properties are directly related to the ability of the ligand to adapt a U-shaped conformation, stabilized by water-mediated

hydrogen bonds between the acylhydrazone groups and the central heterocyclic unit, as was observed in the solid-state structures of **3Aa** and **3Ea** (cf. *RESULTS AND DISCUSSION, Part 1.3*).

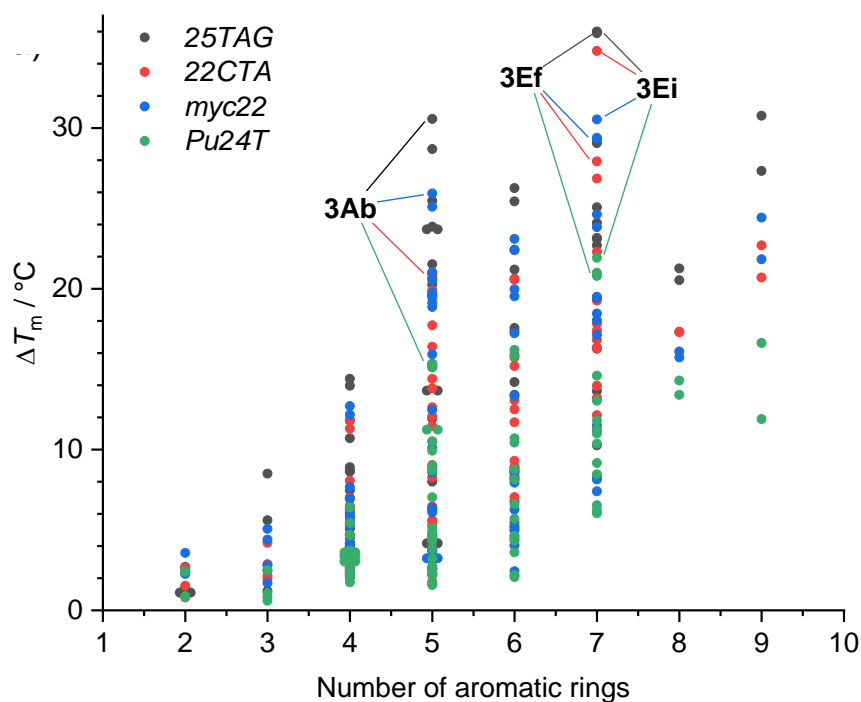


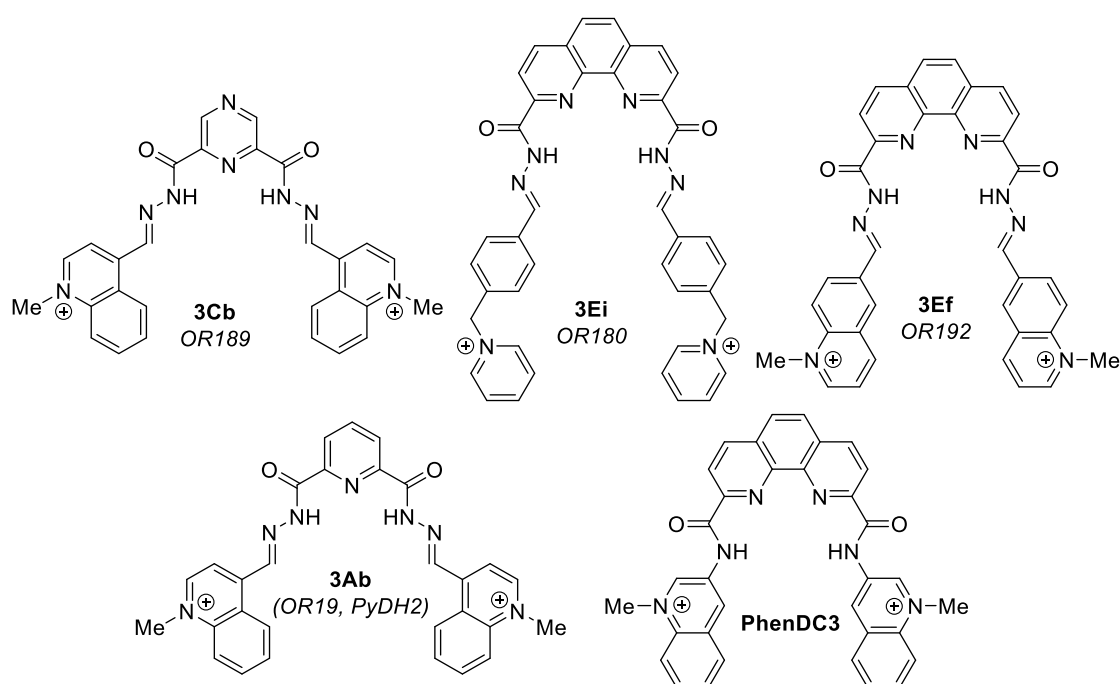
Figure 62. a) Plots of ligand-induced stabilization ( $\Delta T_m$ ) as a function of number of aromatic rings ( $N_{Ar}$ ) for 90 ligands and four G4-DNA substrates.

The nature of lateral moieties (**X** in *Scheme 15*) appears to play a secondary role in G4-stabilizing properties. Thus, 1-methylpyridinium (**X** = **a**) and 4-(pyridinium-1-yl)methylphenyl (**X** = **i**) substituents, generally, resulted in less efficient G4 stabilizers, comparing with quinolinium fragments, *except* when combined with the phenanthroline unit such as in compounds **3Ea** and **3Ei**, respectively. In contrast, in the case of quinolinium substituents (**X** = **b–h**), the G4-binding properties of the ligands were relatively insensitive with respect to the nature or position of the side chain.

Most ligands behaved relatively uniformly with respect to different G4-DNA targets, considering the generally lower propensity of parallel G4 (*myc22* and, particularly, *Pu24T*) to ligand-induced stabilization as compared to *25TAG* and *22CTA* (*Figure 61a–d*). On the other hand, none of compounds had any significant effect on the melting of the hairpin substrate

*hp2* ( $\Delta T_m = 0$  to  $4.0$  °C, *Figure 61e*), demonstrating an excellent level of quadruplex-vs.-duplex selectivity.

Considering the ensemble of the data, compounds **3Ef** and **3Ei** (*Figure 63*) that consistently demonstrated the highest stabilization of most G4 substrates, as well as **3Cb**, representative of moderate-affinity G4 binders of the series, were selected and synthesized in a preparative fashion, yielding analytically pure samples. Their G4-binding properties were studied in more detail and in comparison with compounds **3Ab** (**PyDH2**) and **PhenDC3**.



*Figure 63. Structures of novel bis(acylhydrazone) derivatives selected for detailed studies (3Cb, 3Ei, 3Ef) and structure of positive controls 3Ab (PyDH2) and PhenDC3.*

### 3.3 Hit validation.

#### 3.3.1 FRET-melting

The conditions of FRET-melting experiments employed in the screening step ( $\sim 1.0$   $\mu\text{M}$  of ligand, i.e., five molar equivalents with respect to G4-DNA) were selected in order to compensate for possible differences in ligand concentrations in the “ready-to-screen”

samples. However, these conditions are not appropriate for the ranking of high-affinity binders, since the stabilization effect induced by these compounds ( $\Delta T_m \geq 35$  °C) reaches the experimental limit of the method, and lower concentrations of ligands (0.5  $\mu$ M) are recommended for the assessment of high-affinity binders.<sup>175</sup> For the pure ligands, we studied the ligand-induced stabilization of the four G4-DNA substrates in a range of lower ligand concentrations (0.2–1.0  $\mu$ M). The results (*Figure 64*) demonstrated that, in all cases, almost no stabilization was observed with 0.1 or 0.2  $\mu$ M of ligands. With *25TAG*, compounds **3Ei** and **PhenDC3** yielded an almost maximal stabilization effect ( $\Delta T_m > 30$  °C) already at a 0.4  $\mu$ M concentration, whereas other ligands were slightly less efficient, as similar stabilization was achieved at only higher concentrations (0.8–1.0  $\mu$ M). With *22CTA* and *myc22*, the concentration dependence of the stabilization effect was less steep for all ligands. In all cases, the stabilization effect observed with 1  $\mu$ M of ligand matched the value observed in the screening experiments ('Screen' in *Figure 64*), giving additional proof that this experiment has an excellent tolerance with respect to minor differences in ligand concentration and the presence of potential impurities resulting from the synthesis procedure. Altogether, a clear ranking of ligands could be observed with all four G4-DNA: **3Ei**  $\geq$  **PhenDC3** > **3Ef** > **3Ab** > **3Cb**; the superior activity of compound **3Ei**, compared with PhenDC3, was particularly remarkable in the case of *22CTA* (*Figure 64b*).

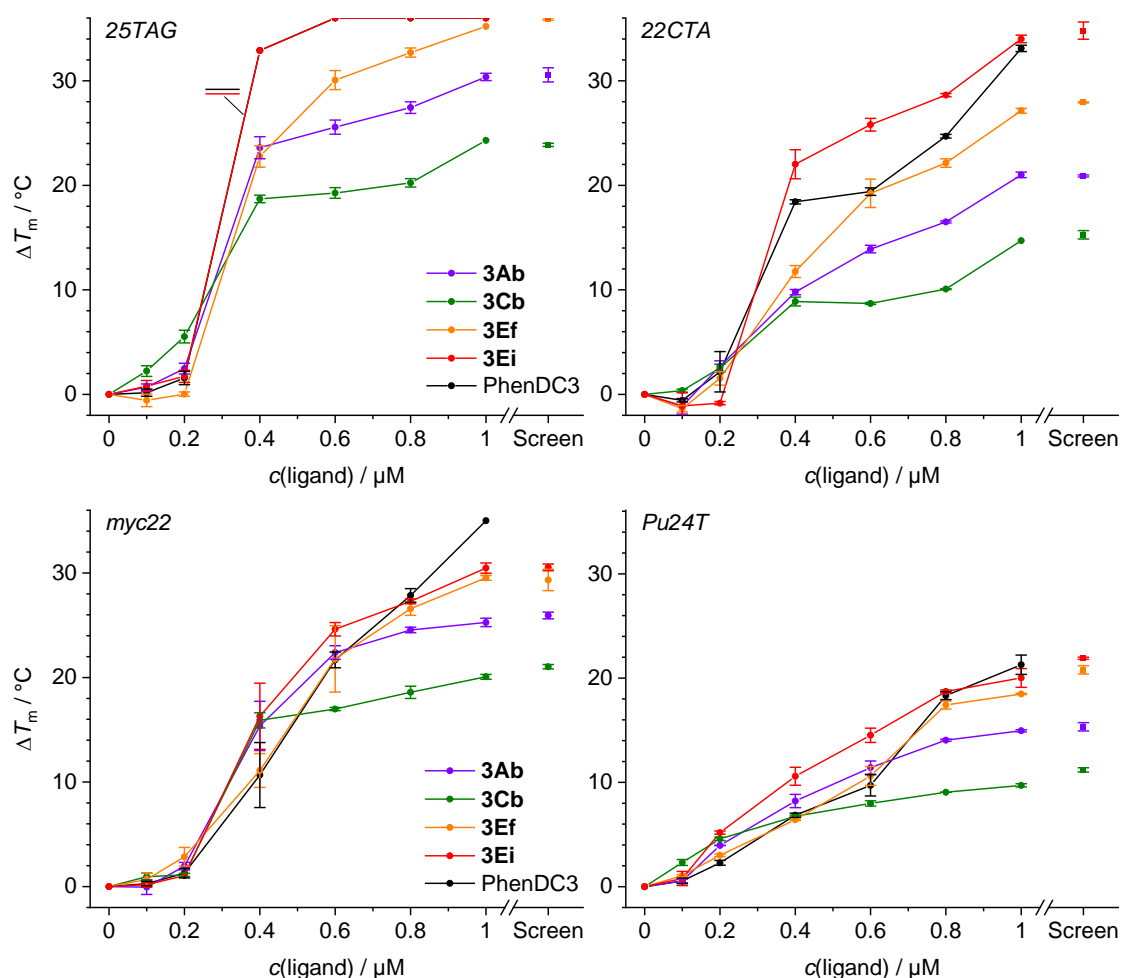


Figure 64. Concentration-dependent stabilization of G4-DNA observed in fluorescence melting experiments. Conditions:  $c(\text{G4-DNA}) = 0.2 \mu\text{M}$ . Data are means  $\pm$  s.d. from three technical replicates. The upper limit of  $\Delta T_m$  detection is 36 °C. The  $\Delta T_m$  value observed in screening experiments with ‘as-synthesized’ ligands (“Screen”) is plotted for comparison.

### 3.3.2 FID assay

Next, the G4-binding properties of these ligands were evaluated by the fluorescent indicator displacement (FID) assay, which assesses the capacity of ligands to displace Thiazole Orange (TO), a probe that is dark in the unbound state but fluorescent in the complex with G4-DNA (cf. Introduction, Part 1.4.3). In this assay, the capacity of the ligands to bind to G4-DNA is quantified by  $\text{DC}_{50}$  values (i.e., the concentration required to decrease the fluorescence of G4-bound TO by half of its initial intensity). An additional G4-DNA substrate (24TTG, a variant of human telomeric sequence forming a *Type 1* hybrid structure in  $\text{K}^+$  conditions) was included

in this assay to complement the *Type 2* hybrid structure (25TAG). Remarkably, the results (Table 14, Figure 65) demonstrated that the relative capacity of ligands to bind to G4-DNA was inverted compared to the one observed in melting experiments. Specifically, pyridine and pyrazine derivatives (**3Ab** and **3Cb**, respectively) were systematically more active than phenanthroline derivatives (**3Ef**, **3Ei**, and **PhenDC3**). Among the latter, compound **3Ei** demonstrated an intriguing selectivity towards anti-parallel G4-DNA (22CTA,  $DC_{50} = 0.23 \mu\text{M}$ ), comparing with all other G4 substrates ( $DC_{50} = 0.31\text{--}1.2 \mu\text{M}$ ), which in line with the high stabilization effect of this compound observed in melting experiments (Figure 64b). It should be noted, however, that  $DC_{50}$  values lower or too close to the concentration of G4-DNA substrate employed in the FID assay (i.e.,  $0.25 \mu\text{M}$ ) should be treated with caution and interpreted merely as an evidence of high affinity of the ligand ( $K_d \leq 0.25 \mu\text{M}$ ). None of tested ligands had the capacity to effectively displace TO from its complex with *hp2* at the tested range of concentrations in this assay and therefore they all have  $DC_{50} > 2.5 \mu\text{M}$  for *hp2* (data not shown).

Table 14. Ligand-induced displacement of TO ( $0.5 \mu\text{M}$ ) from G4-DNA substrates ( $0.25 \mu\text{M}$ ). Data are presented as  $DC_{50}$  values (i.e., concentration of the ligand required to displace 50% of TO) obtained from three technical replicates.

Ligand	G4-oligonucleotide				
	<i>Pu24T</i>	<i>myc22</i>	22CTA	24TTG	25TAG
<b>3Ab (PyDH2)</b>	$0.376 \pm 0.012$	$0.287 \pm 0.005$	$0.240 \pm 0.017$	$0.272 \pm 0.013$	$0.272 \pm 0.016$
<b>3Cb</b>	$0.340 \pm 0.004$	$0.255 \pm 0.007$	$0.219 \pm 0.004$	$0.262 \pm 0.007$	$0.287 \pm 0.010$
<b>3Ef</b>	$0.573 \pm 0.021$	$0.600 \pm 0.013$	$0.348 \pm 0.031$	$0.390 \pm 0.023$	$0.984 \pm 0.044$
<b>3Ei</b>	$1.210 \pm 0.045$	$0.502 \pm 0.002$	$0.227 \pm 0.009$	$0.308 \pm 0.005$	$0.466 \pm 0.025$
<b>PhenDC3</b>	$0.580 \pm 0.013$	$0.559 \pm 0.021$	$0.315 \pm 0.003$	$0.321 \pm 0.017$	$0.451 \pm 0.02$

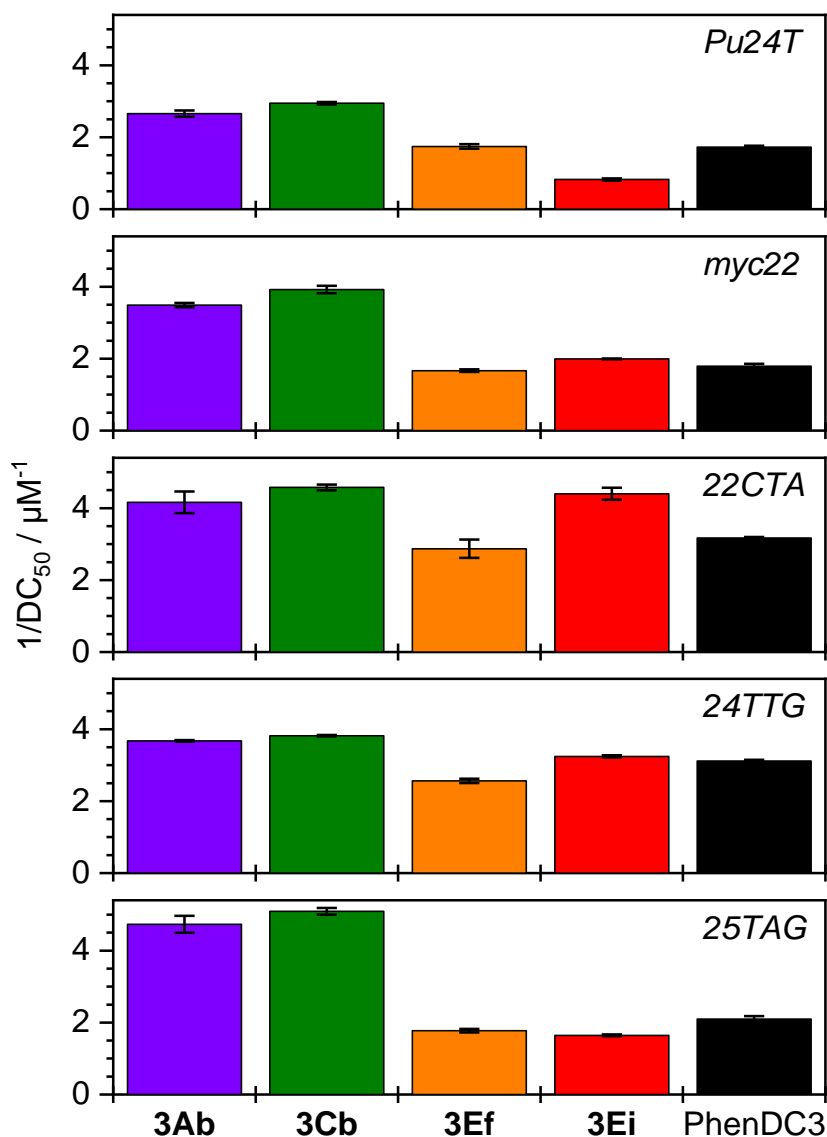


Figure 65. Ligand-induced displacement of TO (0.5  $\mu\text{M}$ ) from G4-DNA substrates (0.25  $\mu\text{M}$ ). Data are presented as reciprocal  $DC_{50}$  values (i.e., concentration of the ligand required to displace 50% of TO) obtained from three technical replicates. The experiments were performed in the K100 buffer (10 mM  $\text{LiAsO}_2\text{Me}_2$ , 100 mM KCl, 1% v/v DMSO).

### 3.3.3 Circular dichroism (CD) spectroscopy.

To evaluate the effect of bis(acylhydrazone) ligands on the conformation of G4-DNA, we recorded CD spectra of unlabelled oligonucleotides in the absence and in the presence of 0.5, 1, and 2 molar equivalents of selected ligands (Figure 66). The results demonstrate that, in the case of 25TAG, all tested ligands induce a change from the *Type 2* hybrid structure into an antiparallel form, characterized by positive peaks at 295 and 240 nm and a negative peak at

260 nm (*Figure 66a*); this change was complete with compounds **3Ab** and **3Cb**, but only partial with compounds **3Ef** and **3Cb**. The situation was opposite with *24TTG* which mostly maintained the *Type 1* hybrid structure in the presence of **3Ab**, **3Cb** and **3Ef**, but underwent a partial shift to the antiparallel form in the presence of compound **3Ei** (*Figure 66b*). Finally, CD spectra of *22CTA* (antiparallel), *myc22* and *Pu24T* (parallel G4) were not affected in the presence of the ligands, indicating the absence of conformational changes (*Figure 66c-e*).



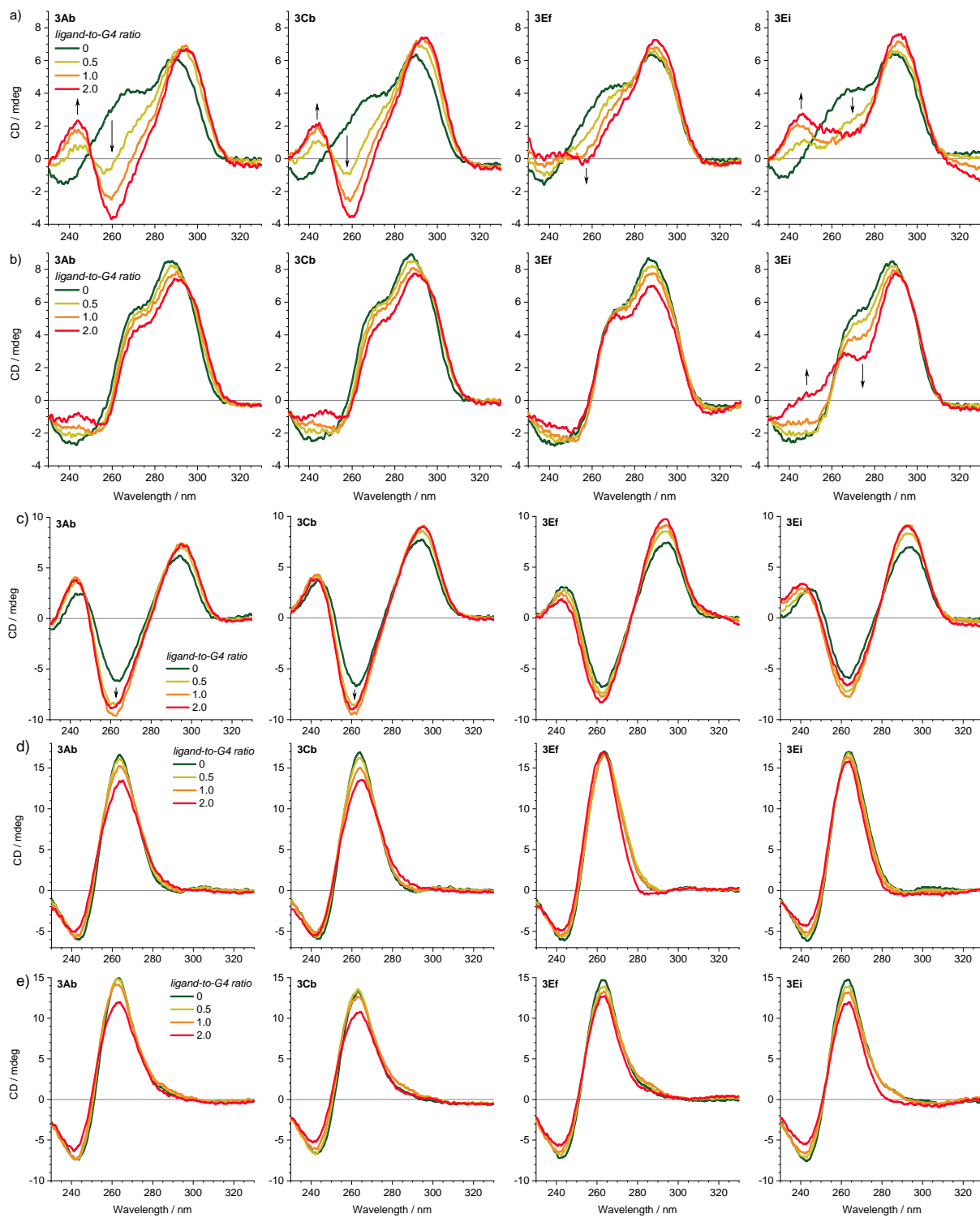


Figure 66. CD spectra of a) 25TAG, b) 24TTG, c) 22CTA, d) *c-myc* and e) Pu24T in the absence (dark green curves) and in the presence of 0.5, 1 and 2 molar equivalents (yellow to red curves) of the indicated ligands. The arrows indicate the ligand-induced changes in CD spectra.  $c(\text{G4-DNA}) = 3 \mu\text{M}$  in K-100 buffer.

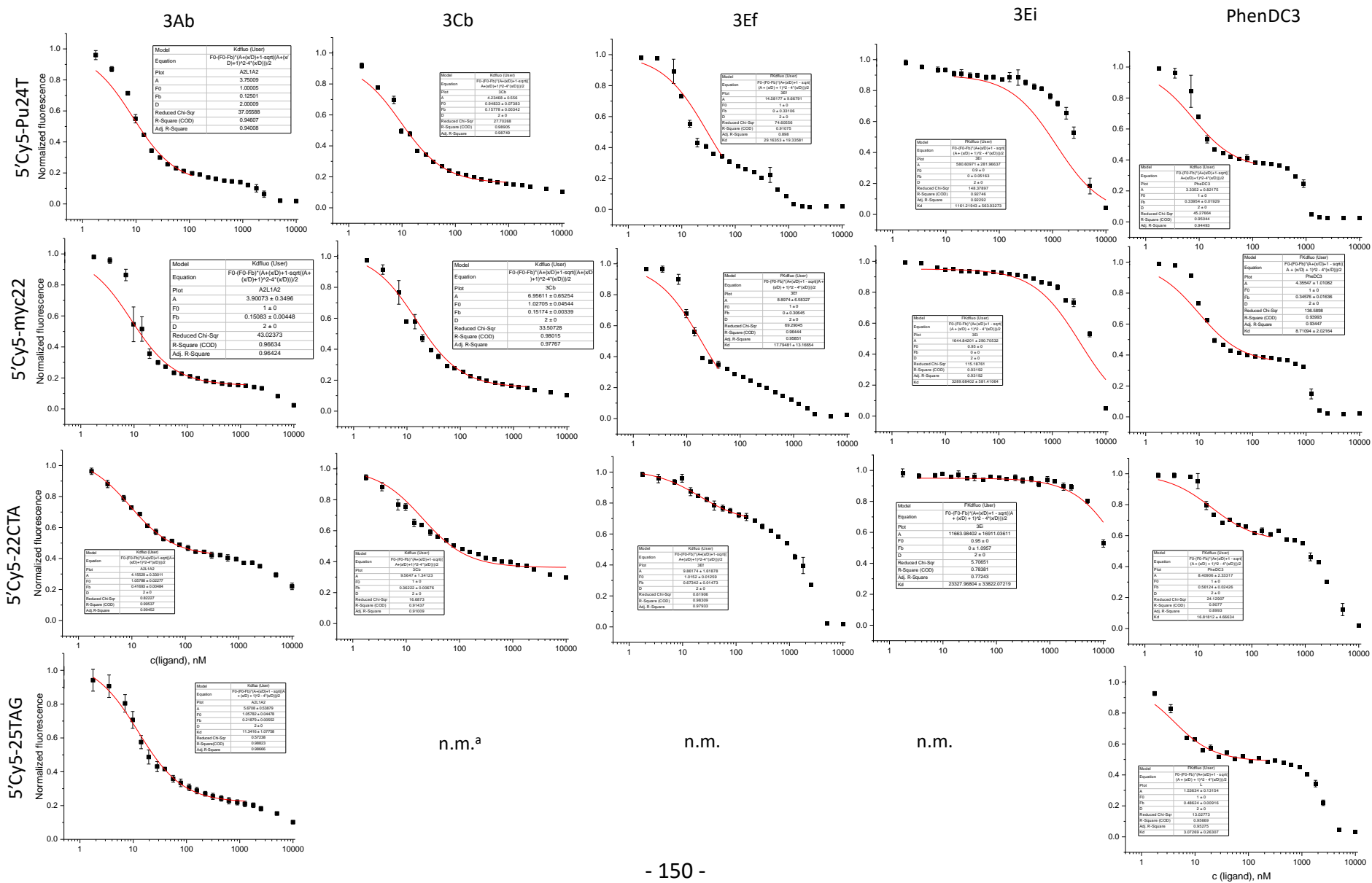
### 3.3.4 Fluorescence titrations.

We performed also fluorescence titration experiments with Cy5-labelled G4-forming sequences: 5'-Cy5-25TAG, 5'-Cy5-Pu24T, 5'-Cy5-myc22 and 5'-Cy5-22CTA. The conditions of the experiments and data treatment are similar to those describes in *RESULTS AND DISCUSSION, Part 2*. The calculated dissociation constants are presented on the *Table 15* and corresponding curves on the *Figure 67*.

*Table 15. Dissociation constants (Kd / nM) of ligands to G4-DNA and a hairpin control (hp2), determined from fluorimetric titrations with 5'-Cy5-labelled oligonucleotides.<sup>a</sup>*

Ligand	G4-oligonucleotide			
	<i>Pu24T</i>	<i>myc22</i>	<i>22CTA</i>	<i>25TAG</i>
<b>3Ab (PyDH2)</b>	7.5 ± 0.5	7.8 ± 0.3	8.3 ± 0.3	11.3 ± 1.1
<b>3Cb</b>	8.5 ± 0.6	13.9 ± 0.7	19.1 ± 1.3	n.m. <sup>a</sup>
<b>3Ef</b>	29.2 ± 19	18 ± 13	19.7 ± 1.6	n.m.
<b>3Ei</b>	1161 ± 564	3290 ± 581	> 10 000	n.m.
<b>PhenDC3</b>	6.7 ± 0.8	8.7 ± 2.0	16.8 ± 4.7	n.m.

<sup>a</sup>) Not measured.



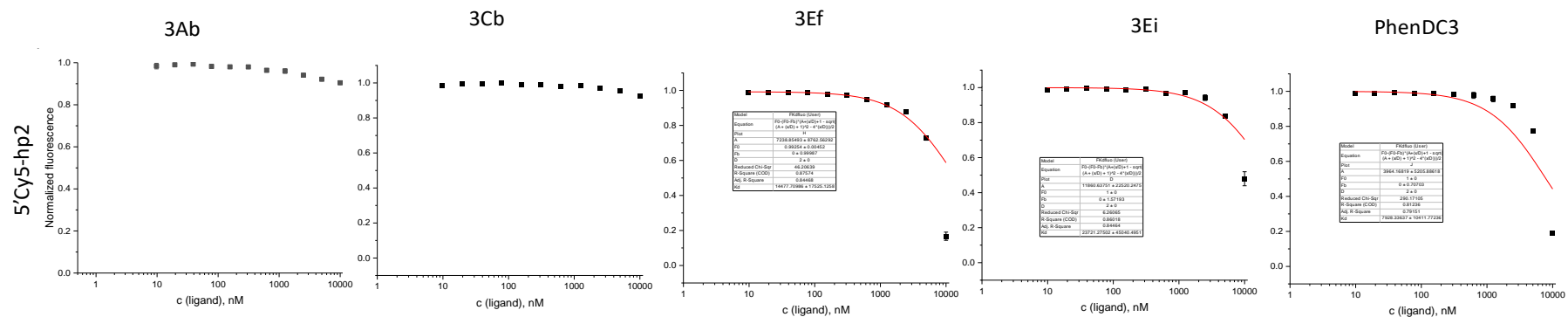
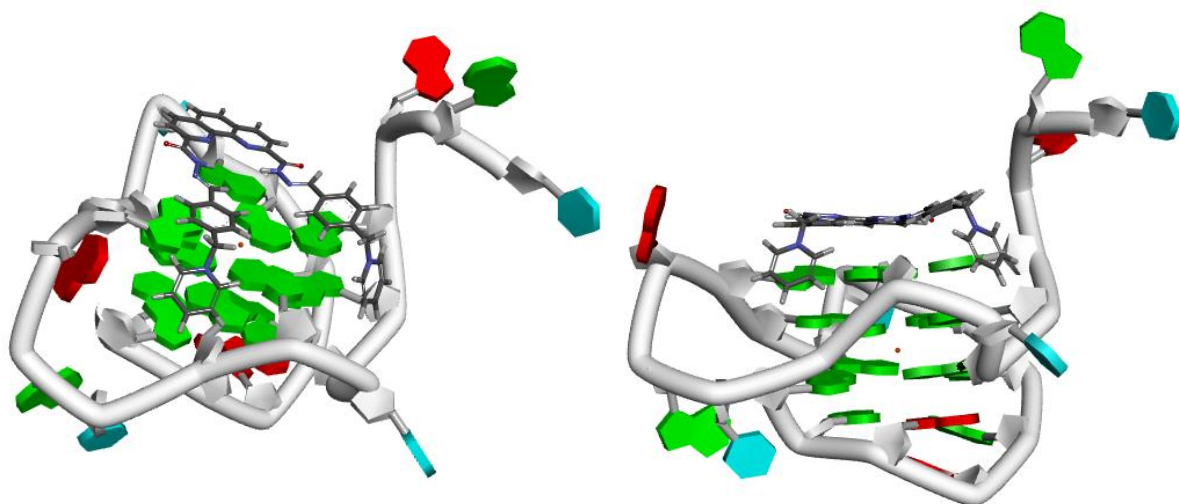


Figure 67. Normalized changes of fluorescence intensity ( $\lambda_{ex} = 590^{BP50}$  nm,  $\lambda_{em} = 675^{BP50}$  nm) upon titrations of ligands (final concentration: 1.8 nM to 10  $\mu$ M) to 5'-Cy5-labelled oligonucleotides ( $c = 2$  nM in 10 mM  $\text{LiAsO}_2\text{Me}_2$ , 100 mM KCl, 0.5 w/v % CHAPS, 0.05 v/v % Triton X-100, pH 7.2). Data are means  $\pm$  s.d. from three independent experiments. Red lines represent fits of experimental data to a 1:1 binding model (Eq. 1). <sup>a)</sup> Not measured yet.

Among the tested compounds, compounds **3Ab** and **3Cb** had very similar and high affinity to G-quadruplexes ( $K_d$  up < 20 nM), similar to **PhenDC3**. Compound **3Ef** had similar affinities (except for *Pu24T*). Very surprisingly, compound **3Ei** did not show efficient fluorescence quenching of none of G4s, including *22CTA*, to which it showed the highest affinity both in FID and in FRET assay. Such a low fluorescence quenching can be explained by two following hypotheses:

- 1) The structure of the compound **3Ei** in the complex may be not favorable for the efficient fluorescence quenching, as the positive charges are situated too far from the central unit so that the central core of **3Ei** is less electron deficient comparing to other compounds of the series. Presumably, the ligand stacks on the G-tetrad of G4 and the positively charged branches are situated in the grooves and are involved in efficient electrostatic interactions with the phosphates. To check this hypothesis, we performed a docking of **3Ei** and *Pu24T* G4 (from PDB 2MGN). Indeed, the positively charged pyridinium substituents were found almost perpendicular to the tetrad and the ligand's plane (*Figure 68*), thus far away from the fluorophore that presumably hovers over the ligand. The absence of the direct contact between the fluorophore and electron-deficient pyridinium units prevents the electron–transfer quenching.



*Figure 68. Results of docking of 3Ei and 2MGN (Pu24T) structure.*

- 2) The second hypothesis that may explain the low fluorescence quenching is that the compound binds to the tetrad (3') opposite to the one close to the fluorophore (i.e. 5'). But in the case of *Pu24T* this would mean that the binding occurs most plausibly with the opening of bottom tetrad that is otherwise hindered because of the snap-back base loop.
- 3) It is also possible that an inefficient quenching is due to different redox potentials ( $E_{red}$ ) for *N*-alkylquinolinium and benzyl pyridinium groups. The other tested compounds have two moieties of *N*-alkylquinolinium that are highly electron deficient. On the contrary, **3Ei** possess pyridinium substituents that may be less electron deficient, in particular, because they have benzyl derivatives as *N*-alkyl substituents. So it may occur that  $E_{red}$  of benzyl pyridinium substituents is not sufficient to quench the fluorescence of Cy5 through an electron-transfer process. To check this hypothesis we are going to repeat the fluorimetric titrations with at least one oligonucleotide, labelled with another fluorophore (6-FAM).

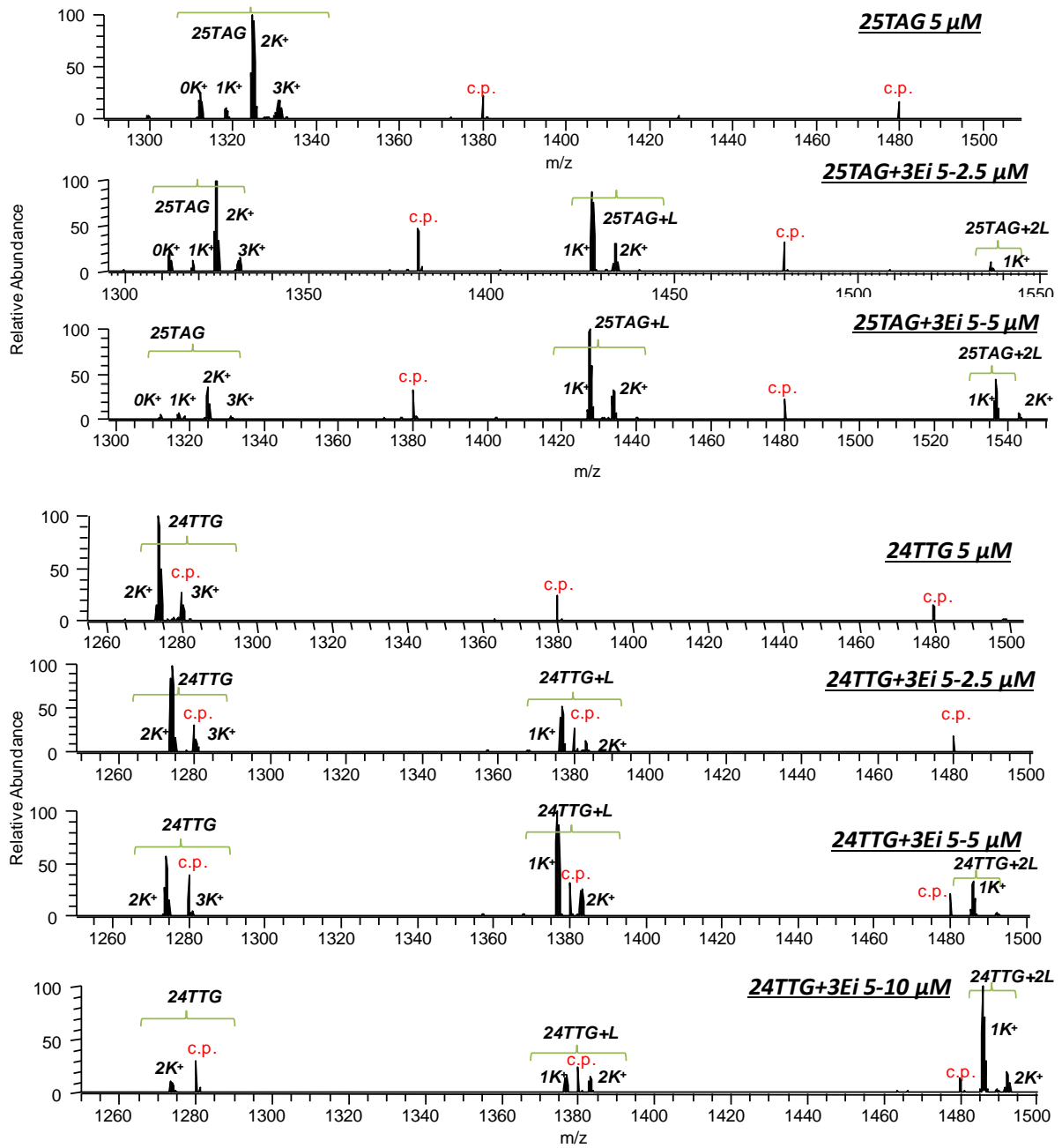
### 3.3.5 Mass spectrometry of G4-DNA-ligand complexes<sup>2</sup>

To further investigate the binding of ligands to selected G4s, we recorded ESI-MS spectra of five G4-DNA oligonucleotides (*24TTG*, *25TAG*, *22CTA*, *myc22* and *Pu24T*) in the presence of selected compounds (**3Ab**, **3Cb**, **3Ef** and **3Ei**). Due to large volume of data we present only spectra with **3Ei** (Figure 69). Interestingly, in the case of hybrid G4 substrates (*24TTG* and *25TAG*), the formation of complexes with **all ligands** was accompanied by ejection one of the two  $K^+$  cations initially present in the structure [DNA +  $2K^+$ ], as evidenced by the appearance of the peaks of complexes [DNA +  $1K^+$  + 1L] and [DNA +  $1K^+$  + 2L], in addition to the complexes that maintained two  $K^+$  ions ([DNA +  $2K^+$  + 1L] and [DNA +  $2K^+$  + 2L]). In contrast, no  $K^+$  ejection was observed with *22CTA* (that maintained one  $K^+$  ion in the complexes with ligand), *myc22* and *Pu24T* (that maintained two  $K^+$  ions). This is reminiscent of the situation observed with

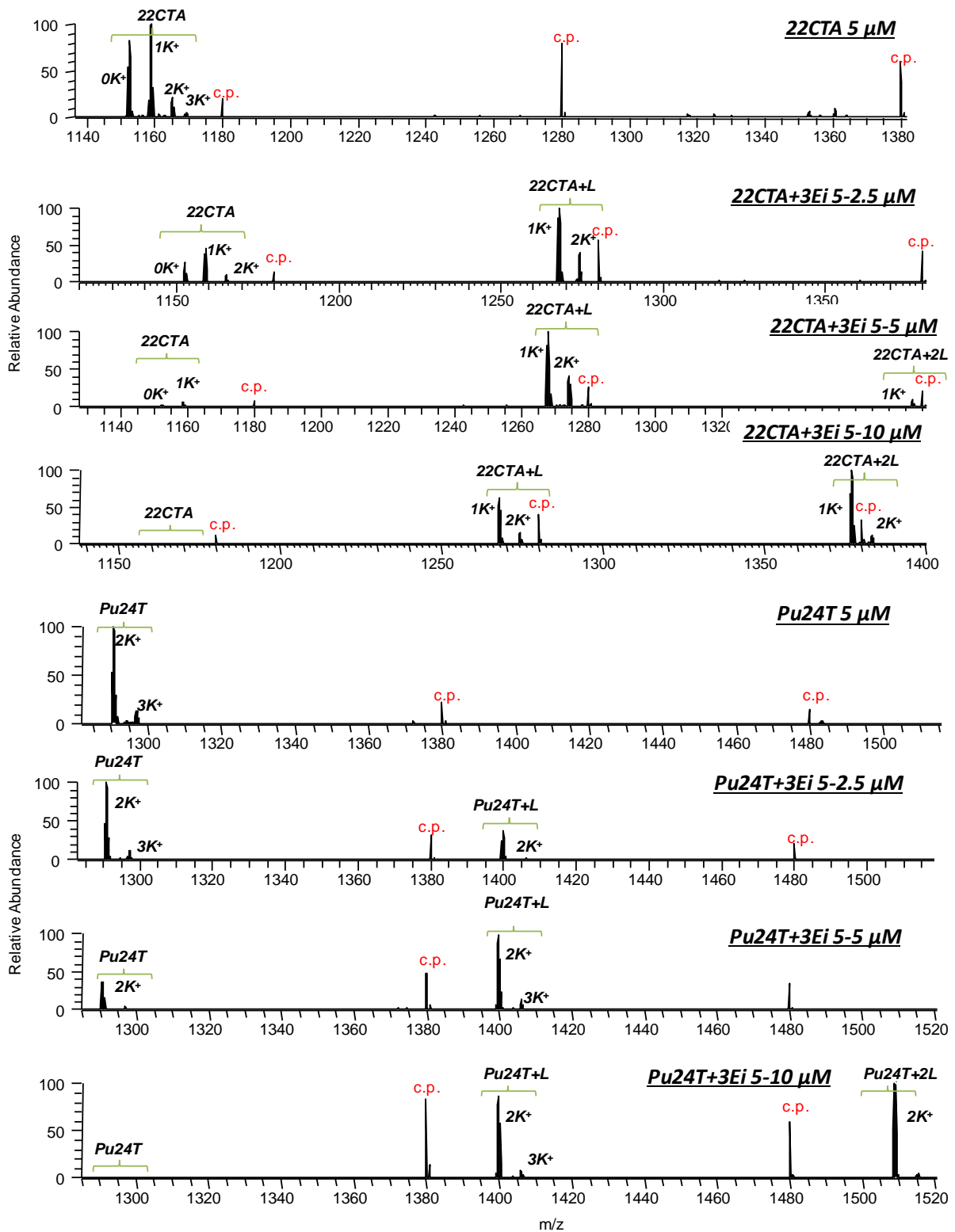
---

<sup>2</sup> Experiments performed in the lab of V. Gabelica, IECB, with assistance of Dr. E. Largy and Dr. F. Rosu.

**PhenDC3, PDC (360A), and pyridostatin (PDS).**<sup>142</sup> Although the details of the ligand-induced conformational change of G4 that accompany the K<sup>+</sup> ejection are not known, this seems to be a general feature of ligands containing a U-shaped bis(carboxamide) or bis(acylhydrazone) motifs. Interestingly, while in CD experiments this change was observed for *25TAG* upon binding of all ligand, *24TTG* was changing its conformation only upon addition of **3Ei**. This can be explained by different experimental condition of the assay: CD spectra were recorded immediately upon ligand addition, mass spectra after two hours of incubation. It is plausible, that sequence *25TAG* readily changes its topology, when for *24TTG* it takes more time for rearrangement. In addition, buffer used for CD contained 100 mM K<sup>+</sup> ions, while for mass spectrometry the buffer contained only 1 mM of K<sup>+</sup> that also could influence the stability of G4. In addition, CD is not a very precise method and gives only general information about the system. For example, the system contains 10% of G4 is in another topology it may not be visible by CD.







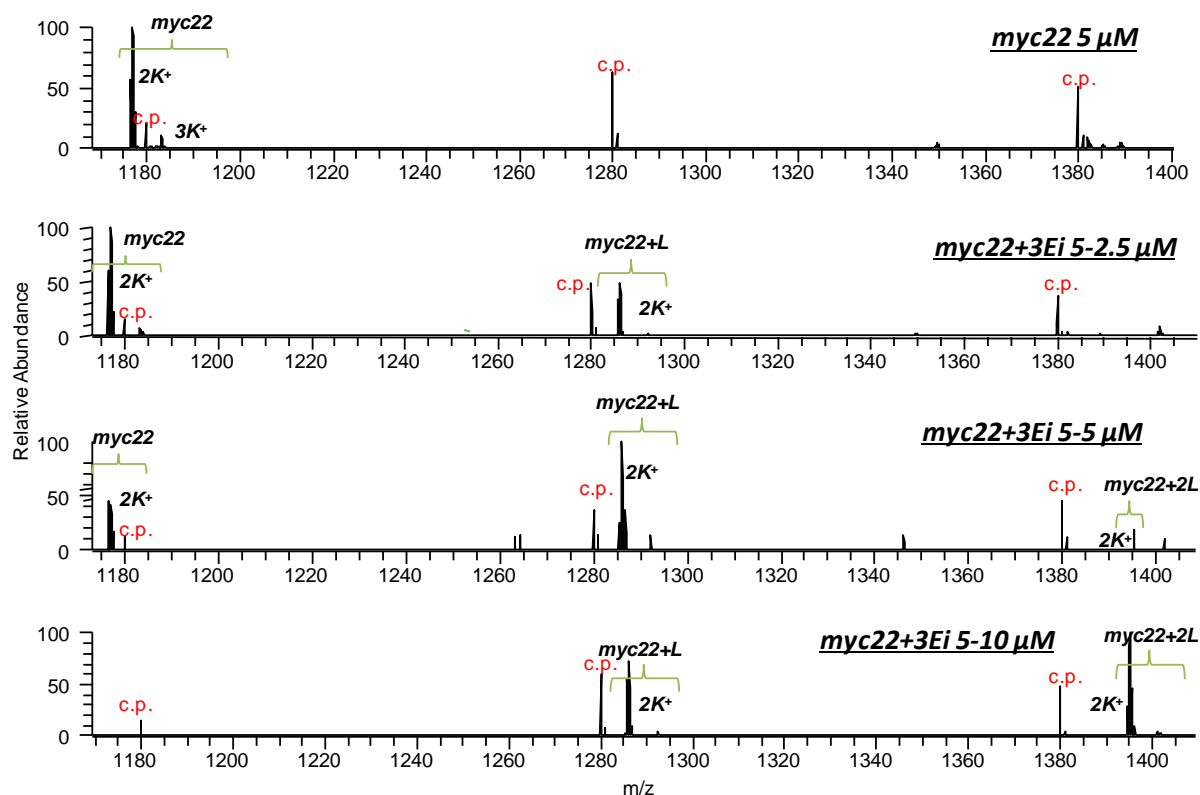


Figure 69. ESI-MS spectra of G-quadruplexes (25TAG, 24TTG, 22CTA, Pu24T and myc22, 5  $\mu\text{M}$ ) in the absence (top) and in the presence of 2.5, 5 and sometimes 10  $\mu\text{M}$  of ligand **3Ei**, in 100 mM TMAA + 1 mM KCl buffer after 2 hours of incubation. *c.p.* = calibration peak ( $dT_6$ ).

The data of mass-spectrometric measurements were exploited to obtain the values of apparent affinity constants that were calculated for two different ratios DNA: ligand: 1:0.5 and 1:1 (Table 16). In some cases,  $K_d$  values, determined from the ratio DNA/ligand 1:0.5, was significantly (10 fold) higher than those from the ratio 1:1. When repeating this experiment after prewashing of the capillary with 2.5  $\mu\text{M}$  solution of the ligand, we obtained the value very similar to one, obtained from 1:1 ratio (<sup>c</sup>) in Table 16). Apparently, there are processes of the saturation of the capillary with ligand that may cause the disruption of the complex and extraction of the ligand from the system. While for 1:1 ratio it does not drastically affect the result, its contribution becomes rather important when the DNA G4 is in excess in the system. In addition, the spectra of systems with 1:1 ratio were recorded after the corresponding 1:0.5 ratio spectra, so it possible that the first injection served as a prewashing for the capillary. For this reason, we tend to consider that the values, obtained from the 1:1 ratio are closer to real  $K_d$ .

Table 16. Apparent dissociation constants ( $K_d$  / nM) determined from mass-spectra of G4-DNA–ligand complexes. Conditions: 100 mM TMAA + 1 mM KCl buffer, incubation for 2 hours.

Ligand	25TAG		24TTG		22CTA		Pu24T		myc22	
	0.5 eq	1 eq.	0.5 eq	1 eq.	0.5 eq	1 eq.	0.5 eq	1 eq.	0.5 eq	1 eq.
<b>3Ab</b>	n.m. <sup>a</sup>	6.6	2 434	461	n.d.	< 1	107	80.8	n.d.	54
<b>3Cb</b>	n.d. <sup>b</sup>	4.6	152	17.6	n.d.	4.0	414	75	194	122
<b>3Ef</b>	56	204	1 059	628	n.d.	52	1 010	58	1 143	214
<b>3Ei</b>	n.d.	n.d.	1 127	978	n.d.	4.0	3 302 <sup>c</sup>	305	2 010	582
<b>PhenDC3</b>	n.m.	< 1	n.m.	75	n.m.	2.1	n.m.	572	n.m.	567

<sup>a)</sup> The spectrum was not measured; <sup>b)</sup> The  $K_d$  value could not be determined since the peak of unbound G4-DNA was undetectable; <sup>c)</sup>  $K_d = 311$  nM after prewashing of capillary with a 2.5  $\mu$ M solution of **3Ei**.

A  $K_d$  was not possible to determine for the 1:0.5 system (G4: ligand) for all ligands with 22CTA oligonucleotide, because the peak area of the complex was higher than the peak area of the oligonucleotide alone. This could be due to an error in determination of oligonucleotide concentration or due to saturation of the capillary with DNA. Another reason for this could be that oligonucleotide alone and in complex can have different responses in the mass spectrometer that can lead to false positive (or negative) results.

Comparing to fluorescence quenching experiments, the dissociation constants, determined by mass spectrometry, were more heterogeneous. All compounds, tightly bind to 25TAG and 22CTA with  $K_d$  less than 7 nM, except **3Ef**, that binds to these G4s with  $K_d$  of 204 and 52 nM, respectively. Considering another G-quadruplex of hybrid topology (24TTG), they bound less tightly with  $K_d$  values up to 1  $\mu$ M (for **3Ei**). Considering the G4s from promoter sequences, ligand bound overall less tightly than to telomeric sequences with the lowest  $K_d$  observed for **3Ab** to myc22 (54 nM). Importantly, high binding of **PhenDC3** to telomeric sequences was observed that is consistent with literature data.<sup>53</sup>

The tight binding of compound **3Ei** to antiparallel sequence 22CTA observed in the FID assay was confirmed by mass spectrometry. Indeed, its affinity to 22CTA was much higher than to parallel sequences from MYC promotor. However, in mass spectrometry the binding of **3Ei** to 24TTG were around two times less strong comparing to parallel G4s. In addition, extremely high affinity of the compound to 25TAG was observed in the mass spectrometry assay.

To measure the  $K_d$  of **3Ei** to **22CTA** and to **25TAG** more accurately, we also recorded spectra of these combination at lower concentration of counterparts (2  $\mu\text{M}$  and 1  $\mu\text{M}$ ). While for **22CTA**, the affinity constants could be measured, that was not possible for **25TAG** (Table 17). However, this data also should be treaded cautiously as the experimental errors are higher for such low concentrations of DNA.

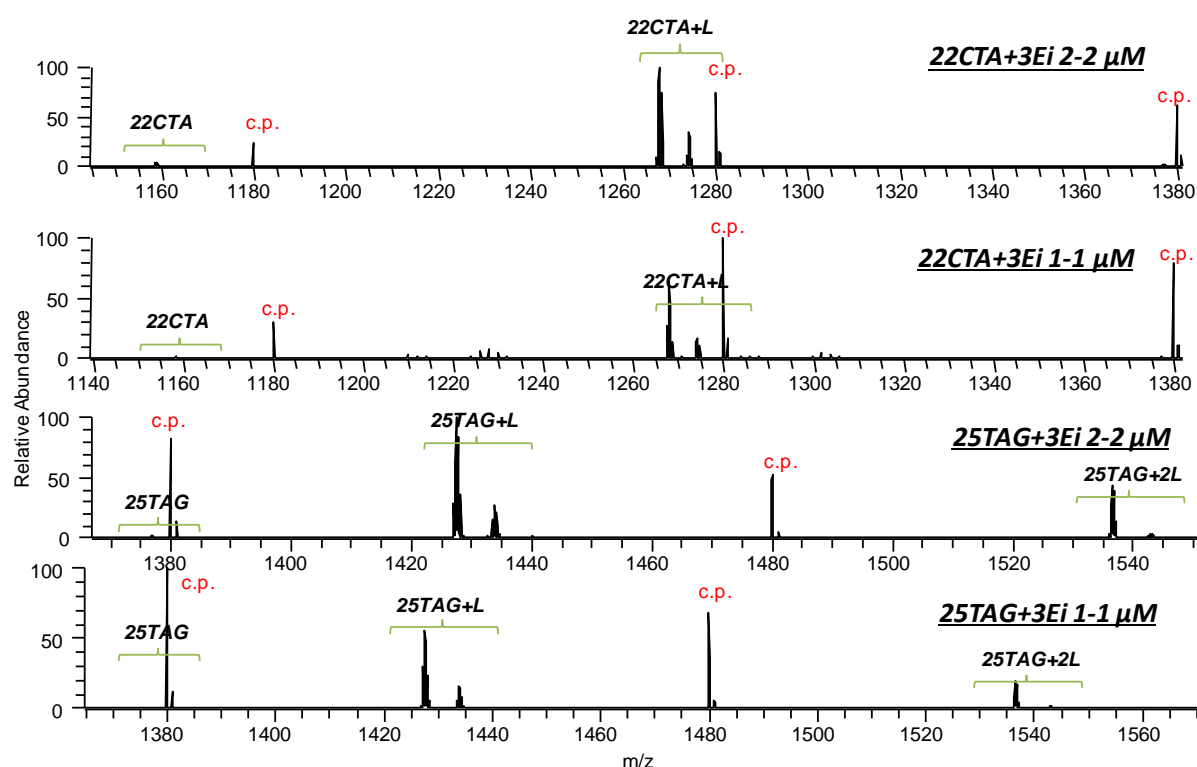


Figure 70. ESI-MS spectra of G-quadruplexes (**25TA** and **22CTA**, 2 or 1  $\mu\text{M}$ ) in the absence (top) and in the presence of equimolar concentration of ligand **3Ei**, in 100 mM TMAA + 1 mM KCl buffer after incubation for 2 hours. *c.p.* = calibration peak ( $dT_6$ ).

Table 17. Apparent dissociation constants ( $K_d$  / nM) determined from mass-spectra of G4-DNA–ligand complexes of **22CTA** with **3Ei**. Conditions: 100 mM TMAA + 1 mM KCl buffer, incubation for 2 hours.

Mixture	$K_d$ , nM	
	<b>22CTA-3Ei</b>	<b>25TAG-3Ei</b>
2 $\mu\text{M}$ G4-DNA + 2 $\mu\text{M}$ <b>3Ei</b>	8.3	n.d.
1 $\mu\text{M}$ G4-DNA + 1 $\mu\text{M}$ <b>3Ei</b>	4.5	n.d.

In order to determine if the experiments were reproducible, we recorded the spectrum of the system **3Ei**–*Pu24T* twice and determined apparent  $K_d$  (Figure 71). The values (48 and 58 nM, respectively) did not differ significantly.

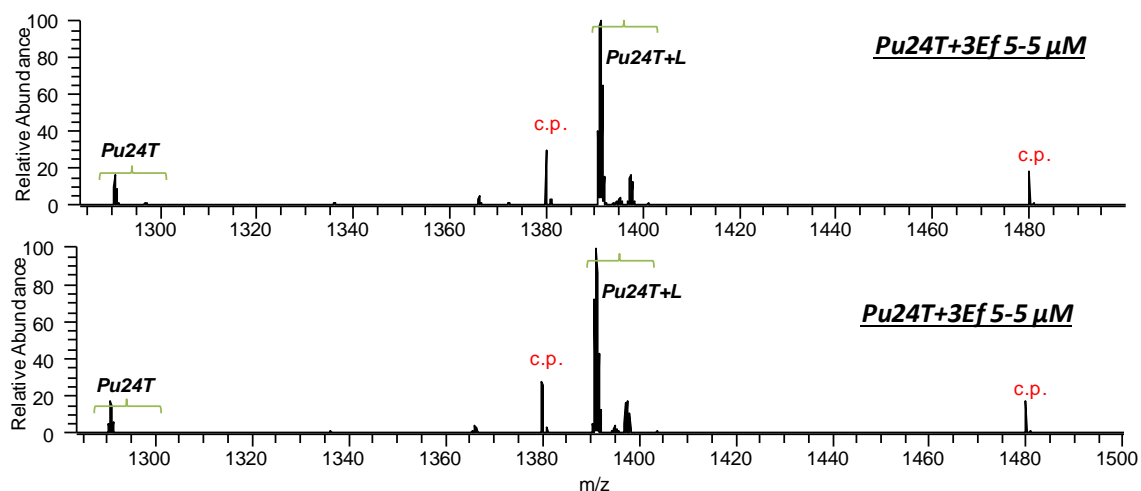
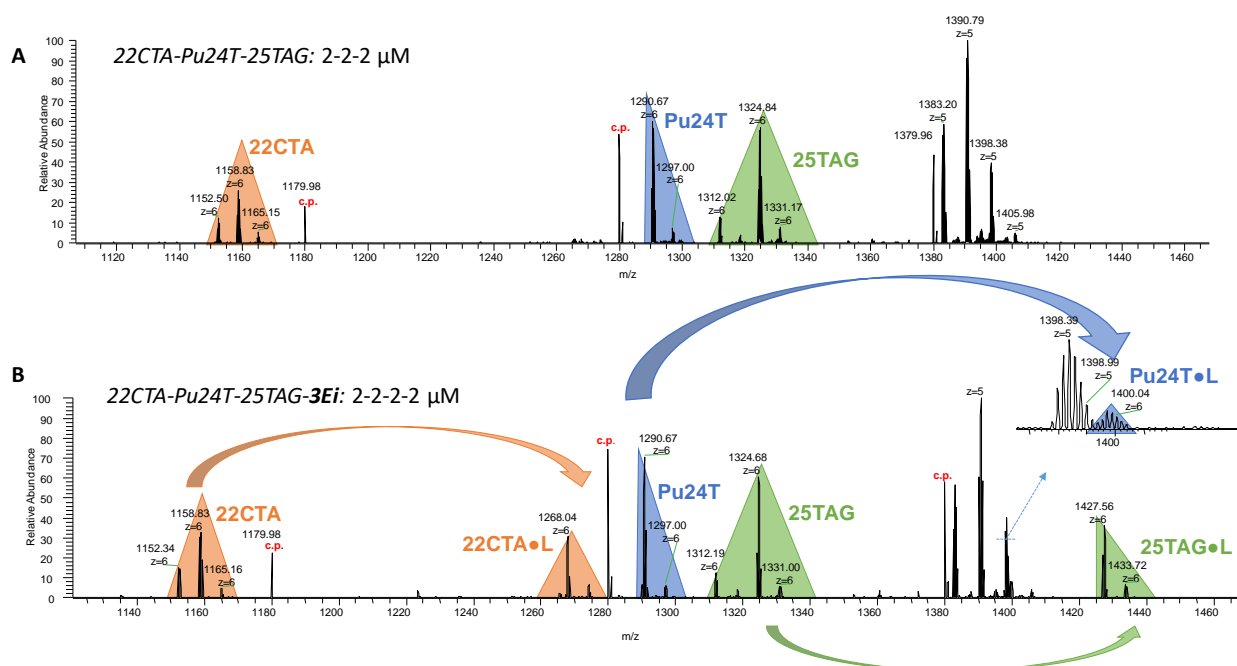


Figure 71. Replicate of the spectrum of *Pu24T* with **3Ef** 5–5  $\mu\text{M}$ . Buffer: 100 mM TMAA + 1 mM KCl; incubation for 2 h.

Finally, to investigate in more detail the selectivity of compound **3Ei** to G4 of antiparallel topology *22CTA*, observed in FID assay, we performed a mass-spectronic competition experiments in which equimolar amounts of DNA-G4 oligonucleotides of three different topologies (*25TAG*: hybrid 2, *22CTA*: antiparallel, *Pu24T*: snapback parallel, 2  $\mu\text{M}$  each) were titrated with compounds **3Ei** and **3Cb**. The oligonucleotide sequences for the experiment were chosen in a way to avoid all possible superimpositions of peaks of oligonucleotides (bound or unbound) at different charge states. We have not performed the competition experiments with compound **3Ef** due to significant mass superimpositions. The peak areas, corresponding to each oligonucleotide alone or its complex, was then calculated at 6<sup>-</sup> charge state and the fraction of bound oligonucleotide ( $\varphi$ ) was calculated using formula (8):

$$\varphi (\%) = \frac{A(DNA \bullet L)}{A(DNA) + A(DNA \bullet L)} \cdot 100, \quad (8)$$

where  $A(\text{DNA})$  and  $A(\text{DNA}+\text{L})$  are peak areas of free DNA-G4 and its complex with ligand, respectively.



**Figure 72.** Design of the competition experiments. A) MS spectrum of the equimolar mixture of three DNA G4 oligonucleotides. B) MS spectrum of the equimolar mixture of three DNA G4 oligonucleotides with ligand. The peaks corresponding to free DNA or its complex with ligands are highlighted in different colors. Conditions: (25TAG), (Pu24T) and (22CTA) = 2  $\mu\text{M}$ ,  $c(\text{ligand}) = 2 \mu\text{M}$ . Buffer: 100 mM TMAA + 1 mM KCl; incubation for 2 h.

The results of this competition experiment (Figure 73) indicate that the ligand systematically binds tighter to telomeric sequences (22CTA and 25TAG) comparing to G4 for MYC promotor Pu24T. At relatively high concentration (3 fold excess), the ligand binds to all G-quadruplexes in the system roughly homogeneously with, however, specifically higher preference for the 22CTA (93% of DNA is bound) comparing to Pu24T (69% DNA is bound). This difference is even better noticeable at concentration of the ligand 4  $\mu\text{M}$  when  $\phi$  (22CTA) is almost two fold bigger than for Pu24T (81.4% and 43.6%, respectively). Finally, this trend becomes even more prominent at equimolar amounts of ligand to G4 mixture when the fraction of 22CTA in the complex is five times bigger than the one of Pu24T (46.1% and 9%, respectively). 25TAG

showed a behavior similar to the one of 22CTA, having consistently high percentage of bound fraction DNA similar to, but consistently lower, than the one of 22CTA. All these data confirm the preferential binding of 3Ei to telomeric G4 of antiparallel topology over a parallel structure (Pu24T). Slightly lower but close tendency of 3Ei to bind to 25TAG can be explained by the same reason: the binding of 3Ei to 25TAG is followed by the ejection of the K<sup>+</sup> cation and change of its conformation to antiparallel. So the affinity of 3Ei to 25TAG is high, but less than for 22CTA as it also comprises the binding of 3Ei to the hybrid-2 G4 structure.

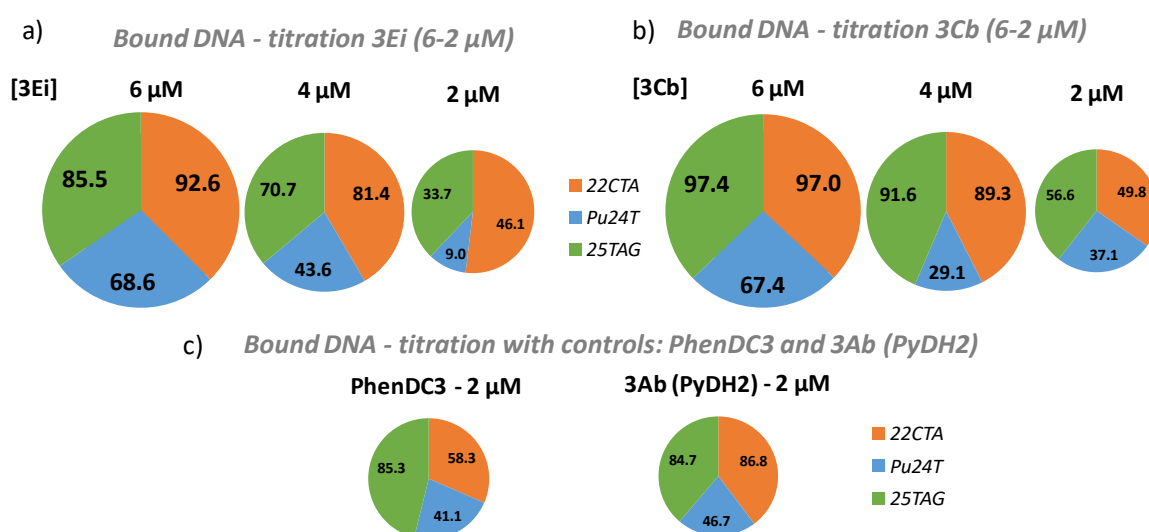


Figure 73. Results of MS-competition experiment, presented as distributions of fractions of ligand-bound G4-DNA ( $\phi$ ). Conditions:  $c(22CTA) = c(Pu24T) = c(25TAG) = 2 \mu M$ ;  $c(\text{ligand}) = 2-6 \mu M$ : a) with 3Ei; b) with 3Cb; c) with control ligands 3Ab (PyDH2) and PhenDC3:  $c(22CTA) = c(Pu24T) = c(25TAG) = c(\text{PhenDC3 or 3Ab}) = 2 \mu M$ . Buffer: 100 mM TMAA + 1 mM KCl; incubation for 2 h.

Compound 3Cb, as well as reference compounds 3Ab and PhenDC3, also preferably bind telomeric sequences rather than Pu24T sequence: even at the concentration of the ligand 2 μM it bound almost homogenously. However, if we recalculate the concentration of ligands 3Cb, 3Ab and PhenDC3 bound to DNA in systems with 2 μM of added ligand, it will be more than 2 μM. For example, in the case of 3Cb, the concentration of the bound DNA and, therefore, the concentration of ligand in the complexes in the bound state will be  $c(22CTA) \times \phi(22CTA) + c(Pu24T) \times \phi(Pu24T) + c(25TAG) \times \phi(25TAG) = 2 \mu M \times (0.498 + 0.371 + 0.566) = 2.87 \mu M$ . If we exclude the error in the concentration of 3Cb, 3Ab and PhenDC3 we need

explain this by the different MS responses of DNA in the bound and unbound states. It also should be mentioned that even a small error in the preparation of the stock solution of the ligand or in the measured concentration of oligonucleotides will affect drastically the results.

### 3.4 Conclusions

In this part, we developed a very simple method for the synthesis and high-throughput screening of bis(acylhydrazones) to estimate their capacity to stabilize G-quadruplexes. Heating of mixtures of building blocks with the catalyst (AcOH) gives the product solution directly in DMSO at approximately 2 mM concentration, which can be directly used in the biophysical experiments without any purification. We used this method to prepare 90 bis(acylhydrazones) with high yield and purity (87% on average). FRET-melting, performed with our combinatorial library and four G4 sequences of different topologies, allowed us to select three novel compounds for further investigation. These compounds were re-synthesized in the preparative manner and their interaction with targets was studied by numerous techniques. First of all, FRET-melting was performed with pure compounds, and we confirmed that  $\Delta T_m$  obtained using “as-synthesized” compounds did not differ from the values obtained with pure samples. We performed FRET-melting using different concentrations of ligands to look at the stabilizing capacities of these compounds more detailed. The FID assay, performed with five G-quadruplexes of different topologies pointed to the selectivity of the compound **3Ei** towards antiparallel telomeric sequence *22CTA*. Fluorimetric titrations with Cy5-labelled G-quadruplexes showed roughly a homogenous response of all compounds to all G4-sequences. The most promising compound **3Ei**, however, showed no significant fluorescence quenching of Cy5 in experiments with all G-quadruplexes, including *22CTA*. In mass spectrometry experiments, however, **3Ei** showed high affinity to telomeric sequences *25TAG* and *22CTA*. Its preferential binding to telomeric antiparallel sequences was also confirmed in competition experiments, where among the three oligonucleotides, (*25TAG*, *22CTA* and *Pu24T*), **3Ei** preferably bound to telomeric sequences, in particular, to *22CTA*.



These results, despite numerous discrepancies in data from different assays, represent an interesting material regarding the advantages and pitfalls of different techniques for selection and characterization of G4-binders. In addition, we consider that compound **3Ei** represents a very promising structure for telomeric targeting due its preferential binding to telomeric antiparallel G-quadruplexes.

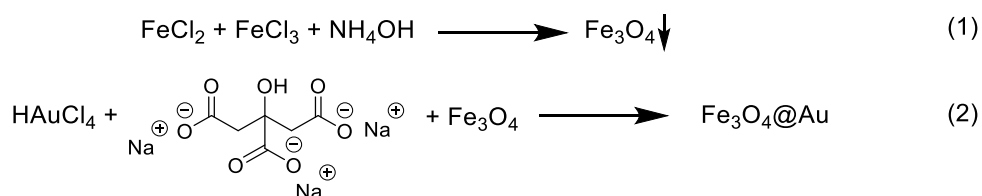
## 4. Ligand pull-down beyond streptavidin-coated magnetic beads

The protocol of ligand pull-down involving streptavidine coated magnetic beads, as described in *RESULTS AND DISCUSSION, Part 2*, has two major disadvantages: the high level of non-specific binding to the beads and their high price. The generally advised beads' pre-washing with poly-L-lysine or biotin did not reduce the amount of non-specific interactions in our hands so for further experiments we decided to use another types of pull-down such as gold-coated magnetic nanoparticles that have been reported as a good support for dynamic combinatorial chemistry<sup>120,128</sup> or solid phase extraction, that has not been used the DNA field before.

### 4.1 Experiments with gold-coated magnetic nanoparticles

In our first experiments we attempted to reproduce the two-step synthesis of Au@Fe<sub>3</sub>O<sub>4</sub> nanoparticles reported by Dash et al. (*Scheme 16*).<sup>1</sup> In the first step, the iron(II)–iron(III) oxide (or ferrous ferric oxide) Fe<sub>3</sub>O<sub>4</sub> nanoparticles are obtained from ferrous and ferric chlorides (FeCl<sub>2</sub> and FeCl<sub>3</sub>, respectively) in ammonium hydroxide solution. They are easily separated from the reaction mixture with the help of the magnet. The second step includes the reduction of chloroauric acid by sodium citrate in the presence of Fe<sub>3</sub>O<sub>4</sub> nanoparticles. As sodium citrate is mixed with Fe<sub>3</sub>O<sub>4</sub> nanoparticles before the addition of chloroauric acid, the reaction of reduction of chloroauric acid should take place at the surface of ferrous ferric oxide and thereby coat magnetic nanoparticles with gold. The desired Au@Fe<sub>3</sub>O<sub>4</sub> nanoparticles are again separated using the magnet and re-suspended in water. However, while the first step was successful and precipitate of Fe<sub>3</sub>O<sub>4</sub> could be easily separated, the magnetic precipitation in the second step did not work well and no visual precipitate was obtained on the magnet. Instead, the resulting solution was rich with floating particles of pink color. We, however, were

able to separate the supernatant after 30 minutes of precipitation of on the magnet and continued using this fraction.



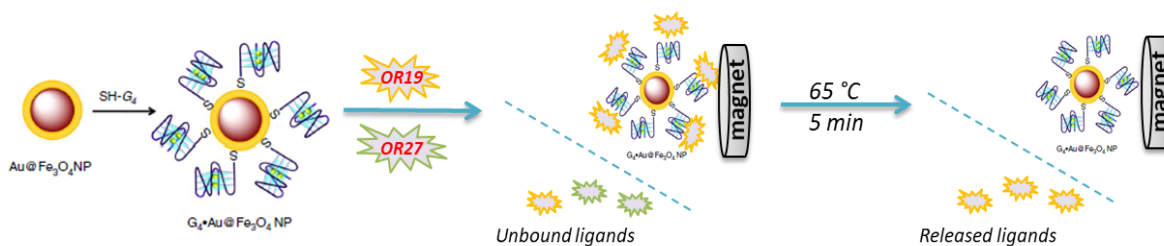
*Scheme 16. Synthesis of gold-coated magnetic nanoparticles*

The obtained nanoparticles were further functionalized with 5'-thiol-capped oligonucleotide 5'-SH-*Pu24T* and 5'-SH-*hp2*. By the measurement of absorbance at 260 nanometers we evaluated the concentrations of bound oligonucleotides in nanoparticles' suspension (*Table 18*).

*Table 18. Absorption measurements of functionalised nanoarticles.*

<b>Oligonucleotide</b>	<b>Maximum of absorbance</b>	<b><math>\epsilon</math></b>	<b>Measured concentration</b>
<i>Pu24T</i>	0.147	248200	0.59 $\mu\text{M}$
<i>hp2</i>	0.047	181700	0.26 $\mu\text{M}$

For the preliminary experiments, we incubated the functionalized nanoparticles with the equimolar mixture of two ligands, **A2-L1-A2 [PyDH2, OR19]** and **A2-L2-A2 [PymDH2, OR27]**, “good” and “poor” G4-ligands, respectively (according to our previous FRET-melting experiments, cf. *RESULTS AND DISCUSSION, Part 1*) for 2 hours (*Scheme 17*). The nanoparticles were then retained with magnet and the supernatant was analyzed by HPLC (fraction “unbound ligands”). The nanoparticles were then re-suspended in the buffer, heated for 5 minutes at 65 °C and separated supernatant was also analyzed by HPLC (fraction “released ligands”).



Scheme 17. Protocol of the experiment with nanoparticles.

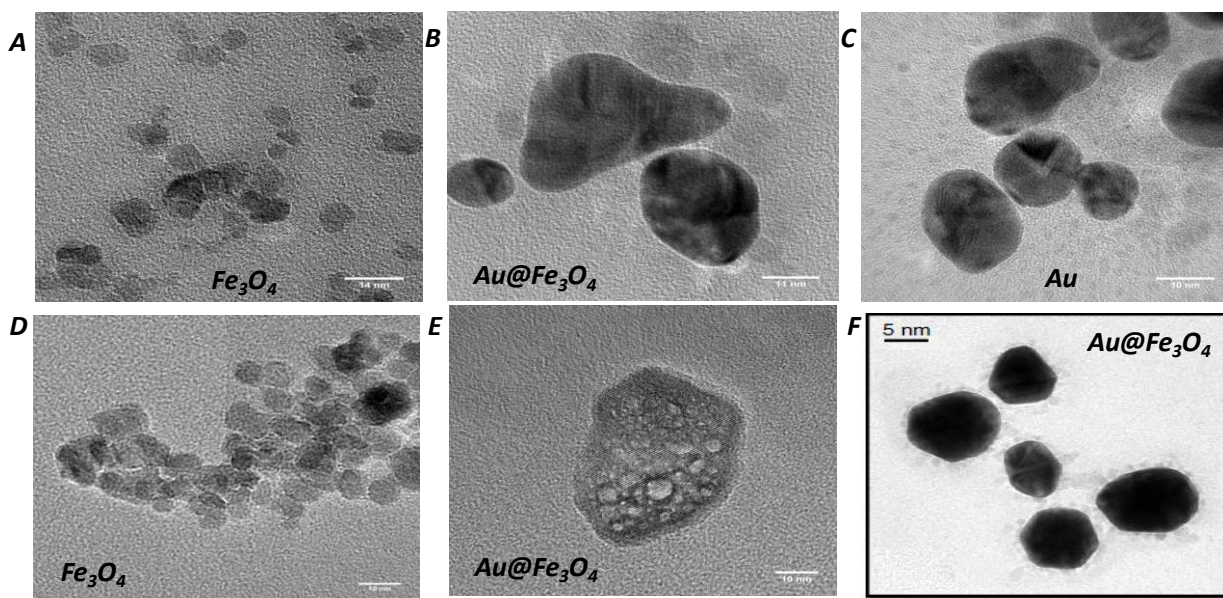
The HPLC analysis of the fractions in experiments with *Pu24T* G-quadruplex indicated the depletion of **A2-L1-A2** in solution of unbound ligands and the presence of this ligand in the released fraction (Figure 74A). In the experiment with the hairpin *hp2*, the solution of unbound ligands contained almost the same amount of both ligands and the released fraction contained only a small amount of **A2-L1-A2** (Figure 74B). The fact, that the “good” ligand **A2-L1-A2** was present in released fraction of both G4 and hairpin functionalized magnetic nanoparticles indicates on some non-specific interactions of **A2-L1-A2** with *hp2*-Au@Fe<sub>3</sub>O<sub>4</sub> that may be due to electrostatic interactions (however, in this case we would expect the mixture of two ligands to be released, as they are equally charged).



Figure 74. HPLC analysis of mixtures of bound and released ligands in the experiments with nanoparticles with attached A) G-quadruplex Pu24T; B) hairpin hp2.

In parallel, the synthesized nanoparticles were analyzed by transmission electron microscopy (TEM). As expected, Fe<sub>3</sub>O<sub>4</sub> nanoparticles were spherical with average radius 7 nm (Figure 75A). However, the nanoparticles obtained in the second step (after coating with gold, Figure 75B) appeared not to have an expected structure. Instead of magnetic nanoparticles coated with gold it is rather likely that we obtained nanoparticles of gold with attached Fe<sub>3</sub>O<sub>4</sub> nanoparticles on their surface. In some gold nanoparticles we could see the grid, characteristic for Fe<sub>3</sub>O<sub>4</sub> nanoparticles that could be an indication that we obtained Fe<sub>3</sub>O<sub>4</sub> nanoparticles coated with gold and then surrounded by Fe<sub>3</sub>O<sub>4</sub> nanoparticles again. We also analyzed pink nanoparticles that were not precipitating on the magnet (Figure 75C). In some cases we also could distinguish the lattice, characteristic for ferromagnetic nanoparticles but in majority the mixture contained gold nanoparticles. For comparison, we also show the micrograph reported by Dash et al. (Figure 75F).<sup>120</sup> We tried to repeat the synthesis, but, unfortunately, obtained

aggregated nanoparticles in the first step (*Figure 75D*) and some porous particles in the second (*Figure 75E*).

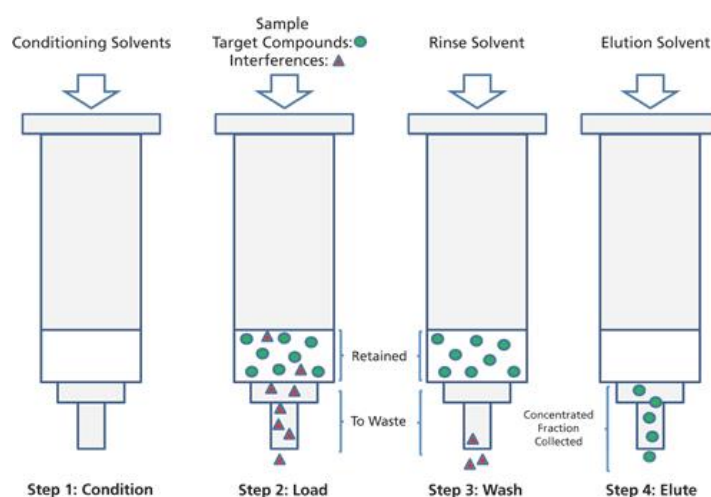


*Figure 75. TEM images of our nanoparticles. First experiment: A)  $Fe_3O_4$ ; B)  $Fe_3O_4@Au$ ; C)  $Fe_3O_4@Au / Au$ , that did not precipitate on the magnet; second batch: D)  $Fe_3O_4$ ; E)  $Fe_3O_4@Au$ , that precipitated on the magnet; F) G4\_ Au@Fe3O4 reported by Dash et al.<sup>120</sup> I thank Sylvain Trepout (Institut Curie) for TEM images of nanoparticles.*

Even though we obtained nanoparticles that are similar to those, reported by Dash et al. (*Figure 75B and F*),<sup>120</sup> it appeared that they do not have a structure needed to perform pull-down experiments. We therefore considered the methods of indirect coating, which imply having the glue material between magnetic core and gold surface.<sup>176</sup> For example, hydrophobic  $Fe_3O_4@Au$  nanoparticles can be obtained by the reduction of  $HAuCl_4$  on the surface of oleylamine and oleic acid capped  $Fe_3O_4$  nanoparticles in chloroform.<sup>177</sup> Hydrophilic core/shell  $Fe_3O_4/Au$  nanoparticles can be obtained by further reduction of  $HAuCl_4$  by sodium citrate in the presence of surfactant in water phase. Another option could be to use commercially available gold coated magnetic nanoparticles.<sup>178</sup> However, after thorough consideration we decided to abandon the experiments with nanoparticles and to focus our attention on the method of solid phase extraction.

## 4.2 Solid phase extraction

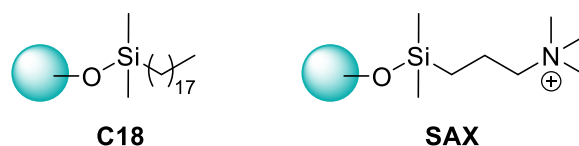
The principle of solid phase extraction (SPE) reside in to selective extraction of analytes from a mixture by absorption on a solid phase. The general steps of SPE include conditioning of the cartridges, load, washing and elution (*Figure 76*). The cartridges are conditioned with organic solvents (acetonitrile or methanol) to wet the pores and then with water or the working buffer, to prepare the stationary phase to interact with analytes. When the sample is loaded onto the cartridges (with a manifold or with a syringe), some molecules will be retained by the stationary phase (specifically or non-specifically) and some other will pass through the cartridge with the percolate. This step has to be performed very slowly in order to make sure that the analytes interact with the stationary phase. In the following step, called washing, the analytes of interest are separated from the non-specifically bound compounds. The choice of washing solvent depends on the mixture and the stationary phase. It is usually performed with big volumes and several times. The last step corresponds to the elution of the analytes. The solvent used to elute the analytes has to disrupt the interaction between the stationary phase and the analyte. The volume used for elution can be small, especially if the sample is at trace levels. Finally, the eluted fraction is analyzed by HPLC or another suitable technique.



*Figure 76. The principle of solid phase extraction method (adapted from <sup>179</sup>).*

### 4.2.1 Preliminary experiments

For our preliminary experiment we selected two Agilent cartridges: non-polar octadecyl (C18) and strong anion-exchange (SAX), with trimethylaminopropyl surface (SAX, *Figure 77*). Both phases are known to retain DNA due to hydrophobic (C18) or electrostatic interactions. We hypothesized that a complex of G4–DNA with a strong ligand could be absorbed on the solid phase without dissociation. The following elution steps, performing in the conditions denaturing the DNA–ligand complex, would release the bound ligand allowing its subsequent analysis by HPLC and / or by LC/MS.



*Figure 77. Structures of a sorbent of two Agilent cartridges, selected for preliminary experiments: C18 and SAX.*

We performed SPE with a mixture of two ligands – “good” G4 ligand **A2-L1-A2 [OR19]** and “poor” **A1-L2-A1 [OR31]** – and caffeine (an additional control). SPE was performed in the presence of *Pu24T* G-quadruplex or *ss6* single strand, as well as without DNA. Both ligands were retained on the C18 cartridge due to hydrophobic interactions, whereas DNA most likely passed through the cartridge with the percolate or during the washing step. Both ligands were found in the elution fractions (data not shown). However, with the SAX cartridge, **A1-L2-A1 [OR31]** was found in percolated and in washing solutions, whereas **A2-L1-A2 [OR19]** was found mostly in the elution step when the experiments were performed in the presence of G4 (*Figure 78*). In the control SPE experiment without DNA both ligands were found in the percolated and washing fractions, however, in eluted fraction only **A2-L1-A2 [OR19]** was present in a rather small amount. The same is true for the single–strand control. If we compare the three eluted fractions (no DNA, G4 and ss), **A2-L1-A2 [OR19]** was found in much higher amount in the eluted fraction from G4 compared to single–strand or control (no DNA) experiment. These results indicate that ligand–G4 complex is indeed retained on the cartridge



and is eluted as the complex is denatured by the eluent, i.e. the mixture of formamide and acetonitrile. Therefore, SAX cartridges were selected for further experiments.

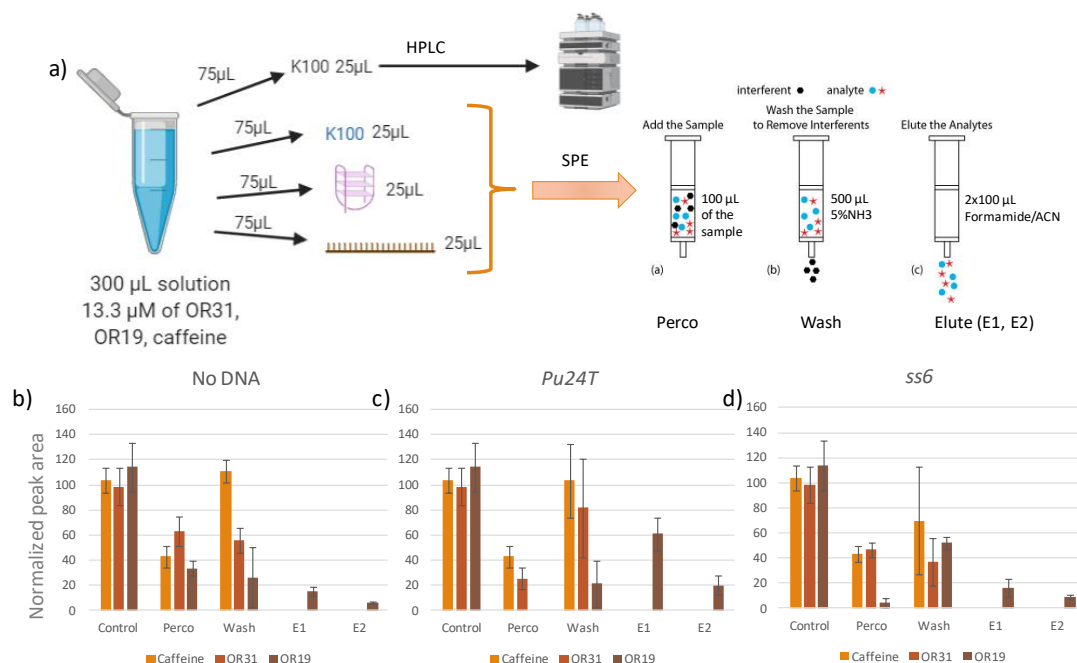


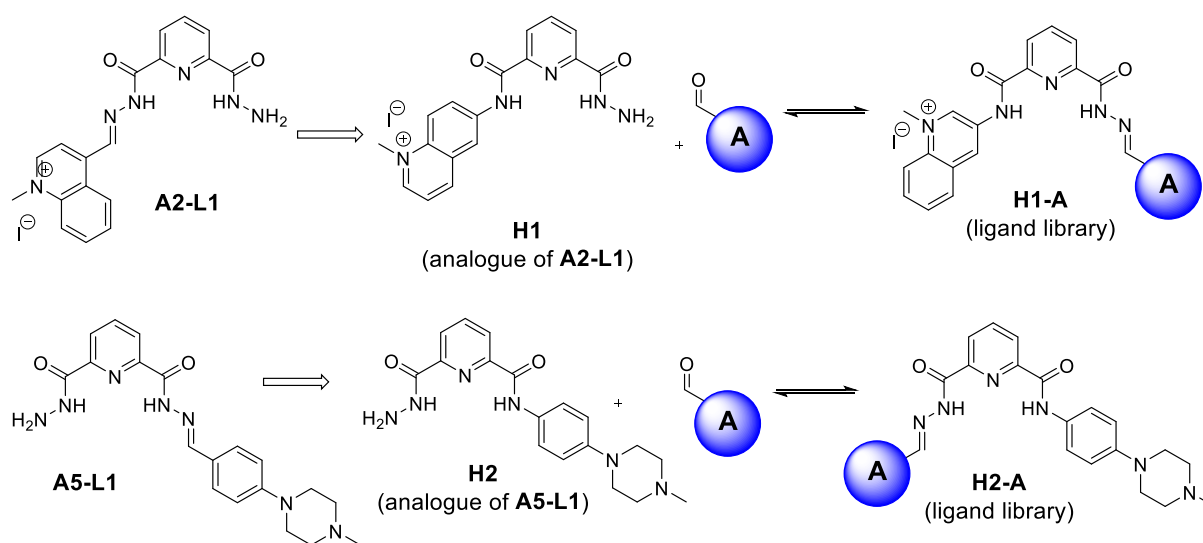
Figure 78. a) Scheme and b–d) results of preliminary experiments presented as relative peak area of corresponding compounds obtained in SPE experiment b) without DNA; in the presence of c) Pu24T G4 or d) single-stranded control ss6. SPE was performed with the mixture of three compounds: **OR19**, **OR31** and caffeine.  $C(\text{caffeine}) = C(\text{OR19}) = C(\text{OR31}) = 10 \mu\text{M}$ ;  $C(\text{DNA}) = 25 \mu\text{M}$ . Experiment is performed in triplicate.

One of advantages of SPE pull-down technique is that no modifications of oligonucleotides are required for these experiments. In addition, the non-specific interaction of cationic ligands with the polycationic cartridge are expected to be minimal.

#### 4.2.2 Design of the experiments

In the previous DCC experiments (*RESULTS AND DISCUSSION, Part 2*) we noticed, that acylhydrazones that contain **A2-L1-Ax** fragment seem to have high affinity to G-quadruplexes. Considering this fact, we sought to develop the novel approach for the screening of ligands of DNA G-quadruplexes. We propose to replace the scaffold **A2-L1** by an amide derivative which, being rather similar to acylhydrazone analogue, guarantees the affinity of ligands that contains

his moiety to G-quadruplexes. In this case, on the one hand, our model ligand should be affine to G-quadruplexes due to **A2-L1** moiety, and, on the other, we can try to achieve the selectivity towards definite G-quadruplex structures by varying the second (aldehyde) counterpart (*Scheme 18*). As we already had in our hands two hydrazides, analogues of **A2-L1** (**H1**) and of **A5-L1** (**H2**), we decided to use both of them for the design of ligand libraries.



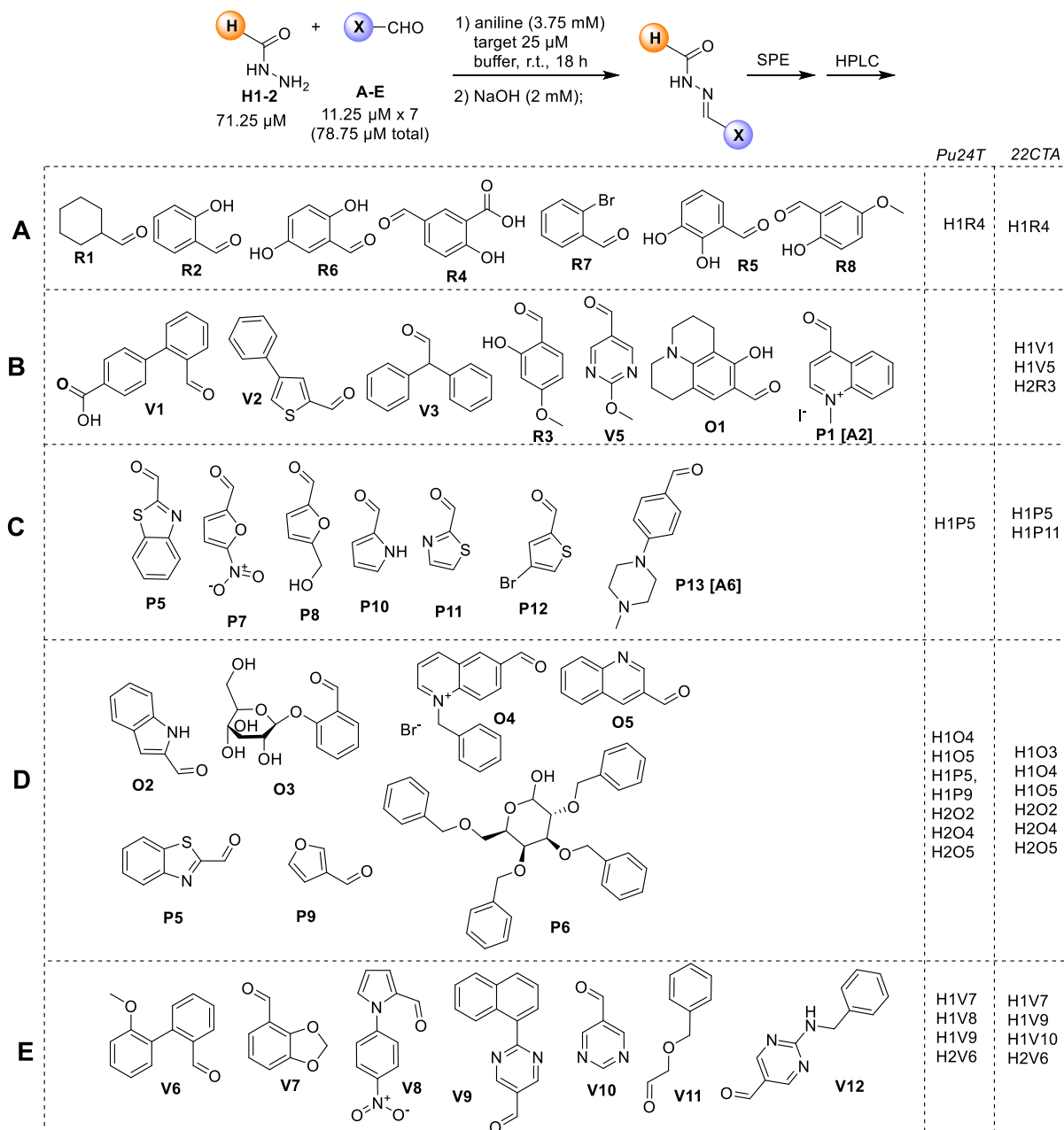
*Scheme 18. The design of ligand libraries for SPE experiments.*

### 4.2.3 Selection of hits

#### First round of selection

The synthesis of acylhydrazides **H1** and **H2** is described previously (cf. *RESULTS AND DISCUSSION, Part 2*). The selection of 34 aldehydes for SPE experiments is shown in *Scheme 19*; most compounds were ordered from commercial suppliers. Some of aldehydes (**A2**, **A5**, **A8**) were already synthesized or used previously. Ten dynamic combinatorial libraries were generated by combining subsets of aldehydes (*A-E*) with two acylhydrazides, **H1** and **H2**, separately (*Scheme 19*). Each of ten combinatorial libraries was generated in the presence of four DNA oligonucleotides: G-quadruplexes *Pu24T* and *22CTA* as well as double-strand *ds26*

and hairpin *hp2*. The control libraries (prepared in the absence of DNA) were analyzed by LC/MS and the peaks of all products were attributed. The libraries, generated in the presence of the targets, were loaded on SAX cartridges, washed, and the eluted fractions were analyzed by HPLC. The use of the same column (Luna Phenomenex 50\*3 mm, dp = 3µm) and chromatographic conditions permitted to attribute the peaks by LC/MS and then to use HPLC for further experiments. Elutes of combinatorial libraries composed from the same building blocks, but templated with different targets, were compared and compounds enriched in the presence of G-quadruplexes compared with double-stranded and hairpin DNA were identified by comparing the peak area of compounds in the corresponding libraries (*Figure 79*).



Scheme 19. Library design, synthesis and selection of acylhydrazones (the labeling of aldehydes is arbitrary).

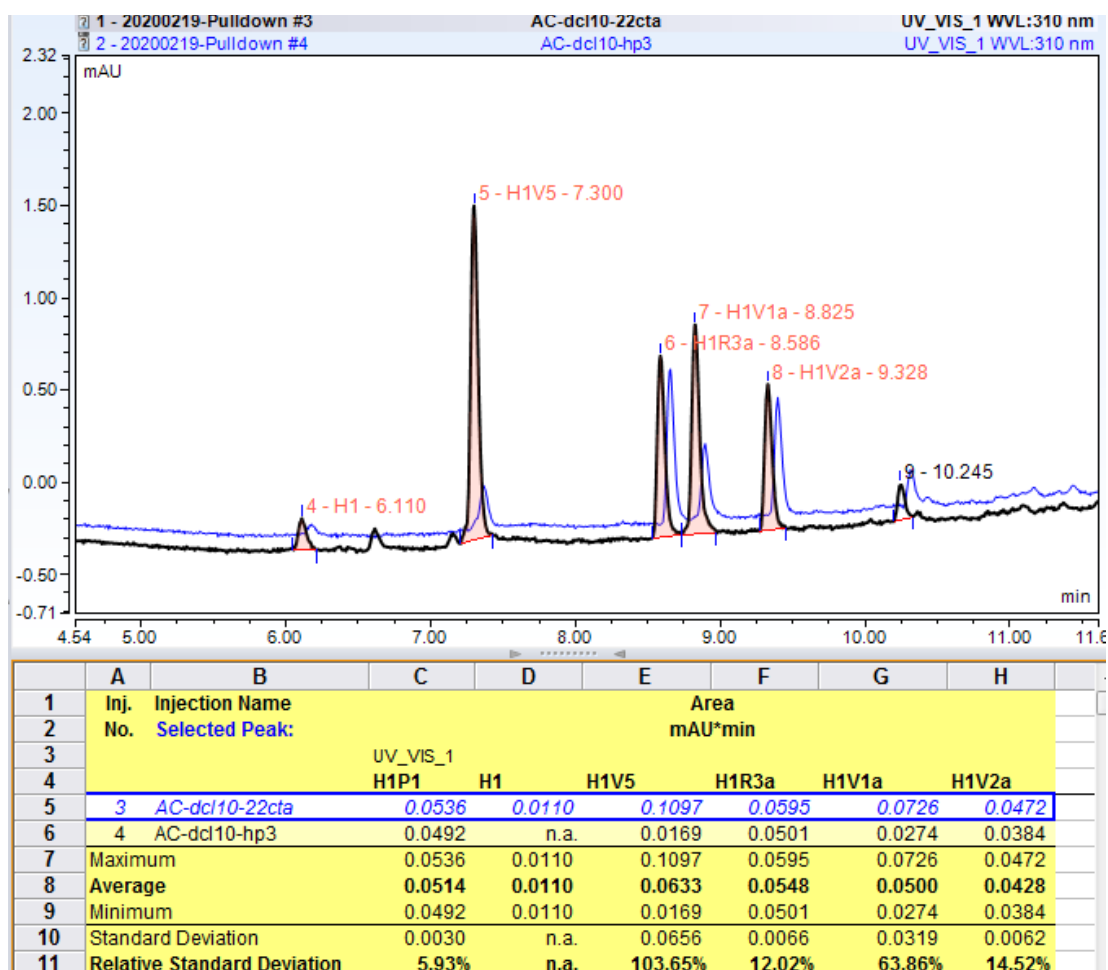


Figure 79. Example of HPLC chromatogram of libraries templated in the presence of 22CTA G4 and hairpin.

Twenty hits were selected in these experiments. These include the adducts of hydrazide **H1** and aldehydes **V1**, **V5**, **V7–V10**, **P5**, **P9**, **P11**, **O3–O5**, as well as the adducts of **H2** with aldehydes **R3**, **R7**, **V7–V9**, **O2**, **O4**, **O5** (Scheme 19). The combinations of these mentioned building blocks were chosen for the second round of selection in which the combinatorial libraries were generated from acylhydrazide **H1** with the aldehyde sets **F–G** and **H2** with sets **H–I** (Chart 10).

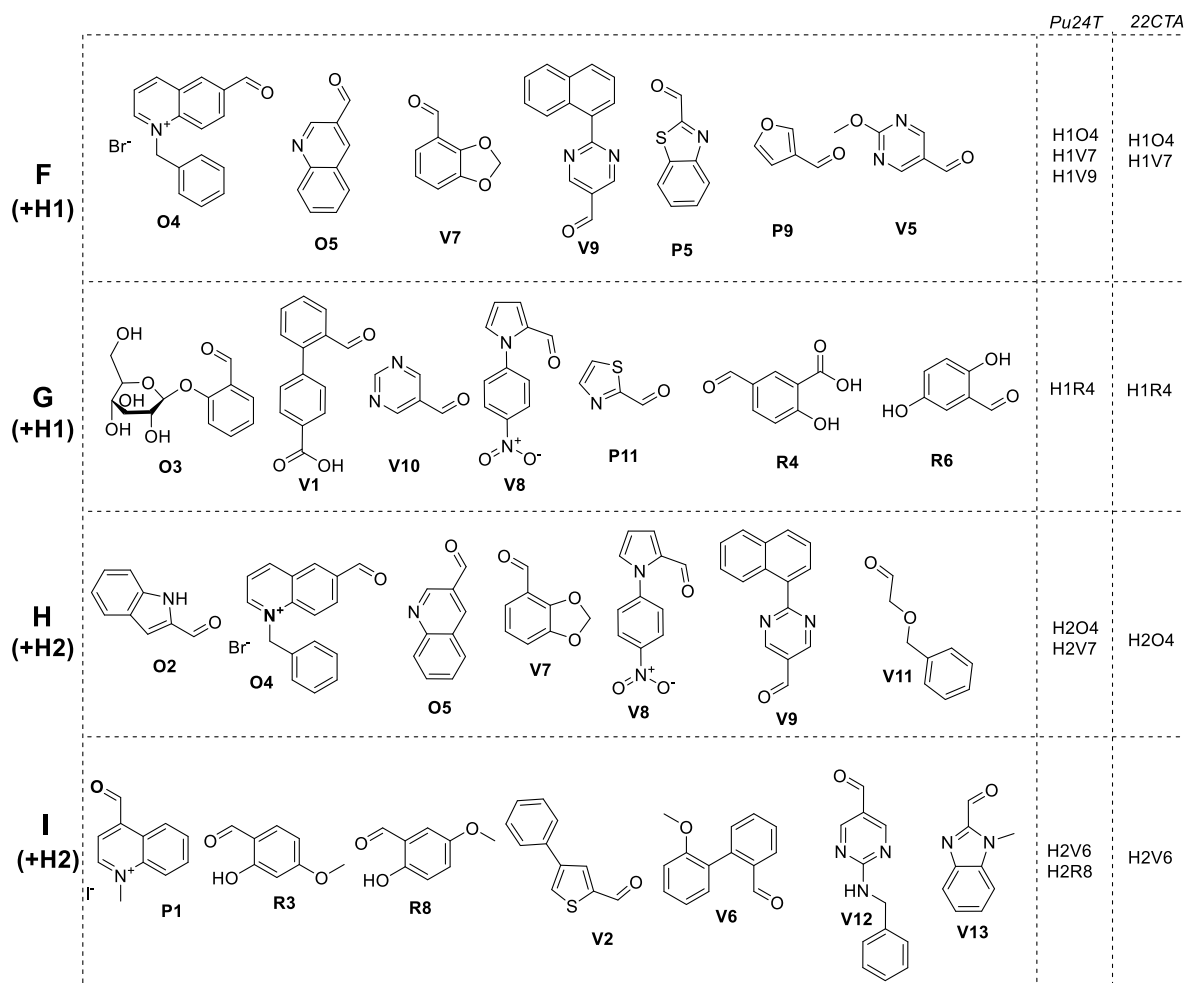


Chart 10. Sets of aldehydes for the second round of selection acylhydrazones from dynamic combinatorial libraries. Sets F and G were mixed with hydrazide **H1** and sets H and I – with **H2**. The labeling of aldehydes is arbitrary.

### Second round of selection

In the second round of selection, 28 compounds were generated in four combinatorial libraries (Chart 10). As in the first round of the selection, each combinatorial library was incubated in the presence of four different DNA oligonucleotides separately (*Pu24T*, *22CTA*, *ds26* and *hp2*) and then extracted on the Agilent SAX cartridge. The elutes were analyzed by HPLC and eight hits were selected: compounds **H1O4**, **H1V7**, **H1R4**, **H2O4**, **H2V6** were retained both by *Pu24T* and *22CTA*, while compounds **H1V9**, **H2V7** and **H2R8** were retained only by *Pu24T* (Figure 80). Surprisingly, in the set of tested compounds there were no products selective exclusively to *22CTA* G-quadruplex.

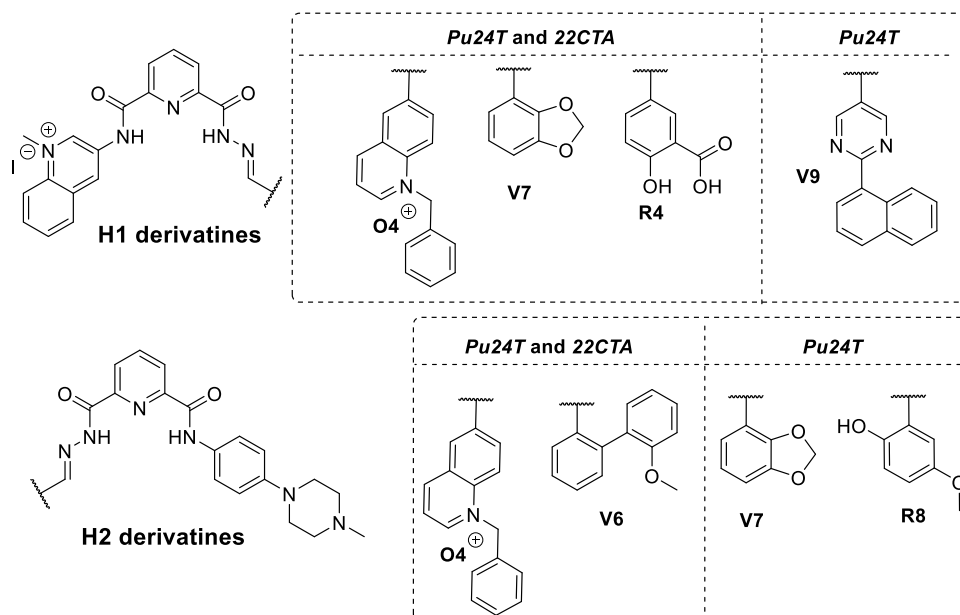


Figure 80. Structures of compounds, captured by both 22CTA and Pu24T and only by Pu24T.

#### 4.2.4 Confirmation of the hits

To evaluate compounds, selected through SPE from combinatorial libraries, we decided first to use the method of parallel synthesis and high-throughput screening of acylhydrazones described in *RESULTS AND DISCUSSION, Part 3* as it is easy to perform. Importantly, in FRET-melting experiments with “as-synthesized” ligands,  $\Delta T_m$  values of these compounds were almost the same as obtained with pure compounds, synthesized in the preparative way.

The building blocks corresponding to selected eight compounds as well as for some negative controls (**H1R8**, **H2P7**, **H2O3**, **H2V9** and **H2R4**) were mixed with the catalyst (acetic acid) in DMSO and heated for two days at 60 °C. The compounds were then analyzed by LC/MS. Except for **H1R4**, almost all products were obtained in excellent ( $\geq 89\%$ ) yields (*Table 19*), as estimated by LC peak areas.

Table 19. Peak areas (%) of the compounds synthesized as “ready-to-screen” solutions. Hits selected from SPE experiments are highlighted in bold.

<b>H1</b> derivatives	Yield [%]	<b>H2</b> derivatives	Yield [%]
<b>H1O4</b>	90	<b>H2O4</b>	100
<b>H1V7</b>	92	<b>H2V7</b>	98
<b>H1V9</b>	90	<b>H2R8</b>	90
<b>H1R4</b>	47	<b>H2V6</b>	97
H1R8	89	H2P7	96
		H2O3	96
		H2V9	97
		H2R4	98

All as-synthesized compounds were then tested in FRET-melting experiments employing double labelled G-quadruplexes *22CTA*, *25TAG*, *KRAS*, *myc22* and *hp2* as a control (Figure 81). FRET-melting experiments were also performed with *Pu24T* G-quadruplex, but due to high standard deviations with  $T_m$  values observed with sequence, this data were discarded. A reproduction of the experiment with *Pu24T* is planned.



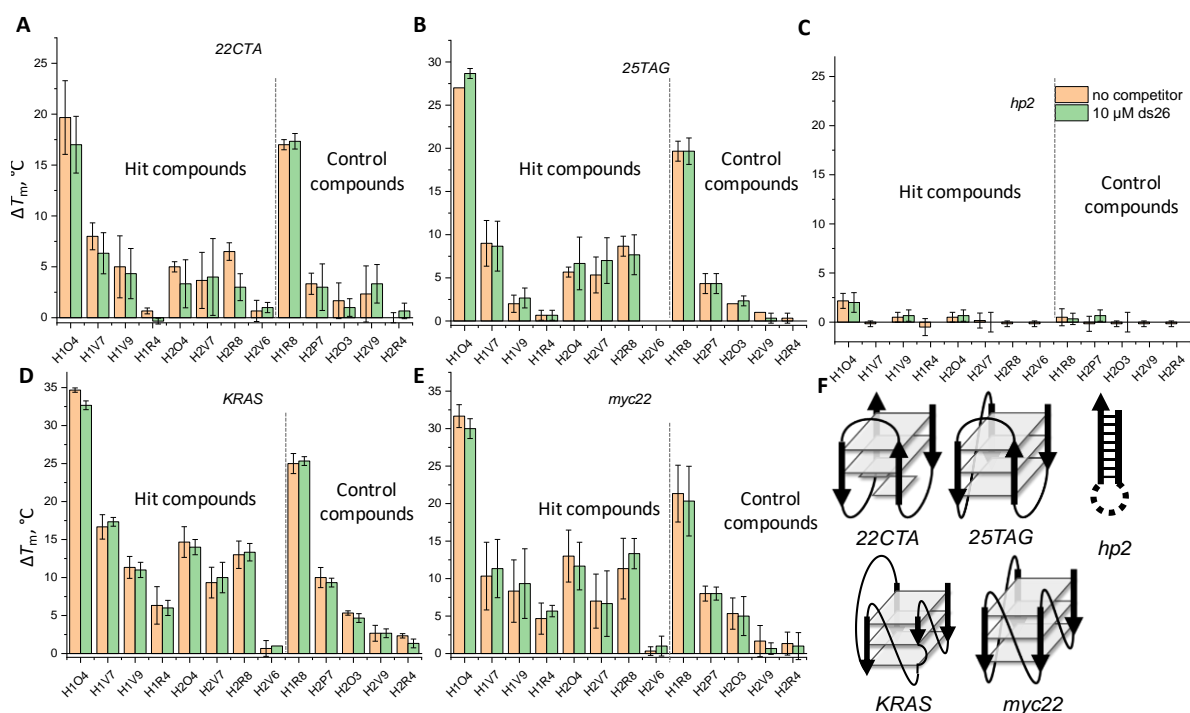
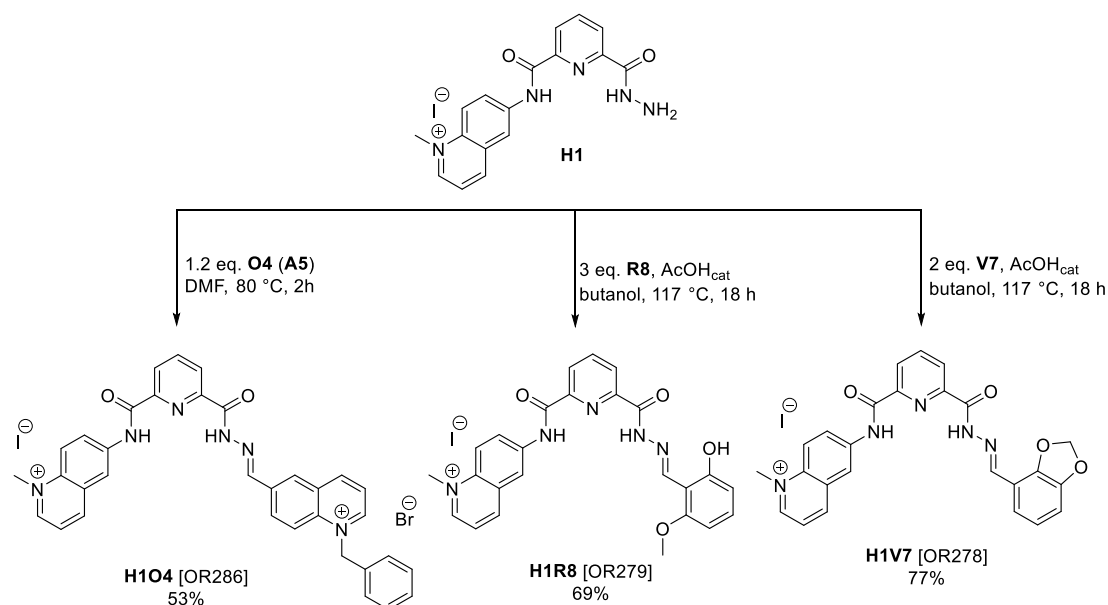


Figure 81. Results of FRET-melting experiment performed with “in situ” synthesized compounds and A) 22CTA, B) 25TAG, C) hp2, D) KRAS, E) myc22. F) Schematic structures of the double-labelled targets, used in this experiment.

As we can see from the results of FRET-melting experiment (Figure 81), “hit” compounds (on the left) generally produced higher  $\Delta T_m$  values than the “controls” (except for H1R8). Among the whole set of tested compounds only two of them, H104 and H1R8, significantly stabilized tested G-quadruplexes. Interestingly, H1R8 was the “control” compound, not selected from combinatorial libraries by SPE. The same tendency was present for all oligonucleotides: H104 systematically gave higher  $\Delta T_m$ , followed by H1R8. These two compounds, as well as H1V7, that provided moderate ( $\Delta T_m \approx 17$  °C) stabilization, but seem to be selective for KRAS G-quadruplex, were re-synthesized in a preparative manner (Scheme 20). The three compounds were obtained in good yields and characterized by  $^1\text{H}$  and NMR spectroscopy.



Scheme 20. Synthesis of lead compounds **H104**, **H1R8** and **H1V7**, selected by FRET-melting experiment with “in situ” synthesized compounds.

Further experiments with synthesized compounds are under the progress. Compared with the symmetric dicationic compounds described in *RESULTS AND DISCUSSION, Part 1*, non-symmetric scaffolds **H1R8** and **H1V7** are interesting in terms of lower charge (1+) and lower insaturation (higher fraction of sp<sup>3</sup>-hybridized atoms).

#### 4.2.5 Conclusions

In this part we tried to substitute the expensive streptavidin coated magnetic beads in the pull-down step of dynamic combinatorial chemistry experimental setup with cheaper options having less non-specific interactions. For this purpose, we first tried to synthesize gold-coated magnetic nanoparticles, as described by Dash et al., but unfortunately, the synthesis was irreproducible in our hands.

We then turned our attention on the method of solid phase extraction. This relatively unexplored method allows to immobilize the complex of the target with bound ligands, wash out the weakly bound ligands, and then analyze the compounds released from the complex. Slightly similar approach is used in the group of Daniela Montesarchio: they use the resin in cartridges that is covalently functionalized with G-quadruplexes, so libraries of ligands are just

loaded on the colon and the compounds that bound to G4s were then identified.<sup>180,181,182</sup> However, this method is not widely used in the G4 field.

The building blocks for the library design were already synthesized in the course of this work or purchased from commercial suppliers. Pseudo-static combinatorial libraries of ten combinations of building blocks were generated in the presence of two G-quadruplexes (*Pu24T* and *22CTA*) and control double-strand and hairpin targets. In two rounds of selection, eight hits were identified. In the third round of selection, these compounds were synthesized “in situ” using the method of synthesis in solution described in *RESULTS AND DISCUSSION, Part 3* and FRET-melting experiment was performed for these compounds with G-quadruplexes of different topologies. Three hits, identified from these experiments, were synthesized preparatively and the study of their interaction with G4s is ongoing.

# 5. Cationic bis(acylhydrazones) as potential drugs for treating EBV-related cancers

The mechanism of immune evasion developed by EBV is described in the INTRODUCTION, (section 1.5). The important role in this mechanism belongs to the process of inhibition of translation of EBNA1 mRNA by the host protein nucleolin which forms complex with G-quadruplexes present in mRNA EBNA1. As mentioned before, **PhenDC3** was found to disrupt the complex of EBNA1 mRNA with nucleolin and therefore enhance the translation EBNA1 mRNA.<sup>40</sup> Compounds, described in the *RESULTS AND DISCUSSION, Part 1* of this work, were also tested for their capacity to regulate the translation of viral genome maintenance protein EBNA1, in the context of collaborative research project with teams of Prof. Marc Blondel (Université de Bretagne Occidentale) and Dr. Robin Fåhraeus (Institut de Génétique Moléculaire).

## 5.1. *In silico* evaluation

The drug-like properties of 20 cationic bis(acylhydrazones) synthesized previously (cf. *RESULTS AND DISCUSSION, Part 1, Table 5*) were evaluated using SwissADME, a free tool allowing to assess the physicochemical descriptors, pharmacokinetics (ADME), and drug-likeness of small molecule.<sup>173</sup> It should be noted that certain methods of *in silico* evaluation are poorly suitable for cationic compounds as the ones described in this work, as illustrated by significant variance of physico-chemical properties predicted by different algorithms for the same compound (e.g. for **PyDH1**, cLogP from -7.75 to 1.83, *Table 20*). According to the established practices, consensus estimation (i.e., the average of five models) was used in evaluation of drug-likeness of compounds.<sup>183</sup> The results of *in silico* assessment (*Table 20*) indicated satisfactory

bioavailability for most bis(acylhydrazones), as given by the combination of six physicochemical descriptors (lipophilicity, molecular weight, polarity, solubility, saturation and number of rotatable bonds). Among those, insufficient saturation (i.e., low fraction of sp<sup>3</sup> carbons) could be identified as the limiting factor for bioavailability of most derivatives, as well as high molecular weight for the derivatives **NaphDH3**, **NaphDH5**, **PhenDH3** and **PhenDH5** ( $M > 700$  Da). With the exception of the latter four derivatives, all bis(acylhydrazones) satisfied the Lipinski's rule, and most derivatives also satisfied the Muegge's rule<sup>184</sup>, indicating a high potential to serve as drugs (*Table 20*). Finally, we employed a gastrointestinal absorption predictor calculated by the BOILED-egg model (*Figure 82*).<sup>185</sup> According to this model, most compounds, again with the exception of **NaphDH3**, **NaphDH5**, and **PhenDH3–5**, had a high rate of passive gastrointestinal absorption; however, none of compounds was expected to cross the blood-brain barrier (*Figure 82*). In conclusion, the results of *in silico* evaluation indicate that the cationic bis(acylhydrazone) scaffolds are compatible with their use as RNA-targeting drugs, while heavy and aromatic substitutes (e.g., R = Bn) should be avoided for the sake of bioavailability, unless indispensable for RNA binding.

Table 20. Selected physicochemical descriptors and computed (SwissADME)<sup>183</sup> ADME parameters, pharmacokinetic properties, and drug-like nature of bis(acylhydrazones)

Molecule	MW	Fract. C(sp <sup>3</sup> )	Rotat. bonds	H-bond accept.	H-bond donors	TPS A (A <sup>3</sup> )	Lipophilicity						ESOL Log S	ESOL Class	GI absorpt.	BBB permeant	Drug-likeness (no. of violations)			PAINS alerts
							iLOGP	XLOGP <sub>3</sub>	WLOGP	MLOGP	Silicos-IT Log P	Consensus Log P					Lipinski	Muegge	Bioavail. score	
PyDH1	403.4	0.1	8	5	2	103.6	-7.75	1.46	0.26	0.4	1.83	-0.76	-3.18	Soluble	High	No	Yes (0)	Yes (0)	0.55	0
PyDH2	503.5	0.07	8	5	2	103.6	-10.7	4.14	2.56	2.25	3.9	0.43	-5.55	Moderately sol.	High	No	Yes (1)	Yes (0)	0.55	1
PyDH3	655.7	0.05	12	5	2	103.6	-8.14	7.33	5.59	4.04	6.77	3.12	-8.29	Poorly sol.	High	No	Yes (1)	No (2)	0.55	0
PyDH4	531.6	0.13	10	5	2	103.6	-8.59	4.74	3.53	2.64	4.72	1.41	-5.94	Moderately sol.	High	No	Yes (1)	Yes (0)	0.55	1
PyDH5	655.7	0.05	12	5	2	103.6	-8.05	7.33	5.59	4.04	6.77	3.14	-8.29	Poorly sol.	High	No	Yes (1)	No (2)	0.55	0
PymDH1	404.4	0.1	8	6	2	116.5	-7.93	0.81	-0.35	-0.18	1.27	-1.28	-2.77	Soluble	High	No	Yes (0)	Yes (0)	0.55	0
PymDH2	504.5	0.07	8	6	2	116.5	-7.73	3.49	1.96	1.7	3.33	0.55	-5.15	Moderately sol.	High	No	Yes (1)	Yes (0)	0.55	1
PymDH3	656.7	0.05	12	6	2	116.5	-5.01	6.68	4.98	3.51	6.18	3.27	-7.89	Poorly sol.	Low	No	Yes (1)	No (2)	0.55	0
PymDH4	532.6	0.13	10	6	2	116.5	-5.37	4.09	2.93	2.08	4.14	1.57	-5.54	Moderately sol.	High	No	Yes (1)	Yes (0)	0.55	1
PymDH5	656.7	0.05	12	6	2	116.5	-4.32	6.68	4.98	3.51	6.18	3.41	-7.89	Poorly soluble	Low	No	Yes (1)	No (2)	0.55	0
NaphDH1	454.5	0.08	8	6	2	116.5	-8.82	2.06	0.81	1.05	2.3	-0.52	-3.91	Soluble	High	No	Yes (0)	Yes (0)	0.55	0
NaphDH2	554.6	0.06	8	6	2	116.5	-6.19	4.74	3.11	2.57	4.35	1.72	-6.27	Poorly soluble	High	No	Yes (1)	Yes (0)	0.55	1
NaphDH3	706.8	0.05	12	6	2	116.5	-6.19	7.93	6.14	4.29	7.2	3.87	-9.00	Poorly soluble	Low	No	No (2)	No (3)	0.17	0
NaphDH4	582.6	0.12	10	6	2	116.5	-5.98	5.34	4.08	2.94	5.17	2.31	-6.66	Poorly soluble	High	No	Yes (1)	No (1)	0.55	1
NaphDH5	706.8	0.05	12	6	2	116.5	-7.24	7.93	6.14	4.29	7.2	3.66	-9.00	Poorly soluble	Low	No	No (2)	No (3)	0.17	0
PhenDH1	504.5	0.07	8	6	2	116.5	-7.36	3.07	1.96	1.29	3.33	0.46	-4.88	Moderately sol.	High	No	Yes (1)	Yes (0)	0.55	0
PhenDH2	604.7	0.06	8	6	2	116.5	-8.05	5.74	4.27	2.74	5.37	2.01	-7.22	Poorly soluble	High	No	Yes (1)	No (2)	0.55	1
PhenDH3	756.9	0.04	12	6	2	116.5	-7.39	8.93	7.29	4.65	8.21	4.34	-9.95	Poorly soluble	Low	No	No (2)	No (3)	0.17	0
PhenDH4	632.7	0.11	10	6	2	116.5	-5.93	6.34	5.23	3.36	6.19	3.04	-7.62	Poorly soluble	Low	No	Yes (1)	No (2)	0.55	1
PhenDH5	756.9	0.04	12	6	2	116.5	-5.78	8.93	7.29	4.65	8.21	4.66	-9.95	Poorly soluble	Low	No	No (2)	No (3)	0.17	0
PhenDC3	550.6	0.06	6	4	2	91.7	-5.15	5.59	4.86	2.96	4.53	2.56	-6.98	Poorly soluble	High	No	Yes (1)	No (1)	0.55	1

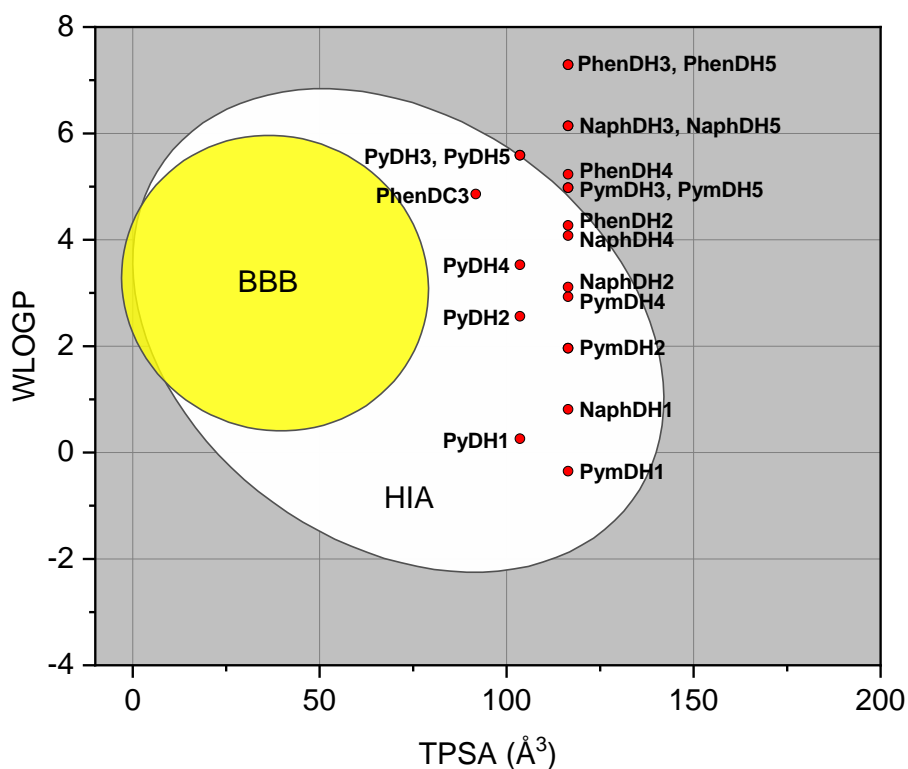


Figure 82. BOILED-egg evaluation<sup>185</sup> of passive gastrointestinal absorption (HIA) and brain (BBB) penetration of bis(acylhydrazones). The white region indicates high probability of passive gastrointestinal absorption, and the yellow region indicates high probability of BBB penetration.

The compounds were also assessed for the presence of fragments associated with pan-assay interfering compounds (PAINS). The presence of quinolinium fragments was identified as potentially troublesome in some (but not all) compounds (Table 20); however, an inspection of the original screen revealed that this fragment was mainly associated with cationic dyes, potentially interfering with colorimetric or fluorimetric assays.<sup>186</sup> Since bis(acylhydrazones) are colorless or faintly colored compounds, this filter was not applied and all compounds were systematically assessed by biophysical and biological tests.

## 5.2 Biophysical studies of interaction of novel ligands with G4 of EBNA1 mRNA and G4-DNA

The interaction of novel bis(acylhydrazone) derivatives with the most frequent G4-forming motif encountered in GAR-encoding sequence of the EBNA1 mRNA (g4-EBNA1, 5'-r(GGGGCAGGA-GCAGGAGGA)-3') was studied *in vitro* using two widely established methods, FRET-melting and FID assay. In our conditions (10 mM KAsO<sub>2</sub>Me<sub>2</sub>, 10 mM KCl, 90 mM LiCl buffer, pH 7.3) and in the absence of ligands, F-g4-EBNA1-T denatured at  $T_m^0 = 58.6$  °C, which is slightly higher than the reported value (54 °C).<sup>187</sup> The results (*Figure 83a*) demonstrate that bis(acylhydrazone) derivatives display great variability with respect to their capacity to bind to and stabilize g4-EBNA1, revealing interesting structure–activity relationships discussed in details below. Specifically, most derivatives of the PyDH and NaphDH families, as well as **PhenDH1**, demonstrated significant stabilization of g4-EBNA1 ( $\Delta T_m = 10$  to 20 °C), and further phenanthroline derivatives (**PhenDH2–5**) demonstrated even higher stabilization of the substrate ( $\Delta T_m = 20$  to 30 °C), comparable to the result obtained with **PhenDC3** ( $\Delta T_m = 30.0$  °C). In contrast, all derivatives of the PymDH family, as well as compounds **PyDH1** and **NaphDH1**, demonstrated low or very low stabilization of G4-RNA ( $\Delta T_m < 10$  °C), illustrating the importance of the nature of heterocyclic residues on the G4-RNA binding properties of ligands. Finally, most derivatives that stabilized g4-EBNA1 also displayed significant level of selectivity with respect to ds-DNA, as their stabilizing effect was almost unaffected by the presence of ds-DNA competitor. The comparison of this data with the results of FRET-melting experiments performed with DNA G-quadruplexes (cf. *RESULTS AND DISCUSSION, Part 1*) demonstrates no preferential stabilization of one or another substrate, since all compounds that stabilized g4-EBNA1 (**PyDH2–5**, **NaphDH2–5**, and all PhenDH derivatives) also strongly stabilized both G4-DNA substrates (with  $\Delta T_m = 15$  to 30 °C). A similar behavior was observed with **PhenDC3**, which stabilized both G4-DNA with  $\Delta T_m$  of over 35 °C. The lack of selectivity is not surprising given the structural similarity of novel ligands to **PhenDC3** (a polyvalent G4 binder) and indicates a similar binding mode, relying on compound stacking with terminal G-tetrads of G4 structures. Obviously, the substituents explored in this work (ethyl or phenyl groups) are not sufficient for achieving a significant level of selectivity between different G4 structures.



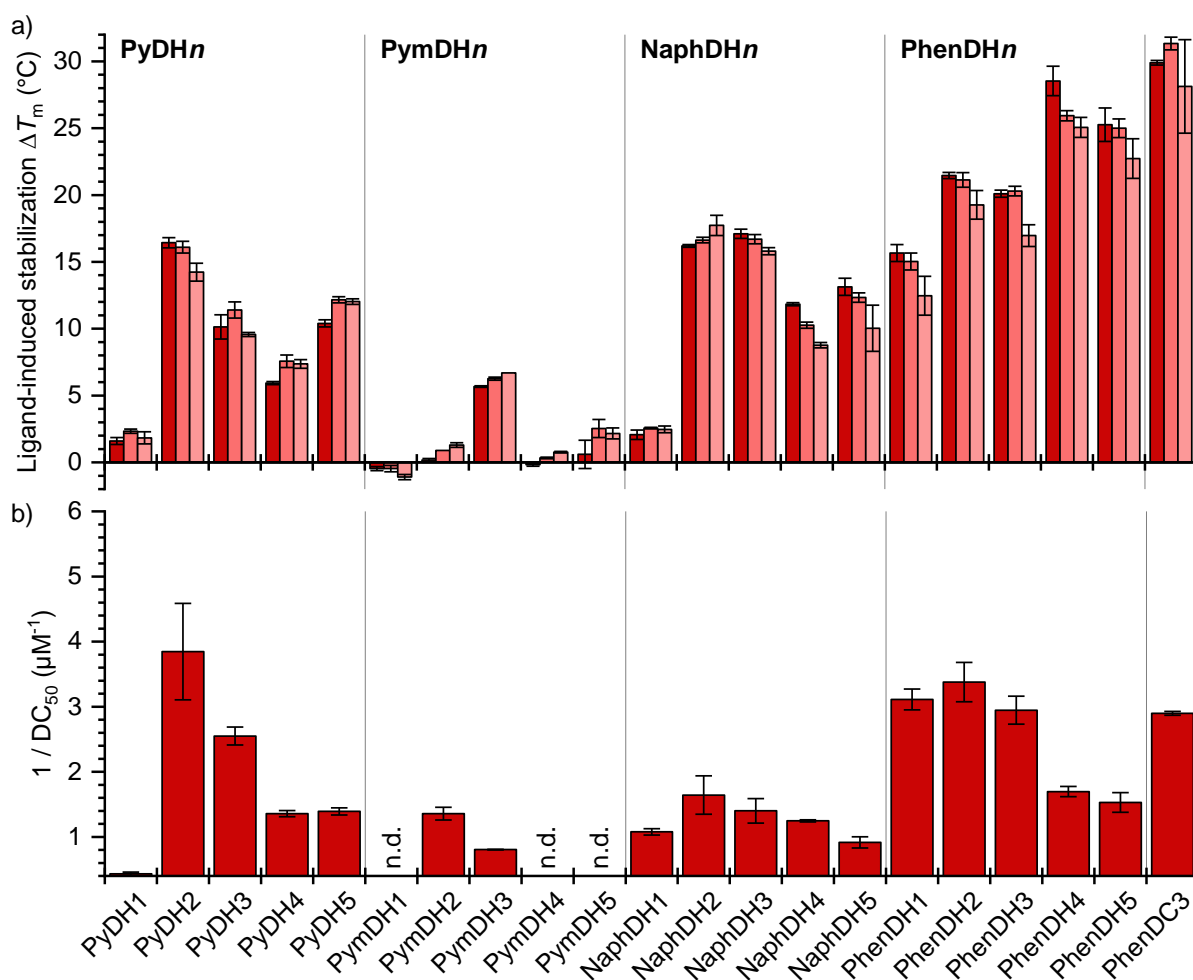


Figure 83. *In vitro* binding of tested compounds to g4-EBNA1. a) Thermal stabilization F-g4-EBNA1-T (0.2 μM) by tested compounds (1.0 μM), assessed by fluorescence melting experiments in the absence (dark red bars) or in the presence of duplex DNA competitor ds26 (red bars: 3 μM, pale red bars: 10 μM). The experiments were performed in K10 buffer (10 mM LiAsO<sub>2</sub>Me<sub>2</sub>, 10 mM KCl, 90 mM LiCl, pH 7.3); data are means ± s.d. from three technical replicates. b) Ligand-induced displacement of TO (0.5 μM) from g4-EBNA1 (0.25 μM). Data are means ± s.d. from three technical replicates; n.d. = no displacement (DC<sub>50</sub> > 2.5 μM). The experiments were performed in K100 buffer (10 mM LiAsO<sub>2</sub>Me<sub>2</sub>, 100 mM KCl, 1 mM EDTA, 1% v/v DMSO).

A complementary information about ligand affinity for g4-EBNA1 was obtained from the FID assay, which assesses the binding of ligands through displacement of a fluorescent probe (Thiazole Orange, TO). In this assay, the apparent ligand affinity is given by the concentration required to displace 50% of the bound probe (DC<sub>50</sub>).<sup>188</sup> The results of the FID assay (Figure 83b and Table 21) indicated high g4-EBNA1 affinity (DC<sub>50</sub> < 0.5 μM) for compounds **PyDH2**, **PyDH3**, as well as **PhenDH1–3**, which was comparable to the result obtained for **PhenDC3** (DC<sub>50</sub> = 0.31 μM). In contrast, the derivatives **PyDH1**, **PymDH1**, **PymDH4** and **PyDH5** were not able to displace TO from g4-EBNA1 (DC<sub>50</sub> > 2 μM), giving evidence of low affinity, whereas other derivatives demonstrated moderate affinity (DC<sub>50</sub> = 0.5 to 2 μM). Of note, none of tested ligands was able to

induce displacement of TO from the double-stranded DNA substrate ds26 ( $DC_{50} > 2.5 \mu\text{M}$  in all cases). Globally, the results obtained with both biophysical methods were in a good agreement, except for a few discrepancies (compounds **PyDH5**, **NaphDH2–5** and **PhenDH4–5** which demonstrated high thermal stabilization of g4-EBNA but only moderate capacity to displace TO,  $DC_{50} = 0.5$  to  $1 \mu\text{M}$ ). Altogether, the results of biophysical studies point to pyridine derivatives **PyDH2** and **PyDH3**, as well as all derivatives of the phenanthroline series, as most promising ligands for g4-EBNA.

*Table 21. Results of the TO displacement assay ( $DC_{50}$ , in  $\mu\text{M}$ ) with tested compounds. Conditions: TO ( $0.5 \mu\text{M}$ ), g4-EBNA1 ( $0.25 \mu\text{M}$ ). Data are means  $\pm$  s.d. from three technical replicates. Buffer: 10 mM  $\text{LiAsO}_2\text{Me}_2$ , 100 mM KCl, 1 mM EDTA, 1% v/v DMSO.*

	DH1	DH2	DH3	DH4	DH5
<b>Py</b>	> 2.5	0.18	0.26	0.72	0.66
<b>Pym</b>	> 2.5	0.59	1.25	2.14	> 2.5
<b>Naph</b>	0.78	0.6	0.64	0.86	1.08
<b>Phen</b>	0.25	0.23	0.25	0.62	0.64
<b>PhenDC3</b>	0.31				

### 5.3. Biological tests<sup>3</sup>

#### 5.3.1. Effect of compounds on GAR-dependent protein expression

To assay the biological activity of the various compounds bis(acylhydrazones), we first determined their effect on the GAR-dependent limitation of EBNA1 expression, which represents a mechanism at the basis of EBNA1/EBV immune evasion. For this purpose, human lung carcinoma H1299 cells were transfected with EBNA1 or EBNA1 $\Delta$ GAr constructs and treated with  $10 \mu\text{M}$  of the indicated compounds or, as a control, with DMSO (the vehicle). Then, the levels of EBNA1 or EBNA1 $\Delta$ GAr was assessed by western blot analysis using an antibody raised against EBNA1 and, as a loading control, an antibody raised against GAPDH. **PhenDC3** was used as a positive control and DMSO as a negative control. In this assay, three types of results were observed (*Table 22*): (i) most of the compounds had no effect (indicated by a “0”); (ii) two

<sup>3</sup> Experiments performed in the lab of Prof. Marc Blondel

compounds, namely **PyDH2** and **PhenDH2**, similarly to **PhenDC3**, led to a GAR-dependent increase in EBNA1 expression, indicated by a “+” (i.e., they increased EBNA1 level while having no effect on EBNA1ΔGAR); and finally (iii) three compounds (**PyDH3**, **PymDH5**, and **PhenDH1**) increased EBNA1 level in a GAR-independent manner indicated by a “x” (i.e., they increased both EBNA1 and EBNA1ΔGAR levels). The western blot results obtained with compounds **PyDH2** and **PhenDH2** are presented on *Figure 84a*. As we were interested in compounds able to interfere with the GAR-dependent limitation of EBNA1 expression and antigen presentation, a mechanism at the basis of EBNA1 immune evasion, we focused only on the compounds that increase EBNA1 level while having no effect on EBNA1ΔGAR, namely **PhenDH2** and **PyDH2**.

*Table 22. Effect of compounds (all tested at 5 μM) on EBNA1 expression in H1299 cells.*

Comp.	Effect <sup>a</sup>	Comp.	Effect <sup>a</sup>	Comp.	Effect <sup>a</sup>	Comp.	Effect <sup>a</sup>
<b>PyDH1</b>	0 (106/93) <sup>b</sup>	<b>PymDH1</b>	0 (80/92) <sup>b</sup>	<b>NaphDH1</b>	0 (112/99) <sup>b</sup>	<b>PhenDH1</b>	x (78/76) <sup>b,c</sup>
<b>PyDH2</b>	+ (162/101) <sup>b</sup>	<b>PymDH2</b>	0 (89/103) <sup>b</sup>	<b>NaphDH2</b>	0 (95/100) <sup>b</sup>	<b>PhenDH2</b>	+ (189/96) <sup>b</sup>
<b>PyDH3</b>	x (134/149) <sup>b</sup>	<b>PymDH3</b>	0 (108/106) <sup>b</sup>	<b>NaphDH3</b>	0 (106/93) <sup>b</sup>	<b>PhenDH3</b>	0 (96/114) <sup>b</sup>
<b>PyDH4</b>	0 (86/98) <sup>b</sup>	<b>PymDH4</b>	x (48/47) <sup>b,c</sup>	<b>NaphDH4</b>	0 (82/111) <sup>b</sup>	<b>PhenDH4</b>	0 (80/96) <sup>b</sup>
<b>PyDH5</b>	x (58/70) <sup>b,c</sup>	<b>PymDH5</b>	x(139/137) <sup>b</sup>	<b>NaphDH5</b>	0 (104/108) <sup>b</sup>	<b>PhenDH5</b>	0 (111/94) <sup>b</sup>

<sup>a</sup> 0: No effect ; +: GAR-dependent increase in EBNA1 expression; x: GAR-independent effect. <sup>b</sup> Number between brackets indicate the quantification of respectively GAR-OVA (numerator) and OVA (denominator) levels, as compared to their respective levels in DMSO-treated cells. <sup>c</sup> Values in italics, which are significantly smaller than 100, point to compounds that are toxic to both GAR-OVA- and OVA-expressing cells.

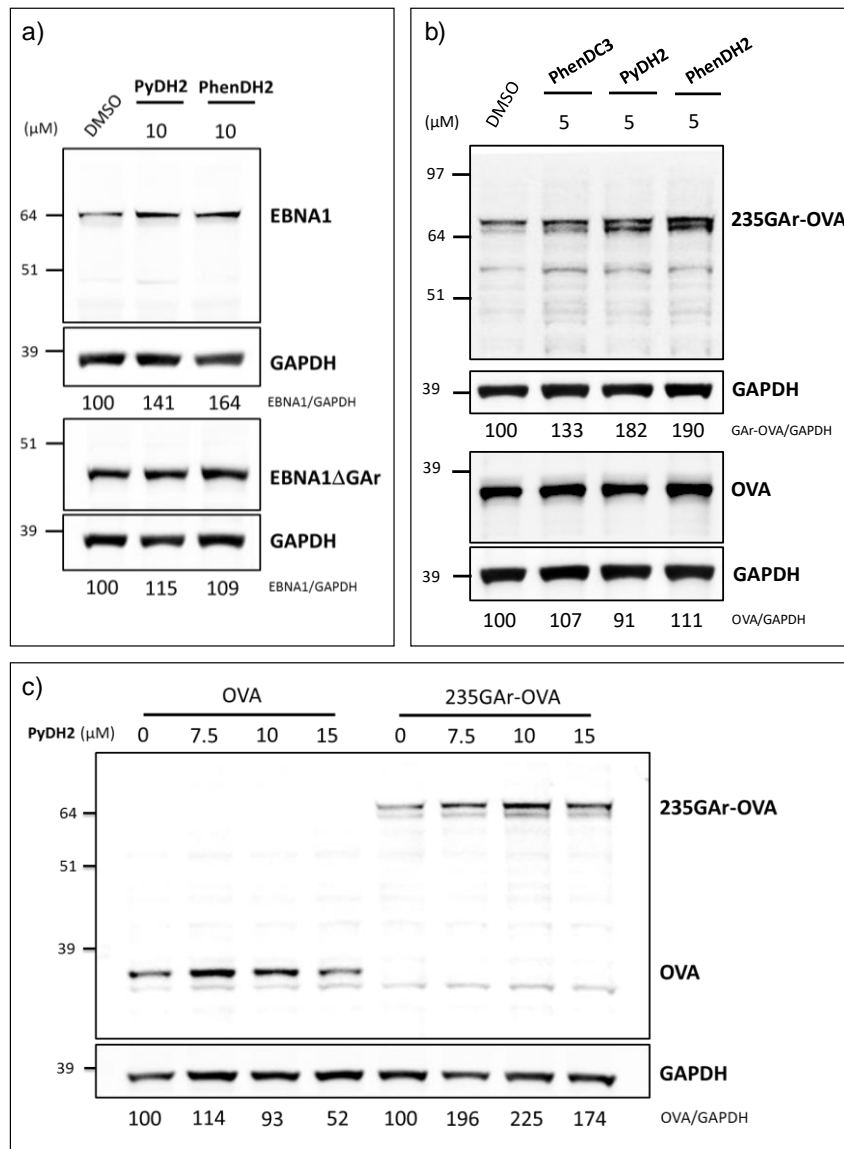


Figure 84. a) Expression of EBNA1 (top panel) or EBNA1ΔGAR (bottom panel) in transfected H1299 cells treated with DMSO (control), **PyDH2** (10 μM), or **PhenDH2** (10 μM) 40 h post-transfection. b) Expression of 235GAR-OVA (top panel) or OVA (bottom panel) in transfected H1299 cells treated with DMSO (control), **PhenDC3**, **PyDH2**, or **PhenDH2** (all at 10 μM) 40 h post-transfection. c) Expression of 235GAR-OVA or OVA in transfected H1299 cells treated with **PyDH2** used at 0 (control), 7.5, 10, or 15 μM concentration. Protein (EBNA1, EBNA1ΔGAR, 235GAR-OVA or OVA) levels were normalized with respect to GAPDH (loading control) and the resulting values indicated below the gels.

On the basis of the compounds effect on the expression of EBNA1 and *in silico* evaluation of their drug-likeness, **PyDH2** and **PhenDH2** were selected for further biological studies that include antigen presentation assay and proximity ligation assay (PLA). For this purpose, we first evaluated the effect of various concentrations of both compounds on expression of ovalbumin (OVA) and 235GAR-OVA in H1299 cells using the same procedure as described above for EBNA1. The OVA/235GAR-OVA system allows to assess the ability of GAR to limit both protein expression and

antigen presentation since 235GAR (a full-length, 235 amino-acid GAR domain), when fused to OVA, strongly limits both its expression and its antigen presentation by the MHC class I pathway. In this way, this model recapitulates the effect of GAR on EBNA1 expression and antigen presentation.<sup>189</sup> Similar to what has been observed with **PhenDC3**,<sup>38</sup> compounds **PyDH2** and **PhenDH2** increased 235GAR-OVA level in a dose-dependent manner, whereas having no significant effect on OVA expression (Figure 84b,c). Together with the data on the effect of compounds on EBNA1 expression, these results confirm the suggested mechanism of biological activity of both compounds, namely, relieving of the GAR-dependent inhibition of protein expression.

### **5.3.2. Inhibition of NCL binding to G4 in EBNA1 mRNA: proximity ligation assay (PLA) and RNA pull-down assay<sup>4,5</sup>**

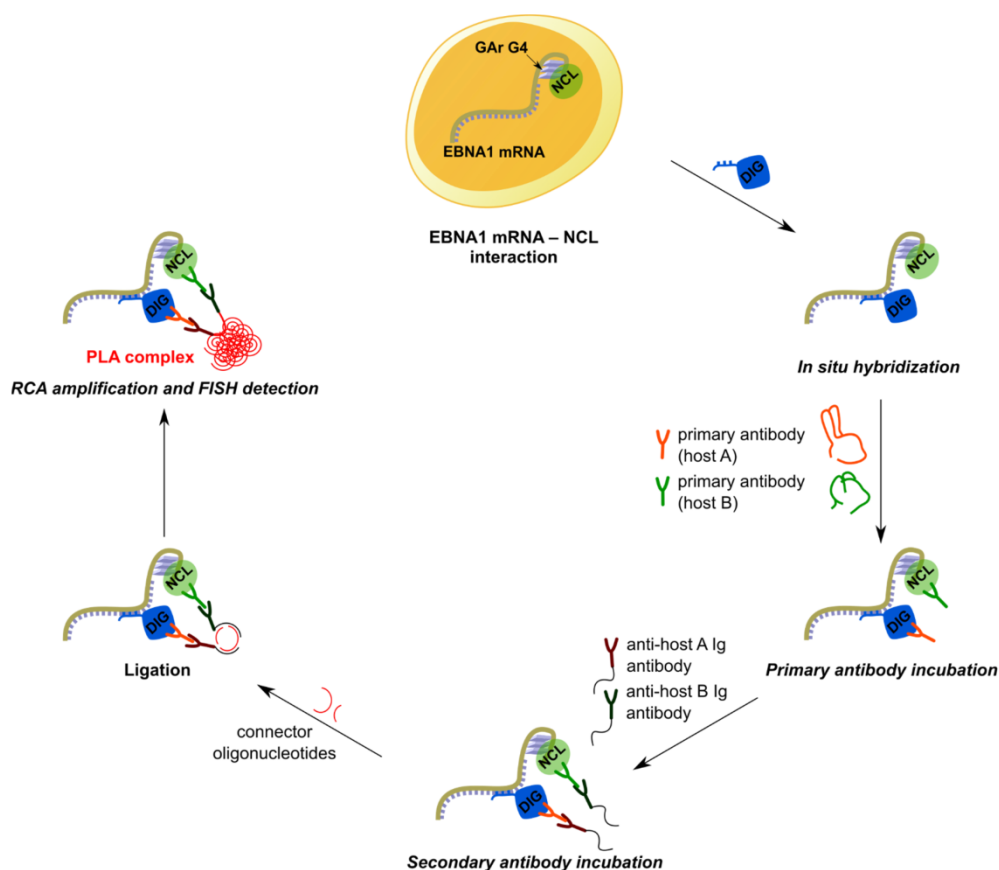
We employed the proximity ligation assay (PLA)<sup>190</sup> to verify that compounds **PyDH2** and **PhenDH2** prevent nucleolin (NCL) to interact with the G4s of the GAR-encoding sequence of EBNA1 mRNA, the mechanism at the basis of the GAR-dependent limitation of EBNA1 expression and antigen presentation. Briefly, PLA is a technique originally developed to detect proteins in close proximity (theoretically, at a maximum distance of 40 nm) and which is based on the use of a pair of antibodies raised in two different species, each targeting one of the two proteins of interest. By using labelled oligomers, PLA has been adapted to the study of protein–DNA<sup>191</sup> and protein–RNA interactions. In the latter case, mRNA of interest (EBNA1) is tagged through *in situ* hybridization with a digoxigenin-labelled DNA probe, followed by incubation with a mouse anti-digoxigenin, whereas NCL is tagged with rabbit anti-nucleolin. If both are located in a close proximity, the subsequent incubation with DNA probe-conjugated anti-rabbit (*plus* probe) and anti-mouse (*minus* probe) antibodies and two connector oligonucleotide probes induces rolling circle amplification, as schematically shown on *Figure 85*. The latter generates a concatemeric DNA product, which is finally detected thorough hybridization with a fluorescently labelled oligonucleotide probe as a distinct bright spot on a micrograph.<sup>192</sup> This way, we previously

---

<sup>4</sup> Experiments performed in the lab of Prof. Marc Blondel

<sup>5</sup> Experiments performed in the lab of Dr. Robin Fåhraeus

demonstrated that NCL interacts with G-quadruplexes formed in the GAR-encoding sequence of EBNA1 mRNA, and that **PhenDC3** inhibits this interaction in Mutu-1 cells as well as in H1299 cells transiently expressing EBNA1.<sup>38,192</sup> Herein, we first exploited H1299 cells transiently expressing EBNA1 following transfection with EBNA1 plasmid. Cells treated with DMSO control demonstrated high level of PLA signals ( $1.83 \pm 0.54$  per cell) further confirming that the EBNA1 mRNA–NCL interaction takes place in, or at the close vicinity of the nucleus (*Figure 86a*), whereas non-transfected cells did not display PLA signals. Treatment with compounds **PyDH2** or **PhenDH2** both at a concentration of 5  $\mu\text{M}$  significantly reduced both the number ( $0.27 \pm 0.11$  and  $0.40 \pm 0.21$  per cell, in cells treated with **PyDH2** and **PhenDH2**, respectively) and the intensity of PLA signals (*Figure 86b–d*), hence confirming the ability of these new derivatives to disrupt the EBNA1 mRNA–NCL interaction, *in fine* leading to the GAR-dependent inhibition of EBNA1 expression and presentation of EBNA1-derived antigenic peptides.



*Figure 85. Principle of the proximity ligation assay (PLA) applied to the study of EBNA1 mRNA–NCL interaction. The DNA probe (50-CTTTCAAACCACCCTCTTTTTTGCGCTGCCCTCCATCAAAAA-igoxigenin-30) was chosen such that hybridization occurs outside the G4-forming region, to avoid eventual interference between G4 formation and RNA–DNA hybridization (adapted from <sup>39</sup>).*

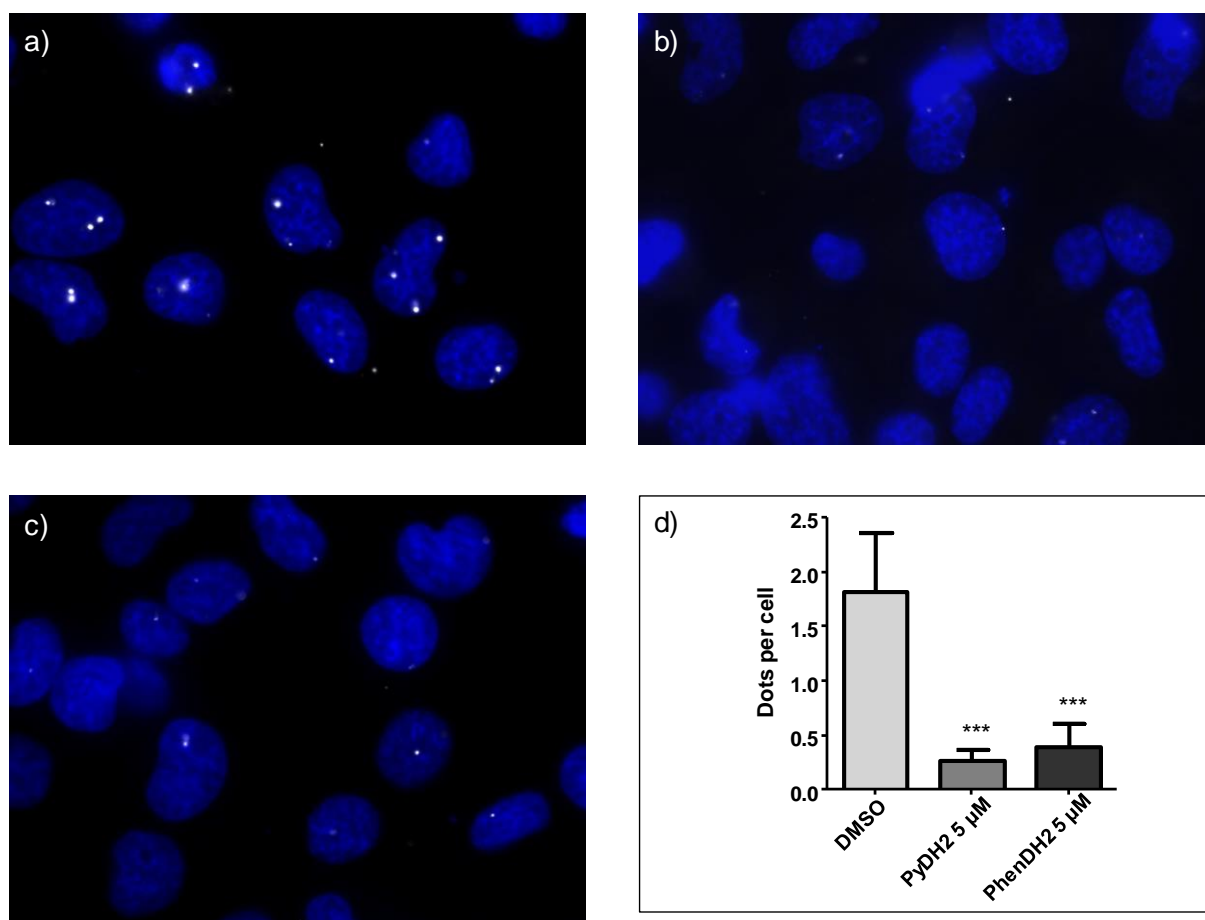


Figure 86. Proximity ligation assay (PLA) performed in H1299 cells transiently expressing EBNA1. a–c) Microscopy images of cells treated with a) DMSO (compound vehicle, negative control), b) **PyDH2** (5 μM), and c) **PhenDH2** (5 μM). Nuclei were revealed by DAPI staining and appear in blue; white dots (PLA signals) indicate interaction between NCL and G4 of EBNA1 mRNA. As previously observed, this interaction mostly takes place in, or at the close vicinity of the nucleus.<sup>38,192</sup> d) Number of nuclear PLA signals (dots) per cell in H1299 cells expressing EBNA1 and treated with DMSO (control), **PyDH2** (5 μM) or with **PhenDH2** (5 μM). Data from two biological replicates, 100 cells per sample were analyzed. \*\*\*,  $p < 0.001$ .

## 5.4 Pharmacological properties of selected drug candidates

The study of pharmacological properties of compounds **PyDH2**, **PhenDH2** were carried out at the Platform of Integrative Chemical Biology of Strasbourg (PCBIS). The results are presented in (Table 23). As it was known from our preliminary data, all compounds are stable and moderately soluble in aqueous buffers. **PyDH2** was found to be slightly hydrophilic while **PhenDH2** that has larger aromatic surface was hydrophobic. In parallel artificial membrane permeability assay

(PAMPA) where the capacity of the compound to pass the lipid membrane is studied, both **PyDH2** and **PhenDH2** had very low permeability that suggests their poor oral bioavailability. However, as it was observed for other drugs with hydrazone scaffold (cf. *AIMS OF THE WORK*) the compounds were stable in plasma (almost no degradation after 2 hours). Compound **PhenDH2** was to be stable in the presence of liver mice microsomes and NADPH that suggests that this compound is not the substrate of cytochromes P450. The amount of **PyDH2** was decreased to 40% suggesting its  $t_{1/2} = 41 \pm 4$  min. However, the same result was observed in the absence of NADPH, so other processes may be involved in the consumption of the compound. The level of plasma protein binding was 90% for **PhenDH2** and only 44% of **PyDH2**. Pharmacokinetics studies suggest  $t_{1/2}$  less than one hour for **PyDH2** and **PhenDH2** (54 and 20 minutes, respectively) that in the combination with low distribution volume suggests low exposure to the compound.

Table 23. Experimentally determined pharmacological properties of compounds **PyDH2**, **PhenDH2** and **PhenDC3**.

Criterion	PyDH2	PhenDH2	PhenDC3
Chemical stability <sup>a</sup>	Stable <sup>g</sup>	stable	stable
Solubility <sup>a</sup>	52 ± 9 µM	17 ± 8 µM	13 ± 2 µM
log D <sup>b</sup>	-0.5 ± 0.4	0.8 ± 0.3	-0.74
PAMPA (log Pe) <sup>c</sup>	-7.58 ± 0.23	-7.59 ± 0.27	-6.4 ± 0.1
PAMPA (% passage) <sup>d</sup>	0	0	4%
PK (mice)	54 min	20 min	37 min
	54 ml/min/kg	6 ml/min/kg	58 mL / min / kg
Plasma stability <sup>e</sup>	Stable ( $t_{1/2} > 120$ min)	Stable ( $t_{1/2} > 120$ min)	Stable ( $t_{1/2} > 120$ min)
Metabolic stability <sup>f</sup>	$t_{1/2} = 41 \pm 4$ min	Stable ( $t_{1/2} > 60$ min)	Stable (> 98% at 120 min)
Plasma protein binding	44 ± 2%	90 ± 0%	74 ± 2%

Conditions: <sup>a</sup> PBS pH 7.4; <sup>b</sup> PBS pH 7.4 / octanol; <sup>c</sup> 16 h, PBS pH 7.4 / dodecane; <sup>d</sup> PBS pH 7.4, 5% DMSO, membrane : dodecane 2% phosphatidylcholine; <sup>e</sup> mice plasma: PBS 1:1, 37 °C, 2 h; <sup>f</sup> Liver mice microsomes, NADPH, 37 °C; <sup>g</sup> No degradation after 24 hours.

While comparing with *in silico* predictions (Table 20), one can say that moderate solubility and low permeability of **PyDH2** and **PhenDH2** were confirmed in pharmacological studies. However, while hydrophilic properties of **PyDH2** were confirmed, compound **PhenDH2**, on the contrary, was found to be hydrophobic. Altogether, poor oral bioavailability (as evidenced by the low membrane permeability) is the key issue that needs to be solved for pharmacological utilization



of these compounds. In this context, intravenous administration or the use of drug delivery systems such as liposomes may be considered in the future.

## 5.5 Conclusions

Among the 20 tested bis(acylhydrazones), **PyDH2** and **PhenDH2** were found to increase EBNA1 expression in a GAR-dependent manner in H1299 cells (*Table 22*). Remarkably, both compounds demonstrated high affinity to g4-EBNA1 according to two biophysical methods (FRET-melting and FID assay, *Figure 83*). These results speak in favor of the expected mechanism for their biological activity, namely, interference with the NCL-based inhibition of mRNA translation by preventing the interaction between NCL and G4 structures in EBNA1 mRNA, as further confirmed by PLA and RNA pull-down assays. Considering the absence of selective targeting of G4-RNA with respect to G4-DNA *in vitro*, the cellular activity of the two compounds can be attributed to the fact that GAR repeat of EBNA1 mRNA contains a cluster of multiple ( $\approx 13$ ) G4-forming sequences, which may be particularly susceptible to ligand-induced effects such as NCL displacement. Moreover, G4-DNA structures form only transiently during DNA transactions such as replication, transcription and recombination, and are otherwise disfavored in the double-stranded DNA context; this is not the case with RNA whose single-stranded nature favors the formation of long-lived secondary motifs such as G4. PLA experiments provided further support this mechanism, demonstrating the ability of compounds to disrupt the interaction between NCL and the GAR-encoding sequence of EBNA1 mRNA in cells. Consistently, both compounds were found to increase expression of 235GAR-OVA and the level of antigen presentation in the GAR-OVA model, while having no effect on cells expressing OVA.

Interestingly, four naphthyridine derivatives (**NaphDH2–5**) demonstrated significant stabilization of g4-EBNA1, according to fluorescence melting experiments (*Figure 83a*), but had no influence on EBNA1 expression (*Table 22*). This behavior is not without precedent, as it was also observed with pyridostatin (**PDS**), another well-studied G4 binder<sup>38</sup>, and indicates that ligand-induced stabilization of a G4 structure is not a prerequisite to prevent its interaction with NCL by a competitive mechanism. In fact, one can imagine that stabilization of G4 structure could even favor their interaction with NCL. Hence, G4 stabilization and ability to interfere with NCL-G4

interaction may be two independent events. In line, all NaphDH derivatives were significantly less active in displacing the fluorescence probe TO from g4-EBNA1 ( $DC_{50} > 0.5 \mu\text{M}$ ), in contrast to the two biologically active derivatives, **PyDH2** and **PhenDH2** (Figure 83b). This also indicates that FID assay is potentially better suited for identification of G4-targeting drugs acting through interference with protein binding to G4 structures. Therefore, the different compounds described here, in addition to constitute promising scaffolds for drugs able to unveil EBV-related tumors to the immune system and more generally to interfere with G4-RNA binders, may also represent useful tools to decipher the mode of interaction between cellular factors and G4.

In this work, we demonstrated that harnessing the bis(acylhydrazone) motif allows facile generation of series of derivatives differing in terms of their physico-chemical properties, drug-like character, G4-binding properties, and biological activity. Specifically, our results demonstrate that the acylhydrazone group does not significantly impart the G4-binding properties of compounds, as compared with carboxamide analogues: instead, the binding to G4 structures seems mostly governed by the nature of central and lateral heterocyclic residues, revealing interesting SARs that can be interpreted in terms of molecular structure and preorganization of ligands. Even though the modifications of the scaffold explored in this work were not sufficient to achieve preferential targeting of viral G4-RNA with respect to G4-DNA *in vitro*, the results of biological assays demonstrate that two compounds, namely **PyDH2** and **PhenDH2**, displayed promising biological activity in EBV-related cellular models, being able to interfere with the GAR-dependent limitation of protein expression and antigen presentation. Moreover, both compounds were significantly less toxic than the prototype drug **PhenDC3** when used at concentrations required for boosting the production of EBNA1-derived antigenic peptides (i.e., 5–10  $\mu\text{M}$ ). Therefore, these compounds represent promising drug candidates for interfering with the immune evasion of EBV. Last but not least, the modular bis(acylhydrazone) scaffold presented here represents a promising platform for the development of novel ligands targeting other therapeutically important G4-RNA and/or G4-DNA structures.



***CONCLUSIONS AND  
PERSPECTIVES***

The principal purpose of this work was to adapt the methodology of dynamic combinatorial chemistry of acylhydrazones to the search of potential G-quadruplex binders. We started with synthesis of 20 model bis(acylhydrazone) compounds and evaluating their properties to stabilize G-quadruplexes in FRET-melting experiments. We also performed docking with 17 compounds of this library in attempts to find a correlation between  $\Delta T_m$  and gScore. The low correlation was found, however, because of numerous drawbacks of the method we did not proceed with docking.

We started DCC experiments by generation of simple dynamic combinatorial libraries followed by their direct HPLC analysis, however, while the HPLC conditions for good separation of all library components (up to 20) were optimized, the analysis of target-templated combinatorial libraries was not successful because of high stability of G4–ligand complexes that did not dissociate in the conditions of HPLC analysis. It could be possible to analyze static libraries of compounds in the presence and in the absence of DNA target and to identify the best binders just by depletion of their peaks in target–templated libraries (because they would be bound and, therefore, depleted on the HPLC chromatogram). However, this very simple and attractive option required the step of the separation of the complex G4–ligand from the library, because while direct injection of the solution that contains DNA–ligand complexes, the peaks of library components became broad and their assignment was no longer possible. Possibly, this issue could be resolved by using another type of HPLC column, but we decided to explore different methods of the separation of the complex of G4–ligand from the mixture of unbound compounds by different pull-down techniques.

First, we successfully applied the method of pull-down based on the interaction of biotin-modified G4-DNA target with streptavidin coated magnetic beads, so that the complex could be separated from the mixture of unbound ligands using a magnet. We successfully introduced the dynamic combinatorial chemistry of acylhydrazone exchange to the G4 field and through the analysis of two dynamic combinatorial libraries, DCL2 and DCL3 (of 14 and 18 components) found a non–symmetric hit **A2-L1-A8**, selectively pulled-down by the G-quadruplex *Pu24T*. However, we encountered some difficulties while trying to synthesize this compound in preparative manner. We could not separate in our conditions **A2-L1-A8** due to the impurities of the symmetric bis(acylhydrazones) **A2-L1-A2** and **A8-L1-A8** that formed during the synthesis and could not be removed, so we designed and synthesized two close analogues (**24a** and **24b**) of this

compound, in which one of acylhydrazone group was substituted with amide group. Even though the enhanced binding capacity of **A2-L1-A8** was not confirmed in the biophysical experiments with its close analogues, the compound **24b** was found to have a moderate affinity ( $K_d$  around 100-200 nM) to *Pu24T* and *25TAG* G-quadruplexes and represents one of rare examples of drug-like ligands of G-quadruplexes.

Regardless the fact that pull-down experiments with streptavidin-coated magnetic beads are widely used by the scientific community, this method is not optimal for DCC because of a high ratio of unspecific interactions and high price of the beads prohibitive for high-throughput analysis. Therefore, we attempted to find a substitute for this method. Firstly, we tried to reproduce the synthesis of gold-coated magnetic nanoparticles, however, in our hands, this approach did not prove reliable. After a thorough search for methods of selective pull-down we assumed that the method of solid phase extraction could serve this purpose perfectly. Therefore, we designed ten combinatorial libraries of seven compounds each and templated them in the presence of two G-quadruplex (*Pu24T* and *22CTA*) and control double-stranded (*ds26* and *hp2*) targets. The experiments of validation of hits, obtained in a few rounds of selection, are still ongoing but we can already presume that the method of pull-down by SPE is a highly promising alternative to pull-down with streptavidin-coated magnetic beads. Firstly, it does not require the modification of G4-oligonucleotides so the selection process is not affected by biotin or any other modification. Secondly, the usage of SAX cartridges with cationic surface leads to a minimization of non-specific interactions of the cartridges' resin with positively charged G4-ligands. Thirdly, the pull-down by SPE can be performed with non-modified oligonucleotides and the price of material for one experiment with SPE is five time cheaper than with streptavidin-coated magnetic beads (2.5 and 28.3 euros, respectively).

In parallel to DCC experiments, we exploited the method of parallel synthesis to generate "ready-to-screen" solutions of compounds; combined with HT screening techniques, such as FRET-melting, this method allows the simple synthesis and fast screening of combinatorial libraries of acylhydrazone compounds. By implementation of this protocol, we synthesized a combinatorial library of 90 bis(acylhydrazones) and screened all of them in a FRET-melting experiment against four G-quadruplexes of different topologies and one hairpin control. The interaction of three hits (identified by FRET-melting experiment) with G-quadruplexes of different topologies was carefully studied and one compound (**3Ei**) was found to be selective to telomeric antiparallel

22CTA G-quadruplex and to 25TAG G4 that adopts antiparallel topology in the complex with **3Ei** [OR180]. In addition, we potentially found an important drawback of isothermal fluorimetric titration method while performing experiments with these hits. We also performed the first competition experiment for study of G4 ligands binding to G-quadruplexes of different topologies by native mass spectrometry.

In a search for biological application of the synthesized compounds, we also studied the interaction of model library of ligands with G4-forming sequence of EBNA1 mRNA of Epstein-Barr virus. In collaboration with colleagues from Université de Bretagne Occidentale and Institut de Génétique Moléculaire, we studied the biological system *in vitro* and *in cellulo* and found that two of our molecules effectively enhance the translation of the highly antigenic protein EBNA1 by binding to G-quadruplexes and disrupting their complexes with nucleolin, that acts as translational regressor in this case.

In summary, two protocols of DCC and one protocol of parallel combinatorial synthesis of acylhydrazones for the search of G4 binders were developed. The method described in *Part 2* of *RESULTS AND DISCUSSION* was the first example of DCC of acylhydrazones in the G4 field. However, a high ratio of non-specific interactions was observed with streptavidin-coated magnetic beads used in this protocol. In this regard, the DCC of acylhydrazones followed by SPE represents a significant improvement of the protocol due to the low amount of non-specific binding observed in the pull-down step. However, this method requires the generation of combinatorial libraries for every target separately, therefore, it may be rather laborious. The method of parallel combinatorial synthesis, on the contrary, is universal and once the compounds are synthesized, they may be tested for their binding to different targets in parallel. In addition, it may be more advantageous when the screening is performed against more complex systems that cannot be studied by DCC (such as disruption of DNA—protein complexes).

This work opens a lot of possibilities for further studies. First of all, the structure of drug-like compound **24b** may be further optimized to enhance G4-binding properties. For example, *N*-methyl substituents of the piperazine and quinolinium rings may be extended or a brunch in the pyridine core may be added (*Figure 87*).

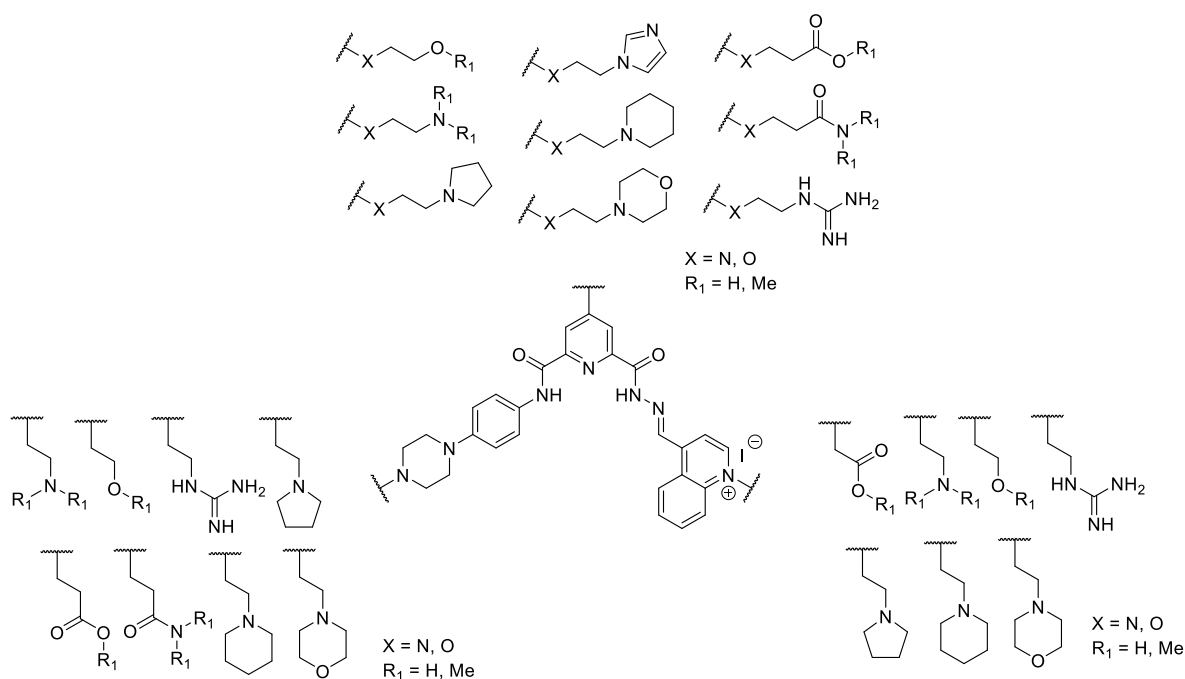


Figure 87. The proposed modifications of the structure of **24b** ligand to enhance its G4 binding properties.

Another interesting study may be started based on the premises, that compound **3Ei** (found to bind tightly to G4s of different topologies, in particular, antiparallel) provided no fluorescence quenching of Cy5 in the fluorescent titration assay. The extensive study of fluorescent response of G4-forming sequences labelled with different fluorophores by the set of structurally different but affine G4 ligands may result in the optimization of the setup of the experiment.

The methods developed in this work can be employed to produce and screen more compounds composed of diverse building blocks to enrich the scope of selective compounds specific for distinct G-quadruplex structures.

More generally, the method of DCC of acylhydrazones is a very powerful technique that can find numerous applications in the search of DNA binders. To start with, it would be interesting to perform DCC for long repeats of G4-forming sequences, as 20-25 base G4-forming oligonucleotides used in this work may not be the best models for intracellular G4s. Indeed, the comparison of set of compounds, captured by a “simple” G4 and by long telomeric sequence may give some insights about mutual organization of G4 units in long G4 clusters. DCC followed by SPE can also be suitable for the search of ligands for *i*-motifs as not many of ligands are reported



for this structure so far. Finally, upon a reasonable design, it may also be suitable for the search of compounds that fit in abasic site of double-stranded DNA.

# ***EXPERIMENTAL SECTION***

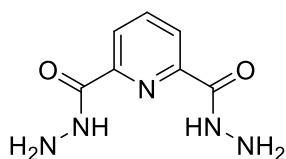
# Synthesis

**Synthesis and characterization:** All commercially available chemicals were reagent grade and used without further purification. NMR spectra were measured with a Bruker Avance 300 spectrometer ( $^1\text{H}$ : 300 MHz,  $^{13}\text{C}$ : 75 MHz) at 25 °C; chemical shifts are given in ppm ( $\delta$ ) values. Multiplicities of  $^{13}\text{C}$  NMR signals were determined from DEPT-135 experiments. The melting points were determined in open-end capillaries with a digital melting point instrument (SMP30, Stuart). Elemental microanalysis of all novel compounds was performed by the *Service de Microanalyse*, CNRS–ICSN, Gif-sur-Yvette, France and *Service Chromato-Masse-Microanalyse*, BioCIS – UMR 8076, Châtenay-Malabry. The purity of final compounds was assessed by LC/MS analysis (Waters Alliance 2695 equipped with a Waters XBridge  $\text{C}_{18}$ -3.5  $\mu\text{m}$  column and a photodiode array detector; eluent A: water with 0.05% TFA, eluent B: MeCN with 0.05% TFA, gradient elution with 2 to 100% of eluent B). Mass spectra (MS, ESI in the positive-ion mode) were recorded with a Waters ZQ instrument (cone voltage: 30 V). In the assignment of MS of salts, *M* always refers to the organic dication.

## RESULTS AND DISCUSSION. Part 1.

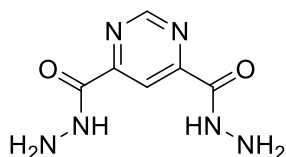
**General procedure for the synthesis of bis(acylhydrazides) (1–4):** A solution of the corresponding dimethyl ester (1.95 g, 10 mmol) and hydrazine hydrate (10.9 mL, 11.2 g, 220 mmol) in ethanol (150 mL) was heated under reflux for 18 h and then cooled to room temperature. The precipitate was filtered, washed twice with ethanol, once with ether, and dried, to give the bis(acylhydrazide) which was sufficiently pure and employed without further purification.

### Pyridine-2,6-dicarbohydrazide (L1):<sup>193</sup>



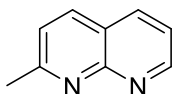
Yield 1.73 g (89%). White solid, m.p. 285–286 °C (lit. 285 °C);  $^1\text{H}$  NMR (300 MHz,  $\text{DMSO-}d_6$ ):  $\delta$  10.63 (s, 2H), 8.13 (s, 3H), 4.63 (br s, 4H);  $^{13}\text{C}$  NMR (75 MHz,  $\text{DMSO-}d_6$ ):  $\delta$  161.9 ( $\text{C}_q$ ), 148.4 ( $\text{C}_q$ ), 139.3 (CH), 123.7 (CH); MS (ESI $^+$ ):  $m/z = 196.2$  [ $M + \text{H}$ ] $^+$ .

**Pyrimidine-4,6-dicarbohydrazide (L2):**



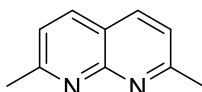
Yield 1.81 g (92%). Pale-yellow solid, m.p. (decomp.) 280 °C;  $^1\text{H}$  NMR (300 MHz,  $\text{DMSO-}d_6$ ):  $\delta$  10.41 (s, 2H), 9.35 (d,  $J = 1.1$  Hz, 1H), 8.35 (d,  $J = 1.1$  Hz, 1H), 4.77 (br s, 4H);  $^{13}\text{C}$  NMR (75 MHz,  $\text{DMSO-}d_6$ ):  $\delta$  160.1 ( $\text{C}_q$ ), 158.4 ( $\text{C}_q$ ), 157.2 (CH), 114.7 (CH); MS (ESI $^+$ ):  $m/z = 197.1$  [ $M + \text{H}$ ] $^+$ .

**2-methyl-1,8-naphthyridine (4):**



The mixture of 2-amino-3-pyridinecarboxaldehyde **3** (10.99g, 90 nmol), propan-2-one (7.95 ml, 108 mmol) and pyrrolidine (8.13 ml, 99 mmol) in ethanol (90 ml) was stirred under Ar at 55°C for 3 h and then cooled to r.t. The volatiles were removed in vacuo. Yield 12.90 g (99%). Yellow solid.  $^1\text{H}$  NMR (300 MHz,  $\text{CDCl}_3$ ):  $\delta$  9.08 (dd,  $J = 4.3, 1.9$  Hz, 1H), 8.15 (dd,  $J = 8.1, 1.9$  Hz, 1H), 7.44 (dd,  $J = 8.1, 4.3$  Hz, 1H), 7.39 (d,  $J = 8.3$  Hz, 1H), 2.82 (s, 3H).

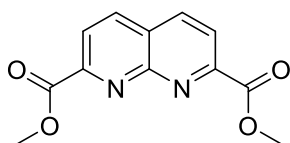
**2,7-dimethyl-1,8-naphthyridine (5):**



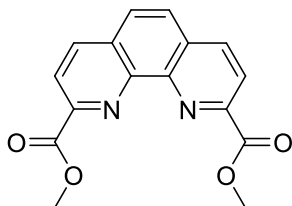
A suspension of 2-methyl-1,8-naphthyridine in anhydrous  $\text{Et}_2\text{O}$  (200 ml) was cooled to -60 °C under Ar and MeLi (75 ml, 0.12 mol) was added drop-wise within approximately 1 h such that inner

temperature didn't exceed  $-50\text{ }^{\circ}\text{C}$ . The reaction mixture was vigorously stirred for 2 h at  $-50\text{ }^{\circ}\text{C}$  in dry ice-acetone bath and then left to come to ambient temperature overnight. The reaction mixture was carefully quenched by dropwise addition of MeOH (10 ml) followed by water (50 ml). The aqueous phase was separated and extracted with Et<sub>2</sub>O (3 \* 50 ml). The combined organic phases were washed with water, brine, dried over anhydrous Na<sub>2</sub>SO<sub>4</sub>, filtered and the solvent was removed under vacuo, leaving the crude 2,7-dimethyl-1,2-dihydro-1,8-naphthiridine. It was dissolved in acetone (30 ml), cooled in an ice bath and oxidized by addition of saturated acetic solution of KMnO<sub>4</sub> until the color of the solution remained violet (approx. 300 ml). The suspension was filtered through a pad of Celite, the filter was washed with acetone and the filtrate was evaporated in vacuo. The residue was dissolved in DCM (150 ml) and washed with 1% aqueous Na<sub>2</sub>CO<sub>3</sub> (2 \* 200 ml), brine, dried over Na<sub>2</sub>SO<sub>4</sub> and finally evaporated to give 2,7-dimethyl-1,8-naphthiridine as a yellow solid. Yield: 77%. <sup>1</sup>H NMR (300 MHz, CDCl<sub>3</sub>):  $\delta$  8.02 (d,  $J$  = 8.2 Hz, 2H), 7.32 (d,  $J$  = 8.3 Hz, 2H), 2.79 (s, 6H).

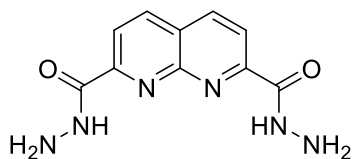
**2,7-dimethyl 1,8-naphthyridine-2,7-dicarboxylate (6):**



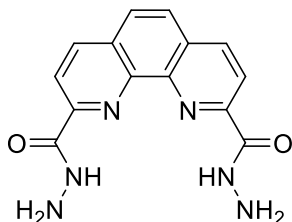
A suspension of 2,7-dimethyl-1,8-naphthyridine **5** (2.23 mg, 14.1 mmol), N-chlorosuccinimide (NCS) (15.62 g, 117 mmol) and benzoyl peroxide (450 mg) in PhCl (400 mL) was refluxed for 2 h and then cooled, filtered, and concentrated in vacuo to give 2,7-Bis(trichloromethyl)-1,8-naphthyridine as colorless crystals. Solution of these crystals in 85% H<sub>2</sub>P<sub>0</sub>4 (25 mL) was heated to 160-170  $^{\circ}\text{C}$ . After 3 h the solution was cooled to 25  $^{\circ}\text{C}$  and MeOH (150 mL) was carefully added. After 12 h of refluxing, most of the MeOH was removed in vacuo, and then CHCl<sub>3</sub>, (150 mL) and saturated aqueous Na<sub>2</sub>CO<sub>3</sub> were carefully added. The solution was stirred vigorously for 15 min, the organic layer was separated, dried over anhydrous Na<sub>2</sub>SO<sub>4</sub> and filtered. The solvent was removed in vacuo to give diester **6**. Yield: 1.77 g (54%). <sup>1</sup>H NMR (300 MHz, CDCl<sub>3</sub>):  $\delta$  8.50 – 8.26 (m, 4H), 4.07 (d,  $J$  = 5.7 Hz, 6H).

**2,9-dimethyl 1,10-phenanthroline-2,9-dicarboxylate (10):**

The solution of 1,10-phenanthroline-2,9-dicarboxylic acid (6.37 g, 23.75 mmol) in MeOH with a drop of concentrated H<sub>2</sub>SO<sub>4</sub> was refluxed overnight and then cooled to r.t. Half of the solvent was removed in vacuo and the rest was added to the concentrate solution of NaHCO<sub>3</sub>. The product was extracted with DCM, organic fraction was washed by water and dried over Na<sub>2</sub>SO<sub>4</sub>. The solvent was removed in vacuo. Yield: 5.84 g (83%). <sup>1</sup>H NMR (300 MHz, DMSO-*d*<sub>6</sub>): δ 8.78 (d, *J* = 8.3 Hz, 2H), 8.44 (d, *J* = 8.3 Hz, 2H), 8.24 (s, 2H), 4.04 (s, 6H).

**1,8-Naphthyridine-2,7-dicarbohydrazide (L3):**

Yield 2.45 g (97%). Pale-yellow solid, m.p. > 290 °C; <sup>1</sup>H NMR (300 MHz, DMSO-*d*<sub>6</sub>): δ 10.00 (s, 2H), 8.74 (d, *J* = 8.4 Hz, 2H), 8.25 (d, *J* = 8.4 Hz, 2H), 4.74 (br s, 4H); <sup>13</sup>C NMR (75 MHz, CD<sub>3</sub>CO<sub>2</sub>D): δ 164.7 (C<sub>q</sub>), 153.1 (C<sub>q</sub>), 141.1 (CH), 126.7 (C<sub>q</sub>), 122.0 (CH); MS (ESI<sup>+</sup>): *m/z* = 247.1 [*M* + H]<sup>+</sup>.

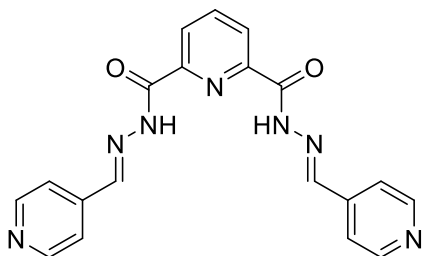
**1,10-Phenanthroline-2,9-dicarbohydrazide (L4):** <sup>194</sup>

Yield 2.38 g (80%). Pale yellow solid, m.p. > 290 °C (lit. 318–325 °C); <sup>1</sup>H NMR (300 MHz, DMSO-*d*<sub>6</sub>): δ 10.77 (s, 2H), 8.69 (d, *J* = 8.3 Hz, 2H), 8.40 (d, *J* = 8.3 Hz, 2H), 8.14 (s, 2H), 4.78 (br s, 4H); <sup>13</sup>C

NMR (75 MHz, DMSO-*d*<sub>6</sub>):  $\delta$  162.9 (C<sub>q</sub>), 149.5 (C<sub>q</sub>), 144.0 (C<sub>q</sub>), 137.9 (CH), 130.1 (C<sub>q</sub>), 127.8 (CH), 121.0 (CH); MS (ESI<sup>+</sup>):  $m/z$  = 297.2 [ $M + H$ ]<sup>+</sup>.

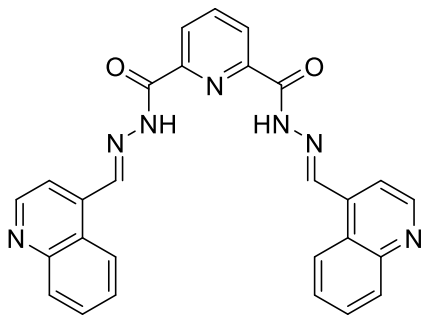
**Typical procedure for the synthesis of bis(acylhydrazone) precursors (12a–c, 13a–c, 14a–c, and 15a–c):** A solution of a bis(acylhydrazide) (**L1–L4**, 2.0 mmol) and a heteroaromatic aldehyde (**11a–11c**, 4.4 mmol) in ethanol (10 mL) was heated under reflux for 18 h. After cooling, the precipitate was collected by filtration, thoroughly washed with ethanol, and dried, to give the corresponding bis(acylhydrazone) which was sufficiently pure and employed without further purification.

**N<sup>2</sup>,N<sup>6</sup>-Bis(pyridin-4-ylmethylene)pyridine-2,6-dicarbohydrazide (12a):**



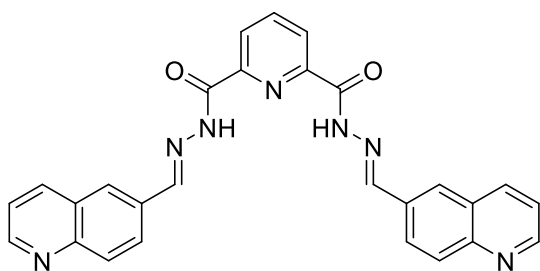
Yield 90%; white powder, m.p. > 290 °C; <sup>1</sup>H NMR (300 MHz, DMSO-*d*<sub>6</sub>):  $\delta$  12.54 (s, 2H), 8.79 (s, 2H), 8.72 (d,  $J$  = 5.9 Hz, 4H), 8.44–8.25 (m, 3H), 7.77 (d,  $J$  = 6.0 Hz, 4H); <sup>13</sup>C NMR (75 MHz, DMSO-*d*<sub>6</sub>):  $\delta$  159.8 (C<sub>q</sub>), 150.4 (CH), 148.0 (C<sub>q</sub>), 147.6 (CH), 141.3 (C<sub>q</sub>), 140.2 (CH), 125.9 (CH), 121.1 (CH); MS (ESI<sup>+</sup>):  $m/z$  = 374.2 [ $M + H$ ]<sup>+</sup>, 187.7 [ $M + 2H$ ]<sup>2+</sup>.

**N<sup>2</sup>,N<sup>6</sup>-Bis(quinolin-4-ylmethylene)pyridine-2,6-dicarbohydrazide (12b):**



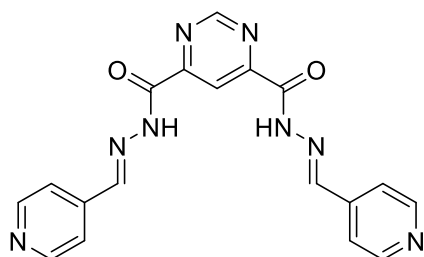
Yield 93%; pale-yellow solid, m.p. > 290 °C; <sup>1</sup>H NMR (300 MHz, DMSO-*d*<sub>6</sub>): δ 12.65 (s, 2H), 9.48 (s, 2H), 9.07 (d, *J* = 4.5 Hz, 2H), 8.91 (dd, *J* = 8.4, 0.7 Hz, 2H), 8.53–8.33 (m, 3H), 8.16 (dd, *J* = 8.3, 0.7 Hz, 2H), 7.99 (d, *J* = 4.5 Hz, 2H), 7.92–7.79 (m, 4H); <sup>13</sup>C NMR (75 MHz, DMSO-*d*<sub>6</sub>): δ 159.7 (C<sub>q</sub>), 150.5 (CH), 148.5 (C<sub>q</sub>), 148.0 (C<sub>q</sub>), 147.3 (CH), 140.3 (CH), 137.1 (C<sub>q</sub>), 129.9 (CH), 129.8 (CH), 127.7 (CH), 125.9 (CH), 124.8 (C<sub>q</sub>), 124.4 (CH), 120.0 (CH); MS (ESI<sup>+</sup>): *m/z* = 474.3 [*M* + H]<sup>+</sup>.

**N'2,N'6-Bis(quinolin-6-ylmethylene)pyridine-2,6-dicarbohydrazide (12c):**



Yield 80%; pale-yellow solid, m.p. > 290 °C; <sup>1</sup>H NMR (300 MHz, DMSO-*d*<sub>6</sub>): δ 12.50 (s, 2H), 9.01–8.94 (m, 4H), 8.52 (d, *J* = 7.8 Hz, 2H), 8.46–8.39 (m, 2H), 8.39–8.28 (m, 5H), 8.13 (d, *J* = 8.8 Hz, 2H), 7.63 (dd, *J* = 8.3, 4.3 Hz, 2H); <sup>13</sup>C NMR (75 MHz, DMSO-*d*<sub>6</sub>): δ 159.6 (C<sub>q</sub>), 151.5 (CH), 149.2 (CH), 148.6 (C<sub>q</sub>), 148.2 (C<sub>q</sub>), 140.1 (CH), 136.5 (CH), 132.3 (C<sub>q</sub>), 129.8 (CH), 129.0 (CH), 128.0 (C<sub>q</sub>), 126.4 (CH), 125.6 (CH), 122.2 (CH); MS (ESI<sup>+</sup>): *m/z* = 474.3 [*M* + H]<sup>+</sup>, 237.7 [*M* + 2H]<sup>2+</sup>.

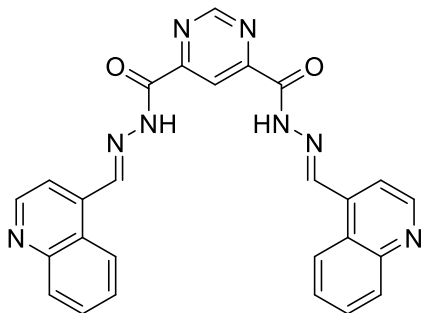
**N'4,N'6-Bis(pyridin-4-ylmethylene)pyrimidine-4,6-dicarbohydrazide (13a):**



Yield 99%; white solid, m.p. > 290 °C; <sup>1</sup>H NMR (300 MHz, DMF-*d*<sub>7</sub>): δ 12.77 (s, 2H), 9.60 (s, 1H), 8.89 (s, 2H), 8.80–8.70 (m, 5H), 7.79 (d, *J* = 6.0 Hz, 4H); <sup>13</sup>C NMR (75 MHz, DMF-*d*<sub>7</sub>): δ 159.7 (C<sub>q</sub>), 159.1 (C<sub>q</sub>), 158.0 (CH), 150.9 (CH), 148.8 (CH), 141.9 (C<sub>q</sub>), 121.6 (CH), 116.8 (CH); MS (ESI<sup>+</sup>): *m/z* = 375.2 [*M* + H]<sup>+</sup>, 188.1 [*M* + 2H]<sup>2+</sup>.

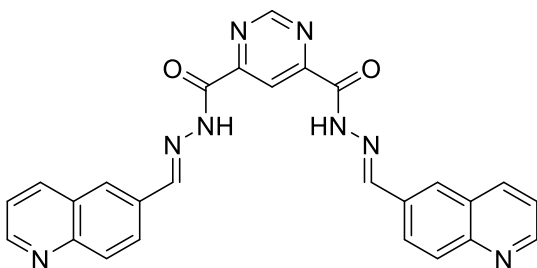


**N'4,N'6-Bis(quinolin-4-ylmethylene)pyrimidine-4,6-dicarbohydrazide (13b):**



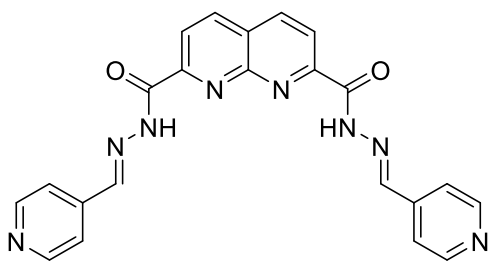
Yield 99%; pale-yellow solid, m.p. (decomp.) 282 °C; <sup>1</sup>H NMR (300 MHz, DMF-*d*<sub>7</sub>): δ 12.89 (s, 2H), 9.68 (d, *J* = 10.9 Hz, 3H), 9.10 (d, *J* = 4.4 Hz, 2H), 8.93–8.79 (m, 3H), 8.18 (d, *J* = 8.2 Hz, 2H), 8.02 (s, 2H), 7.96–7.86 (m, 2H), 7.86–7.76 (m, 2H); <sup>13</sup>C NMR (75 MHz, DMF-*d*<sub>7</sub>): δ 160.0 (C<sub>q</sub>), 157.7 (CH), 150.4 (CH), 149.5 (C<sub>q</sub>), 148.4 (CH), 137.9 (C<sub>q</sub>), 130.3 (CH), 129.5 (CH), 127.6 (CH), 125.6 (C<sub>q</sub>), 124.6 (CH), 123.4 (C<sub>q</sub>), 120.3 (CH), 116.8 (CH); MS (ESI<sup>+</sup>): *m/z* = 475.3 [*M* + H]<sup>+</sup>, 238.2 [*M* + 2H]<sup>2+</sup>.

**N'4,N'6-Bis(quinolin-6-ylmethylene)pyrimidine-4,6-dicarbohydrazide (13c):**



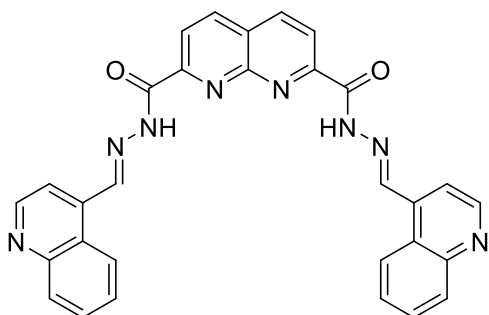
Yield 99%; pale-yellow solid, m.p. > 290 °C; <sup>1</sup>H NMR (300 MHz, DMSO-*d*<sub>6</sub>): δ 12.68 (s, 2H), 9.60 (s, 1H), 9.03–8.87 (m, 4H), 8.65 (s, 1H), 8.51 (d, *J* = 8.0 Hz, 2H), 8.32–8.20 (m, 4H), 8.10 (d, *J* = 8.4 Hz, 2H), 7.66–7.55 (m, 2H).; <sup>13</sup>C NMR (75 MHz, CD<sub>3</sub>CO<sub>2</sub>D): δ 160.2 (C<sub>q</sub>), 158.6 (C<sub>q</sub>), 158.0 (CH), 150.5 (CH), 147.1 (CH), 145.2 (CH), 141.5 (C<sub>q</sub>), 134.7 (C<sub>q</sub>), 131.0 (CH), 130.6 (CH), 129.0 (C<sub>q</sub>), 123.7 (CH), 123.0 (CH), 116.9 (CH); MS (ESI<sup>+</sup>): *m/z* = 475.3 [*M* + H]<sup>+</sup>, 238.2 [*M* + 2H]<sup>2+</sup>.

**N'2,N'7-Bis(pyridin-4-ylmethylene)-1,8-naphthyridine-2,7-dicarbohydrazide (14a):**



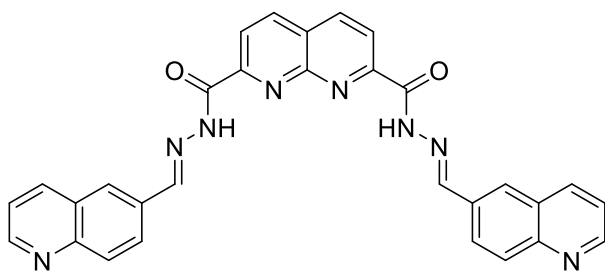
Yield 78%; white solid, m.p. > 290 °C;  $^1\text{H}$  NMR (300 MHz,  $\text{DMSO-}d_6$ ):  $\delta$  12.63 (s, 2H), 8.89 (d,  $J$  = 8.3 Hz, 2H), 8.81–8.58 (m, 6H), 8.42 (d,  $J$  = 8.3 Hz, 2H), 7.71 (d,  $J$  = 4.6 Hz, 4H);  $^{13}\text{C}$  NMR (75 MHz,  $\text{CD}_3\text{CO}_2\text{D}$ ):  $\delta$  162.7 ( $\text{C}_q$ ), 153.5 ( $\text{C}_q$ ), 149.3 ( $\text{C}_q$ ), 147.1 (CH), 145.3 (CH), 141.5 (CH), 127.7 ( $\text{C}_q$ ), 124.8 (CH), 123.1 (CH); MS (ESI $^+$ ):  $m/z$  = 425.3 [ $M + \text{H}$ ] $^+$ .

**N',N'7-Bis(quinolin-4-ylmethylene)-1,8-naphthyridine-2,7-dicarbohydrazide (14b):**



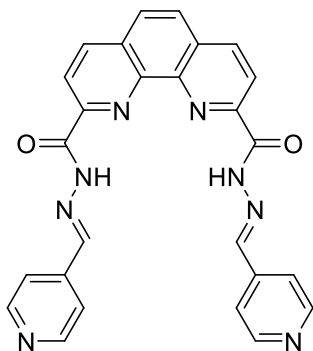
Yield 80%; white solid, m.p. > 290 °C;  $^1\text{H}$  NMR (300 MHz,  $\text{DMSO-}d_6$ ):  $\delta$  12.72 (s, 2H), 9.51 (s, 2H), 9.05 (d,  $J$  = 4.5 Hz, 2H), 8.93 (d,  $J$  = 8.3 Hz, 2H), 8.78 (d,  $J$  = 8.3 Hz, 2H), 8.48 (d,  $J$  = 8.3 Hz, 2H), 8.14 (d,  $J$  = 8.2 Hz, 2H), 7.95 (d,  $J$  = 4.4 Hz, 2H), 7.91–7.84 (m, 2H), 7.84–7.77 (m, 2H);  $^1\text{H}$  NMR (300 MHz,  $\text{CD}_3\text{CO}_2\text{D}/\text{D}_2\text{O}$  1:1 v/v):  $\delta$  9.36 (s, 2H), 9.02 (d,  $J$  = 5.5 Hz, 2H), 8.67–8.55 (m, 4H), 8.36 (dd,  $J$  = 13.8, 7.0 Hz, 4H), 8.11 (d,  $J$  = 8.4 Hz, 2H), 7.99–7.90 (m, 2H), 7.90–7.80 (m, 2H);  $^{13}\text{C}$  NMR (75 MHz,  $\text{CD}_3\text{CO}_2\text{D}/\text{D}_2\text{O}$  1:1 v/v):  $\delta$  162.0 ( $\text{C}_q$ ), 152.7 ( $\text{C}_q$ ), 145.6 (CH), 144.5 (CH), 141.2 (CH), 141.0 ( $\text{C}_q$ ), 134.5 (CH), 130.8 (CH), 127.1 ( $\text{C}_q$ ), 126.4 ( $\text{C}_q$ ), 124.9 (CH), 123.6 (CH), 122.6 (CH), 119.7 (CH); MS (ESI $^+$ ):  $m/z$  = 525.2 [ $M + \text{H}$ ] $^+$ , 263.2 [ $M + 2\text{H}$ ] $^{2+}$ .

**N',N'7-Bis(quinolin-6-ylmethylene)-1,8-naphthyridine-2,7-dicarbohydrazide (14c):**



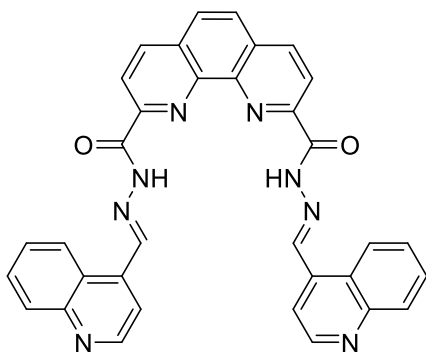
Yield 80%; white solid, m.p. > 290 °C;  $^1\text{H}$  NMR (300 MHz,  $\text{DMSO-}d_6$ ):  $\delta$  12.56 (s, 2H), 9.07–8.74 (m, 6H), 8.52 (d,  $J = 8.3$  Hz, 2H), 8.44 (d,  $J = 8.2$  Hz, 2H), 8.33–8.25 (m, 4H), 8.12 (d,  $J = 8.9$  Hz, 2H), 7.62 (dd,  $J = 7.6, 4.0$  Hz, 2H);  $^{13}\text{C}$  NMR spectrum could not be obtained due to insufficient solubility; MS (ESI<sup>+</sup>):  $m/z = 263.1 [M + 2\text{H}]^{2+}$ .

**N',N'-Bis(pyridin-4-ylmethylene)-1,10-phenanthroline-2,9-dicarbohydrazide (15a):**



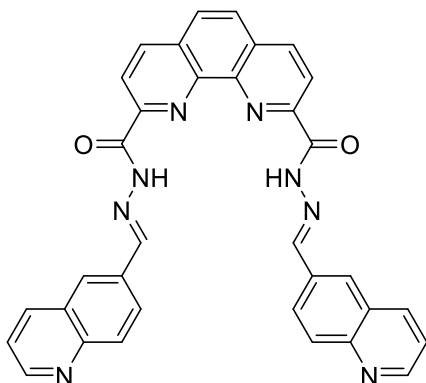
Yield 77%; pale yellow solid, m.p. > 290 °C;  $^1\text{H}$  NMR (300 MHz,  $\text{DMF-}d_7$ ):  $\delta$  12.87 (s, 2H), 9.11 (s, 2H), 8.92 (d,  $J = 8.3$  Hz, 2H), 8.81 (d,  $J = 5.8$  Hz, 4H), 8.67 (d,  $J = 8.3$  Hz, 2H), 8.34 (s, 2H), 7.85 (d,  $J = 5.8$  Hz, 4H); NMR (75 MHz,  $\text{DMF-}d_7$ ):  $\delta$  161.1 ( $\text{C}_q$ ), 150.9 (CH), 149.8 ( $\text{C}_q$ ), 147.9 (CH), 144.6 ( $\text{C}_q$ ), 142.4 ( $\text{C}_q$ ), 139.0 (CH), 131.5 ( $\text{C}_q$ ), 128.8 (CH), 122.1 (CH), 121.6 (CH); MS (ESI<sup>+</sup>):  $m/z = 475.3 [M + \text{H}]^+$ .

**N',N'-Bis(quinolin-4-ylmethylene)-1,10-phenanthroline-2,9-dicarbohydrazide (15b):**



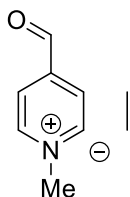
Yield 55%; pale yellow solid, m.p. > 290 °C;  $^1\text{H}$  NMR (300 MHz, DMSO- $d_6$ ):  $\delta$  12.97 (s, 2H), 9.62 (s, 2H), 8.86–8.83 (m, 4H), 8.65 (d,  $J$  = 8.3 Hz, 2H), 8.61 (d,  $J$  = 8.2 Hz, 2H), 8.26 (s, 2H), 7.99 (d,  $J$  = 8.3 Hz, 2H), 7.81 (d,  $J$  = 4.5 Hz, 2H), 7.61–7.53 (m, 2H), 7.10–7.02 (m, 2H);  $^{13}\text{C}$  NMR (75 MHz, DMSO- $d_6$ ):  $\delta$  160.4 (C<sub>q</sub>), 150.1 (CH), 149.0 (C<sub>q</sub>), 148.1 (C<sub>q</sub>), 147.3 (CH), 138.8 (CH), 138.1 (CH), 137.8 (C<sub>q</sub>), 130.9 (C<sub>q</sub>), 129.4 (CH), 129.4 (CH), 128.6 (CH), 126.9 (CH), 124.9 (CH), 124.2 (C<sub>q</sub>), 122.0 (CH), 119.6 (CH); MS (ESI<sup>+</sup>):  $m/z$  = 575.3 [ $M$  + H]<sup>+</sup>, 288.2 [ $M$  + 2H]<sup>2+</sup>.

**N',N'-Bis(quinolin-6-ylmethylene)-1,10-phenanthroline-2,9-dicarbohydrazide (15c):**



Yield 59%; pale yellow solid, m.p. > 290 °C;  $^1\text{H}$  NMR (300 MHz, DMF- $d_7$ ):  $\delta$  12.88 (s, 2H), 9.37 (s, 2H), 9.01 (dd,  $J$  = 4.2, 1.5 Hz, 2H), 8.93 (d,  $J$  = 8.3 Hz, 2H), 8.69 (d,  $J$  = 8.3 Hz, 2H), 8.48 (dd,  $J$  = 8.8, 1.6 Hz, 2H), 8.34 (d,  $J$  = 7.4 Hz, 4H), 8.19 (d,  $J$  = 8.8 Hz, 2H), 8.11 (d,  $J$  = 7.9 Hz, 2H), 7.53 (dd,  $J$  = 8.3, 4.2 Hz, 2H);  $^{13}\text{C}$  NMR (75 MHz, DMF- $d_7$ ):  $\delta$  160.8 (C<sub>q</sub>), 151.8 (CH), 150.0 (C<sub>q</sub>), 149.5 (C<sub>q</sub>), 149.5 (CH), 144.5 (C<sub>q</sub>), 138.9 (CH), 136.6 (CH), 133.6 (C<sub>q</sub>), 131.4 (C<sub>q</sub>), 130.4 (CH), 129.3 (CH), 128.7 (CH), 128.6 (C<sub>q</sub>), 127.0 (CH), 122.5 (CH), 122.0 (CH); MS (ESI<sup>+</sup>):  $m/z$  = 575.3 [ $M$  + H]<sup>+</sup>, 288.2 [ $M$  + 2H]<sup>2+</sup>.

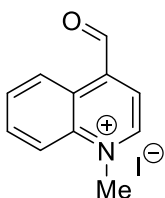
#### 4-Formyl-1-methylpyridinium iodide (A1):<sup>195</sup>



The mixture of pyridine-4-carboxaldehyde (5.4 mL, 6.10 g, 56.9 mmol) and methyl iodide (7.0 mL, 15.95 g, 112.4 mmol) in DCM (20 mL) was stirred for 72 h at room temperature under argon atmosphere. The precipitated solid was filtered and washed twice with DCM, to give **A1** (13.3 g, 94%) as an orange solid. <sup>1</sup>H NMR (300 MHz, D<sub>2</sub>O):  $\delta$  8.80 (d,  $J$  = 6.5 Hz, 2H), 8.12 (d,  $J$  = 6.2 Hz, 2H), 6.21 (s, 1H), 4.39 (s, 3H); <sup>13</sup>C NMR (75 MHz, D<sub>2</sub>O):  $\delta$  160.1 (C<sub>q</sub>), 146.0 (CH), 125.7 (CH), 88.2 (CH), 48.5 (CH<sub>3</sub>); MS (ESI<sup>+</sup>):  $m/z$  (%) = 140.1 (100) [ $M$  + H<sub>2</sub>O]<sup>+</sup>.

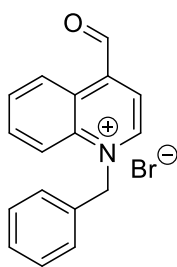
**General procedure for the synthesis of quaternized heterocyclic aldehydes (A2-A5):** The solution of aldehyde (10 mmol) and alkylating agent (100 mmol) in acetone (18 mL) was stirred at 60 °C for 18 h and then cooled to room temperature. The precipitated solid was filtered, washed twice with acetone, once with ether, and dried.

#### 4-Formyl-1-methylquinolinium iodide (A2):



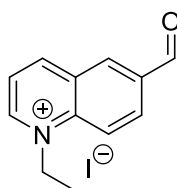
Obtained from quinoline-4-aldehyde and methyl iodide in a 77% yield. Red solid, <sup>1</sup>H NMR (300 MHz, D<sub>2</sub>O):  $\delta$  9.25 (d,  $J$  = 6.1 Hz, 1H), 8.63 (d,  $J$  = 8.4 Hz, 1H), 8.42 (d,  $J$  = 9.0 Hz, 1H), 8.24 (m, 2H), 8.04 (m, 1H), 6.87 (s, 1H), 4.66 (s, 3H); <sup>13</sup>C NMR (75 MHz, D<sub>2</sub>O):  $\delta$  157.8 (C<sub>q</sub>), 150.1 (CH), 139.5 (C<sub>q</sub>), 135.9 (CH), 130.8 (CH), 127.4 (C<sub>q</sub>), 127.1 (CH), 119.5 (CH), 118.5 (CH), 86.8 (CH), 46.3 (CH<sub>3</sub>); MS (ESI<sup>+</sup>):  $m/z$  (%) = 190.3 (100) [ $M$  + H<sub>2</sub>O]<sup>+</sup>, 172.1 (25) [ $M$ ]<sup>+</sup>.

#### 4-Formyl-1-benzylquinolinium bromide (A3):



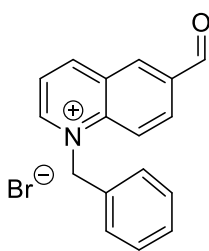
Obtained from quinoline-4-aldehyde and benzyl bromide in an 81% yield.  $^1\text{H}$  NMR (300 MHz,  $\text{D}_2\text{O}$ ):  $\delta$  9.39 (d,  $J = 6.1$  Hz, 1H), 8.63 (d,  $J = 7.9$  Hz, 1H), 8.45–8.25 (m, 2H), 8.15–8.04 (m, 1H), 7.97 (m, 1H), 7.42 (s, 3H), 7.32 (s, 2H), 6.89 (s, 1H), 6.28 (s, 2H);  $^{13}\text{C}$  NMR (75 MHz,  $\text{D}_2\text{O}$ ):  $\delta$  158.7 ( $\text{C}_q$ ), 149.9 (CH), 139.0 ( $\text{C}_q$ ), 136.1 (CH), 133.4 ( $\text{C}_q$ ), 130.8 (CH), 130.1 (CH), 129.9 (CH), 128.1 ( $\text{C}_q$ ), 128.0 (CH), 127.4 (CH), 120.0 (CH), 118.7 (CH), 86.9 (CH), 61.8 ( $\text{CH}_2$ ); MS (ESI $^+$ ):  $m/z$  (%) = 266.2 (100) [ $M + \text{H}_2\text{O}$ ] $^+$ , 248.2 (6) [ $M$ ] $^+$ .

#### 6-Formyl-1-ethylquinolinium iodide (A4):



Obtained from quinoline-6-aldehyde and ethyl iodide in a 73% yield.  $^1\text{H}$  NMR (300 MHz,  $\text{DMSO}-d_6$ ):  $\delta$  10.31 (s, 1H), 9.68 (d,  $J = 4.7$  Hz, 1H), 9.48 (d,  $J = 8.3$  Hz, 1H), 9.08 (d,  $J = 1.7$  Hz, 1H), 8.79 (d,  $J = 9.2$  Hz, 1H), 8.60 (dd,  $J = 9.2, 1.7$  Hz, 1H), 8.31 (dd,  $J = 8.4, 5.8$  Hz, 1H), 5.13 (q,  $J = 7.2$  Hz, 2H), 1.62 (t,  $J = 7.2$  Hz, 3H);  $^{13}\text{C}$  NMR (75 MHz,  $\text{DMSO}-d_6$ ):  $\delta$  191.9 ( $\text{C}_q$ ), 151.5 (CH), 148.7 (CH), 139.7 ( $\text{C}_q$ ), 135.6 ( $\text{C}_q$ ), 134.4 (CH), 132.5 (CH), 129.6 ( $\text{C}_q$ ), 123.5 (CH), 120.3 (CH), 53.5 ( $\text{CH}_2$ ), 15.2 ( $\text{CH}_3$ ); MS (ESI $^+$ ):  $m/z$  (%) = 204.2 (8) [ $M + \text{H}_2\text{O}$ ] $^+$ , 186.1 (100) [ $M$ ] $^+$ .

#### 6-Formyl-1-benzylquinolinium iodide (A5):

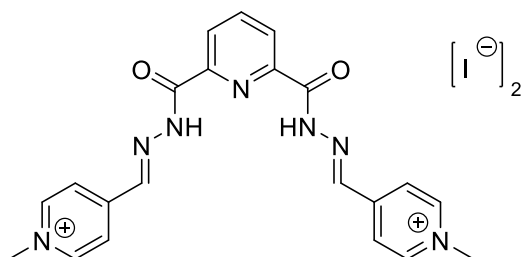


Obtained from quinoline-6-aldehyde and benzyl bromide in an 85% yield.  $^1\text{H}$  NMR (300 MHz,  $\text{DMSO-}d_6$ ):  $\delta$  10.26 (s, 1H), 9.85 (d,  $J = 5.8$  Hz, 1H), 9.58 (d,  $J = 8.3$  Hz, 1H), 9.10 (d,  $J = 1.5$  Hz, 1H), 8.67 (d,  $J = 9.2$  Hz, 1H), 8.54 (dd,  $J = 9.2, 1.7$  Hz, 1H), 8.41 (dd,  $J = 8.4, 5.8$  Hz, 1H), 7.40 (s, 5H), 6.42 (s, 2H);  $^{13}\text{C}$  NMR (75 MHz,  $\text{DMSO-}d_6$ ):  $\delta$  191.7 ( $\text{C}_q$ ), 152.5 (CH), 149.6 (CH), 139.9 ( $\text{C}_q$ ), 135.7 ( $\text{C}_q$ ), 134.3 (CH), 133.6 ( $\text{C}_q$ ), 132.8 (CH), 129.9 (CH), 129.1 ( $\text{C}_q$ ), 128.8 (CH), 127.4 (CH), 123.7 (CH), 120.6 ( $\text{C}_q$ ), 60.2 ( $\text{CH}_2$ ); MS (ESI $^+$ ):  $m/z$  (%) = 266.3 (10) [ $M + \text{H}_2\text{O}$ ] $^+$ , 248.2 (100) [ $M$ ] $^+$ .

**Synthesis of cationic bis(acylhydrazones) (PyDH1, PyDH2). Method A:** A mixture of bis(acylhydrazone) precursor **12a** or **12b** (0.5 mmol), alkyl halogenide (75 mmol) and DMF (3 mL) was sealed in a tube and heated at 40 °C during 18 h. After cooling to room temperature, the precipitate was filtered and washed with DMF and diethyl ether, and dried in vacuum. The crude product was purified by recrystallization from boiling MeCN–H<sub>2</sub>O. *Note: this method gave sufficiently pure products PyDH1 and PyDH2; however, in the case PymDH1, PymDH2, and NaphDH2, the mono-alkylated by-product could not be removed by recrystallization. Method B (below) was therefore preferred for the synthesis of all bis(acylhydrazones).*

**N<sup>2</sup>,N<sup>6</sup>-Bis[(1-methylpyridinium-4-yl)methylene]pyridine-2,6-dicarbohydrazone iodide**

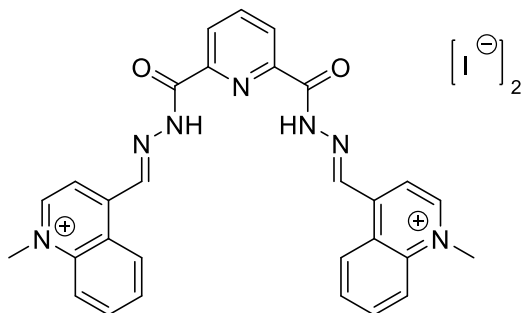
**(PyDH1):**



Yield 279 mg (85%). Orange powder;  $^1\text{H}$  NMR (300 MHz,  $\text{DMSO-}d_6$ ):  $\delta$  12.93 (s, 2H), 9.03 (d,  $J = 5.9$  Hz, 4H), 8.94 (s, 2H), 8.54–8.31 (m, 7H), 4.37 (s, 6H);  $^{13}\text{C}$  NMR (75 MHz,  $\text{DMSO-}d_6$ ):  $\delta$  160.3

(C<sub>q</sub>), 148.8 (C<sub>q</sub>), 147.6 (C<sub>q</sub>), 146.1 (CH), 143.8 (CH), 140.6 (CH), 126.6 (CH), 124.3 (CH), 47.7 (CH<sub>3</sub>); MS (ESI<sup>+</sup>): *m/z* (%) = 402.3 (100) [*M* – H]<sup>+</sup>, 201.7 (41) [*M*]<sup>2+</sup>; purity (LC) 100%.

**N'2,N'6-Bis[(1-methylquinolinium-4-yl)methylene]pyridine-2,6-dicarbohydrazide iodide (PyDH2):**



Yield 360 mg (95%). Orange powder; <sup>1</sup>H NMR (300 MHz, DMSO-*d*<sub>6</sub>): δ 13.03 (s, 2H), 9.74 (s, 2H), 9.54 (d, *J* = 6.3 Hz, 2H), 9.07 (d, *J* = 8.5 Hz, 2H), 8.66–8.56 (m, 4H), 8.52 (d, *J* = 7.0 Hz, 2H), 8.48–8.43 (m, 1H), 8.42–8.33 (m, 2H), 8.27–8.18 (m, 2H), 4.68 (s, 6H); <sup>13</sup>C NMR (75 MHz, DMSO-*d*<sub>6</sub>): δ 160.1 (C<sub>q</sub>), 149.4 (CH), 147.6 (C<sub>q</sub>), 147.0 (C<sub>q</sub>), 143.4 (CH), 140.7 (CH), 139.1 (C<sub>q</sub>), 135.3 (CH), 130.5 (CH), 126.7 (CH), 126.3 (C<sub>q</sub>), 126.0 (CH), 120.0 (CH), 119.0 (CH), 45.7 (CH<sub>3</sub>); MS (ESI<sup>+</sup>): *m/z* (%) = 502.3 (100) [*M* – H]<sup>+</sup>, 616.3 (9) [*M* + CF<sub>3</sub>COO]<sup>+</sup>; purity (LC) 90%.

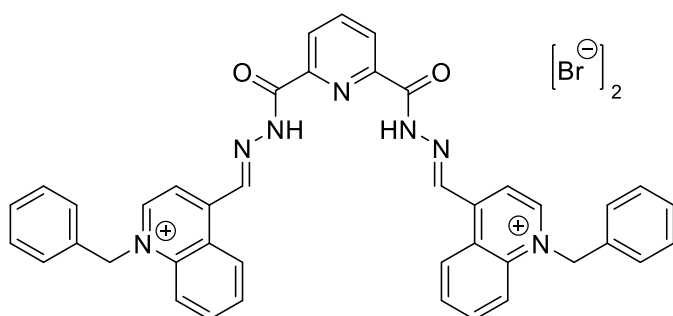
**Synthesis of cationic bis(acylhydrazones). Method B:** The mixture of dicarbohydrazide **L1–L4** (0.5 mmol) and quaternized aldehyde **A1–A5** (1.1 mmol) in DMF (3 mL) was heated at 100 °C (80 °C for NaphDH5 and PhenDH5) for 2 h and then cooled to room temperature. The precipitate was collected by filtration, washed three times with MeCN, once with ether, dried and then additionally recrystallized from MeCN/H<sub>2</sub>O. The yields are indicated for <sup>1</sup>H-NMR spectroscopically pure material prior to the final recrystallization step.

**N'2,N'6-Bis[(1-methylpyridinium-4-yl)methylene]pyridine-2,6-dicarbohydrazide iodide (PyDH1):** Yield 90%. The appearance and spectroscopic properties were identical with those described above; purity (LC): 100%; anal. calcd. for C<sub>21</sub>H<sub>21</sub>I<sub>2</sub>N<sub>7</sub>O<sub>2</sub> × 2 H<sub>2</sub>O (693.3): C 36.38, H 3.63, N 14.14; found: C 36.03, H 3.31, N 14.05.



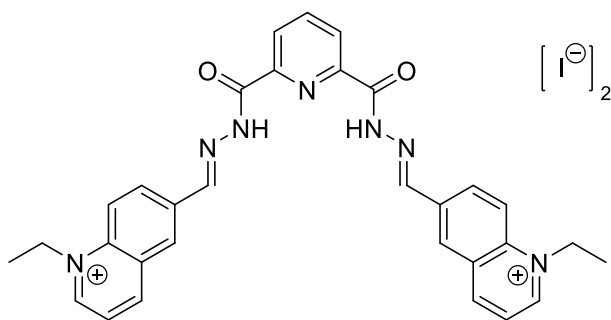
**N<sup>2</sup>,N<sup>6</sup>-Bis[(1-methylquinolinium-4-yl)methylene]pyridine-2,6-dicarbohydrazide iodide (PyDH2):** Yield 76%. The appearance and spectroscopic properties were identical with those described above; purity (LC): 100%; anal. calcd. for C<sub>29</sub>H<sub>25</sub>I<sub>2</sub>N<sub>7</sub>O<sub>2</sub> × 3.5 H<sub>2</sub>O (820.4): C 42.46, H 3.93, N 11.95; found: C 42.46, H 3.59, N 11.60.

**N<sup>2</sup>,N<sup>6</sup>-Bis[(1-benzylquinolinium-4-yl)methylene]pyridine-2,6-dicarbohydrazide bromide (PyDH3):**



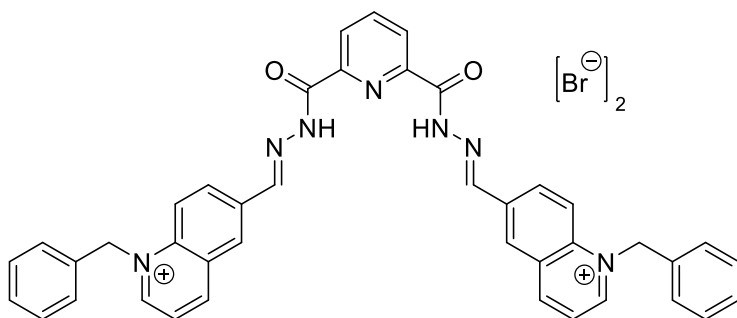
Yield 93%. Yellow-orange powder; <sup>1</sup>H NMR (300 MHz, DMSO-*d*<sub>6</sub>): δ 13.34 (s, 2H), 10.29 (s, 2H), 9.74 (d, *J* = 6.2 Hz, 2H), 9.06 (d, *J* = 8.5 Hz, 2H), 8.66 (d, *J* = 6.2 Hz, 2H), 8.62–8.50 (m, 4H), 8.49–8.38 (m, 1H), 8.31–8.23 (m, 2H), 8.19–8.11 (m, 2H), 7.41 (s, 10H), 6.42 (s, 4H); <sup>13</sup>C NMR (75 MHz, DMSO-*d*<sub>6</sub>): δ 160.0 (C<sub>q</sub>), 149.4 (CH), 148.3 (C<sub>q</sub>), 147.8 (C<sub>q</sub>), 143.1 (CH), 140.4 (CH), 138.2 (C<sub>q</sub>), 135.5 (CH), 134.0 (C<sub>q</sub>), 130.2 (C<sub>q</sub>), 129.2 (CH), 128.9 (CH), 127.4 (CH), 127.2 (C<sub>q</sub>), 127.0 (CH), 126.3 (CH), 120.0 (CH), 118.4 (CH), 59.9 (CH<sub>2</sub>); MS (ESI<sup>+</sup>): *m/z* (%) = 654.4 (34) [*M* – H]<sup>+</sup>, 327.8 (100) [*M*]<sup>2+</sup>; purity (LC): 100%; anal. calcd. for C<sub>41</sub>H<sub>33</sub>Br<sub>2</sub>N<sub>8</sub>O<sub>2</sub> × 1.5 H<sub>2</sub>O (842.6): C 58.44, H 4.31, N 11.64; found: C 58.28, H 4.28, N 11.62.

**N<sup>2</sup>,N<sup>6</sup>-Bis[(1-ethylquinolinium-6-yl)methylene]pyridine-2,6-dicarbohydrazide iodide (PyDH4):**



Yield 92%. Pale brown solid;  $^1\text{H}$  NMR (300 MHz,  $\text{DMSO-}d_6$ ):  $\delta$  12.67 (s, 2H), 9.59 (d,  $J = 5.4$  Hz, 2H), 9.38 (d,  $J = 8.3$  Hz, 2H), 9.07 (s, 2H), 8.85 (s, 2H), 8.74 (q,  $J = 9.6$  Hz, 4H), 8.45 (d,  $J = 7.1$  Hz, 2H), 8.41–8.32 (m, 1H), 8.32–8.16 (m, 2H), 5.14 (q,  $J = 7.0$  Hz, 4H), 1.66 (t,  $J = 7.0$  Hz, 6H);  $^{13}\text{C}$  NMR (75 MHz,  $\text{DMSO-}d_6$ ):  $\delta$  160.0 ( $\text{C}_q$ ), 149.7 (CH), 148.0 ( $\text{C}_q$ ), 147.4 (CH), 147.1 (CH), 140.3 (CH), 138.1 ( $\text{C}_q$ ), 135.2 (CH), 132.3 (CH), 130.1 ( $\text{C}_q$ ), 129.9 (CH), 126.0 (CH), 123.2 (CH), 120.0 (CH), 53.3 ( $\text{CH}_2$ ), 15.3 ( $\text{CH}_3$ ); MS (ESI $^+$ ):  $m/z$  (%) = 644.3 (32) [ $M + \text{CF}_3\text{COO}$ ] $^+$ , 265.8 (100) [ $M$ ] $^{2+}$ ; purity (LC): 100%; anal. calcd. for  $\text{C}_{31}\text{H}_{29}\text{I}_2\text{N}_7\text{O}_2 \times 1.5 \text{H}_2\text{O}$  (821.5): C 45.33, H 4.05, N 11.94; found: C 45.44, H 4.11, N 11.96.

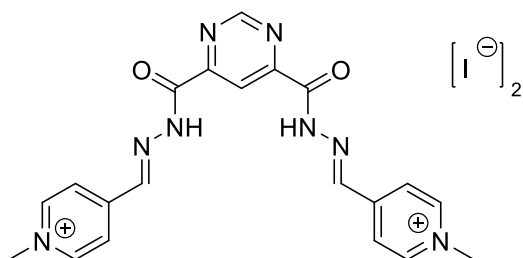
**$\text{N}^2, \text{N}^6$ -Bis[(1-benzylquinolinium-6-yl)methylene]pyridine-2,6-dicarbohydrazide bromide (PyDH5):**



Yield 55%. Pale yellow solid;  $^1\text{H}$  NMR (300 MHz,  $\text{DMSO-}d_6$ ):  $\delta$  12.72 (s, 2H), 9.76 (d,  $J = 5.0$  Hz, 2H), 9.48 (d,  $J = 8.3$  Hz, 2H), 9.12 (s, 2H), 8.85 (s, 2H), 8.73–8.58 (m, 4H), 8.46–8.30 (m, 5H), 7.49–7.34 (m, 10H), 6.41 (s, 4H);  $^{13}\text{C}$  NMR (75 MHz,  $\text{DMSO-}d_6$ ):  $\delta$  159.9 ( $\text{C}_q$ ), 150.6 (CH), 148.3 (CH), 148.1 ( $\text{C}_q$ ), 146.9 (CH), 140.2 (CH), 138.3 ( $\text{C}_q$ ), 135.4 ( $\text{C}_q$ ), 133.8 ( $\text{C}_q$ ), 132.6 (CH), 130.4 ( $\text{C}_q$ ), 129.7 (CH), 129.1 (CH), 128.9 (CH), 127.5 (CH), 126.1 (CH), 123.3 (CH), 120.3 (CH), 60.0 ( $\text{CH}_2$ ); MS (ESI $^+$ ):  $m/z$

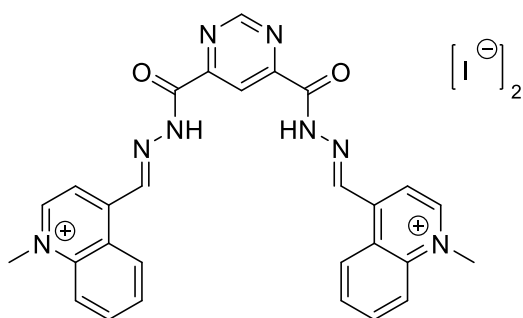
(%) = 768.3 (59)  $[M + CF_3COO]^+$ , 327.8 (100)  $[M]^{2+}$ ; purity (LC): 99%; anal. calcd. for  $C_{41}H_{33}Br_2N_7O_2 \times 1.5 H_2O$  (842.6): C 58.44, H 4.31, N 11.64; found: C 58.74, H 4.42, N 11.56.

**$N^{4},N^{6}$ -Bis[(1-methylpyridinium-4-yl)methylene]pyrimidine-4,6-dicarbohydrazide iodide  
(PymDH1):**



Yield 97%. Yellow-orange solid;  $^1H$  NMR (300 MHz,  $DMSO-d_6$ ):  $\delta$  13.31 (s, 2H), 9.67 (s, 1H), 9.01 (d,  $J = 6.5$  Hz, 4H), 8.86 (s, 2H), 8.64 (s, 1H), 8.38 (d,  $J = 6.4$  Hz, 4H), 4.35 (s, 6H);  $^{13}C$  NMR (75 MHz,  $DMSO-d_6$ ):  $\delta$  159.9 ( $C_q$ ), 158.2 ( $C_q$ ), 157.7 (CH), 148.6 ( $C_q$ ), 146.0 (CH), 144.9 (CH), 124.5 (CH), 119.3 (CH), 47.7 ( $CH_3$ ); MS (ESI $^+$ ):  $m/z$  (%) = 403.3 (35)  $[M - H]^+$ , 202.2 (100)  $[M]^{2+}$ ; purity (LC): 100%; anal. calcd. for  $C_{20}H_{20}I_2N_8O_2 \times H_2O$  (676.3): C 35.52, H 3.28, N 16.57; found: C 35.13, H 3.33, N 16.29.

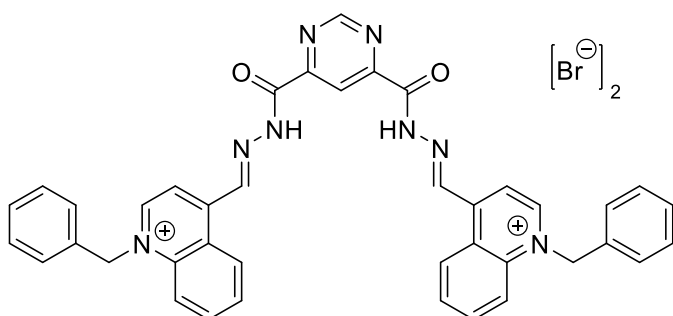
**$N^{4},N^{6}$ -Bis[(1-methylquinolinium-4-yl)methylene]pyrimidine-4,6-dicarbohydrazide iodide  
(PymDH2):**



Yield 92%. Orange solid;  $^1H$  NMR (300 MHz,  $DMSO-d_6$ ):  $\delta$  13.28 (s, 2H), 9.75 (s, 1H), 9.70 (s, 2H), 9.51 (d,  $J = 6.2$  Hz, 2H), 8.86 (d,  $J = 8.4$  Hz, 2H), 8.73 (s, 1H), 8.60 (d,  $J = 8.7$  Hz, 2H), 8.52 (d,  $J = 6.1$  Hz, 2H), 8.40–8.32 (m, 2H), 8.25–8.16 (m, 2H), 4.67 (s, 6H);  $^{13}C$  NMR (75 MHz,  $DMSO-d_6$ ):  $\delta$  159.6

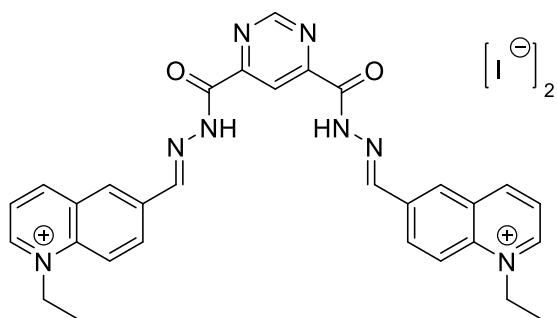
(C<sub>q</sub>), 158.0 (C<sub>q</sub>), 157.8 (CH), 149.4 (CH), 146.7 (C<sub>q</sub>), 144.3 (CH), 139.0 (C<sub>q</sub>), 135.3 (CH), 130.6 (CH), 126.3 (C<sub>q</sub>), 125.6 (CH), 120.0 (CH), 118.8 (CH), 117.0 (CH), 45.7 (CH<sub>3</sub>); MS (ESI<sup>+</sup>):  $m/z$  (%) = 503.3 (12) [ $M - H$ ]<sup>+</sup>, 252.3 (100) [ $M$ ]<sup>2+</sup>; purity (LC): 100%; anal. calcd. for C<sub>28</sub>H<sub>24</sub>I<sub>2</sub>N<sub>8</sub>O<sub>2</sub> × 2.2 H<sub>2</sub>O (798.0): C 42.18, H 3.58, N 14.05; found: C 42.55, H 3.39, N 13.66.

**N<sup>2</sup>,N<sup>6</sup>-Bis[(1-benzylquinolinium-4-yl)methylene]pyridine-2,6-dicarbohydrazide bromide (PymDH3):**



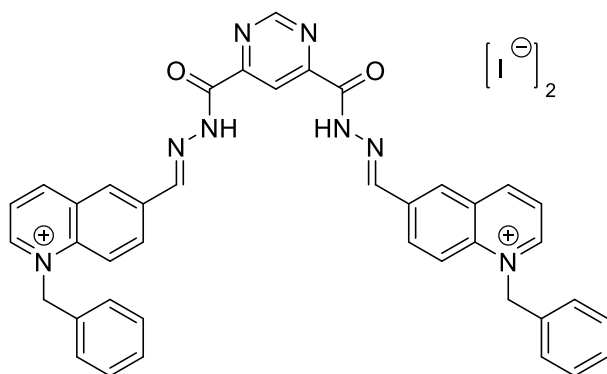
Yield 93%. Brown crystals. <sup>1</sup>H NMR (300 MHz, DMSO-*d*<sub>6</sub>): δ 13.31 (s, 2H), 9.74 (d,  $J = 5.3$  Hz, 5H), 8.85 (d,  $J = 8.7$  Hz, 2H), 8.74 (s, 1H), 8.60 (dd,  $J = 17.3, 7.6$  Hz, 4H), 8.33–8.22 (m, 2H), 8.20–8.12 (m, 2H), 7.40 (s, 10H), 6.42 (s, 4H); <sup>13</sup>C NMR (75 MHz, DMSO-*d*<sub>6</sub>): δ 159.7 (C<sub>q</sub>), 158.1 (C<sub>q</sub>), 157.8 (CH), 149.6 (CH), 147.9 (C<sub>q</sub>), 144.3 (CH), 138.2 (C<sub>q</sub>), 135.5 (CH), 134.0 (C<sub>q</sub>), 130.5 (CH), 129.1 (CH), 128.8 (CH), 127.2 (CH), 127.2 (C<sub>q</sub>), 126.1 (CH), 120.1 (CH), 119.2 (CH), 117.1 (CH), 60.0 (CH<sub>2</sub>); MS (ESI<sup>+</sup>):  $m/z$  (%) = 655.5 (12) [ $M - H$ ]<sup>+</sup>, 328.4 (100) [ $M$ ]<sup>2+</sup>; purity (LC): 100%; anal. calcd. for C<sub>40</sub>H<sub>32</sub>Br<sub>2</sub>N<sub>8</sub>O<sub>2</sub> × H<sub>2</sub>O (834.6): C 57.57, H 4.11, N 13.43; found: C 57.43, H 4.25, N 13.62.

**N<sup>2</sup>,N<sup>6</sup>-Bis[(1-ethylquinolinium-6-yl)methylene]pyridine-2,6-dicarbohydrazide iodide (PymDH4):**



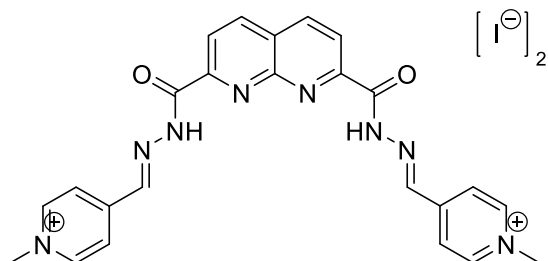
Yield 96%. Orange solid.  $^1\text{H}$  NMR (300 MHz,  $\text{DMSO-}d_6$ ):  $\delta$  12.93 (s, 2H), 9.65 (s, 1H), 9.57 (d,  $J$  = 5.7 Hz, 2H), 9.37 (d,  $J$  = 8.3 Hz, 2H), 9.00 (s, 2H), 8.80–8.61 (m, 7H), 8.25 (dd,  $J$  = 8.3, 5.9 Hz, 2H), 5.12 (q,  $J$  = 7.2 Hz, 4H), 1.64 (t,  $J$  = 7.2 Hz, 6H);  $^{13}\text{C}$  NMR (75 MHz,  $\text{DMSO-}d_6$ ):  $\delta$  159.4 ( $\text{C}_q$ ), 158.6 (CH), 149.7 (CH), 148.0 (CH), 147.5 (CH), 138.2 ( $\text{C}_q$ ), 138.1 ( $\text{C}_q$ ), 135.1 ( $\text{C}_q$ ), 132.3 (CH), 130.1 (CH), 130.1 ( $\text{C}_q$ ), 123.2 (CH), 119.9 (CH), 116.8 (CH), 53.2 ( $\text{CH}_2$ ), 15.3 ( $\text{CH}_3$ ); MS (ESI $^+$ ):  $m/z$  (%) = 266.3 (100) [ $M$ ] $^{2+}$ ; purity (LC): 100%; anal. calcd. for  $\text{C}_{30}\text{H}_{28}\text{I}_2\text{N}_8\text{O}_2 \times 3.2 \text{H}_2\text{O}$  (844.1): C 42.69, H 4.11, N 13.28; found: C 42.34, H 3.84, N 13.25.

**$N'^2, N'^6$ -Bis[(1-benzylquinolinium-6-yl)methylene]pyridine-2,6-dicarbohydrazide bromide (PymDH5):**



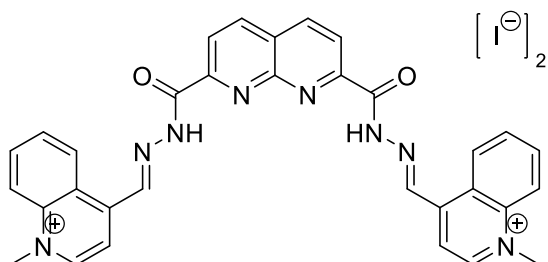
Yield 97%. Pale-yellow solid.  $^1\text{H}$  NMR (300 MHz,  $\text{DMSO-}d_6$ ):  $\delta$  12.91 (s, 2H), 9.74 (d,  $J$  = 5.4 Hz, 2H), 9.62 (d,  $J$  = 1.2 Hz, 1H), 9.46 (d,  $J$  = 8.3 Hz, 2H), 8.96 (s, 2H), 8.80 (s, 2H), 8.67 – 8.55 (m, 5H), 8.34 (dd,  $J$  = 8.3, 5.9 Hz, 2H), 7.47–7.37 (m, 10H), 6.39 (s, 4H);  $^{13}\text{C}$  NMR spectrum could not be obtained due to insufficient solubility; MS (ESI $^+$ ):  $m/z$  (%) = 769.5 (7) [ $M + \text{CF}_3\text{COO}$ ] $^+$ , 328.3 (100) [ $M$ ] $^{2+}$ ; purity (LC): 99%; anal. calcd. for  $\text{C}_{40}\text{H}_{32}\text{Br}_2\text{N}_8\text{O}_2 \times 2 \text{H}_2\text{O}$  (852.6): C 56.35, H 4.26, N 13.14; found: C 56.09, H 4.20, N 13.21.

**N<sup>2</sup>,N<sup>7</sup>-Bis[(1-methylpyridinium-4-yl)methylene]-1,8-naphthyridine-2,7-dicarbohydrazide iodide (NaphDH1):**



Yield 76%. Yellow solid; <sup>1</sup>H NMR (300 MHz, DMSO-*d*<sub>6</sub>): δ 13.14 (s, 2H), 9.01 (d, *J* = 6.0 Hz, 4H), 8.94 (d, *J* = 8.3 Hz, 2H), 8.83 (s, 2H), 8.45 (d, *J* = 8.3 Hz, 2H), 8.39 (d, *J* = 5.9 Hz, 4H), 4.36 (s, 6H); <sup>13</sup>C NMR (75 MHz, DMSO-*d*<sub>6</sub>): δ 161.6 (C<sub>q</sub>), 153.5 (C<sub>q</sub>), 152.5 (C<sub>q</sub>), 148.9 (C<sub>q</sub>), 146.0 (CH), 143.7 (CH), 140.4 (CH), 139.1 (C<sub>q</sub>), 124.4 (CH), 122.1 (CH), 47.7 (CH<sub>3</sub>); MS (ESI<sup>+</sup>): *m/z* (%) = 453.3 (78) [*M* – H]<sup>+</sup>, 227.3 (92) [*M*]<sup>2+</sup>; purity (LC): 100%; anal. calcd. for C<sub>24</sub>H<sub>22</sub>I<sub>2</sub>N<sub>8</sub>O<sub>2</sub> × 0.5 H<sub>2</sub>O (717.3): C 40.19, H 3.23, N 15.62; found: C 40.29, H 3.47, N 15.56.

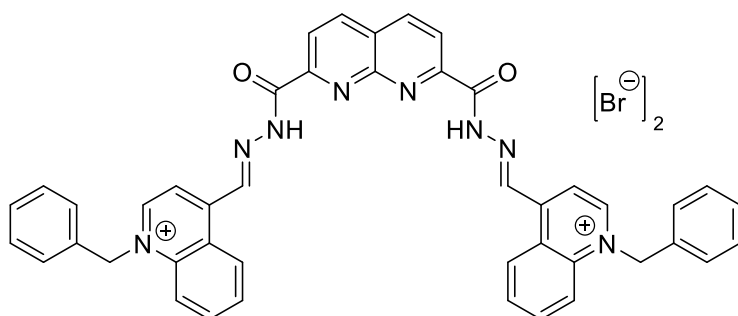
**N<sup>2</sup>,N<sup>7</sup>-Bis[(1-methylquinolinium-4-yl)methylene]-1,8-naphthyridine-2,7-dicarbohydrazide iodide (NaphDH2):**



This compound was obtained in an analytically pure form through a reaction of **3** with **11a** performed in DMSO instead of DMF. Yield 81%. Red solid; <sup>1</sup>H NMR (300 MHz, DMSO-*d*<sub>6</sub>): δ 13.14 (s, 2H), 9.72 (s, 2H), 9.52 (d, *J* = 6.2 Hz, 2H), 8.99 (d, *J* = 8.3 Hz, 2H), 8.90 (d, *J* = 8.4 Hz, 2H), 8.68–8.47 (m, 6H), 8.39–8.32 (m, 2H), 8.25–8.15 (m, 2H), 4.67 (s, 6H); <sup>13</sup>C NMR spectrum could not be obtained due to insufficient solubility; MS (ESI<sup>+</sup>): *m/z* (%) = 553.2 (63) [*M* – H]<sup>+</sup>, 277.3 (100) [*M*]<sup>2+</sup>;

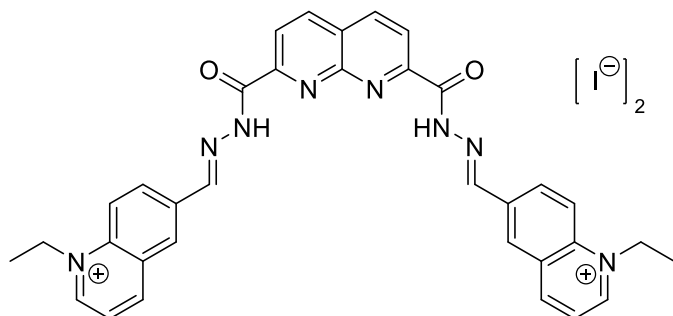
purity (LC): 100%; anal. calcd. for  $C_{32}H_{26}I_2N_8O_2 \times 0.6$  DMSO (855.3): C 46.62, H 3.49, N 13.10; found: C 46.77, H 3.23, N 13.45.

**$N^{2},N^{7}$ -Bis[(1-benzylquinolinium-4-yl)methylene]-1,8-naphthyridine-2,7-dicarbohydrazide bromide (NaphDH3):**



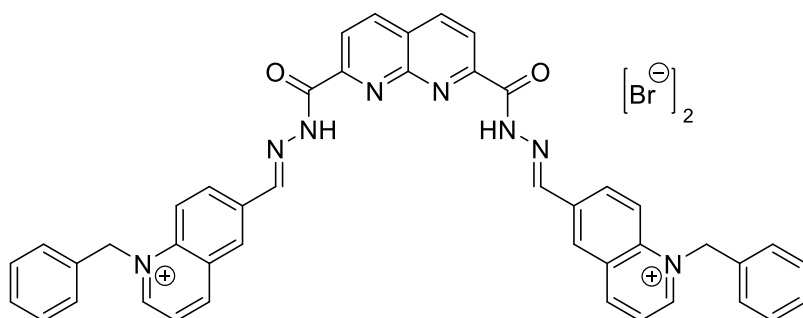
Yield 90%. Yellow solid;  $^1H$  NMR (300 MHz, DMSO- $d_6$ ):  $\delta$  13.19 (s, 2H), 9.77 (s, 4H), 9.00 (d,  $J$  = 8.3 Hz, 2H), 8.90 (d,  $J$  = 8.1 Hz, 2H), 8.66 (d,  $J$  = 6.0 Hz, 2H), 8.56 (dd,  $J$  = 16.8, 8.7 Hz, 4H), 8.32–8.22 (m, 2H), 8.21–8.12 (m, 2H), 7.41 (s, 10H), 6.44 (s, 4H);  $^{13}C$  NMR (75 MHz, DMSO- $d_6$ ):  $\delta$  161.3 (C<sub>q</sub>), 153.2 (C<sub>q</sub>), 152.4 (C<sub>q</sub>), 149.6 (CH), 148.2 (C<sub>q</sub>), 143.1 (CH), 140.6 (CH), 138.2 (C<sub>q</sub>), 135.5 (CH), 134.1 (C<sub>q</sub>), 130.5 (CH), 129.1 (CH), 128.8 (CH), 127.3 (CH), 127.1 (C<sub>q</sub>), 126.1 (CH), 122.0 (CH), 120.2 (CH), 119.0 (CH), 60.0 (CH<sub>2</sub>); MS (ESI<sup>+</sup>):  $m/z$  (%) = 705.4 (50) [ $M - H$ ]<sup>+</sup>, 353.4 (100) [ $M$ ]<sup>2+</sup>; purity (LC): 97%; anal. calcd. for  $C_{44}H_{34}Br_2N_8O_2 \times 3$  H<sub>2</sub>O (920.7): C 57.40, H 4.38, N 12.17; found: C 57.28, H 4.21, N 12.11.

**$N^{2},N^{7}$ -Bis[(1-ethylquinolinium-6-yl)methylene]-1,8-naphthyridine-2,7-dicarbohydrazide iodide (NaphDH4):**



Yield 68%. Pale brown solid;  $^1\text{H}$  NMR (300 MHz,  $\text{DMSO-}d_6$ ):  $\delta$  12.79 (s, 2H), 9.58 (d,  $J = 5.9$  Hz, 2H), 9.38 (d,  $J = 8.6$  Hz, 2H), 8.99 (s, 2H), 8.93 (d,  $J = 8.4$  Hz, 2H), 8.72 (dd,  $J = 20.4, 11.1$  Hz, 6H), 8.45 (d,  $J = 8.3$  Hz, 2H), 8.26 (dd,  $J = 8.3, 5.8$  Hz, 2H), 5.13 (q,  $J = 7.2$  Hz, 4H), 1.65 (t,  $J = 7.1$  Hz, 6H);  $^{13}\text{C}$  NMR (75 MHz,  $\text{DMSO-}d_6$ ):  $\delta$  161.1 ( $\text{C}_q$ ), 153.8 ( $\text{C}_q$ ), 152.5 ( $\text{C}_q$ ), 149.6 (CH), 148.5 ( $\text{C}_q$ ), 147.4 (CH), 146.8 (CH), 140.3 (CH), 138.0 ( $\text{C}_q$ ), 135.3 ( $\text{C}_q$ ), 132.4 (CH), 130.1 ( $\text{C}_q$ ), 129.9 (CH), 123.2 (CH), 121.8 (CH), 119.9 (CH), 53.3 ( $\text{CH}_2$ ), 15.3 ( $\text{CH}_3$ ); MS (ESI $^+$ ):  $m/z$  (%) = 695.4 (15) [ $M + \text{CF}_3\text{COO}$ ] $^+$ , 291.3 (100) [ $M$ ] $^{2+}$ ; purity (LC): 100%; anal. calcd. for  $\text{C}_{34}\text{H}_{30}\text{I}_2\text{N}_8\text{O}_2 \times 1.5 \text{H}_2\text{O}$  (863.5): C 47.29, H 3.85, N 12.98; found: C 47.44, H 4.11, N 12.95.

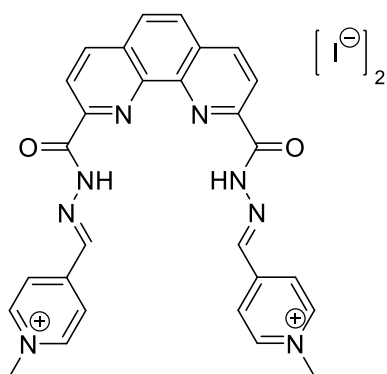
**$\text{N}^{2,7}$ -Bis[(1-benzylquinolinium-6-yl)methylene]-1,8-naphthyridine-2,7-dicarbohydrazide bromide (NaphDH5):**



Yield 80%. Pale brown solid;  $^1\text{H}$  NMR (300 MHz,  $\text{DMSO-}d_6$ ):  $\delta$  12.76 (s, 2H), 9.75 (d,  $J = 5.6$  Hz, 2H), 9.48 (d,  $J = 8.3$  Hz, 2H), 8.98–8.85 (m, 4H), 8.80 (s, 2H), 8.63 (m, 4H), 8.43 (d,  $J = 8.3$  Hz, 2H), 8.35 (dd,  $J = 8.2, 5.9$  Hz, 2H), 7.41 (d,  $J = 6.6$  Hz, 10H), 6.40 (s, 4H);  $^{13}\text{C}$  NMR (75 MHz,  $\text{DMSO-}d_6$ ):  $\delta$  161.1 ( $\text{C}_q$ ), 153.8 ( $\text{C}_q$ ), 152.6 ( $\text{C}_q$ ), 150.6 (CH), 148.3 (CH), 146.7 (CH), 140.3 (CH), 138.3 ( $\text{C}_q$ ), 135.5 ( $\text{C}_q$ ), 133.8 ( $\text{C}_q$ ), 132.7 (CH), 130.6 ( $\text{C}_q$ ), 129.8 (CH), 129.2 ( $\text{C}_q$ ), 128.9 (CH), 127.5 (CH), 123.3 (CH), 121.8 (CH), 120.3 (CH), 60.1 ( $\text{CH}_2$ ); MS (ESI $^+$ ):  $m/z$  (%) = 705.4 (5) [ $M + \text{CF}_3\text{COO}$ ] $^+$ , 353.2 (100) [ $M$ ] $^{2+}$ ; purity (LC): 100%; anal. calcd. for  $\text{C}_{44}\text{H}_{34}\text{Br}_2\text{N}_8\text{O}_2 \times \text{H}_2\text{O}$  (884.6): C 59.74, H 4.10, N 12.67; found: C 59.55, H 4.44, N 12.56.

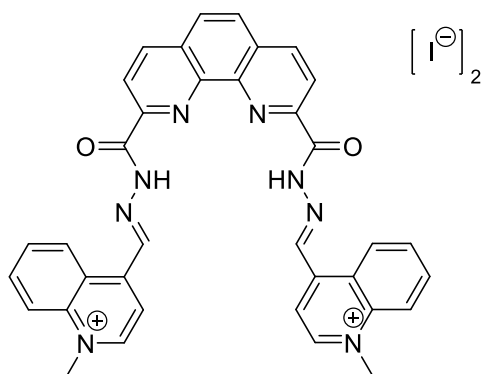
**$\text{N}^{2,9}$ -Bis[(1-methylpyridinium-4-yl)methylene]-1,10-phenanthroline-2,9-dicarbohydrazide iodide (PhenDH1):**





Yield 50%. Yellow solid;  $^1\text{H}$  NMR (300 MHz,  $\text{DMSO-}d_6$ ):  $\delta$  13.38 (s, 2H), 9.04 (s, 6H), 8.89 (d,  $J = 8.1$  Hz, 2H), 8.64 (d,  $J = 8.0$  Hz, 2H), 8.41 (d,  $J = 5.9$  Hz, 4H), 8.32 (s, 2H), 4.41 (s, 6H),  $^{13}\text{C}$  NMR (75 MHz,  $\text{DMSO-}d_6$ ):  $\delta$  161.1 ( $\text{C}_q$ ), 149.3 ( $\text{C}_q$ ), 148.6 ( $\text{C}_q$ ), 146.0 (CH), 144.3 (CH), 143.7 ( $\text{C}_q$ ), 138.9 (CH), 131.1 ( $\text{C}_q$ ), 128.8 (CH), 124.4 (CH), 122.2 (CH), 47.7 ( $\text{CH}_3$ ); MS (ESI $^+$ ):  $m/z$  (%) = 617.3 (20) [ $M + \text{CF}_3\text{COO}$ ] $^+$ , 503.3 (90) [ $M - \text{H}$ ] $^+$ , 252.2 (100%) [ $M$ ] $^{2+}$ ; purity (LC): 100%; anal. calcd. for  $\text{C}_{28}\text{H}_{24}\text{I}_2\text{N}_8\text{O}_2 \times 4.6 \text{H}_2\text{O}$  (841.2): C 39.98, H 3.98, N 13.32; found: C 40.36, H 4.20, N 12.95.

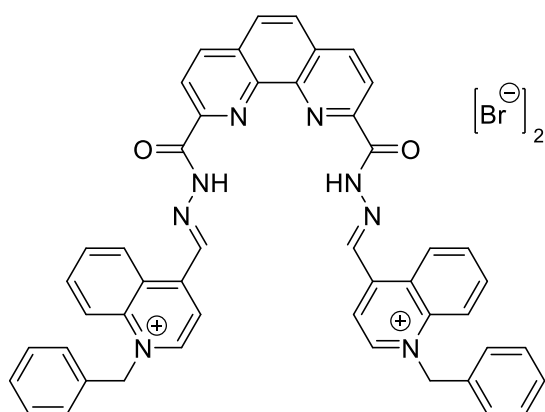
**$\text{N}^2, \text{N}^9$ -Bis[(1-methylquinolinium-4-yl)methylene]-1,10-phenanthroline-2,9-dicarbohydrazide iodide (PhenDH2):**



Yield 78%. Dark-red solid;  $^1\text{H}$  NMR (300 MHz,  $\text{DMSO-}d_6$ ):  $\delta$  13.28 (s, 2H), 9.86 (s, 2H), 9.45 (d,  $J = 6.3$  Hz, 2H), 8.95 (d,  $J = 8.3$  Hz, 2H), 8.71 (d,  $J = 8.3$  Hz, 2H), 8.56 (d,  $J = 6.1$  Hz, 2H), 8.48–8.34 (m, 4H), 8.27 (d,  $J = 9.1$  Hz, 2H), 7.88–7.80 (m, 2H), 7.15–7.08 (m, 2H), 4.62 (s, 6H);  $^{13}\text{C}$  NMR (75 MHz,  $\text{DMSO-}d_6$ ):  $\delta$  160.6 ( $\text{C}_q$ ), 149.3 ( $\text{C}_q$ ), 148.3 ( $\text{C}_q$ ), 147.3 ( $\text{C}_q$ ), 143.6 ( $\text{C}_q$ ), 143.5 (CH), 139.1 (CH), 138.2 ( $\text{C}_q$ ), 134.6 (CH), 131.2 ( $\text{C}_q$ ), 129.2 (CH), 128.8 (CH), 126.1 ( $\text{C}_q$ ), 125.1 (CH), 122.1 (CH), 119.2 (CH), 118.0 (CH), 45.5 ( $\text{CH}_3$ ); MS (ESI $^+$ ):  $m/z$  (%) = 717.4 (14) [ $M + \text{CF}_3\text{COO}$ ] $^+$ , 603.4 (83) [ $M - \text{H}$ ] $^+$ , 302.3

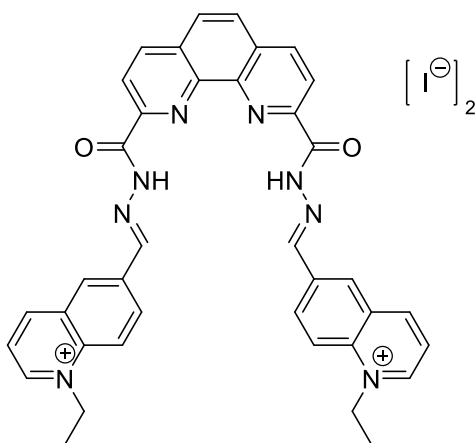
(100)  $[M]^{2+}$ ; purity (LC): 100%; anal. calcd. for  $C_{36}H_{28}I_2N_8O_2 \times 5.7 H_2O$  (961.2): C 44.99, H 4.13, N 11.66; found: C 45.23, H 4.52, N 11.63.

**$N^{2},N^{9}$ -Bis[(1-benzylquinolinium-4-yl)methylene]-1,10-phenanthroline-2,9-dicarbohydrazide bromide (PhenDH3):**



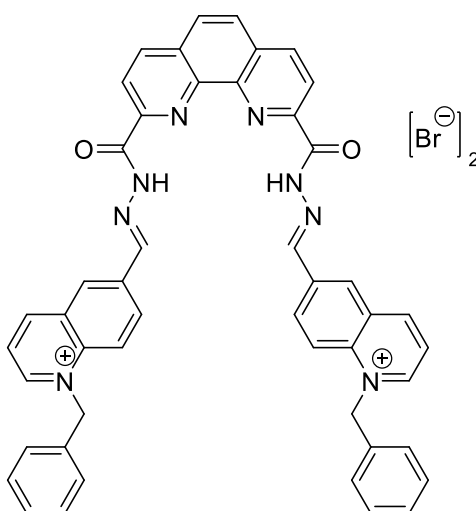
Yield 87%. Pale brown solid;  $^1H$  NMR (300 MHz, DMSO- $d_6$ ):  $\delta$  13.34 (s, 2H), 9.94 (s, 2H), 9.76 (d,  $J$  = 6.0 Hz, 2H), 8.96 (d,  $J$  = 8.2 Hz, 2H), 8.73 (d,  $J$  = 7.7 Hz, 4H), 8.37 (d,  $J$  = 7.9 Hz, 4H), 8.14 (d,  $J$  = 9.2 Hz, 2H), 7.65–7.30 (m, 10H), 7.09–7.00 (m, 2H), 6.84–6.79 (m, 2H), 6.35 (s, 4H);  $^{13}C$  NMR (75 MHz, DMSO- $d_6$ ):  $\delta$  160.7 (C<sub>q</sub>), 149.6 (CH), 148.4 (C<sub>q</sub>), 148.2 (C<sub>q</sub>), 143.7 (C<sub>q</sub>), 143.2 (CH), 139.1 (CH), 137.5 (C<sub>q</sub>), 134.3 (CH), 134.2 (C<sub>q</sub>), 131.2 (C<sub>q</sub>), 129.2 (CH), 128.9 (CH), 128.8 (CH), 127.3 (C<sub>q</sub>), 126.9 (C<sub>q</sub>), 125.7 (CH), 122.2 (CH), 119.4 (CH), 118.3 (CH), 118.3 (CH), 59.8 (CH<sub>2</sub>); MS (ESI<sup>+</sup>):  $m/z$  (%) = 755.4 (20)  $[M - H]^+$ , 378.4 (100)  $[M]^{2+}$ ; purity (LC): 100%; anal. calcd. for  $C_{48}H_{36}Br_2N_8O_2 \times 3.5 H_2O$  (979.7): C 58.85, H 4.42, N 11.44; found: C 58.84, H 4.30, N 11.47.

**$N^{2},N^{9}$ -Bis[(1-ethylquinolinium-6-yl)methylene]-1,10-phenanthroline-2,9-dicarbohydrazide iodide (PhenDH4):**



Yield 72%. Pale brown solid;  $^1\text{H}$  NMR (300 MHz,  $\text{DMSO-}d_6$ ):  $\delta$  13.10 (s, 2H), 9.60 (d,  $J = 5.6$  Hz, 2H), 9.25 (s, 2H), 9.17 (d,  $J = 8.2$  Hz, 2H), 8.89 (d,  $J = 8.3$  Hz, 2H), 8.76 (d,  $J = 2.6$  Hz, 6H), 8.64 (d,  $J = 8.3$  Hz, 2H), 8.32 (s, 2H), 8.22 (dd,  $J = 8.4, 5.9$  Hz, 2H), 5.16 (q,  $J = 7.0$  Hz, 4H), 1.67 (t,  $J = 7.1$  Hz, 6H);  $^{13}\text{C}$  NMR (75 MHz,  $\text{DMSO-}d_6$ ):  $\delta$  160.7 ( $\text{C}_q$ ), 149.6 (CH), 149.0 ( $\text{C}_q$ ), 147.3 ( $\text{C}_q$ ), 147.2 (CH), 143.7 ( $\text{C}_q$ ), 138.7 (CH), 138.0 ( $\text{C}_q$ ), 135.7 ( $\text{C}_q$ ), 132.8 (CH), 130.9 ( $\text{C}_q$ ), 130.1 ( $\text{C}_q$ ), 129.3 (CH), 128.5 (CH), 123.2 (CH), 122.0 (CH), 120.0 (CH), 53.3 ( $\text{CH}_2$ ), 15.4 ( $\text{CH}_3$ ); MS (ESI $^+$ ):  $m/z$  (%) = 745.4 (28) [ $M + \text{CF}_3\text{COO}$ ] $^+$ , 316.2 (100) [ $M$ ] $^{2+}$ ; purity (LC): 98%; anal. calcd. for  $\text{C}_{38}\text{H}_{32}\text{I}_2\text{N}_8\text{O}_2 \times 1.5 \text{H}_2\text{O}$  (904.5): C 50.46, H 3.79, N 12.39; found: C 50.13, H 3.90, N 12.41.

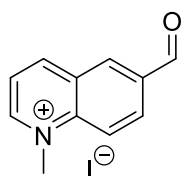
**$\text{N}^{2,9}$ -Bis[(1-benzylquinolinium-6-yl)methylene]-1,10-phenanthroline-2,9-dicarbohydrazide bromide (PhenDH5):**



Yield 91%. Pale brown solid;  $^1\text{H}$  NMR (300 MHz,  $\text{DMSO-}d_6$ ):  $\delta$  13.10 (s, 2H), 9.79 (d,  $J = 5.7$  Hz, 2H), 9.30 (d,  $J = 8.3$  Hz, 2H), 9.21 (s, 2H), 8.87 (d,  $J = 8.3$  Hz, 2H), 8.81 (s, 2H), 8.76–8.65 (m, 4H), 8.63 (d,  $J = 8.3$  Hz, 2H), 8.39–8.28 (m, 4H), 7.53–7.37 (m, 10H), 6.46 (s, 4H);  $^{13}\text{C}$  NMR (75 MHz,  $\text{DMSO-}d_6$ ):  $\delta$  160.7 ( $\text{C}_q$ ), 150.5 (CH), 149.0 ( $\text{C}_q$ ), 148.1 (CH), 147.0 (CH), 143.8 ( $\text{C}_q$ ), 138.7 (CH), 138.3 ( $\text{C}_q$ ), 135.9 ( $\text{C}_q$ ), 133.8 ( $\text{C}_q$ ), 133.0 (CH), 130.9 ( $\text{C}_q$ ), 130.4 ( $\text{C}_q$ ), 129.5 (CH), 129.2 (CH), 128.9 (CH), 128.6 (CH), 127.6 (CH), 123.4 (CH), 122.0 (CH), 120.3 (CH), 60.1 ( $\text{CH}_2$ ); MS (ESI $^+$ ):  $m/z$  (%) = 755.4 (5) [ $M - \text{H}$ ] $^+$ , 378.4 (100) [ $M$ ] $^{2+}$ ; purity (LC): 100%; anal. calcd. for  $\text{C}_{48}\text{H}_{36}\text{Br}_2\text{N}_8\text{O}_2 \times 2 \text{H}_2\text{O}$  (952.7): C 60.51, H 4.23, N 11.76; found: C 60.76, H 4.39, N 11.77.

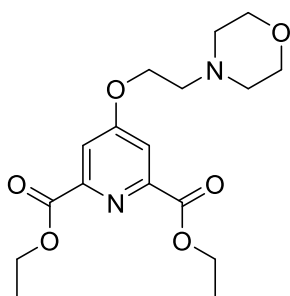
## RESULTS AND DISCUSSION. Part 2.

### 6-Formyl-1-methylquinolinium iodide (A6):



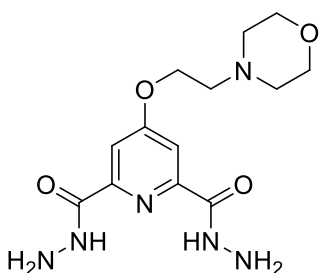
A solution of quinoline-6-aldehyde **11c** (0.39 g, 2.50 mmol) and iodomethane (1.56 mL, 3.55 g, 25.0 mmol) in acetone (25 mL) was stirred at 60 °C for 18 h and then cooled to room temperature. The precipitated solid was filtered, washed twice with acetone, once with ether, and dried, to give **A6** (89%) as a red solid.  $^1\text{H}$  NMR (300 MHz,  $\text{DMSO-}d_6$ ):  $\delta$  10.31 (s, 1H), 9.64 (d,  $J = 5.4$  Hz, 1H), 9.48 (d,  $J = 8.4$  Hz, 1H), 9.08 (d,  $J = 1.4$  Hz, 1H), 8.69 (d,  $J = 9.1$  Hz, 1H), 8.61 (dd,  $J = 9.1, 1.7$  Hz, 1H), 8.30 (dd,  $J = 8.4, 5.8$  Hz, 1H), 4.68 (s, 3H);  $^{13}\text{C}$  NMR (75 MHz,  $\text{DMSO-}d_6$ ):  $\delta$  191.9 ( $\text{C}_q$ ), 152.3 (CH), 148.5 (CH), 140.7 ( $\text{C}_q$ ), 135.7 ( $\text{C}_q$ ), 134.0 (CH), 132.4 (CH), 129.1 ( $\text{C}_q$ ), 123.2 (CH), 120.6 (CH), 45.8 ( $\text{CH}_3$ ); MS (ESI $^+$ ):  $m/z$  (%) = 190.2 (11) [ $M + \text{H}_2\text{O}$ ] $^+$ , 172.2 (100) [ $M$ ] $^+$ .

### Diethyl 4-(2-morpholinoethoxy)pyridine-2,6-dicarboxylate (17):



Diethyl chelidamate **6** (2.00 g, 8.36 mmol) was dissolved in dry acetonitrile (130 mL), and 4-(2-chloroethyl)morpholine hydrochloride (2.02 g, 10.9 mmol) and  $K_2CO_3$  (3.47 g, 25.1 mmol) were added. The resulting mixture was maintained under reflux for 18 h and then cooled, filtered, and evaporated to dryness. The residue was partitioned between DCM (50 mL) and  $H_2O$  (50 mL), and the aqueous layer was extracted with DCM ( $3 \times 50$  mL). The combined organic phases were dried over  $Na_2SO_4$ , filtered through  $SiO_2$  and the solvent was removed under vacuum, to give **7** (3.50 g, 79%) as a yellow viscous liquid that was used in next step without purification.  $^1H$  NMR (300 MHz,  $DMSO-d_6$ ):  $\delta$  7.80 (s, 2H), 4.48 (q,  $J = 7.2$  Hz, 4H), 4.28 (t,  $J = 5.5$  Hz, 2H), 3.76–3.71 (m, 4H), 2.86 (t,  $J = 5.5$  Hz, 2H), 2.63–2.54 (m, 4H), 1.46 (t,  $J = 7.1$  Hz, 6H);  $^{13}C$  NMR (75 MHz,  $DMSO-d_6$ ):  $\delta$  166.3 ( $C_q$ ), 164.1 ( $C_q$ ), 149.7 ( $C_q$ ), 114.1 (CH), 66.4 ( $CH_2$ ), 66.1 ( $CH_2$ ), 61.6 ( $CH_2$ ), 56.6 ( $CH_2$ ), 53.5 ( $CH_2$ ), 14.1 ( $CH_3$ ); MS (ESI<sup>+</sup>):  $m/z$  (%) = 114.0 (100) [4-vinylmorpholin-4-ium]<sup>+</sup>; 353.3 (32) [ $M + H$ ]<sup>+</sup>.

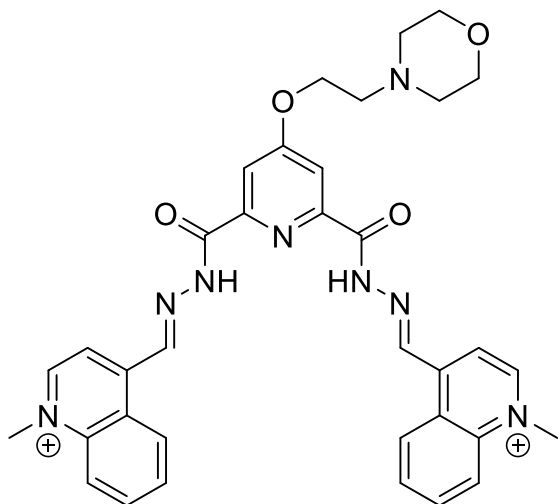
#### 4-(2-Morpholinoethoxy)pyridine-2,6-dicarbohydrazide (**L5**):



A solution of compound **17** (1.25 g, 3.50 mmol) and hydrazine hydrate (1.40 mL, 1.40 g, 28.4 mmol) in methanol (100 mL) was heated under reflux for 18 h and then cooled to room temperature. The precipitate was filtered, washed twice with methanol, once with ether, and dried, to give the bis(acylhydrazide) **L5** as a white solid (0.840 g, 73%) which was sufficiently pure and employed without further purification;  $^1H$  NMR (300 MHz,  $DMSO-d_6$ ):  $\delta$  10.58 (s, 2H), 7.59 (s, 2H), 4.61 (s, 4H), 4.31 (t,  $J = 5.5$  Hz, 2H), 3.64–3.50 (m, 4H), 2.72 (t,  $J = 5.4$  Hz, 2H), 2.48 (s, 4H);

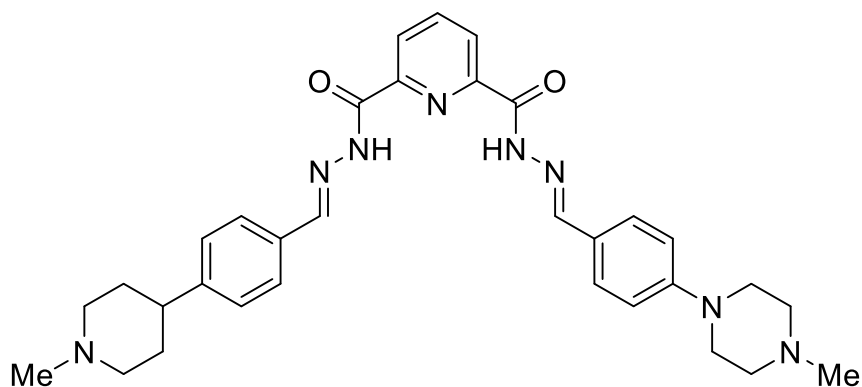
$^{13}\text{C}$  NMR (75 MHz,  $\text{DMSO-}d_6$ ):  $\delta$  166.9 ( $\text{C}_q$ ), 161.7 ( $\text{C}_q$ ), 150.5 ( $\text{C}_q$ ), 109.7 (CH), 66.3 ( $\text{CH}_2$ ), 66.2 ( $\text{CH}_2$ ), 56.6 ( $\text{CH}_2$ ), 53.6 ( $\text{CH}_2$ ); MS (ESI $^+$ ):  $m/z$  (%) = 114.2 (100) [4-vinylmorpholin-4-ium] $^+$ ; 325.3 (12) [ $M + \text{H}$ ] $^+$ .

**$N'^2, N'^6$ -Bis[(1-methylquinolinium-4-yl)methylene]-4-(2-morpholinoethoxy)pyridine-2,6-dicarbohydrazide iodide (A2-L4-A2):**



4-[2-(Morpholin-4-yl)ethoxy]pyridine-2,6-dicarbohydrazide **L5** (162 mg, 0.5 mmol) and 4-formyl-1-methylquinolinium iodide **A2** (329 mg, 1.1 mmol) were mixed in DMF (2 ml) and the reaction mixture was stirred for 2 hours at 100  $^{\circ}\text{C}$ , then cooled and filtered. The collected product was recrystallized from MeCN/ $\text{H}_2\text{O}$  (1:1 v/v) and dried in vacuo over  $\text{P}_2\text{O}_5$ , to give A2-L4-A2 (154 mg, 41%) as a red solid; m.p. (decomp.) 212  $^{\circ}\text{C}$ ;  $^1\text{H}$  NMR (300 MHz,  $\text{DMSO-}d_6$ ):  $\delta$  12.95 (s, 2H), 9.80 (s, 2H), 9.53 (d,  $J = 6.2$  Hz, 2H), 9.06 (d,  $J = 8.4$  Hz, 2H), 8.61 (d,  $J = 8.9$  Hz, 2H), 8.56 (d,  $J = 6.1$  Hz, 2H), 8.43–8.31 (m, 2H), 8.26–8.14 (m, 2H), 7.95 (s, 2H), 4.68 (s, 6H), 4.48 (s, 2H), 3.61 (s, 4H), 2.81 (s, 2H), 2.53 (s, 4H);  $^{13}\text{C}$  NMR (75 MHz,  $\text{DMSO-}d_6$ ):  $\delta$  167.6 ( $\text{C}_q$ ), 159.9 ( $\text{C}_q$ ), 149.6 ( $\text{C}_q$ ), 149.4 (CH), 147.0 ( $\text{C}_q$ ), 143.4 (CH), 139.0 ( $\text{C}_q$ ), 135.3 (CH), 130.4 (CH), 126.3 ( $\text{C}_q$ ), 125.9 (CH), 120.0 (CH), 118.9 (CH), 112.6 (CH), 66.9 ( $\text{CH}_2$ ), 66.2 ( $\text{CH}_2$ ), 56.6 ( $\text{CH}_2$ ), 53.5 ( $\text{CH}_2$ ), 45.7 ( $\text{CH}_3$ ); MS (ESI $^+$ ):  $m/z$  (%) = 631.4 (15) [ $M - \text{H}$ ] $^+$ , 316.3 [ $M$ ] $^{2+}$ , 169.2 (100) [4-cyano-1-methylquinolinium] $^+$ ; anal. calcd. for  $\text{C}_{35}\text{H}_{36}\text{I}_2\text{N}_8\text{O}_4 \times 0.7 \text{H}_2\text{O}$  (899.13): C 46.75, H 4.19, N 12.46; found: C 46.74, H 3.98, N 12.46.

**$N'^2, N'^6$ -Bis[4-(4-methylpiperazin-1-yl)benzylidene]pyridine-2,6-dicarbohydrazide (A8-L1-A8):**

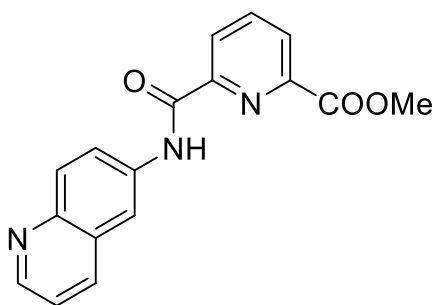


A mixture of pyridine-2,6-dicarbohydrazide **L1** (108 mg, 0.550 mmol) and 4-(4-methylpiperazino)benzaldehyde **A8** (283 mg, 1.38 mmol) in *i*-PrOH (4.0 mL) was heated for 18 h at 100 °C and then cooled to room temperature. The solvent was removed in vacuo, and the product was recrystallized from THF/Et<sub>2</sub>O, to give **A8-L1-A8** (218 mg, 69%) as a yellow solid; <sup>1</sup>H NMR (300 MHz, DMSO-*d*<sub>6</sub>): δ 12.15 (s, 2H), 8.63 (s, 2H), 8.39–8.19 (m, 3H), 7.66 (d, *J* = 8.7 Hz, 4H), 7.03 (d, *J* = 8.8 Hz, 4H), 3.30–3.22 (m, 8H), 2.45 (s, 8H), 2.22 (s, 6H); <sup>13</sup>C NMR (75 MHz, DMSO-*d*<sub>6</sub>): δ 159.1 (C<sub>q</sub>), 152.4 (C<sub>q</sub>), 150.6 (CH), 148.5 (C<sub>q</sub>), 139.9 (CH), 128.6 (CH), 125.2 (CH), 123.7 (C<sub>q</sub>), 114.5 (CH), 54.4 (CH<sub>2</sub>), 47.1 (CH<sub>2</sub>), 45.8 (CH<sub>3</sub>); MS (ESI<sup>+</sup>): *m/z* (%) = 568.6 (14) [*M* + H]<sup>+</sup>; 284.9 (100) [*M* + 2 H]<sup>2+</sup>; anal. calcd. for C<sub>31</sub>H<sub>37</sub>N<sub>9</sub>O<sub>2</sub> × 1.2 H<sub>2</sub>O (589.3): C 63.18, H 6.74, N 21.39; found: C 63.18, H 6.51, N 21.25.

#### **General procedure of synthesis of amides 21a–b:**

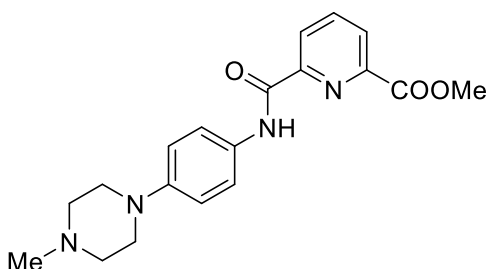
In a 250 mL round-bottom flask, 4-chloropyridine-2,6-dicarboxylic acid (543 mg, 3.00 mmol) was dissolved in a 9:1 (v/v) DCM / DMF mixture (30 mL). A corresponding amine (476 mg, 3.30 mmol), HOBt (53 mg, 0.4 mmol), and EDCI × HCl (719 mg, 8.99 mmol) were added and the mixture was stirred at RT overnight. The completion of the reaction was checked by TLC. When the spot of starting material disappeared, the solution was washed with aqueous NaHCO<sub>3</sub> solution, water, aqueous HCl (0.05 M), water, brine, and concentrated in vacuo. EtOAc was added to the resulting viscous liquid and desired product was collected by filtration.

#### **Methyl 6-[(quinolin-6-yl)carbamoyl]pyridine-2-carboxylate (21a):**



Yield: 645 mg (70%). Pale solid;  $^1\text{H}$  NMR (300 MHz,  $\text{DMSO-}d_6$ ):  $\delta$  10.70 (s, 1H), 8.83 (d,  $J = 2.8$  Hz, 1H), 8.58 (d,  $J = 1.7$  Hz, 1H), 8.44–8.31 (m, 2H), 8.31–8.21 (m, 2H), 8.14 (dd,  $J = 9.1, 2.1$  Hz, 1H), 8.03 (d,  $J = 9.1$  Hz, 1H), 7.52 (dd,  $J = 8.3, 4.2$  Hz, 1H), 3.98 (s, 3H);  $^{13}\text{C}$  NMR (75 MHz,  $\text{DMSO-}d_6$ ):  $\delta$  164.7 ( $\text{C}_q$ ), 162.5 ( $\text{C}_q$ ), 150.4 ( $\text{C}_q$ ), 149.5 (CH), 146.7 (CH), 145.1 ( $\text{C}_q$ ), 139.7 (CH), 135.9 (CH), 135.7 (CH), 129.5 (CH), 128.2 (CH), 127.6 (CH), 125.8 (CH), 124.2 (CH), 121.9 (CH), 116.6 (CH), 52.8 ( $\text{CH}_3$ ); MS (ESI $^+$ ):  $m/z$  (%) = 308.3 (100) [ $M + \text{H}$ ] $^+$ .

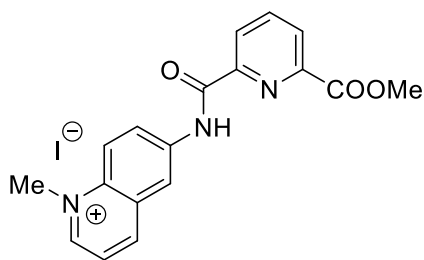
**Methyl 6-[[4-(4-methylpiperazin-1-yl)phenyl]carbamoyl]pyridine-2-carboxylate (21b):**



Yield 776 mg (73%). Yellow solid;  $^1\text{H}$  NMR (300 MHz,  $\text{CDCl}_3$ ):  $\delta$  9.98 (s, 1H), 8.48 (dd,  $J = 7.8, 0.9$  Hz, 1H), 8.25 (dd,  $J = 7.8, 0.9$  Hz, 1H), 8.13–8.00 (m, 1H), 7.71 (d,  $J = 8.9$  Hz, 2H), 6.96 (d,  $J = 9.0$  Hz, 2H), 4.04 (s, 3H), 3.32–3.10 (m, 4H), 2.71–2.48 (m, 4H), 2.36 (s, 3H);  $^{13}\text{C}$  NMR (75 MHz,  $\text{CDCl}_3$ ):  $\delta$  165.1 ( $\text{C}_q$ ), 161.0 ( $\text{C}_q$ ), 150.6 ( $\text{C}_q$ ), 148.5 ( $\text{C}_q$ ), 146.5 ( $\text{C}_q$ ), 138.9 (CH), 130.1 ( $\text{C}_q$ ), 127.4 (CH), 125.6 (CH), 121.4 (CH), 116.6 (CH), 55.2 ( $\text{CH}_2$ ), 53.1 ( $\text{CH}_3$ ), 49.5 ( $\text{CH}_2$ ), 46.3 ( $\text{CH}_3$ ); MS (ESI $^+$ ):  $m/z$  (%) = 355.4 (100) [ $M + \text{H}$ ] $^+$ .

**6-[6-(methoxycarbonyl)pyridine-2-amido]-1-methylquinolin-1-ium iodide (22a):**



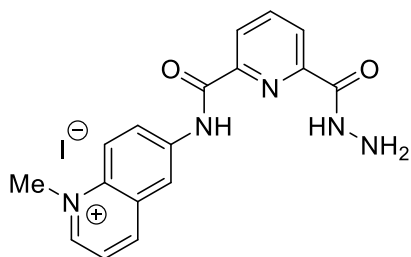


Compound **21a** (386 mg, 1.26 mmol) was dissolved in  $\text{CHCl}_3$  (25 mL). Iodomethane (1.79 mL, 4.07 g, 28.7 mmol) was added and the mixture was stirred in sealed tube at 61 °C for 18 h. The precipitate was filtered, washed twice with acetone and dried in vacuo, to give **22a** (500 mg, 89%) as a yellow solid;  $^1\text{H}$  NMR (300 MHz,  $\text{DMSO-}d_6$ ):  $\delta$  11.17 (s, 1H), 9.39 (d,  $J = 5.6$  Hz, 1H), 9.27 (d,  $J = 8.4$  Hz, 1H), 9.07 (d,  $J = 2.0$  Hz, 1H), 8.67–8.53 (m, 2H), 8.44–8.36 (m, 1H), 8.36–8.29 (m, 2H), 8.13 (dd,  $J = 8.4, 5.8$  Hz, 1H), 4.63 (s, 3H), 3.99 (s, 3H);  $^{13}\text{C}$  NMR (75 MHz,  $\text{DMSO-}d_6$ ):  $\delta$  164.7 ( $\text{C}_q$ ), 163.7 ( $\text{C}_q$ ), 150.1 ( $\text{C}_q$ ), 148.5 (CH), 146.9 ( $\text{C}_q$ ), 146.4 (CH), 139.9 (CH), 139.0 ( $\text{C}_q$ ), 135.3 ( $\text{C}_q$ ), 130.1 ( $\text{C}_q$ ), 129.6 (CH), 128.0 (CH), 126.2 (CH), 122.4 (CH), 120.1 (CH), 117.7 (CH), 52.9 ( $\text{CH}_3$ ), 45.3 ( $\text{CH}_3$ ); MS (ESI<sup>+</sup>):  $m/z$  (%) = 322.3 (100) [ $M + \text{H}$ ]<sup>+</sup>.

#### **General procedure of synthesis of acylhydrazides 23a–b:**

Hydrazine hydrate (0.45 mL, 463 mg, 50.1 mmol) was added to a solution of **22a** or **21b** (0.42 mmol) in MeOH (10 mL). After stirring for 3 h at room temperature the resulting precipitate was filtered, washed with MeOH, acetone and dried in vacuo.

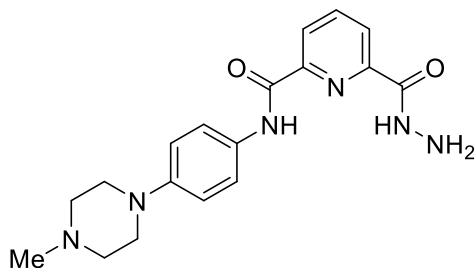
#### **6-[6-(Hydrazinecarbonyl)pyridine-2-amido]-1-methylquinolin-1-ium iodide (23a):**



Yield 170 mg (90%). Pale solid;  $^1\text{H}$  NMR (300 MHz,  $\text{DMSO-}d_6$ ):  $\delta$  11.43 (s, 1H), 10.79 (s, 1H), 9.41 (d,  $J = 5.9$  Hz, 1H), 9.31 (d,  $J = 8.3$  Hz, 1H), 9.04 (s, 1H), 8.63 (s, 2H), 8.40 (dd,  $J = 6.5, 2.2$  Hz, 1H), 8.36–8.22 (m, 2H), 8.16 (dd,  $J = 8.2, 5.7$  Hz, 1H), 4.76 (s, 2H), 4.65 (s, 3H);  $^{13}\text{C}$  NMR (75 MHz,

DMSO-*d*<sub>6</sub>):  $\delta$  162.6 (C<sub>q</sub>), 162.1 (C<sub>q</sub>), 148.7 (C<sub>q</sub>), 148.6 (CH), 147.8 (C<sub>q</sub>), 146.4 (CH), 140.0 (CH), 138.8 (C<sub>q</sub>), 135.4 (C<sub>q</sub>), 130.1 (C<sub>q</sub>), 129.8 (CH), 125.2 (CH), 125.0 (CH), 122.4 (CH), 120.1 (CH), 118.3 (CH), 45.3 (CH<sub>3</sub>); MS (ESI<sup>+</sup>):  $m/z$  (%) = 322.3 (100) [ $M + H$ ]<sup>+</sup>.

**6-(Hydrazinecarbonyl)-*N*-[4-(4-methylpiperazin-1-yl)phenyl]pyridine-2-carboxamide (23b):**

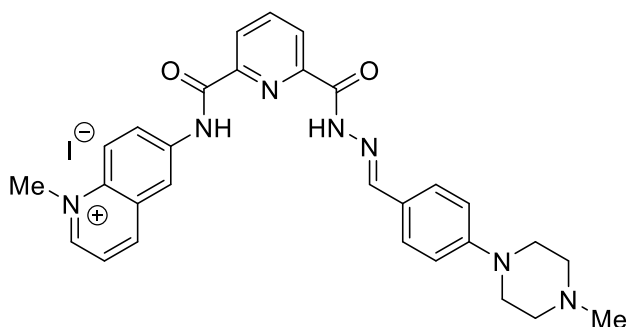


Yield 141 mg (95%). Yellow solid; <sup>1</sup>H NMR (300 MHz, DMSO-*d*<sub>6</sub>):  $\delta$  10.78 (s, 1H), 10.75 (s, 1H), 8.33–8.13 (m, 3H), 7.65 (d,  $J = 9.0$  Hz, 2H), 6.98 (d,  $J = 9.1$  Hz, 2H), 4.69 (s, 2H), 3.15–3.06 (m, 4H), 2.48–2.41 (m, 4H), 2.22 (s, 3H); <sup>13</sup>C NMR (75 MHz, DMSO-*d*<sub>6</sub>):  $\delta$  162.2 (C<sub>q</sub>), 161.2 (C<sub>q</sub>), 149.0 (C<sub>q</sub>), 148.5 (C<sub>q</sub>), 148.0 (C<sub>q</sub>), 139.6 (CH), 129.6 (C<sub>q</sub>), 124.4 (CH), 124.3 (CH), 122.4 (CH), 115.4 (CH), 54.6 (CH<sub>2</sub>), 48.3 (CH<sub>2</sub>), 45.8 (CH<sub>3</sub>); MS (ESI<sup>+</sup>):  $m/z$  (%) = 355.4 (100) [ $M + H$ ]<sup>+</sup>.

**General procedure for the synthesis of compounds 24a and 24b:**

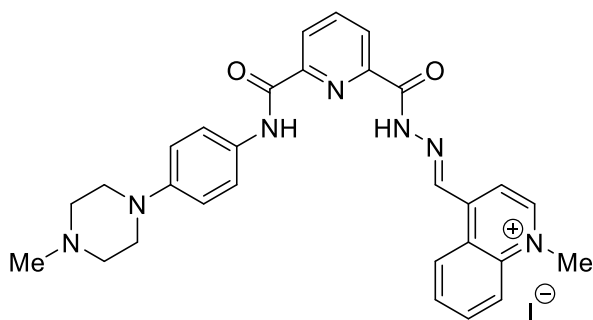
A mixture of acylhydrazide **23a** or **23b** (0.25 mmol), aldehyde **A8** or **A2**, respectively (0.75 mmol), and a catalytic amount of AcOH (10  $\mu$ L) in *n*-PrOH (4 ml) was heated at reflux for 18 h. The precipitate was collected by filtration, washed with MeOH, acetone, and recrystallized from MeCN / H<sub>2</sub>O (1:1 v/v).

**1-Methyl-6-(6-{*N'*-[4-(4-methylpiperazin-1-yl)benzylidene]hydrazinecarbonyl}pyridine-2-amido)quinolin-1-ium iodide (24a):**



Yield 40 mg (22%). Orange solid; m.p. (decomp.) 172 °C;  $^1\text{H}$  NMR (300 MHz,  $\text{DMSO-}d_6$ ):  $\delta$  12.30 (s, 1H), 11.49 (s, 1H), 9.41 (d,  $J = 5.8$  Hz, 1H), 9.33 (d,  $J = 8.3$  Hz, 1H), 9.15 (s, 1H), 8.75–8.59 (m, 3H), 8.50–8.30 (m, 3H), 8.16 (dd,  $J = 8.4, 5.9$  Hz, 1H), 7.68 (d,  $J = 8.5$  Hz, 2H), 7.06 (d,  $J = 8.7$  Hz, 2H), 4.65 (s, 3H), 2.30 (s, 3H);  $^{13}\text{C}$  NMR (75 MHz,  $\text{DMSO-}d_6$ ):  $\delta$  162.6 ( $\text{C}_q$ ), 159.0 ( $\text{C}_q$ ), 152.2 (CH), 150.7 (CH), 148.7 ( $\text{C}_q$ ), 148.6 ( $\text{C}_q$ ), 148.0 ( $\text{C}_q$ ), 146.5 (CH), 140.2 (CH), 138.9 ( $\text{C}_q$ ), 135.9 ( $\text{C}_q$ ), 130.1 ( $\text{C}_q$ ), 129.8 (CH), 128.6 (CH), 125.9 (CH), 125.5 (CH), 123.8 ( $\text{C}_q$ ), 122.4 (CH), 120.2 (CH), 117.8 (CH), 114.6 (CH), 54.2 ( $\text{CH}_2$ ), 46.8 ( $\text{CH}_2$ ), 45.4 ( $\text{CH}_3$ ), 45.3 ( $\text{CH}_3$ ); MS (ESI $^+$ ):  $m/z$  (%) = 508.5 (15) [ $M + \text{H}$ ] $^+$ , 254.9 (100) [ $M$ ] $^{2+}$ ; anal. calcd. for  $\text{C}_{29}\text{H}_{30}\text{I}\text{N}_7\text{O}_2 \times 2 \text{H}_2\text{O} \times 0.5 \text{CH}_3\text{COOH}$  (701.6): C 51.36, H 5.17, N 13.98; found: C 51.26, H 4.89, N 13.72.

***N*<sup>2</sup>-[(1-Methylquinolinium-4-yl)methylene]-6-[[4-(4-methylpiperazin-1-yl)phenyl]carbamoyl]pyridine-2-carbohydrazide iodide (24b):**



Yield 78 mg (49%). Orange solid; m.p. (decomp.) 270 °C;  $^1\text{H}$  NMR (300 MHz,  $\text{DMSO-}d_6$ ):  $\delta$  13.16 (s, 1H), 10.75 (s, 1H), 9.58 (s, 1H), 9.53 (d,  $J = 5.7$  Hz, 1H), 9.04 (d,  $J = 8.4$  Hz, 1H), 8.60 (d,  $J = 8.8$  Hz, 1H), 8.55 (d,  $J = 5.9$  Hz, 1H), 8.45 (d,  $J = 7.1$  Hz, 2H), 8.41 – 8.31 (m, 2H), 8.21 (dd,  $J = 7.5, 7.2$  Hz, 1H), 7.85 (d,  $J = 8.3$  Hz, 2H), 7.12 (d,  $J = 8.4$  Hz, 2H), 4.67 (s, 3H), 3.90–3.00 (m, 4H), 2.85 (s, 3H), 2.67–2.32 (m, 4H);  $^{13}\text{C}$  NMR (75 MHz,  $\text{DMSO-}d_6$ ):  $\delta$  161.0 ( $\text{C}_q$ ), 160.5 ( $\text{C}_q$ ), 149.4 (CH), 149.2 ( $\text{C}_q$ ), 147.4 ( $\text{C}_q$ ), 147.1 ( $\text{C}_q$ ), 146.6 ( $\text{C}_q$ ), 143.3 (CH), 140.3 (CH), 139.0 ( $\text{C}_q$ ), 135.2 (CH), 130.6 ( $\text{C}_q$ ),

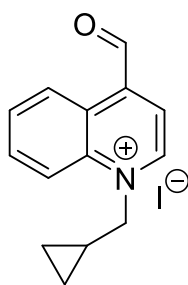
130.5 (CH), 126.2 (C<sub>q</sub>), 126.1 (CH), 125.9 (CH), 125.8 (CH), 122.2 (CH), 120.0 (CH), 119.4 (CH), 116.2 (CH), 52.6 (CH<sub>2</sub>), 46.1 (CH<sub>2</sub>), 45.7 (CH<sub>3</sub>), 42.5 (CH<sub>3</sub>); MS (ESI<sup>+</sup>): *m/z* (%) = 508.4 (100) [*M* + H]<sup>+</sup>, 254.8 (20) [*M*]<sup>2+</sup>; anal. calcd. for C<sub>29</sub>H<sub>30</sub>IN<sub>7</sub>O<sub>2</sub> × HI × H<sub>2</sub>O × 0.3 CH<sub>3</sub>COOH (799.4): C 44.47, H 4.31, N 12.26; found: C 44.66, H 4.43, N 12.18.

## RESULTS AND DISCUSSION. Part 3.

### General Procedure for the Synthesis of Aldehydes:

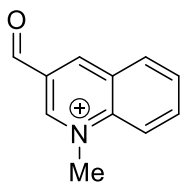
**General Procedure for the Synthesis of Quaternized Heterocyclic Aldehydes (b–i).** The solution of aldehyde (10 mmol) and alkylating agent (100 mmol) in acetone (18 mL) was stirred at 60 °C for 18 h and then cooled to room temperature. The precipitated solid was filtered, washed twice with acetone, once with ether, and dried.

### 1-(cyclopropylmethyl)-4-formylquinolin-1-ium bromide (A9):



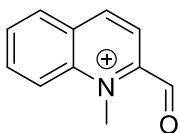
Obtained from quinoline-4-aldehyde (1.2 mmol) and cyclopropylmethyl bromide (24 mmol) in a 12% yield. White solid, <sup>1</sup>H NMR (300 MHz, D<sub>2</sub>O) δ 9.44 (d, *J* = 6.1 Hz, 1H), 8.65 (d, *J* = 8.7 Hz, 1H), 8.55 (d, *J* = 9.0 Hz, 1H), 8.36 – 8.18 (m, 2H), 8.04 (t, *J* = 7.8 Hz, 1H), 6.88 (s, 1H), 4.89 (d, *J* = 7.2 Hz, 3H), 1.72 – 1.49 (m, 1H), 0.83 (dd, *J* = 9.0, 6.0 Hz, 2H), 0.63 (dd, *J* = 9.0, 6.0 Hz, 2H);

### 3-Formyl-1-methylquinolinium iodide (A10):



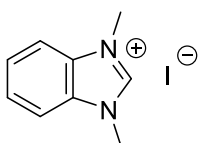
Obtained from quinoline-3-aldehyde (5 mmol) and methyl iodide (100 mmol) in a 86% yield. Orange solid,  $^1\text{H}$  NMR (300 MHz, DMSO-  $d_6$ ):  $\delta$  10.28 (s, 1H), 10.00 (s, 1H), 9.82 (s, 1H), 8.69 (d,  $J$  = 7.5 Hz, 1H), 8.60 (d,  $J$  = 8.8 Hz, 1H), 8.49 – 8.39 (m, 1H), 8.18 (m, 1H), 4.73 (s, 3H);  $^{13}\text{C}$  NMR (75 MHz, DMSO-  $d_6$ ):  $\delta$  189.9 (CO), 149.9 (CH), 148.7 (CH), 139.6 ( $\text{C}_q$ ), 137.9 (CH), 132.2 (CH), 130.9 (CH), 128.91 (s), 128.46 (s), 119.7 (CH), 45.7 ( $\text{CH}_3$ ).

### 4-formyl-1-methylquinolin-1-ium iodide (A11):



To the solution of 2-quinolinecarbaldehyde **11e** (785.8 mg, 5 mmol) in acetone (9 ml) MeI (3.1 ml, 50 mmol) was added and the mixture was stirred at 60 °C for 18 h. No precipitate was identified, so the mixture was additionally stirred at 60 °C for 2 days and the solvent was removed in vacuo. Yield: 160 mg (53%).  $^1\text{H}$  NMR (300 MHz, DMSO-  $d_6$ ):  $\delta$  10.61 (s, 1H), 9.44 (d,  $J$  = 8.4 Hz, 1H), 8.76 (d,  $J$  = 9.2 Hz, 1H), 8.54 (t,  $J$  = 9.2 Hz, 2H), 8.40 (t,  $J$  = 7.8 Hz, 1H), 8.16 (t,  $J$  = 7.5 Hz, 1H), 4.85 (s, 3H).

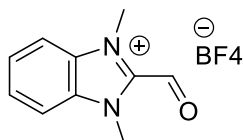
### 1,3-dimethyl-1H-1,3-benzodiazol-3-ium iodide (28):



The solution of 1-methyl-1H-benzimidazole-2-carbaldehyde **27** (80 mg, 0.5 mmol) and iodomethane (6.1 mmol, 0.4 ml) in DCM (2 ml) was stirred 72 hours at r.t. The solvent was

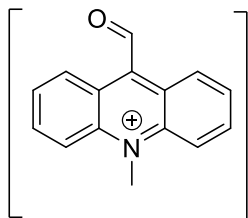
removed in vacuo. Yield: 80 mg (58%), yellow solid.  $^1\text{H NMR}$  (300 MHz,  $\text{DMSO-}d_6$ ):  $\delta$  8.03 – 7.84 (m, 2H), 7.81 – 7.59 (m, 2H), 6.51 (s, 1H), 4.14 (s, 3H), 3.73 (s, 3H).

### 2-formyl-3-methyl-1,3-benzothiazol-3-ium tetrafluoroboranuide (A13):



The mixture of of 1-methyl-1H-benzoimidazole-2-carbaldehyde **27** (160.2 mg, 1 mmol) and trimethyloxonium tetrafluoroborate (177.5 mg, 1.2 mmol) was stirred 62 h at room temperature. No precipitate occurred so temperature was raised to 60 °C. After 18 h no precipitate. The color of solution didn't change (yellow). Temperature was raised to 80 °C. The TLC shows only small amount of starting material. The solvent was evaporated, the precipitate was washed with EtOAc (twice), DCM (twice). Yield: 73 mg (28%).  $^1\text{H NMR}$  (300 MHz,  $\text{D}_2\text{O}$ )  $\delta$  7.87 (dd,  $J = 6.3, 3.2$  Hz, 2H), 7.71 (dd,  $J = 6.3, 3.2$  Hz, 2H), 6.75 (s, 1H), 4.18 (s, 6H).

### 9-formyl-10-methylacridin-10-ium iodide (A14):



Method 1. Acridine-9-carbaldehyde **26** (0.42 g, 2 mmol) was stirred with MeI (1.25 ml, 20 mmol) in acetone at 60 °C for 72 hours. No reaction occurred.

Method 2. A mixture of (62.2 mg, 0.3 mmol) of acridine-9-carboxaldehyde **26**, (83.3 mg, 0.45 mmol) of methyl *p*-toluenesulfonate and 5 ml of acetone was heated at 60 °C under stirring for 12 hours. No reaction occurred.

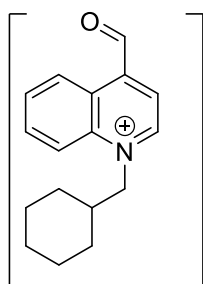
Method 3. A mixture of (62.2 mg, 0.3 mmol) of acridine-9-carboxaldehyde **26**, (83.3 mg, 0.45 mmol) of methyl *p*-toluenesulfonate and 5 ml of toluene was heated at 110 °C under stirring for 12 hours. After cooling to room temperature, the mixture is decanted from the resinous product;

the residue is washed hot with tert-butyl methyl ether and dried in vacuum over sulfuric acid. 46 mg (39 %) of solid. In LC/MS methylated acid. NMR: no pick of aldehyde or hem-diole. The heating of the mixture for 72 hours at 10 °C led to decarbonilation product.

Method 4. To the solution of acridine-9-carbaldehyde (30 mg, 0.14 mmol) in DCM methyl triflate (0.02 ml, 0.18 mmol) was added. The mixture was stirred at room temperature 18 h. Red precipitate was filtered. The product was not identified.

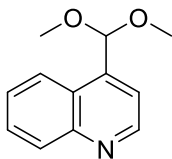
Method 5. 9-Formylacridine (20 mg, 0.10 mmol) was heated in a sealed tube at 100° with dimethyl sulfate (0.4 ml, 4.8 mmol) for 15 min. The solution was cooled and the precipitate was filtered. After recrystallization from Et<sub>2</sub>O/MeCN the product contained still at least 15% of the starting material. <sup>1</sup>H NMR (300 MHz, D<sub>2</sub>O) δ 11.38 (s, 1H), 9.05 (d, *J* = 8.9 Hz, 2H), 8.52 (d, *J* = 9.4 Hz, 2H), 8.33 – 8.16 (m, 2H), 7.92 (dd, *J* = 17.5, 10.3 Hz, 3H), 7.58 (s, 1H).

#### 1-(cyclohexylmethyl)-4-formylquinolin-1-ium bromide (A15):



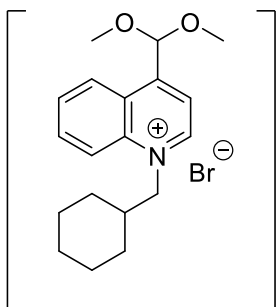
4-Quinolinecarbaldehyde **11b** (157.2 mg, 1 mmol) was dissolved in 1 mL of acetonitrile. One molar equivalent of (bromomethyl)cyclohexane (177.1 mg) was dissolved in 1 mL of acetonitrile and added to the dissolved 4-quinolinecarbaldehyde. The reaction was fitted with a reflux condenser and left to stir at 70 °C for 72 hours, then rotary evaporated under reduced pressure to remove the solvent. The resultant yellow liquid was dissolved in water and the unreacted alkyl halide formed a separate phase at the bottom of the flask that was removed by pipette. The solution was then rotary evaporated under reduced pressure to remove the residual water. <sup>1</sup>H NMR (300 MHz, CDCl<sub>3</sub>) δ 10.53 (s, 1H), 9.22 (d, *J* = 4.2 Hz, 1H), 9.13 – 8.95 (m, 1H), 8.23 (d, *J* = 8.3 Hz, 1H), 7.91 – 7.69 (m, 3H), 3.28 (d, *J* = 6.3 Hz, 0.5H), 1.96 – 0.79 (m, 5H).

#### 4-(dimethoxymethyl)quinoline (30):



To a solution of aldehyde **11b** (1.04 g, 6.6 mmol) in MeOH (65 mL) was added CH(OMe)<sub>3</sub> (3.54 g, 33.4 mmol, 3.65 ml) and pTsOH·H<sub>2</sub>O (0.6 g, 3.3 mmol), followed by refluxing for 24 h. The reaction mixture was evaporated in vacuo and to the resulting residue EtOAc (150 ml) and aqueous NaHCO<sub>3</sub> (75 ml) were added. Organic layer was separated and washed with water (75 mL) then brine, dried over anhydrous Na<sub>2</sub>SO<sub>4</sub>, and evaporated in vacuo resulting in giving yellow liquid and a little amount of residue. The product was dissolved in cyclohexane and filtered through SiO<sub>2</sub>. Cyclohexane was removed in vacuo. Yellow liquid. Yield 1.2 g (90%), <sup>1</sup>H NMR (300 MHz, CDCl<sub>3</sub>) δ 8.94 (d, *J* = 4.4 Hz, 1H), 8.25 (d, *J* = 8.4 Hz, 1H), 8.14 (d, *J* = 8.4 Hz, 1H), 7.80 – 7.67 (m, 1H), 7.67 – 7.51 (m, 2H), 5.95 (s, 1H), 3.37 (s, 6H).

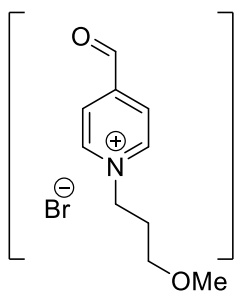
#### 1-(cyclohexylmethyl)-4-(dimethoxymethyl)quinolin-1-ium bromide (31):



To a solution of 4-(dimethoxymethyl)quinoline **30** (214 mg, 1.05 mmol) in DMSO (3 ml) cyclohexylmethyl bromide (0.16 ml, 1.16 mmol) was added. The reaction was stirred 48 hours at room temperature. No desired product was identified.

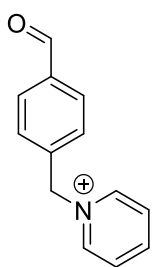
#### 4-formyl-1-(3-methoxypropyl)pyridin-1-ium (A16):





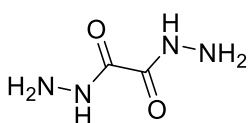
Pyridine-4-carbaldehyde **11a** (94.1  $\mu$ l, 1.00 mmol) and methyl 4-bromobutyrate (284  $\mu$ l, 2.25 mmol) were stirred 18 hour at 60 °C in acetone. **A17** was not identified in the obtained mixture of products.

#### 1-[(4-formylphenyl)methyl]-pyridinium bromide (**A17**):



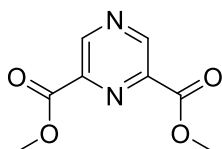
Obtained from 4-bromomethylbenzaldehyde (2.5 mmol, 0.50 g) and pyridine (55 mmol, 4.35 g, 4.45 ml). Yield: 0.60 g (86%). White solid,  $^1\text{H}$  NMR (300 MHz, DMSO- $d_6$ )  $\delta$  10.03 (s, 1H), 9.25 (d,  $J = 5.2$  Hz, 2H), 8.72-8.60 (m, 1H), 8.27-8.15 (m, 2H), 7.98 (d,  $J = 7.6$  Hz, 2H), 7.71 (d,  $J = 7.3$  Hz, 2H), 6.01 (s, 2H);  $^{13}\text{C}$  NMR (75 MHz, DMSO- $d_6$ ):  $\delta$  192.7 (C<sub>q</sub>), 146.3 (CH), 145.1 (CH), 140.4 (C<sub>q</sub>), 136.5 (C<sub>q</sub>), 130.1 (CH), 129.3 (CH), 128.6 (CH), 62.6 (CH<sub>2</sub>).

#### Oxalyl dihydrazide (**G, L6**):



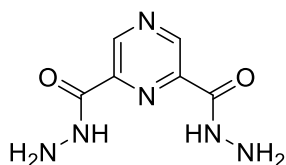
*Oxalyl dihydrazide* was prepared by reacting diethyloxalate (3.22 g, 22 mmol, 3 mL) and hydrazine hydrate (1.99 g, 62.2 mmol, 2 mL) in ethanol (100 ml) at room temperature under stirring for 30 min. The precipitate was filtered and recrystallized from H<sub>2</sub>O in order to obtain white crystalline product. Yield: 2.25 g (87%). <sup>1</sup>H NMR (300 MHz, DMSO-*d*<sub>6</sub>): δ 9.93 (s, 2H), 4.47 (s, 4H);

#### Dimethyl pyrazine-2,6-dicarboxylate (35):

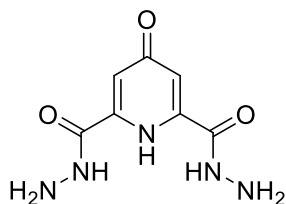


A solution of 2,6-dimethylpyrazine **34** (2.4 g, 22.3 mmol) in pyridine/H<sub>2</sub>O (10:1) (50 mL) was treated with solid SeO<sub>2</sub> (11.1 g, 100.2 mmol) and the resulting suspension was refluxed for 18 hours. The resulting dark red-brown mixture was evaporated to dryness under reduced pressure. H<sub>2</sub>O (80 mL) was added and the solid elemental selenium was filtered off. After evaporation of the red solution to dryness, the resulting dark red brown solid was taken up in MeOH (50 mL), treated with SOCl<sub>2</sub> (1.2 mL) and refluxed for 8 hours. The resulting suspension was filtered whilst hot and the solid was washed with DCM (5×20 mL). The combined organic layers were reduced in volume in vacuo to give the product in the form of pale yellow-orange feathery crystals. <sup>1</sup>H NMR (300 MHz, DMSO) δ 9.40 (s, 2H), 3.96 (s, 6H).

#### Pyrazine-2,6-dicarbohydrazide (C, L11):



*Pyrazine-2,6-dicarbohydrazide (C)* was prepared from 2.5 mmol 2,6-dimethyl pyrazine-2,6-dicarboxylate (490.4 mg) and hydrazine hydrate by the same procedure as for **L1–L4**. Yield: 340 mg (69%). <sup>1</sup>H NMR (300 MHz, DMSO-*d*<sub>6</sub>): δ 10.63 (s, 2H), 9.27 (s, 2H), 4.72 (s, 4H); <sup>13</sup>C NMR (75 MHz, DMSO-*d*<sub>6</sub>): δ 160.7 (C<sub>q</sub>), 144.9 (CH), 142.2 (C<sub>q</sub>); MS (ESI<sup>+</sup>): *m/z* (%) = 197.2 (100) [*M* + H]<sup>+</sup>.

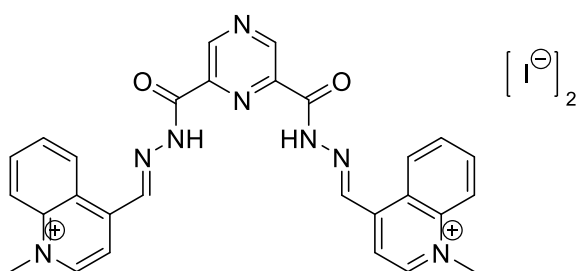
**Chelidamic acid dihydrazide (L12):**

To the solution of 2,6-diethyl 4-hydroxypyridine-2,6-dicarboxylate (478.4 mg, 2 mmol) in 150 ml of ethanol hydrazine hydrate (2.14 ml) was added. The mixture was stirred 18 h at 78 °C and then cooled to r.t. White precipitate was filtered, washed with ethanol then diethyl ether and dried in vacuum over P<sub>2</sub>O<sub>5</sub>. Yield: 349 mg (83%). <sup>1</sup>H NMR (300 MHz, DMSO) δ 10.44 (s, 2H), 7.33 (s, 2H), 4.52 (s, 4H).

**Synthesis of “ready-to-screen” solutions:** 9 mM solutions of building blocks Lx and Ay were added in mixture DMSO – acetic acid. The final concentrations of reagent are the following (Lx) = 2 mM; c (Ly) = 4.2 mM; c(AcOH) = 1 M. The resulting solutions were heated 48 hours in the oven without mixing. When solutions were cooled to r.t., they were analyzed using LC/MS. The peak integration gave the value of purity (and therefore, the yield) of the product.

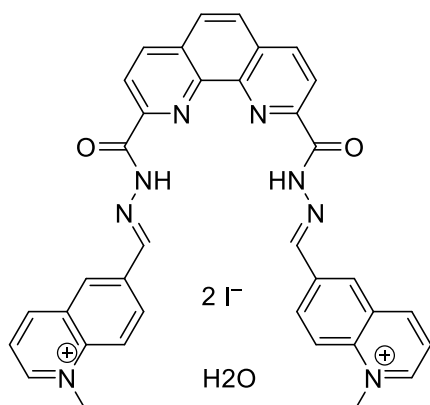
**Synthesis of ‘hits’:** The mixture of dicarbohydrazide **A**, **C** or **E** (0.5 mmol) and quaternized aldehyde **b**, **f** or **i** (1.1 mmol) in DMF (2.5 mL) was heated at 100 °C for 2 h and then cooled to room temperature. The precipitate was collected by filtration, washed three times with MeCN, once with ether, dried and then additionally recrystallized from MeCN/H<sub>2</sub>O. The yields are indicated for <sup>1</sup>H-NMR spectroscopically pure material prior to the final recrystallization step.

**N’2,N’6-Bis[(1-methylquinolinium-4-yl)methylene]pyrazine-2,6-dicarbohydrazide iodide (3Cb, OR189):**



Yield 89%. Red solid, m.p. (decomp.) 277 °C;  $^1\text{H}$  NMR (300 MHz,  $\text{DMSO-}d_6$ ):  $\delta$  13.00 (s, 2H), 9.84 (s, 2H), 9.60 (s, 2H), 9.55 (d,  $J = 6.2$  Hz, 2H), 9.05 (d,  $J = 8.4$  Hz, 2H), 8.60 (d,  $J = 8.9$  Hz, 2H), 8.54 (d,  $J = 6.1$  Hz, 2H), 8.40–8.30 (m, 2H), 8.22–8.12 (m, 2H), 4.69 (s, 6H).;  $^{13}\text{C}$  NMR (75 MHz,  $\text{DMSO-}d_6$ ):  $\delta$  159.3 ( $\text{C}_q$ ), 149.4 (CH), 147.4 (CH), 146.8 ( $\text{C}_q$ ), 143.7 (CH), 141.8 ( $\text{C}_q$ ), 139.0 ( $\text{C}_q$ ), 135.3 (CH), 130.5 (CH), 126.3 ( $\text{C}_q$ ), 125.9 (CH), 120.0 (CH), 119.1 (CH), 45.7 ( $\text{CH}_3$ ); MS (ESI $^+$ ):  $m/z$  (%) = 503.3 (10%) [ $M - \text{H}$ ] $^+$ , 335.2 (15%) [ $M - 4\text{-cyano-1-methylquinolinium}$ ] $^+$ , 252.3 [ $M$ ] $^{2+}$ , 169.2 (100%) [4-cyano-1-methylquinolinium] $^+$ ; purity (LC): 99%; anal. calcd. for  $\text{C}_{28}\text{H}_{24}\text{I}_2\text{N}_8\text{O}_2 \times 2.5 \text{H}_2\text{O}$  (803.39): C 41.86, H 3.64, N 13.95; found: C 41.82, H 3.30, N 13.85.

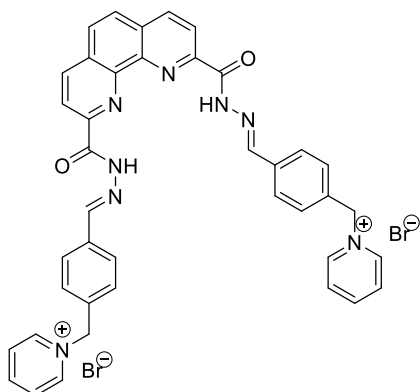
**N',N'-Bis[(1-methylquinolinium-6-yl)methylene]-1,10-phenanthroline-2,9-dicarbohydrazide iodide (3Ef, OR192):**



3.5 excess of aldehyde, 110 °C (because monoacylhydrazone precipitate) Yellow solid, m.p. >290 °C;  $^1\text{H}$  NMR (300 MHz,  $\text{DMSO-}d_6$ ):  $\delta$  13.10 (s, 2H), 9.55 (d,  $J = 5.7$  Hz, 2H), 9.25 (s, 2H), 9.15 (d,  $J = 8.3$  Hz, 2H), 8.89 (d,  $J = 8.3$  Hz, 2H), 8.74 (d,  $J = 8.3$  Hz, 4H), 8.64 (d,  $J = 8.4$  Hz, 4H), 8.32 (s, 2H), 8.23 – 8.13 (m, 2H), 4.70 (s, 6H);  $^{13}\text{C}$  NMR (75 MHz,  $\text{DMSO-}d_6$ ):  $\delta$  160.7 ( $\text{C}_q$ ), 150.4 (CH), 149.0 ( $\text{C}_q$ ), 147.4 (CH), 147.1 (CH), 143.7 ( $\text{C}_q$ ), 139.1 ( $\text{C}_q$ ), 138.8 (CH), 135.7 ( $\text{C}_q$ ), 132.7 (CH), 130.9 ( $\text{C}_q$ ), 129.6 ( $\text{C}_q$ ), 128.9 (CH), 128.5 (CH), 122.9 (CH), 122.0 (CH), 120.3 (CH), 45.7 ( $\text{CH}_3$ ); MS (ESI $^+$ ):  $m/z$  (%) =

302.3 (100)  $[M]^{2+}$ ; purity (LC): 99%; anal. calcd. for  $C_{36}H_{28}Br_2N_8O_2 \times 3 H_2O$  (912.51): C 47.38, H 3.76, N 12.28; found: C 47.16, H 3.59, N 11.91.

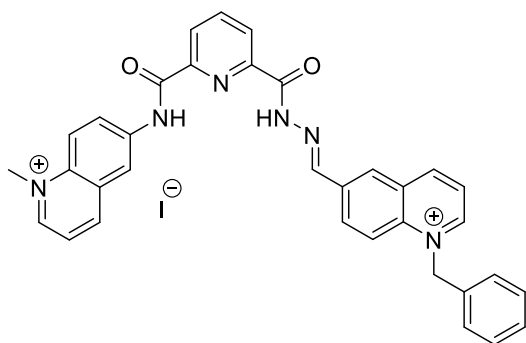
**N',N'-Bis[(1-[(phenyl-4-yl)-methyl]-pyridinium)methylene]-1,10-phenanthroline-2,9-dicarbohydrazide bromide (3Ei, OR180):**



Yield 47%. Pale solid, m.p. (decomp.) 257 °C;  $^1H$  NMR (300 MHz, DMSO- $d_6$ ):  $\delta$  12.79 (s, 2H), 9.44 (d,  $J$  = 5.8 Hz, 4H), 8.95 (s, 2H), 8.81 (d,  $J$  = 8.3 Hz, 2H), 8.69 (t,  $J$  = 7.7 Hz, 2H), 8.56 (d,  $J$  = 8.3 Hz, 2H), 8.31 - 8.18 (m, 6H), 7.86 (d,  $J$  = 8.1 Hz, 4H), 7.79 (d,  $J$  = 8.1 Hz, 4H), 6.07 (s, 4H);  $^{13}C$  NMR (75 MHz, DMSO- $d_6$ ):  $\delta$  160.3 (C<sub>q</sub>), 149.1 (CH), 149.1 (C<sub>q</sub>), 146.2 (CH), 145.0 (CH), 143.7 (C<sub>q</sub>), 138.6 (CH), 136.1 (C<sub>q</sub>), 135.3 (C<sub>q</sub>), 130.8 (C<sub>q</sub>), 129.6 (CH), 128.6 (CH), 128.4 (CH), 127.9 (CH), 121.8 (CH), 62.8 (CH<sub>2</sub>); MS (ESI<sup>+</sup>):  $m/z$  (%) = 657.3 (5%)  $[M + H]^+$ , 116.1 (100%) [5-cyano-2-methylbenzene-1-ylum]; purity (LC): 99%; anal. calcd. for  $C_{40}H_{32}Br_2N_8O_2 \cdot 4 H_2O$  (888.60): C 54.07, H 4.54, N 12.61; found: C 53.96, H 4.34, N 12.78.

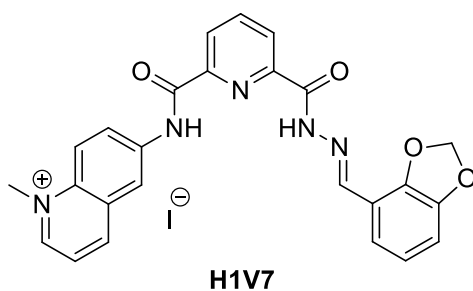
**RESULTS AND DISCUSSION. Part 4.**

**6-(6-{N'-[(1E)-(1-benzylquinolin-1-ium-6-yl)methylidene]hydrazinecarbonyl}pyridine-2-amido)-1-methylquinolin-1-ium bromide iodide (H1O4, OR286):**



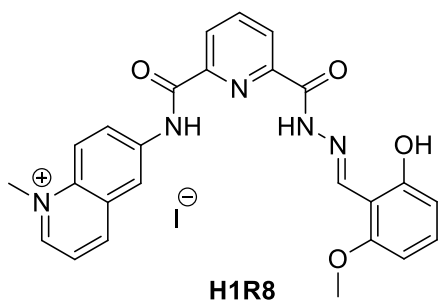
$^1\text{H}$  NMR (300 MHz, DMSO)  $\delta$  12.88 (s, 1H), 11.54 (s, 1H), 9.77 (d,  $J = 5.9$  Hz, 1H), 9.47 (dd,  $J = 14.4$ , 6.9 Hz, 2H), 9.34 (d,  $J = 8.1$  Hz, 1H), 9.21 (s, 1H), 9.15 (s, 1H), 8.88 (s, 1H), 8.80 – 8.58 (m, 4H), 8.55 – 8.45 (m, 2H), 8.44 – 8.32 (m, 2H), 8.24 – 8.09 (m, 1H), 7.44 (s, 5H), 6.42 (s, 2H), 4.68 (s, 3H).

**6-(6-{N'-[(1E)-(2H-1,3-benzodioxol-4-yl)methylidene]hydrazinecarbonyl}pyridine-2-amido)-1-methylquinolin-1-ium iodide (H1V7, OR278)**



$^1\text{H}$  NMR (300 MHz, DMSO)  $\delta$  12.55 (s, 1H), 11.46 (s, 1H), 9.40 (s, 1H), 9.33 (d,  $J = 7.3$  Hz, 1H), 9.17 (s, 1H), 8.94 (s, 1H), 8.67 (s, 2H), 8.55 – 8.40 (m, 2H), 8.38 (d,  $J = 7.3$  Hz, 1H), 8.16 (s, 1H), 7.40 (d,  $J = 7.6$  Hz, 1H), 7.03 (s, 1H), 6.97 (s, 1H), 6.19 (s, 2H), 4.65 (s, 3H).

**6-(6-{N'-[(1E)-(2-hydroxy-6-methoxyphenyl)methylidene]hydrazinecarbonyl}pyridine-2-amido)-1-methylquinolin-1-ium iodide (H1R8, OR279)**



<sup>1</sup>H NMR (300 MHz, DMSO)  $\delta$  12.63 (s, 1H), 11.49 (s, 1H), 10.48 (s, 1H), 9.41 (d,  $J = 5.3$  Hz, 1H), 9.33 (d,  $J = 8.3$  Hz, 1H), 9.17 (s, 1H), 9.00 (s, 1H), 8.66 (s, 2H), 8.54 – 8.41 (m, 2H), 8.41 – 8.28 (m, 1H), 8.22 – 8.11 (m, 1H), 7.26 (s, 1H), 6.93 (t,  $J = 7.7$  Hz, 2H), 4.65 (s, 3H), 3.77 (s, 3H).

# Single-crystal X-ray diffraction analysis

X-ray diffraction data for **PyDH1** ( $2I^-$ ) were collected using a VENTURE PHOTON100 CMOS Bruker diffractometer with Micro-focus IuS source Mo  $K\alpha$  radiation. X-ray diffraction data for **PymDH1** ( $2I^-$ ) were collected using a X8 APEXII CCD Bruker diffractometer with graphite-monochromated Mo  $K\alpha$  radiation. X-ray diffraction data for **PhenDH1** ( $2I^-$ ) were collected using a VENTURE PHOTON100 CMOS Bruker diffractometer with Micro-focus IuS source Cu  $K\alpha$  radiation. All crystals were mounted on a CryoLoop (Hampton Research) with Paratone-N (Hampton Research) as cryoprotectant and then flash-frozen in a nitrogen gas stream at 100 K. For compounds, the temperature of the crystal was maintained at the selected value by means of a 700 series Cryostream (for X8) or N-Helix (for VENTURE) cooling device within an accuracy of  $\pm 1$  K. The data were corrected for Lorentz polarization, and absorption effects. The structures were solved by direct methods using SHELXS-97<sup>196</sup> and refined against  $F^2$  by full-matrix least-squares techniques using SHELXL-2018<sup>197</sup> with anisotropic displacement parameters for all non-hydrogen atoms. All calculations were performed using the Crystal Structure crystallographic software package WINGX<sup>198</sup>. The crystal data collection and refinement parameters are given in Table S3. CCDC 1881844–1881846 contain the supplementary crystallographic data for this paper. These data can be obtained free of charge from the Cambridge Crystallographic Data Centre via <https://www.ccdc.cam.ac.uk/structures/>.



# Biophysical assays

## FRET-melting assay

The assay was performed with a double-labeled oligoribonucleotide (g4-EBNA1, 5'-FAM-r(GGGGCAGGAGCAGGAGGA)-TAMRA-3'), which was annealed prior to experiments (95 °C, 5 min) in K-10 buffer (10 mM LiAsO<sub>2</sub>Me<sub>2</sub>, 10 mM KCl, 90 mM LiCl, pH 7.3) at a concentration of 10 μM. Double-stranded DNA competitor (ds26, 5'-CAATCGGATCGAATTCGATCCGATTG-3') was annealed in the same buffer at a strand concentration of 200 μM. Thermal denaturation runs were performed in 96-well plates with a real-time PCR apparatus (7900HT Fast Real-Time PCR, Applied Biosystems) using a heating ramp of 0.5 °C per minute from 25 to 95 °C; the fluorescence intensity was monitored in the FAM channel. Each well contained 0.2 μM of double-labeled g4-EBNA1, 1 μM of tested compound, and/or 0, 3 or 10 μM of ds26 competitor, in a total volume of 25 μL of K10 buffer. The denaturation temperatures ( $T_m$ ) were determined from the maxima of first-derivative plots of FAM emission intensity vs. temperature, and ligand-induced  $T_m$  shifts ( $\Delta T_m$ ) were calculated as a difference of mean denaturation temperatures in the presence and in the absence of ligands ( $T_m^0 = 58.6$  °C), from three independent experiments. Analogous experiments were performed with G4-forming DNA oligonucleotides F-myc22-T: 5'-FAM-d(TGAGGGTGGGTAGGGTGGGTAA)-TAMRA-3' and F-25TAG-T: 5'-AM-d(TAGGGTTAGGGTTAGGGTTAGGGAA)-TAMRA-3', in K-1 (10 mM LiAsO<sub>2</sub>Me<sub>2</sub>, 1 mM KCl, 99 mM LiCl, pH 7.3) and K-100 (10 mM LiAsO<sub>2</sub>Me<sub>2</sub>, 100 mM KCl, pH 7.3) buffers, respectively.

## TO displacement assay

This assay is performed in a 96-well microplate format as described<sup>188</sup>. The non-labeled g4-EBNA1 oligoribonucleotide was annealed (95 °C, 5 min) in K100 buffer (10 mM KAsO<sub>2</sub>Me<sub>2</sub>, 100 mM KCl, 1 mM EDTA, 1% v/v DMSO) at a concentration of 5 μM and, after cooling, supplemented with TO (10 μM). Every row of a black-bottom, 96-well microplate was filled with K100 buffer (q.s.p. 200 μL per well), pre-folded g4-EBNA1 + TO solution (final concentrations: 0.25 and 0.5 μM, respectively), and an extemporaneously prepared ligand solution (5 μM in the same buffer; final ligand concentration: 0 to 2.5 μM). After 5 min of orbital shaking, fluorescence intensity was

measured with a Fluostar Omega microplate reader (BMG Labtech) using the following parameters: 20 flashes per well, emission / excitation filters: 485 / 520 nm, gain adjusted at 80% of the fluorescence from the most fluorescent well. The experiments were performed in duplicate. The percentage of TO displacement was calculated from the fluorescence intensity ( $F$ ) as  $\%(\text{TO displacement}) = 100 \times (1 - F / F_0)$ , where  $F_0$  is the fluorescence intensity of TO–RNA complex in the absence of ligands. The percentage of displacement was plotted against the concentration of added ligand, and ligand affinity was characterized by the concentration required to decrease the fluorescence of the probe by 50% ( $DC_{50}$ ) after interpolation of the displacement curve. A control experiment was performed at identical conditions with double-stranded DNA substrate ds26.

## Fluorimetric titrations

**Fluorimetric titrations** were performed as described elsewhere.<sup>66</sup> 5'-Cy5-labelled oligonucleotides were dissolved at a concentration of 1  $\mu\text{M}$  in a 10 mM  $\text{LiAsO}_2\text{Me}_2$ , 100 mM KCl buffer, pH 7.2, annealed at 95 °C for 5 min, slowly cooled to room temperature, and then further diluted with the **titration buffer** (10 mM  $\text{LiAsO}_2\text{Me}_2$ , 100 mM KCl, 0.5 w/v % CHAPS, 0.05 v/v % Triton X-100, pH 7.2) to a final concentration of 2.22 nM. Ten serial dilutions (1:1) of ligands in  $\text{H}_2\text{O}$  containing 0.1 vol.% DMSO were prepared starting from two ligand solutions with initial concentrations of 100  $\mu\text{M}$  and 18  $\mu\text{M}$ , respectively, resulting in a total of 22 solutions with different ligand concentrations and two no-ligand controls. The oligonucleotide solution (90  $\mu\text{L}$ ) and ligand solutions (10  $\mu\text{L}$ ) were transferred into 96-well black, flat-bottom polystyrene non-binding (NBS) microplates (Corning) (final oligonucleotide concentration: 2 nM, final ligand concentration: from 10  $\mu\text{M}$  to 1.8 nM). After incubation for 2 h at room temperature, fluorescence intensity was measured on a CLARIOstar Plus microplate reader (BMG) using 590BP50 (excitation), LP639 (dichroic) and 675BP50 (emission) filters and an integration time of 0.5 s per well. The fluorescence intensity was normalized by dividing the raw value by the mean intensity of no-ligand wells.

## CD spectroscopy

CD spectra were recorded with a Jasco J-1500 spectropolarimeter. Spectra were recorded from 3  $\mu\text{M}$  solutions of DNA in buffer K100 without or with 0.5, 1 or 2 equivalents of ligand. Spectra with ligands were recorded without incubation in quartz cuvettes with rectangular cross-section (path length  $1 \times 0.5$  cm), with the beam passing through a path length of 0.5 cm. Parameters used for spectra acquisition: wavelength range, 230–330 nm; scan speed, 50 nm  $\text{min}^{-1}$ ; number of averaged scans, 3; data pitch, 0.5 nm; bandwidth, 2 nm; integration time, 1 s; temperature, 22 °C. Spectra were subsequently corrected for the blank.

## Electrospray mass spectrometry.

Native ESI-MS spectra of DNA and DNA–ligand complexes were obtained using an Exactive ESI–Orbitrap mass spectrometer (Thermo Scientific, Bremen, Germany). ESI spray voltage and capillary voltage were 3.50 kV and –3.5 kV, respectively. Capillary temperature was set to 275 °C. The syringe injection rate was 200  $\mu\text{L}/\text{h}$ . The presented spectra (Figs S4–S5) result from 3-min accumulations (1 scan per 1.1 s). The concentrations of the initial stocks solutions of DNA ( $\sim 1$  mM) were measured by UV absorbance at 260 nm on a Uvikon XS. 50  $\mu\text{M}$  solutions of DNA were annealed in buffer that contained 100 mM trimethylammonium acetate (TMAA, Ultra for HPLC, Fluka analytical) and 1 mM KCl (>99.999%, Sigma) by heating at 85 °C for 10 minutes and slowly cooling to room temperature. Analyzed samples contained 5  $\mu\text{M}$  of G4-DNA, 0 or 5  $\mu\text{M}$  of ligand, and 1  $\mu\text{M}$  of the oligonucleotide 5'-d(TTTTTT)-3' ( $dT_6$ , internal control) in a 100 mM TMAA, 1 mM KCl buffer, and were incubated 2 h before the injection.

Charge state  $6^-$  was examined because it contained fewer nonspecific potassium adducts. We assume that the response factors of free DNA, DNA+L and DNA+2L complexes are equal.

# Protocols of generation of dynamic combinatorial libraries.

## RESULTS AND DISCUSSION. Part 2.

**Capturing protocol for extraction of ligands from DCLs:** Stock solutions of the building blocks were prepared in DMSO/H<sub>2</sub>O (1:1 v/v) at a concentration of 2 mM. Stock solutions of the catalysts **1a–d** were prepared in the **DCC buffer** (100 mM NH<sub>4</sub>OAc, 1.5 mM KCl, pH 6.4) at a concentration of 25 mM. DCLs were prepared by mixing the stock solutions of the building blocks (cf. Table S2) with the DCC buffer. The solutions were supplemented with the catalyst (final concentration: 10 mM) and biotinylated DNA oligonucleotides pre-folded in K1 buffer (5  $\mu$ L, final concentration: 5  $\mu$ M) or an equal volume of the K1 buffer, to a final library volume of 100  $\mu$ L. The libraries were left to equilibrate for 24 h without stirring at ambient temperature. Then, NaOH solution (2 M, 1  $\mu$ L) was added to stop the acylhydrazone exchange. In parallel, 4  $\times$  100  $\mu$ L of streptavidin-coated magnetic beads (Dynabeads MyOne Streptavidin T1, 10 mg mL<sup>-1</sup> suspension, ThermoFisher) were washed with the DCC buffer (2  $\times$  200  $\mu$ L), biotin solution (2 mM, 200  $\mu$ L), DCC buffer (200  $\mu$ L), decanted, and the reaction mixtures were added to the beads. The solutions were incubated for 20 min at ambient temperature without stirring. The supernatant was discarded, and the beads were washed with 3  $\times$  200  $\mu$ L of the DCC buffer. Formamide (100  $\mu$ L) was added and mixed with the beads. Samples were heated at 50 °C for 30 minutes, decanted and the supernatant was subjected to HPLC analysis (injection volume: 5  $\mu$ L). Column: Atlantis T3, 3  $\mu$ m, 3  $\times$  100 mm (Waters), eluent A: H<sub>2</sub>O + 0.01% TFA; eluent B: ACN + 0.01% TFA; flow rate: 0.8 mL min<sup>-1</sup>; detection wavelength: 310 nm. Elution methods were developed as to optimize the separation of DCL components, with a typical gradient as following: 0 min, 5% B; 20 min, 13% B; 25 min, 30% B; 26 min, 80% B; 30 min, 80% B (linear curves).

## RESULTS AND DISCUSSION. Part 4.

### *Experiments with nanoparticles*

#### **Synthesis of nanoparticles:<sup>1</sup>**

Mix  $\text{FeCl}_3$  (4 ml, 1 M) and  $\text{FeCl}_2$  (1 ml, 2M in 2M HCl) and add into diluted  $\text{NH}_3$  solution (50 ml, 0.7 M). The reaction mixture was stirred for 30 min at RT. Then, the precipitate was isolated by magnetic decantation, stirred with diluted  $\text{HClO}_4$  (50 ml, 2 M) and then collected by centrifugation. The residue was finally made up to 50 ml with deionized water. An aqueous solution of  $\text{HAuCl}_4$  (5 mL, 2.0 mg/mL) was mixed into 20 mL of deionized water and boiled for 5 min. Then, 1 mL of  $\text{Fe}_3\text{O}_4$  nanoparticle solution (synthesized by reported procedure) was added into the reaction mixture followed by the addition of sodium citrate (1 mL, 80 mmol). The color of the solution gradually changed from brown to burgundy. The reaction mixture was refluxed under stirring for 5 min. After cooling, the solution was sonicated for 10–15 min and then gold-coated magnetic nanoparticles ( $\text{Au@Fe}_3\text{O}_4$ ) were collected using a magnet, washed three times, and re-dispersed in pure water.

Thiolated DNA (*hp* and *Pu24T*) were diluted with MOPS buffer to the concentration 100  $\mu\text{M}$ . They were annealed (5 min; 95 °C). Before the usage of DNA they were mixed with  $\text{NaBH}_4$ <sub>aq.sol.</sub> (5 mM) 1 eqv for 15 min. 100  $\mu\text{L}$  of NP were added and the mixture was incubated 16 h. After standing for 16 h with  $\text{Au@Fe}_3\text{O}_4$ , nanoparticles were separated using magnet and washed with 20 mM MOPS buffer containing 100 mM KCl (pH 6.5). The volume of the nanoparticle solution was made up to 100  $\mu\text{L}$  with the buffer.

#### **Preliminary experiments:**

To the resulting 100  $\mu\text{L}$  of DNA-NP were added 0.5 mL of MOPS buffer and 5  $\mu\text{L}$  of 10 mM mixture of 2 ligands (*OR19* and *OR27*) in ratio 1:1. The mixture was stirred 2 h on the rotation machine in biophysics lab and then put on magnetic support overnight. The supernatant was isolated, NP were washed twice with 0.5 mL of MOPS (***DNA-unbound***). NP were redispersed in 50  $\mu\text{L}$  of MOPS and heated 5 min at 65 °C. The supernatant (***DNA-released***) was separated after a few min of magnetic support.

#### ***Solid phase extraction***

Preliminary experiments (selection of the cartridge). The experiments were performed with SPE C18, 100 mg, Agilent and SPE SAX, 100 mg, Agilent cartridges.

Four solutions ( $V = 100 \mu\text{L}$  each) that contain 10  $\mu\text{M}$  of **OR31**, **OR19** and 25  $\mu\text{M}$  of DNA were incubated for a few minutes and supplemented with 0.4  $\mu\text{L}$  of standard control (Caffeine, 5 mM

in H<sub>2</sub>O). Cartridges were equilibrated with 1 mL of ACN and 1 mL of buffer (KAsO<sub>2</sub>Me<sub>2</sub> 10 mM, KCl 10 mM, pH 7.2). Sample was added to the cartridge dropwise and the percolated fraction was collected. The cartridge was washed with 500 µL of 5% MeOH +0.1% TFA (1st washing), dropwise and then bound compounds were eluted with ACN/H<sub>2</sub>O + 0.1% TFA (3\*100µL), dropwise. All fractions were analyzed by HPLC. For all further experiments only cartridges SPE SAX, 100 mg, Agilent were used.

### **Library generation/analysis**

Solutions (300 µL) containing 15 µM of aldehydes and 95 µM of acylhydrazide (**H1** or **H2**), 5 mM of catalyst (3.17 µL **Hx**, 0.5µL **Ay**, 60 µL catalyst (aniline, 25 mM in buffer)+ 233.33 µL buffer) were prepared. 75 µL of this solution to four vials and 25 µL DNA in K100 was added (Tube 1 = no SPE, tube 2= no DNA, tube 3 = *Pu24T*, tube 4 = *dT22*). After 12 hours of incubation at room temperature without shaking 1 µL of NaOH 2M was added to stop the reaction.

The cartridge was equilibrated with 1 mL of ACN and 1 mL of buffer. The sample was percolated to the cartridge (dropwise) and washed two times with 500 µL of 5% ACN, dropwise. The bound compounds were eluted with formamide/ACN 50/50 (2\*100 µL) dropwise. The elute fractions were analyzed by LC/MS (Analyse par LC-MS (colonne Luna, Phenomenex 50\*3mm dp 3µm).

The peak of each compound was assigned by LC/MS from control sample (no DNA, no SPE). Elutes were analyzed by HPLC. The areas of corresponding peaks eluted from systems with different targets (G4 or ds/ss control) were compared and the compounds presented in higher quantity in libraries templated with G4 were selected as hits.



# ***REFERENCES***



- (1) Bhattacharyya, D.; Arachchilage, G. M.; Basu, S. Metal Cations in G-Quadruplex Folding and Stability. *Front. Chem.* **2016**, *4* (SEP), 1–14.
- (2) Hardin, C. C.; Perry, A. G.; White, K. Thermodynamic and Kinetic Characterization of the Dissociation and Assembly of Quadruplex Nucleic Acids. *Biopolymers* **2000**, *56* (3), 147–194.
- (3) Guédin, A.; Gros, J.; Alberti, P.; Mergny, J. L. How Long Is Too Long? Effects of Loop Size on G-Quadruplex Stability. *Nucleic Acids Res.* **2010**, *38* (21), 7858–7868.
- (4) Puig Lombardi, E.; Londoño-Vallejo, A. A Guide to Computational Methods for G-Quadruplex Prediction. *Nucleic Acids Res.* **2020**, *48* (1), 1–15.
- (5) Todd, A. K.; Johnston, M.; Neidle, S. Highly Prevalent Putative Quadruplex Sequence Motifs in Human DNA. *Nucleic Acids Res.* **2005**, *33* (9), 2901–2907.
- (6) Belmonte-Reche, E.; Morales, J. C. G4-IM Grinder: When Size and Frequency Matter. G-Quadruplex, i-Motif and Higher Order Structure Search and Analysis Tool. *bioRxiv* **2019**, 532382.
- (7) Berselli, M.; Lavezzo, E.; Toppo, S. QPARSE: Searching for Long-Looped or Multimeric G-Quadruplexes Potentially Distinctive and Druggable. *Bioinformatics* **2020**, *36* (2), 393–399.
- (8) 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.
- (9) Rhodes, D.; Lipps, H. J. Survey and Summary G-Quadruplexes and Their Regulatory Roles in Biology. *Nucleic Acids Res.* **2015**, *43* (18), 8627–8637.
- (10) Norseen, J.; Johnson, F. B.; Lieberman, P. M. Role for G-Quadruplex RNA Binding by Epstein-Barr Virus Nuclear Antigen 1 in DNA Replication and Metaphase Chromosome Attachment. *J. Virol.* **2009**, *83* (20), 10336–10346.
- (11) Rawal, P.; Bhadra, V.; Kummarasetti, R.; Ravindran, J.; Kumar, N.; Halder, K.; Sharma, R.; Mukerji, M.; Das, S. K.; Chowdhury, S. Genome-Wide Prediction of G4 DNA as Regulatory Motifs : Role in Escherichia Coli Global Regulation. *Genome Res.* **2006**, *16*, 644–655.
- (12) Schaffitzel, C.; Berger, I.; Postberg, J.; Hanes, J.; Lipps, H. J.; Plückthun, A. In Vitro Generated Antibodies Specific for Telomeric Guanine-Quadruplex DNA React with Stylonychia Lemnae Macronuclei. *Proc. Natl. Acad. Sci. U. S. A.* **2001**, *98* (15), 8572–8577.
- (13) Biffi, G.; Tannahill, D.; McCafferty, J.; Balasubramanian, S. Quantitative Visualization of DNA G-Quadruplex Structures in Human Cells. *Nat. Chem.* **2013**, *5* (3), 182–186.
- (14) Koepfel, F.; Riou, J. F.; Laoui, A.; Mailliet, P.; Arimondo, P. B.; Labit, D.; Petitgenet, O.; Hélène, C.; Mergny, J. L. Ethidium Derivatives Bind to G-Quartets, Inhibit Telomerase and Act as Fluorescent Probes for Quadruplexes. *Nucleic Acids Res.* **2001**, *29* (5), 1087–1096.
- (15) Wang, Q.; Liu, J. Q.; Chen, Z.; Zheng, K. W.; Chen, C. Y.; Hao, Y. H.; Tan, Z. G-Quadruplex Formation at the 3' End of Telomere DNA Inhibits Its Extension by Telomerase, Polymerase

- and Unwinding by Helicase. *Nucleic Acids Res.* **2011**, *39* (14), 6229–6237.
- (16) Wang, Y.; Patel, D. J. Solution Structure of the Human Telomeric Repeat d[AG3(T2AG3)3] G-Tetraplex. *Structure* **1993**, *1* (4), 263–282.
- (17) Phan, A. T.; Kuryavyi, V.; Luu, K. N.; Patel, D. J. Structure of Two Intramolecular G-Quadruplexes Formed by Natural Human Telomere Sequences in K<sup>+</sup> Solution. *Nucleic Acids Res.* **2007**, *35* (19), 6517–6525.
- (18) Xue, Y.; Kan, Z. Y.; Wang, Q.; Yao, Y.; Liu, J.; Hao, Y. H.; Tan, Z. Human Telomeric DNA Forms Parallel-Stranded Intramolecular G-Quadruplex in K<sup>+</sup> Solution under Molecular Crowding Condition. *J. Am. Chem. Soc.* **2007**, *129* (36), 11185–11191.
- (19) Parkinson, G. N.; Lee, M. P. H.; Neidle, S. Crystal Structure of Parallel Quadruplexes from Human Telomeric DNA. *Nature* **2002**, *417* (6891), 876–880.
- (20) Luu, K. N.; Phan, A. T.; Kuryavyi, V.; Lacroix, L.; Patel, D. J. Structure of the Human Telomere in K<sup>+</sup> Solution: An Intramolecular (3 + 1) G-Quadruplex Scaffold. *J. Am. Chem. Soc.* **2006**, *128* (30), 9963–9970.
- (21) Mendez-bermudez, A.; Hills, M.; Pickett, H. A.; Phan, A. T.; Mergny, J. L.; Riou, J. F.; Royle, N. J. Human Telomeres That Contain (CTAGGG)<sub>n</sub> Repeats Show Replication Dependent Instability in Somatic Cells and the Male Germline. *Nucleic Acids Res.* **2009**, *37* (18), 6225–6238.
- (22) Lim, K. W.; Alberti, P.; Guédin, A.; Lacroix, L.; Riou, J. F.; Royle, N. J.; Mergny, J. L.; Phan, A. T. Sequence Variant (CTAGGG)<sub>n</sub> in the Human Telomere Favors a G-Quadruplex Structure Containing a G·C·G·C Tetrad. *Nucleic Acids Res.* **2009**, *37* (18), 6239–6248.
- (23) Huppert, J. L.; Balasubramanian, S. G-Quadruplexes in Promoters throughout the Human Genome. *Nucleic Acids Res.* **2007**, *35* (2), 406–413.
- (24) 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.
- (25) 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.
- (26) Kumar, N.; Patowary, A.; Sivasubbu, S.; Petersen, M.; Maiti, S. Silencing C-MYC Expression by Targeting Quadruplex in P1 Promoter Using Locked Nucleic Acid Trap. *Biochemistry* **2008**, *47* (50), 13179–13188.
- (27) Hanahan, D.; Weinberg, R. A. Review Hallmarks of Cancer : The Next Generation. *Cell* **2011**, *144* (5), 646–674.
- (28) Le, H. T.; Miller, M. C.; Buscaglia, R.; Dean, W. L.; Holt, P. A.; Chaires, J. B.; Trent, J. O. Not All G-Quadruplexes Are Created Equally: An Investigation of the Structural Polymorphism of the c-Myc G-Quadruplex-Forming Sequence and Its Interaction with the Porphyrin TMPyP4. *Org. Biomol. Chem.* **2012**, *10* (47), 9393–9404.

- (29) Seenisamy, J.; Rezler, E. M.; Powell, T. J.; Tye, D.; Gokhale, V.; Joshi, C. S.; Siddiqui-Jain, A.; Hurley, L. H. The Dynamic Character of the G-Quadruplex Element in the c-MYC Promoter and Modification by TMPyP4. *J. Am. Chem. Soc.* **2004**, *126* (28), 8702–8709.
- (30) 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.
- (31) 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.
- (32) Downward, J. Targeting RAS Signalling Pathways in Cancer Therapy. *Nat. Rev. Cancer* **2003**, *3* (1), 11–22.
- (33) Canon, J.; Rex, K.; Saiki, A. Y.; Mohr, C.; Cooke, K.; Bagal, D.; Gaida, K.; Holt, T.; Knutson, C. G.; Koppada, N.; et al. The Clinical KRAS(G12C) Inhibitor AMG 510 Drives Anti-Tumour Immunity. *Nature* **2019**, *575* (7781), 217–223.
- (34) Cogoi, S.; Paramasivam, M.; Spolaore, B.; Xodo, L. E. Structural Polymorphism within a Regulatory Element of the Human KRAS Promoter: Formation of G4-DNA Recognized by Nuclear Proteins. *Nucleic Acids Res.* **2008**, *36* (11), 3765–3780.
- (35) Kerkour, A.; Marquevielle, J.; Ivashchenko, S.; Yatsunyk, L. A.; Mergny, J. L.; Salgado, G. F. High-Resolution Three-Dimensional NMR Structure of the KRAS Proto-Oncogene Promoter Reveals Key Features of a G-Quadruplex Involved in Transcriptional Regulation. *J. Biol. Chem.* **2017**, *292* (19), 8082–8091.
- (36) Huppert, J. L.; Bugaut, A.; Kumari, S.; Balasubramanian, S. G-Quadruplexes: The Beginning and End of UTRs. *Nucleic Acids Res.* **2008**, *36* (19), 6260–6268.
- (37) Sauer, M.; Juranek, S. A.; Marks, J.; De Magis, A.; Kazemier, H. G.; Hilbig, D.; Benhalevy, D.; Wang, X.; Hafner, M.; Paeschke, K. DHX36 Prevents the Accumulation of Translationally Inactive MRNAs with G4-Structures in Untranslated Regions. *Nat. Commun.* **2019**, *10* (1).
- (38) Lista, M. J.; Martins, R. P.; Billant, O.; Contesse, M. A.; Findakly, S.; Pochard, P.; Daskalogianni, C.; Beauvineau, C.; Guetta, C.; Jamin, C.; et al. Nucleolin Directly Mediates Epstein-Barr Virus Immune Evasion through Binding to G-Quadruplexes of EBNA1 mRNA. *Nat. Commun.* **2017**, *8* (May).
- (39) Granzhan, A.; Martins, R. P.; Fåhraeus, R.; Blondel, M.; Teulade-Fichou, M. P. *Quadruplex-Interacting Compounds for Regulating the Translation of the Epstein-Barr Virus Nuclear Antigen 1 (EBNA1) mRNA: A New Strategy to Prevent and Treat EBV-Related Cancers*, 1st ed.; Elsevier Inc., 2020; Vol. 54.
- (40) Lista, M. J.; Martins, R. P.; Angrand, G.; Quillévéré, A.; Daskalogianni, C.; Voisset, C.; Teulade-Fichou, M. P.; Fåhraeus, R.; Blondel, M. A Yeast Model for the Mechanism of the Epstein-Barr Virus Immune Evasion Identifies a New Therapeutic Target to Interfere with the Virus Stealthiness. *Microb. Cell* **2017**, *4* (9), 305–307.

- (41) Murat, P.; Zhong, J.; Lekieffre, L.; Cowieson, N. P.; Clancy, J. L.; Preiss, T.; Balasubramanian, S.; Khanna, R.; Tellam, J. G-Quadruplexes Regulate Epstein-Barr Virus– Encoded Nuclear Antigen 1 mRNA Translation. *Nat. Chem. Biol.* **2014**, *10*, 358–364.
- (42) Monchaud, D.; Teulade-Fichou, M. P. A Hitchhiker’s Guide to G-Quadruplex Ligands. *Org. Biomol. Chem.* **2008**, *6* (4), 627–636.
- (43) Ohnmacht, S. A.; Neidle, S. Small-Molecule Quadruplex-Targeted Drug Discovery. *Bioorganic Med. Chem. Lett.* **2014**, *24* (12), 2602–2612.
- (44) Duarte, A. R.; Cadoni, E.; Ressurreição, A. S.; Moreira, R.; Paulo, A. Design of Modular G-Quadruplex Ligands. *ChemMedChem* **2018**, *13* (9), 869–893.
- (45) Sun, D.; Thompson, B.; Cathers, B. E.; Salazar, M.; Kerwin, S. M.; Trent, J. O.; Jenkins, T. C.; Neidle, S.; Hurley, L. H. Inhibition of Human Telomerase by a G-Quadruplex-Interactive Compound. *J. Med. Chem.* **1997**, *40* (14), 2113–2116.
- (46) 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.
- (47) Harrison, R. J.; Cuesta, J.; Chessari, G.; Read, M. A.; Basra, S. K.; Reszka, A. P.; Morrell, J.; Gowan, S. M.; Incles, C. M.; Tanious, F. A.; et al. Trisubstituted Acridine Derivatives as Potent and Selective Telomerase Inhibitors. *J. Med. Chem.* **2003**, *46* (21), 4463–4476.
- (48) Cookson, J. C.; Dai, F.; Smith, V.; Heald, R. a; Laughton, C. a; Stevens, M. F. G.; Burger, A. M. Pharmacodynamics of the G-Quadruplex-Stabilizing in Human Tumor Cells Correlates with Telomere Length and Can Be Enhanced , or Antagonized , with Cytotoxic Agents. *Mol. Pharmacol.* **2005**, *68* (6), 1551–1558.
- (49) 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 Res.* **2005**, *65* (4), 1489–1496.
- (50) Perrone, R.; Butovskaya, E.; Daelemans, D.; Palù, G.; Pannecouque, C.; Richter, S. N. Anti-HIV-1 Activity of the G-Quadruplex Ligand BRACO-19. *J. Antimicrob. Chemother.* **2014**, *69* (12), 3248–3258.
- (51) Gilbert-Girard S, Gravel A, Artusi S, Richter SN, Wallaschek N, Kaufer BB, F. L. Stabilization of Telomere G-Quadruplexes Interferes with Human Herpesvirus 6A Chromosomal Integration. *J. Virol.* **2017**, *91* (14), 1–14.
- (52) Bonnat, L.; Bar, L.; Génaro, B.; Bonnet, H.; Jarjayes, O.; Thomas, F.; Dejeu, J.; Defrancq, E.; Lavergne, T. Template-Mediated Stabilization of a DNA G-Quadruplex Formed in the HIV-1 Promoter and Comparative Binding Studies. *Chem. - A Eur. J.* **2017**, *23* (23), 5602–5613.
- (53) Bonnat, L.; Dautriche, M.; Saidi, T.; Revol-Cavalier, J.; Dejeu, J.; Defrancq, E.; Lavergne, T. Scaffold Stabilization of a G-Triplex and Study of Its Interactions with G-Quadruplex

Targeting Ligands. *Org. Biomol. Chem.* **2019**, *17* (38), 8726–8736.

- (54) Campbell, N. H.; Parkinson, G. N.; Reszka, A. P.; Neidle, S. Structural Basis of DNA Quadruplex Recognition by an Acridine Drug. *J. Am. Chem. Soc.* **2008**, *130* (21), 6722–6724.
- (55) Pennarun, G.; Granotier, C.; Gauthier, L. R.; Gomez, D.; Hoffschir, F.; Mandine, E.; Riou, J. F.; Mergny, J. L.; Mailliet, P.; Boussin, F. D. Apoptosis Related to Telomere Instability and Cell Cycle Alterations in Human Glioma Cells Treated by New Highly Selective G-Quadruplex Ligands. *Oncogene* **2005**, *24* (18), 2917–2928.
- (56) Riou, L.; Granotier, C.; Gauthier, L. R.; Cian, A. De; Gomez, D.; Mandine, E. Preferential Binding of a G-Quadruplex Ligand to Human Chromosome Ends. *Nucleic Acids Res.* **2005**, *33* (13), 4182–4190.
- (57) Verga, D.; Hamon, F.; Poyer, F.; Bombard, S.; Teulade-Fichou, M. P. Photo-Cross-Linking Probes for Trapping G-Quadruplex DNA. *Angew. Chemie - Int. Ed.* **2014**, *53* (4), 994–998.
- (58) Betzer, J. F.; Nuter, F.; Chtchigrovsky, M.; Hamon, F.; Kellermann, G.; Ali, S.; Calm ejane, M. A.; Roque, S.; Poupon, J.; Cresteil, T.; et al. Linking of Antitumor Trans NHC-Pt(II) Complexes to G-Quadruplex DNA Ligand for Telomeric Targeting. *Bioconjug. Chem.* **2016**, *27* (6), 1456–1470.
- (59) Yang, P.; De Cian, A.; Teulade-Fichou, M. P.; Mergny, J. L.; Monchaud, D. Engineering Bisquinolinium/Thiazole Orange Conjugates for Fluorescent Sensing of G-Quadruplex DNA. *Angew. Chemie - Int. Ed.* **2009**, *48* (12), 2188–2191.
- (60) Xie, X.; Reznichenko, O.; Chaput, L.; Martin, P.; Teulade-Fichou, M.-P.; Granzhan, A. Topology-Selective, Fluorescent “Light-Up” Probes for G-Quadruplex DNA Based on Photoinduced Electron Transfer. *Chem. - A Eur. J.* **2018**, *24* (48), 12638–12651.
- (61) Cian, A. De; Delemos, E.; Mergny, J.; Monchaud, D. Highly Efficient G-Quadruplex Recognition by Bisquinolinium Compounds. *J. Am. Chem. Soc.* **2007**, *129* (7), 1856–1857.
- (62) Chung, W. J.; Heddi, B.; Hamon, F.; Teulade-Fichou, M. P.; Phan, A. T. Solution Structure of a G-Quadruplex Bound to the Bisquinolinium Compound Phen-DC3. *Angew. Chemie - Int. Ed.* **2014**, *53* (4), 999–1002.
- (63) Lopes, J.; Piazza, A. Le; Bermejo, R.; Kriegsman, B.; Colosio, A.; Teulade-Fichou, M. P.; Foiani, M.; Nicolas, A. G-Quadruplex-Induced Instability during Leading-Strand Replication. *EMBO J.* **2011**, *30* (19), 4033–4046.
- (64) De Cian, A.; Cristofari, G.; Reichenbach, P.; De Lemos, E.; Monchaud, D.; Teulade-Fichou, M. P.; Shin-ya, K.; Lacroix, L.; Lingner, J.; Mergny, J. L. Reevaluation of Telomerase Inhibition by Quadruplex Ligands and Their Mechanisms of Action. *Proc. Natl. Acad. Sci. U. S. A.* **2007**, *104* (44), 17347–17352.
- (65) Rodriguez, R.; M uller, 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.
- (66) Daniel D. Le, Marco Di Antonio, Louis K. M. Chan, and S. B. G-Quadruplex Ligands Exhibit

- Differential G-Tetrad Selectivity. *Chem. Commun.* **2015**, 51 (38), 8048–8050.
- (67) Müller, S.; Kumari, S.; Rodriguez, R.; Balasubramanian, S. Small-Molecule-Mediated G-Quadruplex Isolation from Human Cells. *Nat. Chem.* **2010**, 2 (12), 1095–1098.
- (68) Dhamodharan, V.; Harikrishna, S.; Jagadeeswaran, C.; Halder, K.; Pradeepkumar, P. I. Selective G-Quadruplex DNA Stabilizing Agents Based on Bisquinolinium and Bispyridinium Derivatives of 1,8-Naphthyridine. *J. Org. Chem.* **2012**, 77 (1), 229–242.
- (69) Anas, M.; Sharma, R.; Dhamodharan, V.; Pradeepkumar, P. I.; Manhas, A.; Srivastava, K.; Ahmed, S.; Kumar, N. Investigating Pharmacological Targeting of G-Quadruplexes in the Human Malaria Parasite. *Biochemistry* **2017**, 56 (51), 6691–6699.
- (70) Shan, C.; Tan, J. H.; Ou, T. M.; Huang, Z. S. Natural Products and Their Derivatives as G-Quadruplex Binding Ligands. *Sci. China Chem.* **2013**, 56 (10), 1351–1363.
- (71) Brassart, B.; Gomez, D.; De Cian, A.; Paterski, R.; Montagnac, A.; Qui, K. H.; Temime-Smaali, N.; Trentesaux, C.; Mergny, J. L.; Gueritte, F.; et al. A New Steroid Derivative Stabilizes G-Quadruplexes and Induces Telomere Uncapping in Human Tumor Cells. *Mol. Pharmacol.* **2007**, 72 (3), 631–640.
- (72) Cocco, M. J.; Hanakahi, L. A.; Huber, M. D.; Maizels, N. Specific Interactions of Distamycin with G-Quadruplex DNA. *Nucleic Acids Res.* **2003**, 31 (11), 2944–2951.
- (73) Li, J.; Jin, X.; Hu, L.; Wang, J.; Su, Z. Identification of Nonplanar Small Molecule for G-Quadruplex Grooves: Molecular Docking and Molecular Dynamic Study. *Bioorganic Med. Chem. Lett.* **2011**, 21 (23), 6969–6972.
- (74) Shin-ya, K.; Wierzba, K.; Matsuo, K. Telomestatin, a Novel Telomerase Inhibitor from *Streptomyces Anulatus*. *J. Am. Chem. Soc.* **2001**, 123 (6), 1262–1263.
- (75) Kim, M.; 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.
- (76) Doi, T.; Yoshida, M.; Shin-Ya, K.; Takahashi, T. Total Synthesis of (R)-Telomestatin. *Org. Lett.* **2006**, 8 (18), 4165–4167.
- (77) Amagai, K.; Ikeda, H.; Hashimoto, J.; Kozono, I.; Izumikawa, M. Identification of a Gene Cluster for Telomestatin Biosynthesis and Heterologous Expression Using a Specific Promoter in a Clean Host. *Sci. Rep.* **2017**, 7, 3382.
- (78) Hamon, F.; Largy, E.; Guédin-Beaurepaire, A.; Rouchon-Dagois, M.; Sidibe, A.; Monchaud, D.; Mergny, J. L.; Riou, J. F.; Nguyen, C. H.; Teulade-Fichou, M. P. An Acyclic Oligoheteroaryle That Discriminates Strongly between Diverse G-Quadruplex Topologies. *Angew. Chemie - Int. Ed.* **2011**, 50 (37), 8745–8749.
- (79) Tera, M.; Ishizuka, H.; Takagi, M.; Suganuma, M.; Shin-Ya, K.; Nagasawa, K. Macrocyclic Hexaoxazoles as Sequence- and Mode-Selective G-Quadruplex Binders. *Angew. Chemie - Int. Ed.* **2008**, 47 (30), 5557–5560.

- (80) Chung, W. J.; Heddi, B.; Tera, M.; Iida, K.; Nagasawa, K.; Phan, A. T. Solution Structure of an Intramolecular (3 + 1) Human Telomeric G-Quadruplex Bound to a Telomestatin Derivative. *J. Am. Chem. Soc.* **2013**, *135* (36), 13495–13501.
- (81) Mikuma, T.; Terui, N.; Yamamoto, Y.; Hori, H. A Novel Heme-DNA Coordination Complex and Its Stability. *Nucleic Acids Res. Suppl.* **2002**, No. 2, 285–286.
- (82) Wheelhouse, R. T.; Sun, D.; Han, H.; Han, F. X.; Hurley, L. H. Cationic Porphyrins as Telomerase Inhibitors: The Interaction of Tetra- (N-Methyl-4-Pyridyl)Porphine with Quadruplex DNA [1]. *J. Am. Chem. Soc.* **1998**, *120* (13), 3261–3262.
- (83) Tada-Oikawa, S.; Oikawa, S.; Hirayama, J.; Hirakawa, K.; Kawanishi, S. Dna Damage and Apoptosis Induced by Photosensitization of 5,10,15,20-Tetrakis (n-Methyl-4-Pyridyl)-21h,23h-Porphyrin via Singlet Oxygen Generation. *Photochem. Photobiol.* **2009**, *85* (6), 1391–1399.
- (84) Garcia-sampedro, A.; Tabero, A.; Mahamed, I.; Acedo, P. Multimodal Use of the Porphyrin TMPyP : From Cancer Therapy to Antimicrobial Applications. **2019**, No. ii, 11–27.
- (85) Morris, M. J.; Wingate, K. L.; Silwal, J.; Leeper, T. C.; Basu, S. The Porphyrin TmPyP4 Unfolds the Extremely Stable G-Quadruplex in MT3-MMP mRNA and Alleviates Its Repressive Effect to Enhance Translation in Eukaryotic Cells. *Nucleic Acids Res.* **2012**, *40* (9), 31–34.
- (86) Li, Y.; Geyer, C. R.; Sen, D. Recognition of Anionic Porphyrins by DNA Aptamers. *Biochemistry* **1996**, *35* (21), 6911–6922.
- (87) Arthanari, H.; Basu, S.; Kawano, T. L.; Bolton, P. H. Fluorescent Dyes Specific for Quadruplex DNA. *Nucleic Acids Res.* **1998**, *26* (16), 3724–3728.
- (88) Yett, A.; Lin, L. Y.; Beseiso, D.; Miao, J.; Yatsunyk, L. A. N-Methyl Mesoporphyrin IX as a Highly Selective Light-up Probe for G-Quadruplex DNA. *J. Porphyr. Phthalocyanines* **2019**, *23* (11–12), 1195–1215.
- (89) Nicoludis, J. M.; Barrett, S. P.; Mergny, J. L.; Yatsunyk, L. A. Interaction of Human Telomeric DNA with N-Methyl Mesoporphyrin IX. *Nucleic Acids Res.* **2012**, *40* (12), 5432–5447.
- (90) Perenon, M.; Bonnet, H.; Lavergne, T.; Dejeu, J.; Defrancq, E. Surface Plasmon Resonance Study of the Interaction of: N -Methyl Mesoporphyrin IX with G-Quadruplex DNA. *Phys. Chem. Chem. Phys.* **2020**, *22* (7), 4158–4164.
- (91) Nicoludis, J. M.; Miller, S. T.; Jeffrey, P. D.; Barrett, S. P.; Rablen, P. R.; Lawton, T. J.; Yatsunyk, L. A. Optimized End-Stacking Provides Specificity of N-Methyl Mesoporphyrin IX for Human Telomeric G-Quadruplex DNA. *J. Am. Chem. Soc.* **2012**, *134* (50), 20446–20456.
- (92) Duan, W.; Rangan, A.; Vankayalapati, H.; Kim, M. Y.; Zeng, Q.; Sun, D.; Han, H.; Fedoroff, O. Y.; Nishioka, D.; Rha, S. Y.; et al. Design and Synthesis of Fluoroquinophenoxazines That Interact with Human Telomeric G-Quadruplexes and Their Biological Effects. *Mol. Cancer Ther.* **2001**, *1* (2), 103–120.
- (93) Drygin, D.; Lin, A.; Bliesath, J.; Ho, C. B.; O'Brien, S. E.; Proffitt, C.; Omori, M.; Haddach, M.; Schwaebe, M. K.; Siddiqui-Jain, A.; et al. Targeting RNA Polymerase I with an Oral Small

- Molecule CX-5461 Inhibits Ribosomal RNA Synthesis and Solid Tumor Growth. *Cancer Res.* **2011**, *71* (4), 1418–1430.
- (94) Xu, H.; Di Antonio, M.; McKinney, S.; Mathew, V.; Ho, B.; O’Neil, N. J.; Santos, N. Dos; Silvester, J.; Wei, V.; Garcia, J.; et al. CX-5461 Is a DNA G-Quadruplex Stabilizer with Selective Lethality in BRCA1/2 Deficient Tumours. *Nat. Commun.* **2017**, *8* (205).
- (95) Keating, L. R.; Szalai, V. A. Parallel-Stranded Guanine Quadruplex Interactions with a Copper Cationic Porphyrin. *Biochemistry* **2004**, *43* (50), 15891–15900.
- (96) Dixon, I. M.; Lopez, F.; Tejera, A. M.; Estève, J. P.; Blasco, M. A.; Pratviel, G.; Meunier, B. A G-Quadruplex Ligand with 10000-Fold Selectivity over Duplex DNA. *J. Am. Chem. Soc.* **2007**, *129* (6), 1502–1503.
- (97) Bhattacharjee, A. J.; Ahluwalia, K.; Taylor, S.; Jin, O.; Nicoludis, J. M.; Buscaglia, R.; Brad Chaires, J.; Kornfilt, D. J. P.; Marquardt, D. G. S.; Yatsunyk, L. A. Induction of G-Quadruplex DNA Structure by Zn(II) 5,10,15,20-Tetrakis(N-Methyl-4-Pyridyl)Porphyrin. *Biochimie* **2011**, *93* (8), 1297–1309.
- (98) Ohyama, T.; Kato, Y.; Mita, H.; Yamamoto, Y. Exogenous Ligand Binding Property of a Heme-DNA Coordination Complex. *Chem. Lett.* **2006**, *35* (1), 126–127.
- (99) Terenzi, A.; Lötsch, D.; Van Schoonhoven, S.; Roller, A.; Kowol, C. R.; Berger, W.; Keppler, B. K.; Barone, G. Another Step toward DNA Selective Targeting: Nill and Cull Complexes of a Schiff Base Ligand Able to Bind Gene Promoter G-Quadruplexes. *Dalt. Trans.* **2016**, *45* (18), 7758–7767.
- (100) Reed, J. E.; Neidle, S.; Vilar, R. Stabilisation of Human Telomeric Quadruplex DNA and Inhibition of Telomerase by a Platinum-Phenanthroline Complex. *Chem. Commun.* **2007**, No. 42, 4366–4368.
- (101) Morel, E.; Beauvineau, C.; Naud-Martin, D.; Landras-Guetta, C.; Verga, D.; Ghosh, D.; Achelle, S.; Mahuteau-Betzer, F.; Bombard, S.; Teulade-Fichou, M. P. Selectivity of Terpyridine Platinum Anticancer Drugs for G-Quadruplex DNA. *Molecules* **2019**, *24* (3), 1–19.
- (102) 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), 138–148.
- (103) Calabrese, D. R.; Chen, X.; Leon, E. C.; Gaikwad, S. M.; Phyto, Z.; Hewitt, W. M.; Alden, S.; Hilimire, T. A.; He, F.; Michalowski, A. M.; et al. Chemical and Structural Studies Provide a Mechanistic Basis for Recognition of the MYC G-Quadruplex. *Nat. Commun.* **2018**, *9* (1), 1–15.
- (104) Sparapani, S.; Bellini, S.; Gunaratnam, M.; Haider, S. M.; Andreani, A.; Rambaldi, M.; Locatelli, A.; Morigi, R.; Granaiola, M.; Varoli, L.; et al. Bis-Guanylhydrazone Diimidazo[1,2-a:1,2-c]Pyrimidine as a Novel and Specific G-Quadruplex Binding Motif. *Chem. Commun.* **2010**, *46* (31), 5680–5682.



- (105) Amato, J.; Morigi, R.; Pagano, B.; Pagano, A.; Ohnmacht, S.; De Magis, A.; Tiang, Y. P.; Capranico, G.; Locatelli, A.; Graziadio, A.; et al. Toward the Development of Specific G-Quadruplex Binders: Synthesis, Biophysical, and Biological Studies of New Hydrazone Derivatives. *J. Med. Chem.* **2016**, *59* (12), 5706–5720.
- (106) Amato, J.; Miglietta, G.; Morigi, R.; Iaccarino, N.; Locatelli, A.; Leoni, A.; Novellino, E.; Pagano, B.; Capranico, G.; Randazzo, A. Monohydrazone Based G-Quadruplex Selective Ligands Induce DNA Damage and Genome Instability in Human Cancer Cells. *J. Med. Chem.* **2020**, *63* (6), 3090–3103.
- (107) Shan, C.; Yan, J. W.; Wang, Y. Q.; Che, T.; Huang, Z. L.; Chen, A. C.; Yao, P. F.; Tan, J. H.; Li, D.; Ou, T. M.; et al. Design, Synthesis, and Evaluation of Isaindigotone Derivatives To Downregulate c-Myc Transcription via Disrupting the Interaction of NM23-H2 with G-Quadruplex. *J. Med. Chem.* **2017**, *60* (4), 1292–1308.
- (108) Wang, Y. Q.; Huang, Z. L.; Chen, S. Bin; Wang, C. X.; Shan, C.; Yin, Q. K.; Ou, T. M.; Li, D.; Gu, L. Q.; Tan, J. H.; et al. Design, Synthesis, and Evaluation of New Selective NM23-H2 Binders as c-MYC Transcription Inhibitors via Disruption of the NM23-H2/G-Quadruplex Interaction. *J. Med. Chem.* **2017**, *60* (16), 6924–6941.
- (109) Whitesides, G. M. The Origins and the Future of Microfluidics. *Nature* **2006**, *442* (7101), 368–373.
- (110) Liu, Y.; Jiang, X. Why Microfluidics? Merits and Trends in Chemical Synthesis. *Lab Chip* **2017**, *17* (23), 3960–3978.
- (111) Rakers, V.; Cadinu, P.; Edel, J. B.; Vilar, R. Development of Microfluidic Platforms for the Synthesis of Metal Complexes and Evaluation of Their DNA Affinity Using Online FRET Melting Assays. *Chem. Sci.* **2018**, *9* (14), 3459–3469.
- (112) Litovchick, A.; Tian, X.; Monteiro, M. I.; Kennedy, K. M.; Guié, M. A.; Centrella, P.; Zhang, Y.; Clark, M. A.; Keefe, A. D. Novel Nucleic Acid Binding Small Molecules Discovered Using DNA-Encoded Chemistry. *Molecules* **2019**, *24* (10), 1–14.
- (113) Liu, K. C.; Der, K. R.; Mayer, C.; Adhikari, S.; Wales, D. J. Affinity-Selected Bicyclic Peptide G-Quadruplex Ligands Mimic a Protein-like Binding Mechanism. *J. Am. Chem. Soc.* **2020**, *142*, 8367–8373.
- (114) Fesik, S. W. Discovering High-Affinity Ligands for Proteins : SAR by NMR Author ( s ): Suzanne B . Shuker , Philip J . Hajduk , Robert P . Meadows and Stephen W . Fesik Published by : American Association for the Advancement of Science Stable URL : [Http://Www.Jstor.Or.](http://www.jstor.org) *Science* (80- ). **1996**, *274* (5292), 1531–1534.
- (115) Whittaker, M.; Law, R. J.; Ichihara, O.; Hestekamp, T.; Hallett, D. Fragments: Past, Present and Future. *Drug Discov. Today Technol.* **2010**, *7* (3), e163–e171.
- (116) Tassinari, M.; Lena, A.; Butovskaya, E.; Pirola, V.; Nadai, M.; Freccero, M.; Doria, F.; Richter, S. N. A Fragment-Based Approach for the Development of g-Quadruplex Ligands: Role of the Amidoxime Moiety. *Molecules* **2018**, *23* (8).

- (117) Bosc, D.; Jakhlal, J.; Deprez, B.; Deprez-Poulain, R. Kinetic Target-Guided Synthesis in Drug Discovery and Chemical Biology: A Comprehensive Facts and Figures Survey. *Future Med. Chem.* **2016**, *8* (4), 381–404.
- (118) Di Antonio, M.; Biffi, G.; Mariani, A.; Raiber, E. A.; Rodriguez, R.; Balasubramanian, S. Selective RNA versus DNA G-Quadruplex Targeting by Situ Click Chemistry. *Angew. Chemie - Int. Ed.* **2012**, *51* (44), 11073–11078.
- (119) Rocca, R.; Talarico, C.; Moraca, F.; Costa, G.; Romeo, I.; Ortuso, F.; Alcaro, S.; Artese, A. Molecular Recognition of a Carboxy Pyridostatin toward G-Quadruplex Structures: Why Does It Prefer RNA? *Chem. Biol. Drug Des.* **2017**, *90* (5), 919–925.
- (120) Panda, D.; Saha, P.; Das, T.; Dash, J. Target Guided Synthesis Using DNA Nano-Templates for Selectively Assembling a G-Quadruplex Binding c-MYC Inhibitor. *Nat. Commun.* **2017**, *8* (May), 1–11.
- (121) Lehn, J. M. Dynamic Combinatorial Chemistry and Virtual Combinatorial Libraries. *Chem. - A Eur. J.* **1999**, *5* (9), 2455–2463.
- (122) Frei, P.; Hevey, R.; Ernst, B. Dynamic Combinatorial Chemistry: A New Methodology Comes of Age. *Chem. - A Eur. J.* **2019**, *25* (1), 60–73.
- (123) Hartman, A. M.; Gierse, R. M.; Hirsch, A. K. H. Protein-Templated Dynamic Combinatorial Chemistry: Brief Overview and Experimental Protocol. *European J. Org. Chem.* **2019**, *2019* (22), 3581–3590.
- (124) Whitney, A. M.; Ladame, S.; Balasubramanian, S. Templated Ligand Assembly by Using G-Quadruplex DNA and Dynamic Covalent Chemistry. *Angew. Chemie - Int. Ed.* **2004**, *43* (9), 1143–1146.
- (125) Ladame, S.; Whitney, A. M.; Balasubramanian, S. Targeting Nucleic Acid Secondary Structures with Polyamides Using an Optimized Dynamic Combinatorial Approach. *Angew. Chemie - Int. Ed.* **2005**, *44* (35), 5736–5739.
- (126) Bugaut, A.; Jantos, K.; Wietor, J. L.; Rodriguez, R.; Sanders, J. K. M.; Balasubramanian, S. Exploring the Differential Recognition of DNA G-Quadruplex Targets by Small Molecules Using Dynamic Combinatorial Chemistry. *Angew. Chemie - Int. Ed.* **2008**, *47* (14), 2677–2680.
- (127) Nielsen, M. C.; Ulven, T. Selective Extraction of G-Quadruplex Ligands from a Rationally Designed Scaffold-Based Dynamic Combinatorial Library. *Chem. - A Eur. J.* **2008**, *14* (31), 9487–9490.
- (128) Jana, S.; Panda, D.; Saha, P.; Pantos, G. D.; Dash, J. Dynamic Generation of G-Quadruplex DNA Ligands by Target-Guided Combinatorial Chemistry on a Magnetic Nanoplatfom. *J. Med. Chem.* **2019**, *62* (2), 762–773.
- (129) Circular Dichroism Units and Conversions <https://www.photophysics.com/circular-dichroism/chirscan-technology/circular-dichroism-spectroscopy-units-conversions/>.
- (130) Masiero, S.; Trotta, R.; Pieraccini, S.; De Tito, S.; Perone, R.; Randazzo, A.; Spada, G. P. A

- Non-Empirical Chromophoric Interpretation of CD Spectra of DNA G-Quadruplex Structures. *Org. Biomol. Chem.* **2010**, *8* (12), 2683–2692.
- (131) Waller, Z. A. E.; Sewitz, S. A.; Hsu, S. T. D.; Balasubramanian, S. A Small Molecule That Disrupts G-Quadruplex DNA Structure and Enhances Gene Expression. *J. Am. Chem. Soc.* **2009**, *131* (35), 12628–12633.
- (132) Carvalho, J.; Queiroz, J. A.; Cruz, C. Circular Dichroism of G-Quadruplex: A Laboratory Experiment for the Study of Topology and Ligand Binding. *J. Chem. Educ.* **2017**, *94* (10), 1547–1551.
- (133) Marchand, A.; Rosu, F.; Zenobi, R.; Gabelica, V. Thermal Denaturation of DNA G-Quadruplexes and Their Complexes with Ligands: Thermodynamic Analysis of the Multiple States Revealed by Mass Spectrometry. *J. Am. Chem. Soc.* **2018**, *140* (39), 12553–12565.
- (134) De Cian, A.; Guittat, L.; Kaiser, M.; Saccà, B.; Amrane, S.; Bourdoncle, A.; Alberti, P.; Teulade-Fichou, M. P.; Lacroix, L.; Mergny, J. L. Fluorescence-Based Melting Assays for Studying Quadruplex Ligands. *Methods* **2007**, *42* (2), 183–195.
- (135) Morgan, R. K.; Psaras, A. M.; Lassiter, Q.; Raymer, K.; Brooks, T. A. G-Quadruplex Deconvolution with Physiological Mimicry Enhances Primary Screening: Optimizing the FRET Melt2 Assay. *Biochim. Biophys. Acta - Gene Regul. Mech.* **2020**, *1863* (1), 194478.
- (136) Monchaud, D.; Allain, C.; Bertrand, H.; Smargiasso, N.; Rosu, F.; Gabelica, V.; De Cian, A.; Mergny, J. L.; Teulade-Fichou, M. P. Ligands Playing Musical Chairs with G-Quadruplex DNA: A Rapid and Simple Displacement Assay for Identifying Selective G-Quadruplex Binders. *Biochimie* **2008**, *90* (8), 1207–1223.
- (137) Zuffo, M.; Guédin, A.; Leriche, E. D.; Doria, F.; Pirota, V.; Gabelica, V.; Mergny, J. L.; Freccero, M. More Is Not Always Better: Finding the Right Trade-off between Affinity and Selectivity of a G-Quadruplex Ligand. *Nucleic Acids Res.* **2018**, *46* (19).
- (138) Tanpure, A. A.; Srivatsan, S. G. Conformation-Sensitive Nucleoside Analogues as Topology-Specific Fluorescence Turn-on Probes for DNA and RNA G-Quadruplexes. *Nucleic Acids Res.* **2015**, *43* (22).
- (139) Sontakke, V. A.; Srivatsan, S. G. A Dual-App Nucleoside Probe Reports G-Quadruplex Formation and Ligand Binding in the Long Terminal Repeat of HIV-1 Proviral Genome. *Bioorganic Med. Chem. Lett.* **2020**, *30* (16), 127345.
- (140) Marchand, A.; Gabelica, V. Native Electrospray Mass Spectrometry of DNA G-Quadruplexes in Potassium Solution. *J. Am. Soc. Mass Spectrom.* **2014**, *25* (7), 1146–1154.
- (141) Lecours, M. J.; Marchand, A.; Anwar, A.; Guetta, C.; Hopkins, W. S.; Gabelica, V. What Stoichiometries Determined by Mass Spectrometry Reveal about the Ligand Binding Mode to G-Quadruplex Nucleic Acids. *Biochim. Biophys. Acta - Gen. Subj.* **2017**, *1861* (5), 1353–1361.
- (142) Marchand, A.; Granzhan, A.; Iida, K.; Tsushima, Y.; Ma, Y.; Nagasawa, K.; Teulade-Fichou, M. P.; Gabelica, V. Ligand-Induced Conformational Changes with Cation Ejection upon

- Binding to Human Telomeric DNA G-Quadruplexes. *J. Am. Chem. Soc.* **2015**, *137* (2), 750–756.
- (143) Bhat, V. T.; Caniard, A. M.; Luksch, T.; Brenk, R.; Campopiano, D. J.; Greaney, M. F. Nucleophilic Catalysis of Acylhydrazone Equilibration for Protein-Directed Dynamic Covalent Chemistry. *Nat. Chem.* **2010**, *2* (6), 490–497.
- (144) Rollas, S.; Küçükgülzel, Ş. G. Biological Activities of Hydrazone Derivatives. *Molecules* **2007**, *12* (8), 1910–1939.
- (145) Thota, S.; Rodrigues, D. A.; Pinheiro, P. de S. M.; Lima, L. M.; Fraga, C. A. M.; Barreiro, E. J. N-Acylhydrazones as Drugs. *Bioorganic Med. Chem. Lett.* **2018**, *28* (17), 2797–2806.
- (146) Hartman, A. M.; Elgaher, W. A. M.; Hertrich, N.; Andrei, S. A.; Ottmann, C.; Hirsch, A. K. H. Discovery of Small-Molecule Stabilizers of 14-3-3 Protein-Protein Interactions via Dynamic Combinatorial Chemistry. *ACS Med. Chem. Lett.* **2020**, *11* (5), 1041–1046.
- (147) Newkome, G. R.; Theriot, K. J.; Majestic, V.; Spruell, P. A.; Baker, G. R. Functionalization of 2-Methyl- and 2,7-Dimethyl-1,8-Naphthyridine. *J. Org. Chem.* **1990**, *55* (9), 2838–2842.
- (148) Christopher J. Chandler, L. W. D. and J. A. R. Synthesis of Some 2,9-Disubstituted-1,10-Phenanthrolines. *J. Heterocycl. Chem.* **1981**, *18* (May), 599.
- (149) Plater, M. J.; Barnes, P.; McDonald, L. K.; Wallace, S.; Archer, N.; Gelbrich, T.; Horton, P. N.; Hursthouse, M. B. Hidden Signatures: New Reagents for Developing Latent Fingerprints. *Org. Biomol. Chem.* **2009**, *7* (8), 1633–1641.
- (150) Tan, L.; Lin, W.; Zhu, S.; Yuan, L.; Zheng, K. A Coumarin-Quinolinium-Based Fluorescent Probe for Ratiometric Sensing of Sulfite in Living Cells. *Org. Biomol. Chem.* **2014**, *12* (26), 4637–4643.
- (151) Singh, P.; Kumar, R.; Singh, A. K.; Yadav, P.; Khanna, R. S.; Vinayak, M.; Tewari, A. K. Synthesis and Crystal Structure of Quinolinium Salt: Assignment on Nonsteroidal Anti-Inflammatory Activity and DNA Cleavage Activity. *J. Mol. Struct.* **2018**, *1163*, 262–269.
- (152) Santhiya, K.; Sen, S. K.; Natarajan, R.; Shankar, R.; Murugesapandian, B. D-A-D Structured Bis-Acylhydrazone Exhibiting Aggregation-Induced Emission, Mechanochromic Luminescence, and Al(III) Detection. *J. Org. Chem.* **2018**, *83* (18), 10770–10775.
- (153) Monsen, R. C.; Trent, J. O. G-Quadruplex Virtual Drug Screening: A Review. *Biochimie* **2018**, *152*, 134–148.
- (154) Wirmer-Bartoschek, J.; Bendel, L. E.; Jonker, H. R. A.; Grün, J. T.; Papi, F.; Bazzicalupi, C.; Messori, L.; Gratteri, P.; Schwalbe, H. Solution NMR Structure of a Ligand/Hybrid-2-G-Quadruplex Complex Reveals Rearrangements That Affect Ligand Binding. *Angew. Chemie - Int. Ed.* **2017**, *56* (25), 7102–7106.
- (155) Dai, J.; Carver, M.; Hurley, L. H.; Yang, D. Solution Structure of a 2:1 Quindoline-c-MYC G-Quadruplex: Insights into G-Quadruplex-Interactive Small Molecule Drug Design. *J. Am. Chem. Soc.* **2011**, *133* (44), 17673–17680.

- (156) Ladame, S. Dynamic Combinatorial Chemistry: On the Road to Fulfilling the Promise. *Org. Biomol. Chem.* **2008**, *6* (2), 219–226.
- (157) Ramström, O.; Lehn, J.-M. Drug Discovery by Dynamic Combinatorial Libraries. *Nat. Rev. Drug Discov.* **2002**, *1* (1), 26–36.
- (158) Mondal, M.; Hirsch, A. K. H. Dynamic Combinatorial Chemistry: A Tool to Facilitate the Identification of Inhibitors for Protein Targets. *Chem. Soc. Rev.* **2015**, *44* (8), 2455–2488.
- (159) Durai, C. R. S.; Harding, M. M. Targeting Nucleic Acids Using Dynamic Combinatorial Chemistry. *Aust. J. Chem.* **2011**, *64* (6), 671–680.
- (160) Miller, B. L. DCC in the Development of Nucleic Acid Targeted and Nucleic Acid Inspired Structures. *Top. Curr. Chem.* **2011**, *322*, 107–137.
- (161) Nielsen, M. C.; Ulven, T. Selective Extraction of G-Quadruplex Ligands from a Rationally Designed Scaffold-Based Dynamic Combinatorial Library. *Chem. - A Eur. J.* **2008**, *14* (31), 9487–9490.
- (162) Cordes, E. H.; Jencks, W. P. Nucleophilic Catalysis of Semicarbazone Formation by Anilines. *J. Am. Chem. Soc.* **1962**, *84* (5), 826–831.
- (163) Crisalli, P.; Kool, E. T. Water-Soluble Organocatalysts for Hydrazone and Oxime Formation. *J. Org. Chem.* **2013**, *78* (3), 1184–1189.
- (164) Wendeler, M.; Grinberg, L.; Wang, X.; Dawson, P. E.; Baca, M. Enhanced Catalysis of Oxime-Based Bioconjugations by Substituted Anilines. *Bioconjug. Chem.* **2014**, *25* (1), 93–101.
- (165) Larsen, D.; Pittelkow, M.; Karmakar, S.; Kool, E. T. New Organocatalyst Scaffolds with High Activity in Promoting Hydrazone and Oxime Formation at Neutral PH. *Org. Lett.* **2015**, *17* (2), 274–277.
- (166) Clipson, A. J.; Bhat, V. T.; McNae, I.; Caniard, A. M.; Campopiano, D. J.; Greaney, M. F. Bivalent Enzyme Inhibitors Discovered Using Dynamic Covalent Chemistry. *Chem. - A Eur. J.* **2012**, *18* (34), 10562–10570.
- (167) Reznichenko, O.; Quillévéré, A.; Martins, R. P.; Loaëc, N.; Kang, H.; Lista, M. J.; Beauvineau, C.; González-García, J.; Guillot, R.; Voisset, C.; et al. Novel Cationic Bis(Acylhydrazones) as Modulators of Epstein–Barr Virus Immune Evasion Acting through Disruption of Interaction between Nucleolin and G-Quadruplexes of EBNA1 MRNA. *Eur. J. Med. Chem.* **2019**, *178*, 13–29.
- (168) Miller, M. C.; Buscaglia, R.; Chaires, J. B.; Lane, A. N.; Trent, J. O. Hydration Is a Major Determinant of the G-Quadruplex Stability and Conformation of the Human Telomere 3' Sequence of d(AG 3(TTAG3)3). *J. Am. Chem. Soc.* **2010**, *132* (48), 17105–17107.
- (169) Heddi, B.; Phan, A. T. Structure of Human Telomeric DNA in Crowded Solution. *J. Am. Chem. Soc.* **2011**, *133* (25), 9824–9833.
- (170) Frei, P.; Pang, L.; Silbermann, M.; Eriş, D.; Mühlethaler, T.; Schwardt, O.; Ernst, B. Target-Directed Dynamic Combinatorial Chemistry: A Study on Potentials and Pitfalls as

Exemplified on a Bacterial Target. *Chem. - A Eur. J.* **2017**, *23* (48), 11570–11577.

- (171) Minard, A.; Morgan, D.; Raguseo, F.; Di Porzio, A.; Liano, D.; Jamieson, A. G.; Di Antonio, M. A Short Peptide That Preferentially Binds C-MYC G-Quadruplex DNA. *Chem. Commun.* **2020**, *56* (63), 8940–8943.
- (172) Liu, K. C.; Röder, K.; Mayer, C.; Adhikari, S.; Wales, D. J.; Balasubramanian, S. Affinity-Selected Bicyclic Peptide G-Quadruplex Ligands Mimic a Protein-like Binding Mechanism. *J. Am. Chem. Soc.* **2020**, *142* (18), 8367–8373.
- (173) Daina, A.; Michielin, O.; Zoete, V. SwissADME: A Free Web Tool to Evaluate Pharmacokinetics, Drug-Likeness and Medicinal Chemistry Friendliness of Small Molecules. *Sci. Rep.* **2017**, *7* (January), 1–13.
- (174) Klingele, J.; Moubaraki, B.; Murray, K. S.; Boas, J. F.; Brooker, S. Dinuclear Copper(II) Complexes of Two Homologous Pyrazine-Based Bis(Terdentate) Diamide Ligands. *Eur. J. Inorg. Chem.* **2005**, No. 8, 1530–1541.
- (175) De Rache, A.; Mergny, J.-L. L. Assessment of Selectivity of G-Quadruplex Ligands via an Optimised FRET Melting Assay. *Biochimie* **2015**, *115*, 194–202.
- (176) Moraes Silva, S.; Tavallaie, R.; Sandiford, L.; Tilley, R. D.; Gooding, J. J. Gold Coated Magnetic Nanoparticles: From Preparation to Surface Modification for Analytical and Biomedical Applications. *Chem. Commun.* **2016**, *52* (48), 7528–7540.
- (177) Xu, Z.; Hou, Y.; Sun, S. Magnetic Core/Shell Fe<sub>3</sub>O<sub>4</sub>/Au and Fe<sub>3</sub>O<sub>4</sub>/Au/Ag Nanoparticles with Tunable Plasmonic Properties. *J. Am. Chem. Soc.* **2007**, *129* (28), 8698–8699.
- (178) <https://www.nanoimmunotech.eu/en/Shop/Nanoparticles-/Gold-Coated-Magnetic-Nanoparticles>.
- (179) Zdravkovic, S. Solid-Phase Extraction for the Preparation of Aqueous Sample Matrices for Gas Chromatographic Analysis in Extractable/Leachable Studies. *Pharm. Technol.* **2017**, *41* (5), 54–60.
- (180) Musumeci, D.; Amato, J.; Randazzo, A.; Novellino, E.; Giancola, C.; Montesarchio, D.; Pagano, B. G-Quadruplex on Oligo Affinity Support (G4-OAS): An Easy Affinity Chromatography-Based Assay for the Screening of G-Quadruplex Ligands. *Anal. Chem.* **2014**, *86* (9), 4126–4130.
- (181) Platella, C.; Musumeci, D.; Arciello, A.; Doria, F.; Freccero, M.; Randazzo, A.; Amato, J.; Pagano, B.; Montesarchio, D. Controlled Pore Glass-Based Oligonucleotide Affinity Support: Towards High Throughput Screening Methods for the Identification of Conformation-Selective G-Quadruplex Ligands. *Anal. Chim. Acta* **2018**, *1030*, 133–141.
- (182) Platella, C.; Pirota, V.; Musumeci, D.; Rizzi, F.; Iachettini, S.; Zizza, P.; Biroccio, A.; Freccero, M.; Montesarchio, D.; Doria, F. Trifunctionalized Naphthalene Diimides and Dimeric Analogues as G-Quadruplex-Targeting Anticancer Agents Selected by Affinity Chromatography. *Int. J. Mol. Sci.* **2020**, *21* (6).
- (183) Daina, A.; Michielin, O.; Zoete, V. SwissADME: A Free Web Tool to Evaluate

- Pharmacokinetics, Drug-Likeness and Medicinal Chemistry Friendliness of Small Molecules. *Sci. Rep.* **2017**, *7* (March), 1–13.
- (184) Muegge, I.; Heald, S. L.; Brittelli, D. Simple Selection Criteria for Drug-like Chemical Matter. *J. Med. Chem.* **2001**, *44* (12), 1841–1846.
- (185) Daina, A.; Zoete, V. A BOILED-Egg To Predict Gastrointestinal Absorption and Brain Penetration of Small Molecules. *ChemMedChem* **2016**, 1117–1121.
- (186) Baell, J. B.; Holloway, G. A. New Substructure Filters for Removal of Pan Assay Interference Compounds (PAINS) from Screening Libraries and for Their Exclusion in Bioassays. *J. Med. Chem.* **2010**, *53* (7), 2719–2740.
- (187) Murat, P.; Zhong, J.; Lekieffre, L.; Cowieson, N. P.; Clancy, J. L.; Preiss, T.; Balasubramanian, S.; Khanna, R.; Tellam, J. G-Quadruplexes Regulate Epstein-Barr Virus-Encoded Nuclear Antigen 1 mRNA Translation. *Nat. Chem. Biol.* **2014**, *10* (5), 358–364.
- (188) Largy, E.; Hamon, F.; Teulade-Fichou, M.-P. Development of a High-Throughput G4-FID Assay for Screening and Evaluation of Small Molecules Binding Quadruplex Nucleic Acid Structures. *Anal. Bioanal. Chem.* **2011**, *400* (10), 3419–3427.
- (189) Voisset, C.; Daskalogianni, C.; Contesse, M.-A.; Mazars, A.; Arbach, H.; Le Cann, M.; Soubigou, F.; Apcher, S.; Fåhraeus, R.; Blondel, M. A Yeast-Based Assay Identifies Drugs That Interfere with Immune Evasion of the Epstein–Barr Virus. *Dis. Model. Mech.* **2014**, *7* (4), 435–444.
- (190) Weibrecht, I.; Leuchowius, K.-J.; Clausson, C.-M.; Conze, T.; Jarvius, M.; Howell, W. M.; Kamali-Moghaddam, M.; Söderberg, O. Proximity Ligation Assays: A Recent Addition to the Proteomics Toolbox. *Expert Rev. Proteomics* **2010**, *7* (3), 401–409.
- (191) Gomez, D.; Shankman, L. S.; Nguyen, A. T.; Owens, G. K. Detection of Histone Modifications at Specific Gene Loci in Single Cells in Histological Sections. *Nat. Methods* **2013**, *10* (2), 171–177.
- (192) Prado Martins, R.; Findakly, S.; Daskalogianni, C.; Teulade-Fichou, M.-P.; Blondel, M.; Fåhraeus, R.; Prado Martins, R.; Findakly, S.; Daskalogianni, C.; Teulade-Fichou, M.-P.; et al. In Cellulo Protein-MRNA Interaction Assay to Determine the Action of G-Quadruplex-Binding Molecules. *Molecules* **2018**, *23* (12), 3124.
- (193) Duke, R. M.; O’Brien, J. E.; McCabe, T.; Gunnlaugsson, T. Colorimetric Sensing of Anions in Aqueous Solution Using a Charge Neutral, Cleft-like, Amidothiourea Receptor: Tilting the Balance between Hydrogen Bonding and Deprotonation in Anion Recognition. *Org. Biomol. Chem.* **2008**, *6* (22), 4089.
- (194) Evangelista, R. A.; Pollak, A.; Gudgin Templeton, E. F. Enzyme-Amplified Lanthanide Luminescence for Enzyme Detection in Bioanalytical Assays. *Anal. Biochem.* **1991**, *197* (1), 213–224.
- (195) Barman, S.; Diehl, K. L.; Anslyn, E. V. The Effect of Alkylation, Protonation, and Hydroxyl Group Substitution on Reversible Alcohol and Water Addition to 2- and 4-Formyl Pyridine

Derivatives. *RSC Adv.* **2014**, 4 (55), 28893–28900.

- (196) Sheldrick, G. M. *SHELXS-97, Program for X-Ray Crystal Structure Solution*; University of Gottingen, Germany, 1997.
- (197) Sheldrick, G. M. Crystal Structure Refinement with SHELXL. *Acta Crystallogr. Sect. C Struct. Chem.* **2015**, 71 (1), 3–8.
- (198) Farrugia, L. J. WinGX Suite for Small-Molecule Single-Crystal Crystallography. *J. Appl. Crystallogr.* **1999**, 32 (4), 837–838.





# *Annexes*

**Table A.** Shifts of melting temperature of doubled-labeled oligonucleotides in the presence of “as-synthesized” ligands obtained upon the reaction of corresponding dihydrazides and aldehydes (building blocks).<sup>[a]</sup>

Building blocks		25TAG	25TAG	22CTA	22CTA	MYC22	MYC22	Pu24T	Pu24T	hp2	hp2
competitor		none	ds26 <sup>[b]</sup>	none	ds26	none	ds26	none	ds26	none	ds26
1A	2a	8.5 ± 1.0	8.3 ± 0.2	4.2 ± 1.3	4.5 ± 1.0	5.1 ± 0.3	5.2 ± 0.3	0.6 ± 0.4	2.8 ± 0.4	0.8 ± 0.5	0.0 ± 0.4
1A	2b	30.6 ± 0.7	30.5 ± 1.0	20.9 ± 0.1	17.2 ± 1.2	25.9 ± 0.3	25.0 ± 0.3	15.3 ± 0.4	13.7 ± 0.5	2.6 ± 1.1	0.7 ± 0.3
1A	2c	25.1 ± 0.4	24.9 ± 1.5	17.8 ± 0.6	16.4 ± 0.4	19.5 ± 0.9	17.8 ± 2.5	11.0 ± 0.4	11.1 ± 0.1	2.0 ± 1.2	0.9 ± 0.2
1A	2d	28.7 ± 0.7	28.2 ± 0.1	19.9 ± 0.6	17.9 ± 0.5	25.1 ± 0.3	24.6 ± 0.5	15.1 ± 0.4	14.0 ± 0.3	4.0 ± 1.4	1.4 ± 0.3
1A	2e	21.0 ± 0.9	18.0 ± 0.9	12.1 ± 0.5	9.7 ± 0.3	19.1 ± 0.1	17.5 ± 0.8	7.0 ± 0.7	6.1 ± 0.6	1.2 ± 0.4	0.6 ± 0.3
1A	2f	23.7 ± 0.9	20.3 ± 1.8	15.2 ± 0.3	12.8 ± 0.6	19.6 ± 0.5	17.9 ± 0.6	9.9 ± 0.5	8.8 ± 0.3	2.3 ± 1.1	0.8 ± 0.2
1A	2g	21.5 ± 1.2	20.5 ± 1.3	15.1 ± 0.5	13.1 ± 0.5	18.9 ± 0.9	17.5 ± 0.5	8.7 ± 1.9	8.7 ± 0.2	2.5 ± 1.0	1.3 ± 0.2
1A	2h	23.2 ± 0.5	21.5 ± 1.0	16.3 ± 0.5	15.2 ± 0.2	17.1 ± 0.5	16.1 ± 0.2	11.8 ± 0.5	11.8 ± 0.1	3.6 ± 1.9	2.2 ± 0.4
1A	2i	11.9 ± 3.0	11.9 ± 2.1	11.5 ± 0.8	12.2 ± 0.6	8.6 ± 0.4	8.4 ± 0.6	4.3 ± 0.3	4.5 ± 0.3	0.8 ± 0.2	0.7 ± 0.2
1B	2a	1.2 ± 0.4	1.3 ± 0.2	2.1 ± 1.5	0.9 ± 0.6	1.7 ± 1.4	3.4 ± 0.3	2.5 ± 1.7	3.0 ± 0.2	0.2 ± 0.1	0.0 ± 0.3
1B	2b	4.6 ± 0.4	4.7 ± 0.4	3.7 ± 0.9	3.6 ± 0.6	3.2 ± 0.2	3.2 ± 0.4	2.7 ± 1.1	3.1 ± 0.2	0.7 ± 0.1	0.3 ± 0.3
1B	2c	13.2 ± 0.2	15.3 ± 0.1	13.2 ± 0.7	13.7 ± 0.5	8.3 ± 0.1	9.1 ± 0.2	6.2 ± 0.6	6.7 ± 0.1	1.4 ± 0.5	0.7 ± 0.2
1B	2d	8.9 ± 1.0	9.4 ± 0.2	10.5 ± 1.7	10.0 ± 0.7	6.1 ± 0.0	5.6 ± 0.3	4.7 ± 0.8	4.8 ± 0.2	1.2 ± 0.8	0.5 ± 0.3
1B	2e	4.2 ± 0.6	2.6 ± 0.4	5.6 ± 0.8	4.5 ± 1.7	3.7 ± 0.3	4.3 ± 0.3	4.6 ± 0.1	3.3 ± 0.2	1.5 ± 0.9	0.7 ± 0.2
1B	2f	4.2 ± 0.2	2.5 ± 0.3	3.9 ± 0.9	1.9 ± 0.6	3.2 ± 0.5	3.6 ± 0.3	2.2 ± 1.3	2.0 ± 0.4	1.1 ± 0.7	0.4 ± 0.3
1B	2g	6.2 ± 0.5	3.9 ± 0.4	5.5 ± 1.1	4.3 ± 0.9	3.9 ± 0.2	3.4 ± 0.3	3.2 ± 0.7	3.3 ± 0.2	1.6 ± 0.5	1.0 ± 0.3
1B	2h	10.3 ± 4.7	10.0 ± 5.4	12.1 ± 0.8	11.8 ± 0.6	7.4 ± 1.0	7.6 ± 0.2	6.0 ± 1.2	5.5 ± 0.7	1.4 ± 1.1	1.0 ± 0.6
1B	2i	3.1 ± 1.0	2.0 ± 0.6	2.6 ± 0.6	3.6 ± 0.7	2.3 ± 2.2	3.5 ± 0.5	2.3 ± 0.3	3.0 ± 0.1	0.2 ± 0.2	0.3 ± 0.1
1C	2a	5.6 ± 0.1	6.2 ± 0.6	2.9 ± 0.4	2.9 ± 0.3	4.4 ± 0.6	3.4 ± 0.5	0.8 ± 0.2	0.3 ± 0.4	-0.3 ± 0.0	0.0 ± 0.1
1C	2b	23.9 ± 0.2	23.7 ± 0.8	15.3 ± 0.4	14.4 ± 0.1	21.0 ± 0.2	20.1 ± 0.2	11.2 ± 0.2	11.4 ± 0.5	0.8 ± 0.1	0.7 ± 0.0
1C	2c	22.7 ± 1.7	21.8 ± 0.6	16.8 ± 0.2	16.0 ± 0.2	18.0 ± 0.5	17.4 ± 0.3	10.4 ± 0.3	10.5 ± 0.2	1.4 ± 0.3	1.1 ± 0.0
1C	2d	23.7 ± 0.2	23.5 ± 0.3	16.4 ± 0.4	15.2 ± 0.2	20.5 ± 0.3	20.7 ± 0.3	11.2 ± 0.3	11.4 ± 0.3	1.5 ± 0.0	1.4 ± 0.0
1C	2e	15.3 ± 0.7	14.9 ± 0.8	9.0 ± 0.1	8.2 ± 0.2	15.9 ± 0.2	15.4 ± 0.2	4.9 ± 0.0	4.6 ± 0.2	0.6 ± 0.1	0.7 ± 0.0

1C	2f	20.7 ± 1.0	20.2 ± 0.6	13.8 ± 0.3	12.6 ± 0.3	19.5 ± 0.2	18.9 ± 0.4	9.0 ± 0.3	8.4 ± 0.5	1.1 ± 0.0	0.9 ± 0.0
1C	2g	20.3 ± 0.7	20.3 ± 0.8	14.4 ± 0.3	13.5 ± 0.3	19.6 ± 0.4	19.3 ± 0.6	8.8 ± 0.2	8.4 ± 0.4	1.8 ± 0.1	1.6 ± 0.1
1C	2h	23.1 ± 0.4	22.7 ± 0.7	17.4 ± 0.3	16.4 ± 0.4	19.4 ± 0.4	18.9 ± 0.3	11.2 ± 0.2	11.4 ± 0.3	2.5 ± 0.1	2.2 ± 0.1
1C	2i	13.7 ± 1.1	13.6 ± 0.9	12.6 ± 0.3	11.8 ± 0.6	10.1 ± 0.2	9.5 ± 0.3	4.4 ± 0.3	4.3 ± 0.2	1.3 ± 0.0	1.0 ± 0.2
1D	2a	10.7 ± 1.5	8.8 ± 0.4	6.2 ± 0.3	4.9 ± 0.5	6.9 ± 0.2	5.8 ± 0.5	3.2 ± 0.4	4.4 ± 0.2	0.2 ± 0.5	0.1 ± 0.2
1D	2b	26.3 ± 0.8	26.7 ± 1.4	20.6 ± 0.3	17.4 ± 0.1	23.1 ± 0.4	23.0 ± 0.6	16.2 ± 0.2	15.3 ± 0.3	0.9 ± 0.4	0.4 ± 0.3
1D	2c	21.3 ± 0.8	21.2 ± 1.0	17.3 ± 0.5	16.8 ± 0.6	16.1 ± 0.7	15.7 ± 1.1	14.3 ± 0.6	13.2 ± 0.2	1.1 ± 0.4	0.8 ± 0.2
1D	2d	25.4 ± 0.9	25.0 ± 1.3	20.6 ± 0.4	18.8 ± 0.3	22.4 ± 0.4	22.4 ± 0.9	15.8 ± 0.0	14.8 ± 0.1	1.3 ± 0.5	0.8 ± 0.2
1D	2e	17.4 ± 0.7	15.4 ± 0.6	12.5 ± 0.3	10.5 ± 0.3	17.2 ± 0.3	15.5 ± 0.4	8.1 ± 0.3	6.7 ± 0.1	1.3 ± 0.2	0.8 ± 0.3
1D	2f	22.4 ± 0.6	16.3 ± 1.9	15.7 ± 0.1	11.0 ± 0.2	20.0 ± 0.2	16.1 ± 1.2	10.4 ± 0.1	8.5 ± 0.1	0.8 ± 0.4	0.7 ± 0.1
1D	2g	21.2 ± 1.0	16.9 ± 1.5	15.9 ± 0.3	11.9 ± 0.5	19.5 ± 0.6	15.5 ± 0.6	10.7 ± 0.1	9.0 ± 0.2	1.4 ± 0.5	0.9 ± 0.1
1D	2h	20.5 ± 0.8	19.2 ± 0.9	17.3 ± 0.3	15.8 ± 0.2	15.7 ± 1.4	14.3 ± 0.7	13.4 ± 0.1	12.3 ± 0.1	1.3 ± 0.6	0.9 ± 0.1
1D	2i	16.5 ± 0.4	15.7 ± 0.2	15.2 ± 0.4	13.9 ± 0.2	13.3 ± 1.2	11.4 ± 0.2	6.6 ± 0.4	6.6 ± 0.2	0.5 ± 0.2	0.3 ± 0.4
1E	2a	25.5 ± 0.9	29.2 ± 0.9	17.7 ± 0.8	10.0 ± 0.3	19.8 ± 0.5	12.6 ± 0.8	10.5 ± 0.5	6.9 ± 0.3	1.6 ± 1.0	0.2 ± 0.3
1E	2b	29.3 ± 0.9	29.8 ± 0.8	19.3 ± 0.5	11.9 ± 0.3	23.8 ± 0.5	24.7 ± 0.4	13.0 ± 0.5	7.8 ± 0.7	-0.3 ± 2.3	0.3 ± 0.2
1E	2c	27.3 ± 0.8	29.0 ± 0.8	20.7 ± 1.2	17.5 ± 0.5	21.8 ± 0.9	16.5 ± 1.9	11.9 ± 0.4	9.2 ± 0.3	1.8 ± 0.2	0.8 ± 0.2
1E	2d	29.1 ± 0.7	29.4 ± 0.6	22.3 ± 0.7	16.2 ± 1.7	24.6 ± 0.7	24.1 ± 1.2	14.6 ± 0.3	9.6 ± 0.3	2.2 ± 0.4	0.8 ± 0.1
1E	2e	24.1 ± 0.5	25.8 ± 1.7	16.3 ± 1.4	11.6 ± 0.6	18.5 ± 7.4	11.7 ± 3.8	9.2 ± 0.8	5.8 ± 1.4	1.7 ± 0.7	0.6 ± 0.3
1E	2f	35.9 ± 0.1	34.9 ± 0.6	27.9 ± 0.1	22.6 ± 0.7	29.3 ± 1.0	27.8 ± 2.7	20.8 ± 0.4	15.3 ± 1.0	1.7 ± 0.6	0.3 ± 0.1
1E	2g	36.0 ± 1.1	34.6 ± 1.0	26.9 ± 1.0	22.5 ± 0.4	29.4 ± 0.5	28.4 ± 0.7	21.0 ± 0.2	19.3 ± 0.3	2.7 ± 1.4	1.0 ± 0.3
1E	2h	30.8 ± 0.6	27.2 ± 5.1	22.7 ± 0.5	20.7 ± 0.8	24.4 ± 0.2	19.1 ± 0.7	16.6 ± 2.3	15.4 ± 1.2	1.9 ± 1.5	0.5 ± 0.4
1E	2i	36.0 ± 0.0	36.0 ± 0.0	34.8 ± 0.8	35.9 ± 0.6	30.5 ± 0.3	30.5 ± 0.1	21.9 ± 0.1	21.7 ± 0.3	1.1 ± 0.2	1.1 ± 0.2
1F	2a	2.5 ± 1.5	2.3 ± 0.3	1.9 ± 0.3	1.8 ± 0.3	2.8 ± 0.3	0.0 ± 1.2	1.1 ± 0.5	2.2 ± 0.2	-0.5 ± 0.3	0.0 ± 0.1
1F	2b	13.7 ± 0.8	12.8 ± 0.7	8.3 ± 0.5	7.0 ± 0.3	12.5 ± 0.3	11.6 ± 0.2	5.1 ± 0.1	4.6 ± 0.2	0.2 ± 0.1	0.4 ± 0.1
1F	2c	13.6 ± 1.3	12.5 ± 0.3	11.2 ± 0.3	10.5 ± 0.2	8.1 ± 0.2	6.8 ± 0.5	6.5 ± 0.1	6.1 ± 0.2	0.9 ± 0.1	0.8 ± 0.2
1F	2d	14.0 ± 0.7	13.5 ± 0.4	11.9 ± 0.8	11.1 ± 0.3	12.7 ± 0.3	11.7 ± 0.3	6.4 ± 0.1	6.2 ± 0.2	0.7 ± 0.0	0.7 ± 0.1
1F	2e	14.4 ± 0.8	13.3 ± 0.7	8.1 ± 0.6	7.3 ± 0.4	12.5 ± 0.3	10.5 ± 0.3	5.4 ± 0.1	5.2 ± 0.1	0.6 ± 0.1	0.7 ± 0.1

1F	2f	7.0 ± 1.1	7.2 ± 0.6	5.4 ± 0.3	5.2 ± 0.3	5.8 ± 0.4	4.4 ± 0.3	3.7 ± 0.1	4.4 ± 0.3	0.2 ± 0.0	0.5 ± 0.1
1F	2g	8.0 ± 0.3	6.7 ± 0.9	6.5 ± 0.3	5.8 ± 0.3	6.4 ± 0.3	4.5 ± 0.2	3.8 ± 0.2	3.4 ± 0.2	0.9 ± 0.1	0.8 ± 0.1
1F	2h	16.3 ± 0.7	16.0 ± 0.6	14.0 ± 0.2	13.6 ± 0.2	11.5 ± 0.3	10.6 ± 0.2	8.5 ± 0.3	8.3 ± 0.2	1.6 ± 0.1	1.1 ± 0.1
1F	2i	3.5 ± 1.3	3.7 ± 0.6	3.5 ± 0.3	3.3 ± 0.2	2.4 ± 1.2	2.2 ± 0.6	2.2 ± 0.2	2.3 ± 0.3	-0.2 ± 0.1	-0.2 ± 0.2
1G	2a	1.1 ± 0.9	1.6 ± 0.7	2.6 ± 2.5	2.1 ± 2.5	3.6 ± 0.3	4.5 ± 0.5	2.4 ± 1.1	4.0 ± 0.8	0.3 ± 0.1	0.3 ± 0.2
1G	2b	2.2 ± 0.9	1.7 ± 0.4	3.2 ± 1.6	2.0 ± 0.6	3.7 ± 0.5	4.8 ± 0.5	3.4 ± 0.6	3.5 ± 0.3	1.2 ± 0.4	0.6 ± 0.3
1G	2c	8.7 ± 0.8	7.9 ± 0.6	8.9 ± 0.9	7.9 ± 0.5	5.1 ± 0.2	4.9 ± 0.1	4.6 ± 0.4	4.1 ± 0.3	1.5 ± 0.3	0.8 ± 0.3
1G	2d	4.1 ± 0.9	3.6 ± 0.7	6.3 ± 2.0	4.7 ± 1.2	3.2 ± 0.5	2.1 ± 1.7	3.1 ± 0.7	2.7 ± 0.6	0.8 ± 0.5	0.4 ± 0.3
1G	2e	2.6 ± 1.0	2.5 ± 0.2	3.5 ± 1.8	2.2 ± 0.6	3.4 ± 0.5	4.8 ± 0.6	3.3 ± 0.7	3.6 ± 0.3	1.3 ± 0.3	0.6 ± 0.3
1G	2f	3.1 ± 0.2	2.3 ± 0.4	3.0 ± 1.6	2.0 ± 0.6	2.9 ± 0.2	4.2 ± 1.1	3.0 ± 0.9	3.6 ± 0.5	0.8 ± 0.1	0.6 ± 0.3
1G	2g	3.2 ± 0.8	3.1 ± 1.4	4.0 ± 1.7	2.9 ± 0.5	3.6 ± 0.6	4.6 ± 0.4	3.6 ± 0.4	3.6 ± 0.2	1.8 ± 0.4	1.0 ± 0.3
1G	2h	8.1 ± 0.9	8.0 ± 0.7	9.3 ± 1.5	8.4 ± 0.5	5.0 ± 0.5	4.8 ± 0.1	4.3 ± 0.7	3.8 ± 0.5	1.8 ± 0.3	0.9 ± 0.3
1G	2i	4.5 ± 2.0	2.1 ± 1.6	7.0 ± 3.7	5.3 ± 2.4	3.7 ± 0.7	3.4 ± 0.4	2.9 ± 0.3	3.1 ± 0.6	0.4 ± 0.1	0.7 ± 0.2
1H	2a	1.4 ± 0.7	2.8 ± 1.9	1.5 ± 0.6	1.2 ± 0.3	2.3 ± 0.3	3.0 ± 0.4	0.9 ± 0.2	0.9 ± 0.9	-0.3 ± 0.2	-0.2 ± 0.1
1H	2b	2.5 ± 0.4	4.0 ± 1.6	2.2 ± 0.4	2.1 ± 0.5	3.1 ± 0.3	3.1 ± 0.3	1.7 ± 0.5	2.0 ± 0.4	0.0 ± 0.2	-0.1 ± 0.2
1H	2c	5.2 ± 0.2	6.2 ± 0.3	4.4 ± 0.5	5.6 ± 0.3	2.4 ± 0.3	2.5 ± 0.3	2.1 ± 0.5	2.6 ± 0.4	0.0 ± 0.1	0.2 ± 0.1
1H	2d	5.1 ± 0.5	4.4 ± 0.4	4.7 ± 0.9	5.0 ± 0.8	4.2 ± 0.3	2.9 ± 0.2	3.4 ± 0.2	2.4 ± 0.4	0.1 ± 0.0	0.2 ± 0.1
1H	2e	3.5 ± 0.3	2.7 ± 0.8	2.7 ± 0.4	1.8 ± 0.8	4.0 ± 0.3	4.4 ± 0.1	2.3 ± 0.1	3.7 ± 0.3	0.2 ± 0.1	0.3 ± 0.3
1H	2f	3.0 ± 0.6	1.9 ± 0.5	2.3 ± 0.9	2.2 ± 0.5	3.7 ± 0.3	0.2 ± 1.5	2.2 ± 0.1	1.5 ± 1.1	-0.3 ± 0.0	0.0 ± 0.0
1H	2g	2.0 ± 1.2	3.2 ± 0.6	2.3 ± 0.2	2.6 ± 0.5	3.8 ± 0.5	3.9 ± 0.3	3.0 ± 0.2	3.7 ± 0.3	0.2 ± 0.1	0.5 ± 0.0
1H	2h	8.3 ± 0.8	8.3 ± 0.9	8.6 ± 0.8	9.7 ± 0.2	6.3 ± 0.3	5.8 ± 0.2	4.6 ± 0.4	4.4 ± 0.3	0.6 ± 0.2	0.8 ± 0.1
1H	2i	3.1 ± 0.9	3.5 ± 3.4	2.1 ± 0.7	4.1 ± 0.1	3.5 ± 0.8	3.5 ± 0.2	3.0 ± 0.2	2.7 ± 0.3	0.1 ± 0.2	0.5 ± 0.2
1I	2a	1.1 ± 0.5	1.0 ± 0.3	0.8 ± 0.3	0.6 ± 0.4	2.3 ± 0.4	-1.2 ± 0.2	0.8 ± 0.3	0.1 ± 0.8	-0.4 ± 0.3	-0.2 ± 0.1
1I	2b	2.6 ± 0.8	2.5 ± 0.6	2.7 ± 0.6	2.6 ± 0.1	3.6 ± 0.3	3.6 ± 0.3	2.5 ± 0.3	3.3 ± 0.2	0.2 ± 0.2	0.4 ± 0.1
1I	2c	6.8 ± 1.2	8.2 ± 0.6	7.0 ± 0.1	7.1 ± 0.3	4.0 ± 0.2	3.2 ± 0.1	2.2 ± 0.2	2.7 ± 0.5	0.4 ± 0.1	0.5 ± 0.1
1I	2d	6.4 ± 0.7	5.1 ± 0.8	5.3 ± 0.4	5.7 ± 0.6	3.3 ± 0.2	2.7 ± 0.2	1.6 ± 0.3	2.9 ± 0.2	0.1 ± 0.1	0.3 ± 0.1
1I	2e	3.9 ± 0.2	3.1 ± 0.4	2.6 ± 0.3	2.6 ± 0.3	4.7 ± 0.3	-0.1 ± 0.1	2.4 ± 0.2	3.3 ± 0.2	0.4 ± 0.1	0.5 ± 0.2

1I	2f	3.1 ± 0.6	2.1 ± 0.4	1.8 ± 0.3	2.0 ± 0.2	2.6 ± 0.6	0.4 ± 0.6	1.7 ± 0.2	2.6 ± 0.8	-0.1 ± 0.1	0.2 ± 0.1
1I	2g	3.8 ± 0.3	3.1 ± 0.2	3.7 ± 0.2	3.0 ± 0.6	4.7 ± 0.2	0.1 ± 0.2	3.5 ± 0.2	4.1 ± 0.2	0.9 ± 0.1	0.9 ± 0.1
1I	2h	8.9 ± 0.2	9.0 ± 0.1	8.2 ± 0.3	8.0 ± 0.2	5.5 ± 0.2	4.5 ± 0.1	3.6 ± 0.3	3.9 ± 0.2	0.8 ± 0.1	0.9 ± 0.2
1I	2i	2.6 ± 1.2	2.4 ± 0.7	1.7 ± 0.3	2.3 ± 0.1	3.2 ± 1.2	3.4 ± 0.2	2.2 ± 0.2	2.1 ± 0.2	-0.1 ± 0.1	0.1 ± 0.2
1J	2a	2.7 ± 3.0	1.6 ± 0.2	2.5 ± 1.7	0.7 ± 0.6	2.4 ± 0.6	3.1 ± 0.6	2.4 ± 1.9	3.2 ± 0.3	0.2 ± 0.1	0.3 ± 0.3
1J	2b	8.6 ± 0.5	2.4 ± 0.5	6.0 ± 0.9	4.3 ± 2.3	6.1 ± 0.4	4.3 ± 0.4	2.9 ± 0.9	3.7 ± 0.2	1.4 ± 0.6	0.5 ± 0.1
1J	2c	14.2 ± 0.5	10.0 ± 0.7	13.1 ± 0.8	9.1 ± 0.6	8.6 ± 0.4	5.6 ± 0.1	5.7 ± 0.5	4.4 ± 0.2	1.9 ± 0.6	1.0 ± 0.1
1J	2d	11.8 ± 0.8	4.4 ± 0.8	11.3 ± 1.0	5.6 ± 1.9	7.7 ± 0.5	3.7 ± 0.5	3.7 ± 0.9	3.6 ± 0.2	1.8 ± 0.4	0.5 ± 0.2
1J	2e	7.5 ± 0.2	5.9 ± 0.4	6.3 ± 1.3	5.9 ± 1.6	5.3 ± 0.4	3.8 ± 0.1	3.6 ± 1.0	4.3 ± 0.2	1.6 ± 0.3	0.9 ± 0.2
1J	2f	8.9 ± 0.4	3.4 ± 0.2	5.4 ± 0.9	2.0 ± 0.9	5.8 ± 0.3	3.8 ± 0.4	4.7 ± 1.6	3.9 ± 0.2	2.2 ± 0.7	0.6 ± 0.3
1J	2g	8.7 ± 0.2	4.6 ± 0.3	6.5 ± 0.8	2.4 ± 1.1	5.6 ± 0.4	4.3 ± 0.3	4.7 ± 0.9	4.9 ± 0.1	2.2 ± 0.5	1.2 ± 0.4
1J	2h	13.4 ± 0.2	12.0 ± 1.3	11.7 ± 1.0	9.3 ± 0.2	7.9 ± 0.3	5.7 ± 0.2	8.8 ± 1.1	6.8 ± 0.5	2.1 ± 0.4	1.2 ± 0.3
1J	2i	3.9 ± 2.0	5.4 ± 2.2	4.1 ± 0.3	3.6 ± 0.1	4.1 ± 0.6	4.3 ± 0.4	3.6 ± 0.3	3.6 ± 0.1	1.1 ± 0.1	0.9 ± 0.1

<sup>[a]</sup> Data are means ± s.d. from three independent measurements; conditions: see Table S2 footnote. <sup>[b]</sup> c = 10 μM.

## Resumé en français

Les G-quadruplexes (G4s) sont des structures non-canoniques d'acides nucléiques (ADN et ARN) constituées d'au moins deux quartets de guanines. L'une des propriétés importantes des G4s est leur capacité à former des complexes avec de petites molécules exogènes (« ligands ») et d'influencer ainsi les processus biologiques dans lesquels ils sont impliqués. Ainsi, l'interaction de petites molécules avec certaines structures G4s permettrait de diminuer l'expression de certains oncogènes, d'inhiber la télomérase ou encore d'induire des dommages à l'ADN.

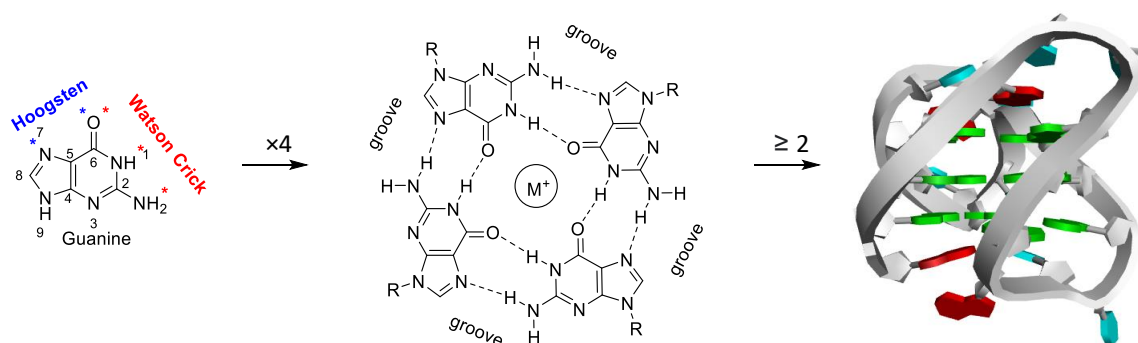
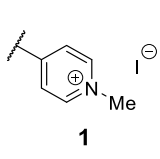
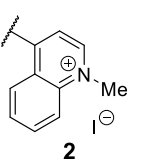
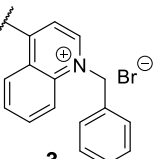
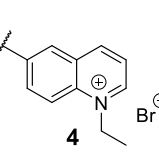
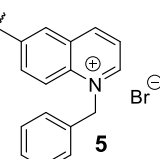
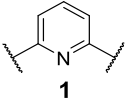
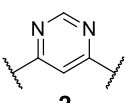
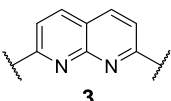
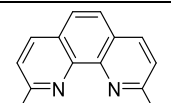


Figure 88. D'une guanine au G-quadruplexe. De gauche à droite: guanine, G-quartet et G-quadruplex (les guanines sont représentées en vert, les adenines en rouge, et les thymines en bleu).

Ce travail vise à développer des méthodologies rapides et simples pour la synthèse et le criblage des molécules afin d'identifier des ligands sélectifs et affins de structures non-canoniques d'acides nucléiques, en particulier des G4s. Plus précisément, ce travail explore la synthèse réversible d'acylhydrazones, jusqu'ici peu appliquée pour le développement de ligands de l'ADN et de l'ARN. Dans un premier temps, une série de 20 bis(acylhydrazones), analogues des ligands **PDC (360A)** et **PhenDC3**, a été obtenue par la synthèse préparative (Tableau 1).

Tableau 1. Structures et acronymes de ligands modèles synthétisés, conditions et rendements de réactions

A L					
	<b>PyDH1</b> , [OR28], A1-L1-A1 90% <sup>a</sup>	<b>PyDH2</b> , [OR19], A2-L1-A2 87% <sup>a</sup>	<b>PyDH3</b> , [OR20], A3-L1-A3 89% <sup>a</sup>	<b>PyDH4</b> , [OR21], A4-L1-A4 92% <sup>a</sup>	<b>PyDH5</b> , [OR40], A5-L1-A5 52% <sup>a</sup>
	<b>PymDH1</b> , [OR31], A1-L2-A1 97% <sup>a</sup>	<b>PymDH2</b> , [OR27], A2-L2-A2 92% <sup>a</sup>	<b>PymDH3</b> , [OR127], A3-L2-A3 93% <sup>a</sup>	<b>PymDH4</b> , [OR129], A4-L2-A4 96% <sup>a</sup>	<b>PymDH5</b> , [OR128], A5-L2-A5 97% <sup>a</sup>
	<b>NaphDH1</b> , [OR29], A1-L3-A1 78% <sup>a</sup>	<b>NaphDH2</b> , [OR30], A2-L3-A2 81% <sup>a</sup>	<b>NaphDH3</b> , [OR22], A3-L3-A3 90% <sup>a</sup>	<b>NaphDH4</b> , [OR23], A4-L3-A4 75% <sup>a</sup>	<b>NaphDH5</b> , [OR42], A5-L3-A5 80% <sup>b</sup>
	<b>PhenDH1</b> , [OR33], A1-L4-A1 76% <sup>a</sup>	<b>PhenDH2</b> , [OR34], A2-L4-A2 78% <sup>a</sup>	<b>PhenDH3</b> , [OR35], A3-L4-A3 87% <sup>a</sup>	<b>PhenDH4</b> , [OR36], A4-L4-A4 72% <sup>a</sup>	<b>PhenDH5</b> , [OR41], A5-L4-A5 91% <sup>b</sup>

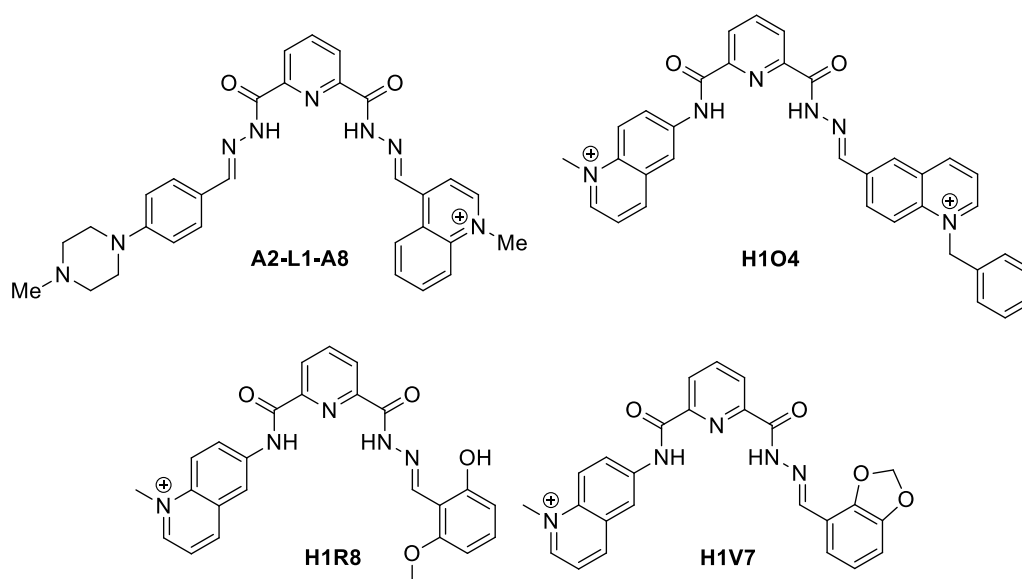
Conditions de synthèse: <sup>a</sup> DMF, 100 °C, 2 h; <sup>b</sup> DMF, 80 °C, 4 h.

Les expériences de dénaturation thermique suivies par fluorescence ont démontré que certains de ces composés avaient une bonne affinité pour l'ADN G4. Ces expériences ont permis de valider le potentiel du motif acylhydrazone pour le développement de ligands des G4s. Ensuite, une méthode de chimie dynamique combinatoire (CDC) a été développée. Cette dernière consiste en la génération de bibliothèques combinatoires comportant jusqu'à 20 composés, suivie par l'isolement des ligands les plus affins par la précipitation avec la cible, préalablement immobilisée sur des billes magnétiques.

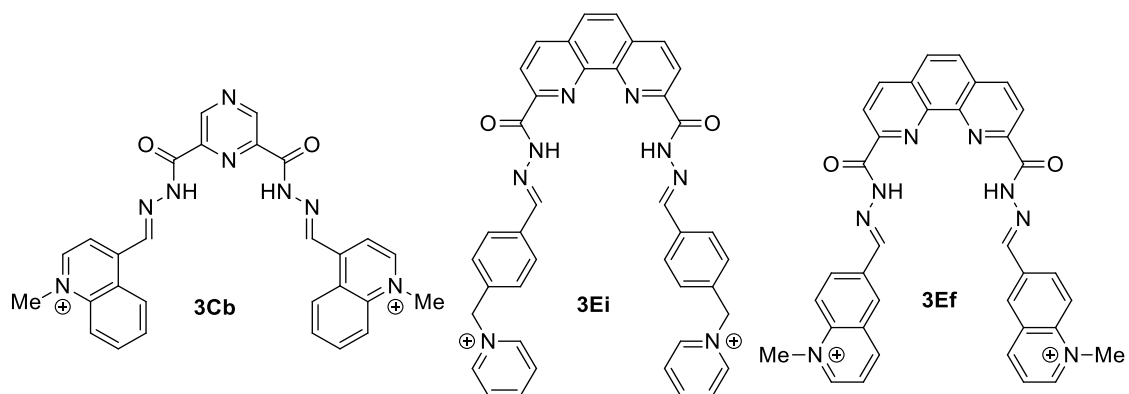
Ainsi, un bis(acylhydrazone) non-symétrique a été identifié comme un ligand prometteur du G4 parallèle *Pu24T*. Cependant, les expériences avec ses proches analogues n'ont pas confirmé son affinité aux G4 augmentée par rapport aux dérivés symétriques. Il a été supposé que les résultats d'expériences de CDC pouvaient être biaisés par des interactions non-spécifiques entre les ligands et les billes magnétiques.



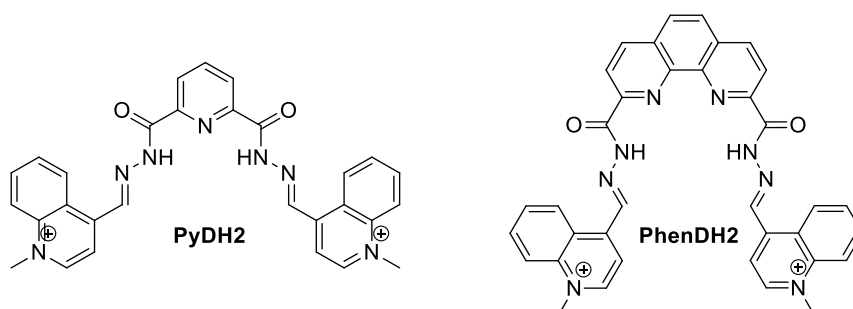
Pour améliorer l'analyse des bibliothèques combinatoires, une nouvelle méthode basée sur l'extraction en phase solide des ligands a été développée et appliquée à deux bibliothèques d'acylhydrazones non-symétriques. Huit hits ont été obtenus à partir de 70 composés générés *in situ*. Trois d'entre eux (**H1O4**, **H1R8** et **H1V7**) ont été sélectionnés pour la synthèse préparative et une étude de l'interaction avec l'ADN G4.



En parallèle, une approche classique de chimie combinatoire a été élaborée. Cette dernière a conduit à la génération d'une bibliothèque combinatoire de 90 dérivés bis(acylhydrazone) sous forme de solutions 2 mM dans le DMSO prêtes à l'emploi. Ces échantillons ont été utilisés directement dans le criblage biophysique contre quatre ADN G4s de trois topologies différentes. Les composés les plus actifs (**3Cb**, **3Ei** et **3Ef**) ont été synthétisés d'une manière préparative et leur interaction avec les G4s a été étudiée en détail par des méthodes biophysiques, y compris la spectrométrie de masse native. Ainsi, au moins un dérivé avec une affinité pour les G4s supérieure à celle de **PhenDC3** et une sélectivité inédite pour la topologie antiparallèle des G4 a été identifié.



Enfin, dans le cadre d'un projet collaboratif (M. Blondel, Université de Bretagne Occidentale), des ligands synthétisés au cours de ce travail ont été étudiés vis-à-vis de leur capacité à moduler d'évasion immune du virus d'Epstein-Barr (EBV). Il a été démontré que certains bis(acylhydrazones) (**PyDH2** et **PhenDH2**) interagissent *in vitro* avec la séquence riche en guanines de l'ARNm codante pour le domaine riche en glycine-alanine (GAR) de la protéine virale EBNA1. Deux de ces dérivés déplacent le facteur de la cellule hôte (nucléoline) de l'ARNm d'EBNA1, conduisant ainsi à la surexpression de la protéine et à la présence exacerbée de peptides antigéniques sur les cellules infectées. Cet effet représente une opportunité thérapeutique pour le traitement des cancers associés à l'EBV.





**Titre :** Les approches de chimie combinatoire pour le développement des ligands de G-quadruplexes de l'ADN et de l'ARN

**Mots clés :** chimie dynamique combinatoire, chimie combinatoire, ligands de l'ADN, G-quadruplexes, chimie organique, chimie supramoléculaire.

**Résumé :** Les G-quadruplexes (G4s) sont des structures non-canoniques d'acides nucléiques (ADN et ARN) constituées d'au moins deux quartets de guanines. L'une des propriétés importantes des G4s est leur capacité à former des complexes avec de petites molécules exogènes (« ligands ») et d'influencer ainsi les processus biologiques dans lesquels ils sont impliqués. Ainsi, l'interaction de petites molécules avec certaines structures G4s permettrait de diminuer l'expression de certains oncogènes, d'inhiber la télomérase ou encore d'induire des dommages à l'ADN.

Ce travail vise à développer des méthodologies rapides et simples pour la synthèse et le criblage des molécules afin d'identifier des ligands sélectifs et affins de structures non-canoniques d'acides nucléiques, en particulier des G4s. Plus précisément, ce travail explore la synthèse réversible d'acylhydrazones, jusqu'ici peu appliquée pour le développement de ligands de l'ADN et de l'ARN. Dans un premier temps, une série de 20 bis(acylhydrazones), analogues des ligands PDC (360A) et PhenDC3, a été obtenue par la synthèse préparative. Les expériences de dénaturation thermique suivie par fluorescence ont démontré que certains de ces composés avaient une bonne affinité pour l'ADN G4. Ces expériences ont permis de valider le potentiel du motif acylhydrazone pour le développement de ligands des G4s. Ensuite, une méthode de chimie dynamique combinatoire (CDC) a été développée. Cette dernière consiste en génération de bibliothèques combinatoires comportant jusqu'à 20 composés, suivie par l'isolement des ligands les plus affins par la précipitation avec la cible, immobilisée sur des billes magnétiques.

Ainsi, un bis(acylhydrazone) non-symétrique a été identifié comme un ligand prometteur du G4 parallèle Pu24T. Cependant, les expériences avec ses proches analogues n'ont pas confirmé son affinité aux G4 augmentée par rapport aux dérivés symétriques. Il a été supposé que les résultats d'expériences de CDC pouvaient être biaisés par des interactions non-spécifiques entre les ligands et les billes magnétiques. Pour améliorer l'analyse des

bibliothèques combinatoires, une nouvelle méthode basée sur l'extraction en phase solide des ligands a été développée et appliquée à deux bibliothèques d'acylhydrazones non-symétriques. Huit hits ont été obtenus à partir de 70 composés générés *in situ*. Trois d'entre eux ont été sélectionnés pour la synthèse préparative et une étude de l'interaction avec l'ADN G4.

En parallèle, une approche classique de chimie combinatoire a été élaborée, ce qui a conduit à la génération d'une bibliothèque combinatoire de 90 dérivés bis(acylhydrazone) sous forme de solutions 2 mM dans DMSO prêtes à l'emploi, avec une pureté moyenne de 87%. Ces échantillons ont été utilisés directement dans le criblage biophysique contre quatre G4s de l'ADN de trois topologies différentes. Les composés les plus actifs ont été synthétisés d'une manière préparative et leur interaction avec les G4s a été étudiée en détail par des méthodes biophysiques, y compris la spectrométrie de masse native. Ainsi, au moins un dérivé avec une affinité pour les G4s supérieure à celle de PhenDC3 et une sélectivité inédite pour le G4 antiparallèle a été identifié.

Enfin, dans le cadre d'un projet collaboratif (M. Blondel, Université de Bretagne Occidentale), des ligands synthétisés au cours de ce travail ont été étudiés vis-à-vis de leur capacité à moduler d'évasion immune du virus d'Epstein-Barr (EBV). Il a été démontré que certains bis(acylhydrazones) interagissent *in vitro* avec la séquence riche en guanines de l'ARNm codante pour le domaine riche en glycine-alanine (GAR) de la protéine virale EBNA1. Deux de ces dérivés déplacent le facteur de la cellule hôte (nucléoline) de l'ARNm d'EBNA1, conduisant ainsi à la surexpression de la protéine et à la présence exacerbée de peptides antigéniques sur les cellules infectées. Cet effet représente une opportunité thérapeutique pour le traitement des cancers associés à l'EBV.

**Title :** Combinatorial chemistry approaches for the development of G-quadruplex DNA and RNA ligands

**Keywords :** dynamic combinatorial chemistry, combinatorial chemistry, DNA ligands, G-quadruplexes, organic chemistry, supramolecular chemistry.

**Abstract :** G-quadruplexes (G4s) are four-stranded structures of nucleic acids (DNA or RNA) that consist of at least two coplanar guanine quartets. An important feature of G4s is their ability to form stable complexes with exogenous small molecules (ligands) and thus influence biological processes in which they are involved. G4 targeting is often associated with oncology, where G4 ligands may suppress the expression of oncogenes, inhibit telomerase, or induce DNA damage in cancer cells.

This work aims to develop methodologies for rapid and simple synthesis and screening of compounds, in order to identify selective and highly affine ligands of given non-canonical structures of nucleic acids, in particular G4s. Specifically, this work exploits the chemistry of reversible synthesis of acylhydrazones, which has been barely applied for the development of DNA or RNA ligands before. First, a small library of 20 cationic bis(acylhydrazones), analogues of the previously reported G4-ligands PDC (360A) and PhenDC3, was obtained by preparative synthesis. Through fluorescence melting experiments it is demonstrated that some of compounds indeed have high affinity to G4-DNA, validating the suitability of the acylhydrazone motif as a scaffold for the development of G4 ligands. Next, a method of dynamic combinatorial chemistry (DCC), which consists in simultaneous one-pot generation of libraries of up to 20 compounds with consecutive pull-down of most affine ligands by bead-immobilized targets (i.e., G4-DNA), was developed. By using this method, a non-symmetrical bis(acylhydrazone) was identified as a promising ligand of a parallel G4-DNA Pu24T. However, biophysical experiments with its close structural analogues did not confirm their preferential binding in comparison with the symmetrically substituted compound. It is proposed that the outcome of DCC experiments may be biased by non-specific interactions of ligands with magnetic beads, leading to false-positive results. In order to improve the analysis of dynamic combinatorial libraries, a novel method based on solid-phase extraction of the G4-

ligand complex was developed and applied to two libraries of non-symmetric acylhydrazones. In a few rounds of selection, 13 hits were obtained out of 70 *in situ* generated compounds. Three of them were selected for preparative synthesis and detailed study of interaction with G4-DNA.

In parallel, a classical combinatorial chemistry approach was developed, resulting in generation of a combinatorial library of 90 individual bis(acylhydrazone) derivatives in the form of ready-to-use 2 mM solutions in DMSO, with an average purity of 87%. These samples were directly used for biophysical screening experiments towards four G4-DNA targets of three different topologies. Three most active compounds were obtained in preparative manner and their interaction with the mentioned biological targets was studied in detail by several biophysical methods, including native mass spectrometry experiments. This way, at least one derivative with a G4-DNA affinity superior to that of PhenDC3 and unprecedented selectivity towards anti-parallel G4-DNA could be identified.

Finally, in the framework of a collaborative project (M. Blondel, University of Western Brittany) the ligands synthesized in this work were studied with respect to their capacity to act as modulators of the immune evasion of Epstein-Barr virus (EBV). Specifically, it was shown that several bis(acylhydrazones) bind *in vitro* to G4-RNA structures formed by the guanine-rich repeat sequence of mRNA encoding for the glycine-alanine rich (GAR) domain of viral genome maintenance protein EBNA1. Moreover, two derivatives were found to displace the host cell factor nucleolin from EBNA1 mRNA, leading to overexpression of EBNA1 protein and a concomitant increase of antigen presentation in EBV-infected cell cultures. This effect represents an interesting therapeutic opportunity for treatment of EBV-related cancers.