

Alma Mater Studiorum – Università di Bologna

DOTTORATO DI RICERCA IN

Scienze Biochimiche e Biotecnologiche

Ciclo XXIX

Settore Concorsuale di afferenza: 03/D1

Settore Scientifico disciplinare: CHIM/08

Exploring Novel Anticancer Approaches: Design, Synthesis and Biological
Evaluation of Small Molecules and NGR Tumor-Homing Peptides

Presentata da: Alessandra Graziadio

Coordinatore Dottorato

Relatore

Prof. Santi Mario Spampinato

Prof.ssa Alessandra Locatelli

Esame finale anno 2017

Contents

Preface	3
Acknowledgements.....	4
Abstract	5
CHAPTER 1.....	6
1. INTRODUCTION	6
1.1. Cancer statistics at a glance	6
1.2. Cancer: an “evolutionary process”	7
1.3. “Cancer hallmarks”	9
1.4. Small molecules targeting DNA: past and future directions	12
1.5. Tumor-homing peptides as vehicles for cancer therapeutics and diagnostics	15
1.6. Aim of the thesis	17
CHAPTER 2.....	19
2. Design and synthesis of novel hydrazone derivatives of diimidazo[1,2 <i>a</i> -1,2 <i>c</i>]pyrimidine as selective G-quadruplex binders and promising anticancer drug candidates.....	19
2.1. Background.....	19
2.1.1. DNA G-quadruplexes: topology, structure and biological roles.....	19
2.1.2. G-quadruplex structures as targets for anticancer therapy.....	22
2.1.3. G-quadruplex binders.....	26
2.2. Aim of the project	32
2.3. Results and discussion.....	34
2.3.1. Chemistry	34
2.3.2. Biophysical studies	42
2.3.3. Molecular docking.....	49
2.3.4. Biological Studies.....	51
2.4. Conclusions.....	54
CHAPTER 3.....	56
3. NGR tumor-homing peptides: structural considerations for effective APN targeting.....	56
3.1. Background.....	56

3.1.1.	Exploring tumor vessels “zip codes”: identification of NGR tumor homing peptides.....	56
3.1.2.	Aminopeptidase N (APN/CD13) as receptor for NGR tumor-homing peptides...	58
3.1.3.	NGR peptide-based agents.....	60
3.1.4.	Interactions between APN and NGR peptides	63
3.2.	Aim of the project	65
3.3.	Results and discussion.....	67
3.3.1.	Chemistry	67
3.3.2.	Biology.....	78
3.4.	Conclusions.....	84
CHAPTER 4.....		86
4.	Experimental section.....	86
4.1.	Experimental data for Chapter 2.....	86
4.1.1.	Chemistry	86
4.1.2.	Biophysical evaluation.....	96
4.1.3.	Computational studies	99
4.1.4.	Biology.....	100
4.2.	Experimental data for Chapter 3.....	102
4.2.1.	Chemistry	102
4.2.2.	Biology.....	114
Abbreviations		116
Bibliography		118
Appendix		129
NMR spectra of compounds described in Chapter 2		129
NMR spectra of compounds described in Chapter 3		138

Preface

The first chapter of this PhD thesis is a brief introduction about the complexity of cancer genetics and biology. Furthermore it describes the two anticancer approaches of interest: DNA G-quadruplex secondary structures as targets for small molecules and tumor-homing peptides as tools for drug delivery into tumors.

The second chapter reports the first project, which has been focused on the development of diimidazopyrimidine hydrazones, as small molecules able to bind DNA G-quadruplex structures, and potential antitumor agents.

The work has been carried out under the supervision of Prof. Alessandra Locatelli at the Department of Pharmacy and Biotechnology, Alma Mater Studiorum-University of Bologna (Italy).

The third chapter describes the second project, based on the synthesis of NGR tumor-homing peptide derivatives and on the investigation of the structural requirements for effective tumor homing targeting.

This research has been developed as PhD visiting student in the research group led by Prof. Matteo Zanda at the Institute of Medical Sciences, University of Aberdeen (Scotland, UK).

The experimental procedures for the chemical synthesis and biological evaluation are reported in the fourth chapter of the thesis.

Acknowledgements

I would like to express my sincere gratitude to my supervisor Prof. Alessandra Locatelli for the great opportunity she gave me and for her guidance and encouragement.

I would like to thank also Prof. Mirella Rambaldi, Prof. Alberto Leoni and Dr. Rita Morigi (University of Bologna) for their help in the project realization.

I would like to thank Prof. Randazzo and his co-workers (University of Naples) for the biophysical studies of the new compounds on several G-quadruplex structures, and Prof. Capranico and co-workers for the biological evaluation.

I wish to express my gratitude to Prof. Zanda and to all the researchers working in his group at University of Aberdeen.

In particular I wish to thank Dr. Monica Piras for the support and constant encouragement during the research stage carried out at University of Aberdeen.

I am very grateful also to Dr. Massimiliano Baldassarre, for the enzymatic assay in cancer cells and for confocal microscopy studies.

Abstract

In anticancer drug discovery DNA secondary structures, called G-quadruplexes (G4s), emerged as targets of interest, owing to their localization in crucial positions of the genome, such as in telomeres and oncogene promoter regions. It has been widely accepted that G4 stabilization by small molecules at these levels, may represent a realistic anticancer strategy. Since a diimidazo[1,2-*a*:1,2-*c*]pyrimidine bis-guanylylhydrazone derivative was found to be a potent stabiliser of several quadruplex DNAs over duplex DNA, a library of analogues was designed and synthesized to shed light on the structural requirements for G4 binding. The newly synthesized hydrazones were evaluated for their ability to stabilize several G4 with different topology and also for their antiproliferative activity. In particular the most promising compound of the series showed the capability to selectively stabilize a specific G4 structure and to inhibit cancer cell growth in the nanomolar range.

Tumor-homing peptides have been considered promising tools for imaging and therapy of tumors over-expressing their target ligand. In this context, the cyclic CNGRC peptide has been extensively used as tumor-homing carrier for its ability to specifically recognise aminopeptidase N (APN/CD13), a zinc-dependent ectoenzyme, upregulated in tumor blood vessels undergoing angiogenesis and in the majority of cancer cells. Considering the potential of APN-directed tumour homing strategies and taking into account the recently reported crystal structure of porcine APN complexed with cCNGRCG, a library of C-terminus and N-terminus functionalized cCNGRC peptides was designed and synthesized with the aim to confirm the importance of N-terminal amine group of the cCNGRC in the APN binding. The synthesized peptides were assayed for their ability to inhibit APN *in vitro*. The C-terminus was identified as the right conjugation site of the cCNGRC peptide, which was consequently conjugated with a fluorescent probe for APN imaging studies in cancer cells.

CHAPTER 1

1. INTRODUCTION

1.1. Cancer statistics at a glance

Cancers are among the leading causes of morbidity and mortality worldwide, with approximately 14 million new cases and 8.2 million cancer related deaths in 2012 (World Cancer Report 2014).

According to Cancer Statistics 2016 by Siegel *et al.*, over the past 3 decades, the 5-year relative survival rate for all cancers combined has increased 20 percentage points in United States.¹

Progress in cancer survival has been most rapid particularly for hematopoietic and lymphoid malignancies owing to improved treatment protocols. The discovery and the application of targeted therapies, such as the use of BCR-ABL tyrosine kinase inhibitors (e.g.: imatinib) allowed that survival for patients with chronic myeloid leukemia increased from 31% in the early 1990s to 63% during 2005 to 2011.²

Despite these considerable progresses, death rates are increasing for cancers of liver, pancreas and uterine corpus. The advances in both diagnosis and therapy of liver and pancreatic cancers are proceeding more slowly. As a consequence these cancer types are two of the most fatal cancers in adults. Moreover, brain cancer represents the leading cause of cancer death in children and adolescents.¹

The fight against cancer still requires intensive clinical and basic research.

1.2. Cancer: an “evolutionary process”

Cancers can be seen as the result of a Darwinian process occurring among cell populations. Indeed, analogous to what happens in the origins of species, cancer development is based on two constitutive processes: the acquisition of heritable genetic mutations in single cells and the natural selection resulting in phenotypic diversity. Cells carrying deleterious mutations, which acquire the capability to proliferate and survive even in stress conditions, are “selected” in cancer pathogenesis. A cancer cell descends from the fertilized egg, carrying a copy of its diploid genome; however, the DNA sequence of a cancer cell genome acquires substantial somatic differences from its progenitor fertilized egg (Fig. 1.1).³

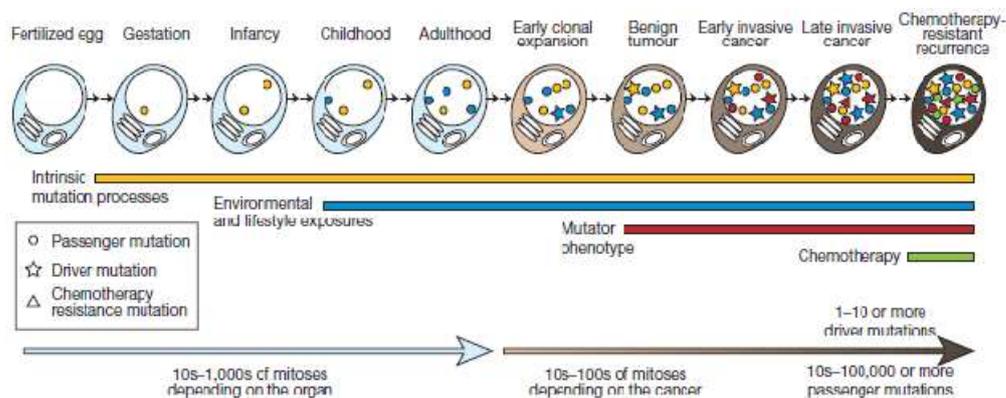


Figure 1.1. Mitotic cell divisions from the fertilized egg to a cancer single cell.³

Somatic mutations in a cancer genome include base substitutions; insertions or deletions of DNA segments and DNA rearrangements. In addition, the cancer cell may have acquired completely new DNA sequences from exogenous sources, such as viruses (e.g.: human papilloma virus, Epstein Barr virus).⁴

DNA in normal cells is continuously damaged by mutagens. Those damages that mislead the DNA reparation machinery, may be turned into fixed mutations. The mutation rates of each somatic mutation differ among the several structural classes, they also vary among cell types and increase upon mutagen exposure. For example tobacco smoke carcinogens exposure has been associated with prevalence of G →T transversions in *TP53* gene found in lung cancer.⁵

However not all acquired somatic abnormalities have a central role in cancer development. The so-called 'driver mutations' are those effectively responsible of oncogenesis.³ A high number of these driver mutations and altered cancer genes have been identified so far. Some of the implicated cancer genes, such as *TP53* and *KRAS* are frequently mutated in different cancer types, whereas others are specific for one cancer type (www.sanger.ac.uk/genetics/CGP/cosmic).

Cancer genome may acquire also epigenetic changes responsible of alterations in chromatin structure and gene expression. Even cancer epigenetic abnormalities can be subjected to natural selection; they confer phenotypic plasticity, which results in cellular heterogeneity and tumor progression. Epigenetic gene regulation mechanisms include histone modifications, which play critical roles in reprogramming a cell's phenotype in response to environmental changes.⁶

Somatic mutations within mitochondrial DNA have been also reported in many tumors. However, further studies are needed to clarify the functional significance of specific mitochondrial mutations in cancer progression.⁷

1.3. “Cancer hallmarks”

Tumors have to be considered not as masses of proliferating cells, but as complex tissues composed of various cell types. Normal cells are recruited to form tumor-associated stroma, thus playing active roles in tumorigenesis. Therefore the biology of a tumor can be understood by studying both the specialized cell types and the “tumor microenvironment” (Fig. 1.2).

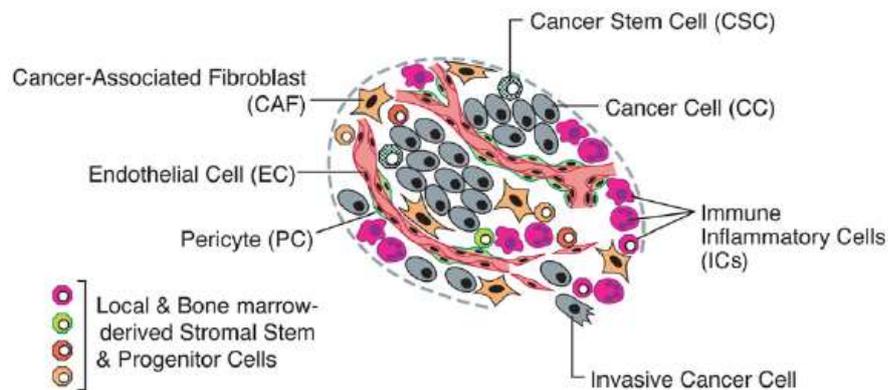


Figure 1.2. The cells of the tumor microenvironment.⁸

Hanahan and Weinberg defined ‘cancer hallmarks’ as the capabilities that cells progressively acquire to become tumorigenic and consequently malignant (Fig. 1.3).^{8,9} These hallmarks of cancer together constitute an organizing principle which can be used for understanding the complexity of cancer biology and the diversity of neoplastic diseases. Cancer cells can acquire the capability to sustain proliferative signaling by producing growth factors, or alternatively by sending signals to normal cells of the tumor associated stroma.¹⁰ Growth factors are able to bind cell-surface tyrosine kinase receptors, thus activating intracellular signalling pathways that regulate progression through the cell cycle as well as cell growth.¹¹

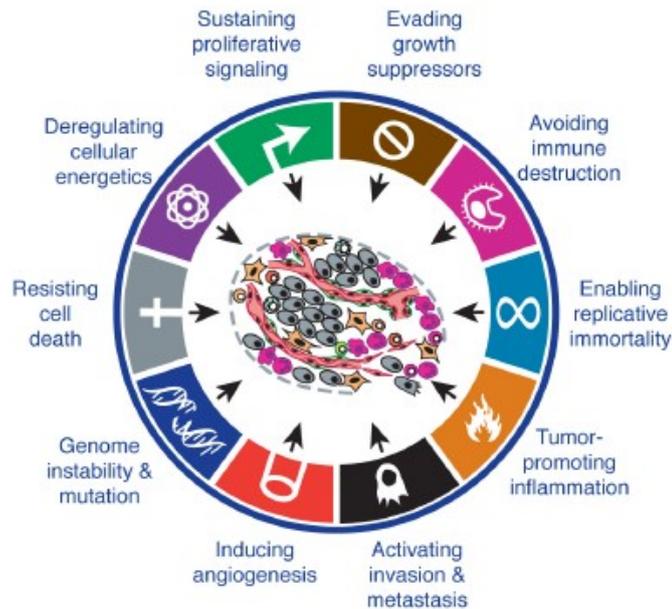


Figure 1.3. Hallmarks of cancer.⁸

Moreover the levels of receptor proteins expressed on cancer cell surface, are often elevated and as a consequence, cells become hyperresponsive to growth factors. In addition somatic mutations in certain human tumors result in constitutive activation of signalling pathways (i.e.: mutations on phosphoinositide 3-kinase (PI3-kinase) hyperactivate mitogen activated Akt/PKB pathways).¹²

In addition cell proliferation is sustained because tumor suppressor genes responsible of the negative regulation of cell proliferation are subjected to mutations in cancer cells. Prototypical tumor suppressors encode the RB (retinoblastoma-associated) and TP53 proteins that operate within regulatory circuits, regulating the cell fate, leading to cell proliferation, or, alternatively, activating senescence and apoptotic programs.¹³ One of the strategies, adopted by tumour cells to circumvent apoptosis, is the loss of TP53 tumor suppressor function.¹⁴

Moreover cancer cells undergo cell immortalization, thanks to the telomerase activity, consisting in the inhibition of the progressive telomere erosion that occurs after each cell cycle.¹⁵

When necrotic cell death occurs in the context of a tumour, proinflammatory signals are released into the tissue microenvironment, and inflammatory cells are recruited,¹⁶ thus

resulting in increased angiogenesis, cancer cell proliferation, and invasiveness. Cancer cells have also the capability of evading the elimination by immune cells.¹⁷

Like normal tissues, tumors require nutrients and oxygen and they need also to evacuate metabolic wastes and carbon dioxide. To this end, the angiogenesis process is of crucial importance to built the necessary tumor-associated vasculature.¹⁸

Reprogramming of cellular energy metabolism is another capability acquired by tumours in order to support continuous cell growth and proliferation. Glucose uptake and glycolytic pathway are markedly increased in many human tumors, in response of activated oncogenes (e.g.: *RAS*, *MYC*) and mutant tumor suppressors (e.g.: *TP53*).¹⁹ Glycolysis can be accentuated under hypoxia; both *RAS* oncoprotein and hypoxia can independently increase the levels of two transcription factors HIF1- α and HIF2- α , responsible of glycolysis upregulation.²⁰

Cancer cells acquire also the capability to attach to other cells and to the extracellular matrix (ECM) proteins, in order to invade tissues. Local invasion is followed by intravasation of cancer cells into blood and lymphatic vessels, consequent extravasation into and colonization of distant organs: this multistep process represents the 'invasion-metastasis cascade' of malignant tumors.²¹

1.4. Small molecules targeting DNA: past and future directions

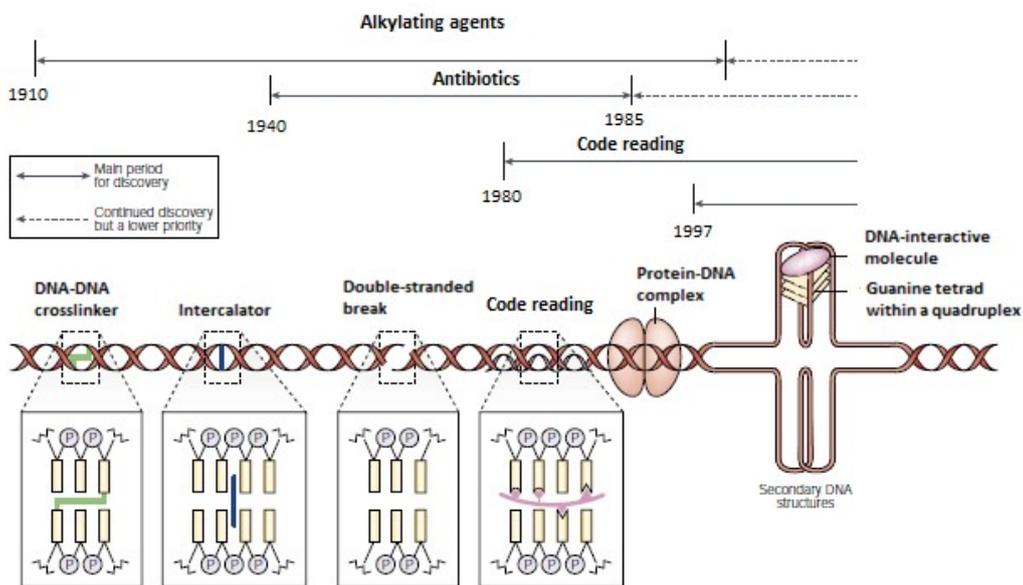


Figure 1.4. Classes of DNA-interactive agents.²²

Research before and during World War II led to the comprehension that nitrogen mustards were able to exert anticancer effects by alkylating DNA, showing a very high reactivity.²³ Therefore DNA started to be perceived as target for therapeutic intervention of cancer and, as a result, a high number of the developed anticancer drugs target DNA with different mechanism of action (Fig. 1.4).²²

Among the synthesized alkylating agents, the aromatic mustards such as chlorambucil and mephalan (Fig. 1.5) are the most effective ones with the tendency to crosslink the two complementary strands of DNA.²⁴

Conversely the agent cisplatin (Fig. 1.5) causes crosslink within a strand of DNA, by binding covalently to adjacent guanines.²⁵ Although showing side effects, which can be overcome by site-specific delivery, some cancers, such as the testicular one, showed a good sensitivity to this drug.

It was surprising that the antibiotic doxorubicin (Fig. 1.5), known as non-specific DNA intercalating agent,²⁶ would have shown higher selectivity for cancer cells than normal cells. Afterwards it was discovered that doxorubicin exerts its anticancer activity by inducing topoisomerase II-associated strand breaks.²⁷ The selectivity for cancer cells

might be dependent on the elevated levels of the DNA associated topoisomerase II in these cells. However these features are counterbalanced by high toxicity and generation of resistance mechanisms.

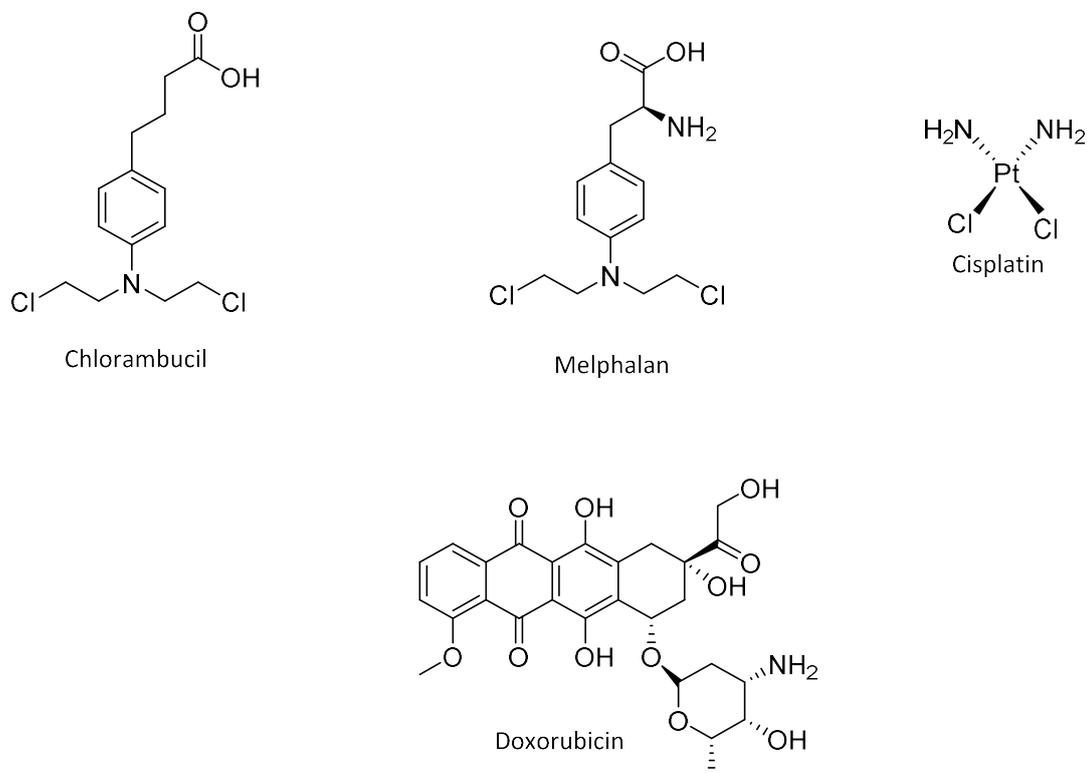


Figure 1.5. Chemical structures of DNA-interactive drugs.

To overcome the toxicity issues of unspecific cytotoxic drugs, much attention was focused on the development of more selective targeted agents.

A noteworthy example is the success of Gleevec against the tyrosine kinase domain of the abl kinase in chronic myelogenous leukaemia (CML).²⁸ However kinase inhibitors and other targeted agents show much less effect in solid tumors; even in CML, mutations in the ATP binding site of the abl kinase led to resistance to Gleevec. Therefore, although the therapeutic window of cytotoxic agents is very narrow, these agents are still employed for the management of tumors and DNA is still considered an interesting target for therapeutic intervention in cancer.²⁹

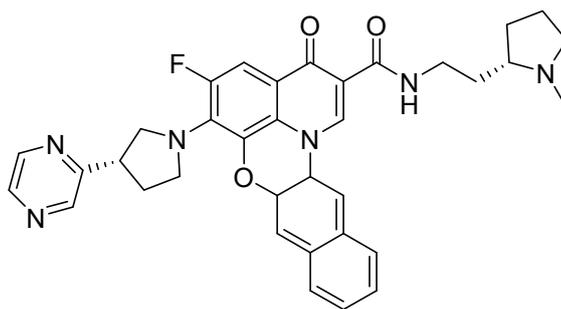
There has been much effort towards the design of DNA sequence-specific molecules that can modulate transcription of particular genes, which would not have the disadvantage of generating resistance at the protein level. Sequence-specific molecules have to recognise

sequences of ca 16–20 nucleotides, available through the polyamide approach by Dervan and co-workers,³⁰ and however they have not yet emerged as potential therapeutic agents.

An alternative and more realistic approach is to target specific DNA secondary sequences, called G-quadruplex. Since these DNA secondary structures found in guanine rich sequences of DNA, were observed, and their biological roles were proposed, they have been considered a new class of molecular targets for DNA-interactive compounds.³¹

The research community showed great interest in these DNA structures because targeting them may have the advantage to combine high target selectivity with the development of small-drug-like molecules.³²

In contrast to traditional duplex DNA binding agents, quarfloxin (Fig. 1.6), a synthetic G4 ligand which entered clinical trials, showed no toxicity and appeared to be well tolerated.³³



Quarfloxin

Figure 1.6. Chemical structure of the G-quadruplex ligand in clinical trials.

In addition, DNA interactive drugs may be used in combination with agents that target molecular receptors to increase selectivity for cancer cells and to reduce side effects on healthy cells. Therefore there are several reasons to expect DNA G-quadruplex to be considered a target of clinical interest for cancer treatment.

1.5. Tumor-homing peptides as vehicles for cancer therapeutics and diagnostics

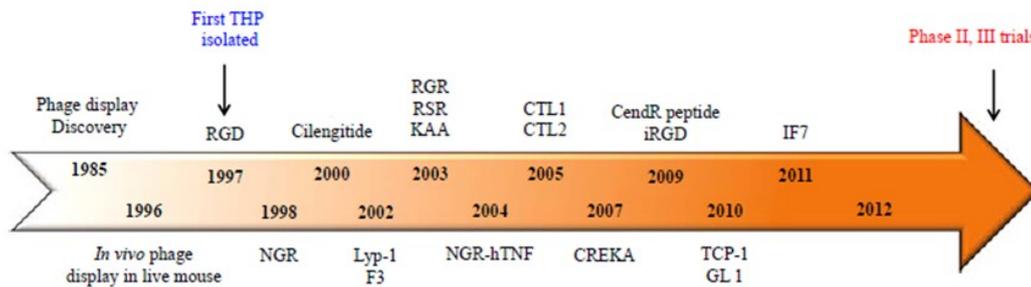


Figure 1.7. Tumor homig peptides journey.³⁴

Recent approaches in anticancer drug design aim to target tumour microenvironment, particularly the tumor vasculature.³⁵ Indeed a body of evidence demonstrate that tumour vessels have unique expression profile of surface protein and for this reason they may represent the ideal target to selectively deliver anticancer drugs into tumours.³⁶

Arap, Pasqualini et Ruoslahti identified by means *in vivo* phage display three peptide motifs capable of homing to tumor vasculature in mice bearing breast carcinoma xenografts: an RGD (arginine-glycine-aspartic acid) peptide motif,³⁷ a GSL (glycine-serine-leucine) motif and a NGR (asparagine-glycine-arginine) motif.³⁸

These peptide ligands are known as tumor homing peptides (THPs) for their ability to specifically bind tumor vasculature and consequently to accumulate in the tumor tissue. These THPs have emerged as promising tools to detect tumor *in vivo* and to deliver anticancer drugs specifically into tumor site, reducing side effects on healthy cells. THPs-based therapy may overcome the disadvantages associated with mono-therapy, such as drug sensitivity and resistance, pharmacokinetics and solubility issues. Considering the potential of THPs, the interest in this field grew esponentially and phage display technology has been extensively applied to identify peptides with homing properties towards various cancer types. Preclinical studies of many THPs-based agents are ongoing, and, interestingly, some of them have been submitted to clinical evaluation (Fig. 1.7).³⁴

Indeed, both RGD and NGR peptides have already entered clinical trials. Cilengitide (Fig. 1.8), a salt of a cyclized RGD pentapeptide acting as an integrin antagonist, is the first RGD-peptide based therapy in phase I, II, and III clinical trials for several cancers.³⁹

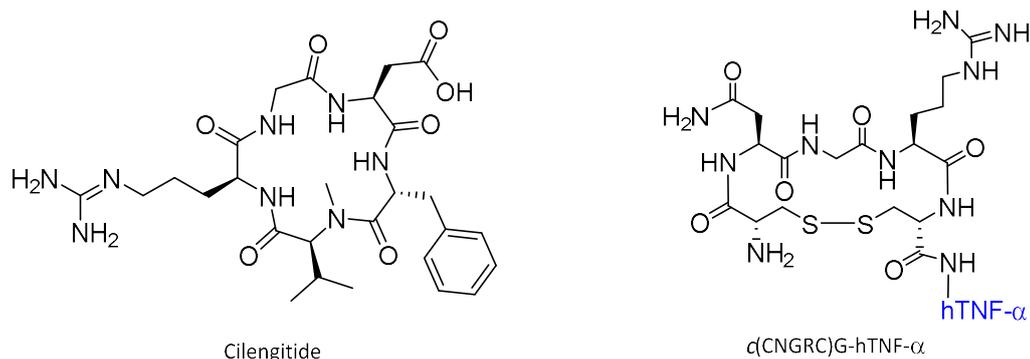


Figure 1.8. RGD and NGR-based agents in clinical trials.

Besides, the tracer [^{18}F]-galacto-RGD can be used for non invasive PET imaging of $\alpha_v\beta_3$ integrin expression in cancer.⁴⁰

As far as NGR-based agents are concerned, NGR peptide combined with hTNF- α (Fig. 1.8) is currently in phase III clinical trials for malignant pleural mesotelioma (www.clinicaltrials.com).

In this scenario the development of hybrid conjugates in which one component is a NGR containing tumour-homing peptide acting as a carrier which targets specifically tumour vessels, while the other one may be the active agent appears as a promising and realistic anticancer strategy. Also NGR might be derivatized with molecular probes for early tumor diagnosis and tumor angiogenesis detection and monitoring upon therapy.

1.6. Aim of the thesis

Since genomic alterations and their phenotypic expression in cancer are unique for each patient, replacing unspecific cytotoxic therapeutics with novel and personalized cancer therapies, represent the challenges for oncology research.

The targeting cancer drugs developed toward specific molecular targets are characterized by less nonspecific toxicity with respect to unspecific cytotoxic drugs. Since they inhibit one key pathway in a tumor, it may not be completely sufficient to impair a hallmark capability, allowing some cancer cells to survive with residual function. Tumor cells can respond by adaptation to the current therapy, through mutation, epigenetic reprogramming or remodelling of the stromal microenvironment, in order to establish the functional capability or to reduce their dependence on a particular hallmark capability.⁸

It may become possible to target several cancer capabilities to prevent adaptive resistance that cancer can develop in response to therapy. Different strategies need to be explored in the search of new anticancer agents in order to obtain more effective therapies for human cancer.

The aim of the present work is focused on two different approaches potentially useful in the clinical management of cancer.

The first one is based on the assumption that DNA secondary structures, known as G-quadruplexes, represent interesting targets for therapeutic intervention in cancer. In particular the stabilization by small molecules of these structures, found in crucial positions of the genome, is perceived as a promising anticancer strategy.

The project was based on the extension of a class of hydrazones, previously reported by Locatelli and co-workers, as G-quadruplex selective binders.⁴¹ The design, the synthesis and the G-quadruplex binding properties of a library of diimidazo[1,2-*a*:1,2-*c*]pyrimidine bis-guanylhydrazone derivatives and analogues will be described in the second chapter of this elaborate.

The second approach is focused on the development of hybrid drugs containing a tumour-homing peptide which targets specifically tumour vessels and that can be used as drug delivery tool. In the context of NGR-tumour homing peptides, the cyclic CNGRC peptide represents a valid carrier to deliver anticancer drugs or imaging probes into tumors.³⁸

The structural requirements of the NGR peptide for effective tumour homing through APN targeting were investigated in this project. The design, the synthesis and the biological evaluation of differently functionalized cCNGRC peptide derivatives will be reported in the third chapter of the thesis.

CHAPTER 2

2. Design and synthesis of novel hydrazone derivatives of diimidazo[1,2*a*-1,2*c*]pyrimidine as selective G-quadruplex binders and promising anticancer drug candidates

2.1. Background

2.1.1. DNA G-quadruplexes: topology, structure and biological roles

G-quadruplexes (G4s) are higher-order nucleic acid arrangements which arise from guanine-rich nucleic acid sequences. G4s are built by tetrads of Hoogsteen hydrogen-bonded guanine bases (G-quartets), that stack together in the manner of base pairs in duplex nucleic acids, showing, however, important differences: they constitute a four stranded structure in which sodium or potassium ions are held centrally between quartets and are coordinated to the O6 substituent of each guanine base. The G-quartets are held together by loops positioned on the exterior. Grooves, formed between the loops, are bounded by negatively charged phosphodiester backbones (Fig. 2.1).^{42–44}

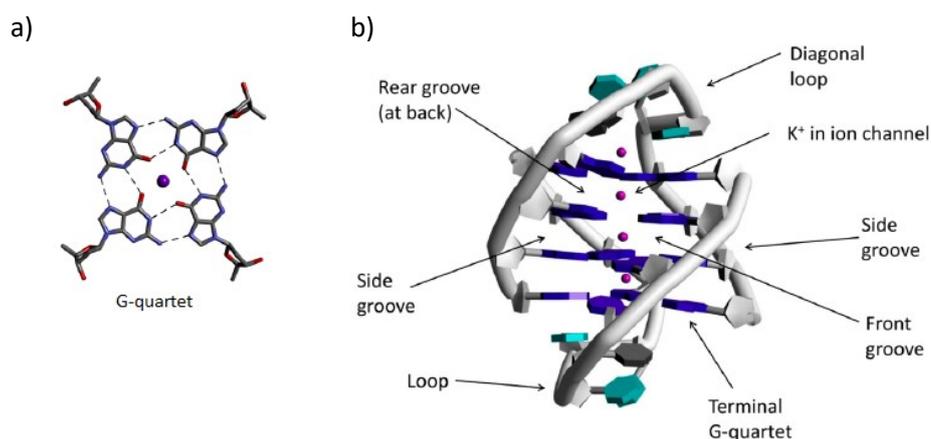


Figure 2.1. a) Stick representation of a guanine G-quartet with hydrogen bonds shown as dashed lines. b) Cartoon representation of a quadruplex (PDB code 1JPQ).⁴⁵

G-quadruplexes adopt several topologies depending on loop length and sequence, and show structural differences, regarding stoichiometry, chains orientation, loops geometry and glycosidic conformation. G4s can be formed from (i) a single strand, (unimolecular quadruplexes), having the general sequence $G_a X_n G_b X_o G_c X_p G_d$ where G_{a-d} represent short guanine (G) tracts and X_{n-p} represent intervening "loop" connections; (ii) two strands (bimolecular quadruplexes); (iii) three or four strands (tetramolecular quadruplexes).^{44,46} The four phosphodiester chains of the quadruplex can be parallel or anti-parallel, depending on the type of loop connecting them. Parallel strands require a connecting loop to link the bottom G-tetrad with the top G-tetrad, resulting in propeller type loops.^{47,48} In anti-parallel quadruplexes lateral (edge-wise)⁴⁹ and diagonal loops⁵⁰ have been observed to link adjacent G-strands and opposite G-strands respectively. Parallel G4s have all guanines in an *anti*-conformation, while the anti-parallel ones have shown both *syn*- and *anti*-conformations. All G4s have the common feature of a core of at least two stacked G-quartets, a central ion channel and four grooves.⁴⁴ The stability of G-quadruplexes strongly depends on monovalent cations. The guanine O6 oxygen atoms create a strong negative channel in which cations are located in a manner depending on the nature of the ion (Fig. 2.1). Na^+ can be located in plane with a G-quartet, or between two G-quartets, while the larger K^+ ion tends to have a central position between two proximal guanine tetrads.⁵¹

A number of quadruplexes have been characterized by biophysical methods (circular dichroism, fluorescence spectroscopy, NMR or crystallography), by either as native structures or as small-molecule complexes.⁵²

G-quadruplexes have been found in several G-rich regions, such as human telomers,⁵³ oncogene promoters⁵⁴ and 5'-untranslated regions (UTR).⁵⁵

The greatest concentration of G-quadruplex motifs has been found in telomers at chromosome ends, owing to the presence of the hexanucleotide repeat sequence d(TTAGGG) and G-quadruplex formation at this level may be involved in telomere maintenance.^{56,57} The telomeric G4s display a variety of folds. 2D-NMR studies in dilute solution in presence of K^+ ion revealed several folds with antiparallel strands, including hybrid topologies in which three of the four backbone strands are parallel and the fourth is in the antiparallel direction (Fig. 2.2 a),⁵⁸⁻⁶⁰ while under crowding and high

concentration conditions the parallel form is prevalent.^{61,62} Even the crystal structures of a human telomeric quadruplex containing K^+ ions show all strands parallel with propeller loops⁴⁷ and this structure may be the favourite one in cellular environments.

In parallel with the developing interest in telomeric G4s in the 1990s, potential quadruplex structures were identified in other positions of the genome, such as gene promoters of human oncogenes and cancer-associated genes, where quadruplex formation may have a role in transcription regulation.⁶³ G-rich sequences have found generally in nuclease hypersensitive/transcriptional activation sites within or upstream the promoter regions of genes including *c-MYC*,^{64,65} *BCL-2*,^{66,67} *c-KIT*,^{68,69} *HIF*,^{70,71} *RET*⁷² and *VEGF*.⁷³ Promoter G4s have showed wide sequence variability. 2D-NMR studies have revealed that the majority of the adopted quadruplexes in promoter regions have the parallel topology, typical of those sequences having at least one short (one or two single-nucleotide) propeller loop (Fig. 2.2 b).

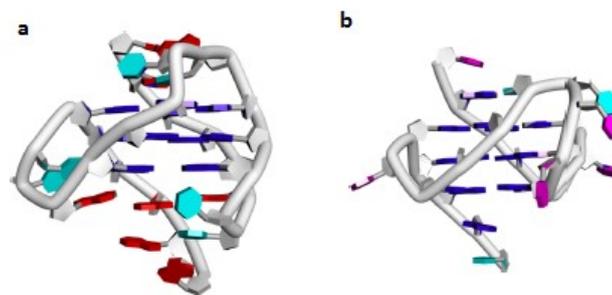


Figure 2.2. a) View of a (3 + 1) hybrid human telomeric quadruplex, (PDB code 2HY9) determined by 2D-NMR.^{45,59} b) View of the parallel quadruplex formed from a sequence in the human *VEGF* promoter (PDB code 2M27), determined by 2D-NMR.^{45,73}

It must also be said that G-quadruplex folding is very likely to occur within the single stranded RNA. It has been shown that G-quadruplex formation in the untranslated region of the *NRAS* oncogene results in inhibition of gene translation⁵⁵ These findings have stimulated a growing interest in this field.

2.1.2. G-quadruplex structures as targets for anticancer therapy

Targeting G-Quadruplexes In Telomeres

Telomeres are located at each end of chromosomes and their role is to protect them from catastrophic events.⁵⁷ In vertebrates, telomeric DNA comprises tandem repeats of the guanine-rich sequence d(TTAGGG), found in the single-strand overhang.⁷⁴ The 3' telomeric single strand DNA is associated with various proteins, such as telomeric-repeat-binding factor 1 and 2 (TRF1 and TRF2), which might be involved in loop-back structures,⁷⁵⁻⁷⁷ while telomere end is capped by protection of telomeres 1 (Pot1).⁷⁸ Telomeric guanine rich sequence is also able to adopt specific G-quadruplexes, which are recognised by several proteins, suggesting a biological role of these DNA secondary structures in telomere regulation.^{79,80}

In normal cells, since DNA polymerase is not able to fully replicate the lagging DNA strand, telomere length progressively decrease after each cell cycle and this phenomenon is known as 'end-replication effect'. Once the telomere shortening reaches the 'Hayflick limit', cells entered a phase of replicative senescence.⁸¹ In striking contrast, in cancer cells telomere length remains constant and this is dependent on the activity of a RNA-dependent DNA polymerase, known as telomerase.⁸² The telomerase catalytic subunit (hTERT in humans) active site provides to the assembly of nascent telomeric DNA onto the 3' telomeric DNA end. The elongation process occurs by hybridizing the terminal 3' end of the telomere on a complementary RNA template which is part of a large 451-nucleotide RNA domain (hTR in humans) of the enzyme.⁸³

Telomerase catalyses telomere extension, leading to cell immortalization: it is overexpressed in around 85% of all human cancers and it is involved in tumorigenesis.^{84,85} Inhibition of telomerase would result in effective anticancer effects, however direct telomerase inhibition correlates with an extended time-lag before telomere length reaches the critically shortening, which would lead to senescence and apoptosis.⁸⁶

An alternative strategy is the indirect inhibition of telomerase, by targeting telomeric G-quadruplex with small molecules (Fig. 2.3).⁸⁷ The enzymatic activity of hTERT can be inhibited by the 3' telomeric single-strand folding into a quadruplex structure, because telomerase can't attach to the single strand DNA. Quadruplex formation in the telomeric

DNA effectively inhibit the hybridization between the DNA and the RNA template.⁸⁸ G4s stabilizing small molecules can inhibit the ability of the telomerase enzyme complex to maintain the integrity of telomeres in cancer cells and this results in rapid DNA damage and consequent cell death, while direct telomerase enzymatic inhibitors induce these effects after the full extent of telomeric DNA shortening. ‘Telomere targeting’ is used to describe the behavior of the well-studied G4-binding compounds RHPS4⁸⁹ and BRACO-19,⁹⁰ which target the single-stranded telomeric DNA overhang in tumor cells, leading to a rapid uncapping of hPOT1 and telomerase, thus resulting in telomerase inhibition and activation of DNA damage responses and cellular apoptosis.⁹¹

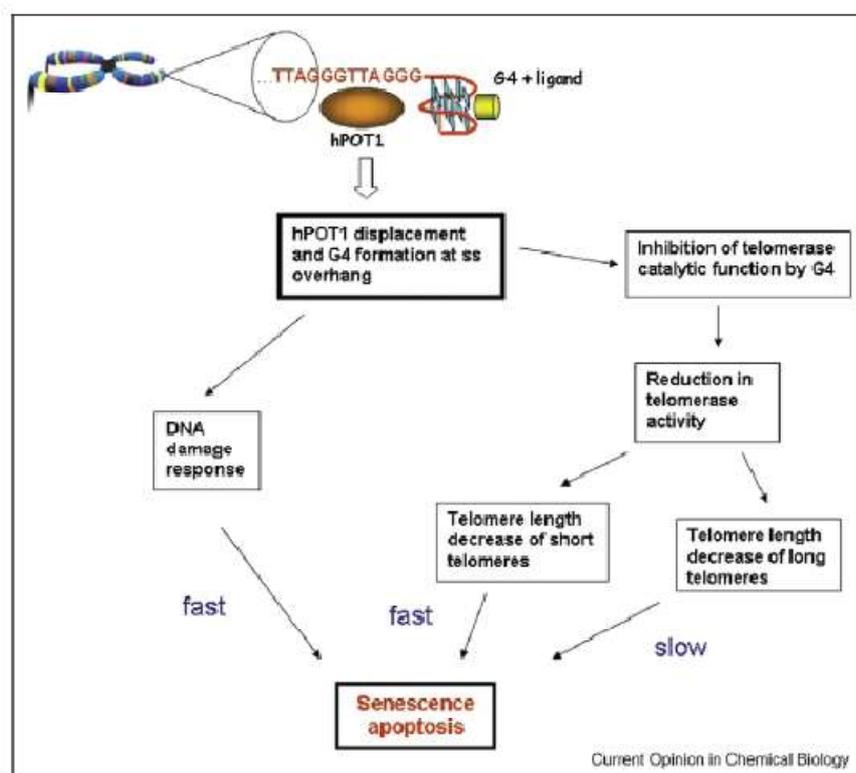


Fig.2.3. Scheme of the cellular consequences of inducing telomeric quadruplex formation by a small-molecule.⁹²

Targeting G-quadruplexes in oncogene promoters

G-quadruplex forming sequences have been found upstream and close to the transcription start site of several genes. These findings supported the hypothesis that quadruplexes play an important role in gene regulation (Fig. 2.4).⁹³

G-quadruplex stabilization at this level by small molecules may alter the transcription of several oncogenes of relevance in human cancer, such as *c-MYC* and *c-KIT*, thus exerting anticancer effects.⁹⁴

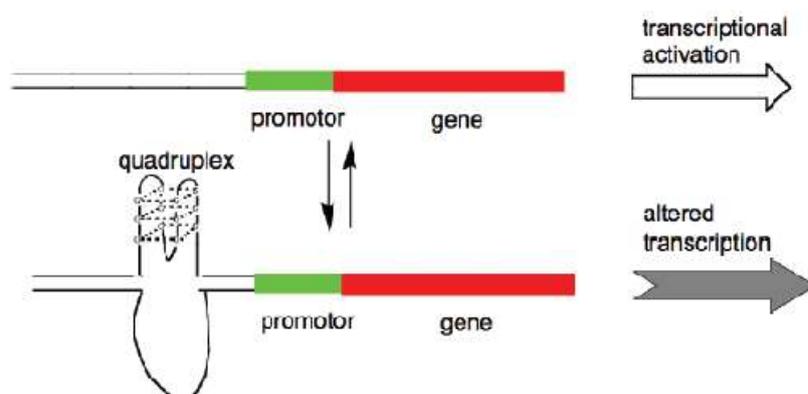


Figure 2.4. Model for transcription modulation via formation of a quadruplex in a promoter region.⁹³

The *c-MYC* oncogene, which encodes a transcription factor involved in cell cycle progression, apoptosis and cellular transformation and whose expression is deregulated in up to 80% of all solid tumours, represents a potential therapeutic target.^{95,96}

However, it has been considered undruggable at the protein levels, owing to its short half-life.⁹⁷ To overcome this issue, targeting *c-MYC* at the gene level may represent a promising anticancer approach. The pivotal publication in this field reported that the transcription of *c-MYC* oncogene could be repressed by targeting a G-quadruplex in the nuclease hypersensitive element (NHE III) in the promoter regions of *c-MYC* with a small molecule (TMPyP4).⁹⁸ It was demonstrated that also quindolines and actinomycin D affect *c-MYC* transcription, by stabilizing G-quadruplex.^{99,100}

The proto-oncogene *c-KIT* encodes a receptor tyrosine kinase that plays a key role in cell proliferation, differentiation and survival.¹⁰¹ Gastrointestinal stromal tumours (GIST) pathogenesis has been correlated to the constitutive activation of *c-KIT*.¹⁰² Kinase inhibitors imatinib (Gleevec; Novartis) and sunitinib (Sutent; Pfizer) target c-KIT at the

protein level. However drug resistance is an emerging issue and a novel approach may be targeting the G-quadruplex in promoter regions of this oncogene.¹⁰³

Two G-quadruplex-forming sequence motifs were identified in the core promoter of the human *c-KIT* oncogene and they are known as c-kit1⁶⁸ and c-kit2.¹⁰⁴

It was shown that an isoalloxazine small-molecule binds both c-kit G-quadruplexes and reduce the levels of *c-KIT* transcription in a cell line.¹⁰⁵

2.1.3. G-quadruplex binders

As stated in the above paragraphs, G4s are considered interesting targets, at both telomeric and oncogene levels, and their stabilization by small molecules represents a promising strategy for therapeutic intervention with focus on cancer. All quadruplex structures are very different from duplex nucleic acids, therefore they may be potentially recognized by molecules in a specific manner.

The growing interest in G-quadruplex structures as targets led to the identification of over 1000 G4-ligands, representing a wide range of chemotypes.^{32,45,106}

Some of G4 ligands have been identified in order to target the grooves or the loops of G4 structures, although the majority of them primarily stack on the external G-tetrads of G4s (Fig. 2.5).¹⁰⁷

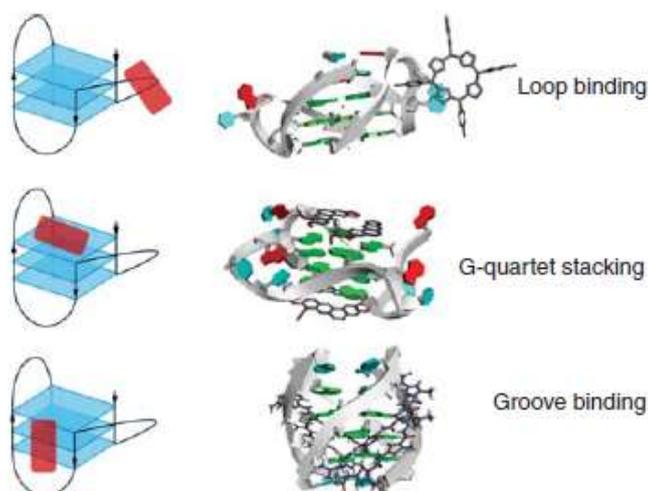


Figure 2.5. Examples of the binding modes of ligand to G-quadruplex.¹⁰⁷

The structural features that facilitate the stacking binding mode are: (i) the presence of an extended planar aromatic scaffold that can efficiently stack on terminal G-tetrads through π - π stacking, (ii) an electron-deficient aromatic core that can enhance stacking interactions on G-tetrads via cation- π interaction, and (iii) the presence of cationic side-chains which can interact with the grooves and loops of G4 and with the negatively charged phosphate backbone.

The G4 ligands can be classified in two main structural groups: i) polycyclic and ii) cyclic/acyclic non-polycyclic ligands.³²

i) Polycyclic G4 ligands

The first identified G-quadruplex interactive ligand able to inhibit telomerase *in vitro* was the amido-anthraquinone derivative reported in Fig 2.6.⁸⁷ Owing to its structure, planarity started to be perceived as an essential feature for effective G4 binding.

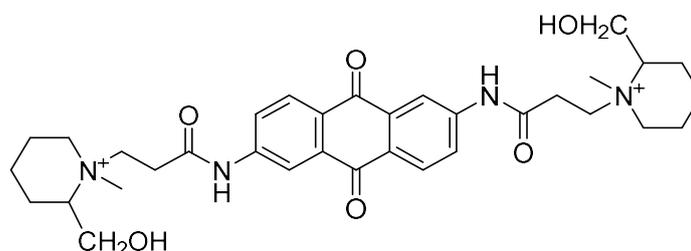


Figure 2.6. Anthraquinone derivative.⁸⁷

The acridine-based compounds such as RHPS4¹⁰⁸ and BRACO-19^{90,109} (Fig.2.7) are potent G4 stabilizers which inhibit telomerase activity. They have shown low micromolar *in vitro* cell growth inhibition activities and also *in vivo* antitumor activity against xenograft models.

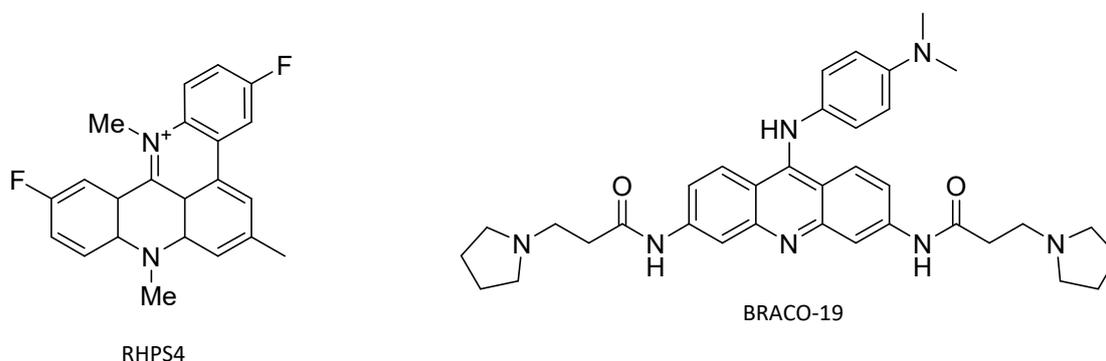
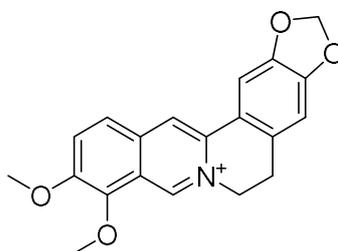


Figure 2.7. Acridine derivatives.

X-ray crystallographic data are available on a human telomeric G4 complex with BRACO-19 and show the BRACO-19 molecule bound to one end of the G4, making π - π interaction with the terminal G-quartet and with side chains held within loops.¹¹⁰ This external binding mode is different from the internal binding of intercalating agents that interact with duplex nucleic acids and this difference may be relevant in terms of G4 selectivity over duplex DNA.

A number of chemotypes have been identified from natural products. Notable examples are the plant alkaloid berberine (Fig. 2.8), which binds human telomeric DNA with end-stacking features analogous to BRACO-19.¹¹¹



Berberine

Figure 2.8. Example of G-quadruplex-interactive natural molecule.

The expansion of the aromatic planar surface of berberine led to the quinolinobenzo-[5,6]-dihydroisoquinolinium scaffold, bearing an additional pyridine ring and amino group (Fig. 2.9) with respect to the natural alkaloid, thus increasing the selectivity for *c-myc* G-quadruplex. The derivative reported in Fig. 2.9 was able to induce the down-regulation of *c-MYC* transcription in cancer cells.¹¹²

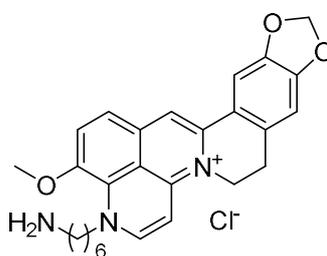


Figure 2.9. Quinolinobenzo-[5,6]-dihydroisoquinolinium derivative.

Tetra-substituted naphthalene diimide derivatives (NDs) with positively charged termini are a distinct group of potent G4-binding agents. The pharmacological properties of ND compounds were improved by replacing two *N*-methyl-piperazine groups by less basic morpholine groups, without affecting quadruplex interaction. The lead compound reported in Fig. 2.10 has shown low nanomolar potency against several cancer cell lines.¹¹³

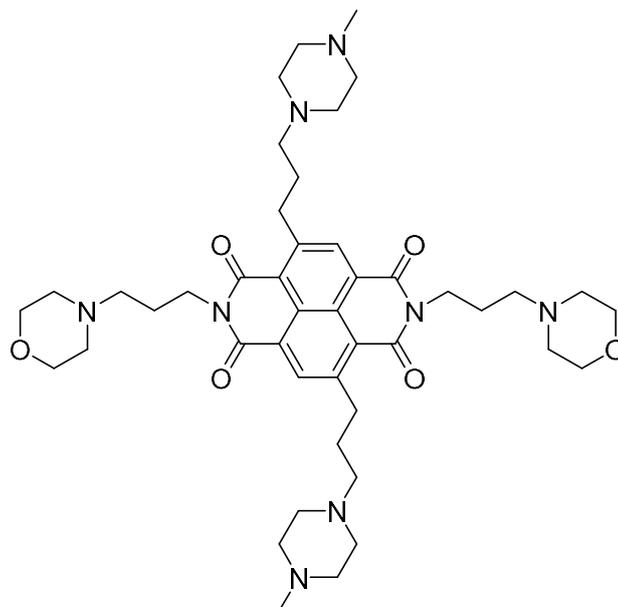


Figure 2.10. Naphthalene diimide derivative.

For *c-myc* promoter quadruplexes only 2D-NMR structures have been determined. A noteworthy example is the monosubstituted quindoline 2:1 complex, which shows the ligand bound to each terminal G-quartet face, with additional stabilization by interaction with 5' and 3' flanking nucleotides (Fig. 2.11).¹¹⁴

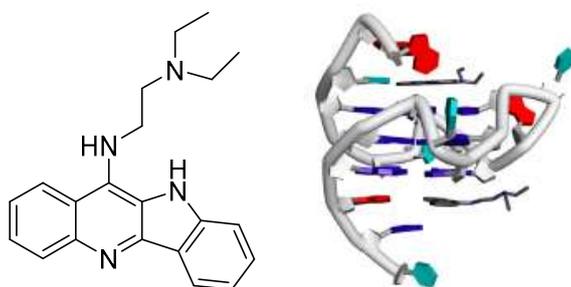


Figure 2.11. View of the 2:1 complex between a monosubstituted quindoline compound and a *c-myc* promoter complex, determined by 2D-NMR (PDB code 2L7V).¹¹⁴

The synthetic compound quarfloxin, (CX-3543) (Fig. 2.12) based on a pentacyclic benzo[*b*]pyrido[3,2,1-*k*]phenoxazine (i.e., a fluoroquinolone core), is the sole G4-binding small molecule that has progressed to clinical trials. It's believed that quarfloxin binds multiple G4s in ribosomal DNA, affecting interaction with the protein nucleolin, leading to ribosome synthesis inhibition and apoptosis.³³

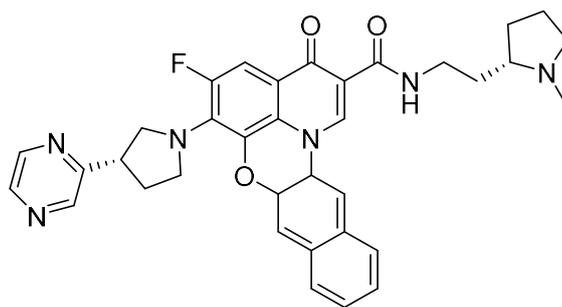


Figure 2.12. Chemical structure of quarfloxin.

ii) *Cyclic/acyclic non-polycyclic G4 ligands*

The macrocycle telomestatin (Fig. 2.13) comprising hepta-oxazole and a single chiral dihydro-thiazole group was isolated from *Streptomyces anulatus*.¹¹⁵ Telomestatin showed high G4 affinity and selectivity, stabilizing the anti-parallel form of the human telomeric G4.¹¹⁶, and strongly inhibited telomerase function and had significant antitumour activity in vitro and in vivo.¹¹⁷ It is atypical G4 stabilizer, as it does not have any protonated side chains.

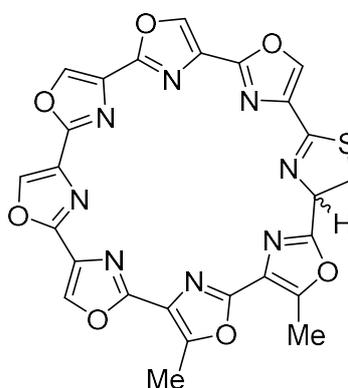


Figure 2.13. Chemical structure of telomestatin.

Recent findings showed that pyridostatin (Fig. 2.14), a G4-selective ligand,¹¹⁸ causes DNA damage in cells predominantly at the SRC gene, which contains a high number of putative G4-forming sequences.¹¹⁹

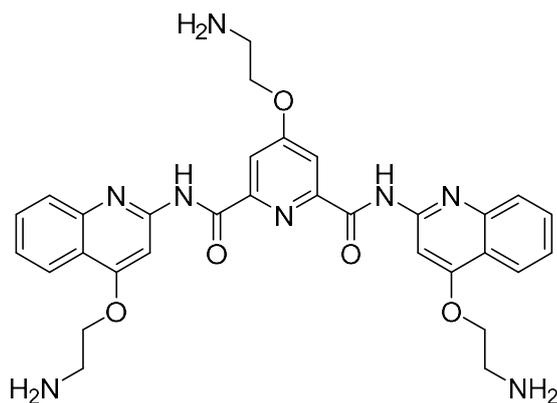


Figure 2.14. Chemical structure of pyridostatin.

In many studies reported so far G-quadruplex interactive small molecules have showed also biological effects with variable degree of success. However, the development of such molecules as anticancer drugs still requires further efforts especially because of the selectivity issue. Higher selectivity for G-quadruplex over duplex DNA is required. Moreover the selective activity towards a single G-quadruplex structure needs further investigations. Another key problem is that, in traditional medicinal chemistry terms, most of the identified molecules are not ideal since they show features which don't match with Lipinski criteria for orally-bioavailable drug-like compounds. Indeed, large planar molecules are not normally considered to be ideal pharmacophores, for their hydrophobicity. Small saturated rings such as pyrrolidine, piperidine or piperazine are commonly used as well as tertiary amines such as -NMe₂ or -NEt₂ to enhance aqueous solubility and G4 affinity, while the number of side-chains can be important for enhancing selectivity between different G4s.

2.2. Aim of the project

Among G-quadruplex binders reported in literature, compound **1** (Fig. 2.15), which emerged from a screening of diimidazo[1,2-*a*:1,2-*c*]pyrimidine derivatives developed by my research group, was identified as a potent stabilizer of several G-quadruplexes over duplex DNA.⁴¹ This compound doesn't show conventional alkylamino substituents, with respect to other G-quadruplex ligands, but instead has two structurally more rigid guanylhydrazone groups. However compound **1** showed a very modest antiproliferative activity probably because it is not readily taken up into most cells.

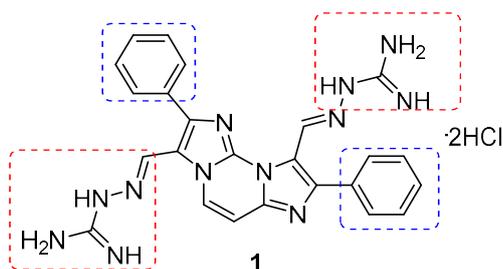


Figure 2.15. Bis-guanylhydrazone derivative of diimidazo[1,2-*a*:1,2-*c*]pyrimidine.

A library of analogues of **1** was designed and synthesized to shed light on the structural requirements for the effective binding to G4s (Fig. 2.16). The effects of structural changes of **1** involving in particular the guanyl chains, the substituents at positions 2 and 8 and the core itself were investigated employing both human telomeric and oncogene promoter G4s with different topologies as targets.

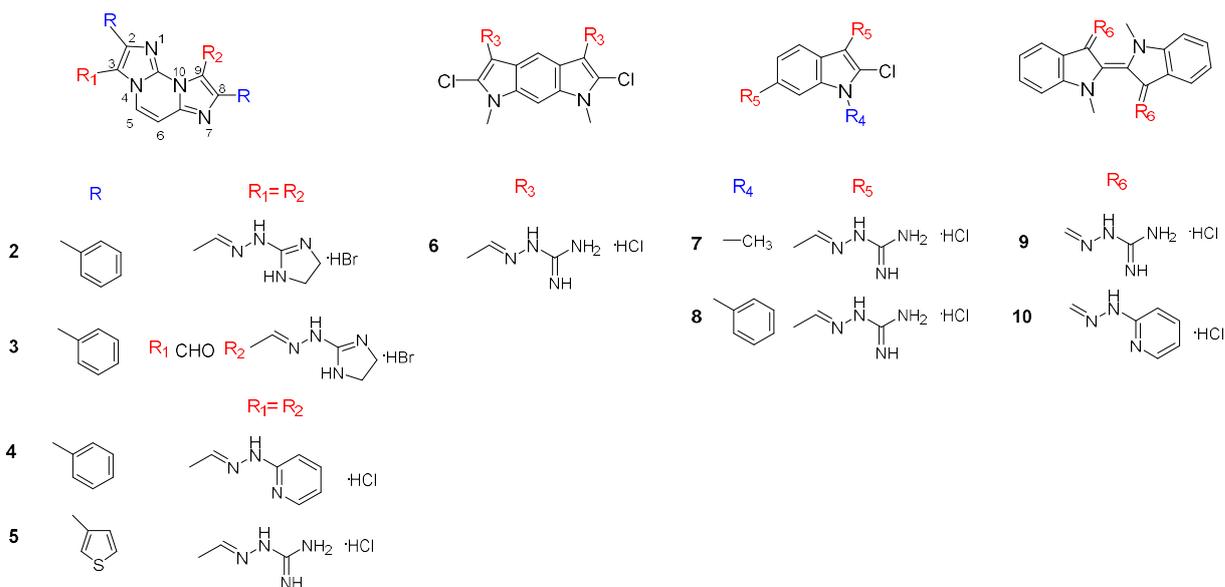


Figure 2.16. Library of designed hydrazone derivatives.

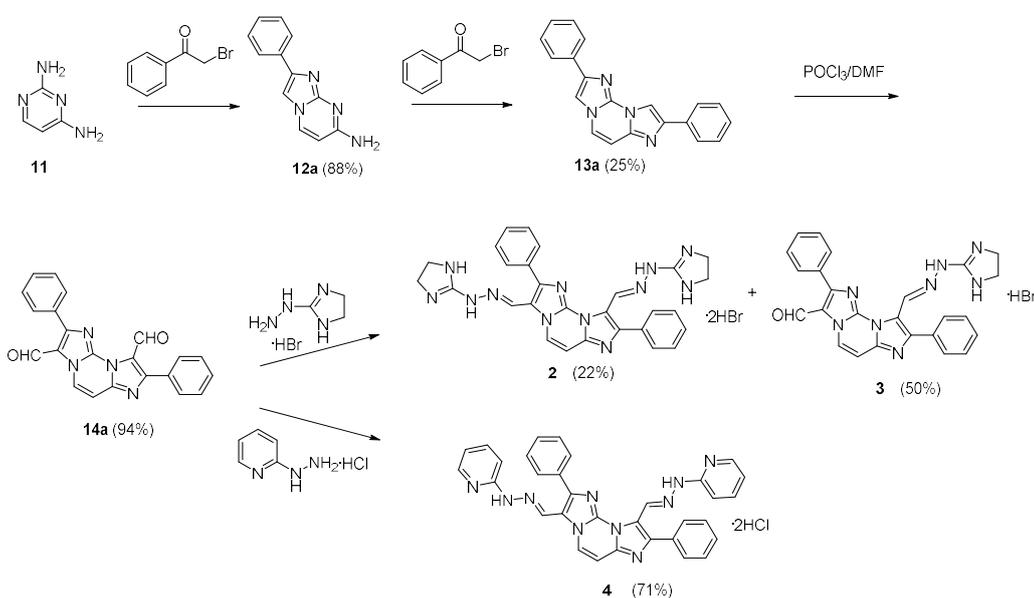
The guanyl chains have been replaced with an imidazoline moiety (compounds **2** and **3**) or hydrazinopyridyl group (compound **4**). The guanyl chains have been maintained, while phenyl rings at positions **2** and **8** have been replaced with the bioisostere thiophene in compound **5**. The tricyclic core has been substituted with different systems: the pyrroloindole nucleus (compound **6**), a chloroindole moiety (compounds **7**, **8**). The planar bisindole system of indigo was introduced, and then functionalized with iminoguanyl or hydrazinopyridyl groups in order to obtain compounds **9** and **10**.

The binding properties of the synthesized compounds to human telomeric and oncogene promoter G4s with different topologies as targets were investigated *in vitro*. The compounds that show the most interesting stabilizing effects were further studied by applying biophysical methodologies. Finally biological studies in cancer cell lines were performed in order to study the association between the G4 stability effect in living cells and potential anticancer activity of selected hydrazone derivatives.

2.3. Results and discussion

2.3.1. Chemistry

The hydrazones **2–5** (Schemes 2.1 and 2.2) and **6–10** (Schemes 2.3–2.5) were synthesized by reaction between an aldehyde or a ketone and the appropriate hydrazine: aminoguanidine hydrochloride, 2-hydrazino-2-imidazoline hydrobromide, or 2-hydrazino-2-pyridine hydrochloride.¹²⁰



Scheme 2.1.

The diimidazopyrimidine core bearing two phenyl rings in positions 2 and 8 (compound **13a**) was obtained in two steps starting from 2,4-diaminopyrimidine **11** in presence of 2-bromoacetophenone. The corresponding bis-aldehyde (compound **14a**) was synthesized through a Vilsmeier reaction in presence of POCl₃ and DMF by means of a procedure described in literature (Scheme 2.1).¹²¹

The next step was the reaction between 2-hydrazino-2-imidazoline and aldehyde **14a** and two products were isolated from the reaction mixture. Indeed, in addition to the bis hydrazone **2**, it was obtained also the mono hydrazone derivative **3** (Scheme 2.1).

NMR experiments were performed to clarify which aldehyde group reacted to give the corresponding mono hydrazone **3**. In particular two NOE experiments confirmed that the formyl group at 9 position was more reactive than that one at position 3. Indeed the irradiation of the CHO at 10.09 ppm provided a NOE correlation with one of the pyrimidine proton at 9.12 ppm (Fig. 2.17) whereas when the CH=N at 9.47 ppm was irradiated, no NOE effect at the pyrimidine proton was observed. These two findings suggest that the unreacted aldehyde is at position 3.

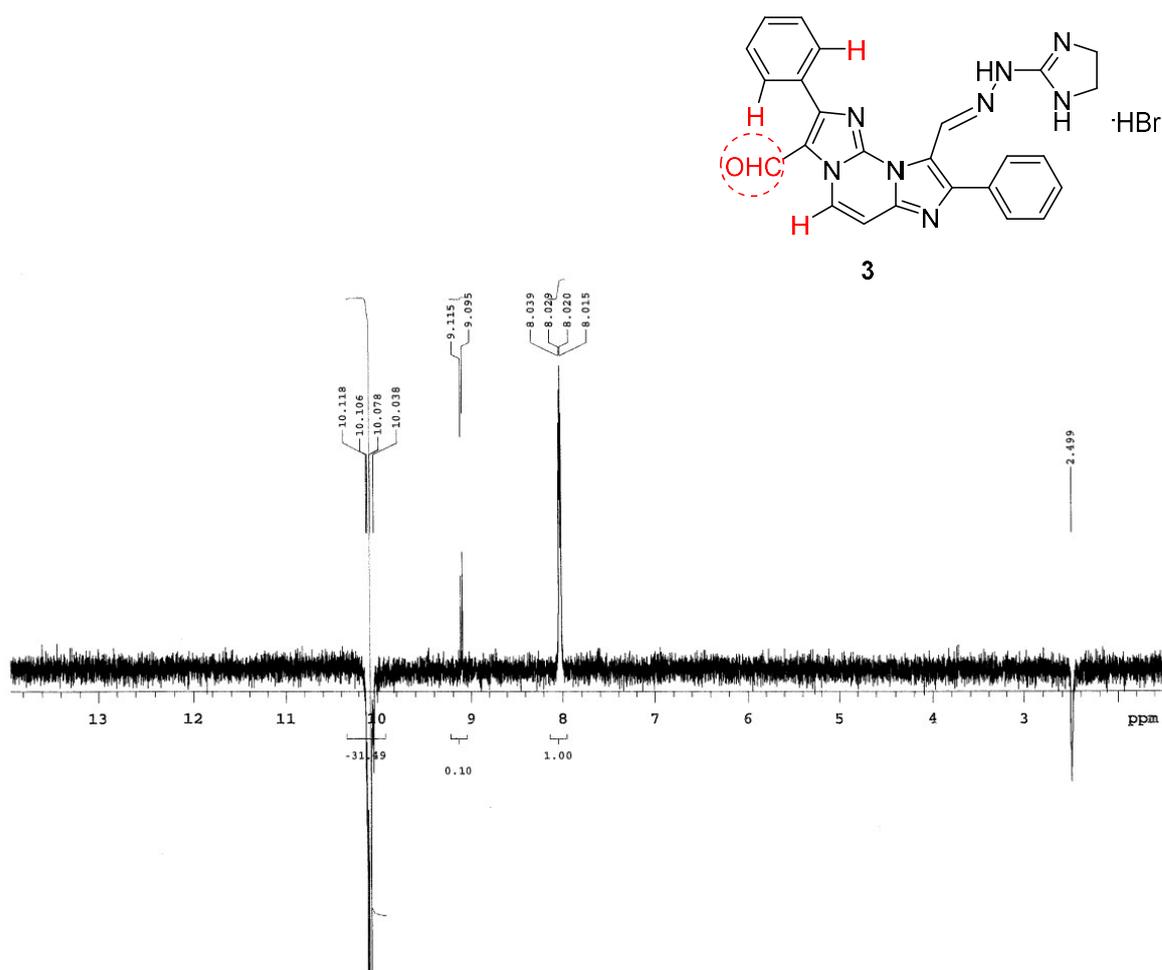


Figure 2.17. NOE experiments on compound **3**: irradiation of formyl group at 10.09 ppm.

On the other hand, when bis-aldehyde **14a** was allowed to react with 2-hydrazino-2-pyridine, only the corresponding bis-hydrazone derivative **4** was obtained (Scheme 2.1).

All the hydrazones have been obtained as hydrochlorides or hydrobromides, in which, as previously reported,¹²² the side chains are protonated and the positive charge is delocalized on the nitrogens of the guanyl or of the imidazoline groups.

Indeed, ¹H NMR spectra of guanylhyazone derivatives display at about 8 ppm a broad signal integrating to eight protons and exchanging after D₂O addition, therefore the two guanyl groups are protonated. Also the ¹H NMR spectrum of compound **3** (Fig. 2.18) show a broad signal at about 8.5 ppm integrating to two protons due to the protonated imidazoline nitrogens. Interestingly, owing to protonation and charge delocalization, the imidazoline ring become symmetric, therefore the protons of the two -CH₂ are chemically equivalent and give only one ¹H NMR signal: a singlet at 3.73 ppm that integrates to four protons.

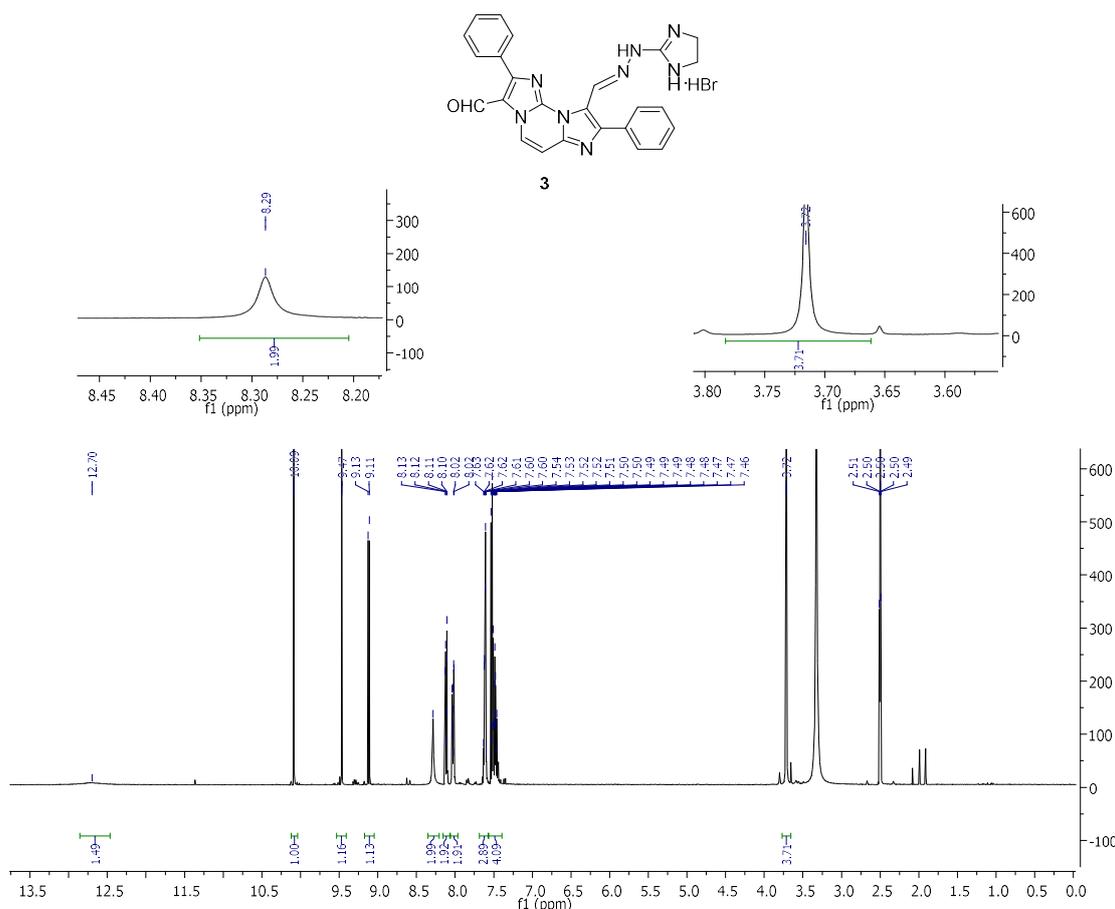
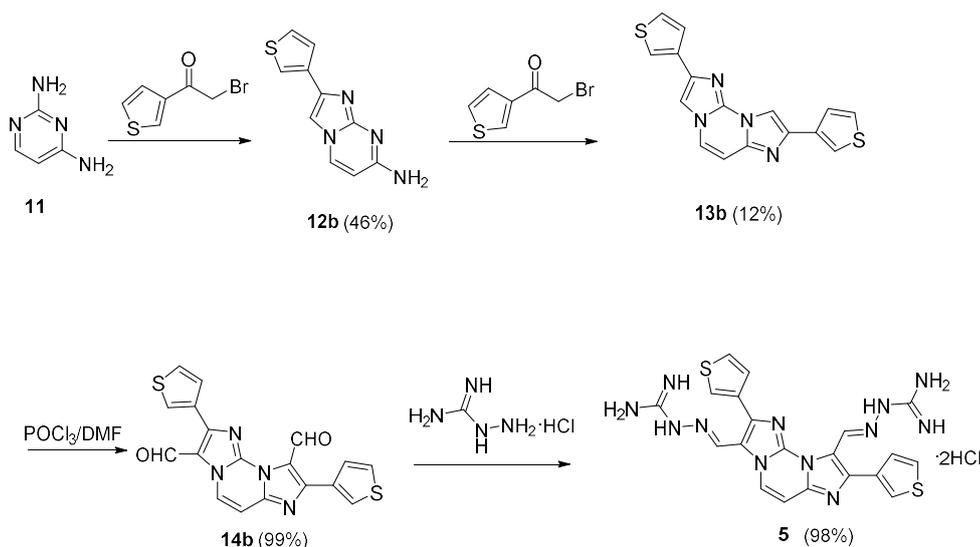


Figure 2.18. ¹H NMR spectrum of compound **3**.

Compound **13b** (Scheme 2.2) was obtained in two steps by reacting 2,4-diaminopyrimidine **11** with 2-bromo-1-(thiophen-3-yl)ethan-1-one. Even when **11** was treated with an excess of 2-bromo-1-(thiophen-3-yl)ethan-1-one in order to obtain **13b** directly, the intermediate **12b** is formed as the hydrobromide and it was necessary to prepare the corresponding base before treating it again with 2-bromo-1-(thiophen-3-yl)ethan-1-one.



Scheme 2.2.

Analogously to compound **12a** (Scheme 2.1),¹²³ and as determined by two NOE experiments, the first cyclization involved the nitrogen at position 1 and the amino group at position 2 and led to the bicyclic compound **12b**.

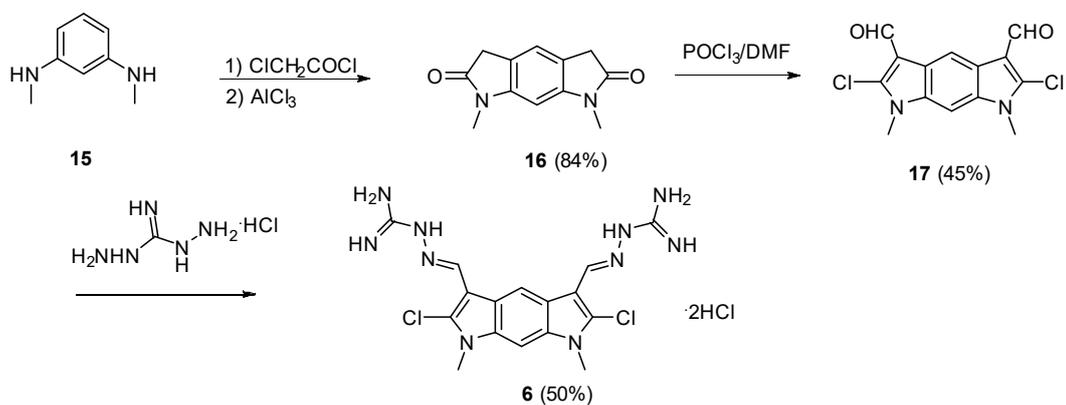
The first NOE experiment confirmed the spatial closeness of the pyrimidine proton at position 6 and the amino group, as suggested by NOE correlation between the doublet at 6.37 ppm (H-6 pyrimidine) and the signals at 8.36 ppm (the other pyrimidine signal) and 7.16 ppm (amino group). The second NOE experiment, performed by irradiating on the doublet at 8.36 ppm, provided a correlation with the signals at 6.37 ppm (H-6 pyrimidine) and 7.74 ppm (imidazole), thus confirming the structure of **12b**.

The second cyclization between **12b** and 2-bromo-1-(thiophen-3-yl)ethan-1-one occurred with lower yield than the first one.

The aldehyde **14b** was obtained by means of the Vilsmeier reaction on compound **13b** performed under reflux in order to promote bis-formylation. In the last step of the

synthetic procedure, aldehyde **14b** reacted with aminoguanidine hydrochloride to give the corresponding bis-hydrazone **5** (Scheme 2.2).

The diimidazopyrimidine core of the lead compound **1** was replaced with the pyrroloindole nucleus in compound **6** (Scheme 2.3).

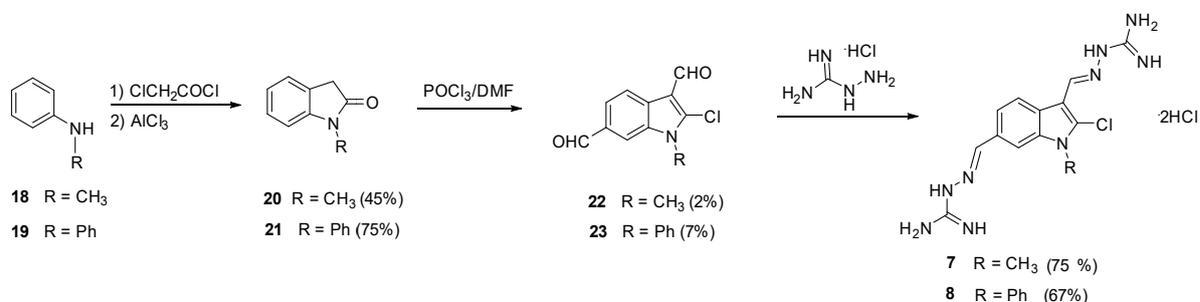


Scheme 2.3.

Starting from *N*¹,*N*³-dimethylbenzene-1,3-diamine (**15**), in presence of chloroacetyl chloride, the acylation of the amino groups was achieved and the intermediate thus obtained was cyclized by Friedel–Crafts alkylation with AlCl_3 .

The resulting pyrroloindole, compound **16**, was then allowed to react under reflux conditions with the Vilsmeier reagent, formed *in situ* to achieve the bis-formylation. The obtained compound **17**, in presence of aminoguanidine hydrochloride, gave the corresponding bis guanylhydrazone **6** in good yield.

A chloroindole moiety was introduced in compounds **7** and **8**, to achieve core simplification. The chloroindole was synthesized following a similar procedure which led to the pyrroloindole scaffold (Schemes 2.3 and 2.4).



Scheme 2.4.

Compounds **20** and **21** reacted with an excess of Vilsmeier reagent formed *in situ* between POCl₃ and DMF, in order to obtain the bis-aldehydes **22** and **23**, which were isolated by column chromatography. The positions of the formyl groups were determined using NOE experiments. The aromatic signals of compound **22** were assigned by irradiating the methyl group at 3.93 ppm. NOE correlation was observed with the signal at 8.24 ppm, therefore this singlet was assigned to the proton at position 7 (indole-7) (Fig. 2.19)

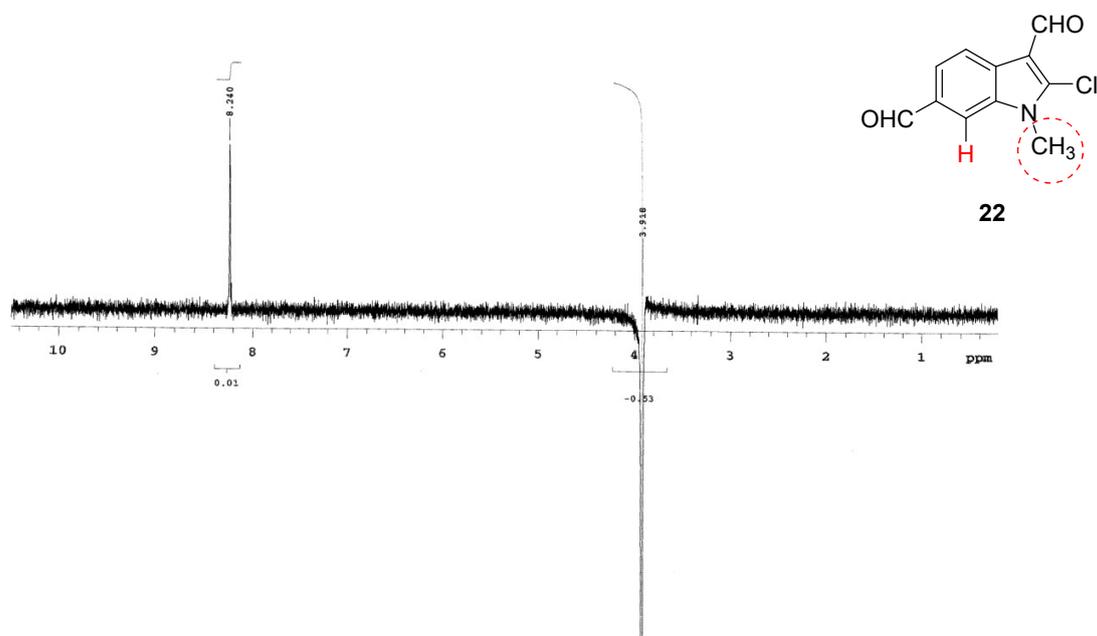


Figure 2.19. NOE experiment on compound **22**: irradiation of methyl group at 3.93 ppm.

Also the formyl signals were irradiated. When the CHO proton at 10.03 ppm was irradiated, no NOE effect was observed, thus this signal was assigned to the CHO group at position 3 (Fig. 2.20 a). On the other hand, irradiation at 10.07 ppm provided a NOE with signals at 8.24 ppm (indole-7) and 7.83 ppm (indole-5) (Fig. 2.20 b), thus indicating that this formyl group is at position 6. Similar NOE experiments performed on compound **23** confirmed that the formyl groups are at positions 3 and 6.

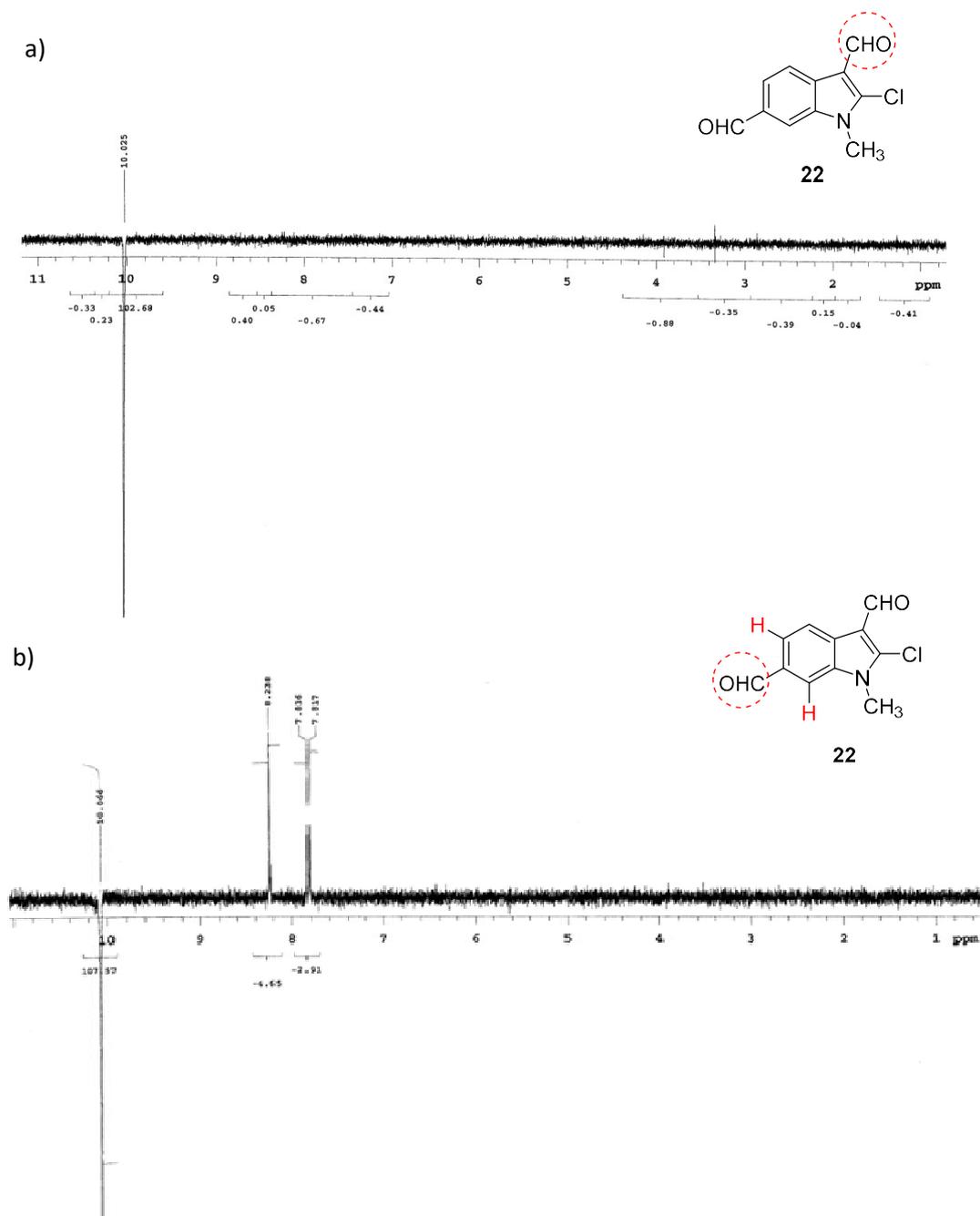
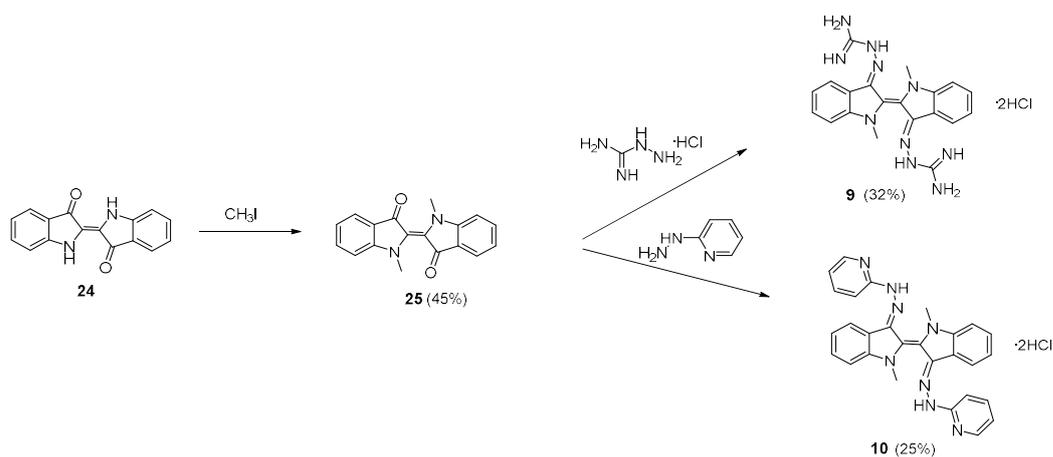


Figure 2.20. NOE experiment on compound **22**. a) Irradiation of formyl group at 10.03 ppm. b) Irradiation of formyl group at 10.07 ppm.

The planar bisindole system of indigo was introduced as central core in compounds **9** and **10**. In order to avoid intramolecular H bonds, which would make the ketones less reactive, it was necessary to functionalize the amino groups of indigo (compound **24**) with methyl groups in presence of methyl iodide (Scheme 2.5) in order to obtain compound **25**, as reported in literature.¹²⁴



Scheme 2.5.

Therefore compound **25** was subsequently used in the reaction with aminoguanidine hydrochloride, or 2-hydrazino-2-pyridine hydrochloride to afford compounds **9** and **10** respectively (Scheme 2.5).

2.3.2. Biophysical studies

*Circular dichroism and FRET studies*¹²⁰

The synthesized compounds **1–10** were evaluated for their ability to interact and to stabilize several G-quadruplex structures with different topology.

To this aim G4-forming sequences able to form parallel, antiparallel, and hybrid G4 structures were selected.

Human telomeric DNA G4s can adopt different topologies depending on the selected sequence and experimental conditions.¹²⁵

Two human telomeric DNA sequences, namely Tel23 and Tel26, able to form the hybrid-1 fold and the hybrid-2 fold respectively, were selected for this study.

However several studies suggest that the parallel G4 fold⁴⁷ is the prevalent one in the cell.¹²⁶ Therefore, the parallel G4 conformation (namely Tel23-p) was also used in this study and it was prepared from Tel23 sample, using high DNA concentration conditions in order to promote the parallel topology.

Moreover the two G4-forming sequences from the nuclease hypersensitive region of the *c-KIT* promoter (c-kit1 and c-kit2) and one from the *c-MYC* promoter (c-myc) were also used.

Circular dichroism (CD) spectroscopy was used to verify the structures adopted by each G4 sample.

Tel23-p, c-kit1, c-kit2, and c-myc displayed a positive band at 264 nm and a negative one around 240 nm in the CD spectrum, which are characteristic of parallel-stranded G4 topologies, according to data reported in literature.^{127,128} On the other hand, Tel23 and Tel26 showed very similar CD spectra having a positive band at 289 with a shoulder at ca. 268 nm and a weak negative band at around 240 nm. These data are consistent with the presence of hybrid structures as major conformations.

Moreover it was observed, by means additional CD experiments, that in presence of 4 mol. equiv. of each compound, no significant variations of CD signal of the analyzed G4 structure were obtained, suggesting an overall preservation of the G4 native topology upon the addition of each ligand.

Also the structure of a duplex DNA (ds12) was also verified by CD spectroscopy.

In the presence of K^+ , the CD spectrum of ds12 was characterized by a positive band at around 280 nm and a negative one at 250 nm and even these bands were not modified upon ligand addition.

CD melting experiments were performed in order to evaluate the stabilizing properties of the newly synthesized compounds. To this aim, the ligand-induced change in the melting temperature (ΔT_m) of G4 and duplex structures were measured, using a 4:1 ligand/DNA so that all possible G4 binding sites were available. CD-melting curves of DNAs in the absence and presence of each ligand were obtained by following the variations of the intensities of CD signals at the wavelength of 264, 290, and 280 nm for parallel G4s, antiparallel G4s, and duplex, respectively.

The results of the CD melting experiments are listed in Table 2.1.

Table 2.1. Ligand-induced thermal stabilization of G4 and duplex DNAs measured by CD melting experiments.

Compound	ΔT_m ($^{\circ}\text{C}$) ^a						
	c-kit1 ^b	c-kit2 ^c	c-myc ^c	Tel23-p ^b	Tel23 ^b	Tel26 ^b	ds12 ^b
1 ^d	>15.0	>20.0	>20.0	>20.0	-10.0	-4.5	0.1
2	1.6	>20.0	>15.0	10.8	-14.0	-8.9	1.3
3	2.7	9.5	>20.0	9.1	-2.4	-3.0	-0.8
4	0.4	0.0	-0.2	1.3	-1.0	-1.6	0.9
5	>15.0	>20.0	>20.0	>20.0	-4.9	-6.5	3.5
6	>15.0	>15.0	>20.0	>20.0	3.0	1.3	ND ^e
7	>15.0	>20.0	>20.0	>20.0	6.0	5.8	ND ^e
8	>15.0	>15.0	>20.0	>20.0	5.6	5.1	ND ^e
9	0.9	8.2	11.0	7.7	-0.5	0.2	-0.8
10	1.5	2.0	1.6	4.0	-2.5	-2.0	-0.7

^a ΔT_m represents the difference in melting temperature [$\Delta T_m = T_m(\text{DNA} + 4 \text{ ligand equiv.}) - T_m(\text{DNA})$]. The T_m values of DNAs alone are: c-kit1 = 69.8 ± 0.5 $^{\circ}\text{C}$; c-kit2 = 62.5 ± 0.5 $^{\circ}\text{C}$; c-myc = 72.9 ± 0.5 $^{\circ}\text{C}$; Tel23-p = 68.5 ± 0.5 $^{\circ}\text{C}$; Tel23 = 65.2 ± 0.5 $^{\circ}\text{C}$; Tel26 = 59.5 ± 0.5 $^{\circ}\text{C}$; ds12 = 71.3 ± 0.5 $^{\circ}\text{C}$. All experiments were duplicated and the values reported are average of two measurements. ^b100 mM KCl buffer. ^c20 mM KCl buffer. ^dLead compound. ^eThese compounds significantly increase the thermal stability of ds12 however ΔT_m values cannot be accurately determined.

On the basis of the obtained data, compounds may be divided in three groups: (i) compounds that exhibit poor G4-stabilizing properties towards any G4 ($\Delta T_m < 5$ $^{\circ}\text{C}$) (compounds **4** and **10**); (ii) compounds showing medium G4-stabilizing effect ($\Delta T_m = 5$ – 15 $^{\circ}\text{C}$) (compound **9**); (iii) potent G4 stabilizers ($\Delta T_m > 15$ $^{\circ}\text{C}$) (compounds **2**, **3**, **5**–**8**). Most of the compounds investigated showed high selectivity for G4 over duplex DNA, analogously

to lead compound **1**. However, compounds **6**, **7**, and **8**, having a different core from compound **1**, were the only ligands which significantly increase duplex stability and therefore were not considered to be selective. Interestingly, compounds **2**, **3**, and **5** preferentially stabilized parallel G4s over antiparallel ones. In particular, compound **3** induced a marked stabilization of c-myc G4 ($\Delta T_m > 20$ °C).

CD melting studies were further performed on the most selective compounds **2**, **3**, and **5** to investigate the effect of increasing concentrations of the ligand on the stability of both c-kit2 and c-myc G4s (Figure 2.21).

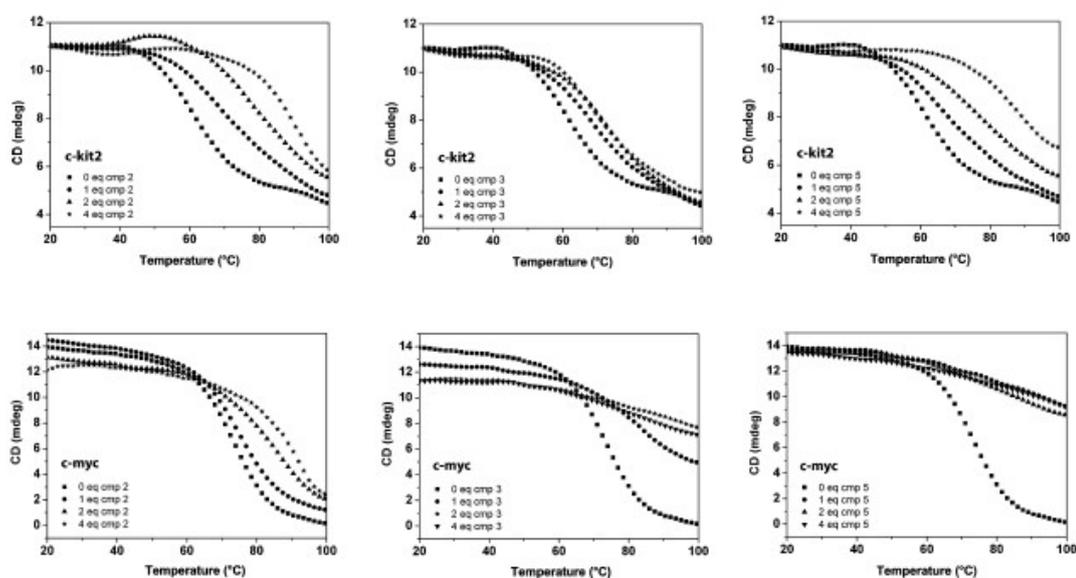


Figure 2.21. CD melting experiments of c-kit2 and c-myc G4s with increasing amounts of compounds **2**, **3**, and **5**.

Three different [ligand]/[DNA] ratio were examined (1:1, 2:1, and 4:1). In the case of c-kit2, a dose dependent increase of thermal stability was observed for all three compounds. In particular, compounds **2** and **5** increased the T_m of the G4 with every addition of ligand, while **3**, which had shown a modest increase in c-kit2 stability, did not induce any further increase of T_m beyond the 2:1 [ligand]/[DNA] ratio. In the case of c-myc G4, a considerable increase of the melting temperature was registered with compound **3** and the stabilization effect ceased at a 2:1 ratio. On the other hand, concentration dependency was observed for **2** (although slight compared to c-kit2), while the high G4-stabilizing effect of **5** on c-myc did not permit to evaluate any dose-dependent effects.

Overall, these results indicate that six out of nine synthesized ligands are potent G4 stabilizers. The diimidazo[1,2-*a*:1,2-*c*]pyrimidine derivatives having a positive charge on the side chain(s) in the used experimental conditions (i.e., **2**, **3**, and **5**) showed high affinity towards some of the investigated G4 structures. Conversely, compound **4**, whose side chain protonation is less favorable at pH 7.0, showed generally poor G4-stabilizing properties. Their duplex-binding activity is negligible. Among compounds with a different core, the positively charged ones (**6**, **7**, and **8**) showed affinity for some of the G4s investigated as well as for duplex DNA. Compounds **9** and **10**, whose side chain protonation is less likely to occur at pH 7.0, showed medium to low G4-stabilizing properties and negligible effects on duplex. These results indicate that the overall charge of compounds is necessary but not sufficient to ensure potent and selective interaction with a specific DNA structure. In addition, CD melting data suggest that two molecules of **3** could bind and stabilize the G4 target.

The G4 stabilizing properties of compounds were also evaluated by *FRET methodology*.¹²⁹To fully compare the results of CD melting experiments obtained for compounds **2**, **3**, and **5**, FRET melting experiments were performed on these ligands under the same buffer conditions used for CD, using as targets the labeled G4-forming oligonucleotides from the *c-KIT* (Fckit2T) and *c-MYC* (FcmycT) promoter regions and the telomeric G4-forming sequence annealed at high concentration to promote the parallel conformation (F21T-p). The results of FRET melting experiments, shown in Table 2.2, are in agreement with those obtained from CD melting studies.

Table 2.2. Ligand-induced thermal stabilization of labeled G4-forming sequences evaluated by FRET melting analyses

Compound	ΔT_m (°C) ^a		
	Fckit2T ^b	FcmycT ^b	F21T-p ^c
2	>20.0	>15.0	>15.0
3	≈15.0	>20.0	7.0
5	>20.0	>20.0	>20.0

^a ΔT_m represents the difference in melting temperature [$\Delta T_m = T_{m(G4 + 4 \text{ ligand equiv.})} - T_{m(G4)}$]. T_m values of G4s alone: Fckit2T = 66.3 ± 1.0 °C; FcmycT = 68.6 ± 1.0 °C; F21T-p = 67.1 ± 1.0 °C. All experiments were duplicated and the values reported are average of two measurements. ^b20 mM KCl buffer. ^c100 mM KCl buffer.

CD melting studies and FRET melting studies revealed that compound **3** is a potent and specific stabilizer of c-myc G4 and therefore its binding mode was further studied in detail by means of NMR spectroscopy.

The well-characterized modified sequence (Myc22) containing two G-to-T mutations at position 14 and 23 was used for the NMR titration experiments.

According to the literature, under the experimental conditions used, the Myc22 DNA sequence forms a single G4 conformation characterized by 12 well-resolved imino protons peaks, corresponding to the 12 guanines involved in the three G-tetrad planes.¹³⁰

Upon addition of compound **3** to the Myc22 solution, both imino and aromatic protons became broad even at higher drug equivalence, indicating a dynamic binding process of **3** to Myc22. The titration was virtually complete at a ligand/DNA ratio of 2:1 (Fig. 2.22).

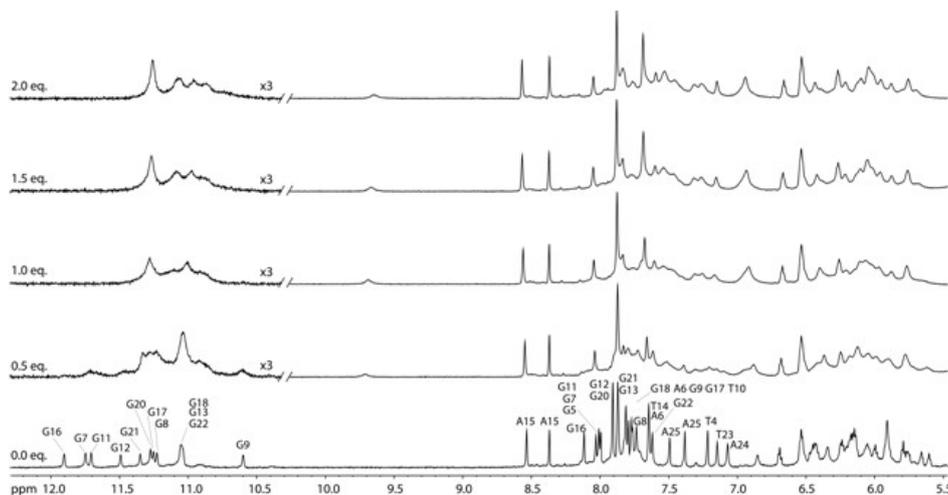


Figure 2.22. Imino and aromatic proton regions of c-MYC promoter G4 (Myc22) in 25 mM phosphate, 70 mM KCl solution (pH 7.0).

Major chemical shift changes were observed for the imino protons belonging to the 5' G-tetrad (G7–G11–G16–G20) as well as for G9 of the 3' G-tetrad (G9–G13–G18–G22). This may be explained by the structural features associated with the 3'- and 5'-faces as well as the flanking sequences. Indeed the 5'-side is more accessible for ligand stacking and the only 5'-flanking residue particularly affected by the binding turned out to be the T4, as shown by the corresponding aromatic proton shift. By contrast, the 3'-face is less accessible. A25 from this side folds back to form a base pair with T23 to cover the

external G-tetrad, while A24 stays above the T23:A25 base pair. However, these latter flanking residues are not able to completely prevent the binding. Upon addition of **3**, the aromatic protons of T23, A24, and A25 were strongly perturbed, suggesting that the ligand can be positioned between the four guanines of the 3'-tetrad and the 3'-flanking bases. The stability of the 3'-end bound ligand complex with Myc22 could be related to the possibility of specific H-bond interactions between the imidazole group of **3** and the 3'-flanking bases. The binding behavior of **3** to the c-kit2 G4 was also evaluated. The modified sequence c-kit2_{T12/T21} (in which G12 and G21 are replaced by two T residues), which is known to fold into a monomeric parallel G4 structure with well-resolved imino proton signals^{131,132} was used for this study. Interestingly, the imino signals belonging to the G-tetrad on the 5'-side (G2, G6, G14, and G18) were equally affected with respect to those belonging to the 3'-terminal G-tetrad, as well as the aromatic protons of C1 and A13 on the 5'-face and those of C9 and T21 on the 3'-face (Fig. 2.23).

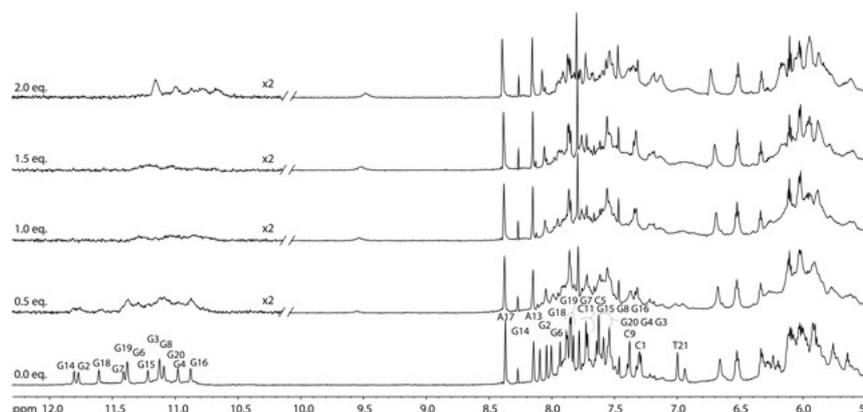


Figure 2.23. Imino and aromatic proton regions of *c-KIT* promoter G4 (c-kit2_{T12/T21}) in 5 mM phosphate, 20 mM KCl solution (pH 6.8) titrated with compound **3**.

These results clearly suggest an involvement of both external G-tetrads in binding, despite the potential pairing between residues C1 and A13 on the 5'-side.¹³¹ It may be hypothesized that **3** is able to interact with both the 5' and 3' G-tetrads, possibly by displacing the C1 and A13 base pairing at the former.

Overall, these findings suggest that compound **3** binds to the 5'- and 3'- terminal G-tetrads of both Myc22 and c-kit2_{T12/T21} G4s. However, some differences in the binding to the two G4s have been observed and they may be due to specific interactions formed between **3** and Myc22.

2.3.3. Molecular docking¹²⁰

To provide a molecular basis for the G4 binding of **3**, molecular docking calculations were also performed. The ligand was docked to the three-dimensional structures of c-kit2 and c-myc G4s determined by solution state NMR spectroscopy.^{132,114}

2:1 (ligand/DNA) complex models were built with the G4 targets, taking into account the binding stoichiometry suggested by previously described NMR studies. Both 5'- and 3'-terminal G-tetrads were considered as potential binding sites for the ligand. The tricyclic diimidazo[1,2-*a*:1,2-*c*]pyrimidine moiety is stacked onto the planes at a 3.4 Å separation at both 5'- and 3'-faces and overlaps part of two guanine bases (G11 and G16 at the 5' G-tetrad; G9 and G13 at the 3' G-tetrad), making extensive π -stacking interactions in the low energy binding pose on c-Myc G4 showed in Fig. 2.24.

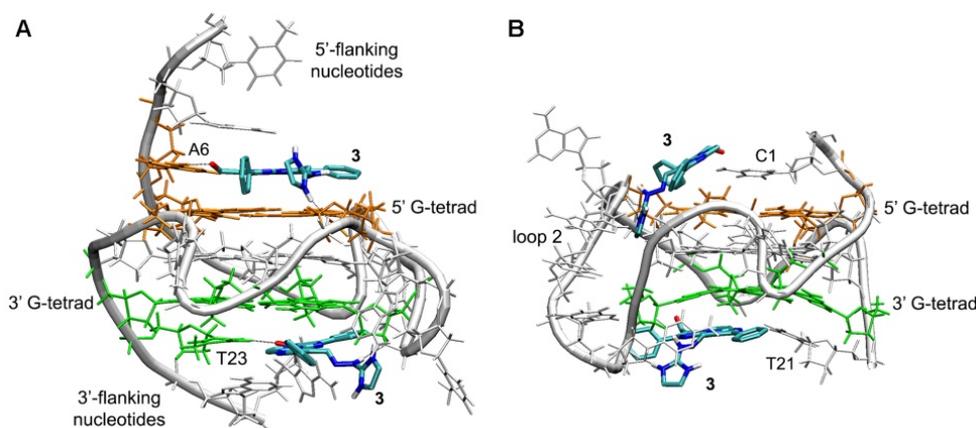


Figure 2.24. Predicted binding poses upon docking of compound **3** onto the (A) c-myc and (B) c-kit2 G4s.

In both cases, one of the two attached phenyl rings overlap part of a guanine of the tetrad, contributing to the stabilization, while the other one is twisted out of the diimidazo[1,2-*a*:1,2-*c*]pyrimidine plane. The 2-hydrazino-2-imidazoline groups of the two ligand molecules make close hydrogen bonds with the oxygen atoms of the backbone phosphate groups of the G4. It is interesting to observe that the formyl group of **3** is, in both cases, hydrogen bonded to the first flanking base (A6 and T23 at 5'- and 3'-end, respectively) that stacks over the G-tetrad, thus forming a plane capping the 5' and 3' external G-tetrads of the c-myc G4.

In the case of c-kit2 G4, different binding poses were observed for the ligand on the two external G-tetrad surfaces (Fig. 2.24). Indeed, the molecular model of **3** bound to the c-kit2 G4 shows that at the 3'-end, the planar aromatic scaffold of the ligand makes π -stacking interactions with the 3'-G-tetrad, while the 2-hydrazino-2-imidazoline group is positioned to participate in a hydrogen bond with a phosphate group. However, by contrast with the c-myc G4, compound **3** does not form any additional interaction to further stabilize the complex. On the other hand, the docking calculations at the 5'-face of the c-kit2 G4 structure reveal that to make hydrogen-bonding/electrostatic contacts between the charged 2-hydrazino-2-imidazoline group and the backbone phosphate groups of the G4, the side chain of compound **3** needs to be located in the deep groove formed by the long loop 2 of the G4. As a consequence, the diimidazo[1,2-*a*:1,2-*c*]pyrimidine core of **3** results in not being parallel to the plane of the terminal G-tetrad. This is probably due to the 5'-3' strand polarity generated by the propeller loop topology of this G4, which does not expose the phosphodiester backbone on this side, thus making only the 3'-end more favorable to interact with ligands with positively charged side chains.

Overall, the docking results reveal that **3** has the appropriate shape and electronic structure to form specific interactions. Interestingly, the binding of **3** could involve the first flanking bases to form a plane covering the external G-tetrads and contributing greatly to the G4 stabilization. This could also explain the selectivity of **3** in increasing the thermal stability of the c-myc G4.

2.3.4. Biological Studies¹²⁰

The most selective G4-interactive agent **3** was studied also at cellular level in comparison with compound **1**. Compound **14a** was used as a negative control because it is unable to bind G4s in vitro.⁴¹

First, a *MTT assay* was performed in order to evaluate the cell killing activity of the three selected molecules in human osteosarcoma U2OS and colon cancer HCT116 cell lines.

HCT116 cells were more sensitive than U2OS cells to the tested compounds, however, both cell lines were more sensitive to compound **3** than other agents (Table 2.3).

Table 2. 3. Cytotoxic activity of selected hydrazone derivatives in human U2OS and HCT116 cells.

Compound	IC ₅₀ (μM) 1 hour		IC ₅₀ (μM) 24 hours	
	U2OS	HCT116	U2OS	HCT116
1	101.0 ± 1.63	35.07 ± 1.06	15.97 ± 1.41	5.12 ± 1.05
3	5.73 ± 1.48	4.18 ± 1.07	0.845 ± 0.001	0.407 ± 0.106
14a	45.23 ± 2.05	16.79 ± 1.06	9.16 ± 1.44	4.06 ± 1.07

The concentrations killing 50% of cells (IC₅₀) are means ± SE of two independent experiments performed in triplicates.

In particular, compound **3** showed high cell killing activity in U2OS cells after 24 h of treatment (IC₅₀ = 0.845 μM). In addition, **3** was 20- and 12-fold more cytotoxic than compounds **1** (IC₅₀ = 15.97 μM) and **14a** (IC₅₀ = 9.16 μM), respectively. Similarly, **3** was more cytotoxic than compounds **1** and **14a** after 1 h of cell treatment (Table 2.3). The sensitivity pattern was very similar in HCT116 cells, with compound **3** showing highest cell killing activity as compared with compounds **1** and **14a** both after 24 and 1 h of treatment (Table 2.3). These results showed that compound **3** is more cytotoxic than the other tested compounds, in the two used cell lines.

Moreover, *immunofluorescence microscopy* was used to visualize nuclear G4 structures in human U2OS cells and to evaluate the G4 stabilizing properties of **3** and **1** in living cells. To this aim the BG4 antibody, a known specific antibody against G4s, was employed.¹³³ U2OS cells were treated for 24 h with compound concentrations very close to the calculated IC₅₀ values.

BG4 showed nuclear staining in untreated cells indicating that G4 structures were present at discrete chromatin sites in the nucleus (Fig. 2.25). Fluorescence patterns were similar in cells treated with the three derivatives. However, the compound effects were very different because compounds **3** and **1**, but not **14a**, markedly increased the number, size, and fluorescence intensity of nuclear spots, thus indicating that they could stabilize G4 structures in the nucleus of living cells.

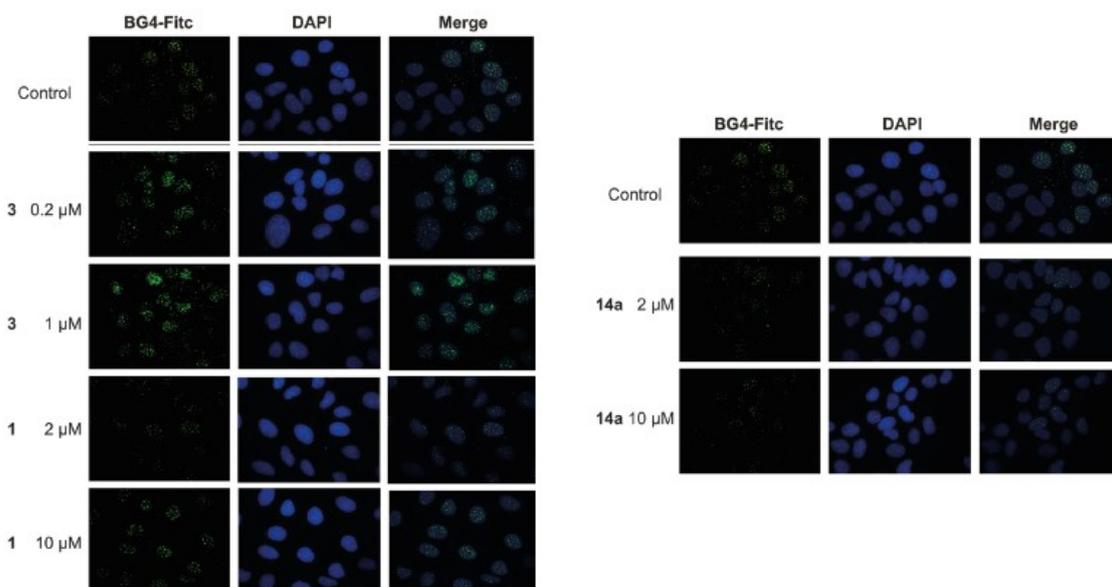


Figure 2.25. Immunofluorescence showing BG4 foci (green) in human osteosarcoma U2OS cell nuclei (blue). Cells were treated with G4 ligands for 24 h and then fixed. G4s and DNA were stained with BG4 antibody and DAPI, respectively.

Measurements of the total fluorescence signal per cell for each sample indicated that **3** was the most potent, indeed an increase of G4 signal was detected even at 0.2 μM (a dose 4-fold lower than IC_{50}), and further increased at 1 μM, of **3**. Compound **1** was able to increase the nuclear G4 signal at 10 μM at comparable levels as **3** at 0.2 μM. In contrast, **14a** did not show any increase in fluorescence intensity at the tested concentrations (Fig. 2.26).

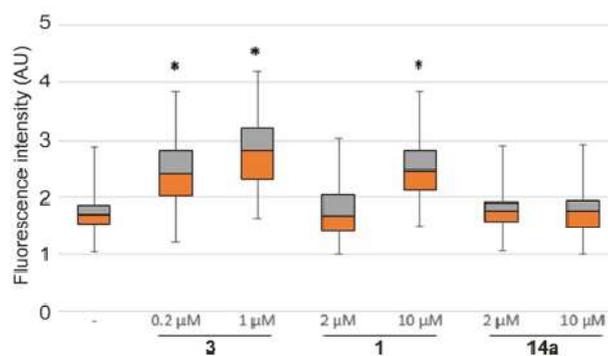


Figure 2.26. Box Plot analysis of G4 signal in the nucleus of control and treated cells. At least two experiments were considered, and 150 nuclei were counted per sample. Statistical significance was determined with the t test, and * indicates a p value less than 10^{-32} .

These data demonstrate that hydrazone derivatives can stabilize G4s in human cancer cells and that **3** is the most potent agent among the studied derivatives. Further investigations at the cellular level are needed to confirm the correlation between G4 stabilization and biological activity of these hydrazone derivatives.

2.4. Conclusions

A series of hydrazone derivatives (**2-10**) of the lead compound **1** was designed and synthesized. CD assays showed that six out of nine compounds are potent G4 stabilizers that retain the pre-existing architecture of the target G4s upon interaction. Compounds having a different core (**6, 7, 8**) from that of compound **1** are not selective for G4s. Compounds **2, 3** and **5** (Fig. 2.27) displayed a marked preference for binding to parallel G4s over duplex and antiparallel G4s.

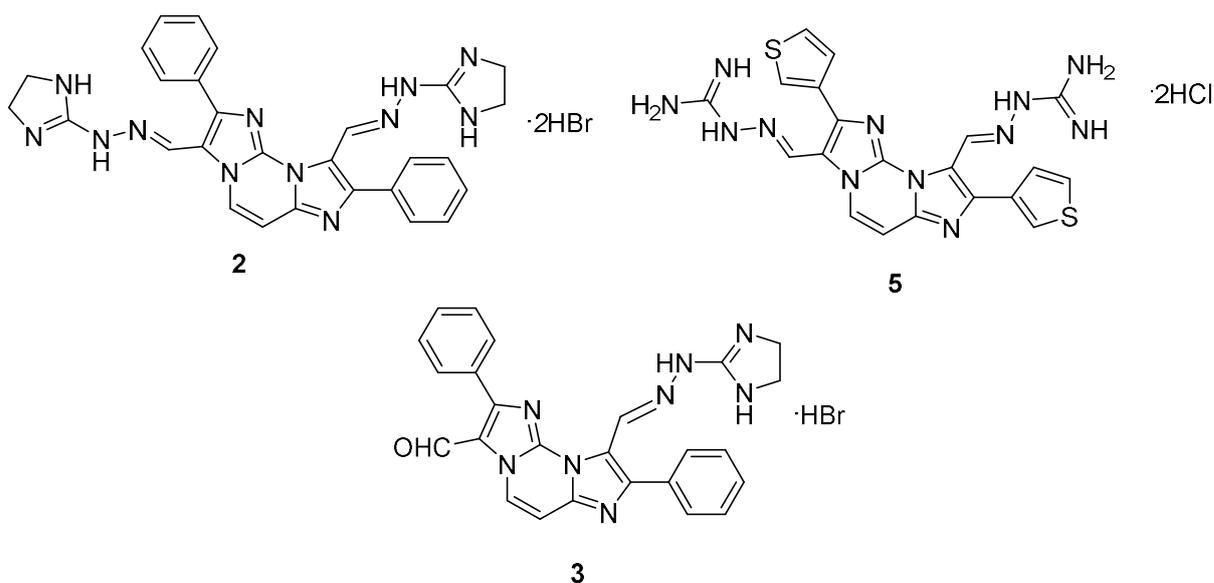


Figure 2.27

The monohydrazone **3** was found to be the most selective of the series. The reduction of positively charged side chains number resulted in enhanced selectivity of this derivative towards c-Myc G4. NMR experiments suggested an end-stacking binding mode at the terminal G-tetrads of the G4 and a 2:1 drug/DNA binding stoichiometry.

Molecular docking results revealed that **3** has the appropriate structure to stack on the planar G-tetrad surfaces of c-myc G4 with the tricyclic diimidazo[1,2-*a*:1,2-*c*]pyrimidine core and, the same time, to make (i) close hydrogen-bonding/electrostatic contact between the charged 2-hydrazino-2-imidazoline group and the oxygen atoms of the backbone phosphate of G4, and (ii) a hydrogen bond interaction between the formyl group at position **3** and the hydrogen-bond donor groups of the first flanking base that stacks on the G-tetrad.

In accordance with the *in vitro* assays, biological experiments demonstrated that hydrazone derivatives can effectively trap G4 structures in the nuclei of cells. Derivative **3** is more effective than the lead compound **1** in inhibiting human U2OS and HCT116 cancer cell growth, as well as in stabilizing G4s in osteosarcoma cells.

The reported results indicate that the diimidazo[1,2-*a*:1,2-*c*]pyrimidine molecular scaffold should be actually considered for the development of highly-selective drug-like G4 ligands. Further structural modifications may be performed to explore selectivity for a single G-quadruplex structure and also to improve biological activity.

CHAPTER 3

3. NGR tumor-homing peptides: structural considerations for effective APN targeting

3.1. Background

3.1.1. Exploring tumor vessels “zip codes”: identification of NGR tumor homing peptides

The angiogenesis process, consisting in the formation of new blood vessels from the existing vasculature, physiologically occurs during embryogenesis and in the adult during female reproductive cycle and wound healing.^{18,134} Folkman in 1971 was the pioneer of the hypothesis that tumor growth and metastasis strictly depend on angiogenesis,¹³⁵ therefore marking the beginning of an intensive research in this field. Nowadays angiogenesis is considered as a hallmark of cancer.⁸ Tumors, defined as “wounds that never heal”,¹³⁶ have lost the balances between positive and negative regulators of angiogenesis and, as a consequence, the ‘angiogenic switch’ is turned on and tumor vessels constantly grow.¹³⁷

The formed tumor vasculature appears quite different from the normal counterpart: it presents fragility, chaotic architecture, high permeability.¹³⁸ Moreover tumor vessels express specific “zip codes”, surface proteins that are not expressed or are expressed at much lower levels in normal vessels.^{36,139}

Indeed elevated expression levels of receptors for VEGF,¹⁴⁰ integrins such as the $\alpha\beta_3$, $\alpha\nu\beta_5$ and $\alpha 5\beta_1$,¹⁴¹ peptidases/proteases such as aminopeptidase N (APN/CD13),¹⁴² have been found on the surface of endothelial cells within tumor vessels. These proteins are considered markers of angiogenesis and are also functionally important in the angiogenic process.^{143–145}

Exploiting the differences in vascular “ZIP codes” between normal and tumour endothelial cells may be the right approach to develop novel anticancer strategies, such as neovasculature-targeted drug delivery systems.¹⁴⁶

In vivo phage display is considered the most effective technology to probe tumor vasculature. It is a high-throughput technique that may be used in both vascular mapping and in the identification of ligands that specifically target a particular vascular bed.¹⁴⁷ A screening of random peptide libraries displayed on phage injected into tumor-bearing mice led to the identification of tumor-homing peptides (THPs), able to target selectively tumor vasculature of a mouse model.

Among various peptides identified by Pasqualini *et al.*, the peptide motif containing the sequence Arg-Gly-Asp (RGD) (Fig. 3.1), is able to recognize the $\alpha\beta_3$ and $\alpha\beta_5$ integrins, which are up-regulated in tumor endothelial cells.³⁷

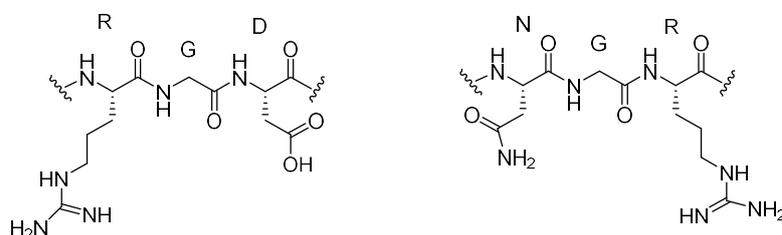


Figure 3.1

In a following publication, the same researchers described that a phage carrying the peptide CNGRCVSGCAGRC, containing the sequence Asn-Gly-Arg (NGR) (Fig. 3.1), which had been identified as a cell adhesion motif,^{148,149} accumulated into tumor angiogenic vessels within human breast carcinoma xenografts. Furthermore the NGR peptide showed higher selectivity toward tumor cells than the RGD peptide.³⁸ A subsequent study revealed that the receptor for NGR peptide is aminopeptidase N, which is overexpressed in angiogenic blood vessels.¹⁴²

3.1.2. Aminopeptidase N (APN/CD13) as receptor for NGR tumor-homing peptides

APN is a cell-surface membrane spanning protein belonging to the M1 family of metalloenzymes, consisting of a large number of zinc-dependent aminopeptidases.¹⁵⁰ It is expressed in some epithelial cells and immune cells.¹⁵¹

With respect to other aminopeptidases, APN seems to have multifunctional roles and it has indeed defined as a “moonlighting ectoenzyme”.¹⁵² APN removes preferentially *N*-terminal neutral amino acids, such as alanine, from oligopeptides. By cleaving the *N*-terminus aminoacid, APN regulates the metabolism of several peptides (e.g.: angiotensin III) involved in important biological processes.¹⁵³ APN acts also as a receptor for some viruses, such as the human coronavirus 229E (HCoV-229E). The viral receptor function results in APN internalization.¹⁵⁴ In addition, it has been proposed that APN exerts some functions as signaling molecule.¹⁵⁵

Several findings revealed that APN is associated with the malignant development and it is involved, in particular, in tumor aggressiveness and metastasis. It has been reported that APN is overexpressed or its enzymatic activity is altered in skin, ovary, thyroid, lung, stomach, colon, kidney, bone and prostate tumors.¹⁵⁶ APN has been linked to increased migratory capacity of tumor cells, chemokine processing and tumor invasion.^{157,158} Anti-APN antibodies block tumor migration, probably because APN has a role in extracellular matrix proteins degradation.¹⁵⁹ However tumor cells that overexpress enzymatically inactive APN also show increase migration and metastasis, therefore APN may influence cell motility in a manner that is not dependent on its enzymatic activity.¹⁶⁰ Moreover APN interacts with cell-surface proteins, playing a role also in cell-cell adhesion.^{155,161}

Pasqualini and co-workers identified APN as the receptor for NGR containing tumor-homing peptides. They showed that phage displaying NGR-peptides were able to bind specifically immunocaptured APN and APN-transfected cells and this binding was inhibited by the cognate soluble peptide; moreover anti-APN antibody inhibited *in vivo* homing of NGR phage into tumors. Furthermore they showed that the only vascular structures with detectable APN are tumor blood vessels and other types of vessels undergoing angiogenesis. So, it was established that APN is a marker of angiogenesis and

it plays also a functional role in tumor angiogenesis. Bestatin, a known APN inhibitor, inhibited angiogenesis *in vivo* and also suppressed tumor growth, showing good antitumor activities. In addition antibodies against APN inhibited tumor angiogenesis.¹⁴²

In a consequent study it has been proposed that the isoform of aminopeptidase N expressed in tumor vessels may be different from that in normal cells, thus explaining the selectivity of NGR tumor-homing peptides.¹⁶²

According to a recent paper reported in literature by Liu et al., APN interacts with the NGR tumor-homing peptide, with a similar recognition mechanism used for interacting with the NGR motifs within ECM proteins (Fig. 3.2). This can explain both the enzymatic-independent tumor cell-motility mediated by APN and the tumor-homing peptide binding to APN.¹⁶³

These findings have prompted an intensive research in both therapy and diagnosis of cancer, in which NGR/APN recognition may be used to develop conjugates between NGR-peptide and an anticancer drug and/or an imaging agent.

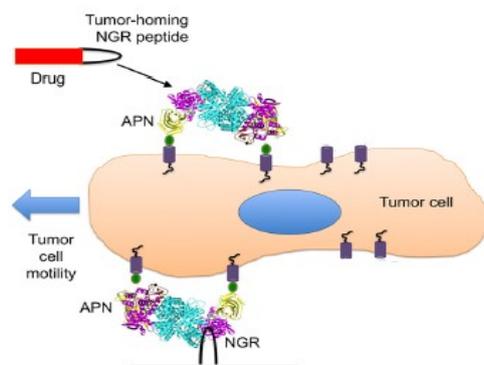


Figure. 3.2. Black loops indicate the NGR motif present in tumor-homing peptide connected to an anti-tumor drug and in ECM proteins.¹⁶³

3.1.3. NGR peptide-based agents

In the last decade tumor-homing peptides have emerged as promising carriers for targeting anticancer and imaging agents into tumors. Peptide-based delivery of compounds has numerous advantages over other delivery systems. Peptides are indeed smaller, penetrate more efficiently into tissues and are less immunogenic than antibodies.

Particularly, much attention was focused on tumor-homing peptides containing the NGR motif. Due to their ability to bind in a specific manner tumor cells and tumor vasculature by interacting with aminopeptidase N, NGR-containing peptides emerged as promising tools to deliver therapeutics and imaging probes into tumors.

As reported by Arap. *et al.*, a pentapeptide containing the NGR sequence flanked by two cysteines forming a disulfide bridge (cCNGRC, Fig. 3.3) is sufficient to home into tumors. Besides, the cyclic CNGRC peptide homed into tumors more effectively than linear peptides. Coupling this peptide with doxorubicin resulted in increased drug efficacy and lowered toxicity in mice.³⁸

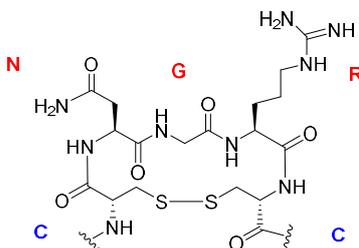


Figure 3.3. Cyclic CNGRC peptide used as tumor homing carrier.

Considering these advantages, many other studies on conjugates between the cCNGRC peptide and small molecules, such as doxorubicin and platinum derivatives (cisplatin and carboplatin), have been reported. The cCNGRC peptide has been derivatized either at the C-terminus or N-terminus, showing a variable grade of success (Fig. 3.4).¹⁶⁴⁻¹⁶⁶

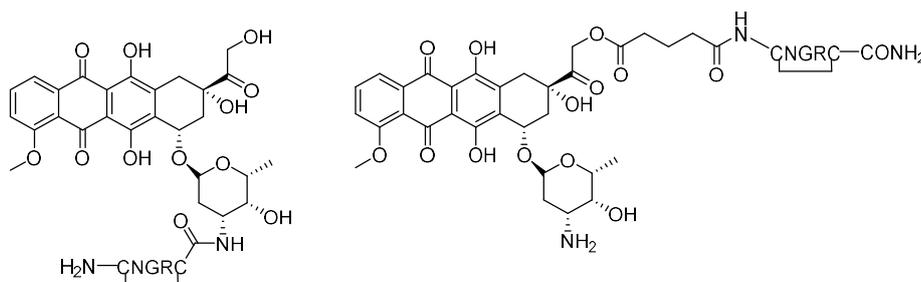


Figure 3.4. Examples of conjugates between cCNGRC peptide and doxorubicin.

NGR peptides found another application in the development of doxorubicin (DOX)-encapsulated liposomes coupled with the NGR motif. In this case improved drug uptake by neuroblastoma tumours and enhanced therapeutic efficacy were observed.¹⁶⁷ This strategy could be effective in tumors characterized by poor vascularisation.

Also the pro-apoptotic KLAKLAKKLAKLAK peptide was conjugated to the C-terminus of the CNGRC peptide and the conjugate selectively killed angiogenic cells.¹⁶⁸

Interestingly, when the c(CNGRC)G peptide was conjugated to the N-terminal domain of (h)TNF- α , increased anti-tumor activity and lowered toxicity of this cytokine were observed in animal models.¹⁶⁹ The cyclic CNGRC coupled to the TNF- α exhibited more than 10-fold higher antitumor activities than the corresponding linear derivative.¹⁷⁰ The c(CNGRC)G-TNF- α construct is able to control tumor growth either alone or in combination with chemotherapeutics.^{171,172} It's noteworthy that c(CNGRC)G-TNF- α based cancer therapy is currently under phase III clinical trials in patients with malignant pleural mesothelioma.¹⁷³

Furthermore NGR tumor homing peptide represents a powerful tool for tumor imaging and to date interesting examples of NGR-based imaging probes have been reported in literature.

The first example of NGR conjugated probes synthesis was reported by Dirksen *et al.*, who extended the C-terminus of cCNGRC with a diethylenetriaminepentaacetic acid (DTPA) ligand complexed to Gadolinium (III) for MRI and with a fluorescent dye (OG 488) for optical imaging, while the N-terminus was maintained in the acetylated form (Fig. 3.5). However no imaging studies were reported for this bimodal agent.¹⁷⁴

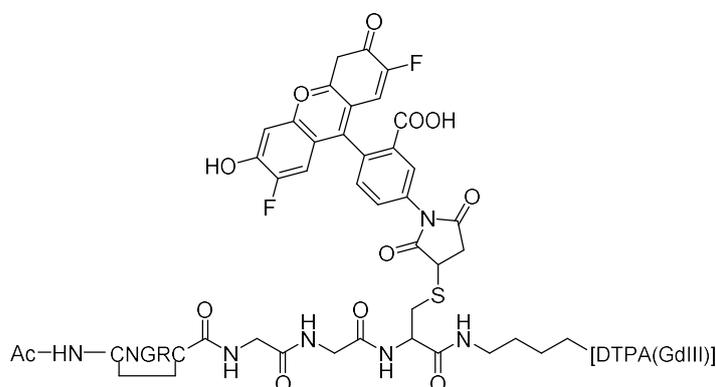


Figure 3.5. The bimodal target-specific contrast agent containing cCNGRC conjugated with gadolinium(III)diethylenetriaminepentaacetic acid (Gd(III)DTPA), and Oregon Green 488 (OG488).

NGR peptide-based imaging was adapted at the quantum dot (QD)-based imaging approach widely adopted in living cells and animal models. When biotin-tagged cNGR was attached to streptavidin-coated QDs, the resulting cNGR-QD was highly specific to APN/CD13-positive cells.¹⁷⁵

Fluorescent dye Cy 5.5 has been conjugated to NGR peptide and the resulting cNGR-Cy 5.5 (Fig. 3.6) differentiates APN/CD13-positive fibrosarcoma cells from APN/CD13-negative MCF-7 cells. cNGR-Cy 5.5 was visualized by fluorescence reflectance imaging not only in tumor region but also in kidney and liver, suggesting the rapid renal excretion of the conjugate *in vivo*.¹⁷⁶

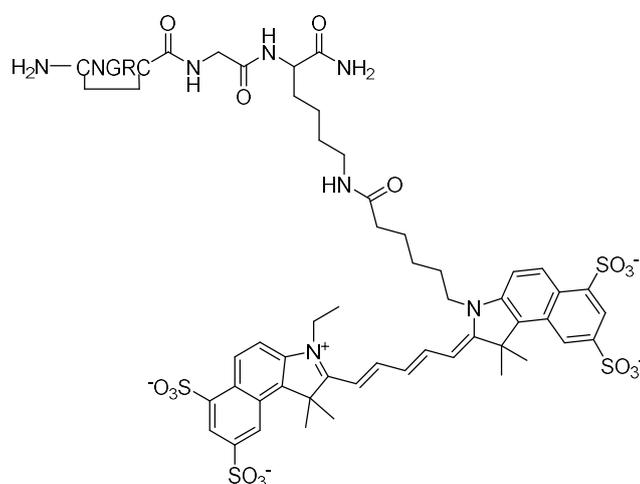


Figure 3.6. Chemical structure of Cy 5.5-labeled cNGRC-peptide.

3.1.4. Interactions between APN and NGR peptides

Despite the potential of APN-directed tumor-homing strategies, the exact structural requirements necessary for APN-NGR recognition and tumor targeting with NGR peptides are still not completely clear.

Chen and co-workers reported the crystal structure of porcine APN (pAPN), which shares a high grade of omology with the human one.¹⁷⁷ APN is a zinc-dependent ectoenzyme consisting in a short *N*-terminal cytoplasmic domain, a single pass transmembrane anchor, a small extracellular stalk and a large *C*-terminal ectodomain containing the active site. The ectodomain of pAPN has a seahorse shape, with 4 domains: head, side, body and tail. The zinc-binding active site is located in the body. A negatively charged channel is located between the head and the body leading to the zinc-binding active site. The substrate poly-alanine (Fig. 3.7) has its *N*-terminal residue anchored to the active site and the rest of the peptide bound to the channel. In the APN active site Gln 208, Glu 350, Glu 406 form three H bonds with the free *N*-terminal -NH₂ group of peptides and other ligands, such as the substrate analogue bestatin (Fig. 3.7). These interactions firmly anchor the *N*-terminus of peptides to the active site for catalysis, thus making APN a *N*-terminal exo-aminopeptidase. Moreover a hydrophobic interaction is formed between a Met349 and the C_β of the *N*-terminal amino acid of peptides and this can explain why APN prefers alanine at the *N*-terminus of peptide substrates. APN binds the rest of the peptides through hydrogen bonds regardless of peptide sequences.¹⁷⁷

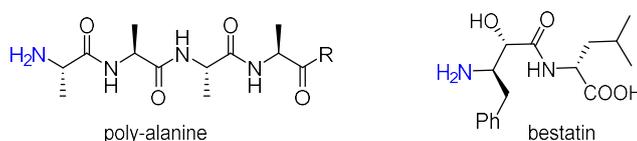


Fig. 3.7. Structures of APN substrate poly-alanine and APN inhibitor bestatin.

In a subsequent publication the same research group reported the crystal structure of porcine APN in complex with the c(CNGRC)G peptide, used as tumor-homing carrier, revealing important structural information. X-ray crystallographic data showed c(CNGRC)G tumor homing peptide binds to the zinc aminopeptidase active site. The two cysteines within c(CNGRC)G form a disulfide bond, thus the NGR region forms a short

loop with a sharp turn. The NGR loop makes sequence-specific interactions with the pAPN active site through the side chains of Asn and Arg. The presumed scissile peptide bond within the *c*(CNGRC)G backbone is stable against APN hydrolysis (Fig. 3.8).¹⁶³

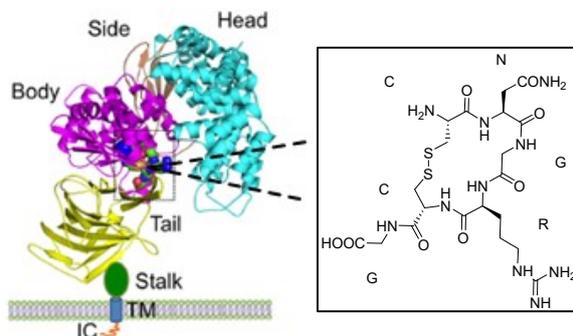


Figure 3.8. Overall structure of the pAPN- *c*(CNGRC)G complex.¹⁶³

As mentioned above, the *N*-terminal amino group of APN natural substrates, such as poly-alanine and inhibitors, such as bestatin, plays a crucial role in the binding process, forming strong interactions with a negatively charged region of the APN catalytic site. X-ray data showed that the H₂N-Cys-Asn fragment of the *c*CNGRC peptide overlaps with the *N*-terminal H₂N-Ala-Ala in APN-bound polyalanine (Fig. 3.9).¹⁶³ These data suggested that the carboxyl terminus on the second cysteine of the tumor-homing peptide plays a less important role in the binding to APN. On the contrary the amino function of *c*CNGRCG may be engaged in the binding process.

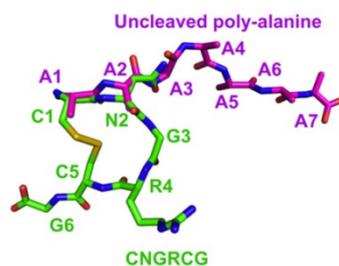


Figure 3.9. Comparison of the conformations of *c*CNGRCG in the crystal of the wild type pAPN-*c*CNGRCG complex and the uncleaved poly-alanine peptide in the crystal of mutant pAPN-polyalanine complex (PDB code 4NAQ).¹⁶³

3.2. Aim of the project

Based on the structural data reported in the above section, the amino-terminus of the cCNGRC peptide may have an important role in APN binding, therefore the carboxyl-terminus should be used as functionalization site for the synthesis of cCNGRC-bioconjugates. Conversely the *N*-terminus should remain in the free NH₂-form.

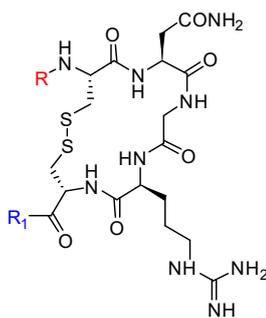
Despite a wide range of linear or cyclic NGR peptides has been reported in literature, this aspect has never been evaluated. Both then *N*-terminus^{165,166,178,179} and *C*-terminus^{38,169} of cCNGRC have been used as anchorage points for bioconjugation, with or without protection of the unused terminus.¹⁷⁹

The objectives of this research were i) the investigation of the role of the free *N*-terminal amine function on the binding properties of cCNGRC peptide towards APN and ii) the identification of the optimal site for cCNGRC derivatization, in order to provide a rationale for the design of APN-targeted cCNGRC constructs.

To this aim, the basic structure of the cyclic peptide with the free *N* and *C* termini (**1**) and a library of cCNGRC peptides bearing differently functionalised *N*-terminus and *C*-terminus was designed and synthesized. The *C*-terminus was modified by introducing an amide group (**2**), the *N*-terminus was masked by acetylation (**3**) the biotin tag was introduced on the *N*-terminus (**4**) and on the *C*-terminus in the derivative **5**, which was further modified by *N*-terminus acetylation to obtain derivative **6** (Fig 3.10).

To synthesize the cyclic peptides, solid phase peptide synthesis-based strategies were explored. Fmoc-6-aminohexanoic acid and Fmoc-Lys(Dde)OH were used as linkers for *N*-terminus and *C*-terminus functionalization, respectively.

The effects of the structural modifications on the ability of derivatives **1–6** to inhibit the enzymatic activity of APN were investigated. Finally a cCNGRC derivative (**7**) bearing a fluorescent probe (Oregon green 488, OG) on the right functionalization site was designed and synthesized for imaging studies in human cancer cells.



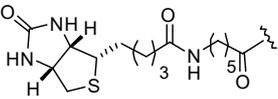
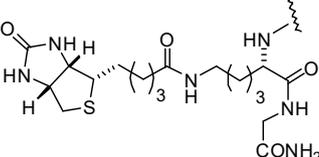
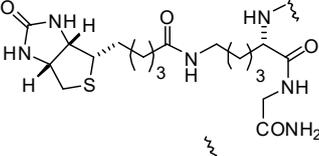
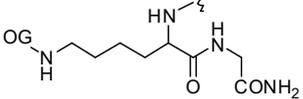
Cmpd.	R (N-terminus)	R ₁ (C-terminus)
1	H	OH
2	H	NH ₂
3	Ac	OH
4		OH
5	H	
6	Ac	
7	H	 OG = Oregon Green 488

Figure 3.10. cNGRC derivatives bearing different moieties on *N*-terminus and *C*-terminus.

3.3. Results and discussion

3.3.1. Chemistry

*General procedure for the synthesis of peptides 1–7*¹⁸⁰

The peptide sequence was assembled in a linear fashion from the C-terminus to the N-terminus via manual SPPS based on Fmoc strategy,¹⁸¹ using a Wang resin to assemble peptides displaying a free carboxylic acid (**1**, **3**, **4**), while a Rink resin was used to synthesize peptide amides (**2**, **5**, **6**, **7**) (Fig. 3.11).

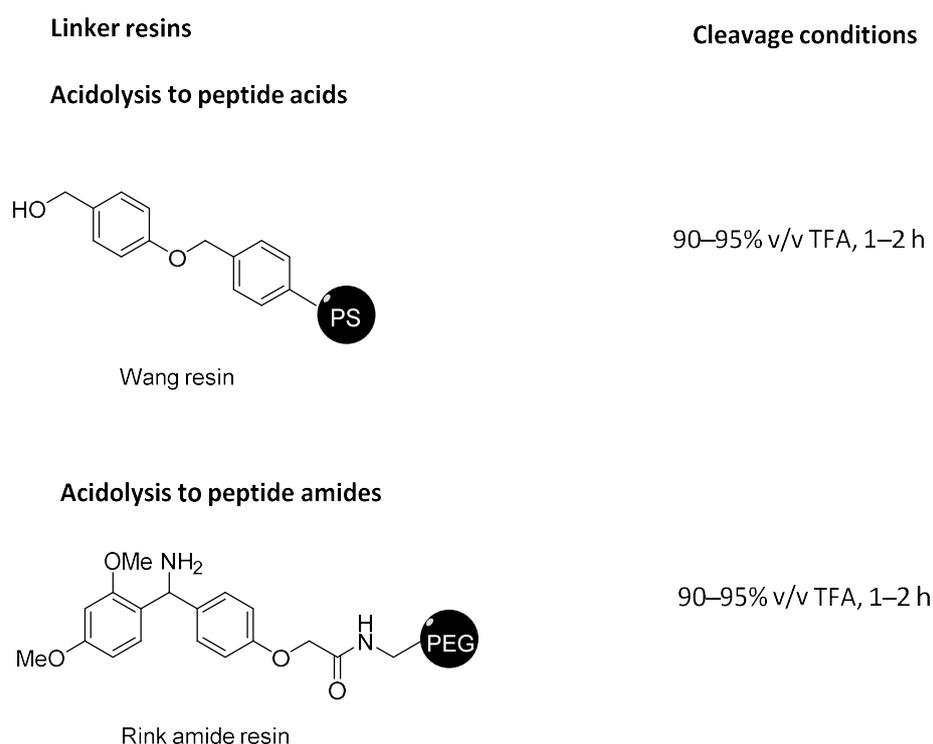


Figure 3.11. Used resins for the synthesis of peptides 1–7.

The Fmoc-protected aminoacids were coupled in the sequence cysteine (Cys), arginine (Arg), glycine (Gly), asparagine (Asn) and cysteine (Cys) for the synthesis of peptides **1**, **2**, **3**. Pbf and Trt side chain protecting groups were selected for their acid lability for Arg and Asn respectively. Fmoc-Ahx-OH was coupled with the last Cys in order to obtain the precursor linear peptide of **4**. On the other hand, the CRGNC sequence was assembled after loading the resin with Fmoc-Gly-OH and coupling with Fmoc-Lys(Dde)OH to synthesize peptides **5**, **6** and **7**.

For the synthesis of C-terminal peptide amides, the first aminoacid was attached using the standard peptide coupling procedure, after Fmoc deprotection of the amino function

of the MBHA Rink resin. The NovaPEG Rink resin was supplied with the free amino group as anchorage point.

On the other hand, the synthesis of peptide acids, was performed on the Fmoc-Cys acetamidomethyl (Acm) preloaded Wang resin to avoid the racemization reaction that is likely to occur during the attachment of cysteine with the resin linker bearing a hydroxyl group.¹⁸² The coupling reactions were carried out by *in situ* activation of the carboxylic function in presence of the uronium/iminium based activating agent HATU and diisopropylethylamine. A large excess of the activated aminoacid (5 equivalents with respect to the resin loading) was dissolved in DMF to ensure effective diffusion of reactants. All the coupling reactions were complete in 2 h as confirmed by a Kaiser test.¹⁸³ After each coupling reaction step, Fmoc protecting group removal from the N-terminus of the growing peptide chain, was achieved by adding a solution of 20% piperidine in DMF. This procedure was repeated twice to ensure effective Fmoc deprotection, which was monitored by thin layer chromatography, since the formation of the dibenzo-fulvene-piperidine adduct strongly adsorbs in the UV range and confirmed by a Kaiser test.

After assembling the aminoacid sequence, simultaneous cleavage of the peptide from the resin and removal of side-chain protecting groups were achieved using a mixture of trifluoroacetic acid (TFA) in presence of alkylsilanes and water as scavengers (TFA/TIPS/H₂O 95:2.5:2.5 v/v/v) to trap carbocations produced by side-chain protecting groups. The disulfide bridge formation was performed on resin only for the synthesis of compound **4**, while it was achieved in solution after the final cleavage of the linear peptides for the synthesis of the other compounds.

All the peptides were purified on a semi-preparative C-18 column on a reverse-phase HPLC and then lyophilized and their >98 % purity was confirmed by reverse-phase LC-MS analysis. The structures were confirmed by NMR spectroscopy.

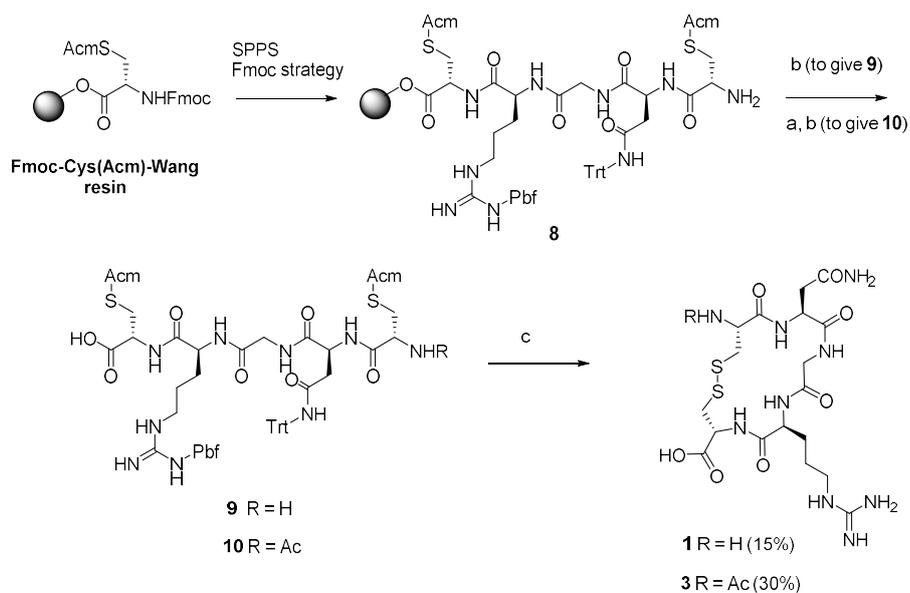
*Synthetic strategy for the synthesis of compounds **1** and **3***

The disulphide bridge formation represented a particular challenge for the chemical synthesis of the cyclic CNGRC peptide derivatives.

One single step conversion of *S*-protected peptides into the corresponding disulphides, is one of the most attractive method for S-S bridge formation.

To this end the acetamidomethyl (Acm) was selected as the -SH protecting group for its potential application. Treatment of peptides containing this protecting group with either thallium trifluoroacetate in TFA and with iodine in AcOH-H₂O mixtures results in the simultaneously deprotection of the thiol functions and ring closure. The latter strategy has been extensively applied in peptide synthesis although side reactions occurring upon iodine treatment. To overcome this problem, the excess of iodine should be quenched. The synthetic strategy adopted for compounds **1** and **3** was the cyclization in solution phase after the cleavage from the Wang resin of the linear free *N*-terminus pentapeptide **9** and acetylated *N*-terminus **10** respectively (Scheme 3.1). For the synthesis of the acetylated *N*-terminus cCNGRC (**3**) the acetylation step was performed on resin before the cleavage of the linear pentapeptide by using a mixture of DCM/Acetic Anhydride/DIPEA.

Iodine was selected as the oxidizing agent and the reaction was performed in a large volume of solvent. High dilution conditions were necessary to promote intramolecular disulphide bond formation and to avoid intermolecular dimerization. The crude pentapeptide was added dropwise to a solution of iodine in AcOH in presence of HCl. The formation of the oxidized product was monitored by RP-HPLC coupled to tandem ESI-MS and stopped after 2 h. Diethylethere was used to simultaneously quench and remove the excess of iodine and to achieve the precipitation of the cyclic pentapeptide, as described in literature.¹⁸⁴



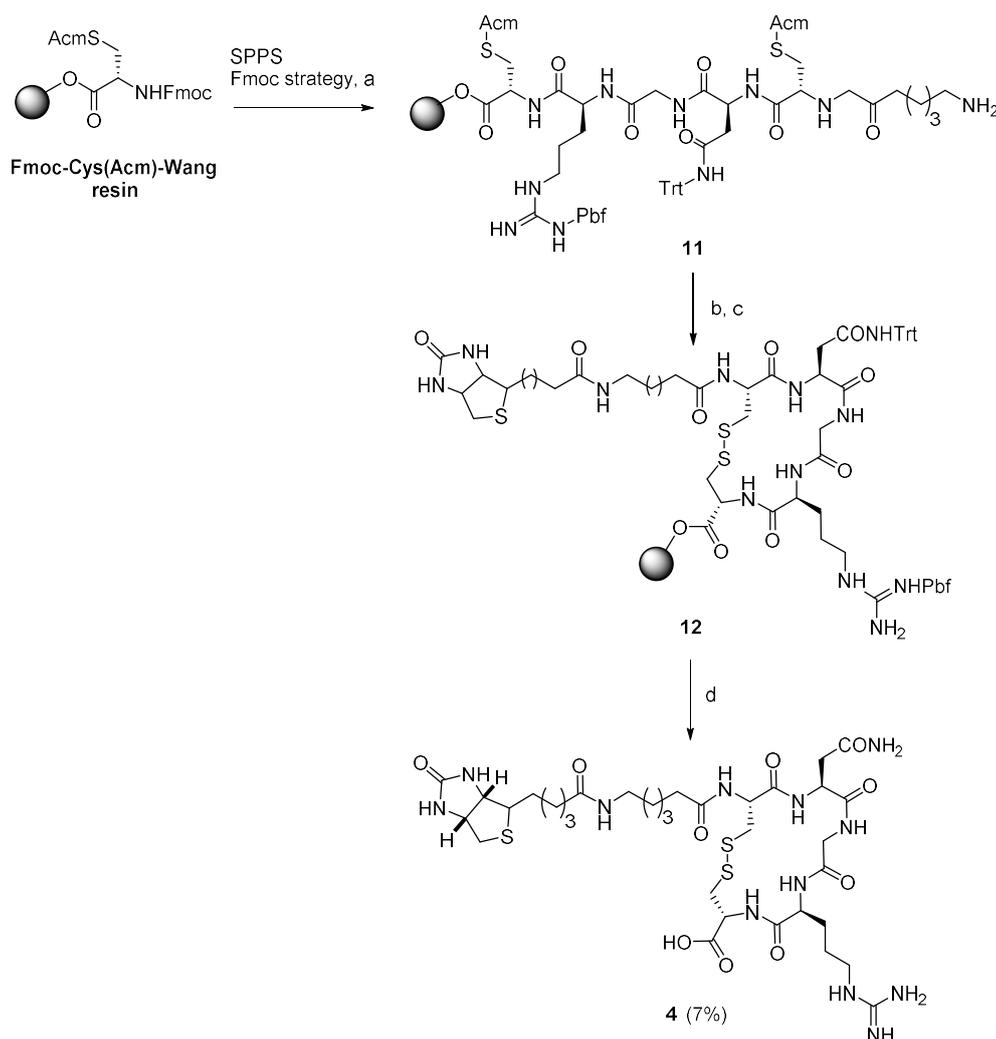
Scheme 3.1 Reagents and conditions: SPPS Fmoc strategy: Fmoc cleavage) 20 % piperidine in DMF, rt, 10 min (x2); coupling) Fmoc-amino acid (5 equiv), HATU (4.9 equiv), DIPEA (10 equiv), DMF, rt, 2 h. Sequence of addition: Fmoc-Arg(Pbf)-OH, Fmoc-Gly-OH, Fmoc-Asn(Trt)-OH, Fmoc-Cys(Acm)-OH. (a) DCM/Acetic Anhydride/DIPEA 20:4:1, rt, 1h (x2); (b) TFA/TIPS/H₂O 95:2.5:2.5, rt, 2 h; (c) I₂ (60 equiv) AcOH, HCl, rt, 2 h.

The cyclic peptides **1** and **3** were obtained in acceptable yields, although the high volume of solvent represented one of the disadvantages of the above described method.

Synthetic strategy for the synthesis of compounds **4**

Derivative **4**, displaying the free terminal carboxylic function and bearing a biotin tag on the *N*-terminus, was assembled on a preloaded Fmoc-Cys(Acm)Wang resin, as with compounds **1** and **3** (Schemes 3.1 and 3.2).

Fmoc-6-aminohexanoic (ϵ -Ahx) acid was coupled with the last cysteine bearing an Acm protecting group and used as a spacer between the CNGRC peptide and the biotin (Scheme 3.2).



Scheme 3.2 Reagents and conditions: SPPS Fmoc strategy: Fmoc cleavage) 20 % piperidine in DMF, rt, 10 min ($\times 2$); coupling) Fmoc-amino acid (5 equiv), HATU (4.9 equiv), DIPEA (10 equiv), DMF, rt, 2 h. Sequence of addition: Fmoc-Arg(Pbf)-OH, Fmoc-Gly-OH, Fmoc-Asn(Trt)-OH, Fmoc-Cys(Acm)-OH; a) (i) Fmoc- ϵ -Ahx-OH (5 equiv), DIC (5 equiv), Oxyma Pure (5 equiv), DMF, rt, 2 h; (ii) 20 % piperidine in DMF, rt, 10 min ($\times 2$); b) I_2 (10 equiv), DMF/H₂O 4:1, rt, 4 h; c) biotin (10 equiv), DIC (10 equiv), Oxyma Pure (10 equiv), DMF, rt, 24 h; (d) TFA/TIPS/H₂O 95:2.5:2.5, rt, 2 h.

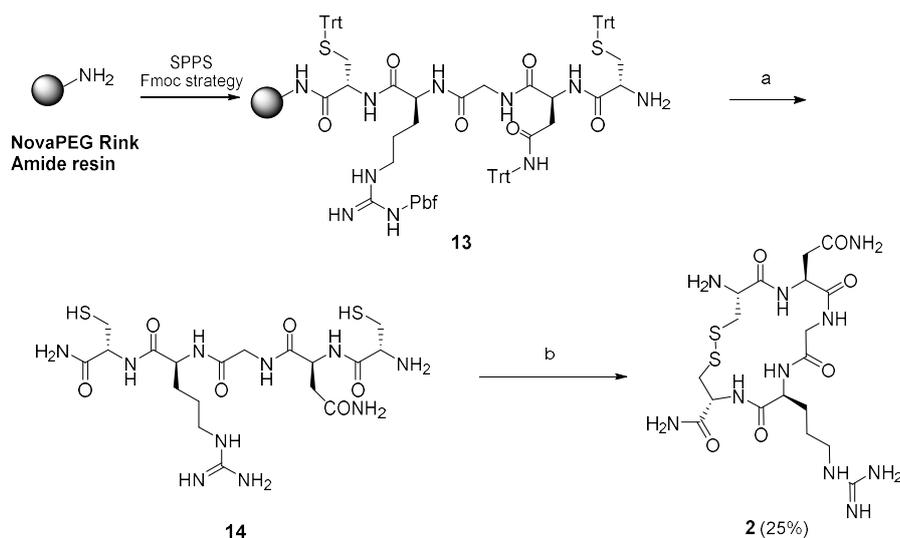
The deprotection/disulphide bridge formation by iodine oxidation was performed on the solid support before the coupling of biotin with the ϵ -amino group of the linker Fmoc-6-Ahx-OH, because of the high susceptibility of biotin to oxidants.

The on-resin cyclization step was performed following a procedure described in literature,¹⁸⁵ by adding to the hexapeptide-resin a solution of 10 equivalents of iodine in a mixture of DMF/H₂O 4:1 v/v.

After extensively washing the peptidyl resin with a 2% solution of ascorbic acid in DMF in order to quench the excess of iodine, the biotinylation step was performed. Biotin was activated in presence of the coupling reagent diisopropylcarbodiimide (DIC) and the nucleophile additive oxyma pure, after warming at 60 °C to achieve complete dissolution of biotin in DMF, as reported in literature.¹⁸⁶ The activated biotin was added to the resin and the resulted mixture was allowed to react for 24 h at room temperature. The consequent cleavage and purification of the crude biotinylated cyclic peptide gave derivative **4** in very low yields.

Synthetic strategy for compound **2**

In order to improve the cyclization step and the final yield, a different synthetic approach was followed for compound **2**: the cleavage of the fully deprotected linear pentapeptide and the following disulphide bond formation starting from the corresponding -SH derivatives (Scheme 3.3).



Scheme 3.3. Reagents and conditions: SPPS Fmoc strategy: Fmoc cleavage) 20 % piperidine in DMF, rt, 10 min ($\times 2$); coupling) Fmoc-amino acid (2 equiv), HATU (1.96 equiv), DIPEA (4 equiv), DMF, rt, 2 h. Sequence of addition: Fmoc-Cys(Trt)-OH, Fmoc-Arg(Pbf)-OH, Fmoc-Gly-OH, Fmoc-Asn(Trt)-OH, Fmoc-Cys(Trt)-OH; a) TFA/TIPS/H₂O 95:2.5:2.5, rt, 2 h; b) NCS (1 equiv), H₂O, rt, 15 min.

To this end the acid-labile trityl (Trt) was selected for the cysteine S-protecting group. The CNGRC sequence **13** was assembled on a NovaPEG Rink Amide resin using HATU as coupling agent as described in the general procedure. After Fmoc removal of the last cysteine, treatment with TFA cleavage mixture gave the fully deprotected linear amide peptide **14** (Scheme 3.3).

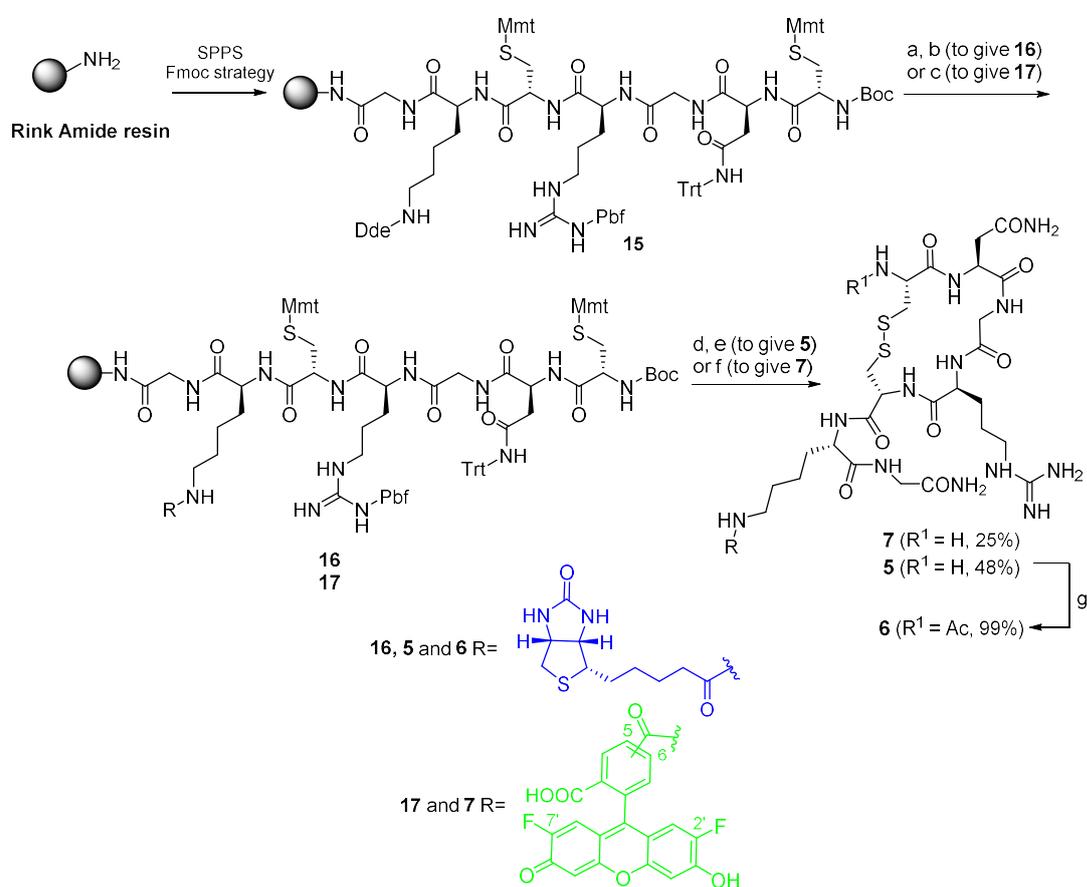
The oxidation of free thiol groups was achieved in aqueous medium by adding one equivalent of *N*-chlorosuccinimide (NCS) in 15 min at room temperature, according to a procedure reported by Postma and Albericio.¹⁸⁷ The crude cyclic peptide was lyophilized, analyzed by LC-MS and subsequently purified by RP-HPLC to afford the pure product in good yield.

This strategy overcame the problems associated with the use of strong oxidants such as iodine either in solid and in solution phases and the pseudodilution limitations of solid-

phase disulphide formation. Furthermore aqueous NCS oxidation resulted in an efficient and rapid synthetic approach.

Synthetic strategy for C-terminus conjugated peptides 5–7

Taking into account the advantages of the synthetic strategy followed for compound **2**, a similar approach based on the use of the acid-labile *p*-methoxytrityl (Mmt) as the cysteine S-protecting group was exploited for conjugates **5–7**. Peptide **16** was assembled on a MBHA Rink resin HL, while peptide **17** was assembled on a NovaPEG Rink amide resin as showed in Scheme 3.4.

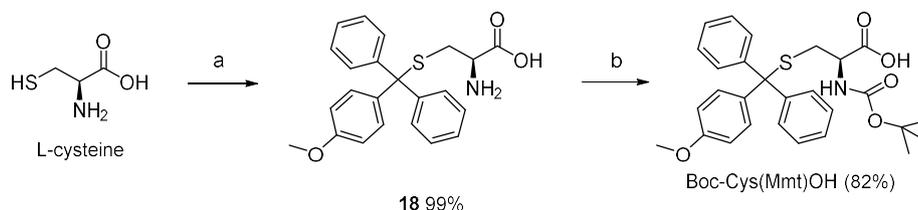


Scheme 3.4. Reagents and conditions: SPPS Fmoc strategy: Fmoc cleavage) 20 % piperidine in DMF, rt, 10 min (x2); coupling) Fmoc-amino acid (2 equiv), HATU (1.96 equiv), DIPEA (4 equiv), DMF, rt, 2 h. Sequence of addition: Fmoc-Gly-OH, Fmoc-Lys(Dde)-OH, Fmoc-Cys(Mmt)-OH, Fmoc-Arg(Pbf)-OH, Fmoc-Gly-OH, Fmoc-Asn(Trt)-OH, Boc-Cys(Mmt)-OH; a) 2% NH₂NH₂ in DMF, rt, 3 min (x2) ; b) biotin (2 equiv), HATU (1.96 equiv), DIPEA (4 equiv), DMF, rt, 2 h (x 2); c) Oregon green 488 (2 equiv), HATU (1.96 equiv), DIPEA (4 equiv), DMF, rt, 2 h; d) TFA/TIPS/H₂O 95:2.5:2.5, rt, 2 h; e) atmospheric O₂, 50 mM PB, pH 7.2, rt, 72 h; f) NCS (1 equiv), H₂O/CH₃CN 9:1, rt, 15 min; g) acetic anhydride (3.4 equiv), H₂O, rt, 10 min.

The assembling of the CRGNC sequence was performed after loading the resin with Fmoc-Gly-OH and coupling with Fmoc-Lys(Dde)OH. The Lys side chain bearing an amino group was selected as the anchorage point for biotin (derivatives **5** and **6**) and fluorescent probe (derivative **7**) conjugation. 1-(4,4-Dimethyl-2,6-dioxacyclohexylidene)ethyl (Dde) was used as protecting group for the -NH_2 function of lysine side chain, for its stability under Fmoc deprotection conditions.¹⁸⁸ On the other hand, Fmoc protecting group is readily deprotected under the classical Dde cleavage conditions (2% v/v hydrazine in DMF).

Therefore an acid-labile Boc protective group was used for the heptapeptide N-terminus in order to achieve an orthogonal protection between the lysine side chain and the α -amino group of the last aminoacid of the sequence.

For this purpose a cysteine bearing the acid-labile Boc and Mmt protective groups, for amino and thiol functions respectively (Boc-Cys(Mmt)-OH), was prepared (Scheme 3.5). 4-Methoxytrityl alcohol readily condensed with the primary -SH group of the cysteine in presence of boron trifluoride (BF_3) in AcOH as catalyst, according to a procedure reported in literature,¹⁸⁹ to afford the intermediate **18** (Scheme 3.4), which subsequently was allowed to react with di-tert-butyl dicarbonate to yield the Boc-Cys(Mmt)-OH derivative.



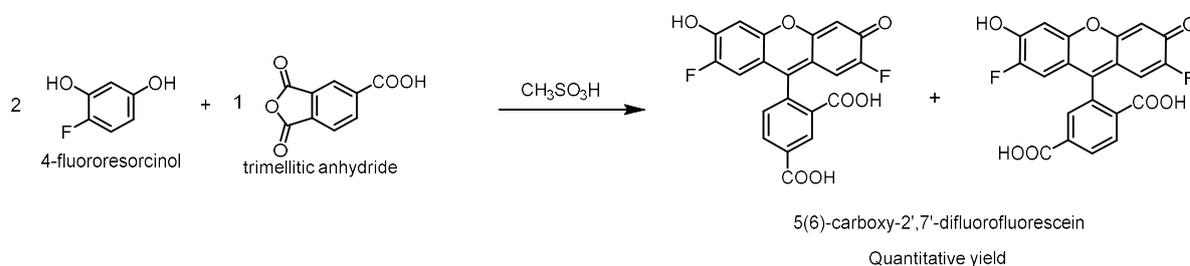
Scheme 3.5. Reagents and conditions a) 4-methoxytrityl alcohol, BF_3 , AcOH; b) Boc_2O , Na_2CO_3 , 1,4-dioxane.

After the last aminoacid of the heptapeptide sequence (**15**) was added, the orthogonal protecting group of lysine (Dde) was removed with hydrazine and conjugation with biotin or 5(6)-carboxy-2',7'-difluorofluorescein (Oregon green 488; OG), was accomplished on-resin through the amino function of the lysine side-chain, using HATU as coupling agent, to give intermediates **16** and **17**, respectively. The coupling reaction was repeated twice to ensure effective peptide conjugation. For the synthesis of intermediate **16**, warming at 80 °C was necessary to completely dissolve biotin in DMF before adding the coupling

agent and the tertiary amine. The solution of activated biotin was then added to the resin at room temperature. After cleavage of the biotinylated precursor **16** from the resin, conjugates **5** and **6** were prepared forming the disulphide bond by air oxidation. Indeed, since in a first attempt treatment of biotinylated linear peptides with NCS resulted in a mixture of undesired by-products, milder oxidation conditions were necessary. The crude linear peptide was dissolved in 50 mM phosphate buffer (pH 7.2) and the solution was stirred while bubbling air through it. The disulphide bridge formation was monitored by LC-MS and was complete in 72 h. After aqueous medium concentration, the resulting white precipitate (compound **5**) was collected and washed several times with H₂O and ACN and no further purification was needed. This synthetic route gave the best final yield. Treatment of compound **5** with acetic anhydride quantitatively afforded derivative **6**, which was purified by semi-preparative RP-HPLC.

Before conjugation of the linear CNGRCKG peptide with OG, the synthesis of the fluorescent dye was performed. OG was selected as the fluorescent probe for its advantages over fluorescein dyes: conjugates of the OG fluorophore are more photostable and moreover their fluorescence is not appreciably quenched, even at high degrees of labelling.

The reaction between 4-fluororesorcinol and trimellitic anhydride in methanesulfonic acid gave a mixture of OG 5- and 6-regioisomers (Scheme 3.6),¹⁹⁰ which was then used for on-resin peptide conjugation (intermediate **17**, Scheme 3.4).



Scheme 3.6

Precursor **17** was cleaved from the solid support and the desired conjugate **7** was obtained performing the cyclization step in solution phase using NCS, in less than 15 minutes (Scheme 3.4). The crude peptide was obtained as a mixture of isomers **7a** e **7b** (Figure 3.12) which were separated by semi-preparative RP-HPLC, using isocratic elution conditions (26% ACN 0.1% TFA in H₂O 0.1 % TFA). The OG conjugated peptide was obtained with a 25 % yield, based on the estimated loading of the resin and appeared as an orange powder.

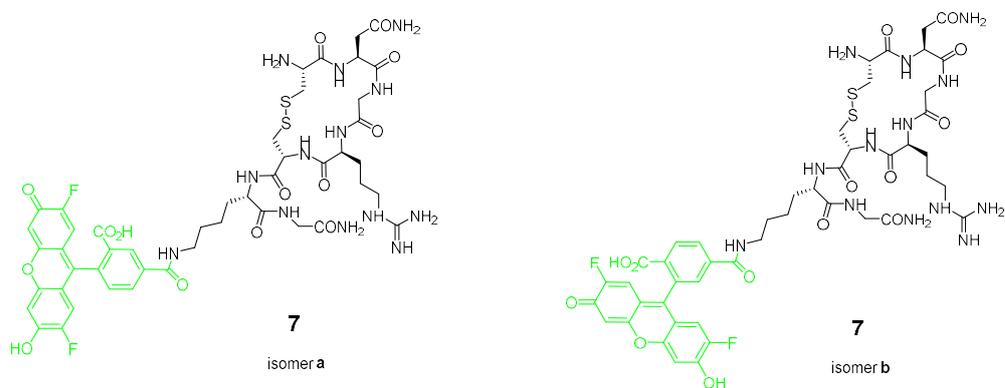


Figure 3.12

3.3.2. Biology

APN inhibition assay¹⁸⁰

The cCNGRC peptide **1** and its derivatives **2–6** were examined for their inhibitory activity toward porcine APN using a spectrophotometric assay commonly employed for screening APN inhibitors.^{177,191,192}

APN obtained from porcine kidneys microsomes, (activity = 15 units/mg) was used for its high grade of omology with the human enzyme.

APN and the substrate L-leucine-*p*-nitroanilide were incubated for 1 h in the presence of gradient concentrations of peptides **1–6** and formation of the product *p*-nitroaniline was measured reading absorbance at 405 nm (Figure 3.13).

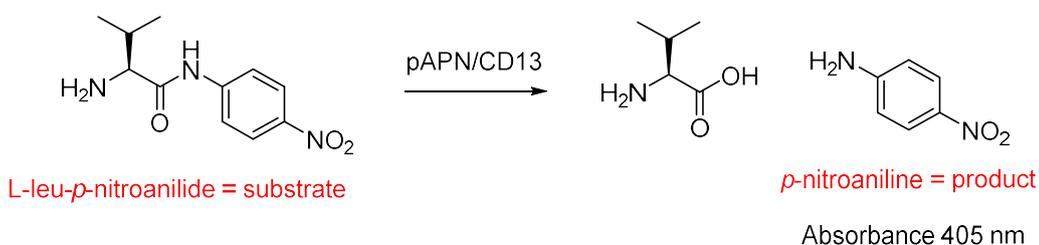


Figure 3.13. Schematic representation of the reaction catalyzed by APN in the biological assay.

The APN inhibitor bestatin was used as a positive control, while biotin as a negative one. The results of the assay show the effects of *C*- or *N*-terminus functionalization on the activity of **1**. In accordance with data previously reported,¹⁹¹ the inhibitory potency of **1** was considerably lower than that of bestatin (about 120-fold). Replacement of the *C*-terminus COOH of **1** with a CONH₂ function (derivative **2**) had negligible effect on the inhibitory potency of the parent peptide (Table 3.1 and Figure 3.14 a, b). Conversely, the activity was markedly impaired upon *N*-terminus acetylation (compound **3**) or by conjugation of amino function to a biotinylated linker (compound **4**). As shown in Figure 3.14 a), at a concentration of 1.25 mM of peptides **3** and **4** the inhibition of APN activity was less than 30%. Interestingly, bioconjugate **5** (consisting of biotin fused through a lysine linker to the *C*-terminus of **1**) showed an increased inhibitory potency relative to the parent peptide **1** (almost 15-fold, Table 3.1). The latter result is in accordance with the IC₅₀ value of 38.7 μM reported in the literature¹⁶³ for the c(CNGRC)G peptide,

suggesting that an additional amino acid directly linked to the C-terminus of cCNGRC peptides can lead to an increase of APN inhibitory potency.

Table 3.1. IC₅₀ values of bestatin and NGR peptides 1-6.^a

compound	IC ₅₀ (μ M)
Bestatin	5.1
1	636
2	620
3	>1000
4	>1000
5	43.3
6	>1000

^a IC₅₀ values were calculated from concentration-response curves by a nonlinear regression analysis using GraphPad Prism Software.

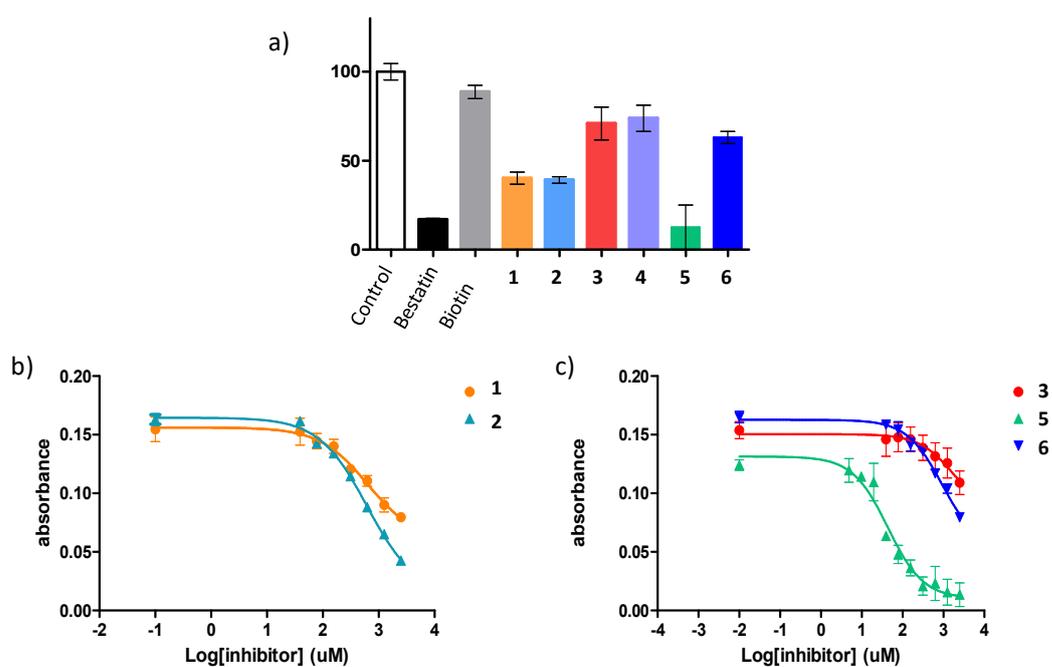


Figure 3.14. Results of enzyme inhibition assays. Effect of bestatin (25 μ M), biotin (1.25 mM) and peptides 1–6 (1.25 mM) on the enzymatic activity of APN. The results are expressed as percentage of APN residual activity relative to the catalytic activity of APN in the absence of any inhibitor taken as 100%; (b) inhibition of APN by peptides 1 and 2; (c) inhibition of APN by peptides 3, 5 and 6. Data are the average of two independent experiments performed in duplicate, with error bars displaying standard deviations.

Similarly to what was previously observed with the *N*-acetylated derivative **3**, acetylation of the free *N*-terminus of peptide **5** significantly impaired the activity of the resulting *C*-biotinylated compound **6** (Figure 3.14 a, c).

Overall, these results show that the *C*-terminus is the optimal “handle” for cCNGRC peptides derivatization, whereas the *N*-terminus must remain in free NH₂ form for achieving higher APN inhibitory activity.

APN inhibition assay on HT1080 cells¹⁸⁰

To further confirm the results of the enzymatic assay in a more physiological system, the inhibitory activities of *C*-derivatives **2** and **5** and the corresponding *N*-acetylated derivatives **3** and **6** toward human APN expressed on the surface of living cells were also evaluated. Human fibrosarcoma cell line (HT1080) is known to express high levels of APN and has been extensively used to measure human APN activity. Adherent HT1080 cells and the substrate L-leucine-*p*-nitroanilide were incubated for 1 h in the presence of gradient concentrations of inhibitors and formation of the product *p*-nitroaniline was measured on a plate reader at 405 nm. Bestatin was included in the test as a positive control. As shown in Figure 3.15, the inhibitory activities of the tested compounds exactly replicated the results obtained with the purified enzyme.

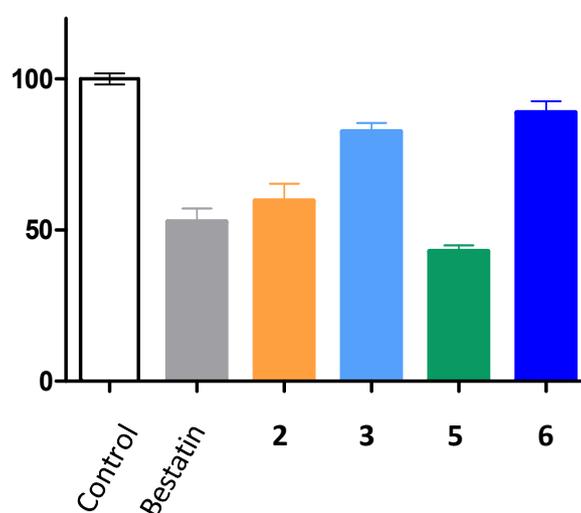


Figure 3.15. Results of APN inhibition assays on cells. Inhibition of APN expressed on HT1080 cells by bestatin (25 μ M) and NGR inhibitors (1.25 mM). The results are expressed as percentage of APN residual activity relative to the untreated control (100%). The data are the average of three replicates \pm standard deviation.

The *N*-acetylated derivative **3** showed reduced activity compared to **2**, while the *C*-biotinylated **5** displayed the highest inhibitory potency of the series. The activity of the *N*-acetylated derivative **6** was significantly reduced, in complete agreement with the results previously obtained with the purified enzyme.

Imaging of APN in HT1080 cells with c(CNGRC)KG-OG (7)¹⁸⁰

Bioconjugate **7**, a cCNGRC peptide carrying the OG 488 as a fluorescent probe at the C-terminus was used to image APN in living cells. Oregon green was selected as fluorescent probe because its high quantum efficiency and resistance to photobleaching make it a good and flexible tool for both wide-field and confocal imaging.¹⁹⁰ To decide the best concentration for APN detection on HT1080, cells were incubated at 0 °C for 1 h, then fixed and imaged under a fluorescent microscope at 20× magnification. The fluorescent signal of **7** was saturable and approached the maximum at a concentration of 100 μM (Figure 3.16 a, b).

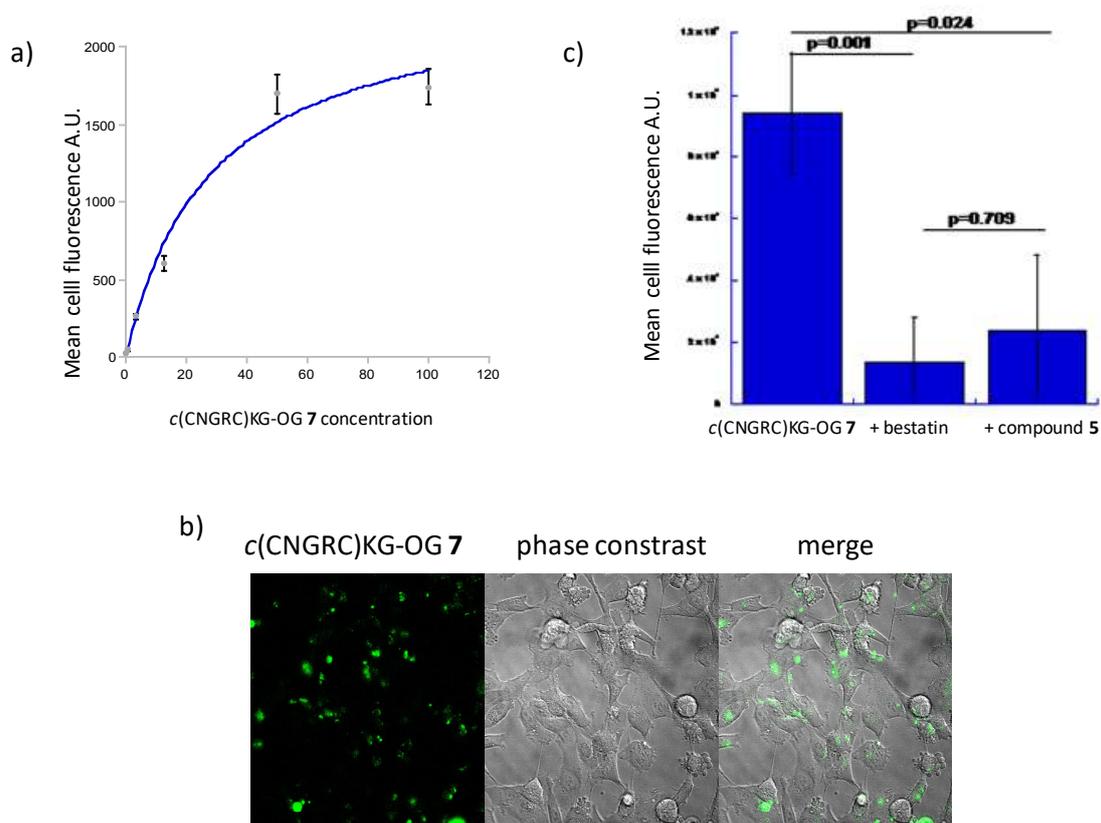


Figure 3.16. Tracer binding studies of c(CNGRC)KG-OG 7 on HT1080 cells. a) Binding curve of c(CNGRC)KG-OG 7 on HT1080 human fibrosarcoma cells. b) Incubation with tracer **7** shows binding to HT1080 cells. c) Quantification of the fluorescent signal present in the cells showing that the binding of tracer **7** to HT1080 cells is effectively inhibited by co-incubation with bestatin (100 μM) and c(CNGRC)KG-biotin **5** (1 mM).

The binding of the fluorescent probe was greatly reduced in the presence of bestatin (100 μ M) or *c*CNGRCKG-biotin **5** (1 mM), confirming that the binding is indeed through APN (Figure 3.16 c).

It has been shown that APN is rapidly internalized after binding on the cell surface.¹⁹³ The internalization is fundamental for some of the APN functions and is exploited by some coronavirus to enter and infect epithelial cells.¹⁹⁴ In addition APN internalization has been used to specifically target cancer cells with NGR conjugated to cytotoxic agents. Also the ability of the fluorescent probe **7** to undergo APN mediated endocytosis was analyzed. As shown in Figure 3.17 the fluorescent conjugate **7** is fully internalized and present in intracellular vesicles after 1 h at 37 °C.

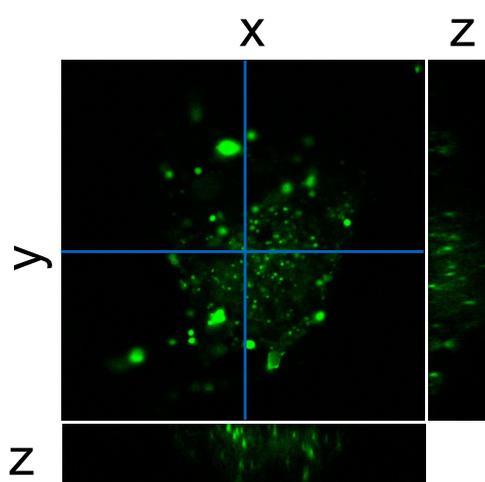


Figure 3.17. Internalization of the conjugate **7** via APN-mediated endocytosis.

3.4. Conclusions

In the context of NGR tumor-homing peptides, the cyclic CNGRC peptide has been used as carrier for delivery of drugs and imaging probes into tumors in several studies. This work was focused on exploiting the structural requirements of this cyclic peptide for effective APN binding. In particular the role of *N*-terminal amine group of the cCNGRC in the APN binding process was unclear. Recognition of the contribution of this group to the binding is of remarkable importance for the rational design of cCNGRC-based bioconjugates as effective tools for tumour targeting. Therefore our attention was focused on the right functionalization site of this cyclic peptide.

To this aim, the basic structure of the cyclic peptide with the free *N* and *C* termini, peptide **1** and a small library of cCNGRC peptides differently functionalised on the *N*-terminus and on the *C*-terminus was synthesized. An efficient solid-phase synthesis of cCNGRC peptides and conjugates carrying different functional tags on either the amino- or the carboxyl-terminus, which could be prepared in a hundreds of mgs scale, has been developed.

The effect of the structural modifications on the APN inhibitory potency was investigated by means an APN inhibition assay either on the purified porcine APN enzyme and in a human cancer cell line (HT1080). Only peptides with the free amino group were able to inhibit APN activity, therefore the free amino terminus is important for the APN recognition and the *C*-terminus is the optimal site for the formation of CNGRC-based derivatives with retained activity towards APN.

A fluorescent CNGRC-based probe, consisting of OG-488 fused through a lysine linker to the *C*-terminus of CNGRC, was synthesised and used to image APN expression on HT1080 cells.

Biological studies showed the cellular uptake of the fluorescent conjugate. It will be used for further molecular imaging studies.

Since (1) the porcine APN residues involved in the binding to cCNGRC peptides are completely conserved between human and porcine APNs, and (2) APN inhibitory activity of peptides **1–6** and cell internalization capacity of **7** were also confirmed on human cells, these results should be translatable to the development of more efficient cCNGRC constructs and conjugates carrying toxic payloads or radioisotopes for pre-clinical and

clinical use. These findings can guide the future development of new NGR bioconjugates and peptidomimetics with high affinity for APN.

CHAPTER 4

4. Experimental section

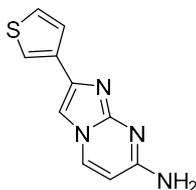
4.1. Experimental data for Chapter 2

4.1.1. Chemistry

General methods

All the compounds prepared were at least 98% pure as determined by combustion analysis. Melting points are uncorrected. TLC was performed on Bakerflex plates (Silica gel IB2-F) and column chromatography on Kieselgel 60 (Merck): the eluent was a mixture of petroleum ether/acetone in various proportions. IR spectra were recorded in nujol on a Nicolet Avatar 320 E.S.P.; ν max is expressed in cm^{-1} . $^1\text{H-NMR}$ spectra were recorded on a Varian MR 400 MHz (ATB PFG probe); the chemical shift (referenced to solvent signal) is expressed in δ (ppm) and J in Hz (abbreviations: ar=aromatic, im=imidazole, ph=phenyl, py=pyridine, pym=pyrimidine, th=thiazole, imi=imidazoline). The IR, $^1\text{H-NMR}$ and $^{13}\text{C-NMR}$ spectra of the compounds are in agreement with the assigned structures. 2,4-diaminopyrimidine **11** is commercially available. The following compounds were prepared according to the literature: 2-bromo-1-(thiophen-3-yl)ethan-1-one,¹⁹⁵ **14a**,¹²¹, **15**,¹⁹⁶, **20**,¹⁹⁷ **21**,¹⁹⁸ and **25**.¹²⁴

*Synthesis of 2-(thiophen-3-yl)imidazo[1,2-a]pyrimidin-7-amine (**12b**)*

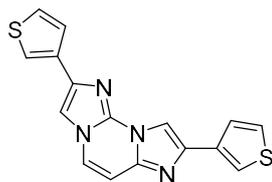


12b

2,4-Diaminopyrimidine **11** (10 mmol) was dissolved in 40 mL of acetone and treated with 2-bromo-1-(thiophen-3-yl)ethan-1-one (15 mmol). The reaction mixture was refluxed for 3h and the resulting precipitate (**12b** hydrobromide) was collected by filtration. The free base was obtained by treatment with 15% NH_4OH . The resulting precipitate was collected by filtration and crystallized from ethanol.

Yield 46%; mp 285-290 °C; IR (cm⁻¹): 3148, 1667, 1087, 846, 720; ¹H-NMR (400 MHz, DMSO-*d*₆): δ = 8.36 (1H, d, pym, *J* =7.6), 7.80 (1H, dd, th, *J* =2.6, *J* =1.2), 7.74 (1H, s, im), 7.57 (1H, dd, th, *J* =5.2, *J* =2.6), 7.48 (1H, dd, th, *J* =5.2, *J* =1.2), 7.16 (2H, s, NH₂), 6.37 (1H, d, pym, *J* =7.6). Anal. Calcd for C₁₀H₈N₄S (MW 216.26): C, 68.90; H, 5.14; N, 4.46. Found: C, 69.04; H, 5.02; N, 4.78.

Synthesis of 2,8-di(thiophen-3-yl)diimidazo[1,2-a:1',2'-c]pyrimidine (13b)

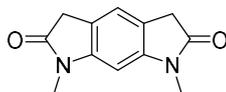


13b

Compound **12b** (5 mmol) was dissolved in ethanol (40 mL), treated with excess of 2-bromo-1-(thiophen-3-yl)ethan-1-one (20 mmol) and kept under reflux for 7h. The resulting compound **13b**.hydrobromide was collected by filtration and treated with 15% NH₄OH in order to prepare the free base which was collected by filtration and crystallized from ethanol.

Yield 12%, mp 245-249 °C dec; IR (cm⁻¹): 1635, 1292, 1193, 852, 724; ¹H-NMR (400 MHz, DMSO-*d*₆): δ = 8.61 (1H, s, im), 8.28 (1H, d, pym, *J* =7.6), 8.09 (1H, s, im), 7.99 (1H, dd, th, *J* =3.0, *J* =1.2), 7.83 (1H, dd, th, *J* =3.0, *J* =1.2), 7.75 (1H, dd, th, *J* =5.0, *J* =1.2), 7.65 (1H, dd, th, *J* =5.0, *J* =3.0), 7.63 (1H, dd, th, *J* =5.0, *J* =3.0), 7.56 (1H, dd, th, *J* =5.0, *J* =1.2), 7.16 (1H, d, pym, *J* =7.6). Anal. Calcd for C₁₆H₁₀N₄S₂ (MW 322.41): C, 59.61; H, 3.13; N, 17.38. Found: C, 60.03; H, 3.67; N, 17.99.

Synthesis of 1,7-dimethyl-5,7-dihydropyrrolo[3,2-f]indole-2,6(1H,3H)-dione (**16**)



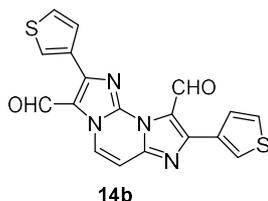
16

Compound **15** (10 mmol) was dissolved in toluene (20 mL) and treated with a solution of NaOH (20 mmol) in H₂O (4 mL). Chloroacetyl chloride (20 mmol) was added dropwise at 0-15°C. The reaction mixture was kept for 3 h at room temperature. The aqueous phase was removed and the organic layer was washed with HCl 2N and with H₂O. Toluene was removed under reduced pressure and the resulting residue was treated with AlCl₃ (10 g) at 160°C for 6 h. The mixture was then poured onto ice. The crude compound **16** was collected by filtration and crystallized from ethanol. Yield 45%; mp 275 °C dec; IR (cm⁻¹): 1705, 1625, 1276, 1025, 641; ¹H-NMR (400 MHz, DMSO-*d*₆): δ = 7.13 (1H, s, ar), 6.76 (1H, s, ar), 3.48 (4H, s, CH₂), 3.14 (6H, s, CH₃). Anal. Calcd for C₁₂H₁₂N₂O₂ (MW 216.24): C, 66.65; H, 5.59; N, 12.96. Found: C, 68.03; H, 6.05; N, 13.11.

General procedure for the synthesis of the aldehydes 14b and 17, 22, 23

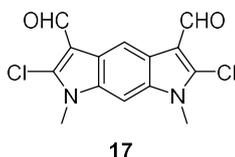
The Vilsmeier reagent was prepared at 0-5°C by dropping POCl₃ (54 mmol) into a stirred solution of DMF (65 mmol) in CHCl₃ (5 mL). The appropriate starting compound (**13b**, **16**, **20**, **21**, 5 mmol) was suspended in CHCl₃ (20 mL). The mixture thus obtained was dropped into the Vilsmeier reagent while maintaining stirring and cooling. The reaction mixture was kept for 3 h at room temperature and under reflux for 2-40 h (according to a TLC test). Chloroform was removed under reduced pressure, the resulting oil was poured onto ice and the precipitate thus obtained was collected by filtration. The crude aldehydes **14b** and **17** were crystallized from ethanol. The aldehydes **22** and **23** were isolated by column chromatography (petroleum ether/acetone, 70/30).

2,8-Di(thiophen-3-yl)diimidazo[1,2-a:1',2'-c]pyrimidine-3,9-dicarbaldehyde(14b).



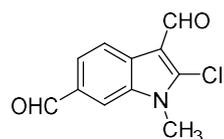
Yield 99%; mp 242-247 °C dec; IR (cm⁻¹): 1662, 1627, 1319, 866, 794; ¹H-NMR (400 MHz, DMSO-*d*₆): δ = 11.38 (1H, s, CHO), 10.25 (1H, s, CHO), 9.21 (1H, d, pym, *J*=7.8), 8.90 (1H, dd, th, *J*=3.0, *J*=1.2), 8.44 (1H, dd, *J*=2.4, *J*=1.2), 7.99 (1H, dd, th, *J*=5.2, *J*=1.2), 7.80 (2H, m, th), 7.68 (1H, dd, th, *J*=5.2, *J*=3.0), 7.57 (1H, d, pym, *J*=7.8). Anal. Calcd for C₁₈H₁₀N₄O₂S₂ (MW 378.43): C, 57.13; H, 2.66; N, 14.81. Found: C, 56.98; H, 2.74; N, 15.01.

2,6-Dichloro-1,7-dimethyl-1,7-dihydropyrrolo[3,2-f]indole-3,5-dicarbaldehyde(17)



Yield 85%; mp 300 °C dec; IR (cm⁻¹): 1644, 1506, 1096, 1049, 718; ¹H-NMR (400 MHz, CF₃COOD): δ = 9.82 (2H, s, CHO), 9.02 (1H, s, ar), 7.79 (1H, s, ar), 4.12 (6H, s, CH₃).
Anal. Calcd for C₁₄H₁₀Cl₂N₂O₂ (MW 309.15): C, 54.39; H, 3.26; N, 9.06. Found: C, 56.22; H, 2.99; N, 8.98.

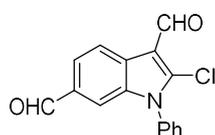
2-Chloro-1-methyl-1H-indole-3,6-dicarbaldehyde(22).



22

Yield 2%; mp 181-183 °C; IR (cm⁻¹):1700, 1686, 1646, 1154, 722; ¹H-NMR (400 MHz, DMSO-*d*₆): δ = 10.07 (1H, s, CHO), 10.03 (1H, s, CHO), 8.24 (1H, d, ind-7, *J* =1.2), 8.23 (1H, d, ind-4, *J* =8.4), 7.83 (1H, dd, ind-5, *J*=8.4, *J* =1.2), 3.93 (3H, s, CH₃). Anal. Calcd for C₁₁H₈ClNO₂ (MW 221.64): C, 59.61; H, 3.64; N, 6.32. Found: C, 60.04; H, 3.87; N, 5.98.

2-Chloro-1-phenyl-1H-indole-3,6-dicarbaldehyde(23).



23

Yield 7%; mp 190-195 °C; IR (cm⁻¹): 1654, 1608, 1151, 827, 725; ¹H-NMR (400 MHz, DMSO-*d*₆): δ = 10.16 (1H, s, CHO), 10.00 (1H, s, CHO), 8.36 (1H, d, ind-4, *J* =8.4), 7.89 (1H, dd, ind-5, *J* =8.4, *J* =1.6), 7.71 (5H, m, ph), 7.66 (1H, d, ind-7, *J* =1.6).

Anal. Calcd for C₁₆H₁₀ClNO₂ (MW 283.71): C, 67.74; H, 3.55; N, 4.94. Found: C, 68.02; H, 3.86; N, 5.08.

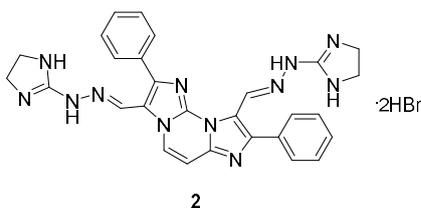
General procedure for the synthesis of hydrazones 2-5 and 6-10

The appropriate aldehyde or ketone (5 mmol) was dissolved in ethanol and treated with two equivalents of :

- 2-hydrazino-2-imidazoline hydrobromide solubilised in ethanol (compounds **2** and **3**);
- 2-hydrazinopyridine solubilized in ethanol and treated with hydrochloridric acid (compounds **4** and **10**);
- aminoguanidinehydrogencarbonatsuspended in ethanol and treated with hydrochloridric acid in order to achieve a solution (compounds **5-9**).

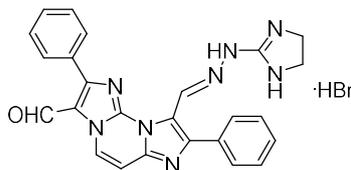
The reaction mixture was refluxed for 15-30 h according to a TLC test and the resulting precipitate was collected by filtration and crystallized from ethanol. Compound **3** was obtained as a precipitate from the filtered solution.

*3,9-Bis((2-(4,5-dihydro-1H-imidazol-2-yl)hydrazono)methyl)-2,8-diphenyldiimidazo[1,2-*a*:1',2'-*c*]pyrimidine dihydrobromide (**2**)*



Yield 22%; mp 300 °C; IR (cm⁻¹): 3360, 1655, 1071, 938, 718; ¹H-NMR (400 MHz, DMSO-*d*₆): δ = 12.50 (2H, broad, NNH), 9.49 (1H, s, CH), 9.30 (1H, d, pym, J=8.0), 8.80 (2H, broad, NH/NH+imi), 8.64 (1H, s, CH), 8.29 (2H, broad, NH/NH+imi), 8.12 (2H, d, ph, J =8.2), 7.84 (2H, d, ph, J =8.2), 7.58 (3H, m, ph), 7.49 (3H, m, ph), 7.37 (1H, d, pym, J=8.0), 3.81 (4H, s, CH₂), 3.72 (4H, s, CH₂); ¹³C-NMR (101 MHz, DMSO-*d*₆): δ = 157.68, 157.07, 146.45, 145.73, 142.16, 140.22, 139.77, 139.73, 132.70, 131.80, 129.51, 129.22, 129.07, 128.97, 128.87, 128.38, 128.00, 117.74, 115.28, 103.69, 42.92, 42.82. Anal. Calcd for C₂₈H₂₆N₁₂.2HBr (MW 692.43): C, 48.57; H, 4.08; N, 24.27. Found: C, 48.93; H, 3.95; N, 24.87.

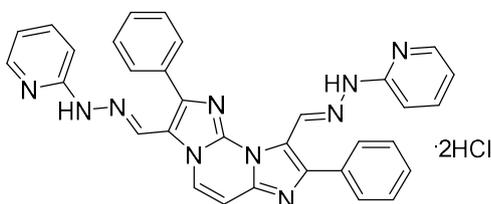
9-((2-(4,5-Dihydro-1H-imidazol-2-yl)hydrazono)methyl)-2,8-diphenyldiimidazo[1,2-a:1',2'-c]pyrimidine-3-carbaldehyde hydrobromide(**3**).



3

Yield 50%; mp 320 °C; IR (cm⁻¹): 3321, 1668, 1625, 774, 699; ¹H-NMR (400 MHz, DMSO-*d*₆): δ = 12.71 (1H, broad, NNH), 10.09 (1H, s, CHO), 9.47 (1H, s, CH), 9.12 (1H, d, pym, *J* = 7.4), 8.29 (2H, broad, NH/NH⁺imi), 8.12 (2H, m, ph), 8.03 (2H, m, ph), 7.61 (3H, m, ph), 7.52 (1H, d, pym, *J* = 7.4), 7.50 (3H, m, ph), 3.72 (4H, s, CH₂); ¹³C-NMR (101 MHz, DMSO-*d*₆): δ = 180.42, 157.71, 152.81, 145.94, 142.39, 140.47, 139.92, 132.51, 131.02, 130.17, 129.63, 129.52, 129.01, 128.95, 128.01, 126.27, 120.24, 117.91, 105.05, 42.93. Anal. Calcd for C₂₅H₂₀N₈O·HBr (MW 529.39): C, 56.72; H, 4.00; N, 21.17. Found: C, 57.03; H, 3.89; N, 21.84.

2,8-Diphenyl-3,9-bis((2-(pyridin-2-yl)hydrazono)methyl)diimidazo[1,2-a:1',2'-c]pyrimidine dihydrochloride (**4**).

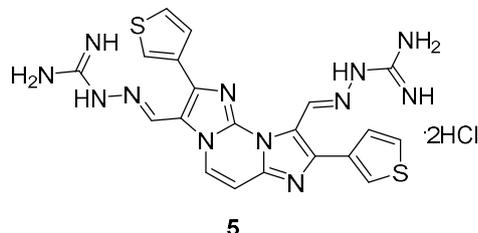


4

Yield 71%; mp 267-270 °C; IR (cm⁻¹): 3380, 1651, 1612, 1276, 707; ¹H-NMR (400 MHz, DMSO-*d*₆): δ = 13.13 (1H, broad, NNH), 12.82 (1H, broad, NNH), 9.64 (1H, s, CH), 9.56 (1H, d, pym, *J* = 8.0), 8.83 (1H, s, CH), 8.22 (1H, d, py, *J* = 6.0), 8.14 (2H, d, ph, *J* = 7.2), 8.09 (1H, d, py, *J* = 6.0), 8.06 (1H, t, py, *J* = 8.4), 7.95 (1H, t, py, *J* = 8.4), 7.93 (2H, d, ph, *J* = 7.2), 7.56 (2H, t, ph, *J* = 7.2), 7.43-7.53 (4H, m, 3H ph+1H py), 7.40 (1H, d, py, *J* = 8.4), 7.28 (1H, d, py, *J* = 8.4), 7.17 (1H, d, pym, *J* = 8.0), 7.08 (1H, t, ph, *J* = 7.2), 7.00 (1H, t, py, *J* = 6.0), 3.75 (2H, broad, NH⁺py); ¹³C NMR (151 MHz, DMSO-*d*₆) δ = 146.08, 144.90, 141.83, 139.66, 133.15, 131.81, 129.31, 129.08, 129.03, 128.87, 128.68, 128.39, 127.91, 118.85, 115.53, 114.96,

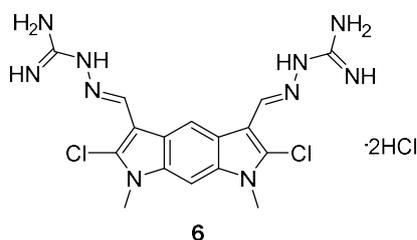
114.59, 111.67, 103.76. Anal. Calcd for $C_{32}H_{24}N_{10} \cdot 2HCl$ (MW 621.52): C, 61.84; H, 4.22; N, 22.54. Found: C, 62.05; H, 4.59; N, 22.87.

2,2'-((2,8-Di(thiophen-3-yl)diimidazo[1,2-a:1',2'-c]pyrimidine-3,9-diyl)bis(methanylylidene))bis(hydrazine-1-carboximidamide) dihydrochloride (5)



Yield 98%; mp 265-269 °C dec; IR (cm^{-1}): 3358, 1675, 1623, 1149, 866; 1H -NMR (400 MHz, $DMSO-d_6$): δ = 12.57 (1H, s, NNH), 12.02 (1H, s, NNH), 9.69 (1H, s, CH), 9.26 (1H, d, pym, J = 8.0), 8.82 (1H, s, CH), 8.47 (1H, dd, th, J = 2.8, J = 1.2), 8.26 (1H, dd, th, J = 2.8, J = 1.2), 7.81 (1H, dd, th, J = 5.2, J = 2.8), 7.78 (1H, dd, th, J = 5.2, J = 1.2), 7.76 (1H, dd, th, J = 5.2, J = 1.2), 7.69 (8H, broad, NH_2/NH_2^+), 7.66 (1H, dd, th, J = 5.2, J = 2.8), 7.28 (1H, d, pym, J = 8.0); ^{13}C -NMR (101 MHz, $DMSO-d_6$): δ = 155.20, 154.67, 142.05, 141.77, 141.27, 139.57, 139.19, 138.43, 134.00, 133.95, 128.56, 128.44, 127.77, 127.45, 126.88, 126.32, 125.72, 117.54, 114.63, 103.26. Anal. Calcd for $C_{20}H_{18}N_{12}S_2 \cdot 2HCl$ (MW 563.49): C, 42.63; H, 3.58; N, 29.83. Found: C, 42.97; H, 3.02; N, 30.05.

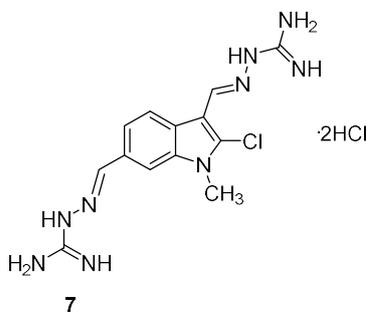
2,2'-((2,6-Dichloro-1,7-dimethyl-1,7-dihydropyrrolo[3,2-f]indole-3,5-diyl)bis(methanylylidene))bis(hydrazine-1-carboximidamide) dihydrochloride (6)



Yield 50%; mp 350 °C dec; IR (cm^{-1}): 3329, 1643, 1516, 1148, 1096; 1H -NMR (400 MHz, $DMSO-d_6$): δ = 11.70 (2H, broad, NNH), 8.68 (1H, s, ar), 8.47 (2H, s, CH), 7.88 (1H, s, ar), 7.60 (8H, broad, NH_2/NH_2^+), 3.88 (6H, s, CH_3); ^{13}C -NMR (101 MHz, $DMSO-d_6$): δ = 154.53, 143.06, 134.61, 130.97, 119.67, 112.71, 105.66, 91.84, 30.65. Anal. Calcd for

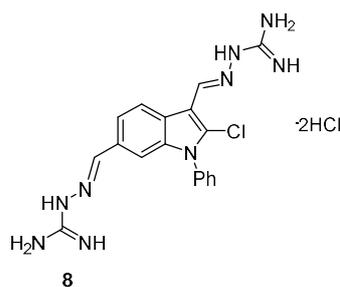
C₁₆H₁₈Cl₂N₁₀·2HCl (MW 494.21): C, 38.89; H, 4.08; N, 28.34. Found: C, 39.03; H, 3.87; N, 28.84.

2,2'-((2-Chloro-1-methyl-1H-indole-3,6-diyl)bis(methanylylidene))bis(hydrazine-1-carboximidamide) dihydrochloride (**7**)



Yield 75%; mp 235-240 °C; IR (cm⁻¹): 3150, 1670, 1605, 1129, 934; ¹H-NMR (400 MHz, DMSO-*d*₆): δ = 12.08 (2H, broad, NNH), 8.38–8.34 (2H, m, ind), 8.30 (1H, s, CH), 8.15 (1H, s, CH), 7.70 (8H, broad, NH₂/NH₂⁺gua), 7.67 (1H, d, ind, *J* = 7.8), 3.85 (3H, s, CH₃); ¹³C-NMR (101 MHz, DMSO-*d*₆): δ = 155.44, 154.86, 147.32, 141.62, 136.15, 131.93, 128.72, 124.42, 121.95, 121.59, 109.60, 106.59, 30.57. Anal. Calcd for C₁₃H₁₆ClN₉·2HCl (MW 406.70): C, 38.39; H, 4.46; N, 31.00. Found: C, 38.78; H, 4.73; N, 29.02.

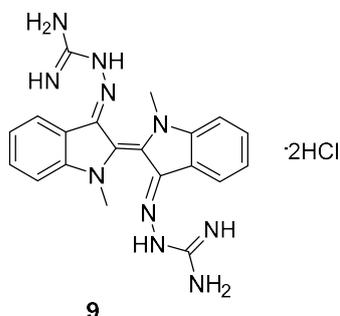
2,2'-((2-Chloro-1-phenyl-1H-indole-3,6-diyl)bis(methanylylidene))bis(hydrazine-1-carboximidamide) dihydrochloride (**8**)



Yield 67%; mp 350 °C dec; IR (cm⁻¹): 3163, 1666, 1631, 1158, 934; ¹H-NMR (400 MHz, DMSO-*d*₆): δ = 12.11 (2H, broad, NNH), 8.48 (1H, d, ind-4, *J* = 8.6), 8.47 (1H, s, CH), 8.25 (1H, s, CH), 7.88 (d, ph, *J* = 8.5 Hz, 2H), 7.85 – 7.60 (m, 10 H), 7.58 (d, ph, *J* = 8.1 Hz, 2H), 7.49 (s, 1H); ¹³C-NMR (101 MHz, DMSO-*d*₆): δ = 155.37, 154.93, 147.71, 141.50, 136.74, 134.31, 131.33, 129.98, 129.55, 129.47, 128.16, 124.54, 122.39, 121.45, 110.57, 108.16.

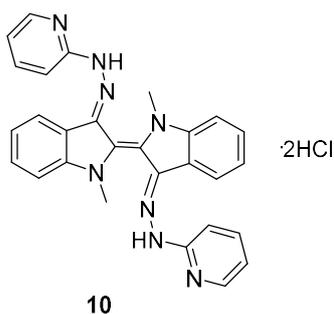
Anal. Calcd for $C_{18}H_{18}ClN_9 \cdot 2HCl$ (MW 468.77): C, 46.12; H, 4.30; N, 26.89. Found: C, 45.97; H, 4.78; N, 27.02.

2,2'-(2E-1,1'-dimethyl-[2,2'-biindolinylidene]-3,3'-diylidene)bis(hydrazine-1-carboximidamide)dihydrochloride (9)



Yield 32%; mp 315-320 °C; IR (cm^{-1}): 3326, 1667, 1571, 1041, 722; 1H -NMR (400 MHz, $DMSO-d_6$): δ = 12.48 (2H, s, NNH), 8.52 (8H, broad, NH_2/NH_2^+), 7.75 (2H, d, ind, $J=7.6$), 7.49 (2H, t, ind, $J=7.6$), 7.18 (2H, t, ind, $J=7.6$), 7.16 (2H, d, ind, $J=7.6$), 3.22 (6H, s, CH_3); ^{13}C -NMR (101 MHz, $DMSO-d_6$): δ = 160.18, 155.88, 144.09, 136.26, 132.08, 123.00, 121.31, 118.88, 109.94, 25.79. Anal. Calcd for $C_{20}H_{22}N_{10} \cdot 2HCl$ (MW 475.38): C, 50.53; H, 5.09; N, 29.46. Found: C, 50.98; H, 4.86; N, 29.78.

2E-1,1'-dimethyl-3,3'-bis(2-(pyridin-2-yl)hydrazono)-2,2'-biindolinylidene dihydrochloride (10)



Yield 25%; mp 255-260 °C; IR (cm^{-1}): 3397, 1685, 1638, 1117, 1043; 1H -NMR (400 MHz, $DMSO-d_6$): δ = 12.84 (2H, s, NNH), 8.32 (2H, m, ind), 7.97 (2H, t, ind, $J=7.6$), 7.67 (4H, m, ind+py), 7.42 (2H, t, ind, $J=7.6$), 7.16 (6H, m, py), 3.72 (2H, broad, NH^+ py), 3.26 (6H, s, CH_3); ^{13}C -NMR (101 MHz, $DMSO-d_6$): δ = 161.25, 154.46, 147.83, 142.02, 139.17, 129.57, 129.31, 122.65, 119.76, 119.04, 118.60, 109.54, 107.53, 25.58. Anal. Calcd for $C_{28}H_{24}N_8 \cdot 2HCl$ (MW 545.47): C, 61.65; H, 4.80; N, 20.54. Found: C, 61.20; H, 5.02; N, 20.12.

4.1.2. Biophysical evaluation

Oligonucleotide synthesis and sample preparation

The DNA sequences were synthesized using standard β -cyanoethylphosphoramidite solid phase chemistry on an ABI 394 DNA/RNA synthesizer (Applied Biosystem) at the 5- μ mol scale. DNA detachment from support and deprotection were performed by treatment with concentrated ammonia aqueous solution at 55 °C for 12 h. The combined filtrates and washings were concentrated under reduced pressure, dissolved in water, and purified by high-performance liquid chromatography (HPLC) on a Nucleogel SAX column (Macherey-Nagel, 1000-8/46), using buffer A consisting of 20 mM $\text{KH}_2\text{PO}_4/\text{K}_2\text{HPO}_4$ aqueous solution (pH 7.0), containing 20% (v/v) CH_3CN , buffer B consisting of 1 M KCl, 20 mM $\text{KH}_2\text{PO}_4/\text{K}_2\text{HPO}_4$ aqueous solution (pH 7.0), containing 20% (v/v) CH_3CN , and a linear gradient from 0% to 100% B for 30 min with a flow rate 1 mL/min. The fractions of the oligomers were collected and successively desalted by Sep-pak cartridges (C-18). The isolated oligomer was proved to be >98% pure by NMR. In particular, the following oligonucleotides have been synthesized: d[$\text{TAGGG}(\text{TTAGGG})_3$] (Tel23) and d[($\text{TTAGGG})_4\text{TT}$] (Tel26) that are two different truncations of human telomeric sequence; two sequences that occur in the nuclease hypersensitive region of the promoter of *c-KIT* oncogene d($\text{AGGGAGGGCGCTGGGAGGAGGG}$) (c-kit1) and d($\text{CGGGCGGGCGCGAGGGAGGGG}$) (c-kit2); the modified c-kit2 promoter sequence d($\text{CGGGCGGGCGCTAGGGAGGGT}$) (c-kit2_{T12/T21}); the *c-MYC* promoter sequence d($\text{TGAGGGTGGGGAGGGTGGGGAAGG}$) (c-myc); the modified *c-MYC* promoter sequence d($\text{TGAGGGTGGGTAGGGTGGGTAA}$) (Myc22) and the self-complementary duplex-forming dodecamer d(CGCGAATTCGCG) (ds12). The concentration of oligonucleotides was determined by UV adsorption measurements at 90 °C using appropriate molar extinction coefficient values ϵ ($\lambda = 260$ nm) calculated by the nearest neighbour model. Samples were heated at 90 °C for 5 min, and then gradually cooled to room temperature overnight. Parallel arrangement of telomeric sequence (Tel23-p) was prepared as previously described.¹⁹⁹

Circular dichroism spectroscopy

Circular dichroism (CD) experiments were recorded on a Jasco J-815 spectropolarimeter equipped with a PTC-423S/15 Peltier temperature controller. All the spectra were recorded at 20 °C in the wavelength range of 230-360 nm and averaged over 3 scans. The scan rate was set to 100 nm/min, with a 1 s response time, and 1 nm bandwidth. Buffer baseline was subtracted from each spectrum. For the CD experiments in presence of metal cations, 10 µM G4 DNA and 15 µM duplex DNA were used. CD spectra of DNA/ligand mixtures were obtained by adding 4 molar equiv. of ligands (stock solutions of ligands were 6 mM in DMSO). DNAs were prepared in 10 mM potassium phosphate (for c-kit1, Tel23 and Tel26) or 10 mM lithium phosphate (for Tel23-p and ds12) buffers (pH 7.0) containing 100 mM KCl. Because the T_m value of the c-myc G4 alone is above 80 °C in buffer containing 100 mM KCl, it is very difficult to evaluate the stabilizing effects of adding ligands, thus a buffer solution containing lower amounts of K^+ (5 mM potassium phosphate, 20 mM KCl) was used for the experiments on c-myc G4. As for c-kit2, it was shown that this sequence forms a dimeric G4 in 100 mM K^+ -containing solution, with a completely different fold compared to the monomeric G4 that is formed in buffer solution containing low amounts of K^+ . Therefore, 5 mM potassium phosphate buffer containing 20 mM KCl was also used in this case.

CD melting were carried out in the 20-100 °C temperature range, at 1 °C/min heating rate by following changes of CD signal at the wavelengths of the maximum CD intensity. CD melting experiments were recorded in the absence and presence of ligands (4 molar equiv.) added to the folded DNA structures. For selected compounds, CD melting experiments were also performed by varying ligand concentration. Three different [ligand]/[DNA] ratio were examined (1:1, 2:1 and 4:1). The melting temperatures (T_m) were determined from curve fit using Origin 7.0 software. ΔT_m values were determined as the difference in melting temperature between the G4-DNA with and without ligands.

FRET melting studies

FRET melting studies were performed by using the following G4 forming sequences: the 5'-FAM-d(GGGCGGGCGCGAGGGAGGGG)-TAMRA-3' (Fckit2T) and 5'-FAM-d(TGAGGGTGGGTAGGGTGGGTAA)-TAMRA-3' (FcmycT) oligonucleotides from the *c-KIT*

and *c-MYC* promoter regions, respectively; and the telomeric F21T G4-forming sequence were annealed at high concentrations to promote the parallel conformation (F21T-p). Labeled oligonucleotides were purchased from Biomers (Germany), purified employing standard HPLC protocols. Such studies were performed under the same buffer and experimental conditions used for the corresponding CD experiments. The FRET melting assay was carried out on a FP-8300 spectrofluorometer (Jasco) equipped with a Peltier temperature controller accessory (Jasco PCT-818) with excitation at 492 nm and detection at 522 nm. Both excitation and emission slit widths were set at 5 nm. A sealed quartz cuvette with a path length of 1 cm was used. The final concentration of the G4s was 0.2 μ M. The fluorescence melting of the G4s was monitored at 1 $^{\circ}$ C/min with and without ligands (at a 4:1 ligand/DNA ratio). Final analysis of the data was carried out using Origin 7.0.

Nuclear magnetic resonance experiments

NMR experiments were performed on a 700 MHz Varian UnityINOVA spectrometer, with data recorded at 25 $^{\circ}$ C. One-dimensional proton spectra of the sample in H₂O were recorded using pulsed-field gradient DPGSE for H₂O suppression. DNA samples were prepared at 0.2-0.5 mM strand concentration in 0.6 mL (H₂O/D₂O 9:1) buffer solution. NMR experiments on G4 forming oligonucleotides from *c-MYC* (Myc22) and *c-KIT* (*c-kit*_{T12/T21}) promoters were performed employing the same buffers and experimental conditions as used for the determination of their 3D structures, in order to avoid different G4 folds or conformational heterogeneity in solution. The solution was either 25 mM KH₂PO₄, 70 mM KCl, 0.2 mM EDTA, pH 7.0 (for Myc22) or 5 mM KH₂PO₄, 20 mM KCl, 0.2 mM EDTA, pH 6.8 (for *c-kit*_{T12/T21}). Aliquots of a stock solution of **3** in DMSO-*d*₆ were added directly to the DNA solution inside the NMR tube. The NMR data were processed on an iMAC running iNMR software (www.inmr.net).

4.1.3. Computational studies

Docking simulations

The solution structures of G4 formed by the human c-kit2 promoter sequence (PDB ID: 2KQH), and of c-myc G4 bound to quindoline molecule (PDB ID: 2L7V) were used as the targets for docking studies. The ligand found in the c-myc complex was removed from the structure to leave empty binding sites. After optimizing the ligand and assigning partial atomic charges, docking calculations were performed with the AutoDock 4.2 program using the Lamarckian genetic algorithm. The all parallel-stranded arrangement with propeller loops linking adjacent parallel strands on opposite surfaces of the G4s, results in accessible outer planar G-tetrad surfaces that, being the potential binding sites for the ligands, were defined as the grid boxes for the docking. The size of the boxes was constrained to $18 \text{ \AA} \times 18 \text{ \AA} \times 18 \text{ \AA}$ in the x, y and z dimensions. Grid maps were generated for each atom type in the ligand using AutoGrid. An active site box was created with a grid spacing of 0.375 \AA . The maximum number of energy evaluations was set to 2.5×10^6 , the maximum number of genetic algorithm operations was set to 2.7×10^4 , the number of individuals in a population was set to be 300, the rates of mutation and crossover were set to 0.02 and 0.8, respectively. When searching the conformational and orientation spaces of ligand with rotatable bonds having full flexibility, the structure of the G4 was kept rigid. The most favorable binding poses were selected on the basis of both the calculated energies and visual inspection. All the figures were drawn using VMD 1.9 software (www.ks.uiuc.edu/Research/vmd).

4.1.4. Biology

Cell cultures and immunofluorescence

Human osteosarcoma U2OS and colon cancer HCT116 cells were cultured in Dulbecco's Modified Eagle Medium (DMEM) (Carlo Erba), supplemented with 1% L-glutamine and 10% FBS (fetal bovine serum). U2OS cells were seeded in 35 mm dishes at a concentration of 100000 cell/ml. 24 h after seeding, cells were treated with **1** (2 μ M and 10 μ M), **3** (0.2 μ M and 1 μ M), or **14a** (2 μ M and 10 μ M). Cells were then fixed in methanol:acetic acid (3:1), permeabilized with 0.1% triton-X100/PBS and blocking in 2% milk/PBS. Immunofluorescence was performed using standard methods with BG4, anti-FLAG (Cell Signaling Technology) and anti-rabbit Alexa 488-conjugated (Invitrogen) antibodies. Nuclei were stained with DAPI (Sigma Aldrich) and coverslips were mounted with Mowiol[®] 4-88 (Sigma Aldrich). Fluorescence signal was determined using ImageJ software with the following formula: Corrected Total Cell Fluorescence (CTCF) = Integrate Density – (Area of selected cell x Mean Fluorescence of Background Readings). Significance was determined by parametric tests "Student's *t* test". BG4 antibody was obtained by transfection of BG4 plasmid (kindly obtained by S. Balasubramanian) in BL21 *E. coli* cells. Then, BG4 protein expression was induced by the autoinduction method as described by Studier. BG4 was purified by using silica-based resin (Protino[®] Ni-IDA) pre-charged with Ni²⁺ ions, eluted with 250 mM Imidazole/PBS pH 8.0. The eluted antibody was concentrated with Amicon Ultra-15 Centrifugal Filter Units (Millipore), and imidazole was finally removed by buffer exchange with PBS pH 8.0 with Amicon Ultra-15 Centrifugal Filter Units.

MTT cell proliferation assay

U2OS and HCT116 cells were cultured in DMEM as above, and were seeded in 24-well at a concentration of 40000 cell/ml. 24 h after seeding, cells were treated with the compound at the indicated concentration and time. After 1 h or 24 h of treatment, agents were removed and cells were further cultured in complete drug-free medium for 1-2 days. Then, thiazolyl blue tetrazolium bromide (MTT) (Sigma Aldrich) was added to each well and incubated for 1 h at 37 °C. Then, medium was removed and 300 μ l of dimethyl

sulfoxide (Sigma Aldrich) was added and incubated for 1 h at room temperature. Then, 100 μ l of the solution was put in 96-well and absorbance at 595 nm was measured using a multiplate reader. The linear regression parameters were determined to calculate the IC_{50} (GraphPad Prism 4.0, Graph Pad Software Inc.).

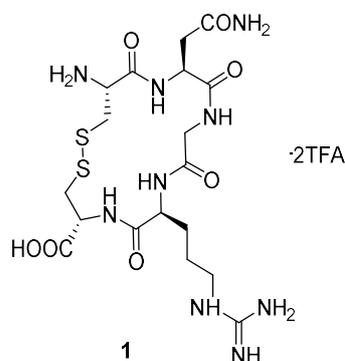
4.2. Experimental data for Chapter 3

4.2.1. Chemistry

General methods

^1H NMR, ^{13}C NMR and ^{19}F NMR spectra were recorded on a Bruker AVANCE III 400 NMR spectrometer. Chemical shifts (δ) are reported in parts per million (ppm) and coupling constants (J) are given in Hertz (Hz). Mass analyses were performed using Agilent 1200 HPLC system coupled to Agilent G6120 single quadrupole detector equipped with an electrospray ionization (ESI) source in direct infusion modality. ESI-MS spectra were recorded in positive mode. RP (reverse phase) HPLC-MS analyses were performed with an Agilent 1200 HPLC system equipped with a DAD and an ESI-MS detector. HPLC conditions for analytical analyses: Phenomenex Luna C18 column, 5 μm , 100 \AA , 250 \times 4.6 mm (L \times ID), inj. volume 20 μL , flow rate 1 mL/min. HPLC conditions for semi-preparative purification: Phenomenex Luna C18 column, 5 μm , 100 \AA , 250 \times 10.0 mm (L \times ID), flow rate 5 mL/min.

Synthesis of cCNGRC 1

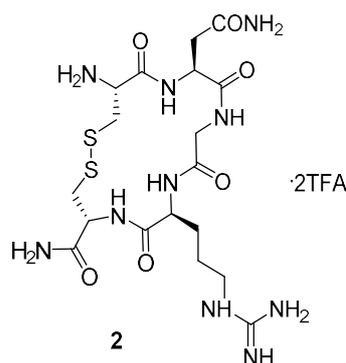


Fmoc-Cys(Acm) Wang resin (500 mg, 0.30 mmol, loading: 0.60 mmol/g) was swollen in DCM for 30 min and then washed with DMF. Fmoc cleavage was carried out using 20 % piperidine in DMF, at rt for 10 min (\times 2). The resin was washed with DMF (3 \times 5 mL) and DCM (3 \times 5 mL). The coupling reaction was performed by adding to the resin a solution of the Fmoc-amino acid (1.5 mmol, 5 equiv), HATU (560 mg, 1.47 mmol, 4.9 equiv), DIPEA (0.51 mL, 3.0 mmol, 10 equiv) in DMF (5 mL), and the mixture was shaken for 2 h. The coupling sequence was: Fmoc-Arg(Pbf)-OH, Fmoc-Gly-OH, Fmoc-Asn(Trt)-OH, Fmoc-

Cys(Acm)-OH. Once the linear pentapeptide was assembled, the *N*-terminal Fmoc protecting group was removed. Cleavage of the pentapeptide from the resin and simultaneous removal of Trt and Pbf protecting groups was performed using a mixture of TFA/TIPS/H₂O 95:2.5:2.5 v/v/v for 2 h. The cleavage mixture was drained off and the resin washed with TFA and DCM. The mixture was concentrated and Et₂O was added to precipitate the crude peptide. The crude linear peptide (270 mg, 0.30 mmol, 1 equiv) was dissolved in AcOH (15 mL) and added dropwise over 3 h to a solution of I₂ (0.12 M) in AcOH (150 mL) and HCl 10 M (1 mL).¹⁸⁴ The cyclization reaction was monitored by RP-HPLC coupled to tandem ESI-MS and stopped after 2 h by adding Et₂O. The resulting suspension was cooled in dry-ice and then centrifuged at 5000 rpm for 3 min, decanted and the resulting white solid was purified by semi-preparative RP-HPLC (Solvent A: CH₃CN 0.1% TFA, Solvent B: H₂O 0.1% TFA, gradient: from 1 to 7% of A in 8 min, flow 5 mL/min, *t_R* = 6 min). The purified compound was lyophilized and obtained as a white powder (35 mg, TFA salt, 15 % yield based on the estimated loading of the resin).

¹H NMR (400 MHz, D₂O): δ = 4.92 (t, *J* = 7.2 Hz, 1H), 4.64 (dd, *J* = 9.9, 3.5 Hz, 1H), 4.31 (t, *J* = 7.3 Hz, 1H), 4.24 – 4.08 (m, 2H), 3.67 (d, *J* = 15.6 Hz, 1H), 3.40 – 3.26 (m, 3H), 3.23 (t, *J* = 6.7 Hz, 2H) 3.12 (dd, *J* = 14.4, 9.9 Hz, 1H), 2.84 (dd, *J* = 15.3, 6.5 Hz, 1H), 2.72 (dd, *J* = 15.5, 8.0 Hz, 1H), 1.91 – 1.61 (m, 4H); ¹³C NMR (101 MHz, D₂O, as a TFA salt, the signal of TFA was not included): δ = 174.20, 173.81, 173.68, 171.57, 170.60, 167.26, 156.74, 53.61, 52.73 (one signal for two carbons, as detected by HSQC analysis), 50.07, 42.43, 41.18, 40.49, 40.43, 36.14, 27.22, 24.21. MS (ESI, *m/z*): calculated for C₁₈H₃₁N₉O₇S₂ 549.18; found 550.2 [M+H]⁺. RP-HPLC: solvent A: CH₃CN 0.1% TFA, Solvent B: H₂O 0.1% TFA, gradient: from 1 to 15% of A in 15 min, *t_R* = 7.6 min, flow: 1 mL/min, λ = 220 nm.

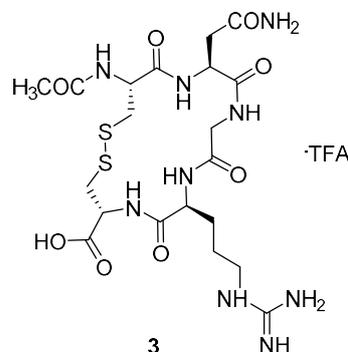
Synthesis of cCNGRC-CONH₂ **2**



Fmoc protected Rink Amide MBHA resin HL (1.0 g, 0.87 mmol, loading: 0.87 mmol/g) was used for this synthesis. Swelling, Fmoc cleavage and washing procedures were performed as described above for compound **1**. The coupling reactions were accomplished by adding to the resin a solution of the Fmoc-amino acid (1.74 mmol, 2 equiv), HATU (650 mg, 1.70 mmol, 1.96 equiv), DIPEA (0.6 mL, 3.48 mmol, 4 equiv) in DMF (5 mL) and shaking the mixture for 2 h. The amino acid coupling sequence was: Fmoc-Cys(Trt)-OH, Fmoc-Arg(Pbf)-OH, Fmoc-Gly-OH, Fmoc-Asn(Trt)-OH and Fmoc-Cys(Trt)-OH. Cleavage of the pentapeptide from the resin and simultaneous removal of Trt and Pbf protecting groups was performed using a mixture of TFA/TIPS/H₂O 95:2.5:2.5 v/v/v for 2 h. The mixture was concentrated and Et₂O was added to precipitate the crude peptide. *N*-chlorosuccinimide (60 mg, 0.45 mmol, 1 equiv) was added to the solution of the crude linear peptide (350 mg, 0.45 mmol, 1 equiv) in H₂O (20 mL) and the reaction was allowed to stir at room temperature for 15 min. The reaction mixture was then frozen in dry ice and freeze dried to obtain the crude cyclic peptide as a white powder. The crude peptide was purified by semi-preparative RP-HPLC (Solvent A: CH₃CN 0.1% TFA, Solvent B: H₂O 0.1% TFA, gradient: from 1 to 7% of A in 7 min, flow: 5 mL/min, t_R = 6.2 min) and lyophilized to give compound **2** as a white powder (170 mg, TFA salt 25 % yield based on the estimated loading of the resin). ¹H NMR (400 MHz, D₂O): δ = 4.92 (t, J = 7.3 Hz, 1H), 4.63 (dd, J = 10.4, 3.4 Hz, 1H), 4.31 (t, J = 7.3 Hz, 1H), 4.20 – 4.13 (m, 2H), 3.67 (d, J = 15.7 Hz, 1H), 3.38 (dd, J = 14.5, 6.8 Hz, 1H), 3.32 (dd, J = 14.5, 3.5 Hz, 1H), 3.27 – 3.18 (m, 3H), 3.05 (dd, J = 14.5, 10.4 Hz, 1H), 2.85 (dd, J = 15.5, 6.7 Hz, 1H), 2.73 (dd, J = 15.5, 7.9 Hz, 1H), 1.94 – 1.57 (m, 4H); ¹³C NMR (101 MHz, D₂O, as a TFA salt, the signal of TFA was not included): δ = 174.24, 174.21, 173.97, 171.62, 170.70, 167.23, 156.74, 53.50, 52.84, 52.69, 50.08, 42.40, 41.18, 40.51, 40.44, 36.03, 27.04, 24.27. MS (ESI, m/z): calculated for C₁₈H₃₂N₁₀O₆S₂548.19; found 549.2

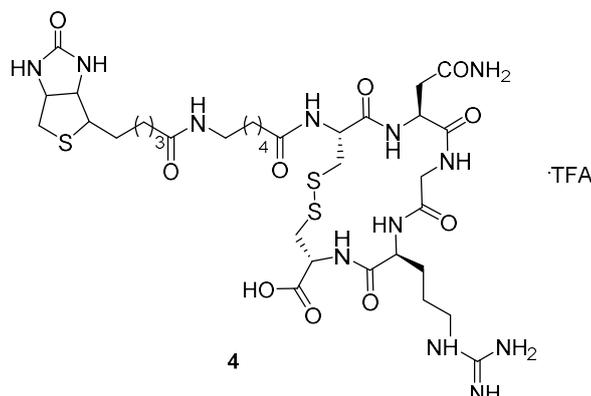
[M+H]⁺. RP-HPLC: solvent A: CH₃CN 0.1% TFA, Solvent B: H₂O 0.1% TFA, gradient: from 1 to 15% of A in 15 min, t_R = 8.0 min, flow: 1 mL/min, λ = 220 nm.

Synthesis of Ac-cCNGRC **3**



The linear precursor **8** (Scheme 3.1) was assembled on the Fmoc-Cys(Acm) preloaded Wang resin (500 mg, 0.30 mmol, loading: 0.60 mmol/g) following the procedure described for compound **1**. After Fmoc cleavage from the last amino acid, acetylation of the *N*-terminus was performed on-resin adding a solution of DCM/acetic anhydride/DIPEA 20:4:1 v/v/v (8 mL) and the resulting suspension was shaken for 1 h (repeated 2 times). Cleavage from the resin was carried out using a solution of TFA/TIPS/H₂O 95:2.5:2.5 v/v/v for 2 h. The mixture was concentrated and Et₂O was added to precipitate the crude peptide. Cyclization was performed following the procedure described for **1**.¹⁸⁴ The crude material was purified by semi-preparative RP-HPLC (Solvent A: CH₃CN 0.1% TFA, Solvent B: H₂O 0.1% TFA, gradient: from 10 to 15% of A in 7 min, flow: 5 mL/min, t_R = 5.7 min) and lyophilized to give compound **3** as a white powder (65 mg, TFA salt, 30 % yield based on the estimated loading of the resin). ¹H NMR (400 MHz, D₂O): δ = 4.74 (dd, *J* = 9.0, 3.6 Hz, 1H), 4.65 (t, *J* = 7.2 Hz, 1H), 4.49 – 4.39 (m, 2H), 4.22 (d, *J* = 16.4 Hz, 1H), 3.73 (d, *J* = 16.4 Hz, 1H), 3.33 (dd, *J* = 15.3, 4.4 Hz, 1H), 3.25 (t, *J* = 6.6 Hz, 2H), 3.21 – 3.18 (m, 1H), 3.15 (dd, *J* = 7.3, 3.8 Hz, 2H), 2.93 (dd, *J* = 15.4, 6.6 Hz, 1H), 2.71 (dd, *J* = 15.4, 7.7 Hz, 1H), 2.05 (s, 3H), 2.00 – 1.62 (m, 4H); ¹³C NMR (101 MHz, D₂O, as a TFA salt, the signal of TFA was not included): δ = 174.78, 174.39, 173.51, 173.41, 172.18, 172.02, 171.24, 156.73, 53.69, 53.39, 52.60, 50.61, 42.60, 40.96, 40.54, 40.47, 35.57, 27.59, 24.31, 21.74. MS (ESI, *m/z*): calculated for C₂₀H₃₃N₉O₈S₂ 591.19 found 592.2 [M+H]⁺. RP-HPLC: solvent A: CH₃CN 0.1% TFA, Solvent B: H₂O 0.1% TFA, gradient: from 10 to 25% of A in 11 min t_R = 6.7 min, flow 1 mL/min, λ = 220 nm.

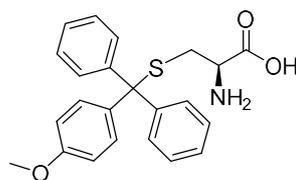
Synthesis of biotin- ϵ -Ahx-cCNGRC **4**



The linear precursor **11** (Scheme 3.2) was assembled on the Fmoc-Cys(Acm) preloaded Wang resin (100 – 200 mesh, 200 mg, 0.14 mmol, loading: 0.70 mmol/g) following the procedure described for compound **1**. After Fmoc cleavage from the last amino acid, a solution of Fmoc- ϵ -Ahx-OH (250 mg, 0.70 mmol, 5 equiv), DIC (0.11 mL, 0.70 mmol, 5 equiv) and Oxyma Pure Novabiochem® (100 mg, 0.70 mmol, 5 equiv) in DMF (3 mL) was added and the suspension was shaken for 2h. Then, the *N*-terminal Fmoc protecting group was removed. The cyclization step was carried out on resin by adding to the resin a solution of I₂ (350 mg, 1.4 mmol, 10 equiv) in DMF/H₂O 4/1 (v/v) (3 mL). After 4 h the solvent was removed and the resin was extensively washed with DMF (2 × 5 mL), 2% ascorbic acid in DMF (2 × 5 mL), DMF (5 × 5 mL), and DCM (3 × 5 mL). Biotin (340 mg, 1.4 mmol, 10 equiv), DIC (0.22 mL, 1.4 mmol, 10 equiv) and Oxyma Pure Novabiochem® (200 mg, 1.4 mmol, 10 equiv) were suspended in DMF (14 mL) at 60°C and stirred for about 40 min until the activated biotin was completely dissolved. After cooling to room temperature, the reaction mixture was added to the cyclic peptide-resin and shaken for 24 h. The resin was washed with DCM (5 × 10 mL) and DMF (5 × 10 mL). Biotin-labelled cyclic peptide was cleaved from the resin, along with the deprotection of the amino acid side chains by adding a mixture of TFA/TIPS/H₂O 95/2.5/2.5 (v/v/v) for 2 h. The mixture was concentrated and Et₂O was added to precipitate the crude peptide. The crude was dissolved in H₂O, purified by semi-preparative RP-HPLC (Solvent A: CH₃CN 0.1% TFA, Solvent B: H₂O 0.1% TFA, gradient: from 10 to 35% of A in 15 min, flow 5 mL/min, t_R=7 min) and lyophilized to give compound **4** as a white powder (9.6 mg, TFA salt, 7 % yield based on estimated loading of the resin). ¹H NMR (400 MHz, D₂O): δ = 4.66 – 4.57 (m, 2H),

4.50 (dd, $J = 8.4, 5.9$ Hz, 1H), 4.46 – 4.35 (m, 2H), 4.22 (d, $J = 16.3$ Hz, 1H), 4.00 – 3.96 (m, 1H), 3.72 (d, $J = 16.4$ Hz, 1H), 3.39 – 2.77 (m, 12H), 2.71 (dd, $J = 15.5, 7.5$ Hz, 1H), 2.40 – 2.23 (m, 4H), 1.96 – 1.30 (m, 16H). MS (ESI, m/z): calculated for $C_{34}H_{56}N_{12}O_{10}S_3$ 888.34; found 889.3 $[M+H]^+$. RP-HPLC: solvent A: CH_3CN 0.1% TFA, Solvent B: H_2O 0.1% TFA, gradient: from 5 to 50 % of A in 15 min, $t_R = 10.46$ min, flow: 1 mL/min, $\lambda = 220$ nm.

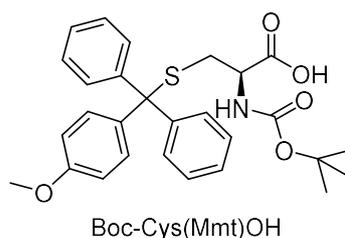
Synthesis of *S*-methoxytrityl-cysteine **18**



18

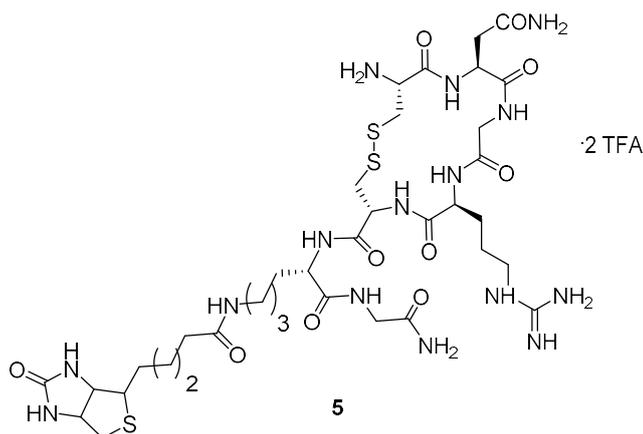
To a stirred suspension of anhydrous L-cysteine (1 g, 8,25 mmol, 1 equiv) in 8 mL of AcOH was added 4-methoxytrityl alcohol (2,8 g, 9,65 mmol, 1.17 equiv). BF_3 -etherate 1,8 mL was then added gradually in order to maintain the reaction mixture at 25 °C. An intense brown coloration developed and the suspended solid dissolved as the addition progressed. The reaction mixture was allowed to react for 2 h and to the solution was then added, with ice-cooling, 20 mL of saturated aqueous AcONa and 45 mL of H_2O . The mixture was stirred for 10-15 minutes and allowed to stand at 0 °C. The resulting gelatinous white solid was collected by decantation and then treated with Et_2O to achieve the precipitation of a white powder. Et_2O was then removed by filtration and the solid was then washed several times with Et_2O to obtain the pure *S*-methoxytrityl-cysteine **18** (3,1 g, quantitative yield). 1H NMR (400 MHz, CD_3OD) $\delta = 7.43$ (d, $J = 7.7$ Hz, 4H), 7.36 – 7.15 (m, 8H), 6.84 (d, $J = 8.9$ Hz, 2H), 3.75 (s, 3H), 3.10 (dd, $J = 8.7, 3.9$ Hz, 1H), 2.82 (dd, $J = 13.3, 3.9$ Hz, 1H), 2.71 (dd, $J = 13.2, 9.1$ Hz, 1H).

Synthesis of Boc-Cys(Mmt)-OH



To a solution of S-methoxytrityl-cysteine **18** (3,3 g, 8,8 mmol, 1 equiv) in 50% saturated aqueous NaHCO₃, Na₂CO₃ (0,9 g, 8,8 mmol, 1 equiv) was added and the resulting suspension was cooled at 0 °C. A solution of Boc₂O (1,9 g, 8,8 mmol, 1 equiv) in 1,4-dioxane 4 mL was then added and the reaction mixture was allowed to react at 0 °C for 30 min and at room temperature for 20 h. The mixture was washed with Et₂O to remove the excess of Boc₂O. The aqueous layer was treated with HCl 2N till pH = 4 and extracted with EtOAc (3 x 50 mL). The combined organic layers were dried over MgSO₄ and evaporated under reduced pressure to obtain a pure white solid (3,3 g, 82 % yield). ¹H NMR (400 MHz, CD₃OD): δ = 7.38 (d, *J* = 7.6 Hz, 4H), 7.27 (t, *J* = 7.5 Hz, 6H), 7.23 – 7.16 (m, 2H), 6.84 (d, *J* = 9.0 Hz, 2H), 4.05 (t, *J* = 6.2 Hz, 1H), 3.78 (s, 3H), 2.56 (d, *J* = 6.3 Hz, 2H), 1.44 (s, 9H).

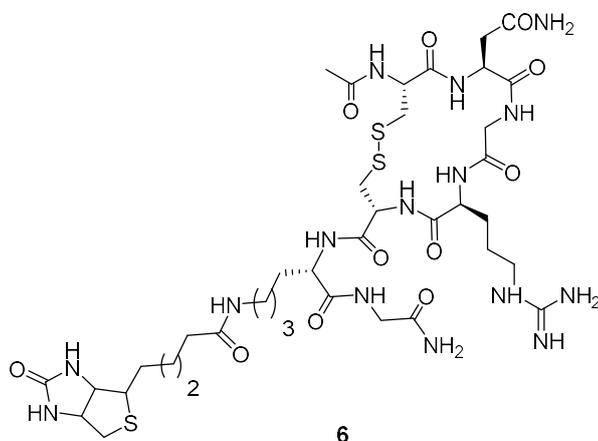
Synthesis of cCNGRCKG-biotin **5**



The linear peptide **15** was assembled on an Fmoc protected Rink Amide MBHA resin HL (1.0 g, 0.87 mmol, loading: 0.87 mmol/g) following the procedure already described for compound **2**. The amino acid coupling sequence was Fmoc-Gly-OH, Fmoc-Lys(Dde)-OH, Fmoc-Cys(Mmt)-OH, Fmoc-Arg(Pbf)-OH, Fmoc-Gly-OH, Fmoc-Asn(Trt)-OH, Boc-Cys(Mmt)-

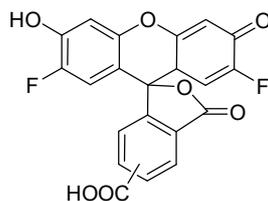
OH. Once the linear heptapeptide **15** was assembled, Dde removal was achieved using a solution (10 mL) of NH_2NH_2 (2%) in DMF for 3 min ($\times 2$). Biotin (420 mg, 1.74 mmol, 2 equiv) was dissolved in DMF (8 mL) and the solution warmed at 80 °C in order to solubilize the solid powder. Once biotin was completely solubilized, the mixture was cooled to rt before adding HATU (650 mg, 1.70 mmol, 1.96 equiv) and DIPEA (0.6 mL, 3.48 mmol, 4 equiv). The mixture was added to the resin and the suspension was shaken for 2 h. The coupling was repeated twice. Cleavage of the biotin labelled peptide **16** from the resin and simultaneous removal of Mmt, Trt and Pbf protecting groups was performed using a mixture of TFA/TIPS/ H_2O 95:2.5:2.5 v/v/v for 2 h. The mixture was concentrated and Et_2O was added to precipitate the crude peptide. The crude peptide (860 mg, 0.72 mmol) was dissolved in 50 mM phosphate buffer, pH 7.4 (250 mL) and the reaction was stirred bubbling air through it. The disulphide bridge formation was monitored by RP-HPLC coupled to tandem ESI-MS and it was complete after 72 h. The aqueous medium was concentrated under reduced pressure and the resulting precipitate was filtrated and washed several times with H_2O and CH_3CN . Compound **5** was obtained as a pure white powder 400 mg, 48% yield, based on the estimated loading of the resin). ^1H NMR (400 MHz, D_2O): δ = 4.79 (overlapped with solvent signal, 1H, as detected by COSY and HSQC experiments) 4.66 – 4.59 (m, 2H), 4.44 (dd, J = 7.7, 4.5 Hz, 1H), 4.40 – 4.34 (m, 1H), 4.31 (dd, J = 8.4, 5.5 Hz, 1H), 4.15 (d, J = 16.0 Hz, 1H), 3.96 (d, J = 17.1 Hz, 1H), 3.90 (d, J = 17.1 Hz, 1H), 3.75 (d, J = 16.2 Hz, 1H), 3.70 (br, 1H), 3.40 – 2.73 (m, 13H), 2.27 (t, J = 7.0 Hz, 2H), 1.96 – 1.30 (m, 16H). MS (ESI, m/z): calculated for $\text{C}_{36}\text{H}_{61}\text{N}_{15}\text{O}_{10}\text{S}_3$ 959.39; found ESI-MS: m/z 960.3 $[\text{M}+\text{H}]^+$. RP-HPLC: solvent A: CH_3CN 0.1% TFA, Solvent B: H_2O 0.1% TFA, gradient: from 5 to 40 % of A in 11 min, t_R = 8.2 min, flow: 1 mL/min, λ = 220 nm.

Synthesis of Ac-cCNGRCKG-biotin **6**



Compound **5** (15 mg, 15.6 μmol , 1 equiv) was dissolved in water (2.5 mL) and treated with acetic anhydride (5.0 μL , 52.9 μmol , 3.4 equiv) at rt for 10 min. The crude peptide was purified by semi-preparative RP-HPLC (Solvent A: CH_3CN 0.1% TFA, Solvent B: H_2O 0.1% TFA, gradient: from 5 to 50 % of A in 15 min, flow: 5 mL/min, t_R = 10.8 min) and lyophilized to give compound **6** as a white powder (17mg, quantitative yield). ^1H NMR (400 MHz, D_2O): δ = 4.70 – 4.61 (m, 3H), 4.52 (t, J = 7.1 Hz, 1H), 4.46 – 4.38 (m, 2H), 4.31 (dd, J = 8.7, 5.5 Hz, 1H), 4.20 (d, J = 16.4 Hz, 1H), 3.96 (d, J = 17.2 Hz, 1H), 3.89 (d, J = 17.3 Hz, 1H), 3.73 (d, J = 16.3 Hz, 1H), 3.38 – 3.09 (m, 9H), 3.01 (dd, J = 13.0, 4.9 Hz, 1H), 2.93 (dd, J = 15.5, 6.5 Hz, 1H), 2.83 – 2.70 (m, 2H), 2.27 (t, J = 7.1 Hz, 2H), 2.06 (s, 3H), 1.95 – 1.36 (m, 16H); ^{13}C NMR (101 MHz, D_2O): δ = 176.63, 174.66, 174.27, 174.25, 173.90, 173.59, 172.11, 171.89, 171.79, 171.44, 165.30, 156.71, 62.10, 60.25, 55.46, 54.03, 53.45, 53.42, 53.20, 50.50, 42.76, 42.03, 40.99, 40.53, 40.33, 39.70, 38.92, 35.64, 35.46, 30.33, 27.87, 27.80, 27.67, 27.51, 25.17, 24.40, 22.45, 21.73. MS (ESI, m/z): calculated for $\text{C}_{38}\text{H}_{63}\text{N}_{15}\text{O}_{11}\text{S}_3$ 1001.4; found 1002.4 $[\text{M}+\text{H}]^+$. RP-HPLC: solvent A: CH_3CN 0.1% TFA, Solvent B: H_2O 0.1% TFA, gradient: from 5 to 50 of A in 15 min, t_R = 10.5 min, flow: 1 mL/min, λ = 220nm.

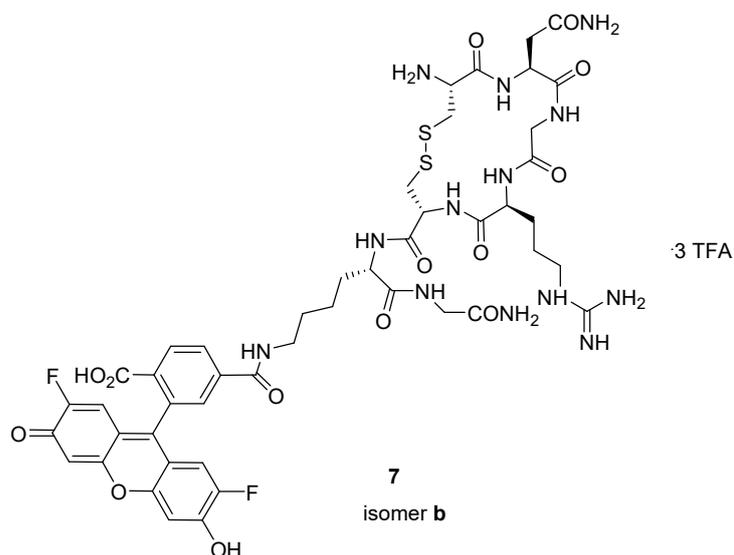
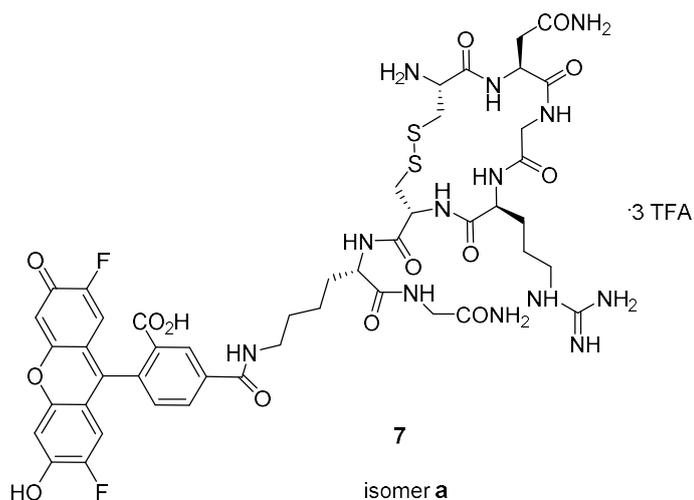
Synthesis of 5(6)-carboxy-2',7'-difluorofluorescein (Oregon green 488)



5(6)-carboxy-2',7'-difluorofluorescein

Trimethylic anhydride (750 mg, 3.9 mmol, 1 equiv) was added to a 1M solution of 4-fluororesorcinol (1 g, 7.8 mmol, 2 equiv) in methanesulfonic acid (7.8 mL). The resulting mixture was heated under dry nitrogen at 80 °C for 24 h. The cooled mixture was poured into 7 volumes of ice water and the resulting precipitate was filtrated, washed with H₂O, dried at 60 °C in vacuo to give 5(6)-carboxy-2',7'-difluorofluorescein as an orange powder (a mixture of 5 and 6 regioisomers was obtained according to ref. 190) (1.6 g, quantitative yield). ¹H NMR (400 MHz, CD₃OD) δ 8.64 (s, 1H), 8.42 (dd, *J* = 8.0, 1.5 Hz, 1H), 8.35 (dd, *J* = 8.0, 1.3 Hz, 1H), 8.14 (d, *J* = 7.7 Hz, 1H), 7.79 (s, 1H), 7.36 (d, *J* = 8.0 Hz, 1H), 6.90–6.85 (m, 4H), 6.46 (d, *J* = 10.2 Hz, 4H); ¹⁹F NMR (376 MHz, DMSO-*d*₆) δ -139.69, -139.73.

Synthesis of cCNGRCKG-OG 7



NovaPEG Rink Amide resin (1.0 g, 0.49 mmol, loading: 0.49 mmol/g) was used for the synthesis of the linear peptide **15**. Swelling, Fmoc cleavage and washing of the resin were performed as described for compound **1**. The coupling reaction was performed by adding to the resin a solution of the Fmoc-aa (2.45mmol, 5 equiv), HATU (913 mg, 2.4 mmol, 4.9 equiv), DIPEA (0.8 mL, 4.9mmol, 10 equiv) in DMF (5 mL), and the mixture was shaken for 2 h. The amino acid coupling sequence was Fmoc-Gly-OH, Fmoc-Lys(Dde)-OH, Fmoc-Cys(Mmt)-OH, Fmoc-Arg(Pbf)-OH, Fmoc-Gly-OH, Fmoc-Asn(Trt)-OH, Boc-Cys(Mmt)-OH. After assembling the linear peptide, the Dde removal was achieved as described for compound **5** and the fluorescent probe OG (as a mixture of 5 and 6 regioisomers) was then coupled by adding to the resin the solution of OG (404 mg, 0.98 mmol, 2 equiv), HATU (365 mg; 0.96 mmol, 1.96 equiv) and DIPEA (0.33 mL, 1.96 mmol, 4 equiv) in DMF

(6 mL) and the resulting red suspension was shaken for 2 h. The resin was extensively washed with DMF (3 x 5 mL), DCM (3 x 5 mL), MeOH containing 0.1% NH₄OH conc. (10 x 5 mL). Cleavage of the OG labelled peptide **17** from the resin and simultaneous removal of Mmt, Trt and Pbf protecting groups were performed using a mixture of TFA/TIPS/H₂O 95:2.5:2.5 v/v/v for 2 h. The mixture was concentrated and Et₂O was added to precipitate the crude peptide. *N*-chlorosuccinimide (53 mg, 0.4 mmol, 1 equiv) was added to the solution of the crude linear peptide (540 mg, 0.4mmol, 1 equiv) in H₂O/CH₃CN 9:1 (40 mL) and the reaction was allowed to stir at room temperature for 15 min. The reaction mixture was then frozen in dry ice and freeze dried to obtain the crude cyclic peptide as an orange powder. The crude peptide was purified by semi-preparative RP-HPLC (Solvent A: CH₃CN 0.1% TFA, Solvent B: H₂O 0.1% TFA, gradient: 26 % of A in 7 min, flow: 5 mL/min, **isomer A**, t_R = 5.6 min; **isomer B**, t_R = 6.1 min) and lyophilized to give compound **7 isomer A** and **isomer B** as an orange powder (180 mg, TFA salt, 25% yield based on the estimated loading of the resin). Isomer A): ¹H NMR (400 MHz, D₂O) δ = 8.45 (s, 1H), 8.10 – 8.03 (m, 1H), 7.34 – 7.25 (m, 1H), 6.82 – 6.75 (m, 2H), 6.67 – 6.56 (m, 2H), 4.91 (t, *J* = 7.1 Hz, 1H), 4.60 (dd, *J* = 8.9, 2.8 Hz, 1H), 4.40 – 4.31 (m, 1H), 4.26 – 4.16 (m, 2H), 4.11 (d, *J* = 15.8 Hz, 1H), 3.95(d, *J* = 17.6 Hz, 1H), 3.88 (d, *J* = 17.2 Hz, 1H), 3.65 (d, *J* = 15.6 Hz, 1H), 3.47 – 2.98 (m, 8H), 2.83 (dd, *J* = 15.3, 6.3 Hz, 1H), 2.73 (dd, *J* = 15.5, 7.7 Hz, 1H), 1.98 – 1.36 (m, 10H); ¹⁹F NMR (376 MHz, D₂O) δ = -75.55 (TFA signal), -134.87. Isomer B): ¹H NMR (400 MHz, D₂O) δ = 8.10 (d, *J* = 6.8 Hz, 1H), 7.96 (d, *J* = 6.5 Hz, 1H), 7.54 (s, 1H), 6.80 – 6.77 (m, 2H), 6.56 (s, 2H), 4.91 (t, *J* = 7.1 Hz, 1H), 4.58 (dd, *J* = 9.1, 3.6 Hz, 1H), 4.27 – 4.16 (m, 3H), 4.11 (d, *J* = 15.7 Hz, 1H), 3.91(d, *J* = 17.2 Hz, 1H), 3.84 (d, *J* = 17.2 Hz, 1H), 3.65 (d, *J* = 15.7 Hz, 1H), 3.40 – 2.98 (m, 8H), 2.84 (dd, *J* = 15.6, 6.4 Hz, 1H), 2.73 (dd, *J* = 15.5, 7.8 Hz, 1H), 1.82 – 1.26 (m, 10H); ¹⁹F NMR (376 MHz, D₂O) δ = -75.55 (TFA signal), -134.87. MS (ESI, *m/z*): calculated for C₄₇H₅₅F₂N₁₃O₁₄S₂ 1127.34; found 1128.4 [M+H]⁺. RP-HPLC: solvent A: CH₃CN 0.1% TFA, Solvent B: H₂O 0.1% TFA, gradient: from 10 to 60% of A in 15 min, isomer A, t_R = 9.9 min; isomer B, t_R = 10.2 min, flow: 1 mL/min, λ = 480nm.

4.2.2. Biology

In vitro APN inhibition assay.

IC₅₀ values against APN were determined using L-leucine-*p*-nitroanilide as substrate and microsomal aminopeptidase from porcine kidney (*p*APN, Sigma) (15 units/mg protein). The assay was performed in a 96-well plate in Phosphate Buffer Saline (PBS, KH₂PO₄ 1.47 mM, Na₂HPO₄ 7.8 mM, NaCl 137 mM, KCl 2.7 mM, CaCl₂ 1.8 mM, MgCl₂ 1.8 mM), pH 7.2, at 37 °C. Bestatin was used as a positive control, while biotin was used as a negative one. 0.5 µg/mL *p*APN and 250 µM L-leucine-*p*-nitroanilide were incubated in 100 µL of PBS, pH 7.2 in the presence of gradient concentrations of inhibitor (the concentration of peptides 1–6 and biotin ranged from 5 µM to 2.5 mM, while the concentration of bestatin from 50 nM to 25 µM). The assay mixture was incubated at 37°C for 1 h. Formation of the product *p*-nitroaniline was monitored by following the change in the absorbance measured at 405 nm with the UV-Vis spectrophotometer (Synergy HT, BioTek). The IC₅₀ value was defined as the concentration of each peptide that led to 50 % of maximal *p*APN catalytic activity.

Cell culture.

Human fibrosarcoma HT1080 cells were cultured in Dulbecco's Modified Essential Medium (Invitrogen) containing 9% fetal bovine serum (SIGMA) and incubated at 37°C in a humidified atmosphere containing 5% CO₂ and passed twice a week.

Inhibitory assay on HT1080 cells.

One day prior the experiment, HT1080 cells from a confluent plate were detached using trypsin/EDTA (Invitrogen) and plated at a 70% confluency in 96 well plate (approximately 60,000 cell/well). 24 hours later the cells were washed twice in PBS and incubated in 100 µl of PBS containing 250 µM APN substrate L-leucine-*p*-nitroanilide) and the different compounds at the indicated concentration. The cells were then incubated for 1.5 h at 37°C and absorbance at 405 measured every 10 minutes.

Binding curve of Fluorescent cCNGRCKG-OG 7 on HT1080 cells.

One day prior the experiment, HT1080 cells from a confluent plate were detached using trypsin/EDTA (Invitrogen) and plated at a 20% confluency in 96 well plate (approximately 20,000 cell/well). After 24 hours the cells were incubated for 60 min on ice with increasing concentration of compound **7**. After incubation the cells were washed three times in PBS and fixed in paraformaldehyde (3.7% in PBS) for 15 minutes, washed three times in PBS and imaged at 20 × magnification. For each field both a DIC and a fluorescent channel were collected.

Displacement of Fluorescent cCNGRCKG-OG 7 on HT1080 cells.

One day prior the experiment, HT1080 cells from a confluent plate were detached using trypsin/EDTA (Invitrogen) and plated at a 20% confluency on glass coverslips (approximately 20,000 cell/well). After 24 h the cells were incubated for 60 min on ice with compound **7** (40 μM) in absence or in presence of bestatin (100 μM) or compound **5** (1 mM). After incubation the cells were washed three times in PBS and fixed in paraformaldehyde (3.7% in PBS) for 15 minutes, washed three times in PBS and imaged at 40 × magnification. For each field both a DIC and a fluorescent channel were collected.

Quantification of fluorescence on cells.

The microscopic fields were acquired using the same settings and exposure time for all the images and we ensured that all the pixels in the images were in the dynamic range of the detector (saturated pixels < 0.1 %). The images were then analyzed using ImageJ open source software. The outline of the cells were manually traced using the DIC channel. The outlines of the cells were then superimposed to the fluorescent channel and the fluorescence of each single cell was measured. For each condition we measured at least 120 cells.

Abbreviations

Ac: acetyl
AcOH: acetic acid
ACN: acetonitrile
Acm: acetamidomethyl
APN, CD13: aminopeptidase N
Arg, R: arginine
Asn, N: asparagine
Boc: tert-butyloxycarbonyl
Boc₂O: di-*tert*-butyl dicarbonate
CD: circular dichroism
cCNGRC: cyclic Cys-Asn-Gly-Arg-Cys peptide
Cy 5.5: cyanine 5.5 fluorescent dye
Cys, C: cysteine
d: deuterium
DCM: dichloromethane
Dde: *N*-(1-(4,4-dimethyl-2,6-dioxocyclohexylidene)ethyl)
DIPEA: *N,N'*-diisopropylethylamine
DIC: *N,N'*-diisopropylcarbodiimide
DMEM: Dulbecco's Modified Eagle Medium
DMF: *N,N*-dimethylformamide
DMSO-*d*₆: hexadeuterodimethyl sulfoxide
ECM: extracellular matrix protein
Et₂O: diethyl ether
EtOAc: ethyl acetate
FBS: fetal bovine serum
Fmoc: fluorenylmethyloxycarbonyl
Fmoc- ϵ -Ahx-OH: Fmoc-6-aminohexanoic acid
FRET: fluorescence resonance energy transfer
G4: G-quadruplex
GIST: gastrointestinal stromal tumors
Gly, G: glycine
HAT: histone acetyltransferase
HATU: 2-(7-Aza-1*H*-benzotriazole-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate
HIF: hypoxia inducible factors
hTER: telomere reverse transcriptase
Lys, K: lysine
MRI: magnetic resonance imaging
Mmt: (4-methoxyphenyl)diphenylmethyl
NCS: *N*-chlorosuccinimide
NHE: nuclease hypersensitivity element
NMR: nuclear magnetic resonance
OG 488: Oregon Green, 5(6)-carboxy-2',7'-difluorofluorescein
Pbf: 2,2,4,6,7-pentamethyldihydrobenzofurane
r.t.: room temperature

THP: tumor-homing peptide
TFA: trifluoroacetic acid
Trt: trityl
TIPS: triisopropylsilane
VEGF: vascular endothelial growth factor

Bibliography

- (1) Siegel, R. L.; Miller, K. D.; Jemal, A. Cancer Statistics, 2016: Cancer Statistics, 2016. *CA. Cancer J. Clin.* **2016**, *66* (1), 7–30.
- (2) Ferdinand, R.; Mitchell, S. A.; Batson, S.; Tumor, I. Treatments for Chronic Myeloid Leukemia: A Qualitative Systematic Review. *J. Blood Med.* **2012**, *3*, 51–76.
- (3) Stratton, M. R.; Campbell, P. J.; Futreal, P. A. The Cancer Genome. *Nature* **2009**, *458* (7239), 719–724.
- (4) Talbot, S. J.; Crawford, D. H. Viruses and Tumours – an Update. *Eur. J. Cancer* **2004**, *40* (13), 1998–2005.
- (5) Olivier, M.; Hussain, S. P.; Caron de Fromentel, C.; Hainaut, P.; Harris, C. C. TP53 Mutation Spectra and Load: A Tool for Generating Hypotheses on the Etiology of Cancer. *IARC Sci. Publ.* **2004**, No. 157, 247–270.
- (6) Sadikovic, B.; Al-Romaih, K.; Squire, J. A.; Zielenska, M. Cause and Consequences of Genetic and Epigenetic Alterations in Human Cancer. *Curr. Genomics* **2008**, *9* (6), 394–408.
- (7) Chatterjee, A.; Mambo, E.; Sidransky, D. Mitochondrial DNA Mutations in Human Cancer. *Oncogene* **2006**, *25* (34), 4663–4674.
- (8) Hanahan, D.; Weinberg, R. A. Hallmarks of Cancer: The Next Generation. *Cell* **2011**, *144* (5), 646–674.
- (9) Hanahan, D.; Weinberg, R. A. The Hallmarks of Cancer. *Cell* **2000**, *100* (1), 57–70.
- (10) Cheng, N.; Chytil, A.; Shyr, Y.; Joly, A.; Moses, H. L. Transforming Growth Factor-Beta Signaling-Deficient Fibroblasts Enhance Hepatocyte Growth Factor Signaling in Mammary Carcinoma Cells to Promote Scattering and Invasion. *Mol. Cancer Res. MCR* **2008**, *6* (10), 1521–1533.
- (11) Lemmon, M. A.; Schlessinger, J. Cell Signaling by Receptor Tyrosine Kinases. *Cell* **2010**, *141* (7), 1117–1134.
- (12) Jiang, B.-H.; Liu, L.-Z. PI3K/PTEN Signaling in Angiogenesis and Tumorigenesis. *Adv. Cancer Res.* **2009**, *102*, 19–65.
- (13) Sherr, C. J.; McCormick, F. The RB and p53 Pathways in Cancer. *Cancer Cell* **2002**, *2* (2), 103–112.
- (14) Sigal, A.; Rotter, V. Oncogenic Mutations of the p53 Tumor Suppressor: The Demons of the Guardian of the Genome. *Cancer Res.* **2000**, *60* (24), 6788–6793.
- (15) Blasco, M. A. Telomeres and Human Disease: Ageing, Cancer and beyond. *Nat. Rev. Genet.* **2005**, *6* (8), 611–622.
- (16) Grivnenkov, S. I.; Greten, F. R.; Karin, M. Immunity, Inflammation, and Cancer. *Cell* **2010**, *140* (6), 883–899.
- (17) Yang, L.; Pang, Y.; Moses, H. L. TGF-Beta and Immune Cells: An Important Regulatory Axis in the Tumor Microenvironment and Progression. *Trends Immunol.* **2010**, *31* (6), 220–227.
- (18) Hanahan, D.; Folkman, J. Patterns and Emerging Mechanisms of the Angiogenic Switch during Tumorigenesis. *Cell* **1996**, *86* (3), 353–364.
- (19) DeBerardinis, R. J.; Lum, J. J.; Hatzivassiliou, G.; Thompson, C. B. The Biology of Cancer: Metabolic Reprogramming Fuels Cell Growth and Proliferation. *Cell Metab.* **2008**, *7* (1), 11–20.
- (20) Semenza, G. L. HIF-1: Upstream and Downstream of Cancer Metabolism. *Curr. Opin. Genet. Dev.* **2010**, *20* (1), 51–56.
- (21) Talmadge, J. E.; Fidler, I. J. AACR Centennial Series: The Biology of Cancer Metastasis: Historical Perspective. *Cancer Res.* **2010**, *70* (14), 5649–5669.
- (22) Hurley, L. H. DNA and Its Associated Processes as Targets for Cancer Therapy. *Nat. Rev. Cancer* **2002**, *2* (3), 188–200.

- (23) Gilman, A.; Philips, F. S. The Biological Actions and Therapeutic Applications of the B-Chloroethyl Amines and Sulfides. *Science* **1946**, *103* (2675), 409–436.
- (24) Goldacre, R. J.; Loveless, A.; Ross, W. C. J. Mode of Production of Chromosome Abnormalities by the Nitrogen Mustards; the Possible Role of Cross-Linking. *Nature* **1949**, *163* (4148), 667–669.
- (25) Bellon, S. F.; Coleman, J. H.; Lippard, S. J. DNA Unwinding Produced by Site-Specific Intrastrand Cross-Links of the Antitumor Drug Cis-diamminedichloroplatinum(II). *Biochemistry (Mosc.)* **1991**, *30* (32), 8026–8035.
- (26) Neidle, S. 3 The Molecular Basis for the Action of Some DNA-Binding Drugs. In *Progress in Medicinal Chemistry*; West, G. P. E. and G. B., Ed.; Elsevier, 1979; Vol. 16, pp 151–221.
- (27) Tewey, K. M.; Rowe, T. C.; Yang, L.; Halligan, B. D.; Liu, L. F. Adriamycin-Induced DNA Damage Mediated by Mammalian DNA Topoisomerase II. *Science* **1984**, *226* (4673), 466–468.
- (28) Druker, B. J. Translation of the Philadelphia Chromosome into Therapy for CML. *Blood* **2008**, *112* (13), 4808–4817.
- (29) Hambley, T. W.; Hait, W. N. Is Anticancer Drug Development Heading in the Right Direction? *Cancer Res.* **2009**, *69* (4), 1259–1262.
- (30) Dose, C.; Farkas, M. E.; Chenoweth, D. M.; Dervan, P. B. Next Generation Hairpin Polyamides with (R)-3,4-Diaminobutyric Acid Turn Unit. *J. Am. Chem. Soc.* **2008**, *130* (21), 6859–6866.
- (31) Han, H.; Hurley, L. H. G-Quadruplex DNA: A Potential Target for Anti-Cancer Drug Design. *Trends Pharmacol. Sci.* **2000**, *21* (4), 136–142.
- (32) Ohnmacht, S. A.; Neidle, S. Small-Molecule Quadruplex-Targeted Drug Discovery. *Bioorg. Med. Chem. Lett.* **2014**, *24* (12), 2602–2612.
- (33) Drygin, D.; Siddiqui-Jain, A.; O'Brien, S.; Schwaebe, M.; Lin, A.; Bliesath, J.; Ho, C. B.; Proffitt, C.; Trent, K.; Whitten, J. P.; Lim, J. K. C.; Hoff, D. V.; Anderes, K.; Rice, W. G. Anticancer Activity of CX-3543: A Direct Inhibitor of rRNA Biogenesis. *Cancer Res.* **2009**, *69* (19), 7653–7661.
- (34) Gautam, A.; Kapoor, P.; Chaudhary, K.; Kumar, R.; Open Source Drug Discovery Consortium; Raghava, G. P. S. Tumor Homing Peptides as Molecular Probes for Cancer Therapeutics, Diagnostics and Theranostics. *Curr Med Chem* **2014**, *21* (21), 2367–2391.
- (35) Prokopiou, E. M.; Ryder, S. A.; Walsh, J. J. Tumour Vasculature Targeting Agents in Hybrid/conjugate Drugs. *Angiogenesis* **2013**, *16* (3), 503–524.
- (36) Ruoslahti, E. Specialization of Tumour Vasculature. *Nat. Rev. Cancer* **2002**, *2* (2), 83–90.
- (37) Pasqualini, R.; Koivunen, E.; Ruoslahti, E. αv Integrins as Receptors for Tumor Targeting by Circulating Ligands. *Nat. Biotechnol.* **1997**, *15* (6), 542–546.
- (38) Arap, W.; Pasqualini, R.; Ruoslahti, E. Cancer Treatment by Targeted Drug Delivery to Tumor Vasculature in a Mouse Model. *Science* **1998**, *279* (5349), 377–380.
- (39) Mas-Moruno, C.; Rechenmacher, F.; Kessler, H. Cilengitide: The First Anti-Angiogenic Small Molecule Drug Candidate. Design, Synthesis and Clinical Evaluation. *Anticancer Agents Med Chem* **2010**, *10* (10), 753–768.
- (40) Beer, A. J.; Niemeyer, M.; Carlsen, J.; Sarbia, M.; Nährig, J.; Watzlowik, P.; Wester, H.-J.; Harbeck, N.; Schwaiger, M. Patterns of $\alpha v \beta 3$ Expression in Primary and Metastatic Human Breast Cancer as Shown by ^{18}F -Galacto-RGD PET. *J. Nucl. Med.* **2008**, *49* (2), 255–259.
- (41) Sparapani, S.; Bellini, S.; Gunaratnam, M.; Haider, S. M.; Andreani, A.; Rambaldi, M.; Locatelli, A.; Morigi, R.; Granaiola, M.; Varoli, L.; Burnelli, S.; Leoni, A.; Neidle, S. 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.
- (42) Gellert, M.; Lipsett, M. N.; Davies, D. R. Helix Formation by Guanylic Acid. *Proc. Natl. Acad. Sci. U. S. A.* **1962**, *48*, 2013–2018.

- (43) Davis, J. T. G-Quartets 40 Years Later: From 5'-GMP to Molecular Biology and Supramolecular Chemistry. *Angew. Chem. Int. Ed Engl.* **2004**, *43* (6), 668–698.
- (44) Burge, S.; Parkinson, G. N.; Hazel, P.; Todd, A. K.; Neidle, S. Quadruplex DNA: Sequence, Topology and Structure. *Nucleic Acids Res.* **2006**, *34* (19), 5402–5415.
- (45) Neidle, S. Quadruplex Nucleic Acids as Novel Therapeutic Targets. *J. Med. Chem.* **2016**, *59* (13), 5987–6011.
- (46) Zhou, J.; Bourdoncle, A.; Rosu, F.; Gabelica, V.; Mergny, J.-L. Tri-G-Quadruplex: Controlled Assembly of a G-Quadruplex Structure from Three G-Rich Strands. *Angew. Chem. Int. Ed.* **2012**, *51* (44), 11002–11005.
- (47) Parkinson, G. N.; Lee, M. P. H.; Neidle, S. Crystal Structure of Parallel Quadruplexes from Human Telomeric DNA. *Nature* **2002**, *417* (6891), 876–880.
- (48) Phan, A. T.; Patel, D. J. Two-Repeat Human Telomeric d(TAGGGTTAGGGT) Sequence Forms Interconverting Parallel and Antiparallel G-Quadruplexes in Solution: Distinct Topologies, Thermodynamic Properties, and Folding/unfolding Kinetics. *J. Am. Chem. Soc.* **2003**, *125* (49), 15021–15027.
- (49) Phan, A. T.; Modi, Y. S.; Patel, D. J. Two-Repeat Tetrahymena Telomeric d(TGGGGTTGGGGT) Sequence Interconverts Between Asymmetric Dimeric G-Quadruplexes in Solution. *J. Mol. Biol.* **2004**, *338* (1), 93–102.
- (50) Haider, S.; Parkinson, G. N.; Neidle, S. Crystal Structure of the Potassium Form of an Oxytricha Nova G-Quadruplex. *J. Mol. Biol.* **2002**, *320* (2), 189–200.
- (51) van Mourik, T.; Dingley, A. J. Characterization of the Monovalent Ion Position and Hydrogen-Bond Network in Guanine Quartets by DFT Calculations of NMR Parameters. *Chem. Weinh. Bergstr. Ger.* **2005**, *11* (20), 6064–6079.
- (52) Pagano, B.; Cosconati, S.; Gabelica, V.; Petraccone, L.; De Tito, S.; Marinelli, L.; La Pietra, V.; Leva, D.; Saverio, F.; Lauri, I.; Trotta, R.; Novellino, E.; Giancola, C.; Randazzo, A. State-of-the-Art Methodologies for the Discovery and Characterization of DNA G-Quadruplex Binders. *Curr. Pharm. Des.* **2012**, *18* (14).
- (53) Reddel, R. R. Telomere Maintenance Mechanisms in Cancer: Clinical Implications. *Curr. Pharm. Des.* **2014**, *20* (41), 6361–6374.
- (54) Qin, Y.; Hurley, L. H. Structures, Folding Patterns, and Functions of Intramolecular DNA G-Quadruplexes Found in Eukaryotic Promoter Regions. *Biochimie* **2008**, *90* (8), 1149–1171.
- (55) Kumari, S.; Bugaut, A.; Huppert, J. L.; Balasubramanian, S. An RNA G-Quadruplex in the 5' UTR of the NRAS Proto-Oncogene Modulates Translation. *Nat. Chem. Biol.* **2007**, *3* (4), 218–221.
- (56) Blackburn, E. H. Structure and Function of Telomeres. *Nature* **1991**, *350* (6319), 569–573.
- (57) Cech, null. Life at the End of the Chromosome: Telomeres and Telomerase. *Angew. Chem. Int. Ed Engl.* **2000**, *39* (1), 34–43.
- (58) Ambrus, A.; Chen, D.; Dai, J.; Bialis, T.; Jones, R. A.; Yang, D. Human Telomeric Sequence Forms a Hybrid-Type Intramolecular G-Quadruplex Structure with Mixed Parallel/antiparallel Strands in Potassium Solution. *Nucleic Acids Res.* **2006**, *34* (9), 2723–2735.
- (59) Dai, J.; Carver, M.; Punchihewa, C.; Jones, R. A.; Yang, D. Structure of the Hybrid-2 Type Intramolecular Human Telomeric G-Quadruplex in K⁺ Solution: Insights into Structure Polymorphism of the Human Telomeric Sequence. *Nucleic Acids Res.* **2007**, *35* (15), 4927–4940.
- (60) Luu, K. N.; Phan, A. T.; Kuryavyi, V.; Lacroix, L.; Patel, D. J. Structure of the Human Telomere in K⁺ Solution: An Intramolecular (3 + 1) G-Quadruplex Scaffold. *J. Am. Chem. Soc.* **2006**, *128* (30), 9963–9970.
- (61) Heddi, B.; Phan, A. T. Structure of Human Telomeric DNA in Crowded Solution. *J. Am. Chem. Soc.* **2011**, *133* (25), 9824–9833.

- (62) Renciuik, D.; Kejnovská, I.; Skoláková, P.; Bednářová, K.; Motlová, J.; Vorlíčková, M. Arrangements of Human Telomere DNA Quadruplex in Physiologically Relevant K^+ Solutions. *Nucleic Acids Res.* **2009**, *37* (19), 6625–6634.
- (63) Simonsson, T.; Kubista, M.; Pecinka, P. DNA Tetraplex Formation in the Control Region of c-Myc. *Nucleic Acids Res.* **1998**, *26* (5), 1167–1172.
- (64) 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 (Mosc.)* **2005**, *44* (6), 2048–2058.
- (65) Mathad, R. I.; Hatzakis, E.; Dai, J.; Yang, D. C-MYC Promoter G-Quadruplex Formed at the 5'-End of NHE III1 Element: Insights into Biological Relevance and Parallel-Stranded G-Quadruplex Stability. *Nucleic Acids Res.* **2011**, *39* (20), 9023–9033.
- (66) Dai, J.; Chen, D.; Jones, R. A.; Hurley, L. H.; Yang, D. NMR Solution Structure of the Major G-Quadruplex Structure Formed in the Human BCL2 Promoter Region. *Nucleic Acids Res.* **2006**, *34* (18), 5133–5144.
- (67) Agrawal, P.; Lin, C.; Mathad, R. I.; Carver, M.; Yang, D. The Major G-Quadruplex Formed in the Human BCL-2 Proximal Promoter Adopts a Parallel Structure with a 13-Nt Loop in K^+ Solution. *J. Am. Chem. Soc.* **2014**, *136* (5), 1750–1753.
- (68) Rankin, S.; Reszka, A. P.; Huppert, J.; Zloh, M.; Parkinson, G. N.; Todd, A. K.; Ladame, S.; Balasubramanian, S.; Neidle, S. Putative DNA Quadruplex Formation within the Human c-Kit Oncogene. *J. Am. Chem. Soc.* **2005**, *127* (30), 10584–10589.
- (69) Fernando, H.; Reszka, A. P.; Huppert, J.; Ladame, S.; Rankin, S.; Venkitaraman, A. R.; Neidle, S.; Balasubramanian, S. A Conserved Quadruplex Motif Located in a Transcription Activation Site of the Human c-Kit Oncogene. *Biochemistry (Mosc.)* **2006**, *45* (25), 7854–7860.
- (70) De Armond, R.; Wood, S.; Sun, D.; Hurley, L. H.; Ebbinghaus, S. W. Evidence for the Presence of a Guanine Quadruplex Forming Region within a Polypurine Tract of the Hypoxia Inducible Factor 1 α Promoter. *Biochemistry (Mosc.)* **2005**, *44* (49), 16341–16350.
- (71) Lombardo, C. M.; Welsh, S. J.; Strauss, S. J.; Dale, A. G.; Todd, A. K.; Nanjunda, R.; Wilson, W. D.; Neidle, S. A Novel Series of G-Quadruplex Ligands with Selectivity for HIF-Expressing Osteosarcoma and Renal Cancer Cell Lines. *Bioorg. Med. Chem. Lett.* **2012**, *22* (18), 5984–5988.
- (72) Tong, X.; Lan, W.; Zhang, X.; Wu, H.; Liu, M.; Cao, C. Solution Structure of All Parallel G-Quadruplex Formed by the Oncogene RET Promoter Sequence. *Nucleic Acids Res.* **2011**, *39* (15), 6753–6763.
- (73) Agrawal, P.; Hatzakis, E.; Guo, K.; Carver, M.; Yang, D. Solution Structure of the Major G-Quadruplex Formed in the Human VEGF Promoter in K^+ : Insights into Loop Interactions of the Parallel G-Quadruplexes. *Nucleic Acids Res.* **2013**, *41* (22), 10584–10592.
- (74) Wright, W. E.; Tesmer, V. M.; Huffman, K. E.; Levene, S. D.; Shay, J. W. Normal Human Chromosomes Have Long G-Rich Telomeric Overhangs at One End. *Genes Dev.* **1997**, *11* (21), 2801–2809.
- (75) van Steensel, B.; de Lange, T. Control of Telomere Length by the Human Telomeric Protein TRF1. *Nature* **1997**, *385* (6618), 740–743.
- (76) van Steensel, B.; Smogorzewska, A.; de Lange, T. TRF2 Protects Human Telomeres from End-to-End Fusions. *Cell* **1998**, *92* (3), 401–413.
- (77) Griffith, J. D.; Comeau, L.; Rosenfield, S.; Stansel, R. M.; Bianchi, A.; Moss, H.; de Lange, T. Mammalian Telomeres End in a Large Duplex Loop. *Cell* **1999**, *97* (4), 503–514.
- (78) Baumann, P.; Cech, T. R. Pot1, the Putative Telomere End-Binding Protein in Fission Yeast and Humans. *Science* **2001**, *292* (5519), 1171–1175.
- (79) Fry, M. Tetraplex DNA and Its Interacting Proteins. *Front. Biosci. J. Virtual Libr.* **2007**, *12*, 4336–4351.

- (80) Simonsson, T. G-Quadruplex DNA Structures Variations on a Theme. *Biol. Chem.* **2005**, *382* (4), 621–628.
- (81) Hayflick, L. The Limited in Vitro Lifetime of Human Diploid Cell Strains. *Exp. Cell Res.* **1965**, *37*, 614–636.
- (82) Kim, N. W.; Piatyszek, M. A.; Prowse, K. R.; Harley, C. B.; West, M. D.; Ho, P. L.; Coviello, G. M.; Wright, W. E.; Weinrich, S. L.; Shay, J. W. Specific Association of Human Telomerase Activity with Immortal Cells and Cancer. *Science* **1994**, *266* (5193), 2011–2015.
- (83) Feng, J.; Funk, W. D.; Wang, S. S.; Weinrich, S. L.; Avilion, A. A.; Chiu, C. P.; Adams, R. R.; Chang, E.; Allsopp, R. C.; Yu, J. The RNA Component of Human Telomerase. *Science* **1995**, *269* (5228), 1236–1241.
- (84) Shay, J. W.; Wright, W. E. Role of Telomeres and Telomerase in Cancer. *Semin. Cancer Biol.* **2011**, *21* (6), 349–353.
- (85) Hahn, W. C.; Counter, C. M.; Lundberg, A. S.; Beijersbergen, R. L.; Brooks, M. W.; Weinberg, R. A. Creation of Human Tumour Cells with Defined Genetic Elements. *Nature* **1999**, *400* (6743), 464–468.
- (86) Hemann, M. T.; Strong, M. A.; Hao, L. Y.; Greider, C. W. The Shortest Telomere, Not Average Telomere Length, Is Critical for Cell Viability and Chromosome Stability. *Cell* **2001**, *107* (1), 67–77.
- (87) 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.
- (88) Zahler, A. M.; Williamson, J. R.; Cech, T. R.; Prescott, D. M. Inhibition of Telomerase by G-Quartet DNA Structures. *Nature* **1991**, *350* (6320), 718–720.
- (89) Phatak, P.; Cookson, J. C.; Dai, F.; Smith, V.; Gartenhaus, R. B.; Stevens, M. F. G.; Burger, A. M. Telomere Uncapping by the G-Quadruplex Ligand RHPS4 Inhibits Clonogenic Tumour Cell Growth in Vitro and in Vivo Consistent with a Cancer Stem Cell Targeting Mechanism. *Br. J. Cancer* **2007**, *96* (8), 1223–1233.
- (90) 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.
- (91) 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.
- (92) Balasubramanian, S.; Neidle, S. G-Quadruplex Nucleic Acids as Therapeutic Targets. *Curr. Opin. Chem. Biol.* **2009**, *13* (3), 345–353.
- (93) Huppert, J. L.; Balasubramanian, S. G-Quadruplexes in Promoters throughout the Human Genome. *Nucleic Acids Res.* **2007**, *35* (2), 406–413.
- (94) Balasubramanian, S.; Hurley, L. H.; Neidle, S. Targeting G-Quadruplexes in Gene Promoters: A Novel Anticancer Strategy? *Nat. Rev. Drug Discov.* **2011**, *10* (4), 261–275.
- (95) Pelengaris, S.; Khan, M.; Evan, G. C-MYC: More than Just a Matter of Life and Death. *Nat. Rev. Cancer* **2002**, *2* (10), 764–776.
- (96) Wierstra, I.; Alves, J. The c-Myc Promoter: Still Mystery and Challenge. *Adv. Cancer Res.* **2008**, *99*, 113–333.
- (97) Gregory, M. A.; Hann, S. R. C-Myc Proteolysis by the Ubiquitin-Proteasome Pathway: Stabilization of c-Myc in Burkitt's Lymphoma Cells. *Mol. Cell. Biol.* **2000**, *20* (7), 2423–2435.
- (98) 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.* **2002**, *99* (18), 11593–11598.

- (99) Ou, T.-M.; Lu, Y.-J.; Zhang, C.; Huang, Z.-S.; Wang, X.-D.; Tan, J.-H.; Chen, Y.; Ma, D.-L.; Wong, K.-Y.; Tang, J. C.-O.; Chan, A. S.-C.; Gu, L.-Q. Stabilization of G-Quadruplex DNA and down-Regulation of Oncogene c-Myc by Quindoline Derivatives. *J. Med. Chem.* **2007**, *50* (7), 1465–1474.
- (100) Kang, H.-J.; Park, H.-J. Novel Molecular Mechanism for Actinomycin D Activity as an Oncogenic Promoter G-Quadruplex Binder. *Biochemistry (Mosc.)* **2009**, *48* (31), 7392–7398.
- (101) Yarden, Y.; Kuang, W. J.; Yang-Feng, T.; Coussens, L.; Munemitsu, S.; Dull, T. J.; Chen, E.; Schlessinger, J.; Francke, U.; Ullrich, A. Human Proto-Oncogene c-Kit: A New Cell Surface Receptor Tyrosine Kinase for an Unidentified Ligand. *EMBO J.* **1987**, *6* (11), 3341–3351.
- (102) Sakurai, S.; Fukasawa, T.; Chong, J. M.; Tanaka, A.; Fukayama, M. C-Kit Gene Abnormalities in Gastrointestinal Stromal Tumors (tumors of Interstitial Cells of Cajal). *Jpn. J. Cancer Res. Gann* **1999**, *90* (12), 1321–1328.
- (103) Gunaratnam, M.; Beltran, M.; Galesa, K.; Haider, S. M.; Reszka, A. P.; Cuenca, F.; Fletcher, J. A.; Neidle, S. Targeting Human Gastrointestinal Stromal Tumor Cells with a Quadruplex-Binding Small Molecule. **2009**, *52* (12), 3774–3783.
- (104) Fernando, H.; Reszka, A. P.; Huppert, J.; Ladame, S.; Rankin, S.; Venkitaraman, A. R.; Neidle, S.; Balasubramanian, S. A Conserved Quadruplex Motif Located in a Transcription Activation Site of the Human c-Kit Oncogene. *Biochemistry (Mosc.)* **2006**, *45* (25), 7854–7860.
- (105) Bejugam, M.; Sewitz, S.; Shirude, P. S.; Rodriguez, R.; Shahid, R.; Balasubramanian, S. Trisubstituted Isoalloxazines as a New Class of G-Quadruplex Binding Ligands: Small Molecule Regulation of c-Kit Oncogene Expression. *J. Am. Chem. Soc.* **2007**, *129* (43), 12926–12927.
- (106) Monchaud, D.; Teulade-Fichou, M.-P. A Hitchhiker's Guide to G-Quadruplex Ligands. *Org. Biomol. Chem.* **2008**, *6* (4), 627–636.
- (107) Yan, Y.; Tan, J.; Ou, T.; Huang, Z.; Gu, L. DNA G-Quadruplex Binders: A Patent Review. *Expert Opin. Ther. Pat.* **2013**, *23* (11), 1495–1509.
- (108) Gowan, S. M.; Heald, R.; Stevens, M. F.; 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.
- (109) Incles, C. M.; Schultes, C. M.; Kempfski, H.; Koehler, H.; Kelland, L. R.; Neidle, S. A G-Quadruplex Telomere Targeting Agent Produces p16-Associated Senescence and Chromosomal Fusions in Human Prostate Cancer Cells. *Mol. Cancer Ther.* **2004**, *3* (10), 1201–1206.
- (110) 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.
- (111) Bazzicalupi, C.; Ferraroni, M.; Bilia, A. R.; Scheggi, F.; Gratteri, P. The Crystal Structure of Human Telomeric DNA Complexed with Berberine: An Interesting Case of Stacked Ligand to G-Tetrad Ratio Higher than 1:1. *Nucleic Acids Res.* **2013**, *41* (1), 632–638.
- (112) Ma, Y.; Ou, T.-M.; Tan, J.-H.; Hou, J.-Q.; Huang, S.-L.; Gu, L.-Q.; Huang, Z.-S. Quinolino-Benzo-[5, 6]-Dihydroisoquinolium Compounds Derived from Berberine: A New Class of Highly Selective Ligands for G-Quadruplex DNA in c-Myc Oncogene. *Eur. J. Med. Chem.* **2011**, *46* (5), 1906–1913.
- (113) Micco, M.; Collie, G. W.; Dale, A. G.; Ohnmacht, S. A.; Pazitna, I.; Gunaratnam, M.; Reszka, A. P.; Neidle, S. Structure-Based Design and Evaluation of Naphthalene Diimide G-Quadruplex Ligands As Telomere Targeting Agents in Pancreatic Cancer Cells. *J. Med. Chem.* **2013**, *56* (7), 2959–2974.

- (114) 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.
- (115) Shin-ya, K.; Wierzba, K.; Matsuo, K.; Ohtani, T.; Yamada, Y.; Furihata, K.; Hayakawa, Y.; Seto, H. Telomestatin, a Novel Telomerase Inhibitor from *Streptomyces Anulatus*. *J. Am. Chem. Soc.* **2001**, *123* (6), 1262–1263.
- (116) Kim, M.-Y.; Vankayalapati, H.; Shin-Ya, K.; Wierzba, K.; Hurley, L. H. Telomestatin, a Potent Telomerase Inhibitor That Interacts Quite Specifically with the Human Telomeric Intramolecular G-Quadruplex. *J. Am. Chem. Soc.* **2002**, *124* (10), 2098–2099.
- (117) Miyazaki, T.; Pan, Y.; Joshi, K.; Purohit, D.; Hu, B.; Demir, H.; Mazumder, S.; Okabe, S.; Yamori, T.; Viapiano, M.; Shin-ya, K.; Seimiya, H.; Nakano, I. Telomestatin Impairs Glioma Stem Cell Survival and Growth through the Disruption of Telomeric G-Quadruplex and Inhibition of the Proto-Oncogene, c-Myb. *Clin. Cancer Res. Off. J. Am. Assoc. Cancer Res.* **2012**, *18* (5), 1268–1280.
- (118) Koirala, D.; Dhakal, S.; Ashbridge, B.; Sannohe, Y.; Rodriguez, R.; Sugiyama, H.; Balasubramanian, S.; Mao, H. A Single-Molecule Platform for Investigation of Interactions between G-Quadruplexes and Small-Molecule Ligands. *Nat. Chem.* **2011**, *3* (10), 782–787.
- (119) Rodriguez, R.; Miller, K. M.; Forment, J. V.; Bradshaw, C. R.; Nikan, M.; Britton, S.; Oelschlaegel, T.; Xhemalce, B.; Balasubramanian, S.; Jackson, S. P. Small-Molecule-Induced DNA Damage Identifies Alternative DNA Structures in Human Genes. *Nat. Chem. Biol.* **2012**, *8* (3), 301–310.
- (120) Amato, J.; Morigi, R.; Pagano, B.; Pagano, A.; Ohnmacht, S.; De Magis, A.; Tiang, Y.-P.; Capranico, G.; Locatelli, A.; Graziadio, A.; Leoni, A.; Rambaldi, M.; Novellino, E.; Neidle, S.; Randazzo, A. 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.
- (121) Andreani, A.; Granaiola, M.; Leoni, A.; Locatelli, A.; Morigi, R.; Rambaldi, M.; Giorgi, G.; Garaliene, V. Potential Antitumor Agents. 34.(1) Synthesis and Antitumor Activity of Guanylhydrazones from imidazo[2,1-*b*]thiazoles and from diimidazo[1,2-*α*: 1,2-*c*]Pyrimidine. *Anticancer Res.* **2004**, *24* (1), 203–211.
- (122) Andreani, A.; Rambaldi, M.; Locatelli, A.; Bossa, R.; Fraccari, A.; Galatulas, I. Potential Antitumor Agents. 21. Structure Determination and Antitumor Activity of imidazo[2,1-*b*]thiazole Guanylhydrazones. *J. Med. Chem.* **1992**, *35* (24), 4634–4637.
- (123) Andreani, A.; Granaiola, M.; Leoni, A.; Locatelli, A.; Morigi, R.; Rambaldi, M.; Giorgi, G.; Salvini, L. Cancer Fighting Cancer: Synthesis of the New Heterocyclic System diimidazo[1,2-*a*:1,2-*c*]pyrimidine. *Arkivoc* **2002**, *2002* (11), 32.
- (124) Russell, G. A., K. G. Reactions of Resonance Stabilized Carbanions. XXXI. Oxidation of Carbanions. 4. Oxidation of Indoxyl to Indigo in Basic Solution. *J. React. Reson. Stab. Carbanions XXXI Oxid. Carbanions 4 Oxid. Indoxyl Indigo Basic Solut.* **1969**, *91*, 3851–3859.
- (125) Dai, J.; Carver, M.; Yang, D. Polymorphism of Human Telomeric Quadruplex Structures. *Biochimie* **2008**, *90*, 1172.
- (126) Xue, Y.; Kan, Z.; Wang, Q.; Yao, Y.; Liu, J.; Hao, Y.; Tan, Z. Human Telomeric DNA Forms Parallel-Stranded Intramolecular G-Quadruplex in K⁺ Solution under Molecular Crowding Condition. *J. Am. Chem. Soc.* **2007**, *129* (36), 11185–11191.
- (127) 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.
- (128) Karsisiotis, A. I.; Hessari, N. M.; Novellino, E.; Spada, G. P.; Randazzo, A.; Webba da Silva, M. Topological Characterization of Nucleic Acid G-Quadruplexes by UV Absorption and Circular Dichroism. *Angew. Chem. Int. Ed Engl.* **2011**, *50* (45), 10645–10648.

- (129) Mergny, J. L.; Maurizot, J. C. Fluorescence Resonance Energy Transfer as a Probe for G-Quartet Formation by a Telomeric Repeat. *Chembiochem Eur. J. Chem. Biol.* **2001**, *2* (2), 124–132.
- (130) 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 (Mosc.)* **2005**, *44* (6), 2048–2058.
- (131) Kuryavyi, V.; Phan, A. T.; Patel, D. J. Solution Structures of All Parallel-Stranded Monomeric and Dimeric G-Quadruplex Scaffolds of the Human c-kit2 Promoter. *Nucleic Acids Res.* **2010**, *38* (19), 6757–6773.
- (132) Hsu, S.-T. D.; Varnai, P.; Bugaut, A.; Reszka, A. P.; Neidle, S.; Balasubramanian, S. A G-Rich Sequence within the c-Kit Oncogene Promoter Forms a Parallel G-Quadruplex Having Asymmetric G-Tetrad Dynamics. *J. Am. Chem. Soc.* **2009**, *131* (37), 13399–13409.
- (133) 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.
- (134) Carmeliet, P. Mechanisms of Angiogenesis and Arteriogenesis. *Nat. Med.* **2000**, *6* (4), 389–395.
- (135) Folkman, J. Tumor Angiogenesis: Therapeutic Implications. *N. Engl. J. Med.* **1971**, *285* (21), 1182–1186.
- (136) Dvorak, H. F. Tumors: Wounds That Do Not Heal. Similarities between Tumor Stroma Generation and Wound Healing. *N. Engl. J. Med.* **1986**, *315* (26), 1650–1659.
- (137) Bergers, G.; Benjamin, L. E. Tumorigenesis and the Angiogenic Switch. *Nat. Rev. Cancer* **2003**, *3* (6), 401–410.
- (138) Carmeliet, P.; Jain, R. K. Angiogenesis in Cancer and Other Diseases. *Nature* **2000**, *407* (6801), 249–257.
- (139) Ruoslahti, E. Vascular Zip Codes in Angiogenesis and Metastasis. *Biochem. Soc. Trans.* **2004**, *32* (Pt3), 397–402.
- (140) Korpelainen, E. I.; Alitalo, K. Signaling Angiogenesis and Lymphangiogenesis. *Curr. Opin. Cell Biol.* **1998**, *10* (2), 159–164.
- (141) Kim, S.; Bell, K.; Mousa, S. A.; Varner, J. A. Regulation of Angiogenesis in Vivo by Ligation of Integrin alpha5beta1 with the Central Cell-Binding Domain of Fibronectin. *Am. J. Pathol.* **2000**, *156* (4), 1345–1362.
- (142) Pasqualini, R.; Koivunen, E.; Kain, R.; Lahdenranta, J.; Sakamoto, M.; Stryhn, A.; Ashmun, R. A.; Shapiro, L. H.; Arap, W.; Ruoslahti, E. Aminopeptidase N Is a Receptor for Tumor-Homing Peptides and a Target for Inhibiting Angiogenesis. *Cancer Res.* **2000**, *60* (3), 722–727.
- (143) Ferrara, N. Vascular Endothelial Growth Factor: Molecular and Biological Aspects. *Curr. Top. Microbiol. Immunol.* **1999**, *237*, 1–30.
- (144) Brooks, P. C.; Clark, R. A.; Chersesh, D. A. Requirement of Vascular Integrin Alpha v Beta 3 for Angiogenesis. *Science* **1994**, *264* (5158), 569–571.
- (145) Bhagwat, S. V.; Lahdenranta, J.; Giordano, R.; Arap, W.; Pasqualini, R.; Shapiro, L. H. CD13/APN Is Activated by Angiogenic Signals and Is Essential for Capillary Tube Formation. *Blood* **2001**, *97* (3), 652–659.
- (146) Corti, A.; Pastorino, F.; Curnis, F.; Arap, W.; Ponzoni, M.; Pasqualini, R. Targeted Drug Delivery and Penetration into Solid Tumors. *Med. Res. Rev.* **2012**, *32* (5), 1078–1091.
- (147) Bábíčková, J.; Tóthová, L.; Boor, P.; Celec, P. In Vivo Phage Display — A Discovery Tool in Molecular Biomedicine. *Biotechnol. Adv.* **2013**, *31* (8), 1247–1259.
- (148) Koivunen, E.; Wang, B.; Ruoslahti, E. Isolation of a Highly Specific Ligand for the Alpha 5 Beta 1 Integrin from a Phage Display Library. *J. Cell Biol.* **1994**, *124* (3), 373–380.
- (149) Koivunen, E.; Gay, D. A.; Ruoslahti, E. Selection of Peptides Binding to the Alpha 5 Beta 1 Integrin from Phage Display Library. *J. Biol. Chem.* **1993**, *268* (27), 20205–20210.
- (150) Hooper, N. M. Families of Zinc Metalloproteases. *FEBS Lett.* **1994**, *354* (1), 1–6.

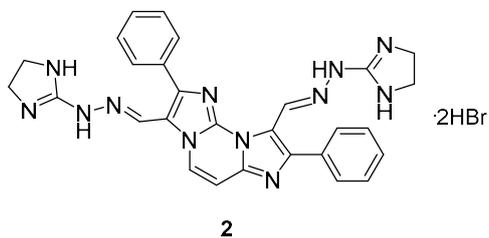
- (151) Olsen, J.; Kokholm, K.; Norén, O.; Sjöström, H. Structure and Expression of Aminopeptidase N. *Adv. Exp. Med. Biol.* **1997**, *421*, 47–57.
- (152) Mina-Osorio, P. The Moonlighting Enzyme CD13: Old and New Functions to Target. *Trends Mol. Med.* **2008**, *14* (8), 361–371.
- (153) Danziger, R. S. Aminopeptidase N in Arterial Hypertension. *Heart Fail. Rev.* **2008**, *13* (3), 293–298.
- (154) Nomura, R.; Kiyota, A.; Suzaki, E.; Kataoka, K.; Ohe, Y.; Miyamoto, K.; Senda, T.; Fujimoto, T. Human Coronavirus 229E Binds to CD13 in Rafts and Enters the Cell through Caveolae. *J. Virol.* **2004**, *78* (16), 8701–8708.
- (155) Mina-Osorio, P.; Shapiro, L. H.; Ortega, E. CD13 in Cell Adhesion: Aminopeptidase N (CD13) Mediates Homotypic Aggregation of Monocytic Cells. *J. Leukoc. Biol.* **2006**, *79* (4), 719–730.
- (156) Wickström, M.; Larsson, R.; Nygren, P.; Gullbo, J. Aminopeptidase N (CD13) as a Target for Cancer Chemotherapy. *Cancer Sci.* **2011**, *102* (3), 501–508.
- (157) Fujii, H.; Nakajima, M.; Saiki, I.; Yoneda, J.; Azuma, I.; Tsuruo, T. Human Melanoma Invasion and Metastasis Enhancement by High Expression of Aminopeptidase N/CD13. *Clin. Exp. Metastasis* **1995**, *13* (5), 337–344.
- (158) Kehlen, A.; Lendeckel, U.; Dralle, H.; Langner, J.; Hoang-Vu, C. Biological Significance of Aminopeptidase N/CD13 in Thyroid Carcinomas. *Cancer Res.* **2003**, *63* (23), 8500–8506.
- (159) Saiki, I.; Fujii, H.; Yoneda, J.; Abe, F.; Nakajima, M.; Tsuruo, T.; Azuma, I. Role of Aminopeptidase N (CD13) in Tumor-Cell Invasion and Extracellular Matrix Degradation. *Int. J. Cancer* **1993**, *54* (1), 137–143.
- (160) Chang, Y.-W.; Chen, S.-C.; Cheng, E.-C.; Ko, Y.-P.; Lin, Y.-C.; Kao, Y.-R.; Tsay, Y.-G.; Yang, P.-C.; Wu, C.-W.; Roffler, S. R. CD13 (aminopeptidase N) Can Associate with Tumor-Associated Antigen L6 and Enhance the Motility of Human Lung Cancer Cells. *Int. J. Cancer* **2005**, *116* (2), 243–252.
- (161) Mina-Osorio, P.; Winnicka, B.; O’Conor, C.; Grant, C. L.; Vogel, L. K.; Rodriguez-Pinto, D.; Holmes, K. V.; Ortega, E.; Shapiro, L. H. CD13 Is a Novel Mediator of Monocytic/endothelial Cell Adhesion. *J. Leukoc. Biol.* **2008**, *84* (2), 448–459.
- (162) Curnis, F.; Arrigoni, G.; Sacchi, A.; Fischetti, L.; Arap, W.; Pasqualini, R.; Corti, A. Differential Binding of Drugs Containing the NGR Motif to CD13 Isoforms in Tumor Vessels, Epithelia, and Myeloid Cells. *Cancer Res.* **2002**, *62* (3), 867–874.
- (163) Liu, C.; Yang, Y.; Chen, L.; Lin, Y.-L.; Li, F. A Unified Mechanism for Aminopeptidase N-Based Tumor Cell Motility and Tumor-Homing Therapy. *J. Biol. Chem.* **2014**, *289* (50), 34520–34529.
- (164) Anizon, F.; Boyle, F. T.; Fisher, J.; Kocienski, P. J. Synthesis and Characterisation of a Doxorubicin-CNGRC Peptide Conjugate That Targets Tumour Vasculature. *Synthesis* **2002**, *2002* (18), 2733–2736.
- (165) van Hensbergen, Y.; Broxterman, H. J.; Elderkamp, Y. W.; Lankelma, J.; Beers, J. C. C.; Heijn, M.; Boven, E.; Hoekman, K.; Pinedo, H. M. A Doxorubicin-CNGRC-Peptide Conjugate with Prodrug Properties. *Biochem. Pharmacol.* **2002**, *63* (5), 897–908.
- (166) Ndinguri, M. W.; Solipuram, R.; Gambrell, R. P.; Aggarwal, S.; Hammer, R. P. Peptide Targeting of Platinum Anti-Cancer Drugs. *Bioconjug. Chem.* **2009**, *20* (10), 1869–1878.
- (167) Pastorino, F.; Brignole, C.; Marimpietri, D.; Cilli, M.; Gambini, C.; Ribatti, D.; Longhi, R.; Allen, T. M.; Corti, A.; Ponzoni, M. Vascular Damage and Anti-Angiogenic Effects of Tumor Vessel-Targeted Liposomal Chemotherapy. *Cancer Res.* **2003**, *63* (21), 7400–7409.
- (168) Ellerby, H. M.; Arap, W.; Ellerby, L. M.; Kain, R.; Andrusiak, R.; Rio, G. D.; Krajewski, S.; Lombardo, C. R.; Rao, R.; Ruoslahti, E.; Bredesen, D. E.; Pasqualini, R. Anti-Cancer Activity of Targeted pro-Apoptotic Peptides. *Nat. Med.* **1999**, *5* (9), 1032–1038.

- (169) Curnis, F.; Sacchi, A.; Borgna, L.; Magni, F.; Gasparri, A.; Corti, A. Enhancement of Tumor Necrosis Factor Alpha Antitumor Immunotherapeutic Properties by Targeted Delivery to Aminopeptidase N (CD13). *Nat. Biotechnol.* **2000**, *18* (11), 1185–1190.
- (170) Colombo, G.; Curnis, F.; De Mori, G. M. S.; Gasparri, A.; Longoni, C.; Sacchi, A.; Longhi, R.; Corti, A. Structure-Activity Relationships of Linear and Cyclic Peptides Containing the NGR Tumor-Homing Motif. *J. Biol. Chem.* **2002**, *277* (49), 47891–47897.
- (171) Santoro, A.; Rimassa, L.; Sobrero, A. F.; Citterio, G.; Sclafani, F.; Carnaghi, C.; Pessino, A.; Caprioni, F.; Andretta, V.; Tronconi, M. C.; Finocchiaro, G.; Rossoni, G.; Zanoni, A.; Miggiano, C.; Rizzardi, G.-P.; Traversari, C.; Caligaris-Cappio, F.; Lambiase, A.; Bordignon, C. Phase II Study of NGR-hTNF, a Selective Vascular Targeting Agent, in Patients with Metastatic Colorectal Cancer after Failure of Standard Therapy. *Eur. J. Cancer Oxf. Engl. 1990* **2010**, *46* (15), 2746–2752.
- (172) Sacchi, A.; Gasparri, A.; Gallo-Stampino, C.; Toma, S.; Curnis, F.; Corti, A. Synergistic Antitumor Activity of Cisplatin, Paclitaxel, and Gemcitabine with Tumor Vasculature-Targeted Tumor Necrosis Factor-Alpha. *Clin. Cancer Res. Off. J. Am. Assoc. Cancer Res.* **2006**, *12* (1), 175–182.
- (173) ClinicalTrials.gov. NGR015: Randomized Double-Blind Phase III Study of NGR-hTNF Plus Best Investigator's Choice (BIC) Versus Placebo Plus BIC in Previously Treated Patients With Advanced Malignant Pleural Mesothelioma (MPM). *www.clinicaltrials.gov* **2017**.
- (174) Dirksen, A.; Langereis, S.; de Waal, B. F. M.; van Genderen, M. H. P.; Meijer, E. W.; de Lussanet, Q. G.; Hackeng, T. M. Design and Synthesis of a Bimodal Target-Specific Contrast Agent for Angiogenesis. *Org. Lett.* **2004**, *6* (26), 4857–4860.
- (175) Buehler, A.; van Zandvoort, M. A. M. J.; Stelt, B. J.; Hackeng, T. M.; Schrans-Stassen, B. H. G. J.; Bennaghmouch, A.; Hofstra, L.; Cleutjens, J. P. M.; Duijvestijn, A.; Smeets, M. B.; de Kleijn, D. P. V.; Post, M. J.; de Muinck, E. D. cNGR: A Novel Homing Sequence for CD13/APN Targeted Molecular Imaging of Murine Cardiac Angiogenesis in Vivo. *Arterioscler. Thromb. Vasc. Biol.* **2006**, *26* (12), 2681–2687.
- (176) von Wallbrunn, A.; Waldeck, J.; Höltnke, C.; Zühlendorf, M.; Mesters, R.; Heindel, W.; Schäfers, M.; Bremer, C. In Vivo Optical Imaging of CD13/APN-Expression in Tumor Xenografts. *J. Biomed. Opt.* **2008**, *13* (1), 011007.
- (177) Chen, L.; Lin, Y.-L.; Peng, G.; Li, F. Structural Basis for Multifunctional Roles of Mammalian Aminopeptidase N. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109* (44), 17966–17971.
- (178) Chen, K.; Ma, W.; Li, G.; Wang, J.; Yang, W.; Yap, L.-P.; Hughes, L. D.; Park, R.; Conti, P. S. Synthesis and Evaluation of ⁶⁴Cu-Labeled Monomeric and Dimeric NGR Peptides for MicroPET Imaging of CD13 Receptor Expression. *Mol. Pharm.* **2013**, *10* (1), 417–427.
- (179) Wei, Y.; Yin, G.; Yin, H.; Yan, D.; Ma, C.; Huang, Z.; Liao, X.; Yao, Y.; Chen, X. Inhibiting Effects of a Cyclic Peptide CNGRC on Proliferation and Migration of Tumor Cells In Vitro. *Int. J. Pept. Res. Ther.* **2013**, *19* (2), 163–173.
- (180) Graziadio, A.; Zanda, M.; Frau, S.; Fleming, I. N.; Musolino, M.; Dall'Angelo, S.; Baldassarre, M.; Piras, M. NGR Tumor-Homing Peptides: Structural Requirements for Effective APN (CD13) Targeting. *Bioconjug. Chem.* **2016**, *27* (5), 1332–1340.
- (181) Chan, W. C.; White, P. D. *Fmoc Solid Phase Peptide Synthesis A Practical Approach*, 2nd edn. Series editor Hames, B.D. Oxford University Press.; 2004.
- (182) Bodanszky, M. Side Reactions in Peptide Synthesis. In *Principles of Peptide Synthesis; Reactivity and Structure Concepts in Organic Chemistry*; Springer Berlin Heidelberg, 1984; pp 158–201.
- (183) Kaiser, E.; Colecott, R. L.; Bossinger, C. D.; Cook, P. I. Color Test for Detection of Free Terminal Amino Groups in the Solid-Phase Synthesis of Peptides. *Anal. Biochem.* **1970**, *34* (2), 595–598.

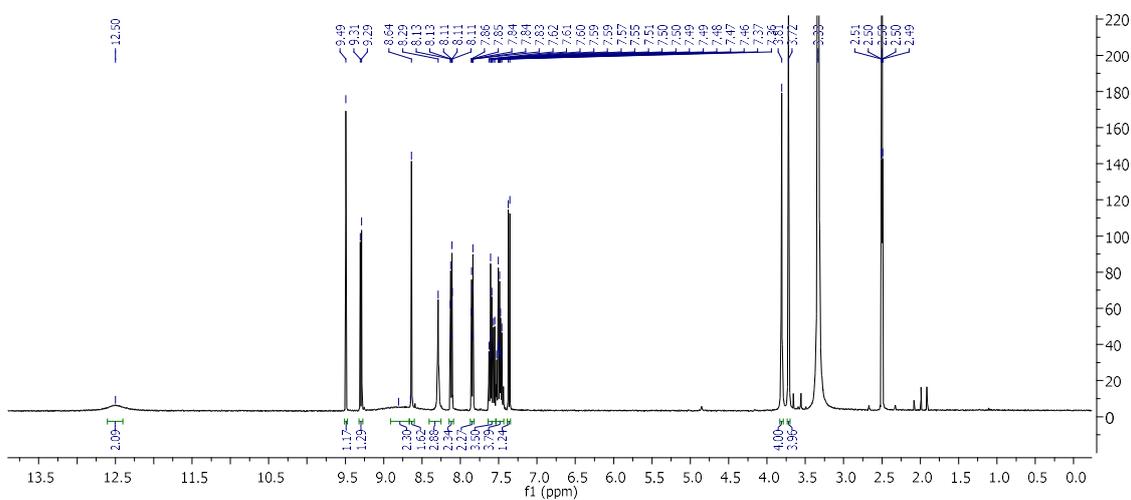
- (184) Zhang, S.; Lin, F.; Hossain, M. A.; Shabanpoor, F.; Tregear, G. W.; Wade, J. D. Simultaneous Post-cysteine(S-Acm) Group Removal Quenching of Iodine and Isolation of Peptide by One Step Ether Precipitation. *Int. J. Pept. Res. Ther.* **2008**, *14* (4), 301–305.
- (185) Tian, X.; Wickstrom, E. Continuous Solid-Phase Synthesis and Disulfide Cyclization of Peptide–PNA–Peptide Chimeras. *Org. Lett.* **2002**, *4* (23), 4013–4016.
- (186) Song, A.; Wang, X.; Zhang, J.; Marik, J.; Lebrilla, C. B.; Lam, K. S. Synthesis of Hydrophilic and Flexible Linkers for Peptide Derivatization in Solid Phase. *Bioorg. Med. Chem. Lett.* **2004**, *14* (1), 161–165.
- (187) Postma, T. M.; Albericio, F. N-Chlorosuccinimide, an Efficient Peptide Disulfide Bond-Forming Reagent in Aqueous Solution. *RSC Adv.* **2013**, *3* (34), 14277.
- (188) Bycroft, B. W.; Chan, W. C.; Chhabra, S. R.; Hone, N. D. A Novel Lysine-Protecting Procedure for Continuous Flow Solid Phase Synthesis of Branched Peptides. *J. Chem. Soc., Chem. Commun.* **1993**, No. 9, 778–779.
- (189) Zee-Cheng, K. Y.; Cheng, C. C. Experimental Antileukemic Agents. Preparation and Structure-Activity Study of S-Tritylcysteine and Related Compounds. *J. Med. Chem.* **1970**, *13* (3), 414–418.
- (190) Sun, W.-C.; Gee, K. R.; Klaubert, D. H.; Haugland, R. P. Synthesis of Fluorinated Fluoresceins. *J. Org. Chem.* **1997**, *62* (19), 6469–6475.
- (191) Plesniak, L. A.; Salzameda, B.; Hinderberger, H.; Regan, E.; Kahn, J.; Mills, S. A.; Teriete, P.; Yao, Y.; Jennings, P.; Marassi, F.; Adams, J. A. Structure and Activity of CPNGRC: A Modified CD13/APN Peptidic Homing Motif. *Chem. Biol. Drug Des.* **2010**, *75* (6), 551–562.
- (192) Su, L.; Cao, J.; Jia, Y.; Zhang, X.; Fang, H.; Xu, W. Development of Synthetic Aminopeptidase N/CD13 Inhibitors to Overcome Cancer Metastasis and Angiogenesis. *ACS Med. Chem. Lett.* **2012**, *3* (12), 959–964.
- (193) Miki, T.; Takegami, Y.; Okawa, K.; Muraguchi, T.; Noda, M.; Takahashi, C. The Reversion-Inducing Cysteine-Rich Protein with Kazal Motifs (RECK) Interacts with Membrane Type 1 Matrix Metalloproteinase and CD13/aminopeptidase N and Modulates Their Endocytic Pathways. *J. Biol. Chem.* **2007**, *282* (16), 12341–12352.
- (194) Yeager, C. L.; Ashmun, R. A.; Williams, R. K.; Cardellicchio, C. B.; Shapiro, L. H.; Look, A. T.; Holmes, K. V. Human Aminopeptidase N Is a Receptor for Human Coronavirus 229E. *Nature* **1992**, *357* (6377), 420–422.
- (195) Kipnis, F.; Soloway, H.; Ornfelt, J. 2-Acyloxyacetylthiophenes. *J Am Chem Soc* **1949**, *71*, 10–11.
- (196) Daszkiewicz, Z.; Domanski, A.; Kyziol, J. B. A Simple Route to N-Methylarylamines. *Chem Pap* **1993**, *47*, 109–113.
- (197) Hodges, R.; Shannon, J. S.; Jamieson, W. D.; Taylor, A. Chemical and Biological Properties of Some Oxindol-3-Ylidene Methines. *Can J Chem* **1968**, *46*, 2189–2194.
- (198) Kisteneva, M. S. Azomethine Dyes from Oxindole Derivatives. *Zh Obshch Khim* **1956**, *26*, 1169–1175.
- (199) Pagano, B.; Amato, J.; Iaccarino, N.; Cingolani, C.; Zizza, P.; Biroccio, A.; Novellino, E.; Randazzo, A. Looking for Efficient G-Quadruplex Ligands: Evidence for Selective Stabilizing Properties and Telomere Damage by Drug-like Molecules. *ChemMedChem* **2015**, *10* (4), 640–649.

Appendix

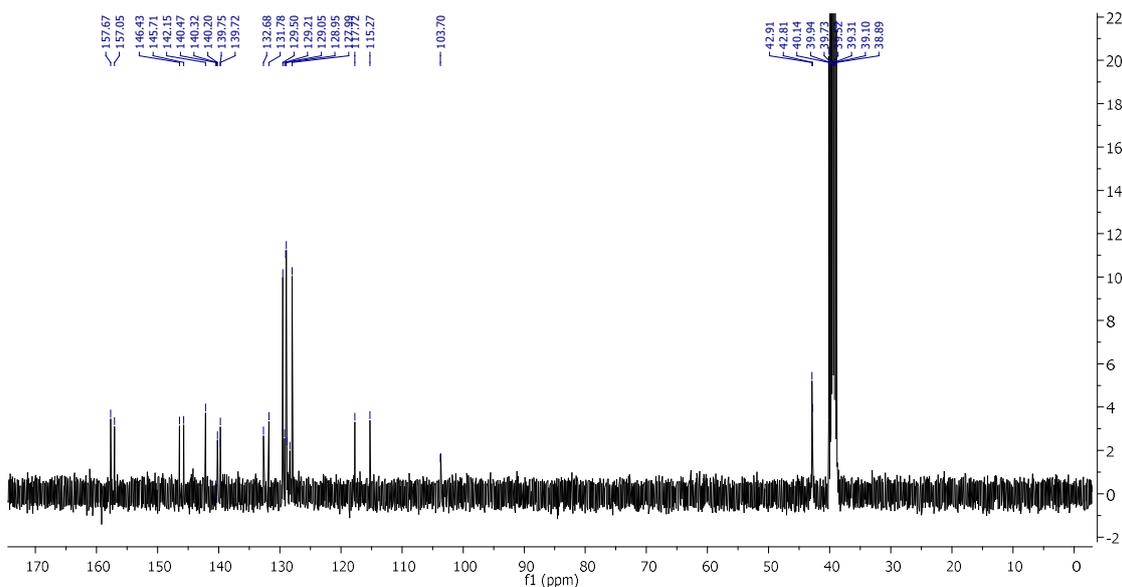
NMR spectra of compounds described in Chapter 2

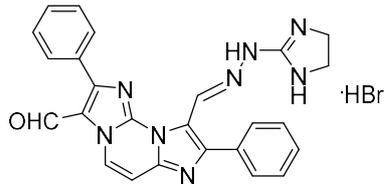


^1H NMR DMSO- d_6



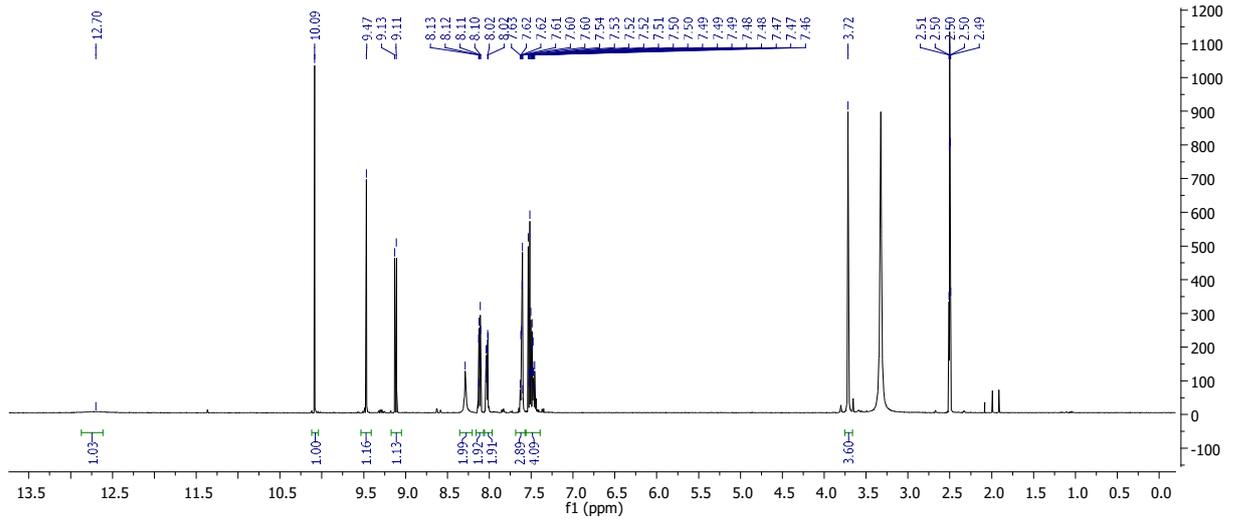
^{13}C NMR DMSO- d_6



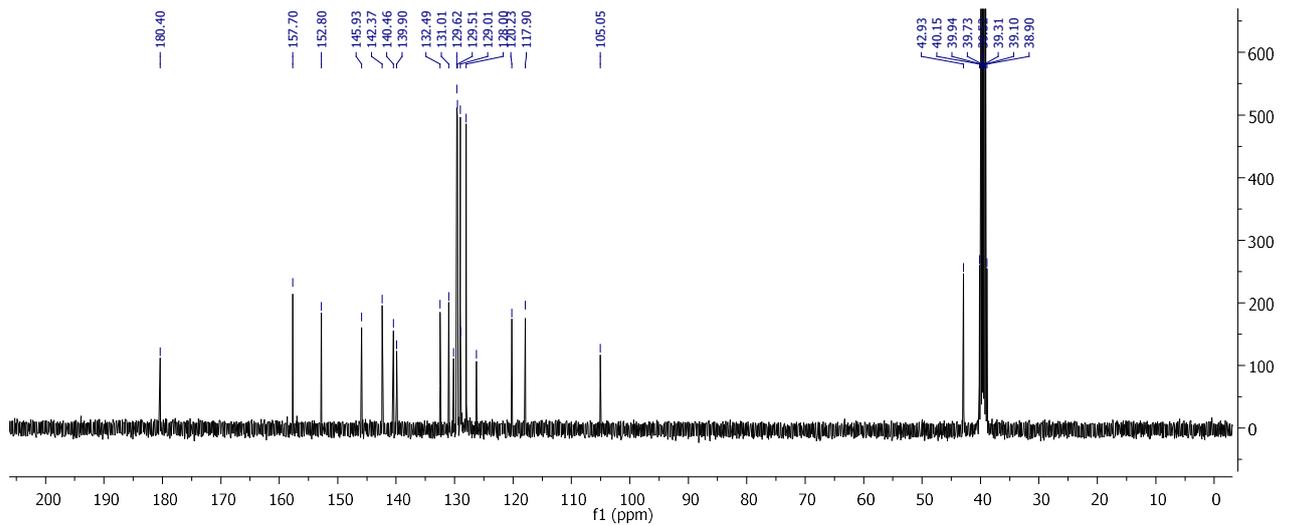


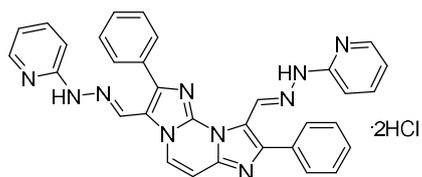
3

^1H NMR DMSO- d_6



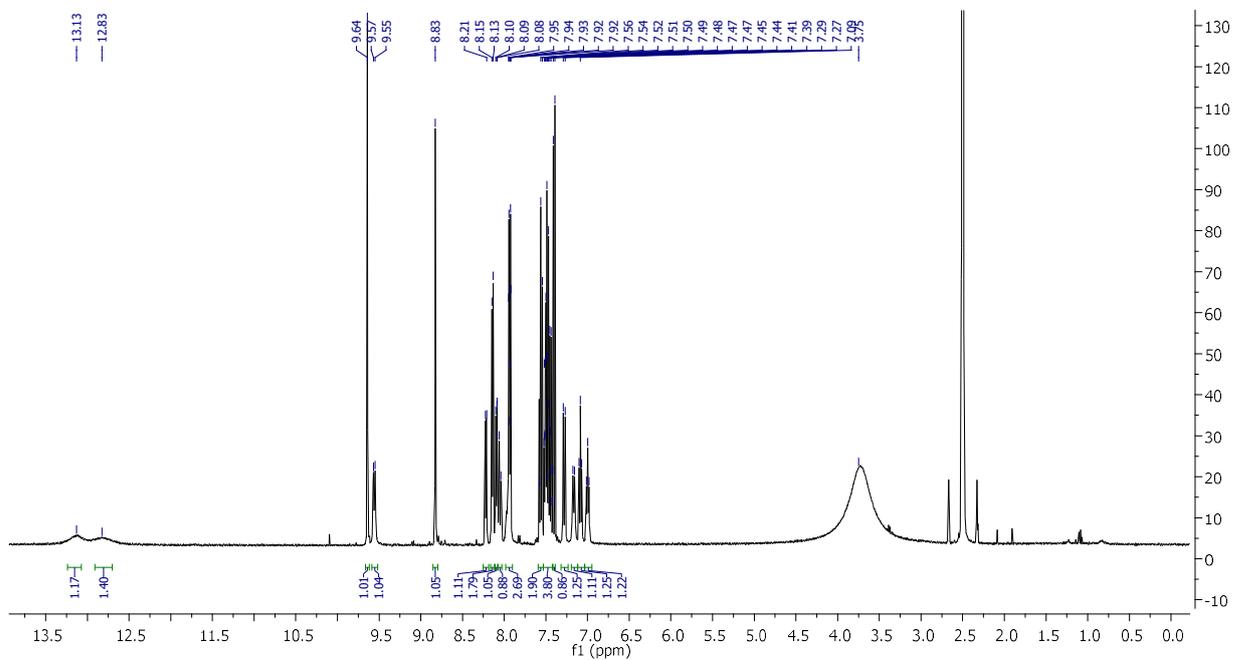
^{13}C NMR DMSO- d_6



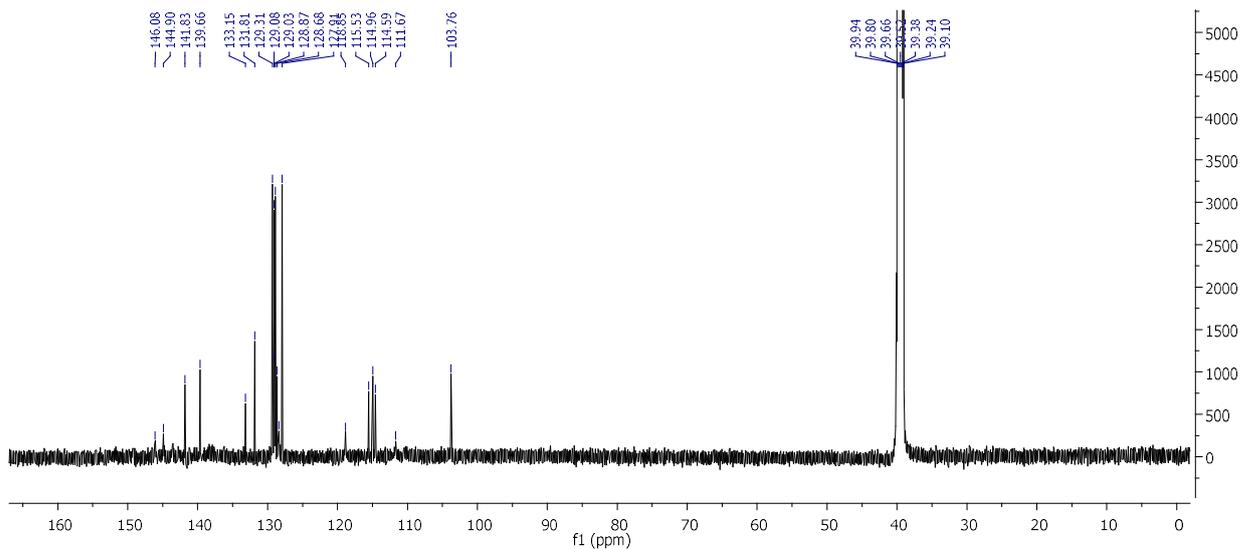


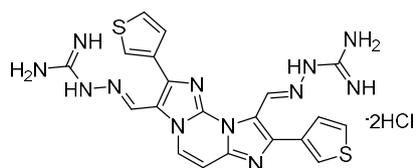
4

$^1\text{H NMR DMSO-}d_6$



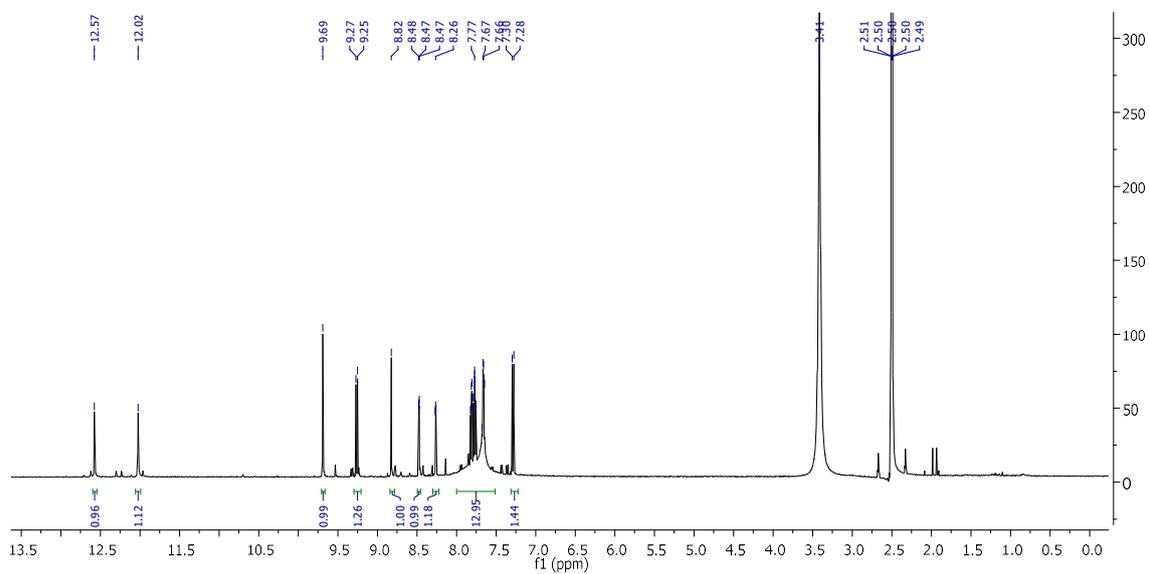
$^{13}\text{C NMR DMSO-}d_6$



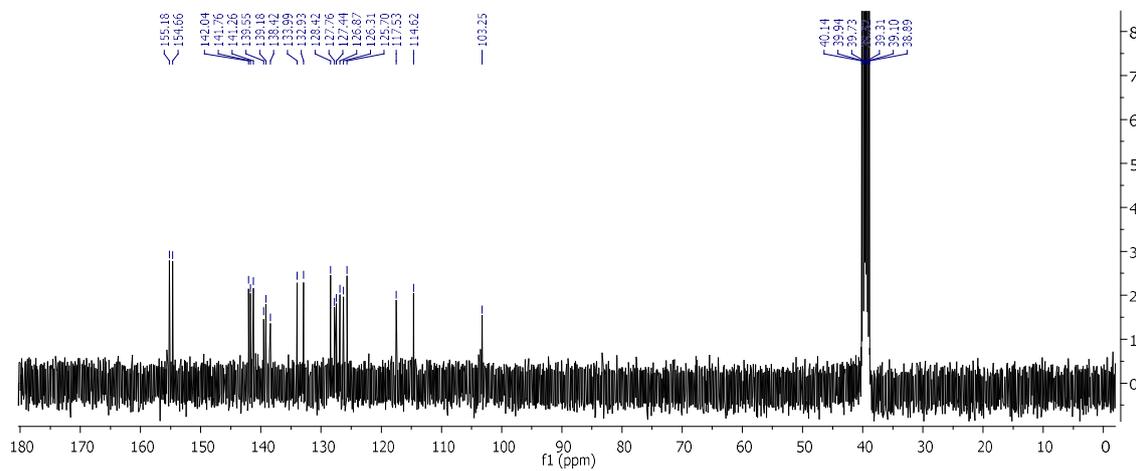


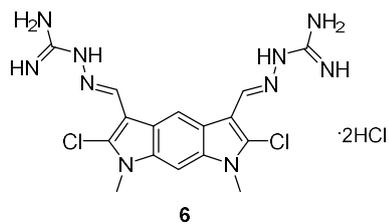
5

$^1\text{H NMR DMSO-}d_6$

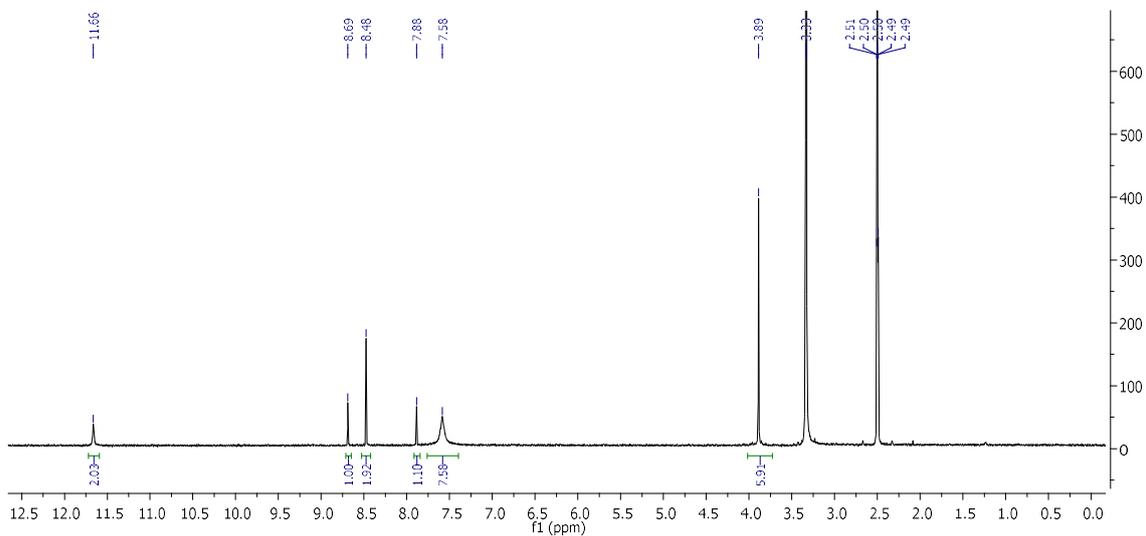


$^{13}\text{C NMR DMSO-}d_6$

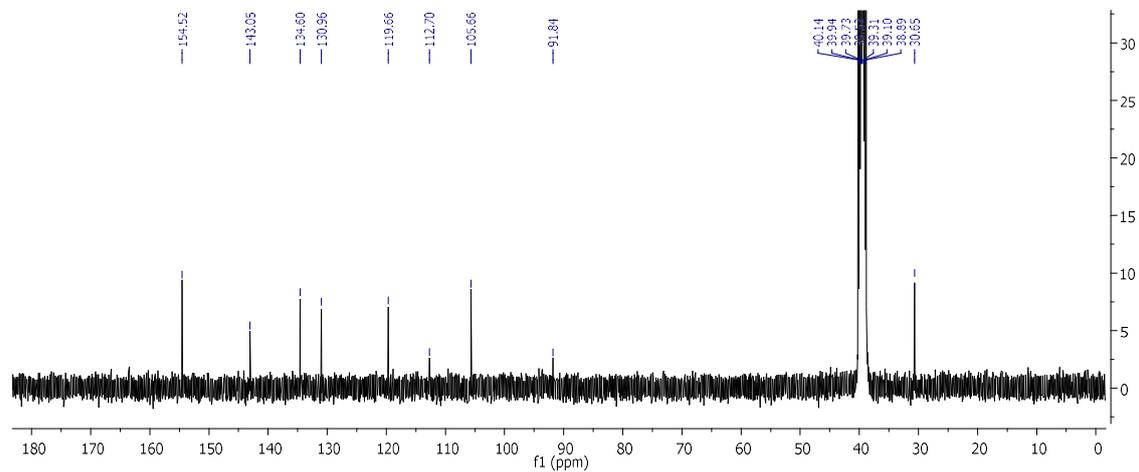


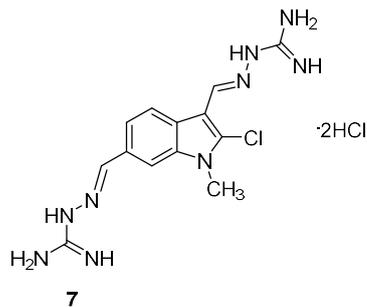


¹H NMR DMSO-*d*₆

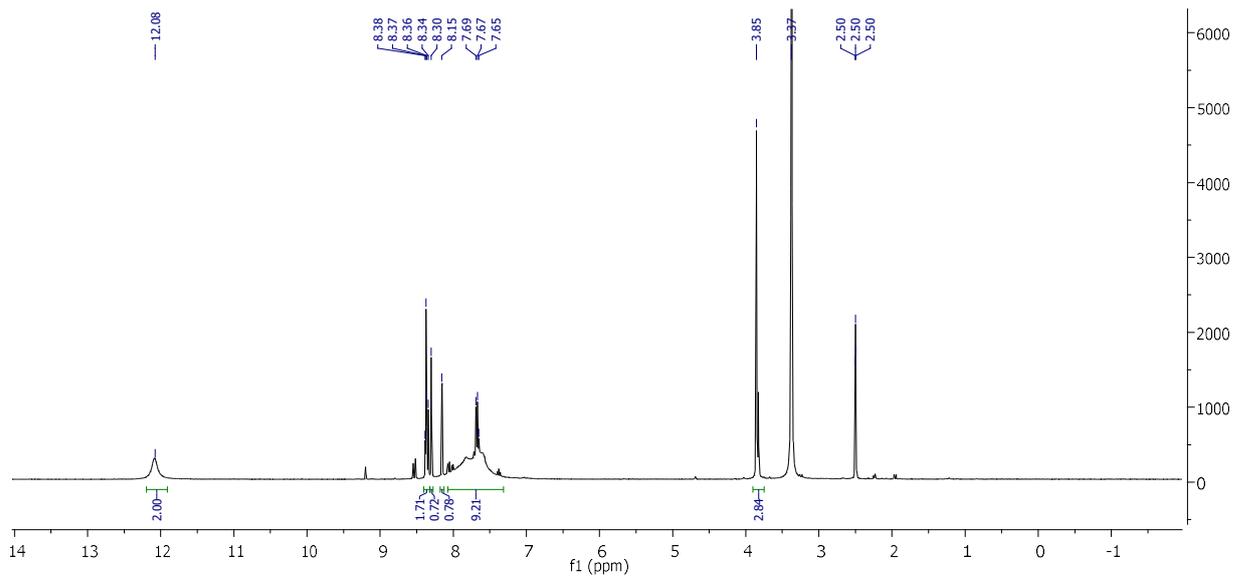


¹³C NMR DMSO-*d*₆

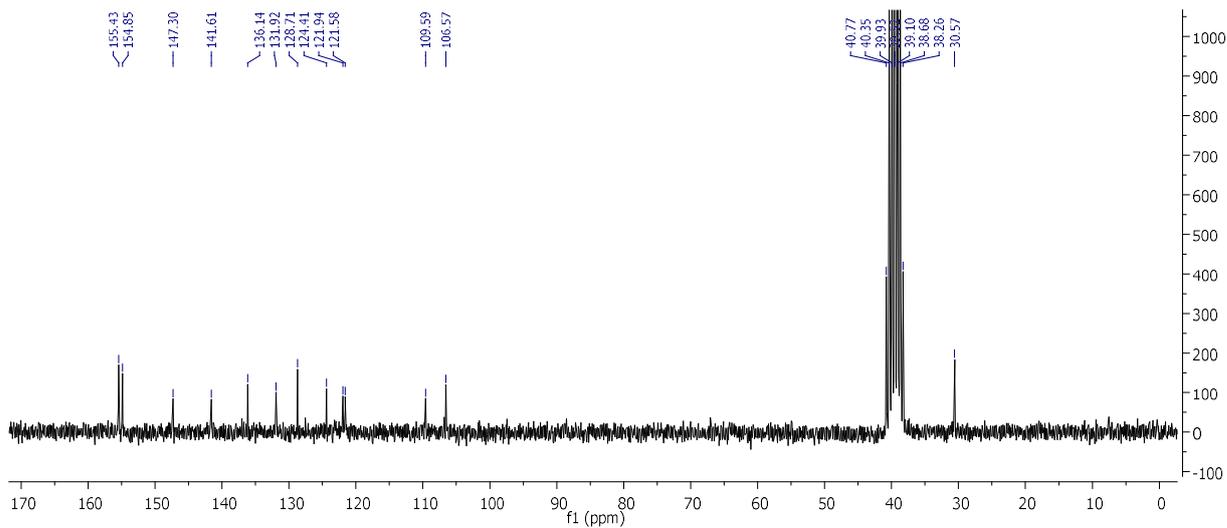


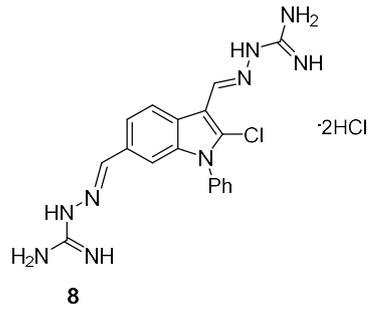


¹H NMR DMSO-d₆

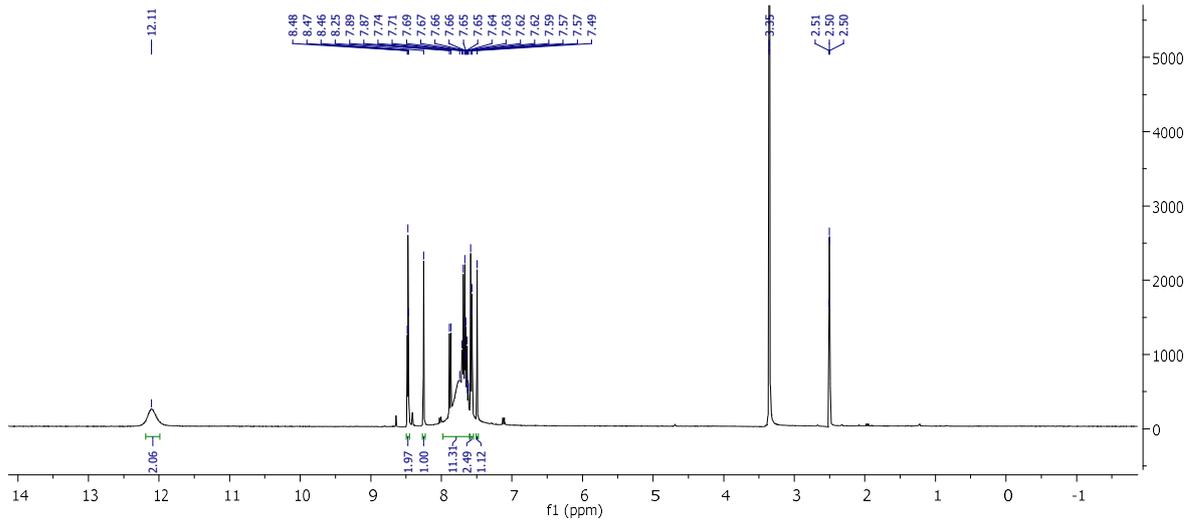


¹³C NMR DMSO-d₆

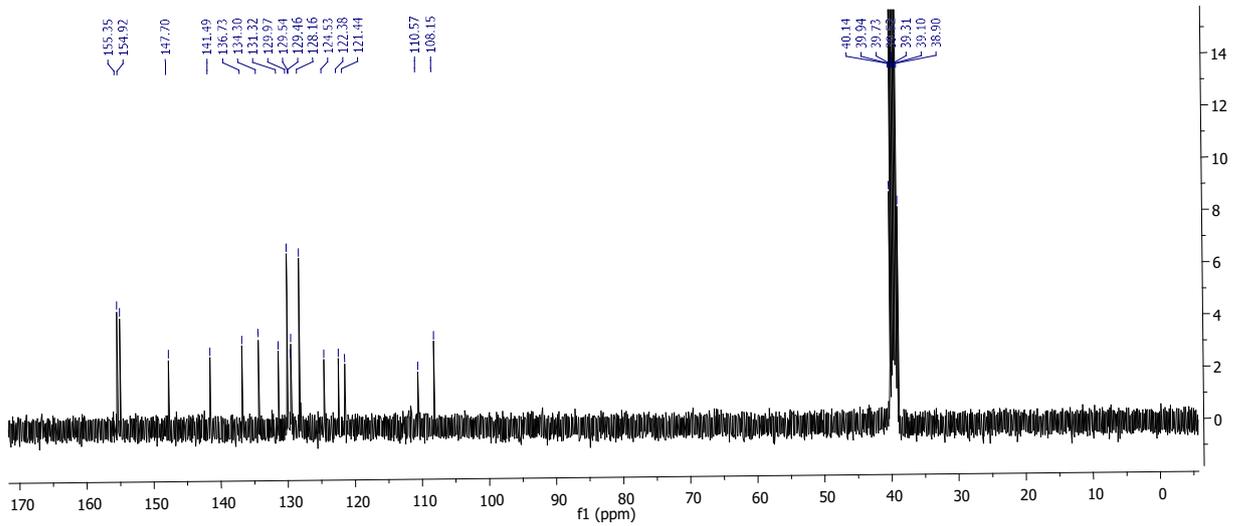


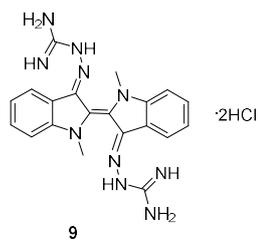


¹H NMR DMSO-*d*₆

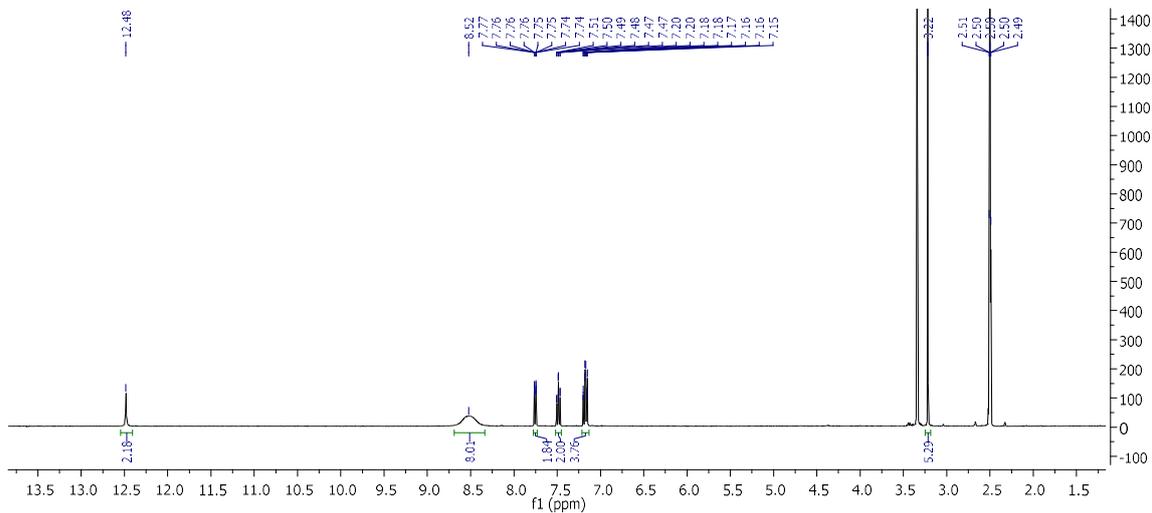


¹³C NMR DMSO-*d*₆

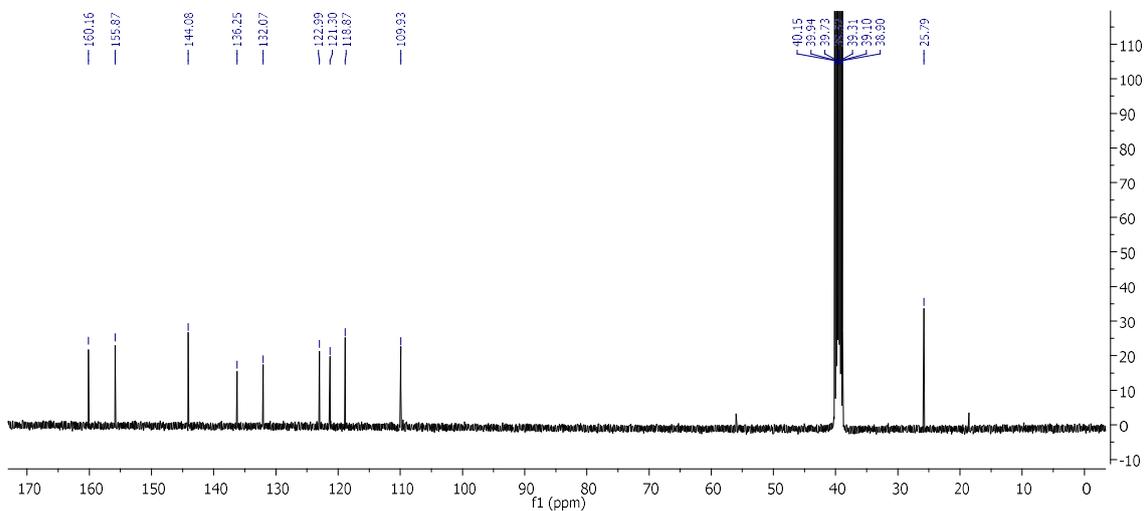


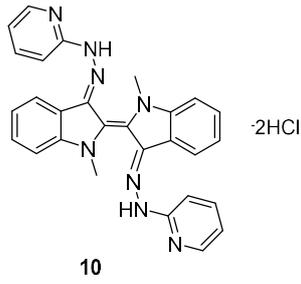


¹H NMR DMSO-*d*₆

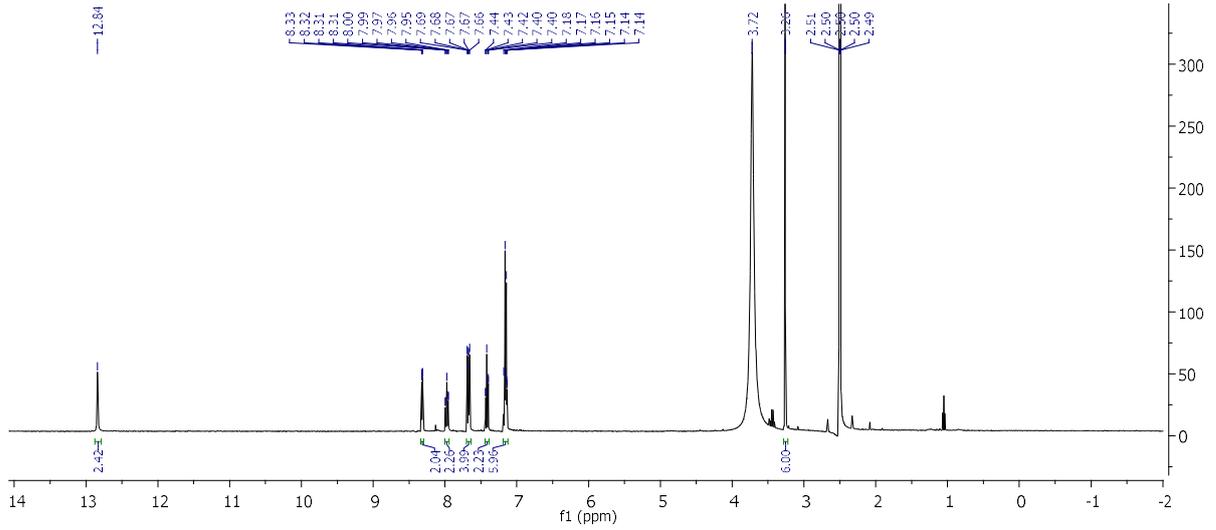


¹³C NMR DMSO-*d*₆

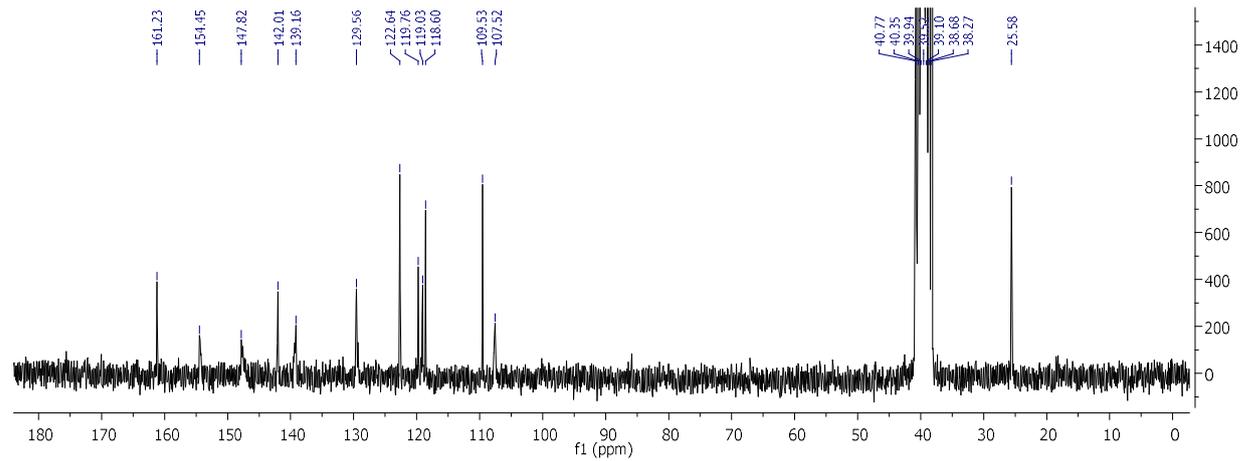




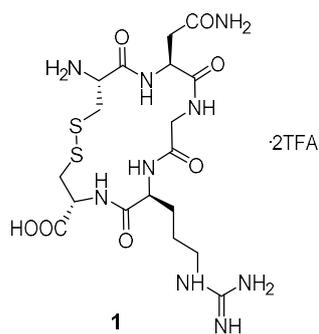
¹H NMR DMSO-*d*₆



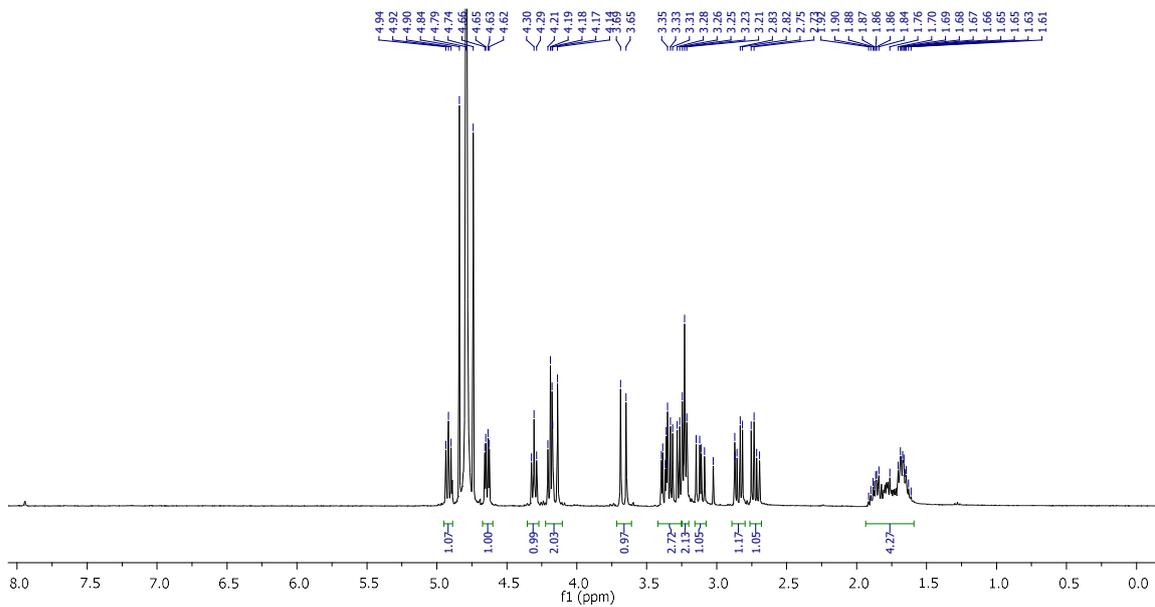
¹³C NMR DMSO-*d*₆



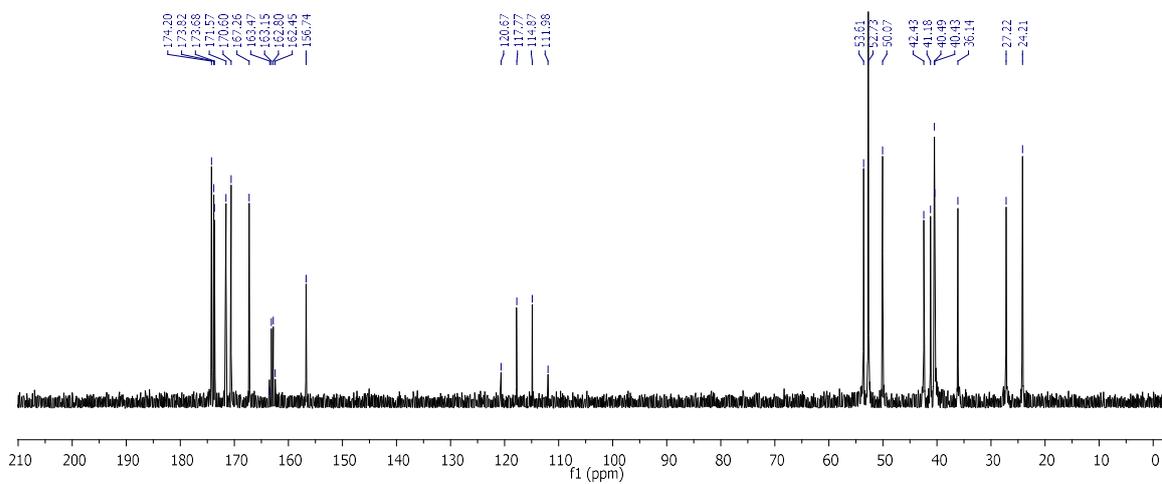
NMR spectra of compounds described in Chapter 3

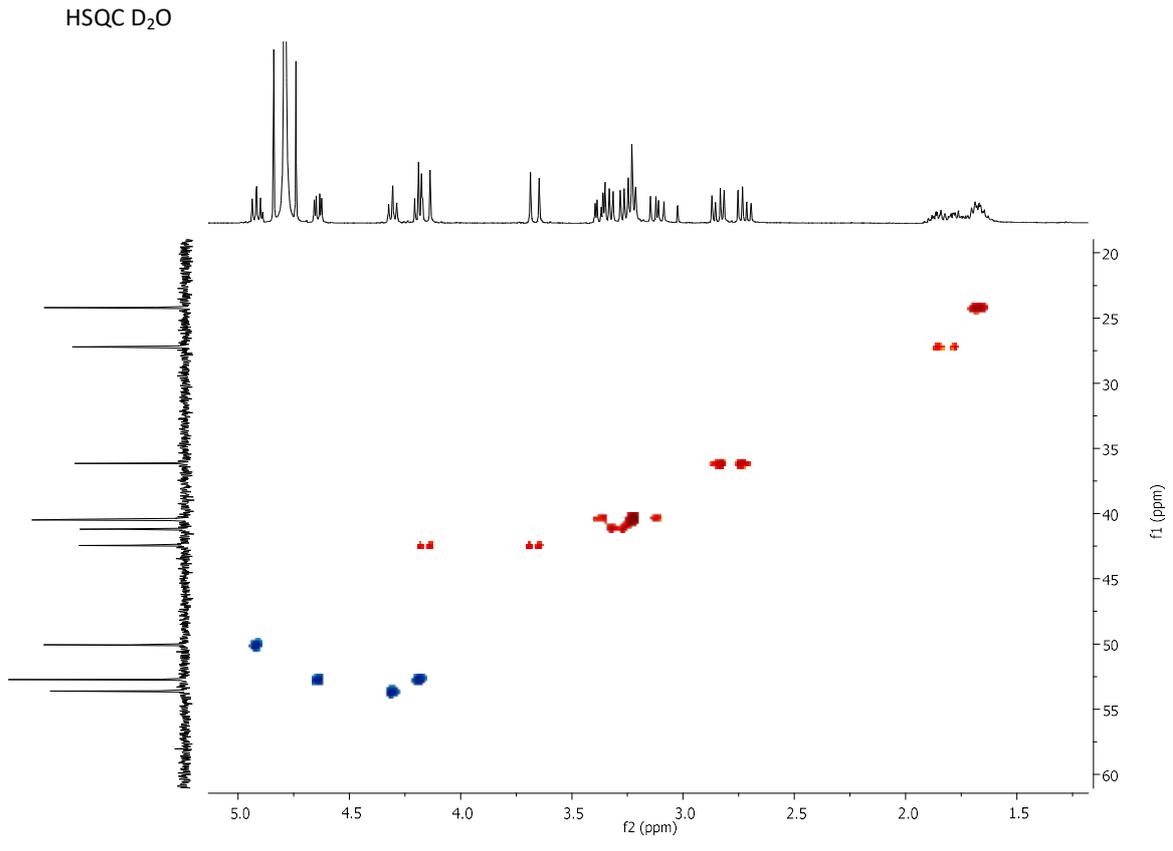
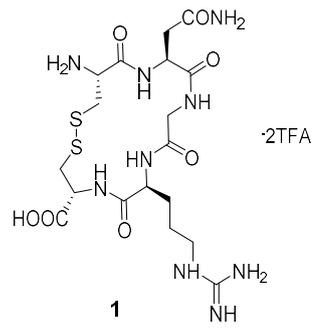


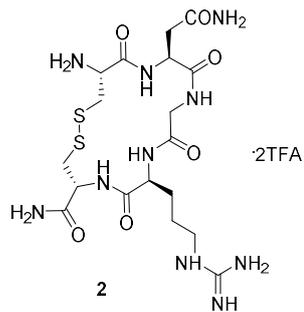
^1H NMR D_2O



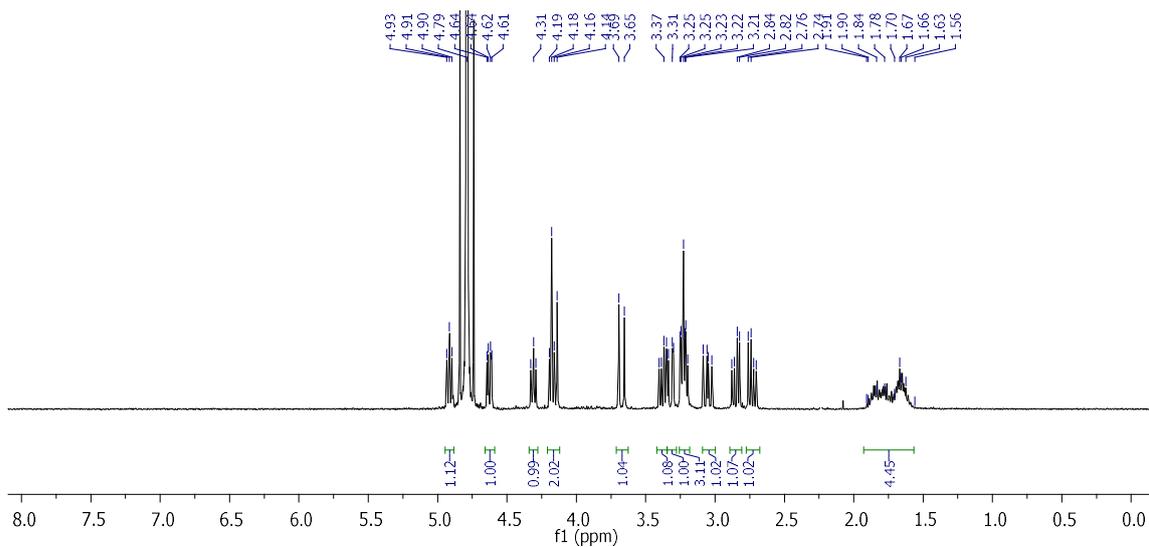
^{13}C NMR D_2O



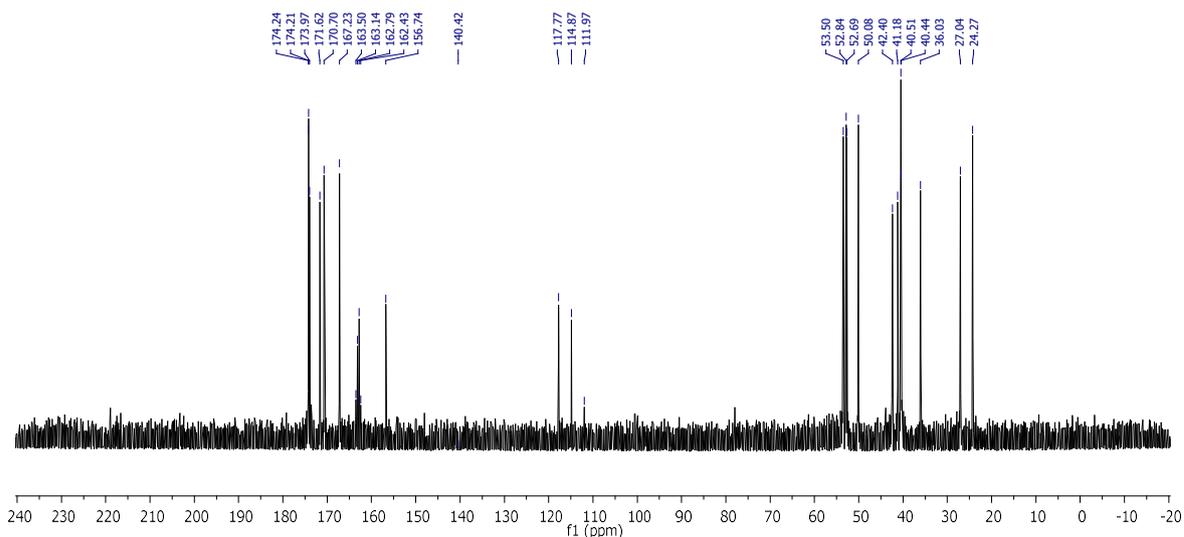


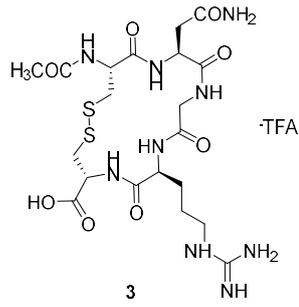


¹H NMR D₂O

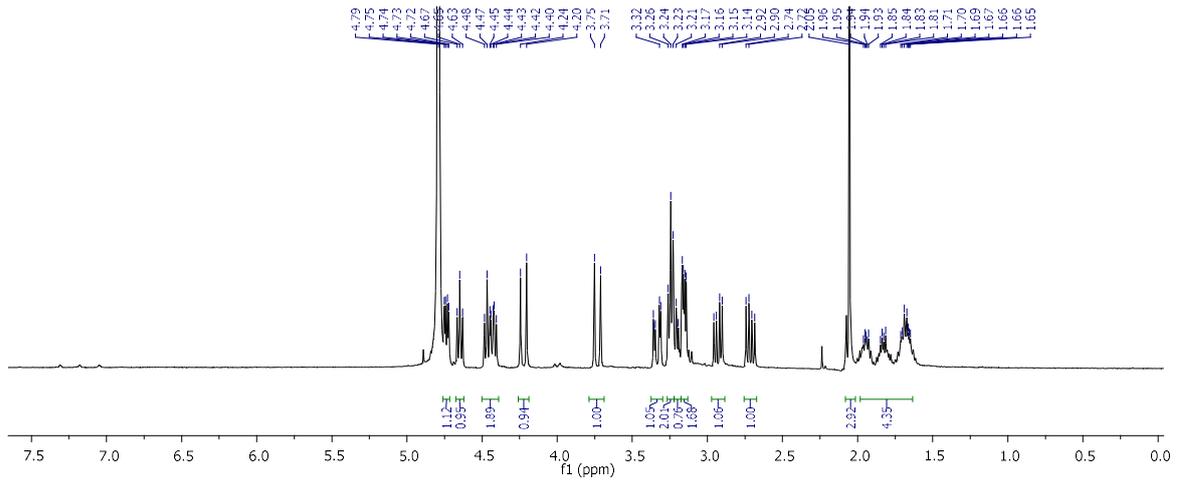


¹³C NMR D₂O

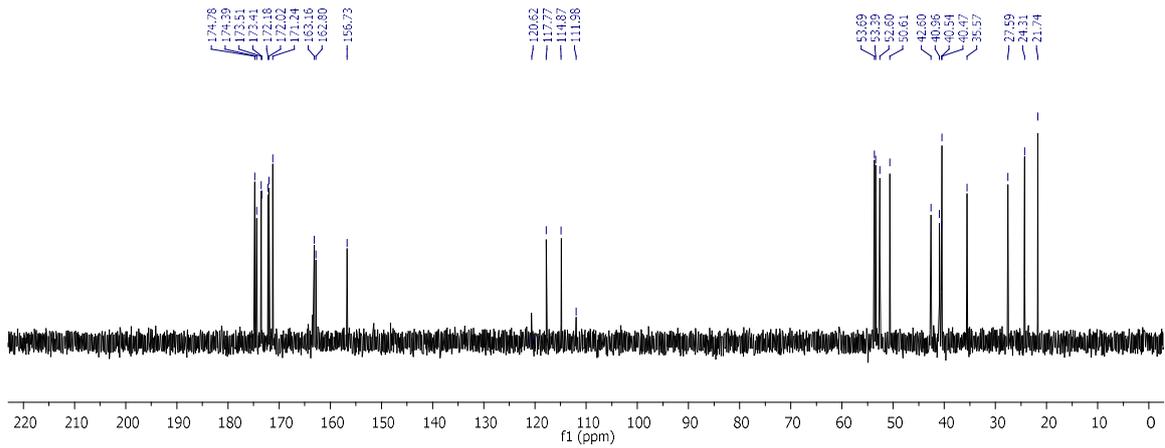


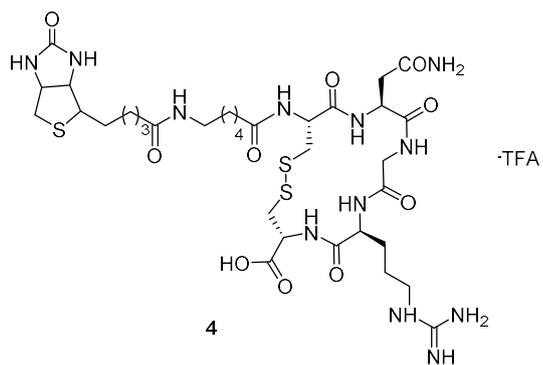


¹H NMR D₂O

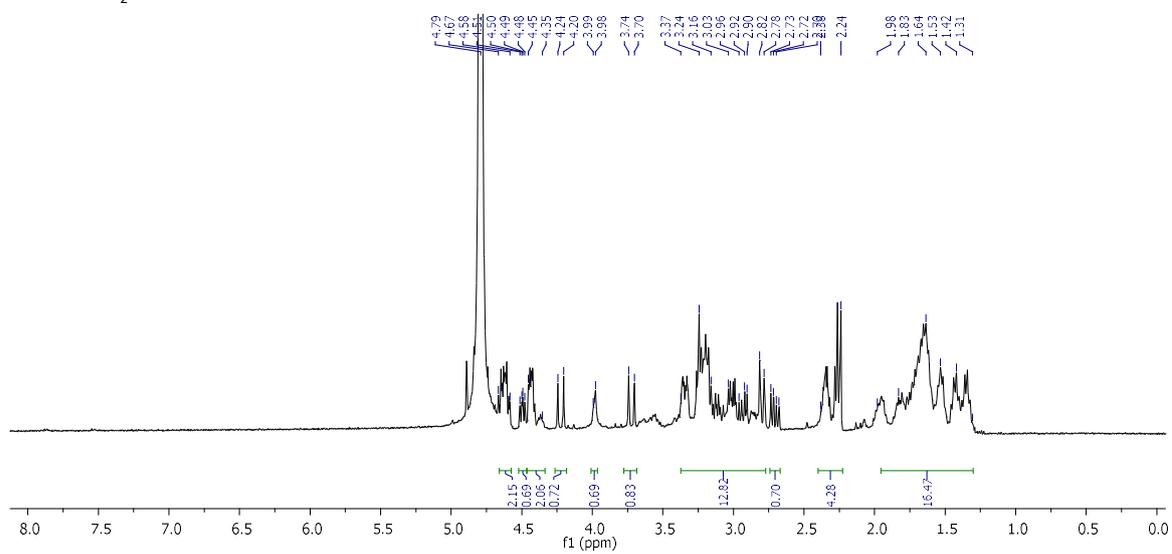


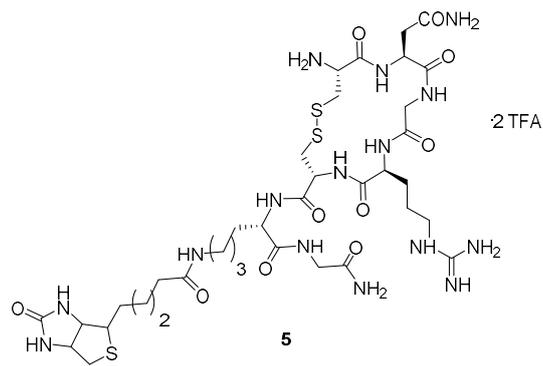
¹³C NMR D₂O



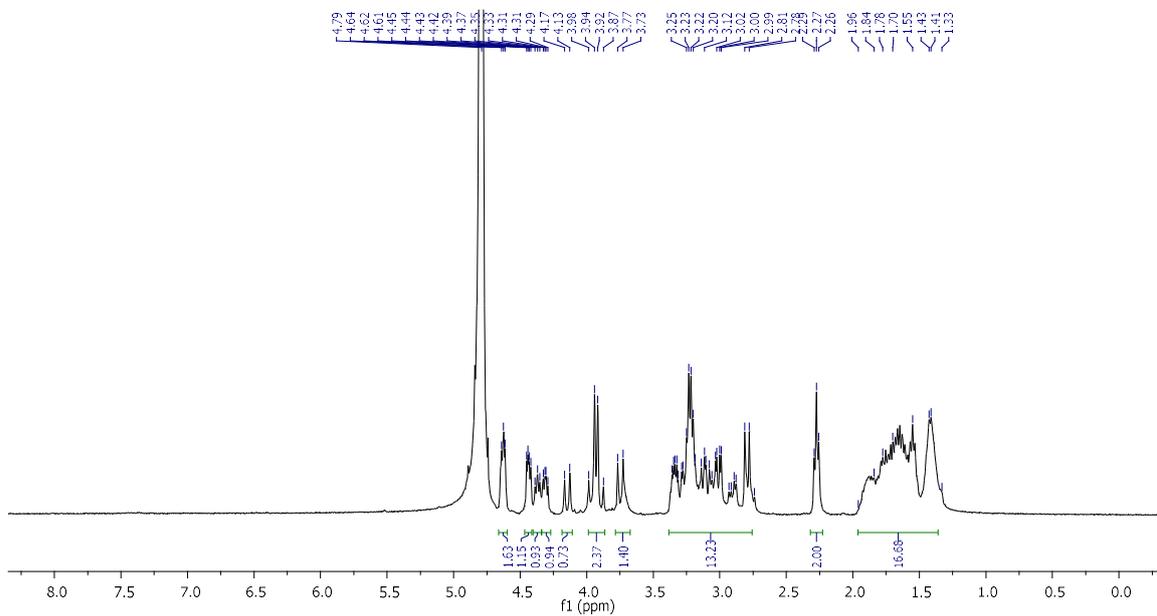


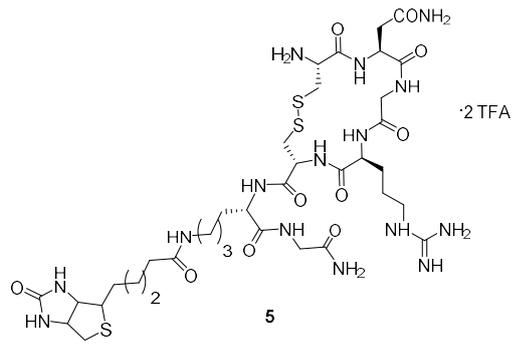
¹H NMR D₂O



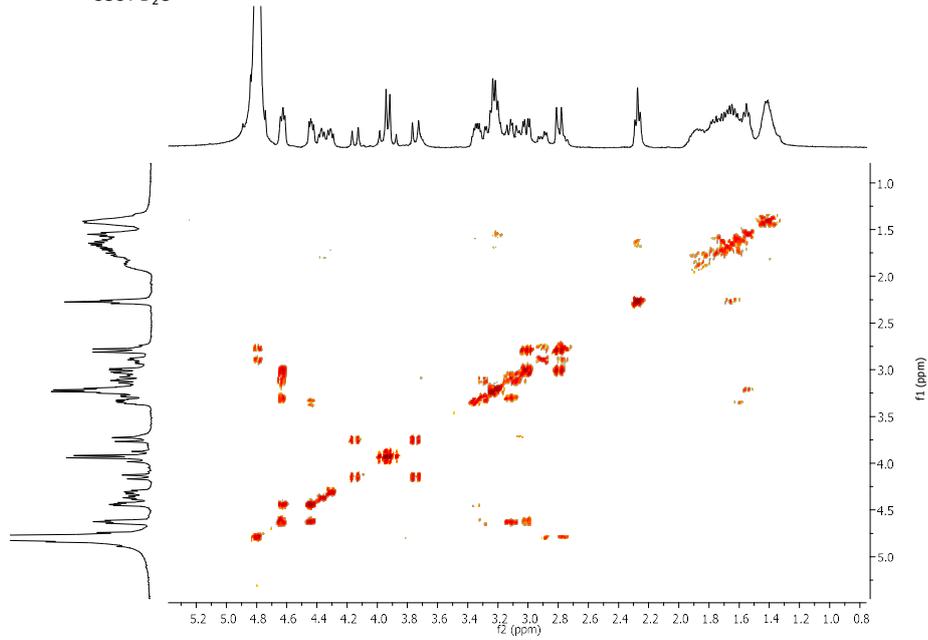


$^1\text{H NMR D}_2\text{O}$

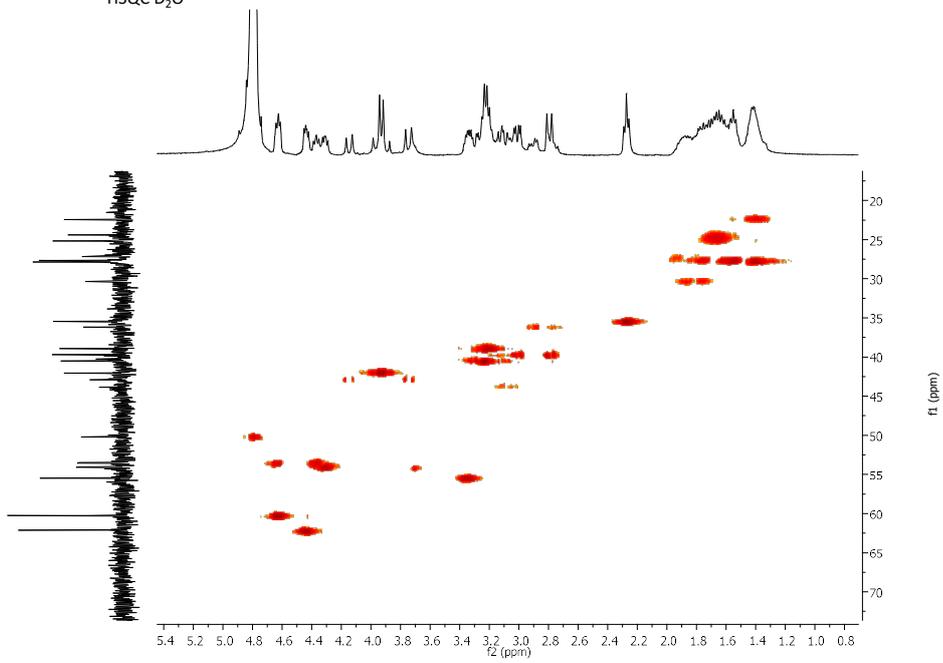


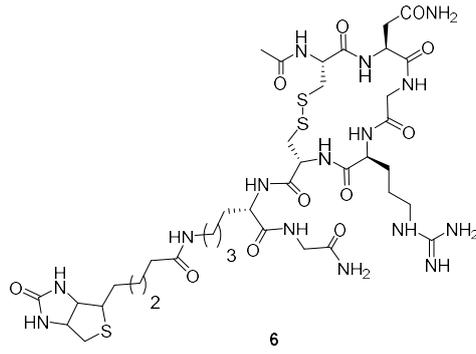


COSY D₂O

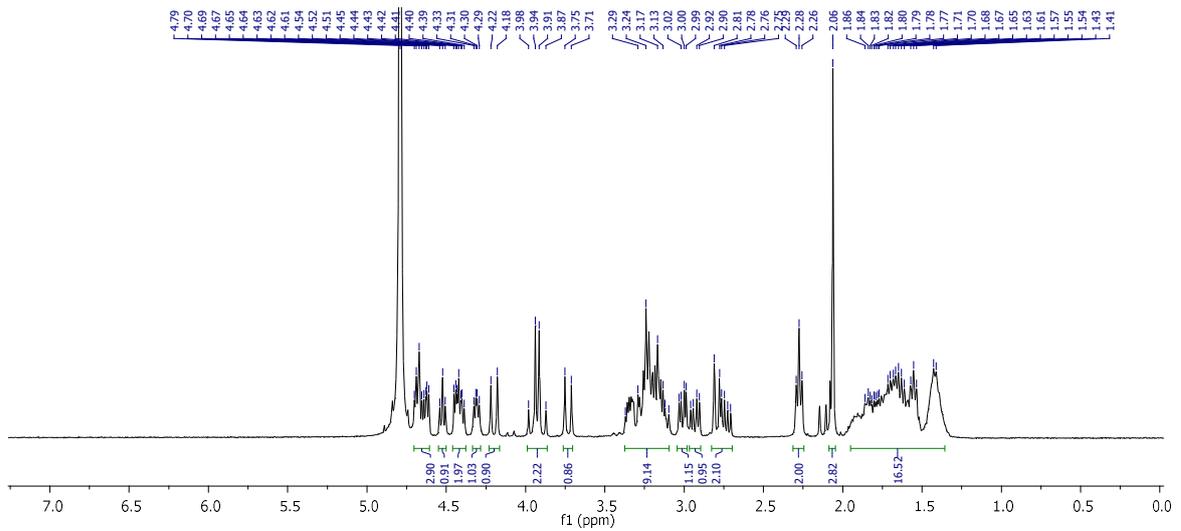


HSQC D₂O

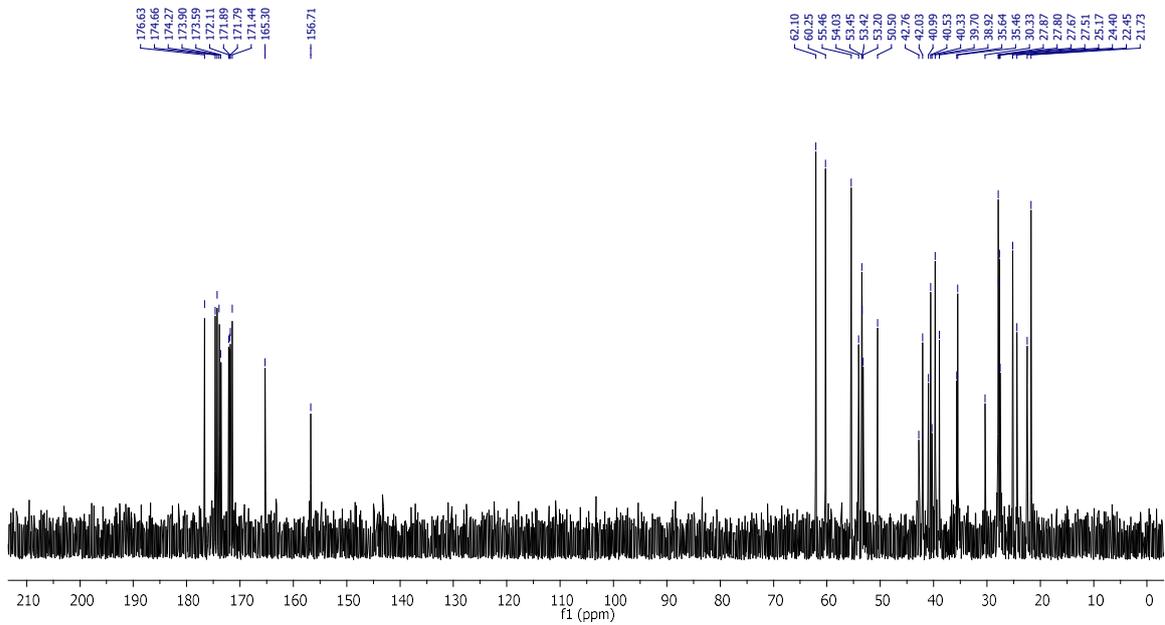


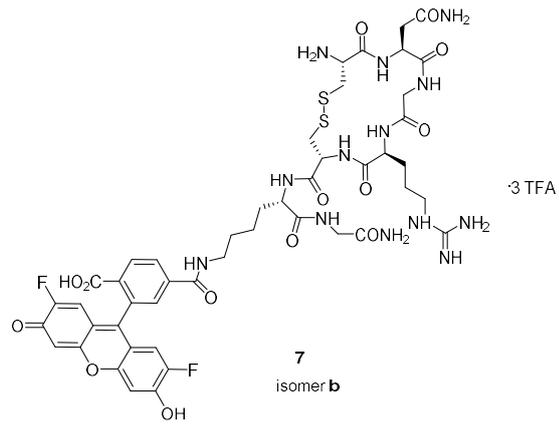


¹H NMR D₂O

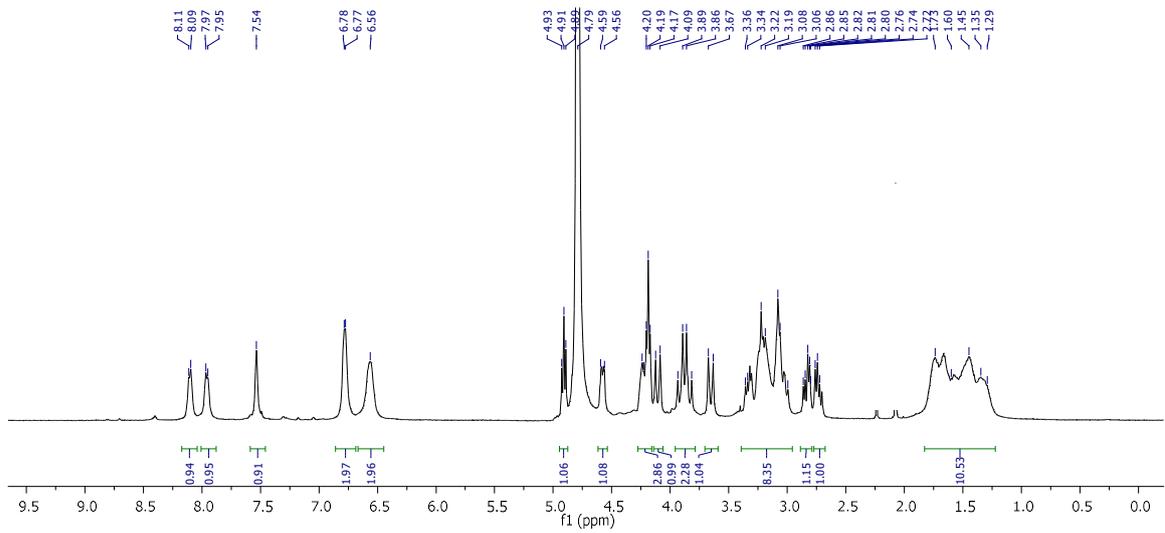


¹³C NMR D₂O





¹H NMR D₂O



¹F NMR D₂O

