UNIVERSITÀ
DEGLI STUDI
FIRENZE

DOTTORATO DI RICERCA IN SCIENZE CHIMICHE

CICLO XXX

COORDINATORE Prof. BAGLIONI PIERO

Studi strutturali sugli addotti formati da composti naturali e sintetici con avvolgimenti non canonici di DNA

Structural investigations on the adducts formed by natural and synthetic compounds with non-canonical DNA foldings

Settore Scientifico Disciplinare CHIM/03

Dottorando
Dott. Papi Francesco

(firma)

Tutore
Prof. Bazzicalupi Carla

(firma)

Coordinatore
Prof. Baglioni Piero

(firma)

Anni 2014/2017

Thesis for the degree of Doctor of Philosophy in Chemical Sciences

Candidate Francesco Papi

Structural investigations on the adducts formed by natural and synthetic compounds with non-canonical DNA foldings

XXX PhD cycle

2014-2017

Supervisor
Professor Carla Bazzicalupi

Co-supervisors
Professor Paola Gratteri
Doctor Marta Ferraroni

PhD Program Coordinator
Professor Piero Baglioni

Ai miei genitori

ABSTRACT

Non-canonical DNA structures are involved in fundamental biological processes as replication, transcription and repair. Their dysregulation is indeed connected to the development of several human diseases, including cancer. As more and more information about their existence and function in living cells are documented, such DNA structures have emerged as promising therapeutic targets. In the last decades, the G-quadruplex folding has caught the attention of scientists because of its implication in the origin and growth of various cancer forms. The stabilization of G-quadruplex structures at human telomeres is thought to be particularly attractive as it might lead to the identification of potential drug candidates with wide-spectrum anticancer activity and reduced side effects in comparison to classical chemotherapies. The research project underlying this thesis concerns the structural investigation on the interaction of non-canonical DNA foldings, especially of the human telomeric G-quadruplex, with natural and synthetic compounds in order to select potential anticancer drugs. The characterization of ligand-DNA adducts has been carried out primarily by X-ray crystallography which provides detailed structural information. In addition, alternative techniques, as CD spectroscopy and *in silico* calculations, have supplied complementary data with particular reference to the formation of adducts in solution.

CONTENTS

1 INTRODUCTION

1.1 DNA, AS MUCH IMPORTANT AS COMPLEX

Deoxyribonucleic acid (DNA) is the biomacromolecule that stores all the information necessary to the outliving, development, reproduction and death for all known living organisms as well as for many viruses. Because of its fundamental importance, DNA is safely conserved within living cells, either in eukaryotic nuclei or in prokaryotic nucleoids, in tight and ordered organizations as chromosomes and chromatin of eukaryotes.
DNA holds the information required for life in genes, that are molecular units showing both coding ($\rightarrow$ proteins) and regulatory sequences. It is still surprising that the majority of DNA (more than 98% in humans) is formed by non-coding repetitive sequences which, however, may have significant importance in the genome function and maintenance, as lately discussed about telomeres [1]. The expression of genes is regulated by means of many different mechanisms in according to the cell needs.
Being the source of all biological information, damages on DNA can take to severe problems. Mutations can have different causes including replication errors, chemicals, viruses, UV and ionizing radiations which lead to casual modifications of the DNA sequences and so unpredictable scenarios. In order to minimize the damages, living organisms have developed various repair DNA processes which rely on the ability of specific proteins to fix the modifications encountered in the DNA sequence [2].
Unfortunately, not all DNA damages can be repaired and serious

diseases, including tumours, are likely to arise. The attention of the thesis is focused on the identification of potential anticancer drugs which is a challenging task because cancer is actually a family grouping different illnesses with common features that are abnormal cell growth and altered cell functionalities. Up to now there have been identified more than 100 cancer diseases in human beings that differ in the organs they originate from, the kind of cells they affect and their physiology [3].
Nowadays, cancer is fought by a combination of chemotherapy, targeted therapy, radiation therapy, surgery and lately immunotherapy [3]. Chemotherapy is still crucial in anticancer treatments, however its application is often merely based on the higher metabolism of malignant cells compared to the healthy ones. In fact, anticancer drugs usually exert cytotoxic activity towards all the cells of the organism, but, because of the higher metabolism, tumour cells are in general the most targeted. Despite the positive action against cancer cells, chemotherapeutic agents usually show side effects which can sometimes be severe. More efficient anticancer drugs with reduced side effects are thus demanding [4].
Non-canonical DNA foldings are promising therapeutic targets for the identification of better anticancer drugs as they are involved in important processes and exert high biological activity [5]. During the years of PhD, different compounds have been evaluated as DNA binding molecules of appealing non-canonical foldings with the aim to select potential drug candidates. In order to properly discuss the results obtained, an introduction on the complexity of DNA structures has to be covered.

1.1.1 Building blocks of DNA

Since its isolation in 1869 [6], DNA has caught great interest for its fundamental importance in living organisms but also for its pecu-

liar properties. DNA is indeed a biopolymer with different levels of organization that span from short random coil oligonucleotides to extremely complex and ordered structures like chromosomes and chromatin [7]. The first step in understanding the DNA properties is the analysis of its building blocks, the nucleotides.
DNA nucleotides are formed by a nitrogenous base, a β-D-2′-deoxyribose residue and a phosphate group (figure 1). If the phosphate group is not present, the molecule is termed nucleoside. Bases can be either purines (guanine, adenine) or pyrimidines (cytosine, thymine). Besides the classical ones, other kinds of nitrogenous bases have been pointed out to be present in the DNA sequence. These bases can originate from DNA damages but it has been demonstrated that a significant amount of them is naturally produced by living organisms. The motive of the existence of non-classical bases is still debated, however it seems they might be implicated in important biological processes, as the considerable role of 5-methylcytosine in the gene transcription suggests [8, 9].

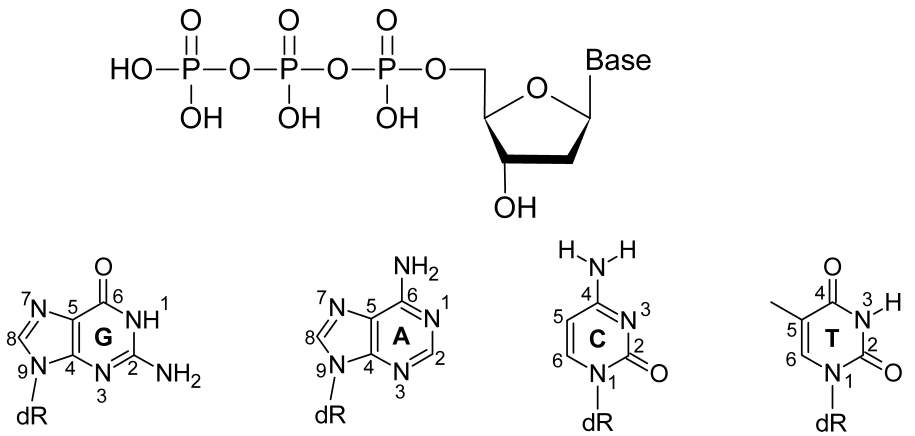


Figure 1: Structure of a DNA nucleotide. The nitrogenous base can be either guanine (**G**), adenine (**A**), thymine (**T**) or cytosine (**C**).

DNA is formed by linear chains termed strands in which nucleotides are connected by means of phosphate and sugar groups, the so called DNA backbone. By convention, the order of the sequence is 5′ ⟶ 3′ (figure 2), where the 5′ end is defined as that extremity which lacks of a nucleotide in the 5′ position while the 3′ end is that extremity without a nucleotide in the 3′ position. The directionality of the macromolecule arises from the extension of DNA strands in a 5′ to 3′ way by DNA polymerases.

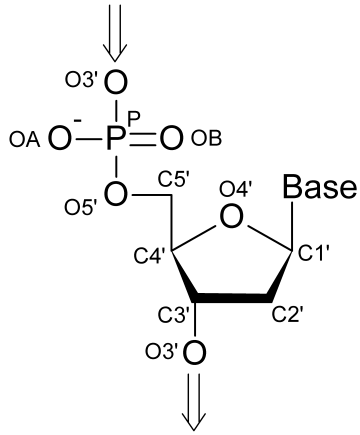


Figure 2: The directionality of DNA strands is 5′ ⟶ 3′ by convention.

Because DNA polynucleotides show great flexibility, an accurate analysis of their conformations is quite useful. An important aspect in this regard is the conformation of the deoxyribose residue. The sugar is indeed puckered and different types of puckering can be identified as a consequence of steric hindrances that restrain the torsional angles in the ring. When one ring atom is out of the plane of the other four, the pucker is termed envelope. More commonly, two atoms deviate from the plane of the other three, being these two at opposed sides of the plane. It is usual for one of the two atoms to have a larger deviation from the plane

than the other, resulting in a twist conformation. The direction of atomic displacement from the plane is also important. If the major deviation is on the same side as the base and the C4′-C5′ bond, then the atom involved is termed *endo*. If it is on the opposite side, it is called *exo*. The most commonly observed puckers in crystal structures of isolated nucleosides and nucleotides are either close to C2′ *endo* or C3′ *endo* types (figure 3) [10, 11].

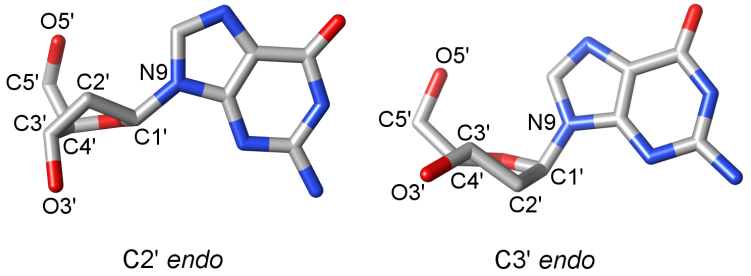


Figure 3: C2′ *endo* and C3′ *endo* sugar puckering for guanine nucleosides. Carbon coloured in grey, oxygen in red and nitrogen in blue.

The glycosidic bond is defined as the covalent bond between the sugar and the base in each nucleotide. It is termed β-anomeric because the base is on the same face of the ring plane as the 5′ hydroxyl substituent. Analogously, the glycosidic torsional angle χ is described. In the case of purines, the angle is defined by the O4′-C1′-N9-C4 atoms, instead in the case of pyrimidine bases it is defined by the O4′-C1′-N1-C2 atoms. The χ angle shows two main low energy ranges referred as *anti* ($-120° < \chi < 180°$) and *syn* ($0° < \chi < 90°$) conformations (figure 4) [10, 11]. Purine nucleotides have no particular preference for *syn* over *anti* configurations with the exception of guanine nucleotides which usually show *syn* conformations [10, 11]. Pyrimidine nucleotides usually present *anti* conformations [10, 11].
A more complete depiction should consider also the overall back-

bone torsional angles which are indeed found in certain low energy ranges [10, 11]. As they are not useful to the analyses addressed in the thesis, they will not further discussed.
The conformational parameters defined are really helpful in the description of ordered secondary DNA structures.

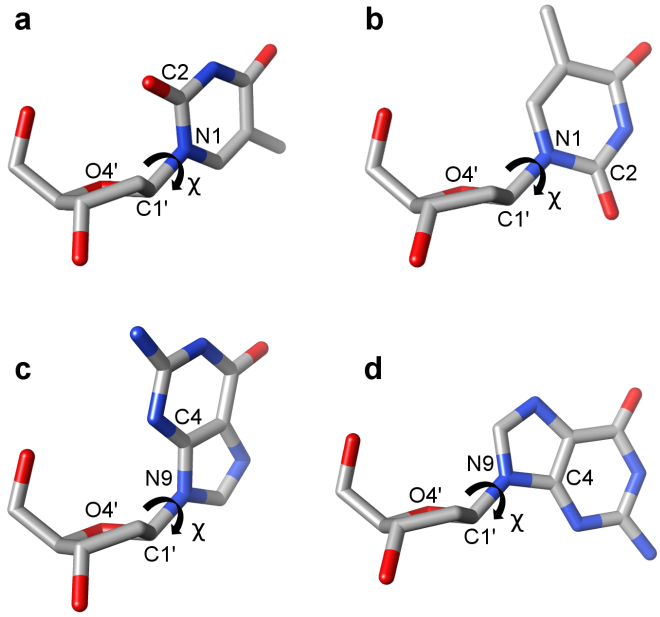


Figure 4: *Syn* and *anti* conformations for the glycosidic angle χ in thymine (a *syn*, b *anti*) and guanine (c *syn*, d *anti*) nucleosides. Carbon coloured in grey, oxygen in red and nitrogen in blue.

1.1.2 Classical DNA structures

The structural complexity of DNA arises from its tendency to form different foldings in which alternative non-covalent interactions are recognized, regarding in particular the nitrogenous bases. The strict connection between non-covalent interactions

and DNA foldings was plainly established by James Watson and Francis Crick (basing on the X-ray diffraction experiments of Rosalind Franklin) with the identification of the right-handed double helix DNA structure formed by antiparallel oriented strands (figure 5) [12]. This structure is usually referred as the duplex B-type DNA and it is the classical DNA structure par excellence. As it describes local arrangements of DNA sequences, the B-type DNA has to be considered a secondary structure of the macromolecule.

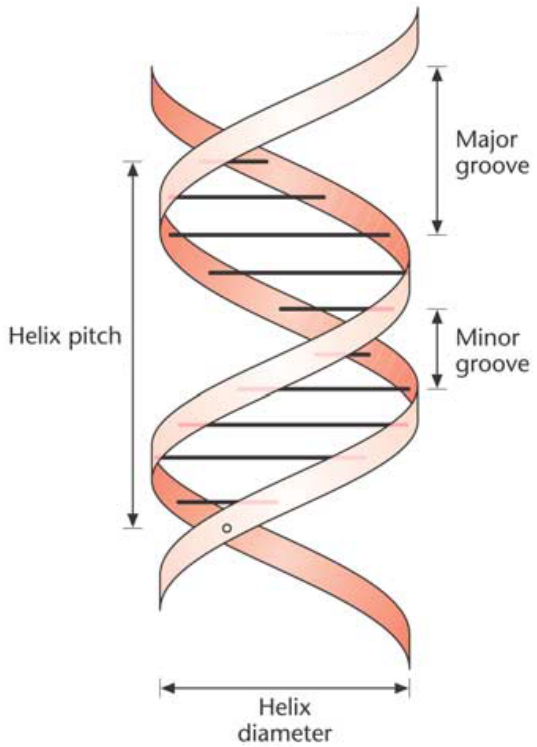


Figure 5: Schematic representation of the Watson-Crick B-DNA structure. Adapted from reference [13].

The double helix organization shows numerous non-covalent interactions between residues both belonging to the same strand and to different strands [14]. The interactions are essentially hydrogen bonds and π-π stacking involving the nitrogenous bases. While the latter are non-directional and non-specific, the former show great specificity and directionality which allows molecular recognition between the bases. In fact, the B-type DNA structure is defined by the formation of specific base pairings between adenine and thymine, involving two hydrogen bonds and indicated as A=T, and between guanine and cytosine, in which three hydrogen bonds are formed as highlighted by the notation G≡C (figure 6). These couplings have been named Watson-Crick base pairings and referred as classical base pairings [14].
Hydrogen bond interactions show donor-acceptor distances around 2.9 Å while the inter-planar distances between the aromatic rings of the bases are about 3.4 Å [10, 11]. As regards the π-π stacking, the face to face and parallel-displaced are the most common arrangements for the bases [10, 11].

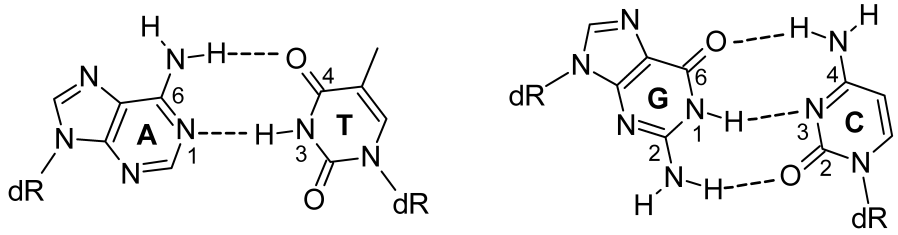


Figure 6: A=T and G≡C base pairings in the Watson-Crick double helix.

Both theoretical and experimental studies have indicated that the singular formation of the G≡C pairing is more energetically favoured than the A=T pairing as the formation energies are approximately -29 and -15 kcal/mol, respectively [15–17]. Conversely, the formation energies of π-π stacking interactions lie between -17

and -10 kcal/mol [15, 18]. The energies involved into these kinds of interactions are so comparable. Nevertheless, hydrogen bonds are thought not to be the main stabilizing factor in the DNA double helix structure [19]. Base-stacking plays indeed the determinant role in the stabilization of the folding due to the high cooperativity of the interactions established along the strands. In more detail, when the stabilization contributes of the base pairings are considered, it emerges how the A=T pairing is destabilizing whereas the G≡C pairing does not substantially affect the stability of the B-DNA helix [20]. Theoretical calculations have definitely indicated that London dispersion forces are crucial for the formation of the double helix DNA structure [20, 21].
The DNA backbone is quite resistant to cleavage so it offers a good protective and supporting architecture to the overall structure. The backbone is directed outward from the core of the structure where the bases are localized and gives rise to the so called grooves. DNA grooves are corrugations in the surface of the helix which wind along, parallel to the backbone. Grooves are classified as major and minor in dependence of their width and depth [10, 11]. Essentially, the major groove occurs where the backbones of strands are far apart, instead the minor groove is observed where they are close together. Because of the negatively charged phosphate groups at physiological pH, cations tend to bind to the backbone so nullifying the overall negative charge of DNA [22].
The structure proposed by Watson and Crick is the most common secondary structure observed for DNA in living cells [10, 11]. It has clarified how DNA stores the genetic information through the (A-T-C-G) quaternary code and has also suggested the semiconservative replication mechanism of DNA. Furthermore, as the non-covalent interactions are copious and cooperative, the double helix structure is quite stable and allows the safe conservation of the genome. Last but not least, its good flexibility favour the pack-

ing of the macromolecule in chromosomes and chromatin [7, 23]. Along with the B-type, DNA has been reported to fold in alternative double helix structures, namely A-DNA and Z-DNA (figure 7) [24, 25]. The former has been discovered by Rosalind Franklin and it is essentially a more compact form of B-DNA found for certain purine sequences in dehydrating conditions. On the other hand, Z-DNA is quite different as the antiparallel strands give rise to a left-handed double helix. The formation of this folding is usually unfavourable unless in the presence of an alternating *syn* - C3′ *endo* and *anti* - C2′ *endo* nucleotide framework [10, 11]. Sequences rich in cytosine and guanine under high salt conditions (taking to a minimization of electronic repulsion between close backbones) are particularly valuable for the formation of Z-DNA.

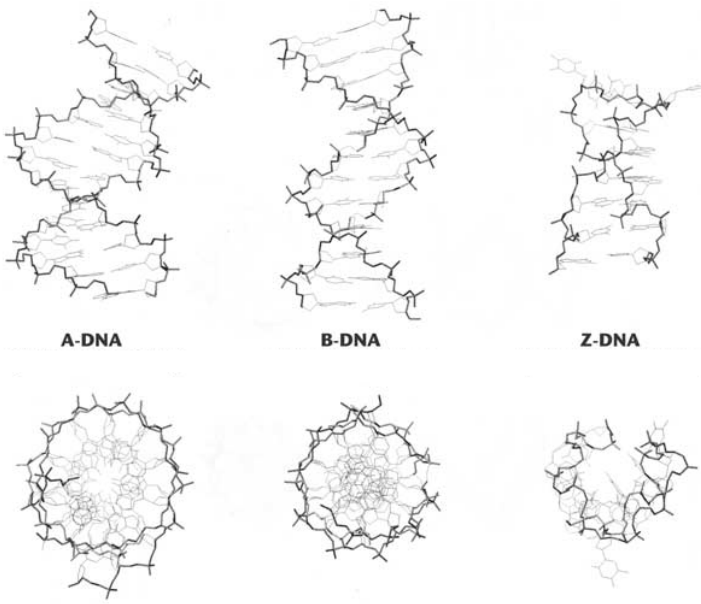


Figure 7: Structures of A, B and Z double-stranded DNA. Adapted from reference [13].

In order to better understand the differences between the three helices, some average structural parameters are reported in table 1 [10, 13]. A-DNA shows greater helix diameter and number of bases per turn than B-DNA. It displays also a more pronounced inclination of bases with respect to the helix axis and, as a consequence, the major groove is deeper while the minor groove is less deep than those observed in B-DNA. Conversely, Z-DNA is more stretched than the other foldings, as suggested by greater helix pitch and smaller diameter. Its backbone exhibits a characteristic zig-zag order due to the alternating *syn* and *anti* conformations. Additionally, the minor groove is quite deeper than the major one.

Table 1: Structural parameters of DNA double helices. M=major, m=minor.

	B-DNA	**A-DNA**	**Z-DNA**
Helix handedness	Right-handed	Right-handed	Left-handed
Diameter of helix (Å)	20	23	18
Number of bases per turn	10	11	12
Axial rise (Å)	3.4	2.6	3.7
Helix pitch (°)	34	28	45
Normal inclination of bases (°)	-6	20	7
Helical twist (°)	36	33	-30
Glycosidic bond conformation	*anti*	*anti*	*anti* / *syn*
Sugar pucker	C2′ *endo*	C3′ *endo*	C2′ *endo* / C3′ *endo*
Grooves depth (Å)	8.5 (M) 8.2 (m)	13.0 (M) 2.6 (m)	3.7 (M) 13.8 (m)
Grooves width (Å)	11.6 (M) 6.0 (m)	2.2 (M) 11.1 (m)	8.8 (M) 2.0 (m)

The biological functions of A-DNA are largely unknown, although some studies have shed light on its presence and functions in viruses [26]. Z-DNA, on the contrary, seems to be involved into genetic transcription and its formation has been evidenced in connection to enzyme binding as described for the Z-DNA-binding protein 1, also called DLM-1 protein [27, 28]. Actually, B-type DNA is the energetically favoured structure for a random sequence in physiological conditions hence studies on alternative helix foldings appear to be quite challenging [10, 13].

The existence of other similar foldings as the C-type and E-type DNA has been reported along the years, but these structures are generally formed by synthetic sequences thus they are thought not be relevant for DNA *in vivo* [10, 13]. Nevertheless, the actual polymorphism and structural complexity of the biomacromolecule has been even more confirmed. Further examples of DNA foldings are those concerning non-helical structures which are termed non-canonical.

1.1.3 Non-canonical DNA foldings

Despite the significant stability of the Watson-Crick double helix, DNA shows several secondary organizations in which this structure is partially or totally missing. Non-canonical foldings have been evidenced to be present in living cells and their existence, even transient, is presumably connected to the regulation of various processes [5, 29, 30]. Thus, these foldings might be of great help in unravelling important biological mechanisms and also offer feasible therapeutic routes for the treatment of severe diseases as cancer [31].
Palindromic sequences are representative examples of DNA tracts that are in equilibrium between different foldings. As the antiparallel strands show the same bases sequence albeit oppositely 5′ end - 3′ end oriented, palindromic sequences can give rise to stem-loop arrangements (also known as hairpin loops, figure 8) where the bases related to the palindrome are involved in Watson-Crick hydrogen bonds while other bases are completely unpaired [32]. The biological relevance of such structures is closely related to the corresponding RNA foldings which play a significant role in the regulation of genetic transcription as for 5′ untranslated regions (UTRs) [32, 33].

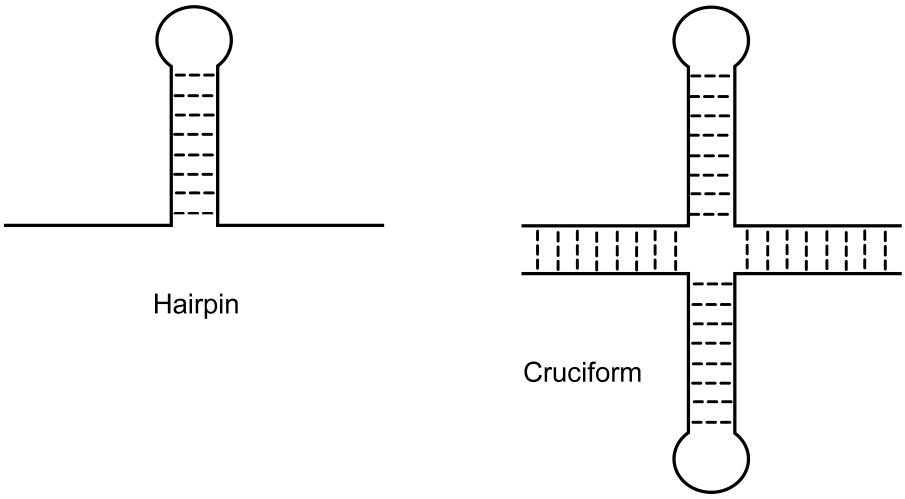


Figure 8: Schematic representations of hairpin and cruciform DNA structures.

Holliday junctions are an interesting case of DNA strands involved both in helix structures and non-canonical arrangements (figure 9). The junctions are indeed formed by four double-stranded strands joined together [34, 35]. The resulting branches can adopt different conformations, depending on the sequences and on salt conditions of solutions [36]. It is easy to imagine how Holliday junctions play a determinant role in genetic recombination and repairing processes and in fact they are promising therapeutic targets [37–39].

Hairpin structures and Holliday junctions are DNA foldings still showing Watson-Crick base pairings, A=T and G≡C. Different base pairings have however been reported and they are usually referred as of Hoogsteen type, a general class of bases arrangements unforeseen by the Watson-Crick model. Actually, the number of Hoogsteen arrangements observed in DNA structures is quite great and it is connected to the formation of non-canonical

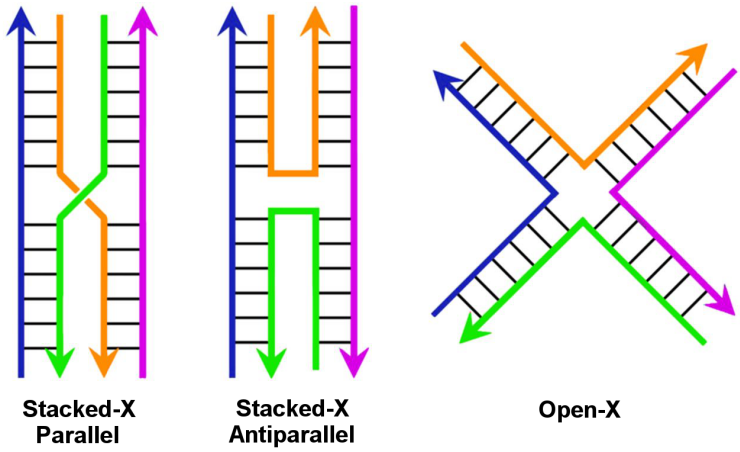


Figure 9: Schematic representations of Holliday junctions. Adapted from reference [39].

DNA foldings [10, 11].

In the early 1960s Hoogsteen described for the first time some non-classical interactions (hereinafter referred by the • symbol) regarding A•T and G•C^+ base pairings then named after him (figure 10) [40, 41]. These arrangements involve in hydrogen bonding different atoms with respect those involved in the Watson-Crick base pairs (figure 6). It is noteworthy to highlight that purines implicated in such pairings show *syn* instead *anti* glycosidic conformation.

The formation of non-canonical base pairings is usually referred to base mismatches which are thought to be dangerous for living organisms as they may promote mutations and eventually cancer [42]. Typical examples reported for mismatches arisen during cell replication are G•A, G•T, A•A, G•G, T•T, C•C, T•C and A•C base pairings [10, 11]. Cells produce in this regard enzymes able to repair mismatches, although the complete restoration of base pairings is not assured [43]. The presence of mismatches makes the

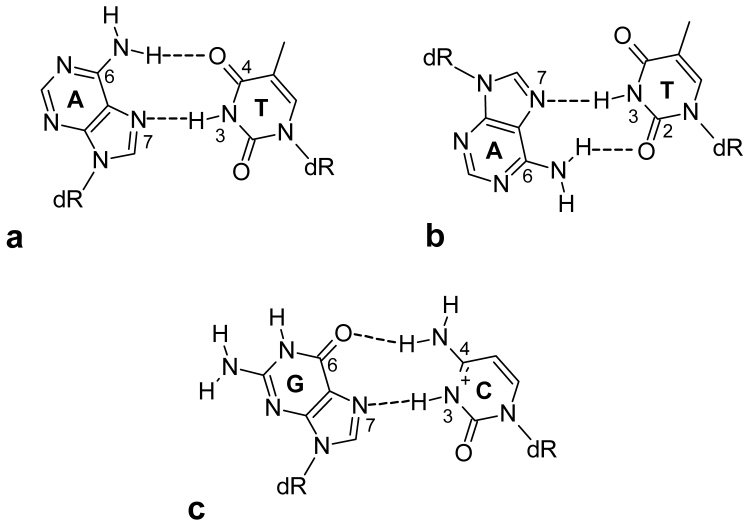


Figure 10: Hoogsteen (a) and reverse Hoogsteen (b) A•T base pairings, Hoogsteen G•C^+ (c) base pairing.

double helix structure less stable and so different non-canonical structures are likely to originate [42].
DNA triplexes arise from the association of three DNA strands involving the formation of triads of bases interacting by non-Watson-Crick hydrogen bonds (figure 11) [44]. The triads are either composed of pyrimidine-purine-pyrimidine or purine-purine-pyrimidine bases [10, 11]. Though many possible triads have been reported, the most valuable are the T=A•T and C≡G•C^+ triplets (figure 12). The formation of the latter can be triggered by lowering the pH to 5 or lower so allowing a fraction of cytosine to get protonated on its N3 atom.

Triple helices show quite different morphologies and sugar conformations depending on the sequences and the conditions explored. Overall, the close proximity of negatively charged phosphate groups make triplexes not very stable in comparison with

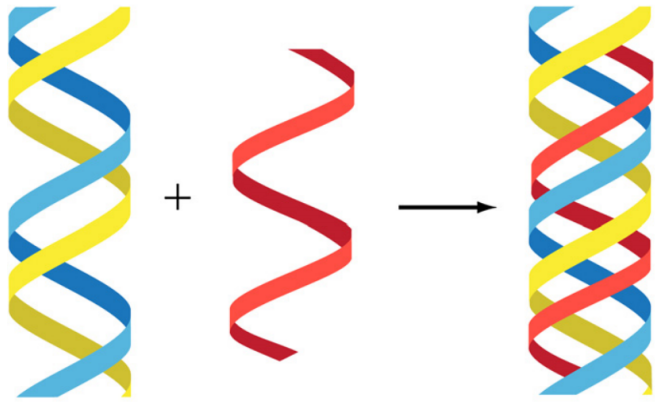


Figure 11: Schematic representation of triple helix DNA. Adapted from reference [45].

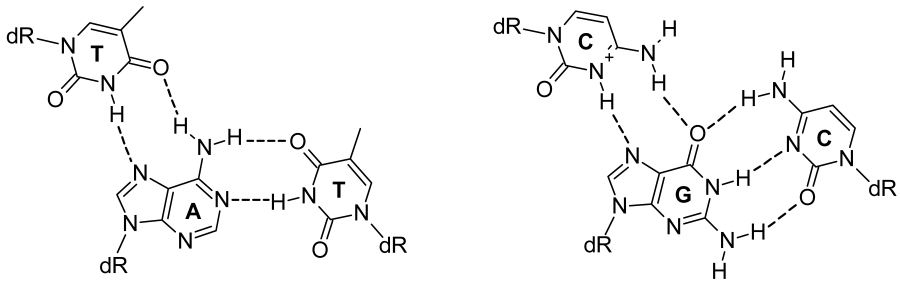


Figure 12: T=A•T and C≡G•C$^+$ base pairings in DNA triplexes.

the related double helices [44].
The interest in DNA triplexes relies in their ability to exert regulatory activity in genetic transcription, replication and recombination [45, 46]. Their presence is also connected to genome instability as confirmed by the specific binding to the p53 tumour suppressor protein [47]. Thus, triple helices offer promising routes for genetic therapy, in particular for the specific inhibition of particular genes [44, 45, 48].
The i-motif structure is a further DNA folding found in quite acidic solutions for cytosine-rich sequences where some cytosines' N3 get protonated. The organization results from the intercalation of two double helices thanks to $C \bullet C^+$ base pairings showing three hydrogen bonds (figure 13) [49]. Actually, the i-motif structure is a four-stranded arrangement and it is no coincidence that complementary guanine-rich sequences are able to fold into four-stranded foldings themselves (see below). The biological relevance of i-motif relies in the regulation of genetic transcription as suggested by the significant number of proteins that specifically bind cytosine-rich telomeric tracts [49].

1.1.4 G-quadruplex DNA structures

Guanine self-assemblies have been observed since the 19^{th} century [50] but a general understanding regarding these phenomena was missing until the work published in 1962 by Gellert and coworkers which demonstrated the arrangement of guanosine monophosphate in quartets of coplanar guanines interacting through non-classical hydrogen bonds, the so-called G-quartets or G-tetrads (figure 14) [51]. Later on it was found that poly-guanine sequences were able to originate four-stranded helix structures with the G-quartets stacked on one another [52, 53]. The interest in guanine assemblies was then triggered by the discovery that the terminal

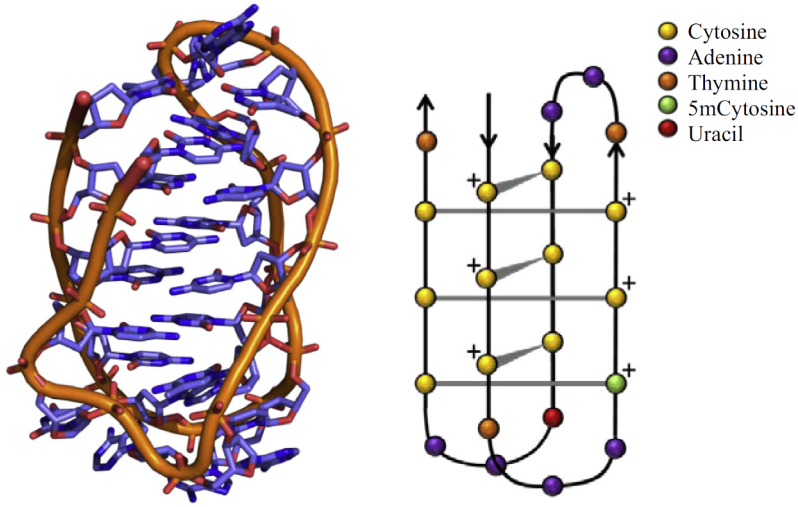


Figure 13: Example of an i-motif DNA structure. Adapted from reference [49].

tracts of eukaryotic chromosomes, the telomeres, are composed of guanine-rich sequences able to fold into four-stranded structures usually referred as quadruplexes, tetraplexes or simply G4 [54]. These non-canonical DNA structures arise from the stacking of at least two G-quartets one on another in a right-handed twist fashion [55, 56].

Positively charged ions play an active role in the stabilization of the quadruplex folding as they are placed in the internal channel formed by the assembly, where there is a negative electronic density (figure 14) [57–59]. The coordination of cations is carried out by guanine O6 atoms which are headed towards the interior of the cavity [60]. The G4 stability is greatly influenced by the nature of ions embedded with particular reference to their charge and dimension [61, 62]. In general, the stabilization, or eventual destabilization, depends on the given G4 structure. Apart from

certain conditions [63, 64], the presence of metal ions is however necessary for the formation of the quadruplex. As regards monovalent cations, the stabilization trend usually observed is $K^+ > Na^+, NH_4^+, Rb^+ >> Li^+, Cs^+$ [61, 62, 65]. K^+ ions are the most representative cations in intracellular medium (about 100 mM in mammalian cells) and it might not be a coincidence that they exert the major stabilization of the folding.

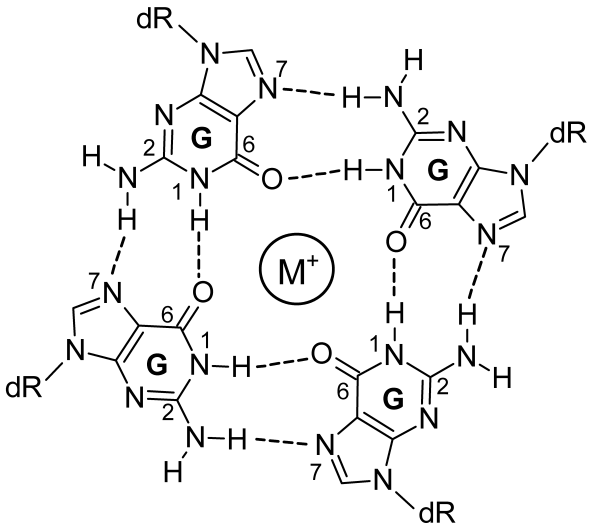


Figure 14: Formation of guanine tetrads in presence of metal cations.

G-quadruplexes are polymorphic and various topologies have been reported [55, 56]. First of all, tetraplexes show different molecularities as they can either be intramolecular, that are originated by one single strand, or intermolecular, so being dimers or tetrameters. The assemblies differ in strands orientation and possibly in the types of loops, which are the tracts connecting different strands (present just in the case of monomeric or dimeric G_4). In general, three main topologies can be identified by taking

into account mutual strands orientation that are parallel, parallel/antiparallel and antiparallel G-quadruplexes (figure 15).

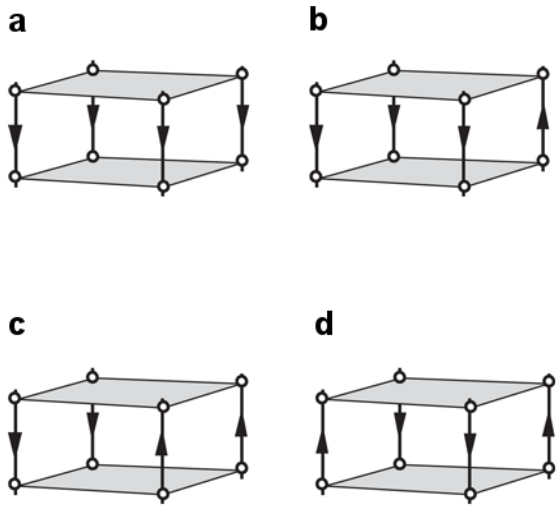


Figure 15: G-quadruplex strands orientations: a) parallel; b) parallel/antiparallel; c,d) antiparallel. Adapted from reference [56].

While the parallel topology displays all *anti* guanines, the antiparallel and the mixed parallel/antiparallel arrangements have guanines in both *syn* and *anti* conformations. When present, loops are also useful for the classification of different topologies as they are closely correlated to the strands orientation (figure 16) [55, 56]. Loops can be:

- Diagonal, connecting two opposing antiparallel strands
- Lateral or edge-wise, connecting two adjacent antiparallel strands

- Propeller or double-chain-reversal, connecting two adjacent parallel strands
- V-shaped, connecting two corners of a G-quartet core in which a support column is lacking

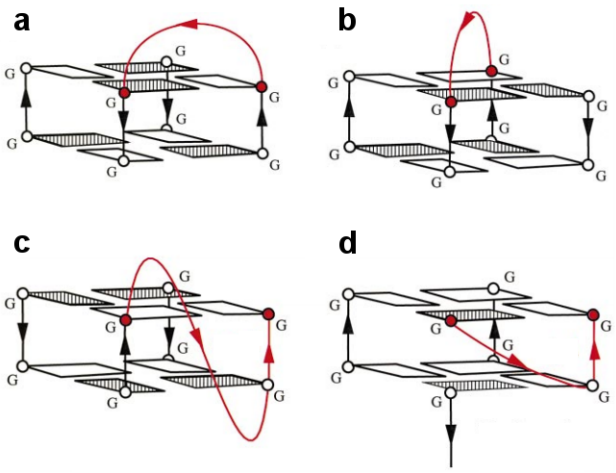


Figure 16: G-quadruplex loops: a) edge-wise or lateral; b) diagonal; c) double-chain-reversal or propeller; d) V-shaped. Adapted from reference [66].

G4 structures show four grooves whose width and depth depend on the loops framework and the overall topology. In general, grooves are classified as narrow, medium and wide [55, 56].
The relative stability of G-tetraplexes depends on many factors such as the nucleotide sequence [67], the length of loops [68–71], the ions enclosed in the structure [61, 62], the concentration of DNA [72] and the flanking residues at the 5′ end and 3′ end [73]. Molecular crowding agents influence the quadruplex topology, too [74]. These chemicals are able to induce both large excluded volume and lower water activity within solutions, so they

are often used to simulate cell-mimicking conditions (they are usually added at about 40% v/v or 400 mg/mL concentration). On the whole, conformational equilibria affect G4 foldings and the actual topology of a given sequence under certain conditions should always be experimentally checked [75, 76].
The folding of guanine-rich sequences into G-quadruplexes over double helices or other foldings seems not to be generally favoured [77, 78], although dehydrating and molecular crowding conditions tend to stabilize G4 structures [79, 80]. Most likely the formation of non-canonical DNA foldings in cells is strictly regulated by cellular mechanisms (e.g. chromatic structure, binding proteins) and this could be especially true for tetraplexes which show slow unfolding kinetics in comparison to other DNA arrangements [81].
Since the statement 'If G-quadruplexes form so readily in vitro, Nature will have found a way of using them *in vivo*' by A. Klug over 30 years ago, huge efforts have been made to demonstrate the existence of G-quadruplexes in living cells. The research in this field has grown notably in the last decades and interdisciplinary approaches have been developed to reach this goal and to better understand how and why such structures might be active *in vivo*.
Theoretical studies have tried to evaluate the number of sequences able to fold into G-quadruplex within the human genome in the first place [82, 83]. Considering only intramolecular quadruplexes, the general formula can be written:

$$\mathbf{G}_m\mathbf{L}_n\mathbf{G}_m\mathbf{L}_o\mathbf{G}_m\mathbf{L}_p\mathbf{G}_m$$

where **m** is the number of guanines involved in the G-tetrads while **n**, **o** and **p** are the numbers of loops bases (potentially including guanines as well). It has been commonly assumed that **m**=3-5 and **n,o,p**=1-7 since otherwise the resulting foldings would likely show significant lower stability [69, 71].

The following formulas have been proposed for sequences able to form dimeric and tetrameric G-tetraplexes, respectively:

$$\mathbf{L}_n\mathbf{G}_m\mathbf{L}_o\mathbf{G}_m\mathbf{L}_p$$

$$\mathbf{G}_m\mathbf{L}_n\mathbf{G}_m \quad \text{or} \quad \mathbf{L}_n\mathbf{G}_m\mathbf{L}_o$$

Using the parameters cited, about 376,000 putative G-quadruplex sequences have been identified in the mid 2000s [82, 83]. As the years passed, the algorithms underlying these studies have revealed their limits as stable G-quadruplexes with relatively long loops have been reported as well as the presence of false positive (sequences that satisfy the requirements but do not form G-quadruplexes *in vitro*) and false negative (sequences which are able to fold into G-quadruplex but do not meet the criteria devised) in the results of genome wide searches [84–86]. More recent studies have indeed predicted a higher number of sequences able to form G-tetraplexes. An accurate high-resolution sequencing-based method has pointed out about 716,000 distinct G4 structures in the human genome [87], in agreement with subsequent theoretical studies [85]. Interestingly, a significant amount of tetraplexes appears to show long loops or bulged structures [87, 88]. Potential G-quadruplexes are largely concentrated in the telomeres, the terminal regions of chromosomes, and in oncogene promoters [82, 83, 89–92]. Mitochondrial DNA and some immunoglobulin switch regions show putative G4 sequences, too [92, 93]. Analogous observations have been made for other eukaryotic species and also for bacteria and viruses [94]. Throughout these organisms, potential G4 sequences have been found in functional regions with a high conservation indicating selection pressure to retain such guanine-rich tracts at specific sites.

G-quadruplexes formed by human telomeric sequences and guanine-rich tracts from proto-oncogenes are very interesting non-canonical DNA foldings and a massive amount of studies has

provided information about their structure and possible biological functions. Since human telomeric G-quadruplexes cover an important part in the research activity underlying the thesis, their features are discussed in the following section.

Oncogene promoters are regions of genes that are involved in the uncontrolled growth and proliferation of malignant cells [95]. Normally, they take part in the regulation of the cell life cycle by preventing apoptosis, that is programmed cell death, until the natural outcome. However, when they are mutated or over-expressed, their action results in the immortalization and proliferation of cancer cells [95, 96]. The existence of dozens of proto-oncogenes has been reported since the 1970s and they are likely to act together in tumours [95, 97, 98]. Remarkable cases of oncogene promoters able to fold into G4 structures are c-Myc, c-Kit, Bcl-2, H-ras, K-ras, B-Raf, TERT and VEGF (table 2) [99–112]. The formation of G-quadruplex structures at the guanine-rich tracts of promoters is responsible of decreasing transcription activity of the genes [113, 114]. Small molecules able to stabilize such foldings are hence thought to be potential wide-spectrum anticancer drugs [86].

Among the sequences able to fold into G4 structures, thrombin binding aptamers are an important case of study. Thrombin is an enzyme designed to manage fibrin formation and platelet aggregation during vascular hemostasis, however it plays an important role in the development of cancer, in particular during metastasis, when its receptors are over-expressed [115]. The thrombin binding aptamer TBA has been identified as inhibitor of thrombin and hence as potential anticancer drug [116]. Interestingly, TBA forms a peculiar antiparallel G-quadruplex, usually referred as chair-type due to its shape [117–122], and its stabilization by small molecules represents a viable route in anticancer drug discovery [123, 124]. Alternative thrombin aptamers, such as HD22, have shown promising therapeutic properties and are widely used

for biosensing applications [123–127].
Actually, an increasing number of putative G4 sequences has been reported whose stabilization might open stimulating perspectives for the treatment of various diseases, as the fragile X syndrome (promoter of the FMR1 gene) [128, 129] and AIDS (HIV-1 long terminal repeat promoter) [130, 131].

Table 2: Sequences of interest able to fold into G-quadruplex and relative native structures deposited in the Protein Data Bank.

ID	Representative sequence	PDB entries	References
c-Myc	$TG_4AG_3TG_4AG_3TG_4A_2G_2$	1XAV, 2A5P, 2LBY	[99–101]
c-Kit1	$AG_3AG_3CGCTG_3AG_2AG_3$	2O3M, 3QXR, 4WO2, 4WO3	[102–104]
c-Kit2	$CG_3CG_3CGCGAG_3AG_4$	2KQH, 2KQG, 2KYP, 2KYO	[105, 106]
Bcl-2	$AG_4CG_3CGCG_3AG_2A_2G_5CG_3A$	2F8U	[107]
H-ras1	$TCG_3T_2GCG_3CGCAG_3CACG_3CG$	(na)	(na)
H-ras2	$CG_4CG_4CG_5CG_5CG$	(na)	(na)
K-ras	$AG_3CG_2TGTG_3A_2TAG_3A_2$	5I2V	[108]
B-Raf	$G_3CG_4AG_5A_2G_3A$	4H29	[109]
TERT	$AG_4AG_4CTG_3AG_3C$	2KZD, 2KZE	[110]
VEGF	$CG_4CG_3C_2G_5CG_4T$	2M27, 2M53	[111, 112]
TBA	$G_2T_2G_2TGTG_2T_2G_2$	148D, 1QDH, 1QDF, 1C32, 1C34 1C35, 1C38, 1RDE, 2IDN	[118–122]
HD22	$GTC_2GTG_2TAG_3CAG_2T_2G_4TGAC$	(na)	(na)
FMR1	$GCG_2T_3GCG_2$	1A6H	[128]

The existence of tetraplex structures in living cells has been recently proven by using G4-specific antibodies [132–134]. The formation of these non-canonical foldings has been demonstrated to be modulated according to the cell cycle having its maximum during the S phase when DNA replication occurs [133]. G-quadruplexes are found in higher densities close to origins of replication, especially in proto-oncogenes, regulatory genes and tumour suppression genes [82, 83, 89–92]. It is noteworthy that elevated levels of G4 structures have been observed in human stomach and liver cancer tissues [135].
It is supposed that tetraplexes may form from unpaired residues during the DNA replication, transcription or repair when the dou-

ble helix is temporary unwounded. Chaperones might also induce G4 folding at least when the kinetics is relatively slow, as observed in ciliates [94, 136]. The G4 formation seems to be required for the initiation of DNA replication as they indicate the localization of replication sites [94]. On the other hand, G4 structures behave at the same time as obstacles to the replication process and they are usually unwound by specific helicases, otherwise they could take to genome instability [94, 137]. FANCJ, XPB, XPD, Pif1, RTEL1, BLM, RECQ4, WRN and PIF1 are notable enzymes involved in the unfolding of such non-canonical structures [138–140]. In particular, the FANCJ helicase is of great interest as it is able to specifically recognize G-tetraplexes and promotes their unfolding during the DNA replication process ($5\prime \rightarrow 3\prime$ directionality) [141]. The downregulation of FANCJ is indeed correlated to the observation of a higher number of G-quadruplexes in living cells [134, 140].
Digressing for a moment, G-quadruplexes have also been observed for RNA guanine-rich sequences in living cells [142–144]. The biological functions of such structures are still matter of investigation, however they should concern the regulation of various processes concerning DNA (telomere elongation, recombination, transcription) but also post-transcriptional mechanisms (e.g. pre-mRNA processing, mRNA turnover, targeting and translation) [145]. An interesting case of putative RNA G4 sequences is that of the $5\prime$ end untranslated regions (UTRs) of many genes where their involvement in the regulation of translation has been demonstrated [146]. Also, telomeric RNA is able to form parallel G-quadruplexes which are thought to exert various functions even through the formation of RNA-DNA hybrids at the telomere [147].
Coming back to our topic, recent studies have pointed out the actual formation of about 10,000 G4 structures in human chromatin of an immortalized cell line by means of an immunoprecipitation technique based on the G4-specific BG4 antibody in combination

with high-throughput sequencing [137]. Quadruplexes are mainly found in regulatory regions, with particular mention of promoters and 5′ UTRs of oncogenic genes, where the presence of such non-canonical foldings is correlated to increased transcriptional activities [137, 146]. The strict difference between theoretical estimations of putative G4 sequences and the actual prevalence of G-tetraplexes in living cells probably relies on the restrictions due to heterochromatin packing and complex regulatory mechanisms [137]. Still, it emerges that putative G4 sequences are largely not folded in G-quadruplexes and such structures are concentrated in genes related to the development of tumours [137].

Human telomeric G-quadruplex

Telomeres are the non-coding DNA regions found at the termini of eukaryotic chromosomes (figure 17). Their main function is to cap chromosomes so protecting them from aberrant activation of the DNA damage response (DDR) and end-fusion with other chromosomes [148]. Telomeres are also actively involved in the regulation of cell life cycle and their instability is connected to the origin and development of cancers [148, 149]. Due to their biological importance, extensive researches have been carried out in order to better understand their functions that culminated with the discoveries by E.H. Blackburn, C.W. Greider and J.W. Szostak which were awarded the Nobel Prize in Physiology in 2009.
Various enzymes are deputed to telomere protection and maintenance (figure 17) [150, 151]. Prime roles are covered by the shelterin complex, the PARP1 enzyme and the CST complex, whose importance in mammals is however solely confined to conditions of replication stress. Telomere protection is also provided by the telomeric RNA (TERRA), that is the transcription product of telomeric DNA. TERRA tasks have not been figured out yet, although

it seems that telomeric RNA behaves as a molecular scaffold for proteins that govern telomere functions [139, 152, 153].

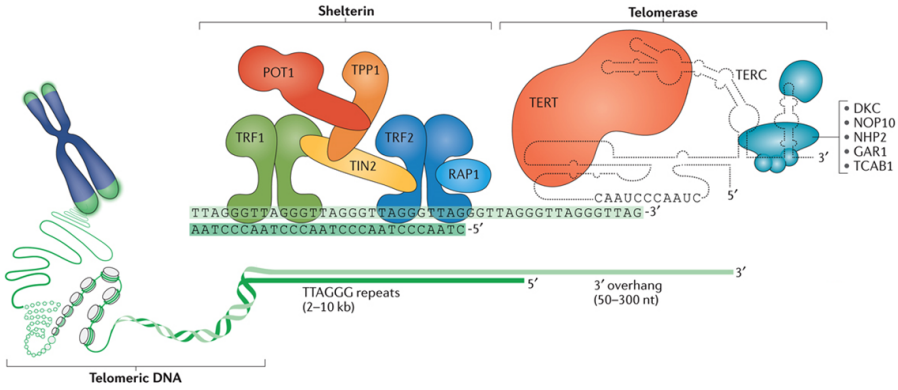


Figure 17: Organization and structure of human telomeres. Adapted from reference [151].

The length of the telomere is an important biological marker as it decreases along with successive cell divisions (due to the requests of the Okazaki fragments synthesis) until it reaches a critical threshold which triggers cell senescence and then apoptosis [148, 149, 154]. While apoptosis indicates the resulting programmed cell death, senescence is a transient state in which cells show metabolic activities but they are not able to replicate any more as they have approached the so called Hayflick limit (about 50-80 cellular divisions).
In humans, the average telomere length diminishes since the birth to the old age from about 10 kilo bases to 2 kilo bases (50-300 bases shortening per division, critical telomere length approximately 6-8 kilo bases) [151, 155]. The length decreasing has however to be modulated over the life time. When needed, cells are indeed able to produce the telomerase complex whose reverse transcriptase (TERT) adds GGTTAG repeats at the 3′ single-stranded overhang

in order to prevent senescence and apoptosis [148]. Telomerase activity is normally downregulated during the organism development by inhibition of TERT, thus allowing the natural cell life cycle to occur [151].
Telomerase acts as a significant regulator of cell life cycle. When its activity is altered as when it is over-expressed, cells gain virtual immortalization because their telomeres length does not reach the limit that triggers senescence and apoptosis [156]. This situation is observed in about 85% of cancer forms along with the inactivation of p53 and pRb tumour suppressor proteins [149, 157–160]. Telomerase activation is frequently caused by mutations in TERT and also genomic rearrangements, as in neuroblastoma [161–163]. Anyway, a minority of cancers, about 10% of forms reported, shows an alternative lengthening of telomeres (ALT) system without the involvement of telomerase [149, 151, 164, 165].
Telomeres are composed of tandem repeats that in humans are TTAGGG copies [94, 150]. They show 2,000-10,000 double-stranded base pairs and a terminal overhang of 50-300 nucleotides at the 3′ end (figure 17) which could fold into G-quadruplex multimers [166, 167]. Actually, the terminal overhang of telomeres is usually associated to binding proteins as POT1 of the shelterin complex that prevent the formation of such non-canonical DNA structures [168]. Telomeric G4 structures can however be observed especially during the S phase, when DNA replication occurs [133]. The formation of telomeric G-quadruplexes relies on their high, not yet fully understood, biological activity in important processes as in replication, where they block access to DNA, behaving as telomere capping agents, but at the same time they indicate the initiation sites [94, 169].
The attention of researchers towards telomeric G-quadruplexes was triggered when it was reported that these secondary foldings are able to inhibit telomerase activity and lead to telomere

uncapping, so opening promising therapeutic perspectives [170–172]. Despite the ability of telomerase to, at least partially, recognize and unwind G4 structures [169], small molecules could indeed stabilize G4 structures at the telomere and so act as wide-spectrum anticancer drugs with reduced collateral effects in comparison to classical therapies [86, 173].

The stabilization of telomeric tetraplexes is not the only viable route for the inhibition of telomerase as different approaches have been developed, largely concerning the inhibition of reverse transcriptase TERT and the telomerase RNA component TERC thanks to small molecules, antisense oligonucleotides, targeted immuno and gene therapies [174–176]. Nevertheless, the direct inhibition of telomerase might lead to cell resistance after excessive telomere shortening, whereas G-quadruplex stabilizers could overcome this problem [176].

Huge efforts have been made in the identification of small molecules able to induce and stabilize the human telomeric G-quadruplex [86] which however shows high polymorphism. In order to properly discuss about G4 ligands, the characteristics of human telomeric G-tetraplexes should in the first place be considered.

Telomeric G-quadruplexes exhibit high polymorphism in molecularities, strand orientations and types of TTA loops, as determined by X-ray crystallography and NMR spectroscopy [55, 56]. As telomeres are most likely to fold into intramolecular structures, alternative molecularities will not be further discussed although the considerations made could be extended to such foldings. Intramolecular tetraplexes show various conformations which can be approximately grouped into three main topologies: antiparallel, parallel and hybrid (figure 18). These groupings mainly differ in the orientations of strands and associated kinds of loops, as shown in table 3.

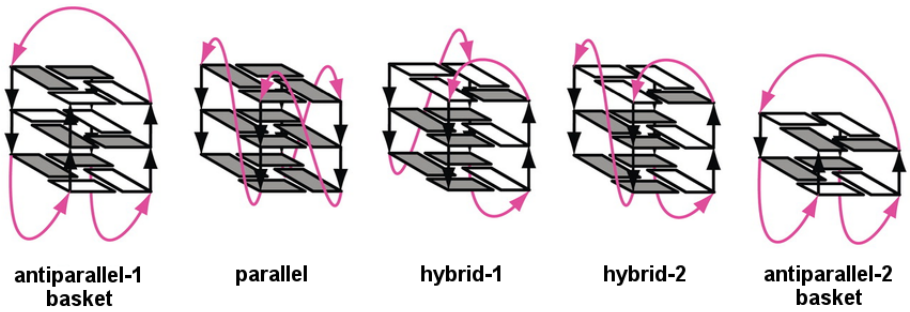


Figure 18: Topologies of intramolecular human telomeric G-quadruplexes. Guanines coloured either in grey (χ *anti*) or in white (χ *syn*). Adapted from reference [177].

Table 3: Generic features of intramolecular G-quadruplexes formed by human telomeric sequences, determined by X-ray crystallography and NMR spectroscopy. PDB codes of native structures are reported, too.

Topology	Strands	Loops	PDB entries	References
Antiparallel-1 basket	antiparallel	lateral diagonal	143D	[178]
Antiparallel-2 basket	antiparallel	lateral diagonal	2KF7, 2KF8, 2KKA	[179, 180]
Antiparallel-3	antiparallel	lateral propeller	2MBJ	[181]
Parallel	parallel	propeller	1KF1, 2LD8	[182, 183]
Hybrid-1	parallel antiparallel	lateral propeller	2GKU, 2HY9, 2E4I, 2JSM 2JSK, 2MWZ, 5MBR	[184–189]
Hybrid-2	parallel antiparallel	lateral propeller	2JPZ, 2JSL, 2JSQ	[187, 190]

The antiparallel-1 topology [178] can easily be stabilized in sodium containing solutions, while the antiparallel-2 [179, 180], the parallel [182, 183] and the hybrids structures [184–191] are more stable in potassium solutions. Human telomeric G-quadruplexes thus show higher polymorphism in potassium solutions [191, 192], although an additional antiparallel structure, featuring lateral and propeller loops, has been recently observed in sodium solutions [181]. An important factor for the resulting stability of the arrangements is the nature of flanking bases which can form caps on the external guanine-tetrads [193]. Also, the antiparallel-2 basket topology featuring only two guanine tetrads shows surprising great stability due to the numerous base pairings and stacking interactions in which the loops bases are involved [179, 180]. Anyway, the *in vitro* relative stability of different topologies is not obvious to predict for a given sequence as a number of factors are concerned (cations, additives, crowding agents, state of aggregation) [61, 74, 194, 195]. For example, the well studied $AG_3(TTAGGG)_3$ sequence shows a basket topology in sodium solutions [178], hybrids conformations in potassium solutions [191, 192, 194, 196] and a parallel arrangement both in crystals and in molecular crowded solutions [182, 197], especially where polyethylene glycol has been added [198]. As a matter of fact, G-quadruplexes undergo conformational equilibria and a mixture of conformers are usually found in solution [192, 193, 199], although single configurations can be stabilized, as by base modifications [200]. Consequently, it is very challenging to establish which are the "biological relevant" G-quadruplex topologies for human telomeric sequences.

Nevertheless, recent studies have tried to point out which human quadruplex topologies might more probably be found at the telomere. Interesting results have been obtained by using in-cell NMR spectroscopy on isotopically labelled human telomeric se-

quences injected in Xenopus laevis oocytes, that are eukaryotic cells of reference [201]. In more detail, high resolution information about the G4 structure is usually attained by performing *ex vivo* NMR experiments on Xenopus laevis egg extracts. Human telomeric sequences in such environments are found to adopt a mixture of conformations in equilibrium that are similar to those present in dilute conditions [177]. The preferential topology seems to be hybrid-1, although extensive data on a number of labelled sequences are still lacking [202].

On the other hand, telomeric oligonucleotides have been studied in presence of crowding agents able to simulate cell-mimicking conditions characterized by a large excluded volume and a lower water activity [74]. Polyethylene glycol (PEG) is probably the most common chemical used for this purpose, however it has been reported how PEG and other agents selectively stabilize the parallel topology because of the significant dehydration [177, 198]. The parallel topology seems to be thermodynamically unfavoured in living cells, although more conclusive data are missing.

Telomeres are composed of thousands of TTAGGG tandem repeats, thus they could form in principle multiple G-quadruplexes at the same time, also referred to multimers G4 structures or higher order G4 structures. Studies on long human telomeric sequences, $(TTAGGG)_n$ with n=4-20, have in fact shown the formation of tandem tetraplexes, usually made of three guanine tetrads, by means of various techniques including gel electrophoresis, analytical ultracentrifugation, atomic force microscopy and electrospray ionization mass spectrometry (ESI-MS) [166, 203, 204]. The mutual interaction between different G4 structures is still unclear (figure 19) [56, 204]. A common idea concerns the substantial independence of the foldings and predicts a "beads-on-a-string" organization. An alternative proposal consists, conversely, in the formation of assemblies of G-tetraplexes interacting by means of their

loops or their G-tetrads giving rise to different architectures (e.g. either 5′-5′, 5′-3′ or 3′-3′ end stacking). Parallel G-quadruplexes could efficiently give rise to such organizations, as observed in molecular crystals [205]. Clearly, it is feasible that human telomeric DNA could be arranged according to a combination of the two models.

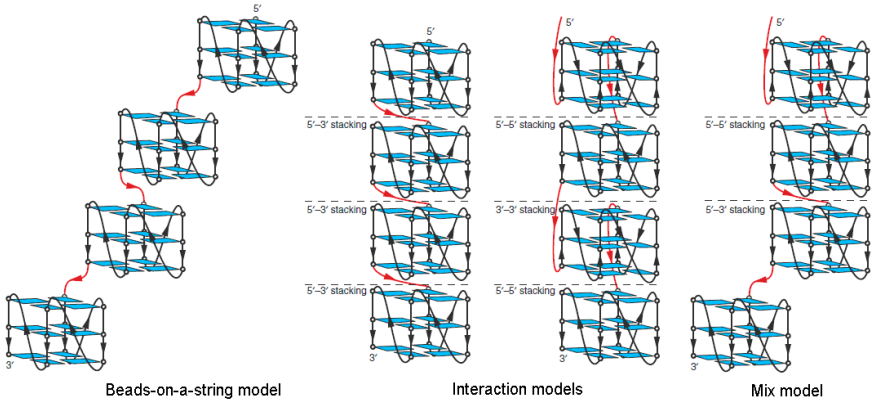


Figure 19: Main hypotheses for the organization of tandem G-quadruplexes at the telomere. Adapted from reference [56].

Recent studies have indicated the beads-on-a-string model to be the most reliable organization of human telomeric DNA. In-cell NMR experiments have pointed out the formation of antiparallel-2 and hybrid-2 structures in long telomeric sequences, thus showing a quite different behaviour with respect to short telomeric sequences forming single tetraplexes [202]. A slight different stability for terminal and inner G_4 structures within long sequences has also been evidenced. Lately, the formation of adjacent quadruplexes in long strands has been demonstrated to be more energetically expensive than the folding of a tetraplex along with duplex tracts, as highlighted by gel electrophoretic, calorimetric and spectroscopic data [206]. As a consequence, long telomeric se-

quences do not tend to form quadruplexes in the maximum possible number, so confirming the different behaviour of long telomeric sequences in comparison to short ones. Recently, the formation of G4 structures in long sequences has been described using a combined gel electrophoretic and spectroscopic approach in terms of energetic differences associated to the localization in the sequence of the resulting non-canonical foldings [167]. In more detail, the stability of tetraplexes at the sequence termini appear to be fairly modulated by their flanking residues unlike that of inner G4 units which is less affected. Furthermore, the global stability of these higher order architectures is not influenced by the length of the nucleotide tracts separating adjacent G4 structures, thus the overall organization is thought to be actually composed of independent G4 units, that is the assumption at the basis of the beads-on-a-string model. The unfolding kinetics observed seem to confirm this finding [167]. Accordingly to previous results, a certain degree of interaction is however displayed in tandem G-quadruplexes, probably because of the different properties of the tracts connecting adjacent G4 units which can be either terminal, and so quite "free", or "shared", and so restricted in some way [167].

As the structure of long telomeric sequences is matter of intense debate, experiments are usually carried out on short sequences able to fold into a single G-quadruplex. In general, studies on DNA foldings are performed on model systems, that are short sequences able to fold into little but representative DNA structures. Model systems are really helpful in simulating the DNA properties with easy handling, both in vial and *in vitro*. Such model systems are extensively used by researchers, as for the search of ligands able to interact and stabilize DNA structures which is the issue of the following section.

1.2 SMALL MOLECULES AS DNA BINDING LIGANDS

Non-canonical DNA structures represent interesting targets for the development of potential therapeutic strategies for various diseases, especially cancer. Such structures might be the target of small molecules able to bind and stabilize them thus causing important biological responses [31, 207]. Despite the rising of successful immunotherapy treatments, chemotherapeutic agents are still crucial in fighting cancer. The growing information about the mechanisms underlying tumour origin and progress make the identification of small molecules of better drug-like properties easier and easier.

The idea of targeting DNA by small molecules has spread since the discovery of the biological activity of cis-platin and related complexes which are able to covalently bind to double helix DNA leading to alteration of local DNA structure and functions [208]. Such compounds are illustrative examples of highly cytotoxic chemicals lacking in selectivity for the recognition of malignant cells, thus taking to severe collateral effects [209–211]. In fact, covalent ligand binding involves alkylation, base modifications, strands cross-linking and strand breakage which result in mutations and loss of biological functions. For this reason, the non-covalent and reversible binding of small molecules to DNA is thought to be a promising route in order to enhance the efficiency of anticancer drugs as well as their selectivity towards cancer over healthy cells. Such requests could be fulfilled by targeting non-canonical DNA foldings which are involved in various biological processes that are often altered in tumours [5, 9, 29, 31].

G-quadruplexes are particularly interesting non-canonical DNA structures as they are most likely involved in the regulation of important biological processes [94]. The numerous clues collected

about their functions in living cells support this hypothesis. The idea of targeting G-quadruplexes by means of small molecules has already brought interesting results which have confirmed the correlation between strong G-quadruplex binding and relevant anticancer activity for a number of ligands [86, 212].
Telomeric G4 targeting is thought to be a much valuable route for the design of potential anticancer drugs. The stabilization of these structures could indeed result in the inhibition of telomerase, an enzyme normally absent in healthy cells but over-expressed in about 85% of tumours and involved in the immortalization of cancer cells [86, 113]. This therapeutic perspective might lead to the identification of anticancer agents effective against a wide range of cancers, in contrast to classical chemotherapeutics which usually are compelling with a limited number of pathologies. For these reasons, huge efforts have been made in this direction as confirmed by the massive literature available.
The targeting of human telomeric G-quadruplexes by small molecules covers most of the research activity results, hence the thesis has a particular focus on this topic. The concerns addressed for telomeric DNA binders can however be extended to the targeting of alternative non-canonical DNA foldings, in particular to different G4 foldings as those formed at the guanine-rich tracts of proto-oncogenes.
Desirable G4 ligands should show selective recognition of telomeric tetraplexes over other G4 foldings and in general over alternative DNA arrangements, as the ubiquitous double helix [213, 214]. These demands can be met by taking into account the structural diversity of the secondary organizations of DNA. Despite the quadruplex/duplex selectivity can be achieved quite easily, the favoured interaction of ligands with telomeric tetraplexes over e.g. proto-oncogene G4 structures is still challenging [86]. In order to better understand the problems related to the issue, the char-

acteristics of reported G-quadruplex binding ligands should be discussed. Small molecules generally bind G4 structures through five main modes:

- Stacking on external G-tetrads
- Involvement in G-tetrads
- Groove binding
- Loops binding
- Loops/G-tetrad binding

The stacking on external quartets, the involvement in G-tetrads and the quartet-loops binding involve the formation of π-π stacking interactions which are favoured for ligands showing diffused aromatic surfaces. Groove binding and loops binding usually require positively charged groups able to interact with the DNA backbone. Hydrogen bonds occurred between ligands and targets could further strengthen the stability of the adducts. In general, small molecules able to strongly bind telomeric G-quadruplexes should feature different functional groups able to develop such non-covalent interactions [86, 213].
The most common binding mode is the stacking on external G-tetrads, as confirmed by the great number of reported G4 ligands showing extended planar aromatic conjugation. As regards this point, the presence of positive charges or at least electron-deficient aromatic (e.g. heteroaromatic) moieties enhance the overall interaction [86, 213]. No intercalation has ever been described, probably because of the almost incomparable stability associated to the G-tetrads stacking. This observation prompted some authors to design a molecule able to assembly in an artificial G-tetrad able to induce and bind quadruplex foldings [215]. The involvement of ligands in the G-tetrad formation is a quite rare

case and only peculiar chemicals as peptide nucleic acids (PNA) show this behaviour [216]. The idea of forming even artificial G-quadruplexes able to inhibit telomerase have led to guanine-rich oligonucleotides (GROs) characterized by telomeric repeats [176, 217, 218].

Overall, the selective quadruplex over duplex recognition is usually achieved by exploiting the much greater aromatic surface featured by tetraplexes which can involve the external G-tetrads into strong π-π stacking interactions with ligand molecules carrying extended aromatic surfaces (interactions of several tens of kJ mol^{-1}) [86, 213, 219, 220]. In addition, G-quadruplexes also differ from double helices in the characteristics of their grooves and in the presence of loops, thus ligands carrying groups able to interact with grooves and loops might increase the quadruplex over duplex selectivity [86, 213, 214, 221]. Notable examples in this regard are BRACO-19 and PIPER (figure 20) characterized by diffused aromatic moieties able to stack on the external G-tetrads and cationic side-arms which can bind to the grooves of G-quadruplexes [222, 223]. On the other hand, interesting G4 groove binders have also been reported (figure 20), such as a Dystamicin A derivative [224, 225] and the crescent-shaped small molecule TOxaPy (figure 20), which is able to specifically interact with the antiparallel-1 basket topology [226].

The recognition of telomeric G-quadruplexes over other tetraplex arrangements is more challenging as the ligand-G4 binding usually involved non-specific hydrophobic interactions, although slight differences in the quartet surface area, nature of loops and grooves allow the design of, at least partially, specific ligands [86, 227]. Despite the developments of various approaches, it seems almost impossible to identify G4 binders able to selectively interact with certain sequences, considering that less than 1% of putative G-tetraplexes in the genome has been described and also that it is

most likely that small molecules might target multiple quadruplexes in different genes at the same time, so exerting a wide complex activity, especially in presence of deficient repairing mechanisms [86]. An illustrative example in this regard is the polycyclic perylene derivative Emicoron (figure 21) which stabilizes telomeric as well as c-Myc and Bcl-2 proto-oncogene G4 structures, showing an interesting anticancer activity *in vivo* [228].

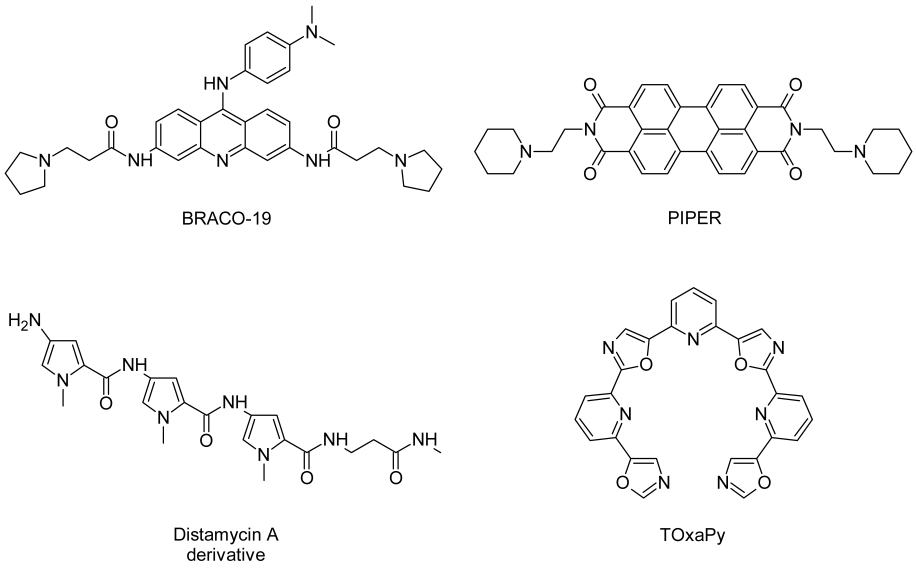


Figure 20: Representative examples of ligands able to stack on external G-tetrads (BRACO-19, PIPER) and to bind to the G-quadruplex grooves (Dystamicin A derivative, TOxaPy).

Nevertheless, a viable route for the identification of ligands selective for telomeric G4 could be the targeting of higher order organizations at the telomeres which feature potential binding sites at the quadruplex-quadruplex junctions most likely unique in the genome [204, 229]. Tandem G-quadruplex formations might be the actual biological targets of small molecules and their study is thus essential, as confirmed from the different binding properties

of some compounds in presence of putative sequences forming single and tandem G4 structures [229]. Up to now, a few compounds have been reported to bind tandem human telomeric G-quadruplexes, even with interesting selectivity for the interaction with tandem than with single G4 [230–235].
The search of telomeric G-quadruplex ligands acting as potential anticancer drugs has produced interesting results in the last decade [86, 175, 176, 236, 237]. Molecular scaffolds as quinoline, acridine, anthraquinone, phenanthroline, porphyrine and perylene have led to the design of binders able to strongly interact especially with the external G-tetrads. Phen-DC3, RHPS4 (active *in vivo*), Quarfloxin (active *in vivo* and currently under phase 2 clinical trials), TMPyP4, Pyridostatin, BRACO-19 (active *in vivo*) and PIPER are notable examples in this regard (figure 20,21). Natural molecules as Telomestatin and Berberine have also shown valuable G-quadruplex binding properties and telomerase inhibition activity. Through the binding of telomeric G4 structures, Telomestatin (active *in vivo*), Pyridostatin and RHPS4 trigger DNA damage response (DDR) pathways, especially in combination with PARP1 and WRN inhibitors [86].

Biological evaluations have suggested that the inhibition of telomerase through G4 ligand binding occurs along with the alteration of telomere structure and function [86, 175]. In fact, the binding is followed by telomere uncapping by a number of proteins which provide its maintenance and biological activity. As a result, repairing processes are triggered, but due to deficiencies in cancer cells (e.g. mutations in DNA repair proteins BRCA1 and BRCA2), damages are not often restored. People often refer to telomere-induced senescence and apoptosis where the inhibition of telomerase is just a part of more complex processes of response. Beyond the research underlying the identification of potential anticancer drugs, G4 interacting small molecules are also used as flu-

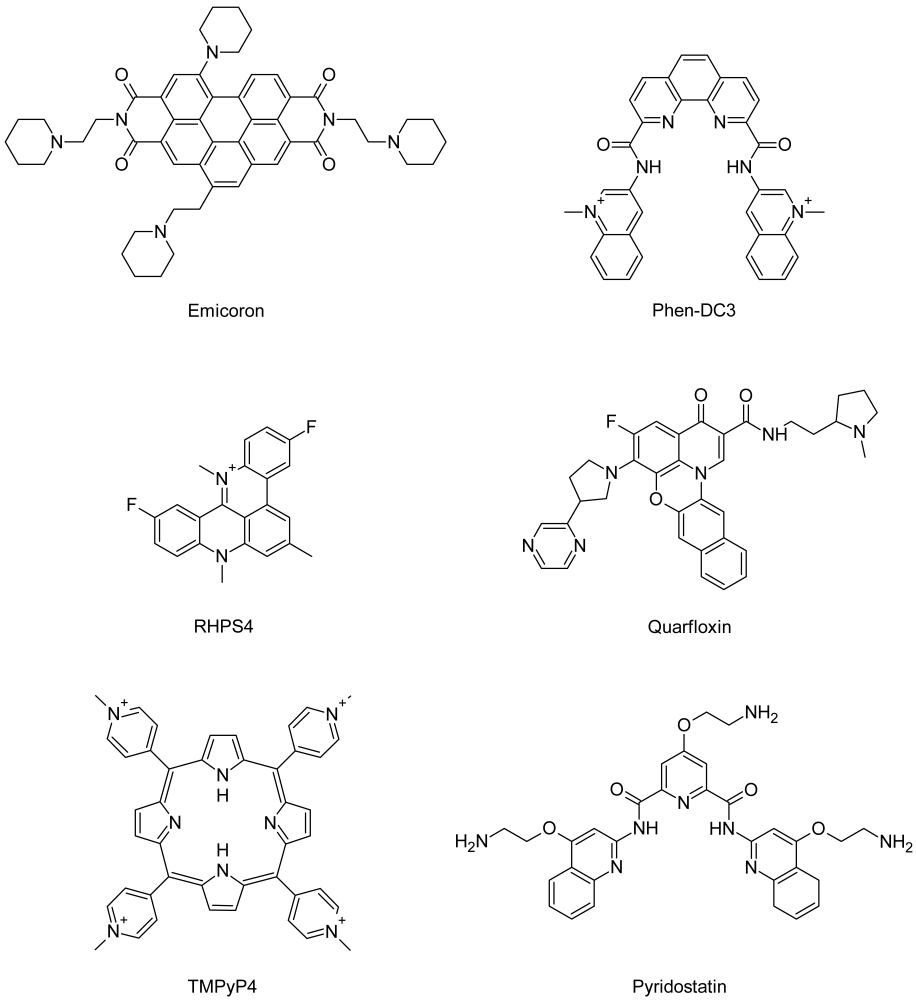


Figure 21: Remarkable examples of G4 binding ligands.

orescent probes in order to investigate the nature of G-quadruplex structures in living cells [238–240]. Such compounds possess chromophores with useful optical properties for imaging. Metal complexes as those of ruthenium are particular investigated in this regard [239, 241, 242]. Bioimaging experiments exploiting fluorescent molecules have confirmed the actual existence of G-quadruplexes *in vitro*, their location in the chromosomes, and also their increasing number upon addition of G4 ligands [133, 134]. One of the main task of the fluorescent probes is to distinguish among different DNA foldings and different strategies have been developed to meet this issue. Antibodies able to selectively recognize G4 structures as well as selective small molecule binders are used. For example, an isoquinoline-Ir^{III} complex has displayed good optical properties together with significant affinity and selectivity for G-quadruplex over other foldings due to the isoquinoline-based groove binder [243].

1.3 RESEARCH AIMS

The research activity carried out during the PhD years has concerned the investigation of ligands able to interact and stabilize non-canonical DNA foldings, in particular human telomeric G-quadruplexes, in order to select valuable candidates in anticancer drug discovery. The achievement of detailed structural data about the interaction of small molecules and target structures is crucial for the design of efficient DNA binders and for a better understanding of the binding properties of the targets themselves. X-ray crystallography has been the main technique used along the years to shed light on the interactions in ligand-DNA adducts, due to the accuracy of results. Complementary information have been obtained from spectroscopic characterization and *in silico* simula-

tions performed on the adducts. Additional biological evaluations have been performed by collaborators in order to verify the telomerase inhibition activity for selected compounds.
The choice of the studied ligands has been usually motivated on the basis of previously reported biological activities of the compounds along with the interest for the investigation of the binding properties of attractive scaffolds. Both natural and synthetic compounds have been considered for the work.

1.3.1 Natural compounds

Natural products are a fundamental source of medicines since millennia. Despite the advent of synthetic drugs, they are still important in drug discovery because of their large chemical diversity, wide range of pharmacological activities and valuable pharmacokinetic properties [244–246]. The attention has been focused on some natural benzylisoquinoline alkaloids extracted from plants of the traditional Ayurvedic and Chinese medicines, namely Berberine, Coptisine, Chelerythrine and Sanguinarine (figure 22), which can affect various biological processes and show promising anticancer activity [247–249].

As shown in figure 22, the structurally related pairs Berberine/Coptisine and Chelerythrine/Sanguinarine differ just in the number of dioxolo rings and in fact their biological activities are quite similar. Berberine shows anti-microbial, anti-inflammatory, anti-arrhythmic, anti-ischemic, cholesterol-lowering and anti-hyperlipidemic properties and it is currently used for the treatment of type-2 diabetes mellitus and of digestive problems caused by bacterial infections [250, 251]. It has also shown neuroprotective effects thought to be useful for the treatment of Alzheimer's and Parkinson's diseases and also of multiple sclerosis [252, 253]. Coptisine displays comparable activities [254–257]. Sanguinarine has been

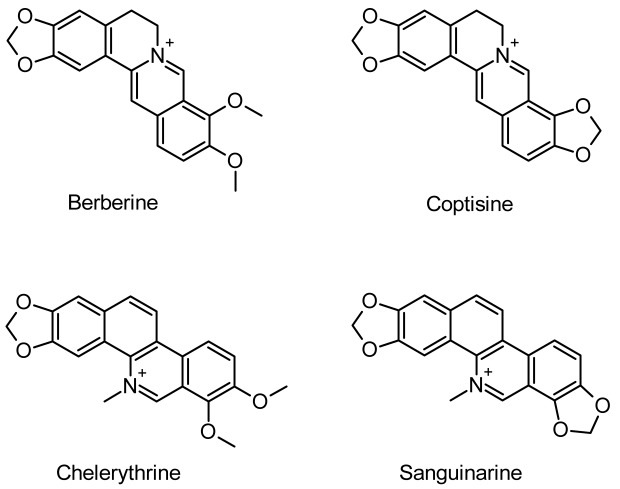


Figure 22: Natural alkaloids considered for the research activity underlying the thesis.

mainly known for its anti-microbial, anti-inflammatory, neuroprotective and cardiotonic properties as well as Chelerythrine which is a potent and selective inhibitor of protein kinase C [258–261].
The alkaloids show promising anticancer activity through the inhibition of cell proliferation in tumours and the induction of cell cycle arrest [262–269]. The anticancer properties observed for the compounds have been associated to DNA and RNA binding as well as the interaction with proteins, including telomerase, DNA topoisomerase and p53, overall leading to DNA damage, chromatin alteration, autophagy and apoptosis [250, 251, 267, 270–274]. In addition, Berberine can downregulate gene transcription by inhibiting the association between the TATA binding protein and the TATA box in gene promoters [275].
Several works have highlighted the interaction of the alkaloids with non-canonical DNA foldings, in particular with G-quadruplexes formed by telomeric and oncogene promoter sequences

[276–283], though scarce data have been reported for Coptisine [284, 285]. The compounds show good features for G4 binding because of the extended aromatic surfaces, hydrogen bond acceptors and also positively charged nitrogen atoms. In fact, Berberine, Chelerythrine and Sanguinarine (as in iminium form, pKa=7.4) can strongly bind to G-quadruplex foldings, in particular to human telomeric G4 structures [286]. The alkaloids are able to stack on the external G-tetrads although their binding with grooves and loops have also been reported in the case of G4 structures of hindered quartets by caps and loops, as for the antiparallel basket-1 human telomeric G-quadruplex [286–288]. The ability to bind grooves and loops has been evidenced in particular for Sanguinarine and Chelerythrine which can bind multiple G-quadruplex structures at the junctions, respectively, of tandem tetraplexes from long human telomeric sequences and of ligand-promoted human telomeric aggregates [230, 282]. Berberine and Sanguinarine have shown interesting telomerase inhibition activity along with the inhibition of proto-oncogenes, as also observed for Chelerythrine that is able to downregulate VEGF, Bcl-2 and K-ras expressions [280, 286, 289–292]. The alkaloids are thought to interact also with RNA G-quadruplex structures as recently reported for the binding of Chelerythrine to human telomeric RNA which has been demonstrated to be even stronger with respect to the binding to human telomeric DNA [281].

The aim of the research activity was to supply a more comprehensive analysis of the interactions of the alkaloids with non-canonical DNA foldings, with particular attention on how slight structural differences, such as the number of dioxolo rings, can affect the binding properties of the compounds. As the molecules behave as interesting telomeric G4 binders, major efforts were devoted to the study of the adducts formed with human telomeric sequences. X-ray crystallography, spectrophotometric characterization (thanks

to a collaboration with Professor Claudia Sissi from the University of Padua, Italy) and *in silico* calculations were used for this purpose. On the other hand, crystallization screenings were performed for the compounds bounded to short DNA oligonucleotides which are able to give rise to local non-canonical arrangements, as previous crystallization trials for telomeric DNA adducts yielded suitable crystals for X-ray diffraction analysis only in the case of Berberine.

1.3.2 Berberine derivatives

As natural products display attractive properties, semi-synthetic modifications are alluring in order to further improve their biological activity and decrease eventual side effects. The alkaloids just described have shown promising activities and many attempts have in particular been reported for modifications on the Berberine skeleton. As regards the binding of DNA structures, the modifications on the Berberine scaffold have largely involved the C13 and C9 positions (figure 23) [293]. Various groups have been added to these positions through linkers, even additional Berberine molecules [294], and successful improvements have been observed in dependence of the linker length.
Derivatives of Berberine on the 9-position and 13-position have shown enhanced DNA binding properties associated to bettered anticancer activities concerning telomerase inhibition and poisoning of DNA topoisomerases [267, 293, 295–299]. The inhibition of telomerase has been associated to strong binding of the derivatives to human telomeric G-quadruplexes, as supported by the good quadruplex/duplex selectivity displayed by various derivatives. In fact, it has been suggested that additional groups could interact with grooves, especially when carrying positive charges, and so increasing the selectivity for the interaction with telomeric

G-quadruplexes over different DNA foldings.
The interest in Berberine derivatives has led to the investigation of a series of analogs featuring phenylalkyl, diphenylalkyl and pyridinealkyl moieties on the 13-position (figure 23), synthesized by our collaboration partner Naxospharma srl, pharmaceutical industry in Milan (Italy) [300]. The addition of aromatic groups to such position through alkyl linkers may strengthen the binding to human telomeric G-quadruplexes. Also, the presence of positively charged nitrogen atoms in the pyridil derivatives may play an important role in establishing the ligand-DNA interactions.

Some of the derivatives studied have shown improved cytotoxic activities against human colon and breast cancer cell lines with respect to the natural precursor, overall leading to DNA damage, cell cycle arrest, autophagy and apoptosis [301–306]. Similarly, diphenylalkyl derivatives behave as catalytic inhibitors of topoisomerases, presumably through the binding to DNA, with better inhibition activity with respect to Berberine [307]. Overall, the binding to DNA is supposed to be the most likely responsible for the anticancer activity displayed by the compounds.
The research objective was to provide a general picture of the interaction of the derivatives with human telomeric sequences, in particular by highlighting how the additional groups affect the G4 binding properties with respect to the natural precursor Berberine. The work was primarily carried out by means of X-ray crystallography. While the investigation of pyridil analogs is still in progress, the analysis of the interaction of phenylalkyl and diphenylalkyl modified Berberine was also extended to spectrophotometric characterization and, thanks to a collaboration with Doctor Ivana Scovassi from the CNR in Pavia (Italy), to evaluations of the cytotoxic activity of the derivatives on cancer cells lines.

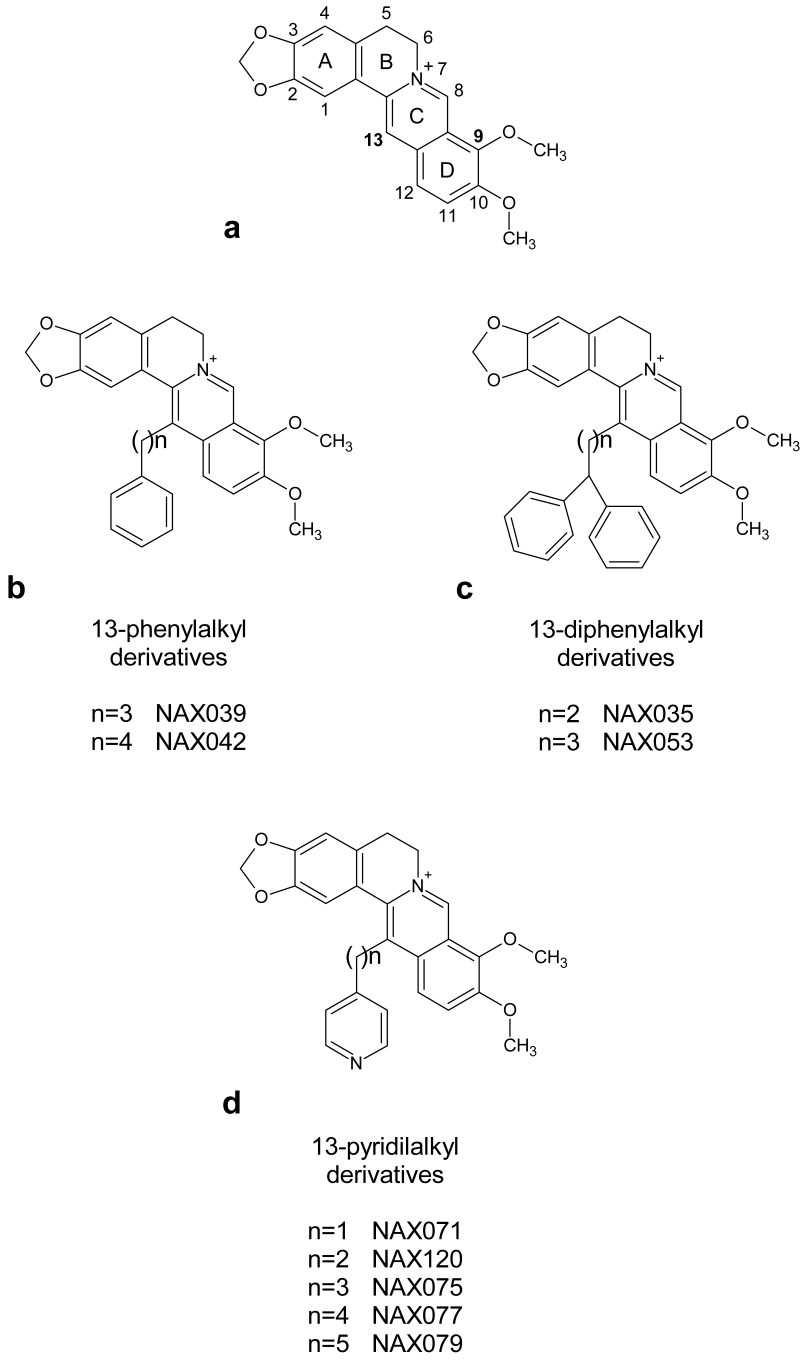


Figure 23: Natural alkaloid Berberine (a) and the semi-synthetic derivatives (b,c,d) studied as human telomeric G-quadruplex binders.

1.3.3 Au^{I} complexes

Since the discovery of the anticancer activity of cis-platin, huge efforts have been made in order to identify metal-based drugs with better anti-tumour properties, reduced side effects and lowered drug resistance. In addition, the chemistry of metals offer very interesting opportunities to tune structural and electronic characteristics, exploiting unique optical, magnetic and catalytic properties [239, 241, 242]. In fact, metal complexes have been largely investigated both as potential drugs and fluorescent probes. The research activity during the PhD program has focused its attention on different metal compounds able to interact with non-canonical DNA foldings.

In the first place, gold complexes were investigated because of their promising anticancer properties also effective against cancer cells resistant to cis-platin together with reduced drawbacks [308–310]. Because of the remarkable oxidizing properties of Au^{I} and Au^{III}, gold compounds should be stable in physiological conditions in order to avoid the non-selective oxidation of biological species [308–310]. Hence, suitable organic ligands have to be taken into account. A useful approach is the stabilization of such gold oxidation states with the creation carbon-gold direct bonds, as by using N-heterocyclic carbene (NHC) scaffolds [311]. Several gold NHC complexes have shown significant anticancer activity by interactions with multiple targets including DNA and various proteins [309, 312, 313]. The cytotoxicity displayed is in particular associated to mitochondrial damage through the inhibition of thioredoxin reductases, a class of seleno-enzymes involved in the growth and proliferation of malignant cells [309, 313]. Furthermore, the binding to DNA seems to be responsible for the inhibition of the telomerase enzyme.

Caffeine-based NHCs bound to two-coordinated Au^{I} (AuNHC **I**, figure 24) were designed with the purpose of favouring the interac-

tion of a flat, aromatic and positively charged compound with the external tetrads of G-quadruplexes [314]. The choice of such NHC scaffold was motivated by the close similarity between guanine and caffeine together with the relatively low cost of synthesis. The AuNHC **I** complex, $[Au(9\text{-methylcaffein-8-ylidene})_2]^+$, has shown indeed quite strong binding to quadruplex foldings with high selectivity over alternative DNA structures, including duplex DNA [314]. The compound displays interesting anti-proliferative activity *in vitro* against human ovarian cancer cell lines with modest side effects on healthy cells [315, 316]. The biological properties of AuNHC **I** are presumably associated to G4 binding.
Similar biological activities have been observed for AuNHC **II** (figure 24), $[Au(bis(1\text{-butyl-}3\text{-methyl-imidazole-}2\text{-ylidene}))]^+$, which is characterized by reduced aromatic surface and bulky alkyl side groups. The compound is cytotoxic *in vitro* against ovarian cancer cells resistant to cis-platin and its mechanism of action seems to rely on metalation of metalloproteins as the copper chaperone Atox1 [316, 317]. Multiple targets have however been hypothesized for the compound, including DNA binding.

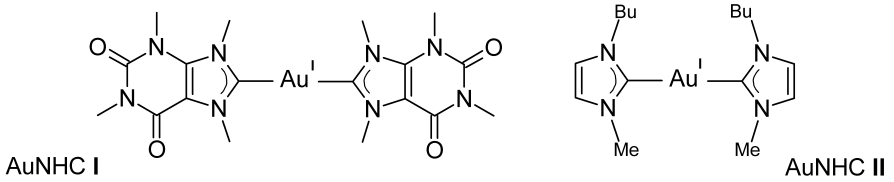


Figure 24: Au^I metal complexes studied as G4 stabilizing ligands.

Overall, AuNHC **I** and AuNHC **II** display good features for the binding to human telomeric sequences and the former, especially, has shown promising anticancer activity accompanied by a very good quadruplex over duplex selectivity. During the PhD years, the interaction of these Au^I complexes with telomeric G-quadruplex was more deeply investigated. X-ray crystallographic

studies were carried in order to shed light on the formation ligand-DNA adducts. Additional solution studies performed on AuNHC **I** revealed the interesting binding behaviour of the compound which was explained in more detail by *in silico* calculations. The telomerase inhibition activity of the compound was also evaluated, thanks to a collaboration with Prof. Donato Colangelo from the Università del Piemonte Orientale (Italy).

1.3.4 Au^{III} complex Auoxo6

Four-coordinated Au^{III} complexes display a more pronounced oxidative character and their development was in fact initially motivated to the identification of potential alternatives for cis-platin resistant tumours [310]. Complexes stable in cellular conditions have lately been studied as anticancer drugs acting through the selective inhibition of thioredoxin reductases [310, 318, 319]. This is the case of the square planar Auoxo6 complex (figure 25), $[(bipy^{2Me})_2 Au_2(\mu\text{-}O)_2]^{2+}$ ($bipy^{2Me}$=6,6′-dimethyl-2,2′-bipyridine). Auoxo6 is a dinuclear gold complex that exerts promising anticancer activity thought to be achieved by interaction with proteins and DNA, accompanied by oxidative damages due to reduced stability in physiological conditions [310, 320, 321]. The complex can strongly bind and induce human telomeric G-quadruplex structures whereas it weakly interacts with double helix DNA [322].

Auoxo6 has shown promising telomeric G-quadruplex binding properties accompanied by significant anticancer activity, possibly connected to telomerase inhibition. Detailed structural information about the ligand-G4 interaction could be useful for a better understanding of the binding behaviour of the compound. For this reason, crystallization screenings were performed but unfortunately any suitable crystal for X-ray diffraction analysis was yielded. Thus, a combined NMR and *in silico* approach has been

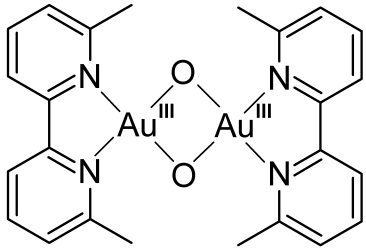


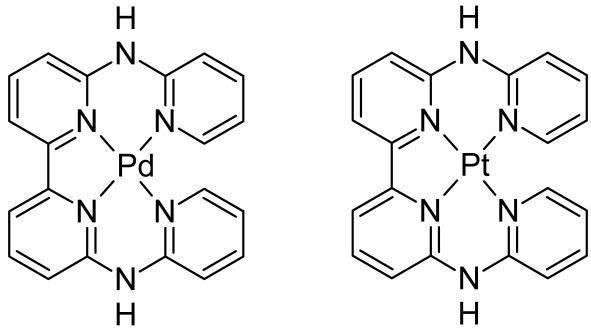
Figure 25: The Au^{III} complex Auoxo6 studied as G-quadruplex stabilizing ligand.

carried out on the adducts formed by Auoxo6 with human telomeric sequences, in collaboration with Prof. Harald Schwalbe from the University of Frankfurt (Germany).

1.3.5 Palladium and platinum complexes

Aside from gold compounds, alternative metal compounds have also been taken into account. Platinum and palladium are ideal metals for the formation of positively charged square planar complexes able to interact with non-canonical DNA foldings. Recently, $[Pd(bapbpy)]^{2+}$ and $[Pt(bapbpy)]^{2+}$ complexes (figure 26) (bapbpy=6,6′-bis(2-aminopyridyl)-2,2′-bipyridine) have been developed in Prof. Sylvestre Bonnet group from Leiden University and they have shown very interesting anticancer properties associated to the disruption of mitochondrial functions. Despite the biological mechanisms of the complexes have not been clarified yet, it is likely that the compounds bind to DNA. Nevertheless, because of the intention of obtain a patent, more detailed information can not be provided here. Anyway, structural studies about the interaction of $[Pt(bapbpy)]^{2+}$ and $[Pd(bapbpy)]^{2+}$ with non-canonical DNA structures could be useful to give a molecular explanation for the biological behaviour observed. Crystallization screenings were

therefore carried out for the adducts formed by the metal complexes and human telomeric sequences as well as short oligonucleotides able to give rise to local non-canonical DNA arrangements.



Pd(bapbpy) Pt(bapbpy)

Figure 26: $[Pd(bapbpy)]^{2+}$ and $[Pt(bapbpy)]^{2+}$ complexes investigated as non-canonical DNA binders.

2 METHODS

2.1 X-RAY CRYSTALLOGRAPHY

X-ray crystallography is the major technique able to determine the structure of a given system at atomic level. The great limitation of the method is associated to the crystalline state of the samples which could be not merely indicative of the behaviour of the studied system in solution. Furthermore, the picture provided by this technique is essentially static. Anyway, crystallography provides unique information about intermolecular and intramolecular interactions which are quite useful for the study of ligand-DNA adducts.
Obtaining suitable crystals for X-ray diffraction analysis is usually the most challenging step, especially in the case of DNA samples. Following the collection of X-ray reflections, diffraction pattern are analysed and structure factors calculated. Several approaches have been developed to solve the phase problem and supply good initial models of the crystallized system to be refined. In the following sections, the different stages of the process of solving the structure of a biomacromolecule are described.

Crystallization

A crystal is by definition an ordered and periodic lattice of physical entities. Such ordered phase is formed by species involved in intermolecular interactions leading to spontaneous self assem-

bly. This phenomenon is observed from supersaturated solutions where stable crystal nuclei can form. Otherwise also aggregates and amorphous precipitates are observed (figure 27).

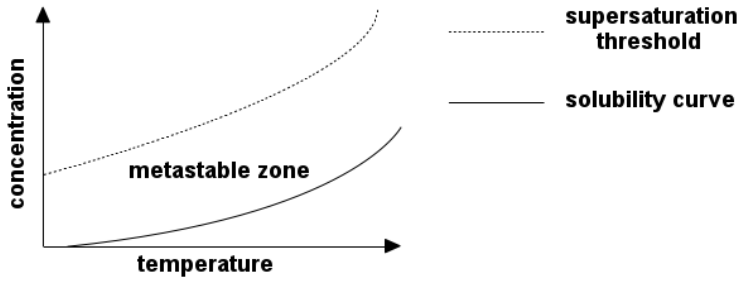


Figure 27: Dependence of macromolecule solubility on temperature.

The crystallinity of the aggregates is influenced by many factors including temperature, concentration of crystallizing species, pH, ionic strength and additives [323, 324]. In general, the quality of the crystals is dependent also by the time of crystal growth, as more ordered aggregates are usually observed for longer crystallization times [324].
The crystallization of macromolecules is a challenging task because of their scarce tendency to spontaneously crystallize. The supersaturation is achieved by the addition of a substance, generally called precipitant, to the acqueous solution of the macromolecules at controlled pH. Crystal screenings are generally performed for the identification of good crystallization conditions. Such screenings concern the testing of a number of solutions which differ in the content of precipitants, salts, buffer and additives [323]. In the crystallization of biomacromolecules, the most used precipitant agents are salts like ammonium sulphate, polyethylene glycol (PEG) of different molecular weights and organic molecules such as ethanol and MPD, 2-methyl-2,4-pentanediol (figure 28). Crystallization screenings are often based on sparse matrices where

the components of the solutions are varied on the basis of the available data on the crystallization of similar molecular systems. The most promising crystallization conditions can be optimized in order to obtain crystals of better quality. The dimensions of macromolecular crystals are generally in the range 0.1-0.5 mm.

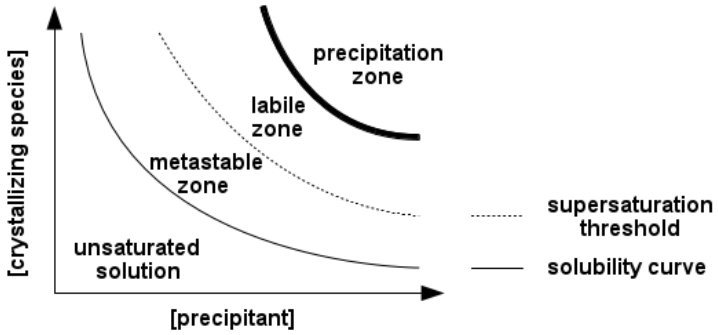


Figure 28: Dependence of macromolecule solubility on precipitants concentration.

Different methods are used for performing crystallization trials [323, 324]. The most common is the vapour diffusion method which is carried out on a drop containing the macromolecule equilibrated with a reservoir. The process is favoured by the lower content of precipitant agents in the drop with respect to the reservoir which cause a significant evaporation of the solvent from the drop thus possibly inducing the supersaturation. The little amounts of macromolecule requested to perform a great number of crystallization trials have made the method widely used. As shown in figure 29, a few variants of the vapour diffusion method have been developed with particular reference to the position of the drop in relation to the reservoir.

Alternative crystallization methods are microdialysis, microbatch and the free interface diffusion method [323, 324]. Microdialysis is based on the diffusion of low molecular weight compounds from

Figure 29: Variants of the vapour diffusion method: sitting drop (drop on a support), hanging drop (drop on a glass slide) and sandwich drop (drop between two surfaces).

the solution containing the species to crystallize to a reservoir thanks to a semipermeable membrane. Microbatch crystallization is conversely carried out by mixing the species with a solution rich in precipitants which immediately induces the supersaturation. A little amount of inert oil, usually paraffin, is added on the solution in order to prevent from rapid dehydration. On the contrary, the free interface diffusion method is performed by using capillaries in which the migration of solvent from the solution containing the species to crystallize to a phase rich in precipitant agents takes place.

Cryoprotecting solutions

X-ray diffraction experiments concern the structural investigation of matter at atomic level. The usage of ionizing radiation however promotes the formation of radicals which can disrupt the chemical structure of the sample, resulting in damage to the crystals. In order to minimize the diffraction decay and improve the overall diffraction data quality, the experiments are usually carried out on samples brought to low temperature, namely 100 K for macromolecular crystals. This choice is also useful because the withdrawal of samples from the crystallization wells could take to

a rapid degradation of crystals due to loss of solvent. This process is particularly relevant for macromolecular crystals that are composed at 40-60% v/v by water [323]. X-ray diffraction analyses at low temperature are nowadays widely performed along with the use of cryoprotecting solutions [325].
Cryo solutions have the same composition of mother solutions where crystals formed excepting for the addition of viscous chemicals like glycerol, PEG400, MPD or ethylene glycol, or a higher concentration of these compounds in the cases where they are used as precipitating agents. Crystals prior to be mounted for a diffraction experiment are picked up from the plates and soaked in a cryoprotecting solution. Raising the concentration of precipitants or viscous compounds promotes the formation of a glassy-like matrix at low temperature preventing the appearance of spurious peaks (Debye rings, actually) in the diffraction pattern due to the formation of ice crystals [323, 324].

X-ray diffraction analysis

The achievement of structural information by X-ray crystallography relies on the crystal diffraction phenomenon. This process takes place when electromagnetic waves of a given wavelength hit an atomic system and they are scattered into the entire solid angle by electronic clouds of atoms. The waves are scattered from every entity of the lattice and undergo interference. Because of the ordered nature of crystal lattices, electromagnetic radiations can be observed only at specific directions where positive interference occurs, as shown by the single crystal diffraction pattern reported in figure 30. The wavelength of the radiation used is usually in the range 0.4-2.5 Å, which falls in the X-ray spectral region [326]. These wavelengths are comparable to the inter-atomic dis-

tances, thus working as a probe to investigate the atomic structure of matter.

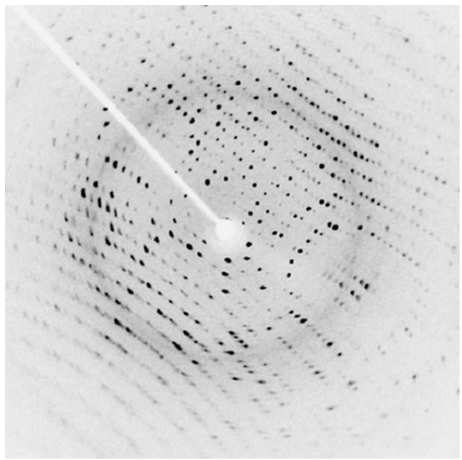


Figure 30: Representative X-ray diffraction pattern.

In order to better discuss the X-ray diffraction experiments performed, a few concepts have to be provided.
Firstly, the interaction of X-rays with the electronic clouds of atoms concerns different processes, including scattering and absorption [326]. The former is actually the phenomenon underlying diffraction and could be either coherent (Thomson) or incoherent (Compton), depending on the retain or the slight decrease in the radiation wavelength upon hitting atoms, respectively.
When the radiation of wavelength λ interacts coherently with the atoms, every entity of the lattice behaves as an emitter of waves featuring different phases as a consequence of different optical paths. A nice picture of X-rays scattering by crystal lattices is supplied by the Bragg law. The crystal lattice is imagined as cut by sheaves of parallel planes separated by a distance d (figure 31). The reflection of the waves by one of this family of parallel planes

leads to the radiation diffraction which depends on the angle of incidence θ:

$$n\lambda = 2d \sin\theta \tag{1}$$

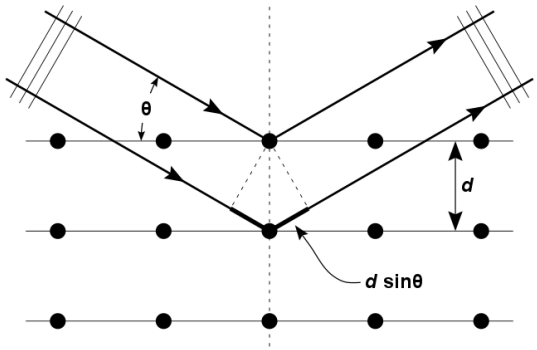


Figure 31: X-rays diffraction by parallel and equidistant crystal planes.

The Bragg law suggests that, keeping λ constant, X-rays diffracted by a given family of planes would be observed only at a determined angle. The maximum resolution achieved by diffraction experiments corresponds to the lowest inter-planar *d* spacing detected. The Bragg law is quite useful for the understanding of the diffraction analysis, although the features of the pattern observed are dependent on the nature of the entities that makes the crystal, on the crystal symmetry and on the crystal orientation with respect to the X-ray source.
A diffraction experiment is carried out by means of a goniometer, which is a device able to rotate into the entire solid angle in terms of either eulerian (θ,φ,χ) or kappa (ω,φ,κ) coordinates (figure 32). The sample is usually mounted on a loop which is placed onto the goniometer. In more detail, the loop is formed by a hollow microtube bounded to a nylon loop that is able to trap drops containing crystals. During the diffraction experiments, the

goniometer is rotated according to predefined ranges of the angular coordinates. The diffracted rays produce diffraction patterns which are recorded as images, usually at a bidimensional detector as the notable charge-coupled device (CCD).

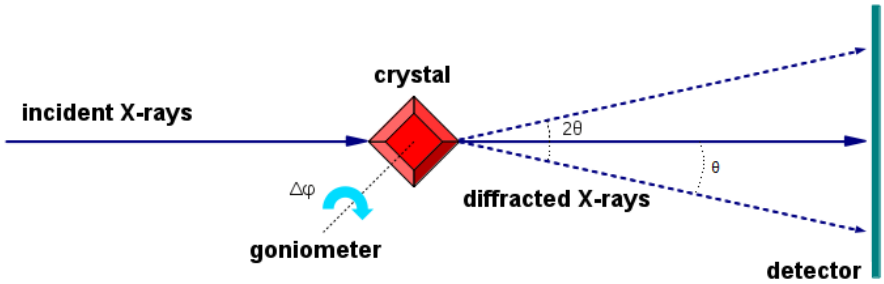


Figure 32: Typical setup of a single crystal X-rays diffraction experiment.

Diffraction data are analysed by specific software which evaluate the peaks intensity and determine the crystal symmetry and cell parameters through the indexing of reflections, their integration and scaling [326]. This analysis can be challenging, in particular for macromolecular crystals which are characterized by rather big cells (compared to those of small organic molecules) and give rise to a significant number of reflections. The collection of a large angular portion of the reciprocal space is particularly important in this field, as multiple independent acquisitions for each reflection are useful for the subsequent structural solution. The overall redundancy (multiplicity of independent measurements) and completeness (the number of unique reflections observed relative to the theoretical number) are important parameters. In general, multiplicity should be at least 2-3, while good completeness about 90-100% [323].
A useful parameter for the evaluation of diffraction data quality

is R_{sym} which compares the intensity of symmetry related reflections:

$$R_{sym}(I) = \frac{\sum_{hkl} \sum_i | I_i(hkl) - \overline{I(hkl)} |}{\sum_{hkl} \sum_i I_i(hkl)} \quad (2)$$

where the summations run over the i observations of a given reflection. $\overline{I(hkl)}$ is the average intensity of the i observations.
Other valuable indicators of diffraction data quality are the signal/noise ratio of reflections $I(hkl)/\sigma(I(hkl))$ and the Pearson's correlation coefficient $CC_{1/2}$ which allows a better estimation of the effective maximum resolution of data.

X-ray sources

X-rays sources deserve a special mention as regards the experimental apparatus used for the diffraction experiments. Conventional sources as sealed X-ray tubes and rotating anode tubes rely on the white light emission by a metal target upon incidence of accelerated electrons (30-60 kV) [326]. The light intensity is determined by the acceleration of electrons and by the nature of the metal that makes the target. In more detail, the higher the atomic number of the metal, the more intense is the resulting white light. Actually, the radiation emitted by the target is characterized by the specific spectral series of the metal, whose emission lines can be selected by a monochromator and used as X-ray probe of certain wavelength.
Unfortunately, conventional X-ray sources can not always afford valuable performances because macromolecular crystallography often requires highly brilliant radiations. Crystals of macromolecules display indeed scarce diffraction properties due to their low lattice order and reduced crystal dimensions. The much higher brilliance of synchrotron radiation with respect to those of tra-

ditional sources (table 4) has however led to significant developments in the field.

Table 4: Relative intensity and brilliance of X-ray sources [327]. Brilliances are reported in photons/mm^2/s.

X-ray source	Relative intensity	Approximate brilliance
Conventional sealed tube	1	$0.6 \cdot 10^9$
Microfocus sealed tube	6	$3.6 \cdot 10^9$
Early rotating anode	7	$4.4 \cdot 10^9$
Microfocus rotating anode	37	$2.2 \cdot 10^{10}$
Microfocus rotating anode/advanced optics	117	$7.0 \cdot 10^{10}$
High brilliance rotating anode/advanced optics	167	$1.0 \cdot 10^{11}$
Synchrotron (APS 14-ID-B)	3173	$1.9 \cdot 10^{12}$

Synchrotrons are accelerators of charged particles like electrons or positrons. They are composed of a circular storage ring in which the particles are injected from a linear accelerator or a booster synchrotron, and then accelerated by potent electric fields to near light speed while powerful magnets control the beam trajectory. The acceleration of charged particles causes the emission of electromagnetic radiations characterized by high brilliance, polarization, intensity, stability and low beam divergence. As the synchrotron radiation spectrum covers from microwaves to X-rays, monochromators are needed in order to select the desidered radiation wavelength. The broad spectrum is however useful because it allows to tune the wavelength acting only on the monochromators, as opposed to conventional sources which produce radiations of constant wavelength. Synchrotron have also favoured the feasibility of anomalous scattering experiments (see following section). The wavelength usually chosen is close to 1 Å because of the low radiation absorption along the optic path by the air and crystal itself [323].
X-ray diffraction experiments at synchrotron are usually carried out over the time of minutes, contrary to conventional sources which request several hours. Anyway, the usage of such highly

brilliant X-ray sources has renewed the problem of ionization damage which determines a rapid degradation of crystal, regardless various expedients such as the adoption of cryoprotecting solutions.

The phase problem

The analysis of diffraction patterns allows to shed light on the structure and the disposition of the physical entities in the crystal. The experimental diffraction images are indeed correlated to the electron densities which X-rays have encountered passing through the crystal. In order to properly discuss the issue, the amplitudes of the waves coherently scattered by a crystal unit cell of N atoms are mathematically described as structure factors F:

$$F = \sum_{n=1}^{N} f_n e^{2\pi i \boldsymbol{r}_n \cdot \boldsymbol{r}^*} \tag{3}$$

where f_n is the atomic scattering factor, which is approximately proportional to the atomic number of the n^{th} atom, while $\boldsymbol{r}_n$ indicates the position of the n^{th} atom with respect to the origin of the unit cell and $\boldsymbol{r}^* = \lambda^{-1}$ ($\mathbf{s}$ - $\mathbf{s}_0$) accounts for the scattering of the X-ray radiation of wavelength λ and direction associated to the unit vector $\mathbf{s}_0$ in the direction associated to the unit vector $\mathbf{s}$. The diffracted wave is so described by the structure factor which mathematically expresses the interference of the coherently scattered radiation by the N atoms of the cell through a summation of the contribution from each atom, that differs in amplitudes (f_n) and phases (phase difference equal to $2\pi \boldsymbol{r}_n \cdot \boldsymbol{r}^*$ for the scattering wave of the n^{th} atom with respect to the origin of the unit cell). The previous equation can be rewritten in terms of the Miller in-

dices *hkl* for a given family of planes which reflects X-rays in a certain direction:

$$F_{hkl} = \sum_{n=1}^{N} f_n e^{2\pi\imath(hx_n+ky_n+lz_n)} \tag{4}$$

Interestingly, $| F_{hkl} | = | F_{\text{-}h\text{-}k\text{-}l} |$ for coherent scattering processes in the absence of anomalous dispersion (Friedel law). The electronic density of the atoms in the unit cell of volume V can be obtained from the Fourier synthesis of F_{hkl} and is described by the continuous function $\rho(x, y, z)$:

$$\rho(x, y, z) = \frac{1}{V} \sum_{h=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} F_{hkl}\ e^{-2\pi\imath(hx+ky+lz)} \tag{5}$$

The square of the structure factor F_{hkl} is proportional to the intensity of diffracted waves, that is the physical quantity measured during the X-rays diffraction experiments. The phase problem arises from the loss of F_{hkl} phases for the diffracted waves, as the measurements only provide diffraction intensities that are real physical quantities.
Since the development of the theoretical description of crystal diffraction, huge efforts have been made in order to solve the phase problem by mathematical algorithms [326]. Several methods are nowadays widely used which are described as follows.
First of all, the Patterson function should be introduced as it is involved in the majority of phasing methods.

$$P(x, y, z) = \sum_{hkl} | F_{hkl} |^2\ e^{-2\pi\imath(hx+ky+lz)} \tag{6}$$

This mathematical construction concerns an inter-atomic vector map created by squaring the structure factor amplitudes and setting all phases to zero. The vector map contains a peak for each

pair of atoms in the unit cell, with the highest peak at (0,0,0) coordinates, the origin peak, which is the sum of the peaks relative to every atom^{nt}-atom^{nt} vector. The Patterson map is quite useful for the localization of heavy atoms in the crystal cell as the heavy atoms inter-atomic peaks are the highest since the intensity of each peak depends on the atomic number of the atoms involved.
Molecular replacement (MR) is the method which was applied for the solution of the crystal structures of the adducts studied in this thesis. The MR method allows the calculation of phases from coordinates previously reported for a homologous structure by using Patterson-based algorithms. Briefly, Patterson maps are derived from the data of the unknown structure and from the coordinates of a previously solved homologous structure, and then a search is carried out to achieve a maximum overlap between the two maps, corresponding to the correct orientation and position of the model within the unit cell. The procedure involves a rotation and a translation search of the initial model in the unit cell. In the rotation search, the unknown Patterson map is compared to the map obtained from the homologous structure in different orientations. The subsequent translation search concerns the correct positioning of the model in the asymmetric unit by moving the model, calculating a new Patterson map, and comparing it to the unknown-derived Patterson map. The translation search shows a high computational cost, hence fast translation functions are nowadays more commonly used. The MR output lists a number of solutions scored on the basis of parameters as the probabilistic scores TFZ and LLG as well as the number of clashes. Molecular replacement is the most commonly used method for the solution of the phase problem, in view also of the increasing number of macromolecular structures deposited in the PDB every year [323].
When the wavelength of X-rays is close to the absorption edge of an element, anomalous scattering can be observed. The energy

range of synchrotron X-rays radiations allows the investigation only of the anomalous scattering of heavy atoms. The synchrotron radiation is particularly useful for the collection of the anomalous signal as the X-ray wavelength can be easily tuned. Significant anomalous scattering takes to interesting features in the diffraction patterns, as being $|F_{hkl}| \neq |F_{-h-k-l}|$, thereby $|F_{hkl}| - |F_{-h-k-l}|$ differences can be used to overcome the phase problem through different approaches [326].

Single and multiple wavelength anomalous dispersion (SAD,MAD) methods concern the analysis of anomalous scattering caused by heavy atoms soaked or cocrystallized with the macromolecule. The different intensities associated to the so-called Bijvoet pairs $|F_{hkl}|, |F_{-h-k-l}|$ are exploited for the localization of heavy atoms in the structure and the structure factors phasing, thanks to Patterson-based algorithms [323]. SAD and MAD require just one crystal for the data collection and are hence competitive with respect to MR for the solution of macromolecular crystal structures.

Alternative phasing approaches rely on the obtaining of native and heavy atoms containing isomorphous crystals (single and multiple isomorphous replacements, SIR and MIR). The comparison of the diffraction patterns of the crystals allows the determination of the heavy atoms positions in the crystal cell and so the structure factors phasing. These methods are not much used nowadays, primarily because of the challenging production of isomorphous derivatives and the great developments in the application of the synchrotron radiation to macromolecular crystallography (i.e. SAD,MAD), although the combined uses of the isomorphous crystals approaches with anomalous scattering experiments (SIRAS,MIRAS) are still employed [323, 328].

Refinement and validation of crystal structures

The solution of the phase problem leads to an initial structural model of the crystal structure which has to be refined and validated before publication. The refinement goal is to improve the agreement between the model and the experimental diffraction data through the optimization of accord between observed and calculated structure factors, namely F_o and F_c. Commonly used parameters in this regard are R_{factor} and $R_{freefactor}$:

$$R_{factor} = R = \frac{\sum_{hkl} || F_o | - w | F_c ||}{\sum_{hkl} | F_o |} \tag{7}$$

$$R_{free} = \frac{\sum_{hkl \subset T} || F_o | - w | F_c ||}{\sum_{hkl \subset T} | F_o |} \tag{8}$$

where w is a scale factor. Lower the parameters, better is the agreement between structural model and diffraction data. In general, acceptable R_{free} values should be close to 0.250 or lower (depending on the resolution of the data), while R_{factor} should be a few percent lower [323]. Regardless the similar equation, an important difference relies in how the parameters are calculated. The R_{factor} is calculated on all the reflections used for the refinement, instead R_{free} is defined on a randomly selected set of data, usually 5% of the complete data set, which is then excluded from the refinement. R_{free} provides a more reliable indication of the quality of a solution and progress of the refinement in comparison to R_{factor} [323].
The development of graphic programs has allowed to visualize electronic density maps to further validate the quality of the model and phases calculated. [F_o-F_c] and [$2F_o$-F_c] electron density maps are commonly used. The latter increases when calculated in points close to atoms, while the former shows positive and negative electron densities which respectively indicate where atoms should not

be placed and where atoms should be added in the model.
The most common methods used for the refinement are the maximum likelihood and the simulated annealing [323]. The goal of maximum likelihood refinements is to determine the parameters of the model (e.g. bond distances, angles and torsions) in order to best explain the observed data, that is to give the highest value of likelihood between observed and calculated structure factors. On the other hand, the simulated annealing addresses the existence of multiple minima for the macromolecules, which largely arise from the high dimensionality of the parameters space. The many local minima of the target make difficult to shift the atomic coordinates enough to correct errors in the initial model. The simulation of the structure annealing at a certain temperature (even 2000-4000 K) and its subsequent cooling, usually performed by molecular dynamics (see next section), is thus performed with the purpose to escape from local minima and yield a family of conformations to be selected on the basis of the constraints imposed by the observed diffraction data and parameters as the R_{factor} [323, 326].
The validation of crystal structures consists in the analysis of various parameters concerning the agreement of the model with experimental data and also the accord with previously reported structures. As regards this last issue, the comparison of structures is carried out thanks to visual graphic programs and through the calculations of parameters regarding bond lengths and bond angles as the root mean square deviations (RMSD):

$$RMSD_{bonds} = \sqrt{\frac{\sum \delta d_i^2}{N_{atoms}}} \tag{9}$$

$$RMSD_{angles} = \sqrt{\frac{\sum \delta \alpha_i^2}{N_{atoms}}} \tag{10}$$

where δd_i and $\delta \alpha_i$ are respectively the i^{th} deviations of bond length and angle from the ideal values determined from high-resolution data (small molecules structures). RMSD parameters should be hopefully close to zero, however $RMSD_{bonds} \leqslant 0.02$ and $RMSD_{angles} \leqslant 1.9$ are considered acceptable [323].

2.2 COMPUTATIONAL STUDIES

Molecular modelling concerns the simulation of molecular systems *in silico*. Different theoretical approaches have been developed to shed light on the properties of studied systems. Simulations are commonly used to provide complementary results to those yielded by X-ray crystallography, and modelling can exploit the results from other experimental techniques, such as spectrophotometric analyses of the samples in solution. In the field of biological macromolecules, the attention has been particularly focused on the simulation of target-ligand interaction at different levels of theory.

Quantum mechanical calculations

Accurate methods for the investigation of molecular systems rely on quantum mechanical (QM) calculations which deal with the approximated solution of the Schrödinger equation:

$$\hat{H}\psi(r,t) = E\psi(r,t) \tag{11}$$

where $\psi(r,t)$ are space and time dependent wavefunctions describing the system and E are the corresponding energy levels. In

atomic units, the Hamiltonian $\hat{H}$ of a system of n electrons and m nuclei can be expressed as:

$$\begin{aligned}\hat{H} = &-\frac{1}{2}\sum_{i=1}^{n}\nabla_i^2 - \sum_{A=1}^{m}\frac{1}{2M_A}\nabla_A^2 - \sum_{i=1}^{n}\sum_{A=1}^{m}\frac{Z_A}{r_{iA}} \\ &+ \sum_{i=1}^{n}\sum_{j>i}^{n}\frac{1}{r_{ij}} + \sum_{A=1}^{m}\sum_{B>A}^{m}\frac{Z_A Z_B}{R_{AB}}\end{aligned} \tag{12}$$

where Z is the nuclear charge, M_A is the A^{th} nucleus/electron mass ratio, R_{AB} is the distance between A and B nuclei, r_{ij} is the distance between i and j electrons and r_{iA} is the distance between the i electron and the A nucleus. The terms of equation 12 refer respectively to the kinetic energy of electrons and nuclei, the attractive electron-nucleus interaction and the repulsive electron-electron as well as nucleus-nucleus interactions. On the other hand, the square of $\psi(r,t)$ wavefunction is associated to the probability to find the electron in the volume (V+δV), that is correlated to the electron density.

The Schrödinger equation can be analytically solved only for simple systems, such as the hydrogen atom, hence approximations are usually made. The most common is the Born-Oppenheimer approximation which takes into account the much slower motion of nuclei with respect to electrons as a consequence of quite different masses. Nuclei are treated as fixed while electrons move around them. The kinetic nuclear term of equation 12 can thus be neglected, leading to a purely electronic Hamiltonian:

$$\hat{H}_{el} = -\frac{1}{2}\sum_{i=1}^{n}\nabla_i^2 - \sum_{i=1}^{n}\sum_{A=1}^{m}\frac{Z_A}{r_{iA}} + \sum_{i=1}^{n}\sum_{j>i}^{n}\frac{1}{r_{ij}} \tag{13}$$

and the associated Schrödinger equation for electronic motion:

$$(\hat{H}_{el} + V_{NN})\psi_{el} = U\psi_{el} \tag{14}$$

where V_{NN} accounts for the internuclear repulsion:

$$V_{NN} = \sum_{A=1}^{m} \sum_{B>A}^{m} \frac{Z_A Z_B}{R_{AB}} \tag{15}$$

The energy U shown in equation 14 is the electronic energy including internuclear repulsion. The Born-Oppenheimer approximation yields electronic wave functions and energies which depend parametrically on the nuclear configuration chosen for the calculations and allows a more easily solution of the Schrödinger equation.
A useful method for the calculation of electronic wave functions and energies is the Hartree-Fock approach, which is based on the computation of accurate self-consistent field wave functions. The energy E_{HF} for a closed shells system is:

$$E_{HF} = 2\sum_{i=1}^{n/2} H_{ii} + \sum_{i=1}^{n/2} \sum_{j=1}^{n/2} (2J_{ij} - K_{ij}) + V_{NN} \tag{16}$$

where the electron core integral H_{ii} accounts for the kinetic and electron-nucleus attractive energies, while J_{ij} and K_{ij} are termed coulomb and exchange integrals, respectively, and contain electron-electron interaction terms. The sums run over the $n/2$ occupied orbitals for the molecule of n electrons. As regards open shell systems, the formulas are similar but a bit more complicated, so they are not here introduced.
The purpose of the Hartree-Fock method is to calculate orbitals which minimize E_{HF} by solving iteratively the equation:

$$\hat{F}(1)\psi_i(1) = \epsilon_i\psi_i(1) \tag{17}$$

where ϵ_i is the orbital energy and $\hat{F}$ is the one-electron Fock operator defined as:

$$\hat{F}(1) = \hat{H}(1) + \sum_{j=1}^{n/2} [2\hat{J}_j(1) - \hat{K}_j(1)] \tag{18}$$

where $\hat{H}$ includes the kinetic energy and the electron-nuclei attraction, while $\hat{J}$ and $\hat{K}$ are the coulomb and exchange operators, referring respectively to the electron-electron repulsion and the antisymmetric nature of electron wave functions due to the electronic spin. The main limits of the Hartree-Fock approach concern the underestimation of the electron correlation underlying these latter integrals which is usually referred to the (electronic) exchange-correlation energy [329, 330].
A key development in the calculation of accurate self-consistent wave functions was the expression of molecular orbitals as linear combinations of atomic orbitals:

$$\psi_i = \sum_{\mu} C_{\mu i} \chi_\mu \tag{19}$$

where ψ_i is the i^{th} molecular orbital, $C_{\mu i}$ are appropriate coefficients and χ_μ are atomic orbitals. The basis set of atomic orbitals used for the calculations should be as greater as possible in order to accurately describe the system [330]. Among the several sets proposed, the linear combination of gaussian functions is particularly valuable. Table 5 shows some basis sets commonly used for quantum mechanical calculations. The minimum basis set concerns functions for only electron occupied orbitals, which can be described by two or more functions per orbital (double zeta, triple zeta, etc.), even for only the valence shell (split-valence sets). The accuracy of the basis set can be further improved by the addition of polarization functions accounting for virtual orbitals, either for

only non-hydrogen atoms (notation *) or also including hydrogen atoms (notation **). In some cases, as those concerning anions, more diffused polarization functions (which go to zero slowly with the distance) should be used, either for only non-hydrogen atoms (notation +) or also including hydrogen atoms (notation ++). Anyway, when heavy atoms are concerned, the calculations can be sped up by using pseudo-potentials which circumvent the complicated effects of motion of the core electrons through the definition of an effective potential, thus drastically reducing the calculations requested for the solution of the Schrödinger equation [330].

Table 5: Representative basis sets used for quantum mechanical calculations.

ID	Orbitals involved	Representative basis sets
Minimum set	occupied	STO-3G, STO-4G, STO-6G
Split-valence set	occupied	3-21G, 4-31G, 6-311G
Polarization functions set	occupied, virtual	3-21G*, 3-21G**, 6-311G*
Diffuse functions set	occupied	3-21G+, 3-21G++
Diffuse functions set	occupied, virtual	3-21G*+, 6-311G*+, 6-311G**+

Besides the Hartree-Fock approach, alternative methods have been developed in order to solve approximated Schrödinger equations. Among them, the density functional theory (DFT) has emerged as much valuable as it supplies accurate results with a relatively low computational time [329]. The DFT approach is based on the Hohenberg-Kohn theorem which demonstrates how the electronic molecular properties of the system, as the wavefunctions and the molecular energy, in the non-degenerate ground state are uniquely determined by the ground state electron probability density ρ_0. The ground state molecular energy can be indeed expressed as a functional of the ground state electron probability density, that is $E_0=E_0[\rho_0]$. The great advantage of DFT relies in the calculation of the system energy from ρ_0 which is a function dependent on only three spatial coordinates instead on the $3n$ spatial and n

spin coordinates associated to the electronic wave function of a molecule of n electrons. As a result, the time cost of DFT *ab initio* calculations is reduced in comparison to different approaches as the Hartree-Fock method. For a molecule of n electrons, the electronic Hamiltonian can be expressed as:

$$\hat{H} = -\frac{1}{2}\sum_{i=1}^{n}\nabla_i^2 + \sum_{i=1}^{n} v(r_i) + \sum_{i=1}^{n}\sum_{j>i}^{n}\frac{1}{r_{ij}} \tag{20}$$

where $v(r_i)$ is the potential of interaction between the i^{th} electron and the nuclei:

$$v(r_i) = -\sum_{A=1}^{m}\frac{Z_A}{r_{iA}} \tag{21}$$

which is commonly called external potential. Given $v(r_i)$ and the number of electrons n, the electronic wavefunctions and energies are determined as solutions of the electronic Hamiltonian (equation 20). Hohenberg and Kohn demonstrated that $\rho_0(r)$ determines vice versa the external potential and the number of electrons in the ground state of the system.

The second Hohenberg-Kohn theorem provides a practical tool for DFT calculations as it shows how the ground state electron density minimizes the energy functional $E_0[\rho_0]$ that corresponds to the ground state energy, thus ground state molecular properties can be in principle calculated if the ground state electronic density probability $\rho_0(r)$ is known.

Kohn and Sham theorized a method for DFT calculations based on a fictitious reference system of n non-interacting electrons under an external potential, with the purpose of determining a ground state electron density $\rho_s(r)$ equal to the desired electronic density

$\rho_0(r)$. Kohn and Sham derived the following electronic Hamiltonian for the reference system:

$$E[\rho] = T_s[\rho] + V[\rho] + J[\rho] + E_{xc}[\rho] \tag{22}$$

where the $T_s[\rho]$ functional accounts for the kinetic energy, $V[\rho]$ is the external potential functional, $J[\rho]$ and $E_{xc}[\rho]$ are respectively the electron-electron repulsion and exchange-correlation functionals. The Kohn-Sham equation describes the system in terms of non-exact kinetic and electron-electron repulsion functionals, but allows the calculation of $\rho_s(r)$. In more detail, the application of the variational method to equation 22 leads to n mono-electronic equations:

$$\left[-\frac{1}{2}\nabla^2 + v(r) + v_J(r) + v_{xc}(r)\right]\psi_i = \left[-\frac{1}{2}\nabla^2 + v_{eff}\right]\psi_i = \epsilon_i\psi_i \tag{23}$$

which are formally mono-electronic equations for electrons under an effective potential v_{eff}. The relation between the electron density ρ and the wavefunctions ψ_i is:

$$\rho = \rho_s = \sum_{i=1}^{n} |\psi_i|^2 \tag{24}$$

DFT calculations manage to recursively calculate $\rho(r)$ and $v_{eff}(r)$ until self-consistency is achieved. Unfortunately, the actual nature of the v_{eff} functional, in particular of the electronic exchange-correlation functional, is actually unknown. For this reason, huge efforts have been made in the identifications of suitable functionals able to accurately approximate the exchange-correlation functional and so correctly describe the system [329, 330]. B3LYP is one of the most commonly used functionals in this regard [331]. This hybrid functional mixes the exact Hartree-Fock electron exchange

term and alternative gradient-corrected ($\nabla\rho$) exchange and correlation functionals by using three parameters obtained from *ab initio* and empirical data (B3LYP: Becke-3 parameters, Lee-Yang-Park functional). Overall, the accuracy of DFT calculations is essentially influenced by the functional as well as the basis sets used [330].

Molecular mechanics calculations

Molecular mechanics (MM) allows the simulation of molecular systems under the laws of classical newtonian physics. Clearly, MM investigations do not deal with the evaluation of electron properties, but they provide useful information about the dynamics of nuclei. Assuming the permanence in the electronic ground state, the conformation of molecules is determined by the relative positions of the nuclei which define potential energy surfaces.
MM calculations are useful for the identification of energetically stable conformation of the system, taking into account that the chemical groups present in many molecules are characterized by similar geometries and behaviour. This concept results in the definition of atom types which are referred to specific elements involved in distinct chemical bonds. For example, there are different atom types for aliphatic and aromatic carbon atoms. Molecules are treated as an ensemble of rigid spheres connected through spring-like bonds, where each atom features defined radius, charge and hybridisation.
The potential energy of the system is calculated through a parametric function dependent on nuclear coordinates, that is the force field (FF). MM simulations allow to evaluate the energy of the system as a function of nuclear coordinates and so to identify the most stable geometries which correspond to potential energy surface minima. The steepest descent and conjugate gradient algorithms (first derivative $\partial E/\partial r$) are commonly used to minimize

the system energy as well as the truncate Newton-Raphson (second derivative $\partial^2 E/\partial r^2$) which is particularly suited for large ensembles of atoms [330, 332].
Several force fields have been reported in order to better describe various molecular systems. FF are composed of bond stretching, bending, torsion, Van der Waals and coulombic terms, as shown in the following equation for OPLS_2005, a force field commonly used for the simulations of biological macromolecules [333–335].

$$\begin{aligned} V(r_n) = & \sum_{\text{bonds}} K_b(r - r_0)^2 + \sum_{\text{angles}} K_a(\theta - \theta_0)^2 \\ & + \sum_{\text{torsions}} \left\{ \frac{V_1}{2}[1 + \cos(\phi - \phi_1)] + \frac{V_2}{2}[1 - \cos(2\phi - \phi_2)] \right. \\ & \left. + \frac{V_3}{2}[1 + \cos(3\phi - \phi_3)] + \frac{V_4}{2}[1 - \cos(4\phi - \phi_4)] \right\} \\ & + \sum_{i<j} \left\{ 4\epsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^{6} \right] + \frac{q_i q_j e^2}{r_{ij}} \right\} f_{ij} \end{aligned} \tag{25}$$

where K_b, r_0, K_a, θ_0, V_1, V_2, V_3, V_4, $\phi_1, \phi_2, \phi_3, \phi_4$, ϵ and σ are atom type parameters regarding bond stretching, bond bending, bond torsions and non-bonding interactions, respectively. r_0 and θ_0 are the equilibrium values of the bond length r and angle θ, whereas $\phi_1, \phi_2, \phi_3, \phi_4$ are phases of the dihedral angle ϕ. The parameters are generally derived from observations on small organic molecules that are more tractable for experimental studies and quantum mechanical calculations, in particular experimental properties of liquids are taken into account for the OPLS force field [333–335]. The f_{ij} coefficients account for non-covalent interactions, for example they are usually set to zero for 1,2 (atoms connected by a covalent bond) and 1,3 interactions, while set to

0.5 for 1,4 interactions and 1.0 for all the other cases [333].
While the properties of organic systems are generally well reproduced by force fields, the treatment of metal complexes by MM methods is quite problematic because of the variability in coordination number, coordination geometry, bond length and occasionally in the case of not well-defined bonds (e.g. in metallocenes). In fact, common force fields usually lead to distortions in the coordination geometry and are not able to correctly simulate systems containing metals. Different approaches have been developed to handle these issues, as the modification of FF neglecting bending terms involving the metal-ligand bonds instead including 1,3 non-bonded interactions [330]. Additionally, universal force fields (UFF) based on atomic constants have been proposed to improve the treatment of metal geometries, although the results obtained are less accurate than those achieved by commonly used force fields [330, 336].
An important issue in the simulation of systems containing metals is the accurate estimation of atom charges, that is crucial for the calculation of electrostatic terms of the force field and so for the overall reliability of the simulation. MM methods are able to assign atomic partial charges which however do not supply trustworthy values when metal ions are present. Preliminary QM calculations are thus useful for an accurate evaluation of atomic charges through the simulation of the electrostatic potential (ESP) of the system. Briefly, a grid is defined around the system studied and the electrostatic potential at the grid is calculated. A series of point charges are then assigned to the atoms of the molecule in order to better reproduce the potential observed on the grid [330]. QM simulations provide accurate atomic charges and are also useful for obtaining good starting molecular geometries for the subsequent calculations. This approach can be also applied on systems that do not contain metal ions in order to improve the

accuracy of MM simulations.
When the binding of a small molecule to a biological target has to be simulated, the molecular modelling offer the opportunity to perform molecular docking procedures [330, 332]. These methods rely on the definition of properly centred grids of points where the electrostatic and Van der Waals potential due to the grid-contained atoms is calculated. These grids are developed for the biomolecule on a case by case basis. Algorithms then explore possible orientations of the ligand molecules on the grid and the resulting poses are scored, for instance, in terms of interaction energy as calculated by FF. In general, the interaction energy ΔE is defined as:

$$\Delta E = E_{RL} - E_R - E_L \tag{26}$$

where E_{RL}, E_R and E_L are the adduct, receptor and ligand energy, respectively. Docking procedures are able to point out the most probable binding sites of small molecules, their orientation at the site and also allow the energetic comparison of the poses calculated. In the case of biological macromolecules, the receptor is usually kept rigid for the docking while the ligand can be either considered rigid or flexible.
Anyway, the interaction of small molecules bounded to biotargets, especially in the case of nucleic acids, can be hampered by residues of the macromolecule which do not allow an accurate simulation of the binding. However, among the available methods, NVT molecular dynamics (MD) is useful in the treatment of biomolecules and their adducts, and provides the possibility to simulate the temporal evolution of the adduct at the desired temperature and so to explore the conformational space [330, 332]. The technique is based on the iterative calculation of the Hook forces acting on atoms in motion involved in chemical bonds (average atom kinetic energy dependent on the temperature of the system), according to the FF used, and the analysis of atomic tra-

jectories. MD algorithms solve motion equations derived from classical physics:

$$F = m \cdot a \tag{27}$$

where F is the force applied to a particle of mass m and acceleration a. The atomic trajectories are calculated through the integration of the previous equation for the i^{th} atom:

$$\frac{F_i}{m_i} = \frac{\partial^2 x_i}{\partial t^2} \tag{28}$$

So, knowing the mass and the forces associated to each atom, the atomic acceleration, velocity and resulting position can be obtained by solving the following equation:

$$F_i = -\frac{\partial V}{\partial x_i} = \frac{m_i \partial^2 x_i}{\partial t^2} \tag{29}$$

where V is the potential energy. In order to simulate the molecular dynamics, the previous equation has to be iteratively integrated in a defined time step δt which is short enough to consider the forces acting on atoms as constant. The choice of the time step is quite important as it is crucial for the accuracy and the computational cost of the MD simulations. A good compromise is to set δt as 1-2 fs allowing a good sampling for the dynamics of the system (notably, C-H oscillation period about 10 fs) together with a reasonable computational cost [330, 332].
When large systems are considered, the motion equations display complex terms, as V depends on all the nuclear coordinates, causing the calculations to request a great amount of time. In order to perform simulations within an acceptable time, MD algorithms have developed to numerically solve the integrals, as shown by

the widely used Verlet algorithm which is based on third order Taylor expansion:

$$x_i(t + \delta t) \approx 2x_i(t) - x_i(t - \delta t) + \frac{F_{x_i}(t)}{m_i}\delta t^2 \tag{30}$$

Prior to the simulations, initial velocities are randomly assigned to atoms and let equilibrated over a determined time which allows the bond to vibrate according to the desired temperature. The equilibration time should be as much longer as larger is the system studied. During the molecular dynamics calculations, the configurations of the system are sampled at determined time intervals, and then undergo MM minimization in order to remove eventual conformational stress. The temperature and the overall energy of the system are kept constant, although they vary instantaneously during the simulations.

Molecular modelling is a valuable approach for the simulation of biomacromolecules and their interaction with small molecules. Significant limitations of the methods emerge however when the system studied contains heavy metal ions. As already discussed, force fields are not suited to provide accurate atomic charges and to manage coordination geometries, even applying some geometrical restraints. Thus, higher levels of theory are often required when treating biomolecule-metal complex adducts. QM methods are particularly useful in this regard, still they display elevated computational cost and they are solely confined to the treatment of small molecule ligands. The search of a compromise has led to the development of QM/MM hybrid methods which supply accurate and time efficient calculations. Nevertheless, exploring the potential surface of these systems remains a very challenging issue.

QM/MM hybrid calculations

Quantum mechanical calculations are undoubtedly more accurately than molecular mechanics methods, however they require high computational cost even for reduced ensemble of atoms [330]. The request of high accuracy for *in silico* calculations on large systems as biomacromolecules has led to the development of QM/MM hybrid approaches in which different regions of the system are either studied at QM or MM level of theory [337]. These methods allow to properly simulate the binding of small molecules on large receptors within an acceptable computational time, and it is much more useful when metal complexes are involved [338]. The definition of QM and MM regions determines a QM/MM boundary in which mechanical and electronic effects from both the regions are conveyed. The choice of the QM/MM boundary is thus quite critical and greatly affects the results obtained [330]. The total energy of the system can be defined as:

$$E_{tot} = E_{MM} + E_{QM} + E_{QM/MM} \tag{31}$$

where MM and QM terms account for theoretically isolated regions while the QM/MM term expresses the interaction energy of the regions for electrostatic, Van der Waals and chemical bond contributions. As regards the calculations, the numerical solution of the Schrödinger equation and associated partial differential equations usually takes advantage of the pseudo-spectral method which can considerably speed up the simulations.
When the QM/MM boundary includes covalent bonds between atoms belonging to different regions, different strategies have been developed in order to maintain a correct atom valency and molecular geometry. The most common strategy is to use capping atoms, usually dummy hydrogen atoms, that are considered for the QM

calculations while they are neglected for the treatment of the MM region [338].

2.3 SPECTROPHOTOMETRY

X-ray crystallography and *in silico* calculations supply accurate structural information, however a more general picture of the interaction of small molecules with DNA foldings often requests complementary data. Spectrophotometry provides interesting evidences on the formation of ligand-DNA adducts in solution, thus the research activity has also concerned experiments carried out by UV-Vis absorption and circular dichroism spectroscopy.

UV-Vis absorption spectroscopy

UV-Vis absorption spectroscopy deals with the absorption of electromagnetic radiations in the ultra violet (100 Å $< \lambda <$ 300 nm) and visible (300 nm $< \lambda <$ 750 nm) spectral regions by the system examined. The absorption is associated to excitation of valence shell electrons of chemical groups. UV-Vis spectroscopy can be used to monitor the chemical reactions which affect the electronic properties of the system, as the interaction of ligand molecules with DNA structures.
In general, when an electromagnetic radiation of intensity ι_0 passes through a sample, a different intensity ι is observed. Transmittance T and absorbance A are defined as:

$$T = \frac{\iota}{\iota_0} \tag{32}$$

$$A = \log \frac{\iota_0}{\iota} = \log \frac{1}{T} \tag{33}$$

The Lambert-Beer equation describes the linear correlation between the absorbance A and the concentration c of absorbent chemical species at a certain radiation wavelength λ:

$$A = \epsilon_\lambda bc \tag{34}$$

where b is the sample width crossed by the radiation and ϵ is a coefficient associated to the electronic transition probability for the absorbing species which is dependent on radiation wavelength, temperature, solvent, pH, concentration and solution solutes (salts, in particular). When c is expressed in molarity and b in cm, ϵ is termed molar attenuation coefficient or molar extinction coefficient or also molar absorptivity. Interestingly, the Lambert-Beer equation is valid only for certain ranges of concentration.

As regards DNA, the macromolecule shows strong absorption in the region 275-240 nm due to the aromatic pyrimidine and purine ring systems of the bases. The intensity at 260 nm is usually considered for analytical evaluations.

The description of ligand-receptor interactions is often modelled on the equilibrium:

$$L + R \leftrightarrows L - R \tag{35}$$

leading to the definition of association and dissociation binding constants:

$$K_a = \frac{[LR]}{[L][R]} \tag{36}$$

$$K_d = \frac{[L][R]}{[LR]} \tag{37}$$

which are connected through the equation $K_a = 1/K_d$. Among a number of equations that are commonly used for the analysis of

spectroscopic data, a valuable method for studying ligand-DNA interactions [280, 282, 339, 340] concerns the titration of a ligand solution with concentrated DNA aliquots and the construction of the binding isotherm by the equation:

$$c_0\left(\frac{\Delta A}{\Delta A_{max}}\right)^2-(c_0+c_p+K_d)\left(\frac{\Delta A}{\Delta A_{max}}\right)+c_p=0 \qquad (38)$$

where c_0 is the initial concentration of the ligand, c_p is the concentration of the oligonucleotide and ΔA is the variation in absorbance for the ligand at a given wavelength upon DNA addition. ΔA_{max} is the maximum increase observed for ΔA which refers to the achievement of ligand saturation. In more detail, $\Delta A=(\Delta\epsilon)bc$, in according to the Lambert-Beer equation, and in particular $\Delta\epsilon=\epsilon_b-\epsilon_f$, where ϵ_b is the molar extinction coefficient of the bound ligand, instead ϵ_f is the molar attenuation coefficient of the free ligand. The determination of ΔA_{max} concerns the equation:

$$\frac{1}{\Delta A}=\frac{1}{\Delta A_{max}}+\frac{1}{c_p-c_0} \qquad (39)$$

The linear plot $1/\Delta A$ vs $1/(c_p-c_0)$ yields ΔA_{max} from the intercept. Subsequently, the linear plot $\Delta A/\Delta A_{max}$ vs c_p is used to estimate K_d from the abscissa intercept of 0.5. Lastly, the ligand-DNA stoichiometry can be approximately assessed from the break in the straight lines resulting from the fitting of the initial and terminal titration profile. The stoichiometry is derived by dividing the DNA concentration at the break by c_0.

Circular dichroism spectroscopy

Electromagnetic waves can split into a left-handed and a right-handed circular components that have the same intensity, frequency and phase. The circular nature of the radiation concerns the ro-

tation of the electric field vector at constant angular velocity in the time scale which results in a spiral pattern in the space scale. This behaviour of light can be observed when electromagnetic radiations interact with structural non-uniformities at molecular level, as chiral molecules and arrangements of chiral residues, notably in the case of DNA. The circular dichroism spectroscopy investigates such kind of interaction by evaluating the different absorptions of left-handed and right-handed polarized circular radiations belonging to the UV-Vis spectrum.
When a monochromatic polarized radiation of wavelength λ interacts with an optically active system, a rotation of angle α in the polarization plane is observed. This angle can be described as a function of the different refractive indices of left-handed (n_{left}) and right-handed (n_{right}) components through the equation:

$$\alpha = \frac{(n_{\mathrm{left}} - n_{\mathrm{right}})\pi b}{\lambda} \tag{40}$$

where b is the optical path. The plot α vs λ is called optical rotatory dispersion (ORD). When absorption of the radiation occurs, the resulting light beams exhibit elliptical polarization because the circularly polarized left-handed and right-handed components are absorbed to a different degree (circular dichroism, CD). The extent of the differential absorption is measured by the ellipticity angle θ which is defined as a function of the magnitudes of the electric field vectors of the right-handed and left-handed circular components:

$$\tan\theta = \frac{\mathrm{E}_{\mathrm{right}} - \mathrm{E}_{\mathrm{left}}}{\mathrm{E}_{\mathrm{right}} + \mathrm{E}_{\mathrm{left}}} \tag{41}$$

Furthermore, the absorption A can be analytically evaluated by the Lambert-Beer law:

$$\Delta A = A_{\mathrm{left}} - A_{\mathrm{right}} = (\epsilon_{\mathrm{left}} - \epsilon_{\mathrm{right}})bc = (\Delta\epsilon)bc \tag{42}$$

where $\Delta\epsilon$ is the differential dichroic absorption. Analogously to what stated for UV-Vis absorption spectroscopy, c is the concentration of absorbing species while b is the optical path. If $\Delta\epsilon > 0$, then $\theta > 0$. The equation just introduced can also be rewritten as:

$$\theta = [\theta]^T_\lambda bc \tag{43}$$

The plot θ vs λ is the circular dichroism plot. When c is expressed in molarity and b in cm, $[\theta]^T_\lambda$ is called molar ellipticity of the species at wavelength λ and temperature T. Similarly to the molar attenuation coefficient, this coefficient depends on the radiation wavelength, temperature, solvent, solution pH and solution solutes (salts, in particular). ΔA, that is $\Delta\epsilon$, is usually the physical quantity measured in modern instruments.
CD is a useful tool for the investigation of DNA secondary structures, as the asymmetric orientation of bases gives rise to dichroic bands in the UV-Vis spectral region [341]. For instance, the human telomeric G-quadruplex display in particular intense bands, largely caused by *syn*/*anti* guanine conformations, which are informative for the G4 topologies [342–346]:

- Antiparallel: absolute maximum at 295 nm, relative maximum at 245 nm, absolute minimum at 265 nm
- Parallel form: absolute maximum at 260 nm, absolute minimum at 240 nm
- Parallel/antiparallel topology: absolute maximum at 290 nm, relative maximum at 268 nm, absolute minimum at 240 nm

CD spectroscopy can be used to monitor how ligand binding affects the G-quadruplex topology and also to record thermal melting curves, especially at wavelengths where the signal is high. The evaluation of binding constants can be attained by performing CD

titrations and fitting experimental data by equations as that previously reported for UV-Vis absorption spectroscopy. A robust, yet rough, method for the analysis of ligand-DNA interactions is the Benesi-Hildebrand plot [347–350]:

$$\frac{1}{(\theta - \theta_0)} \; vs \; \frac{1}{c_{ligand}} \tag{44}$$

where $(\theta - \theta_0)$ is difference in ellipticity observed for G4 DNA upon addition of ligand aliquots to a concentration c_{ligand}.

3 EXPERIMENTAL SECTION

This chapter provides information about the experiments personally carried out during the PhD years. The chemical products used for the experiments have been purchased from Sigma-Aldrich with the exception of synthetic DNA oligonucleotides (table 6), bought from Jena Bioscience (Jena, Germany) at HPLC purity, and ligand molecules. Synthetic and semi-synthetic compounds (table 7) have been synthesised by collaborators in according to literature procedures [306, 317, 322, 351–353], while natural alkaloids (table 7) were purchased from Shanghai Trust & We, Ltd. (Shanghai, China).

Table 6: DNA sequences used for the experiments during the PhD years.

ID	sequence
Tel12	d[$TAG_3TTAGGGT$]
Tel23	d[$TAG_3(TTAGGG)_3$]
Tel24	d[$(TTAGGG)_4$]
Tel24′	d[$TAG_3(TTAGGG)_3T$]
Tel26	d[$A_3G_3(TTAGGG)_3A_2$]
wtTel26	d[$(TTAGGG)_4T_2$]
CGTACG	d[CGTACG]
CGATCG	d[CGATCG]

Table 7: Small molecule ligands studied during the PhD years.

ID	IUPAC name
Berberine	9,10-Dimethoxy-5,6-dihydro[1,3]dioxolo[4,5-g]isoquinolino[3,2-a]isoquinolin-7-ium chloride
Coptisine	6,7-Dihydro[1,3]dioxolo[4,5-g][1,3]dioxolo[7,8]isoquinolino[3,2-a]isoquinolin-5-ium chloride
Sanguinarine	13-Methyl[1,3]benzodioxolo[5,6-c][1,3]dioxolo[4,5-i]phenanthridin-13-ium chloride
Chelerythrine	1,2-Dimethoxy-12-methyl[1,3]benzodioxolo[5,6-c]phenanthridin-12-ium chloride
NAX039, NAX042	13-arylalkyl Berberine chloride
NAX035, NAX053	13-diarylalkyl Berberine chloride
NAX071, NAX075, NAX077, NAX079, NAX120	13-pyridilalkyl Berberine chloride
AuNHC **I**	$[Au(9\text{-methylcaffein-8-ylidene})_2]BF_4$
AuNHC **II**	$[Au(bis(1\text{-butyl-3-methyl-imidazole-2-ylidene}))]PF_6$
Auoxo6	$[(bipy^{2Me})_2Au_2(\mu\text{-}O)_2](PF_6)_2$ $bipy^{2Me}$=6,6′-dimethyl-2,2′-bipyridine
$[Pd(bapbpy)]^{2+}$	$[Pd(bapbpy)](PF_6)_2$ bapbpy=6,6′-bis(2-aminopyridyl)-2,2′-bipyridine
$[Pt(bapbpy)]^{2+}$	$[Pt(bapbpy)](PF_6)_2$ bapbpy=6,6′-bis(2-aminopyridyl)-2,2′-bipyridine

DNA oligonucleotides have been annealed prior to the experiments. Briefly, the lyophilized samples were dissolved in buffer and heated in a water bath under magnetic stirring up to 90-95°C for at least 15 minutes and then let cool overnight to room temperature. Human telomeric sequences were usually dissolved in 0.02 M potassium cacodylate buffer (pH 6.5), 0.05 M KCl to a concentration of either 0.001 M (Tel23, Tel24, Tel24/) or 0.002 M (Tel12). On the other hand, CGTACG and CGATCG sequences were dissolved in 0.01 M Tris sulphate, 0.001 M EDTA (pH 7.2) to a concentration of 0.002 M. Stock solutions of the ligand molecules were prepared in DMSO at a concentration of 0.002-0.02 M, depending on the solubility of the compounds. MilliQ water from Millipore Water System (Millipore, USA) was used for the preparation of all solutions, including those of the crystallization screenings, which were also filtered through Sartorius Stedim Biotech filters of 0.2 μm.

3.1 X-RAY CRYSTALLOGRAPHY

Crystallization screenings for the adducts formed by DNA and ligand molecules were performed using the sitting drop vapour diffusion method. The screenings used for human telomeric DNA are reported in table 8 while the conditions used for the short sequences CGTACG and CGATCG are shown in table 9. Both the screenings were adapted from previously reported works [354, 355].

Table 8: Crystallization screening used for the crystallization of telomeric G-quadruplex adducts. MPD=2-Methyl-2,4-PentaneDiol, PEG=polyethylene glycol.

Condition	Buffer	Salts	Precipitant
1	0.05 M K Cacodylate pH 6.5	0.1 M KCl	20% v/v MPD
2	0.05 M K Cacodylate pH 6.5	0.3 M KCl	10% v/v MPD
3	0.05 M K Cacodylate pH 6.5	0.1 M KF, 0.025 M $MgSO_4$	25% v/v MPD
4	0.05 M K Cacodylate pH 6.5	0.2 M KI	25% v/v MPD
5	0.05 M Na Cacodylate pH 6.5	0.1 M KF, 0.02 M $(NH_4)_2SO_4$	15% v/v MPD
6	0.05 M Na Cacodylate pH 6.5	0.3 M NaCl	30% v/v MPD
7	0.05 M Na Cacodylate pH 6.5	0.1 M NaI	35% v/v MPD
8	0.05 M K Cacodylate pH 6.5	0.3 M KF, 0.02 M $MgSO_4$	35% v/v MPD
9	0.05 M Na Cacodylate pH 6.5	0.15 M NaF, 0.15 M KF	40% v/v MPD
10	0.05 M K Cacodylate pH 6.5	0.1 M KCl, 0.1 M $(NH_4)_2SO_4$	20% v/v MPD
11	0.05 M Na Cacodylate pH 6.5	0.1 M NaI, 0.1 M NaF, 0.1 M NaCl	30% v/v MPD
12	0.05 M K Cacodylate pH 6.5	0.2 M KCl, 0.05 M KI	25% v/v PEG400
13	0.05 M K Cacodylate pH 6.5	0.1 M KCl, 0.05 M $MgCl_2$	10% v/v PEG400
14	0.05 M K Cacodylate pH 6.5	0.1 M Li_2SO_4	20% v/v PEG400
15	0.05 M K Cacodylate pH 6.5	0.3 M KI	15% v/v PEG400
16	0.05 M Na Cacodylate pH 6.5	0.3 M NaI	15% v/v PEG400
17	0.05 M K Cacodylate pH 6.5	0.2 M KCl, 0.05 M Li_2SO_4	1.8 M $(NH_4)_2SO_4$
18	0.05 M K Cacodylate pH 6.5	0.05 M Li_2SO_4, 0.05 M $MgSO_4$	20% v/v isopropanol
19	0.05 M Na Cacodylate pH 6.5	0.2 M NaCl	20% v/v PEG3000
20	0.05 M Na Cacodylate pH 6.5	0.2 M NaCl, 0.1 M NaF	10% v/v PEG3000
21	0.05 M K Cacodylate pH 6.5	0.05 M KI, 0.05 M $MgCl_2$	15% v/v PEG3000
22	0.05 M Na Cacodylate pH 6.5	0.1 M NaCl, 0.05 M $MgCl_2$	1 M $(NH_4)_2SO_4$
23	0.05 M Na Cacodylate pH 6.5	0.05 M Li_2SO_4	1 M $(NH_4)_2SO_4$

Table 9: Crystallization screening used for the crystallization of CGTACG and CGATCG adducts. MPD=2-Methyl-2,4-PentaneDiol.

Condition	Buffer	Salts	Precipitant	Additive
1	0.04 M K Cacodylate pH 6.0	0.08 M KCl, 0.02 M $MgCl_2$	10% v/v MPD	0.012 M spermine chloride
2	0.04 M K Cacodylate pH 6.0	0.08 M KCl	10% v/v MPD	0.012 M spermine chloride
3	0.04 M Na Cacodylate pH 6.0	0.08 M NaCl, 0.02 M $MgCl_2$	10% v/v MPD	0.012 M spermine chloride
4	0.04 M Na Cacodylate pH 6.0	0.08 M NaCl	10% v/v MPD	0.012 M spermine chloride
5	0.04 M Li Cacodylate pH 6.0	0.02 M $MgCl_2$, 0.08 M LiCl	10% v/v MPD	0.012 M spermine chloride
6	0.04 M Li Cacodylate pH 6.0	0.08 M LiCl	10% v/v MPD	0.012 M spermine chloride
7	0.04 M K Cacodylate pH 6.0	0.08 M KCl, 0.02 M $CaCl_2$	10% v/v MPD	0.012 M spermine chloride
8	0.04 M Na Cacodylate pH 6.0	0.08 M NaCl, 0.02 M $CaCl_2$	10% v/v MPD	0.012 M spermine chloride
9	0.04 M Li Cacodylate pH 6.0	0.02 M $CaCl_2$, 0.08 M LiCl	10% v/v MPD	0.012 M spermine chloride
10	0.04 M Na Cacodylate pH 6.0	0.012 M KCl, 0.08 M NaCl, 0.02 M $MgCl_2$	10% v/v MPD	0.012 M spermine chloride
11	0.04 M K Cacodylate pH 6.0	0.08 M KCl, 0.012 M NaCl	10% v/v MPD	0.012 M spermine chloride
12	0.04 M Na Cacodylate pH 6.0	0.08 M NaCl, 0.02 M $BaCl_2$	10% v/v MPD	0.012 M spermine chloride
13	0.04 M K Cacodylate pH 6.0	0.08 M KCl, 0.02 M $BaCl_2$	10% v/v MPD	0.012 M spermine chloride
14	0.04 M Li Cacodylate pH 6.0	0.08 M $SrCl_2$	10% v/v MPD	0.012 M spermine chloride
15	0.04 M K Cacodylate pH 7.0	0.08 M KCl, 0.02 M $MgCl_2$	10% v/v MPD	0.012 M spermine chloride
16	0.04 M K Cacodylate pH 7.0	0.08 M KCl	10% v/v MPD	0.012 M spermine chloride
17	0.04 M Na Cacodylate pH 7.0	0.08 M NaCl, 0.02 M $MgCl_2$	10% v/v MPD	0.012 M spermine chloride
18	0.04 M Na Cacodylate pH 7.0	0.08 M NaCl	10% v/v MPD	0.012 M spermine chloride
19	0.04 M Na Cacodylate pH 7.0	0.012 M KCl, 0.08 M NaCl, 0.02 M $MgCl_2$	10% v/v MPD	0.012 M spermine chloride
20	0.04 M K Cacodylate pH 7.0	0.08 M KCl, 0.012 M NaCl	10% v/v MPD	0.012 M spermine chloride
21	0.04 M Na Cacodylate pH 7.0	0.08 M NaCl, 0.02 M $BaCl_2$	10% v/v MPD	0.012 M spermine chloride
22	0.04 M K Cacodylate pH 7.0	0.08 M KCl, 0.02 M $BaCl_2$	10% v/v MPD	0.012 M spermine chloride
23	0.04 M Li Cacodylate pH 7.0	0.02 M $MgCl_2$, 0.08 M $SrCl_2$, 0.04 M LiCl	10% v/v MPD	0.012 M spermine chloride
24	0.04 M Li Cacodylate pH 7.0	0.08 M $SrCl_2$, 0.04 M LiCl	10% v/v MPD	0.012 M spermine chloride
25	0.04 M Na Cacodylate pH 7.0	0.02 M $MgCl_2$, 0.08 M $SrCl_2$	10% v/v MPD	0.012 M spermine chloride
26	0.04 M K Cacodylate pH 7.0	0.08 M $SrCl_2$	10% v/v MPD	0.012 M spermine chloride

Crystallization trials were performed using Greiner® 96-well plates designed for sitting drop experiments. The formation of adducts was achieved by mixing solutions of annealed DNA and ligands usually in 1:1 stoichiometric ratio. The solutions were then incubated for at least 20 minutes at 25°C. Every drop was formed by 1muL of the adduct solution and 1muL of each crystallization condition. In the case of the double helix crystallization screening, the reservoir contained 100 µL of 30% v/v MPD. On the other hand, in the case of the quadruplex crystallization screening, the 100 µL reservoirs were composed either of the same condition used in the drop or of solutions with increasing amounts of the precipitant (35% and 50% v/v MPD, 30% and 45% v/v PEG400, 35% and 50% v/v isopropanol, 1.8 M and 2.4 M $(NH_4)_2SO_4$, 30% and 45% v/v PEG3000). At the end of drops deposition, the plates were hermetically sealed thanks to transparent adhesive tape and stored in an incubator (3001 model type by Rumed®) at 23°C.
The plates were periodically checked by using an optical microscope. Crystals have been observed for most of the adducts studied and X-ray diffraction analyses were performed on them. Several crystals showed good diffraction and the X-ray structures of the relative adducts were successfully solved. The crystallization conditions from which the crystals were obtained are reported in table 10 along with the cryoprotecting solutions used.
The crystals were analysed at a X-ray diffractometer with the experimental parameters reported in tables 11, 12. All the experiments were carried out at the cryo temperature of 100 K. The analysis of diffraction images (indexing, integration, scaling) was performed by using the XDS software [356].

Table 10: Crystallization conditions and cryoprotecting solutions used for the DNA-ligand adducts. Only the increased concentration of precipitant or viscous chemicals is reported for the cryo solutions.

Adduct	Condition	Reservoir	Cryo
Tel12-Coptisine	0.05 M Na Cacodylate pH=6.5, 0.15 M NaF 0.15 M KF, 40% v/v MPD	0.05 M Na Cacodylate pH=6.5, 0.15 M NaF 0.15 M KF, 40% v/v MPD	not needed
CGTACG-Chelerythrine	0.04 M Na Cacodylate pH= 6.5, 0.012 M KCl, 0.08 M NaCl 0.02 M $MgCl_2$, 10% v/v MPD, 0.012 M spermine chloride	30% v/v MPD	30% v/v MPD
CGTACG-Coptisine	0.04 M Na Cacodylate pH= 6.0, 0.08 M KCl, 0.02 M $MgCl_2$ 10% v/v MPD, 0.012 M spermine chloride	30% v/v MPD	30% v/v MPD
CGATCG-Coptisine	0.04 M Na Cacodylate pH= 6.0, 0.08 M KCl, 0.02 M $BaCl_2$ 10% v/v MPD, 0.012 M spermine chloride	30% v/v MPD	20% v/v MPD
Tel12-NAX053	0.05 M Na Cacodylate pH=6.5, 0.05 M Li_2SO_4 1 M $(NH_4)_2SO_4$	2.4 M $(NH_4)_2SO_4$	30% v/v glycerol
Tel12-NAX075	0.05 M K Cacodylate pH=6.5, 0.05 M Li_2SO_4 0.05 M $MgSO_4$, 20% v/v isopropanol	0.05 M K Cacodylate pH=6.5, 0.05 M Li_2SO_4 0.05 M $MgSO_4$, 20% v/v isopropanol	30% v/v glycerol
Tel12-NAX077	0.05 M K Cacodylate pH 6.5, 0.1 M Li_2SO_4 20% v/v PEG400	45% v/v PEG400	40% v/v PEG400
Tel12-NAX120	0.05 M Na Cacodylate pH=6.5, 0.1 M NaI 35% v/v MPD	0.05 M Na Cacodylate pH=6.5, 0.1 M NaI 35% v/v MPD	not needed
Tel23-AuNHC **I**	0.05 M K Cacodylate pH=6.5, 0.05 M Li_2SO_4 25% v/v PEG400	45% v/v PEG400	40% v/v PEG400
Tel24-AuNHC **II**	0.05 M K Cacodylate pH=6.5, 0.05 M Li_2SO_4 0.05 M $MgSO_4$, 20% v/v isopropanol	0.05 M K Cacodylate pH=6.5, 0.05 M Li_2SO_4 0.05 M $MgSO_4$, 20% v/v isopropanol	30% v/v glycerol
CGTACG-$[Pt(bapbpy)]^{2+}$	0.04 M Na Cacodylate pH 7.0, 0.02 M $MgCl_2$, 0.08 M $SrCl_2$ 10% v/v MPD, 0.012 M spermine chloride	30% v/v MPD	25% v/v MPD

Table 11: Data collection statistics for X-ray diffraction experiments carried out on DNA-ligand crystals, part one.

Adduct	X-ray wavelength (Å)	Space group	Lattice	Cell parameters (Å,°)	Resolution range (Å)	Experimental apparatus
Tel12-Coptisine	1.002	$P\,4_3\,2_1\,2$	tetragonal	a=b=41.57, c=66.16 $\alpha = \beta = \gamma = 90$	35.20-1.55	ID29 beamline ESRF
CGTACG-Chelerythrine	1.000	$P\,3_2\,21$	trigonal	a=b=30.23, c=119.25 $\alpha = \beta = 90, \gamma = 120$	20.00-2.10	XRD1 beamline ELETTRA
CGTACG-Coptisine	1.542	$P\,3_2\,21$	trigonal	a=b=30.28, c=118.33 $\alpha = \beta = 90, \gamma = 120$	30.00-2.44	Oxford Diffraction instrument (Cu X-rays source)
CGATCG-Coptisine	1.542	$P\,6_5\,22$	hexagonal	a=b=26.58, c=77.09 $\alpha = \beta = 90, \gamma = 120$	14.78-2.71	Oxford Diffraction instrument (Cu X-rays source)
Tel12-NAX053	1.000	$P\,4_1\,2_1\,2$	tetragonal	a=b=41.13, c=71.20 $\alpha = \beta = \gamma = 90$	35.62-1.70	ID23-1 beamline ESRF
Tel12-NAX075	1.000	$P\,4_1\,2_1\,2$	tetragonal	a=b=41.11, c=68.39 $\alpha = \beta = \gamma = 90$	35.23-1.60	XRD1 beamline ELETTRA
Tel12-NAX077	0.9726	$P\,4_2\,2_1\,2$	tetragonal	a=b=42.10, c=34.15 $\alpha = \beta = \gamma = 90$	29.77-2.04	ID30B beamline ESRF
Tel12-NAX120	0.9726	$P\,4_2\,2_1\,2$	tetragonal	a=b=41.68, c=34.31 $\alpha = \beta = \gamma = 90$	29.47-2.04	ID30B beamline ESRF
Tel23-AuNHC **I**	0.9798	$I\,2_1\,2_1\,2_1$	orthorhombic	a=47.13, b=51.29, c=58.77 $\alpha = \beta = \gamma = 90$	30.00-1.89	BM30A beamline ESRF
Tel24/-AuNHC **II**	0.8729	$C\,2$	monoclinic	a=36.60, b=71.37, c=27.05 $\alpha = \gamma = 90, \beta = 92.42$	35.69-1.80	ID23-2 beamline ESRF
CGTACG-$[Pt(bapbpy)]^{2+}$	1.0718	$P\,3_2\,21$	trigonal	a=b=30.21, c=117.44 $\alpha = \beta = 90, \gamma = 120$	26.16-2.30	ID30B beamline ESRF

Table 12: Data collection statistics for X-ray diffraction experiments carried out on DNA-ligand crystals, part two.

Adduct	Unique reflections	Completeness (%)	Redundancy	R_{sym}	I(*hkl*)/ σ(I(*hkl*))	$CC_{1/2}$
Tel12-Coptisine	8899	99.6	18.7	11.3	15.22	99.7
CGTACG-Chelerythrine	4064	99.0	5.3	0.089	13.39	na
CGTACG-Coptisine	2642	99.5	13.7	0.108	17.48	na
CGATCG-Coptisine	559	99.1	11.1	0.097	16.2	99.9
Tel12-NAX053	7183	99.1	25.5	6.2	20.83	99.9
Tel12-NAX075	8224	99.4	23.1	8.2	19.8	99.4
Tel12-NAX077	2179	98.9	9.5	4.6	21.4	99.8
Tel12-NAX120	2137	99.8	22.8	9.6	26.0	99.8
Tel23-AuNHC **I**	5821	97.8	12.8	6.9	15.68	99.9
Tel24/-AuNHC **II**	6398	99.0	5.6	10.0	8.0	99.6
CGTACG-[Pt(bapbpy)]$^{2+}$	5333	99.1	9.8	15.2	8.4	99.4

Using the anomalous signal, coordinates of the heavy atoms in the crystal cells were obtained in the case of metal complexes with the program Shelx [357]. The X-rays wavelength used for the detection of platinum anomalous scattering was the wavelength corresponding to the L-III edge (CFOM=0.429, pseudo-free CC=0.523). The diffraction experiments carried out on the gold complexes supplied only low representative data for the gold anomalous scattering which however provided useful clues for the positioning of ligand molecules in the electron density maps.
The structures of the adducts were solved in each case by the molecular replacement method using the program Phaser [358] and Molrep [359]. Suitable coordinates of previously reported structures without all the heteroatoms (table 13) were used as search models. The models were refined with the program Refmac5 [360] from the CCP4 program suite [361]. In the case of the CGTACG structures, the refinement also considered the presence of hemihedrally twinning (Chelerythrine adduct: fraction 0.46, law -h -k l; Coptisine adduct: fraction 0.21, law h+k -k -l; $[Pt(bapbpy)]^{2+}$ adduct: fraction 0.26, law -h -k l), as suggested by the phenix.xtriage program from the Phenix program suite [362]. Manual rebuilding of the models was performed using the program Coot [363]. The ligands coordinates were designed by the build module of Maestro [364] and library descriptions were obtained thanks to libcheck from the CCP4 program suite [361] and to eLBOW [365] from the Phenix program suite [362]. Refinement statistics for the crystal structures solved are reported in the Discussion and Result chapter.

Table 13: Initial models used for the solution of DNA-ligand crystal structures by molecular replacement.

Adduct	PDB code
Tel12-Coptisine	1K8P
CGTACG-Chelerythrine	3NP6
CGTACG-Coptisine	3NP6
CGATCG-Coptisine	3FT6
Tel12-NAX053	4P1D
Tel12-NAX075	5CDB
Tel12-NAX077	5CDB
Tel12-NAX120	5CDB
Tel23-AuNHC **I**	3R6R
Tel24/-AuNHC **II**	3R6R
CGTACG-[Pt(bapbpy)]$^{2+}$	3NP6

3.2 COMPUTATIONAL STUDIES

In silico calculations were carried out by using the Schrödinger® program suite [366]. The accuracy of the procedures employed has been proved in a number of cases, as for the simulation of the CGATCG-Coptisine adduct where the calculated ligand pose, QM/MM optimized, was in very good agreement with that determined by experimental crystallographic data [367].

Adducts of natural alkaloids Chelerythrine and Coptisine

The interaction of natural alkaloid Chelerythrine and Coptisine with human telomeric G-quadruplex was simulated at different levels of theory. Ligand molecules were designed by the build module of Maestro [364] and the atomic electrostatic charges were calculated at the B3LYP/6-311G$**+$ (split-valence basis set) level of theory by fitting them to an electrostatic potential calculated using the Jaguar software [368]. The binding ability of the alkaloids was investigated toward the structures of the human telomeric sequences Tel26 d[$A_3G_3(TTAGGG)_3A_2$] (PDB code 2HY9 [185]) and

wtTel26 d[$(TTAGGG)_4T_2$] (PDB code 2JPZ [190]), able to fold in hybrid-1 and hybrid-2 G4 structures, respectively. Rigid docking calculations were performed using Glide [369] with the DNA structures kept fixed in their original conformations. Six different grids, two of them centred nearby the guanine platforms and the other ones nearby the lateral grooves, were prepared for each DNA folding. For each ligand-DNA adduct, poses obtained from all the grids were collected and submitted to a minimization procedure (MacroModel software [370], OPLS_2005 FF [333–335], maximum iteration 3000 cycles, 0.05 kJ/Åmol convergence criterion, guanines restrained) and then sorted as a function of their energy content. Only poses with energy no more than 5 kcal/mol above the minimum were selected. In the case of the results from grids centred nearby the guanine platforms, additional molecular dynamics procedures were carried out (OPLS_2005 FF [333–335], Verlet algorithm, simulation time 4.5 ns, T = 300 K, integration step 1.5 fs, guanines restrained) by means of the Impact software [371].

Adducts of the AuNHC I complex

The binding ability of AuNHC **I** has been investigated towards hybrid-1 (Tel26 d[$A_3G_3(TTAGGG)_3A_2$], PDB code 2HY9 [185]), hybrid-2 (wtTel26 d[$(TTAGGG)_4T_2$], PDB code 2JPZ [190]) and antiparallel-1 basket (Tel22 d[$AG_3(TTAGGG)_3$], PDB code 2MCC [372]) topologies which were modified according to the Tel23 human telomeric sequence, d[$TAG_3(TTAGGG)_3$]. When needed, potassium ions were added in-between the G-tetrads, and each resulting structure was minimized by using the Impact software [371] (OPLS_2005 FF [333–335]). The starting coordinates for the metal complex were obtained from the crystal structure of [Au(9-methylcaffein-8-ylidene)$_2$]BF_4 (CCDC code 885458) [315]. The metal com-

plex cation was fully optimized at the B3LYP/6-311G** + (split-valence basis set) level of theory and the atomic electrostatic charges were calculated by fitting them to an electrostatic potential calculated using the Jaguar software [368]. Rigid docking calculations were performed using Glide [369] with the DNA structures kept fixed in its original conformation. The grids were centred nearby the 5′ end and the 3′ end guanine platforms and also in the four grooves of the G-quadruplex structures. Poses obtained were collected and submitted to Impact [371] molecular dynamics procedures (OPLS_2005 FF [333–335], Verlet algorithm, simulation time 5.0 ns, T = 298.15 K, integration step 1.0 fs), freezing the metal complex coordinates and restraining the guanine residues. Finally, the coordinates provided by this recursive procedure were submitted to a QM/MM calculation (QSite software [373] - metal complex treated at the B3LYP/LACVP level of theory, LACVP: 6-31G basis set for H-Ar atoms while LANL2DZ basis set for heavier elements, OPLS_2005 FF [333–335], all DNA constrained with the exception of the guanine and TTA residues of the binding sites) and then to a QM optimization procedure at the B3LYP/LACVP level of theory, restraining all the DNA atoms by using Jaguar program [368].

Crystal structure of an AuNHC II adduct

The crystal structure of the adduct formed by AuNHC **II** and the Tel24′ human telomeric sequence d[$TAG_3(TTAGGG)_3T$] shows disordered ligand positions which have allowed the localization of only the gold atom of the metal complex in the structure. *In silico* calculations were used to predict the most probable poses of the ligand in the binding site, taking into account the localization of gold atoms. The starting coordinates for the metal complex were obtained from the crystal structure of [Au(bis(1-butyl-3-methyl-

imidazole-2-ylidene))]PF_6 (CCDC code 936581) [317]. The ligand was quantum mechanically optimized as in both *cis* and *trans* conformations at the B3LYP/LACV3P$**+$ level of theory and the atomic electrostatic charges were calculated by fitting them to an electrostatic potential by using the Jaguar software [368]. Rigid docking calculations were performed using Glide [369] with the DNA structure kept fixed in its original conformation. The grids were centred nearby the end G-tetrad where the gold atoms were found. Each ligand molecule was docked one by one on the target restraining the gold atom positions. The resulting adducts were submitted to a QM/MM procedure (QSite software [373] - metal complex treated at the B3LYP/LACVP level of theory, OPLS_2005 FF [333–335], all DNA frozen).

Predicted NMR structure of an Auoxo6 adduct

The binding of Auoxo6 to the human telomeric sequence wtTel26 d[$(TTAGGG)_4T_2$], able to fold into hybrid-2 G4 structure [190], was investigated by *in silico* calculations, taking into account the NMR results which pointed out the significant variation of chemical shifts for the residues at the 5′ end G-tetrad upon ligand addition. Potassium ions were added in-between the G-tetrads, and each resulting structure was minimized by using the Impact software [371] (OPLS_2005 FF [333–335]). The starting coordinates for the metal complex were obtained from the crystal structure of $[(bipy^{2Me})_2Au_2(\mu\text{-}O)_2](PF_6)_2$ (CSD entry QIYZII). The cationic metal complex was optimized at the B3LYP/LACV3P$**+$ level of theory and the atomic electrostatic charges were calculated by fitting them to an electrostatic potential calculated using the Jaguar software [368]. Rigid docking calculations were performed using Glide [369] with the DNA structures kept fixed in its original conformation. The grid was centred nearby the 5′ end guanine

platform. The poses were then scored by evaluating the distance between gold and potassium ions of the channel. Selected poses were submitted to Impact [371] molecular dynamics procedures (OPLS_2005 FF [333–335], Verlet algorithm, simulation time 5.0 ns, T = 298.15 K, integration step 1.0 fs) with suitable restraints for ligand geometry, by means of the Impact software [371]. The obtained coordinates of DNA were then used to carry out a second rigid docking procedure of the QM optimized complex (Glide software [369]), using ligand-based newly calculated grids. The poses finally obtained were submitted to a QM/MM calculation (QSite software [373] - metal complex treated at the B3LYP/LACVP level of theory, OPLS_2005 FF [333–335], all DNA constrained with the exception of the guanine and TTA residues of the binding sites).

Crystal structure of a [Pt(bapbpy)]$^{2+}$ adduct

The coordinates of the adduct formed by [Pt(bapbpy)]$^{2+}$ and the oligonucleotide CGTACG, solved by X-ray crystallography, were used to confirm the stable poses found for the ligand at the binding site. The structure was optimized at QM/MM level of theory (QSite software [373] - metal complex and CGGC quartets of the binding site treated at the B3LYP/LACVP level of theory, OPLS_2005 FF [333–335], constrained all remaining residues).

3.3 SPECTROPHOTOMETRY

Phenylalkyl and diphenylalkyl Berberine derivatives

The interaction of 13-phenylalkyl and 13-diphenylalkyl derivatives of natural alkaloid Berberine with the human telomeric sequence sequence Tel24 d[(TTAGGG)$_4$] was investigated by UV-Vis absorp-

tion and CD spectroscopy.
UV-Vis absorption titrations were performed for ligand solutions with concentrated Tel24 aliquots in 10 mM potassium phosphate buffer pH 7, 150 mM KCl (temperature 25°C) on a Thermo Scientific Evolution 220 UV-vis spectrophotometer (Waltham, MA, USA). Matched quartz cells (Hellma, Germany) of 1 cm path length were used. The Tel24 concentration was determined by using ϵ_{260nm} = 244,600 $M^{-1}cm^{-1}$ while for Berberine and its phenylalkyl and diphenylalkyl semi-synthetic derivatives ϵ_{345nm} = 22,500 M^{-1} cm^{-1} and ϵ_{341nm} = 21,500 $M^{-1}cm^{-1}$ were respectively used, in agreement with previously published data [280, 340, 374–376]. Briefly, small aliquots of Tel24 solution were added to a solution of ligand at known concentration (about 5 μM) until saturation was reached. Equal additions were made to a reference cell during each titration. Readings were noted 5 min after each addition and subsequent mixing in order to guarantee the homogeneous adduct formation. The parameters used for recording spectra are listed as follows: spectral range 400-240 nm, scan speed 100 nm min^{-1}, bandwidth 1 nm, integration time 0.6 s, data interval 1 nm. The variation in absorbance for the ligands at around 340 nm was used for the estimation of dissociation constants and stoichiometry for the G4-ligand binding by using the method described in the UV-Vis absorption spectroscopy section [280, 282, 339, 340].
CD spectroscopy was also used to shed light on the Tel24 topological changes upon ligand binding. Spectra were recorded on a Jasco J-810 Spectropolarimeter (Jasco Cooperation, Tokyo, Japan) equipped with a Peltier temperature controller (model JWJTC-484). Matched quartz cells (Hellma, Germany) of 0.1 cm path length were used. Solutions at known concentration of Tel24 (56.3 μM) were prepared raising the ligand concentration up to quadruplex:ligand ratio of 1:4. Solutions were prepared in 90 mM LiCl, 10 mM KCl and 10 mM lithium cacodylate buffer at pH 7. For each

solution, CD spectra at 25°C and thermal melting curves were recorded. The technical parameters of the experiments are listed as follows. CD spectra: spectral range 400-230 nm, scan speed 100 nm min^{-1}, bandwidth 0.5 nm, data pitch 0.5 nm, time constant 1 s, 6 accumulations. Thermal melting curves: λ=290 nm, temperature slope 90°C h^{-1}, bandwidth 0.5 nm, data pitch 0.2°C, delay time 0 s, time constant 1 s.

Gold metal complex AuNHC I

The interaction between the gold metal complex AuNHC **I** and the human telomeric sequence Tel23 d[$TAG_3(TTAGGG)_3$] was investigated by CD spectroscopy. UV-Vis absorption spectroscopy was not suited for the studies as the absorption bands of ligand and DNA largely overlapped.
CD spectra were recorded on a Jasco J-810 Spectropolarimeter (Jasco Cooperation, Tokyo, Japan) equipped with a Peltier temperature controller (model JWJTC-484). Matched quartz cells (Hellma, Germany) of 0.1 cm path length were used. All experiments were carried out in 90 mM LiCl, 10 mM KCl and 10 mM lithium cacodylate buffer pH 7.2. Briefly, solutions of 60 μM Tel23 were prepared in absence or presence of 240 μM AuNHC **I** (Tel23:ligand stoichiometric ratio 1:4) keeping the DMSO concentration fixed at 3% v/v (DMSO used for the preparation of AuNHC **I** stock solution). The solutions were stored in incubator at 25°C and CD spectra were collected on them at different times of incubation at 25°C. A second set of experiments was then performed by preparing solutions of 60 μM Tel23 in presence of different concentrations of AuNHC **I** (DMSO concentration 3% v/v) and keeping them in incubator at 25°C for 72 h. CD spectra were collected at 25°C and the changes at 268 nm were analysed by Benesi-Hildebrand plot [347–350], already described in the Circular Dichroism sec-

tion, considering three distinct binding events. The values in ellipticity at 268 nm were blank-corrected prior to the fitting. Thirdly, thermal melting curves were recorded for analogous solutions at known ligand:Tel23 stoichiometric ratios (0, 0.5, 1, 1.5, 2, 2.5, 3, 3.5, 4, 5) at 290 nm. The technical parameters of the experiments are listed as follows. CD spectra: spectral range 400-220 nm, scan speed 100 nm min^{-1}, bandwidth 0.5 nm, data pitch 0.5 nm, time constant 1 s, 6 accumulations. Melting curves: temperature range 25-90°C, monitor wavelength 290 nm, temperature slope 60°C h^{-1}, bandwidth 0.5 nm, data pitch 0.2°C, delay time 0 s, time constant 1 s.

4 RESULTS AND DISCUSSION

The stabilization of non-canonical DNA structures, as human telomeric G-quadruplexes, is a promising route in modern drug discovery, especially for the identification of efficient anticancer agents with reduced side effects in comparison to classical chemotherapeutics. The achievement of detailed structural data about the ligand-DNA interactions is crucial for a better understanding of the binding properties of the targets and also to give suggestions for the design of further ligand candidates. Due to the accuracy of results, X-ray crystallography has been the principal technique used to shed light on the ligand-DNA interactions. Additional *in silico* calculations have been performed in order to complement crystallographic data when needed, especially for the simulation of adducts formed in solution, where different G-quadruplex topologies are found with respect to the solid state. Spectrophotometric characterization has been carried out to supply informative data about the formation of adducts in solution and so expanding the general picture emerged from structural investigations. Profitable scientific collaborations with other research groups have been developed in order to perform solution studies and biological assays, able to assess the cytotoxic activity of selected compounds against cancer cells lines. The results obtained during the PhD project are discussed in the following sections.

4.1 NATURAL ALKALOIDS

Natural alkaloids Berberine, Coptisine, Sanguinarine and Chelerythrine (figure 33) are biologically active substances whose anticancer activity is thought to be connected to protein and DNA binding, with particular mention of the G-quadruplex DNA folding [276–283]. A general picture of the interaction of these alkaloids with G4 DNA is, however, still lacking, in particular for Coptisine and Chelerythrine. The attention of the research activity was caught by these two compounds which are structural analogs of the widely studied Berberine and Sanguinarine, respectively. The interest concerned the potential binding properties of Coptisine and Chelerythrine towards human telomeric DNA which could result in potent telomerase inhibition activity. In order to supply an accurate view of the interaction of the alkaloids with human telomeric sequences, a multi-approach investigation has been carried out during the PhD project, including spectrophotometric characterization, *in silico* calculations and X-ray crystallography [377].

4.1.1 Solution studies

Studies on the ligand-G4 interaction in solution were performed in the research group of Professor Claudia Sissi from the University of Padua, thanks to a scientific collaboration. In the first place, fluorescence melting assays confirmed the strong interaction between the alkaloids and telomeric G-quadruplexes in potassium solutions, and pointed out the G4 thermal stabilization trend Sanguinarine>Chelerythrine≈Coptisine>Berberine (figure 34).
The Sanguinarine scaffold is thus more efficient in the stabilization than that of Berberine, and in addition it appears how the closure of the two methoxy functions into a dioxolo ring (figure

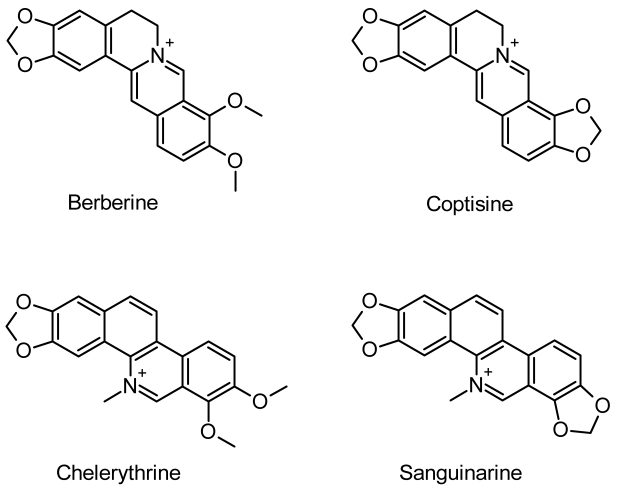


Figure 33: Natural alkaloids Berberine, Coptisine, Sanguinarine and Chelerythrine.

33) improves the adducts stability for both the structural analogs Sanguinarine-Chelerythrine and Berberine-Coptisine. This observation is in according with a previously reported evidence indicating that the presence of benzodioxole groups in alkaloid derivatives promotes the interaction with G-quadruplex structures from human telomeric DNA [279, 284].
Further experiments in the same conditions pointed out the interesting features of the alkaloids for the interaction with different DNA foldings (figure 34). While the compounds significantly interact with G-quadruplex foldings, they scarcely stabilize duplex DNA, hence confirming their promising G4 binding properties. Additionally, the alkaloids display better stabilization of the hybrid telomeric topologies found in potassium solutions with respect to the antiparallel folding observed in sodium solutions, in according with precedent results reported for Chelerythrine [281]. Although Sanguinarine induces the greatest thermal stabilizations,

the alkaloid is the less capable to distinguish different DNA structures, probably because of its significant propensity to act as efficient intercalator, thanks to its planar conformation, and also as groove binder [273, 287, 378]. The more selective behaviour observed for the other alkaloids likely relies in their ability to mainly interact through π-π stacking with extensive aromatic surfaces formed by DNA bases thus leading to a relevant discrimination of quadruplex over duplex foldings. Overall, Chelerythrine and Coptisine are the most interesting compounds of the two different alkaloid scaffolds considered.

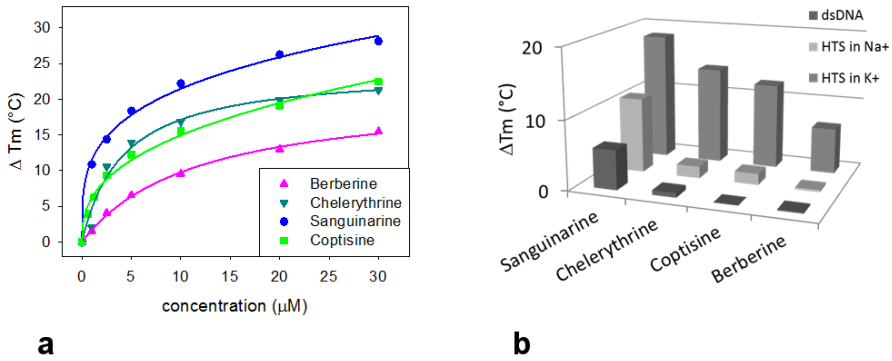


Figure 34: Fluorescence melting assays: a) thermal stabilization of the human telomeric sequence HTS, d[Dabcyl-$AG_3(TTAGGG)_3$T-Fam], in presence of increasing concentrations of the alkaloids in 10 mM LiOH, 50 mM KCl, pH 7.5; b) shift of the melting temperature induced by 5 μM ligands on calf thymus DNA and on HTS either in 50 mM KCl or 50 mM NaCl.

In order to better characterize the interaction of Chelerythrine and Coptisine with the human telomeric topologies found in potassium solutions, fluorescence intercalator displacement (FID) assays were carried out using sequences able to predominantly fold into hybrid-1 and hybrid-2 G4 structures (table 14), namely Tel26

d[$A_3G_3(TTAGGG)_3A_2$] and wtTel26 d[$(TTAGGG)_4T_2$] [185, 190]. Briefly, a solution at known concentration of DNA and of the dye thiazole orange was prepared and the corresponding fluorescence spectrum was acquired in the absence and presence of increasing concentrations of the tested alkaloids. The percentage of dye displacement was calculated as θ = 100 - $[(F/F_0) \cdot 100]$, where F_0 is the fluorescence in the absence of ligand and F the fluorescence recorded at each point of titration. The plot θ vs the compound concentration gives EC_{50} as the half maximal effective concentration, which supplies an indication of the binding strength of the ligands towards the G-quadruplexes.
The alkaloids were able to displace the dye and interact with both the topologies. In more detail, the assays highlighted a weaker binding of Coptisine to both the structures with respect to Chelerythrine, although Coptisine showed a more discriminatory binding to the G-quadruplex forms considered.

Table 14: EC_{50} derived by FID for the interaction of Chelerythrine and Coptisine with the hybrid-1 and hybrid-2 human telomeric foldings in 10 mM Tris, 50 mM KCl, pH 7.5. Thiazole orange was used as a probe, excitation wavelength 501 nm. Each titration was repeated at least in triplicate.

Ligand	Hybrid-1 (μM)	Hybrid-2 (μM)	Ratio H2/H1
Chelerythrine	2.59±0.10	3.75±0.24	0.7
Coptisine	12.01±0.82	5.93±0.36	2.0

CD titrations were performed to evaluate the binding of Chelerythrine and Coptisine to the tested sequences (figure 35). The variations of the G4 bands centred at 295 nm and 260 nm upon ligand addition confirmed the formation of adducts, however small differences in the spectra at saturating ligand concentrations suggested that the alkaloids bind hybrid G-quadruplexes in a slight different manner.

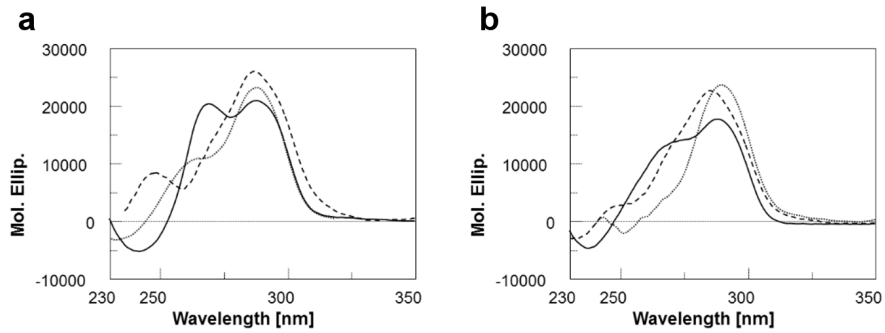


Figure 35: CD spectra of 4 μM (a) Tel26 (hybrid-1) and (b) wtTel26 (hybrid-2) in 10 mM Tris, 50 mM KCl, pH 7.5, in the absence (solid lines) or in the presence of 40 μM Chelerythrine (dashed lines) or Coptisine (dotted lines).

4.1.2 *In silico* calculations

The demand for structural data about the interaction of Chelerythrine and Coptisine with telomeric G-quadruplex resulted in the development of *in silico* study plans. Ligand molecules were designed by means of the build module of Maestro [364], while the NMR structures of Tel26 d[$A_3G_3(TTAGGG)_3A_2$] (PDB code 2HY9 [185]) and wtTel26 d[$(TTAGGG)_4T_2$] (PDB code 2JPZ [190]) sequences, able to fold in hybrid-1 and hybrid-2 G4 structures, respectively, were taken into account for the simulations.
Firstly, ligand molecules were QM optimized at the B3LYP/6-311-G$**+$ level of theory and atomic electrostatic charges were calculated. The achievement of accurate charges is quite important for the accuracy of *in silico* simulations. Because QM charges can give far better results than those calculated by MM methods, they were used for all the subsequent calculations. Rigid docking calculations were carried out through the definition of grids centred nearby the external G-tetrads and the grooves of the targets. For

each adduct, poses obtained from all the grids were collected and submitted to a MM minimization procedure (OPLS_2005 FF) and then sorted as a function of their energy content. This task allows the energetic comparison of ligand poses from different grids which otherwise could not have been confronted, hence allowing to point out the most probable binding sites of the ligands on the targets. Selected poses in the proximity of external G-tetrads underwent additional molecular dynamics procedures in order to favour a better allocation of the ligands. In fact, the lowest energy structures pointed out the localization of the ligands at the 5′ end G-tetrad (figure 36 and 37) which is considerably hindered by bases from loops and flanking residues [185, 190]. In fact, the dynamics are indispensable for the ligand to insinuate in the capping architectures and eventually reach the external G-tetrads.

Despite the not completely planar conformation (naphthalene and benzodioxolo groups form a dihedral angle of about 13°), Chelerythrine is able to efficiently interact through π-π stacking at the 5′ end quartet of the hybrid-1 structure, especially with G4 and G22 residues (inter-planar distances 3.4-3.9 Å, figure 36). The ligand is also involved in stacking interactions with the A3 residue belonging to the flanking bases. It is noteworthy the presence of a salt-bridge between the positively charged nitrogen atom of the alkaloid and the phosphate oxygen of the G22 residue (inter-atomic distance 3.5 Å) which contributes to overall adduct stability. On the other hand, Chelerythrine tends not to directly stack on the 5′ end G-tetrad of the hybrid-2 folding and it is sandwiched between T2 and A15 residues (inter-planar distances 3.4-3.7 Å, figure 36) assuming a similar conformation with respect that calculated for the interaction with Tel26 (naphthalene and benzodioxolo groups form a dihedral angle of about 10°).
As shown in figure 37, Coptisine shows a quite bent conformation when bounded to hybrid-1 (quinoline and benzodioxolo groups

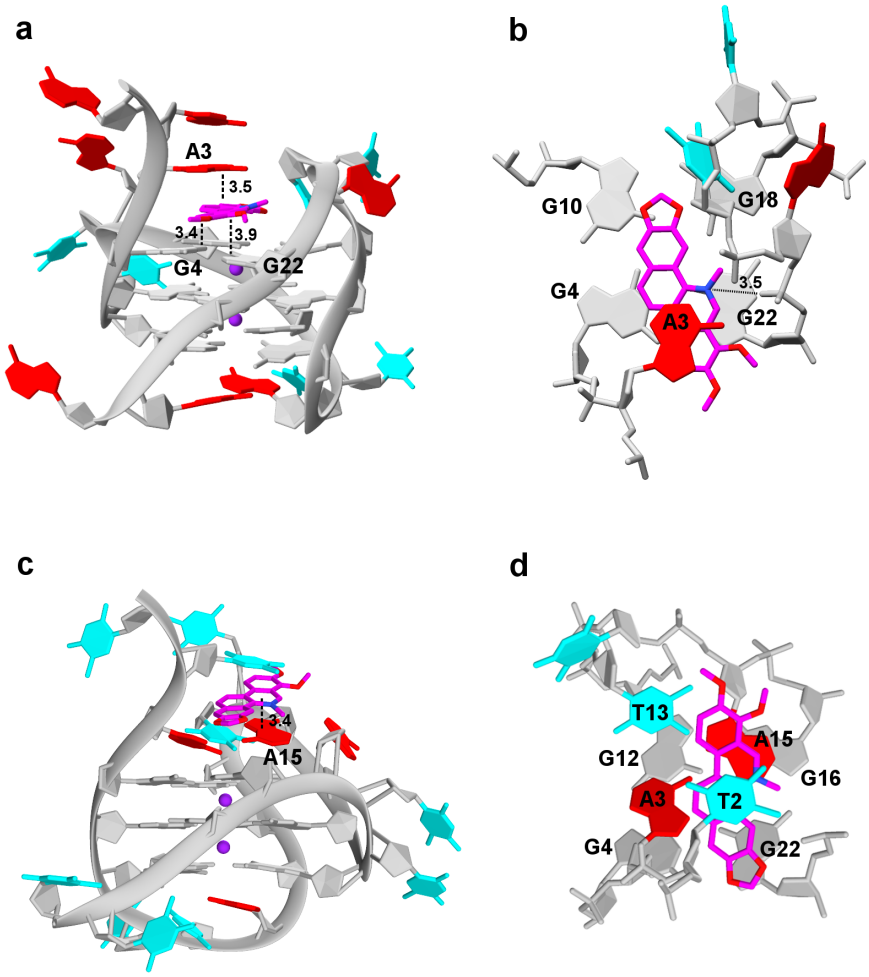


Figure 36: Calculated structures of the adducts formed by Chelerythrine with hybrid-1 (a,b) or hybrid-2 (c,d) G-quadruplex. Adenine, thymine, guanine and potassium ions coloured in red, cyan, gray and purple, respectively.

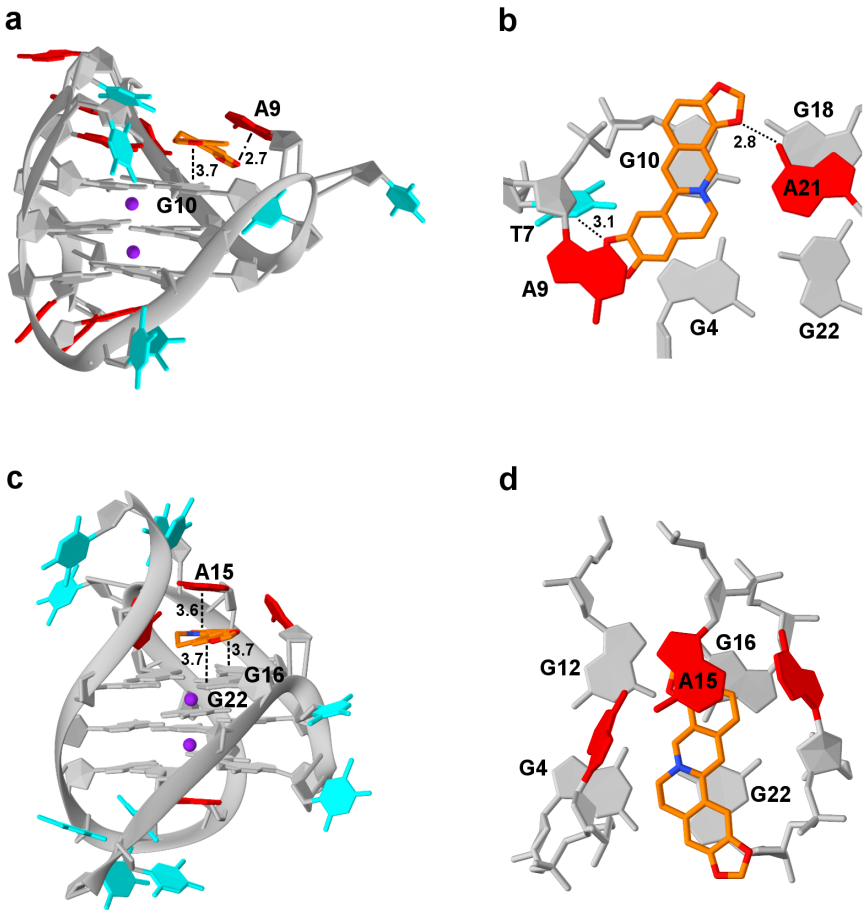


Figure 37: Calculated structures of the adducts formed by Coptisine with hybrid-1 (a,b) or hybrid-2 (c,d) G-quadruplex. Adenine, thymine, guanine and potassium ions coloured in red, cyan, gray and purple, respectively.

form a dihedral angle of about 23°) which prevents optimal stacking at the 5′ end G-tetrad, as the only significant π-π stacking interaction is observed with the G10 residue (inter-planar distance 3.7 Å). However, the ligand interacts through the dioxolo moieties by hydrogen bonds with T7 and A21 residues (inter-atomic distances 3.1 Å and 2.8 Å, respectively) and also by a CH$\cdots\pi$ contact with the A9 residue (carbon-centroid distance 2.7 Å). Conversely, Coptisine shows a less bent conformation in the hybrid-2 adduct (quinoline and benzodioxolo groups form a dihedral angle of about 18°, figure 37). The alkaloid interact by π-π stacking with G16 and G22 residues belonging to the 5′ end quartet (inter-planar distances 3.7 Å). Similar to Chelerythrine, the alkaloid is involved in stacking interactions also with the A15 residue of the flanking bases (inter-planar distance 3.6 Å). Overall, Coptisine seems to be better suited for the binding of the human telomeric hybrid-2 structure, in according to the FID results which evidenced a remarkable selectivity of Coptisine for hybrid-2 over hybrid-1 topology.

4.1.3 X-ray crystallography

Crystallization screenings were performed for the adducts formed by Coptisine and Chelerythrine and the human telomeric sequences Tel12 d[$TAG_3TTAGGGT$] and Tel23 d[$TAG_3(TTAGGG)_3$] folded into G-quadruplex. Crystals suitable for X-ray diffraction analysis were obtained only for the adduct formed by Tel12 and Coptisine. The crystal structure was subsequently solved and the refinement statistics are reported in table 15.

The crystal structure shows parallel bimolecular G-quadruplexes and ligand molecules in 1:1 stoichiometric ratio. The quadruplex units are head-to-tail disposed and give rise to columns growing along the 4-fold crystallographic axis (figure 38). Each G4 unit is

Table 15: Crystal data and refinement statistics for the Tel12-Coptisine structure.

Space group	P $4_3 2_1 2$
Crystal system	tetragonal
Cell dimension (Å)	a=b=41.57, c=66.16
Cell angles (°)	$\alpha = \beta = \gamma = 90$
Resolution range (Å)	35.20-1.55
Rfactor	0.232
Rfree	0.259
RMSD bonds (Å)	0.009
RMSD angles (°)	2.472
PDB code	4P1D

composed of three stacked G-tetrads (inter-planar distances 3.4 Å) of two symmetry independent chains and features TTA propeller loops which can be described as type 5 and type 1 (subtype 2), according to the analysis reported by Neidle and coworkers [379]. The type 5 arrangement involves the stacking of the second thymine (TTA) on the adenine whereas the first thymine (TTA) is directed towards the G-tetrads. On the other hand, the type 1 (subtype 2) configuration shows the stacking of the first thymine on the adenine, while the second thymine is approximately stacked on the external face of the adenine. The type 1 loop arrangement has been observed in the majority of the reported crystal structures of the human telomeric G-quadruplex and it appears to be a stable and energetically favourable arrangement which is independent on the crystal packing mode and has also been observed for human telomeric sequences in solution by NMR spectroscopy [379]. In fact, it is hypothesized that such loop configuration could be biologically relevant for telomeric G4 DNA.
Similarly to previously reported structures [380], potassium ions are found in the internal channel of the tetraplex located between two adjacent G-quartets at 2.7-3.0 Å apart from the guanine O6 atoms which determine an antiprismatic coordination sphere.

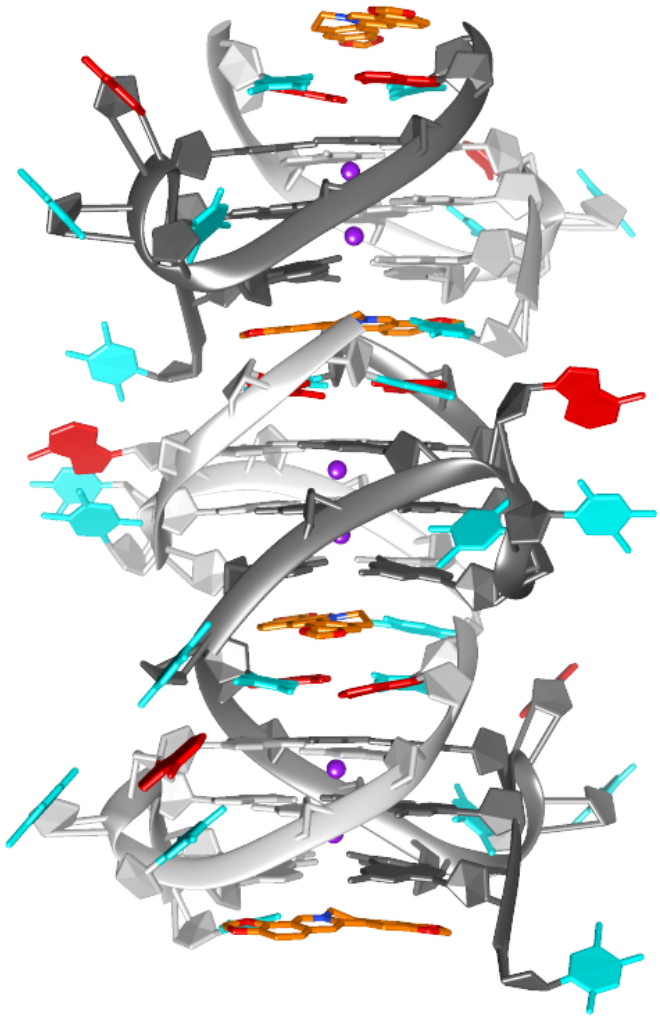


Figure 38: Columnar organization formed by stacked units in the crystal packing of the Tel12-Coptsine crystal structure.

As shown in figure 39, the binding site of the ligand (omit map, figure 40) is found at the junction between adjacent G4 units and it is defined by the 3′ end G-tetrad and the TATA quartet formed at the 5′ end by adenine and thymine bases interacting via Watson-Crick hydrogen bonds (distances range from 2.7 Å to 3.1 Å). Unlike the guanines of the G-tetrad which are almost perfectly coplanar, adjacent thymine and adenine involved in the TATA quartet form significant dihedral angles, in the range 14-23°. Interestingly, this kind of quartets arrangement has already been observed in the crystal structure of the G4 binder BRACO-19 bounded to a human telomeric sequence [381]. The quite planar conformation of Coptisine (quinoline and benzodioxolo groups form a dihedral angle of about 12°), often observed for protoberberine alkaloids at the solid state due to crystal packing forces [382–384], allows the molecule to interact through π-π stacking with bases from the quartets (inter-planar distances 3.4 Å), in particular with G5, G23, A2 and A14 residues, and also through CH···O hydrogen bonds with the T12 residue which is directed towards the ligand (interatomic distances 3.3-3.4 Å). The asymmetric localization of the alkaloid on the quartets makes the positively charged nitrogen atom of the molecule be located approximately in line with potassium ions in the internal channel of the quadruplex, where a negative electronic density is present.

The crystal structure solved shows Coptisine molecules located at the junctions between adjacent G-quadruplexes so defining a higher order organization which could be actually found at human telomeres, where dozens of TTAGGG repeats can form multiple G4 structures. As previously shown in figure 19, tandem G-quadruplexes could stack one on another at their 3′ and 5′ ends, in a fashion similar to what observed in the Tel12-Coptisine structure. Furthermore, recently reported crystal structures have highlighted the formation of analogous thymine and adenine organizations at

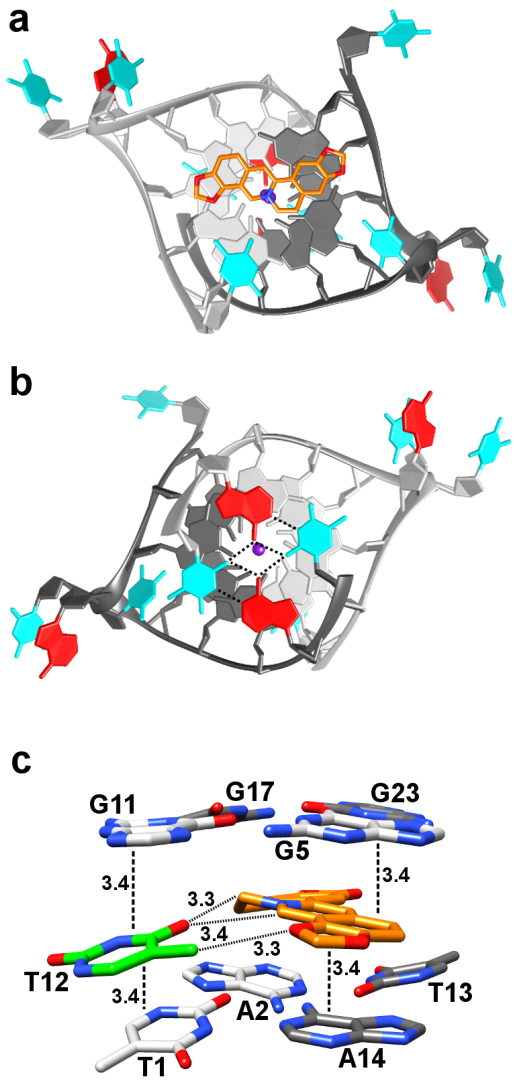


Figure 39: Crystal structure of the Tel12-Coptisine adduct. Top view of the 3′ end G-tetrad (a) and of the TATA quartet at the 5′ end (b). View of the binding site (c). Dotted lines in (b) indicate hydrogen bonds in the TATA quartet (distances range from 2.7 Å to 3.1 Å). The two symmetry independent Tel12 strands are light and dark grey coloured. Thymine and adenine residues are coloured in cyan and red, respectively.

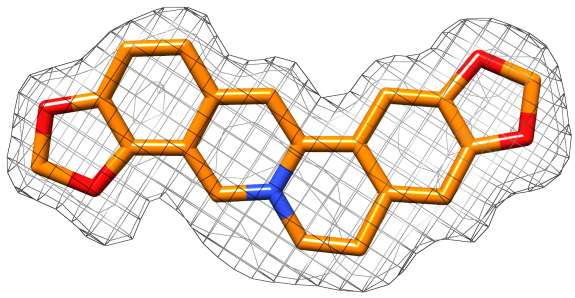


Figure 40: Coptisine shown within the electron density map [F_o-F_c] at 2.0σ level.

telomeric quadruplex/duplex junctions, thus supporting the idea that TATA tetrads could really be found at potential binding sites of human telomeres [385].
The structure obtained for Coptisine allows a comparison with that previously solved for the natural alkaloid Berberine bounded to the human telomeric sequence Tel23 d[$TAG_3(TTAGGG)_3$] [288]. Berberine was localized at two binding sites, one defined by two 5′ end G-tetrads and another by the 3′ end G-tetrad and a TA base pair from a symmetry related G-quadruplex. Both the sites accommodated two coplanar Berberine molecules. Upon inspection of the 3′ end binding site in the two crystal structures (figure 41), it can be observed how Berberine and Coptisine interact similarly by π-π stacking with two guanines of the quartet (interplanar distances 3.4 Å), although the orientation of the molecules is quite different. While the positively charged nitrogen atom of Berberine is directed towards the TTA loops, Coptisine lies on the quartet having the charged nitrogen above the internal channel of the G-quadruplex, where a negative electron density is present. As a consequence, two Berberine molecules can fill the binding site, instead in the case of Coptisine just a ligand molecule and an additional thymine residue are able to stack on the tetrad. Such

residue is in contact with Coptisine through CH···O hydrogen bond involving one of the dioxolo rings (inter-atomic distances 3.3-3.4 Å, figure 41). These findings support the idea that dioxolo rings are important for establishing favourable ligand-DNA interactions and allowing a better allocation of the ligands in the binding sites.

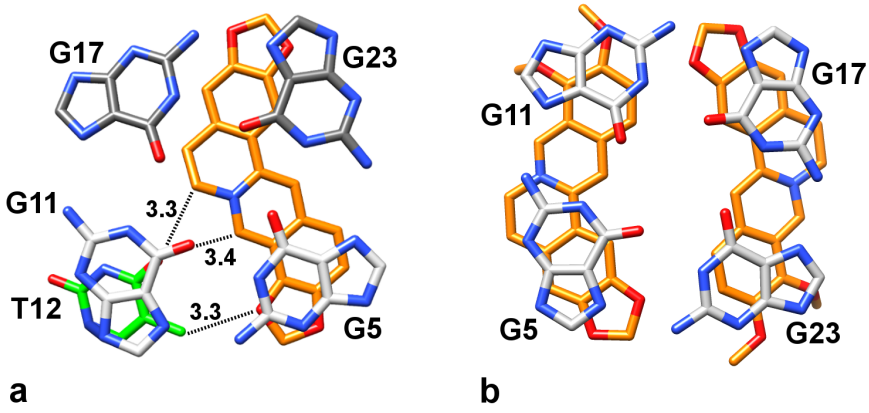


Figure 41: Binding site of Coptisine (a) and Berberine (b) at the 3′ end G-tetrad in their respective crystal structures with Tel12 and Tel23.

The studies on the interaction of Sanguinarine, Chelerythrine, Coptisine and Berberine with human telomeric G-quadruplex have confirmed the promising binding properties of these compounds and also highlighted the interesting features of Chelerythrine and Coptisine to selectively stabilize the topologies found in physiological conditions. Such alkaloids are able to bind hybrid-1 and hybrid-2 tetraplexes and the structures of the adducts have been calculated *in silico* in order to rationalize the slight differences observed for the binding strengths (table 14). The crystal structure of the adduct formed by Coptisine and the human telomeric sequence Tel12 has provided further information regarding the

binding mode of these compounds. From these data, it emerges how the dioxolo rings play an important role in the formation of DNA adducts for these scaffolds as these groups are involved into favourable interactions and presumably allow a better allocation of the molecules in the binding sites. Furthermore, the different binding strengths and thermal stabilizations shown by the ligands are likely correlated to the degree of planarity of such scaffolds that is however greatly influenced by the interactions in which the alkaloids are involved, as shown by the crystal structure of Tel12-Coptisine adduct.

4.1.4 Interaction with short oligonucleotides, X-ray crystallography

The ability of biologically relevant sequences to give rise to non-guanine quartets has been highlighted in several works [128, 386–388] regarding in particular the formation of alternative homo-quartets (involving a single kind of nitrogenous base) and various hetero-quartets, such as GCGC, GTGT and ATAT [109, 389, 390]. This significant structural variability is due to the high flexibility of the DNA backbone and also to the multiple hydrogen bond donor and acceptor atoms in purines (figure 42). Adenines display indeed donor and acceptor atoms on both the Watson-Crick (see also figure 6) and Hoogsteen (figure 10) faces, while guanines show donors only on the Watson-Crick face, albeit they have acceptors on their Hoogsteen face.

The formation of base mismatches is usually associated to DNA instability [42], although, for instance, the existence of transient Hoogsteen base pairs having comparable stability with respect to Watson-Crick pairings has been reported [391]. Actually, the search of small molecules able to preferentially bind mismatch sites over fully matched DNA tracts is a well-established research

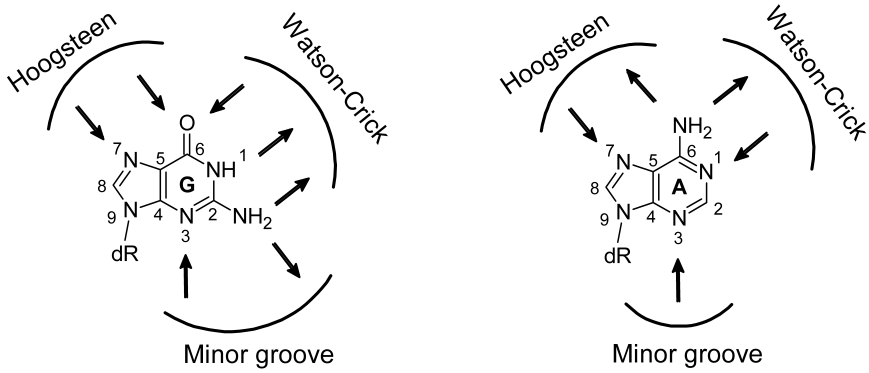


Figure 42: Watson-Crick, Hoogsteen and minor groove faces in purines, guanine **G** and adenine **A**.

area in view of their potential as drug candidates or tools for chemical biology *in vivo* [392, 393].
These suggestions prompted the investigation of Chelerythrine and Coptisine for the interaction with d[CGTACG] and d[CGATCG] oligonucleotides, able to give rise to local non-B DNA arrangements. Crystallization screenings for the ligand-DNA adducts were carried out in order to supply a more general picture of the ligands binding properties. The crystal structures of the adducts CGTACG-Chelerythrine, CGTACG-Coptisine and CGATCG-Coptisine were successfully solved (table 16) [367].

Looking at first at the CGTACG adducts, it can be easily appreciated the great similarity of the Coptisine and Chelerythrine crystal structures which show double-stranded DNA units and ligand molecules in 2:1 stoichiometric ratio (figure 43). The crystal packings reveal the organization of helices in columns growing along the [100], [010] and [1-10] crystallographic directions. The asymmetric units contain two double-stranded helices where the central four base pairs give rise to a continuous array of eight stacked pairings adopting a B-like motif, while the terminal bases of each

Table 16: Crystal data and refinement statistics for the structures of CGTACG-Chelerythrine, CGTACG-Coptisine and CGATCG-Coptisine adducts.

	CGTACG-Chelerythrine	CGTACG-Coptisine	CGATCG-Coptisine
Space group	P 3_221	P 3_221	P 6_522
Crystal system	trigonal	trigonal	hexagonal
Cell dimension (Å)	a=b=30.23 c=119.25	a=b=30.28 c=118.33	a=b=26.58 c=77.09
Cell angles (°)	$\alpha=\beta=90$ $\gamma=120$	$\alpha=\beta=90$ $\gamma=120$	$\alpha=\beta=90$ $\gamma=120$
Resolution range (Å)	20.00-2.10	30.00-2.44	14.78-2.71
Rfactor	0.220	0.261	0.267
Rfree	0.314	0.307	/
RMSD bonds (Å)	0.026	0.008	0.013
RMSD angles (°)	1.906	1.955	2.922
Twin law	-h -k l	h+k -k -l	/
Twin fraction	0.46	0.21	/
PDB code	4D9Y	4D9X	4L5K

helix are involved in different kinds of contacts. In more detail, the junction between contiguous double helices shows the insertion of the 3′ end G12 residues (chain B,D) in the minor groove of adjacent duplexes in the column, while the 5′ end C5 residues (chain A,C) were disordered and so not introduced in the model. In both the structures, the ligand molecules are found at the interface between helices of adjacent asymmetric units, located in an intercalation site defined by six bases. As shown in figure 44 and 45, the binding sites display a non-canonical arrangement composed of the 3′ end C5 and G6 residues (chain A,C) and also of the 5′ end G8 residues (chain B,D), analogously to what observed in alternative CGTACG adducts by other authors [394]. Interestingly, G6 and G8 (chains A,D and C,B) interact through hydrogen bonds involving the minor groove faces of guanines (inter-atomic distances 2.7-3.3 Å). Both the alkaloids (omit maps, figure 46) interact through π-π stacking with G6 and G8 residues (chain C and B,D respectively) and Chelerythrine in addition with the C5

residue (chain A) (inter-planar distances 3.3-3.5 Å). The positively charged nitrogen atoms of the ligands are close to the bases and seem to strengthen the overall binding. It is noteworthy that the 5′ end C7 residues (chain B,D) establish contacts with adjacent not parallel DNA columns in each structure.

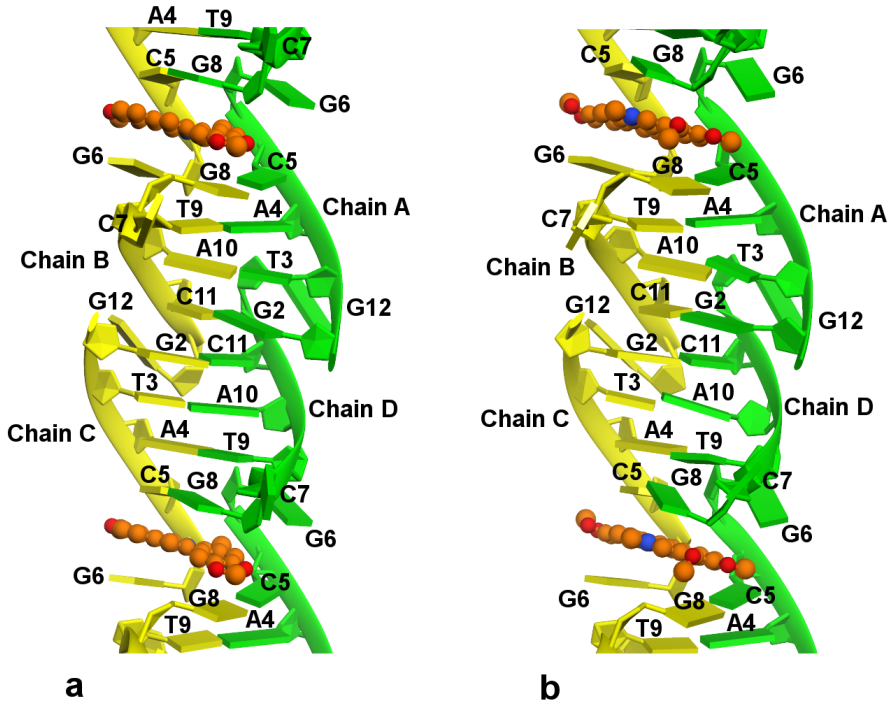


Figure 43: Crystal structures of CGTACG bounded to Coptsine (a) and Chelerythrine (b).

The crystal structures obtained for Chelerythrine and Coptisine bounded to the CGTACG sequence are isomorphic and isostructural with those previously solved for the structural related alkaloids Berberine and Sanguinarine (figure 33) [395]. The binding of the alkaloids at the interfaces between adjacent helices, where

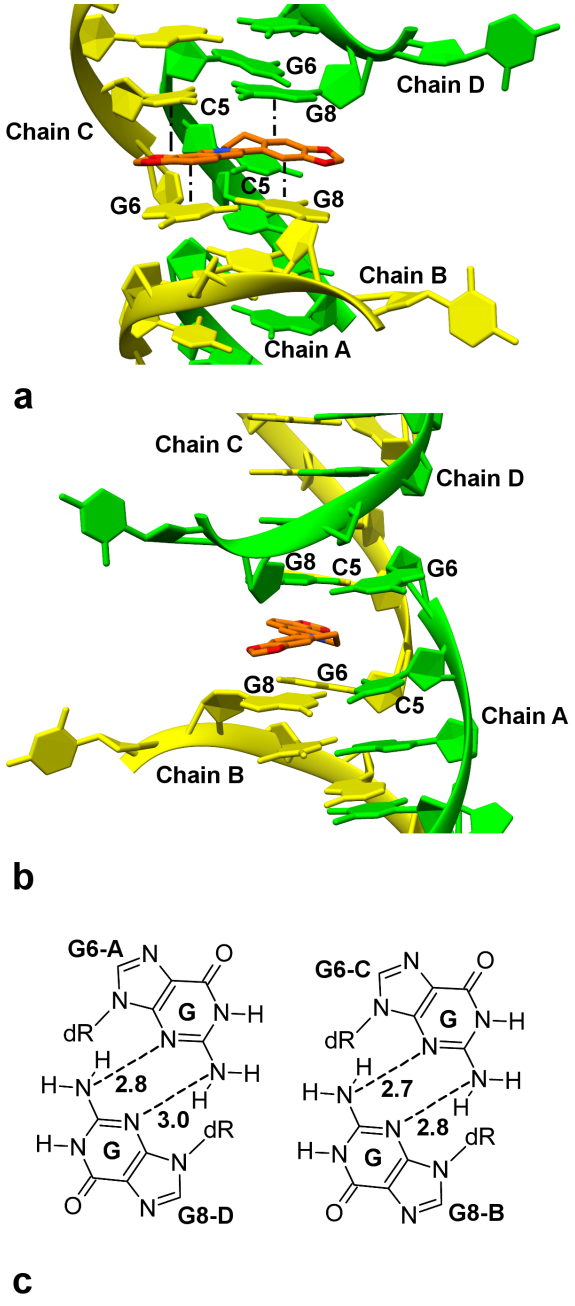


Figure 44: Binding site of Coptisine from frontal and lateral views (a,b). Details of G•G interactions at the site (c).

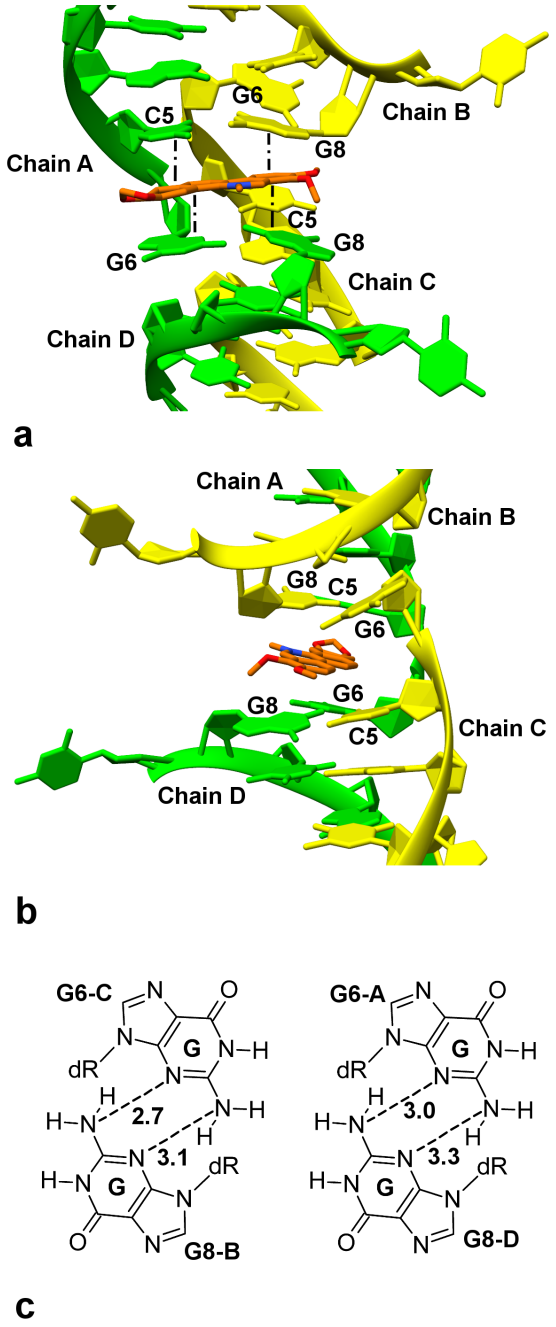


Figure 45: Binding site of Chelerythrine from frontal and lateral views (a,b). Details of G•G interactions at the site (c).

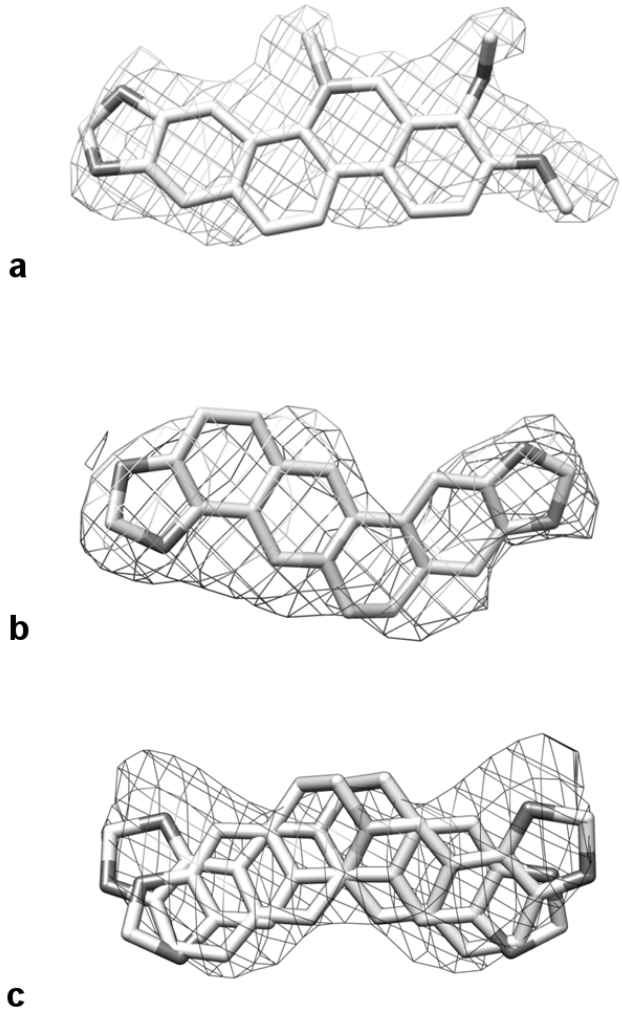


Figure 46: Chelerythrine (a, CGTACG adduct) and Coptisine (b, CGTACG adduct; c, CGATCG adduct) shown within the electron density map [F_o-F_c] at 1.5σ level.

six-bases arrangements are observed, is indeed quite similar. It is questionable if such organizations are largely caused by crystal packing forces than by the binding of the alkaloids itself, however the crystallization of the CGTACG sequence with anthraquinone and acridine derivatives in alternative space groups and crystal packings [396] seems to be suggesting that the positions observed for the alkaloids in the structures are truly representative of their binding modes. It is noteworthy that Coptisine is quite bent when bounded to CGTACG (quinoline and benzodioxolo groups form a dihedral angle of about 18°) while Chelerythrine, Sanguinarine and also Berberine are essentially planar.

Moving on to the next structure, the CGATCG-Coptisine adduct displays double-stranded DNA helices and ligand molecules in 1:2 stoichiometric ratio (figure 47). The strands of the helices are related by a two-fold symmetry axis and give rise to two symmetry equivalent binding sites at the CpG steps. The ligand undergoes intercalation in each binding site and its orientation is orthogonal both to the plane formed by the bases and to the axis of the helix, which adopts a distorted B-type configuration. The two-fold symmetry defines the formation of CGGC quartets at the binding sites in which the C1, G2, C5 and G6 residues from adjacent double helices are involved (figure 48). Interestingly, the G2 and G6 residues interact through their minor groove faces by hydrogen bonds (guanines form a dihedral angle of about 32°, inter-atomic distances 2.5-2.7 Å) while C1-G2 and C5-G6 base pairs are not coplanar (dihedral angle of about 46°). Coptisine is stacked between the CGGC quartets and shares two symmetry related positions (figure 47). The ligand interacts by π-π stacking in particular with G2 and G6 residues (inter-planar distances 3.5 Å, figure 48), thanks to its quite bent conformation (quinoline and benzodioxolo groups form a dihedral angle of about 27°). Additional O$\cdots\pi$ interactions are also engaged by the dioxolo groups with C1 and

C5 residues (atom-centroid distance 2.9 Å, figure 48). Overall, the binding of Coptisine seems to be driven by Van der Waals and electrostatic forces, which involve the positively charged nitrogen atom of the alkaloid.

The crystal structures of the adducts formed by Chelerythrine and Coptisine with the CGTACG and the CGATCG sequences show local non-canonical arrangements where guanines are involved in hydrogen bonds by their minor groove faces. Guanines show fairly good propencity to interact with adenines through their minor groove faces (figure 42), resulting in the formation of G•A mismatches which give very stable structures when present in tandem repeats [397]. Conversely, an isolated G•G mismatch has only been observed in the crystal structure of the d[GCGAAAGCT] sequence [398]. It is noteworthy how G•G interactions are present in tandem in the present CGATCG-Coptisine structure and give rise to a quadruplex platform. A similar CGGC organization has been observed at the junctions between stacked helices featuring CGAA fragments, hence confirming the potential local formation of such assemblies in B-like helices [399]. Interestingly, the platform in both the structures shows an antiparallel (up-down-up-down) arrangement (figure 48). Also, the bases of the quartets are similarly not coplanar and the greater inclination observed in the CGATCG-Coptisine adduct (the Watson-Crick G≡C base pairs from the interacting symmetry related double helices form a dihedral angle of about 46°, while the same angle is about 30° in the previously reported structure) could be explained as a consequence of ligand intercalation.

Recalling the results concerning human telomeric G-quadruplex foldings, it appears that the alkaloids adapt their conformation in order to establish favourable interactions with DNA bases. When the stacking on planar platforms of bases is possible, less bent conformations seem to be favoured. The planar conformations dis-

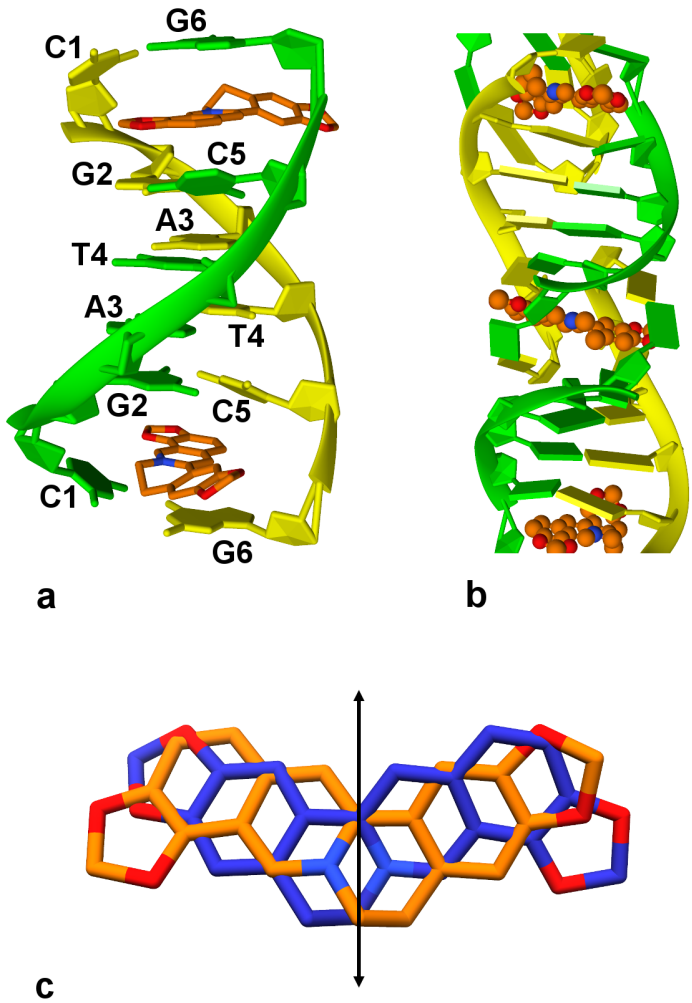


Figure 47: Crystal structure of the CGATCG-Coptisine adduct (a) and columnar organization of the DNA helices (b). Details of the ligand disorder at the binding site (c).

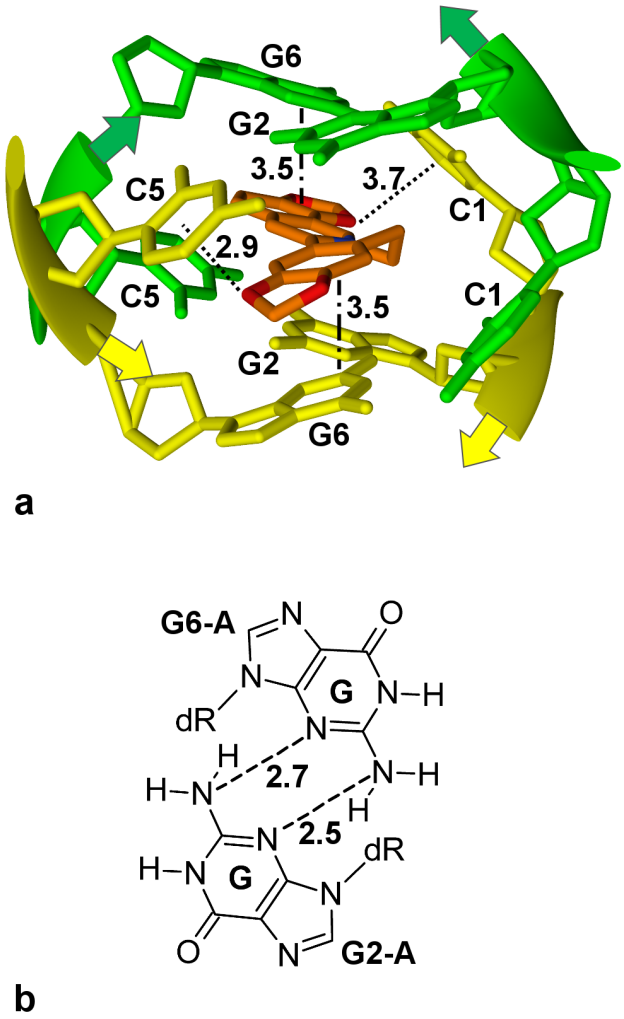


Figure 48: Binding site of Coptisine on the CGGC quartets (a) together with details of local G•G interactions (b).

played by Berberine and analogous compounds at the solid state support this hypothesis [382–384]. On the other hand, bent ligand conformations have been observed in association with non-planar organizations of bases, such as the CGGC quartet in the CGATCG-Coptisine structure and the TATA tetrad in the Tel12-Coptisine adduct. The ability of Coptisine to easily adopt different conformations has been proven by performing QM calculations on 0°, 11° and 30° bent conformers which have shown an overall small energy gap ΔE=0.271 kcal/mol between the lower energy 30° and the higher energy 0° conformers.

In summary, the crystal structures solved for the CGTACG and CGATCG adducts highlight the local distortion of B-like helices associated to the binding of the natural alkaloids Coptisine and Chelerythrine. The ligands are found at non-canonical arrangements of bases which suggest that the compounds are not very suited, most likely due to their elongated shape, to behave as classical intercalators (whose binding sites are defined by only two base pairs of the intercalated helix, see for example [400]). Coptisine and Chelerythrine, as previously observed for Berberine and Sanguinarine [395], appear to preferably interact with extensive aromatic surfaces, as those of CGGC quartets or of GGGG tetrads from G-quadruplex structures. Anyway, the presence of dioxolo groups seems to be important in the allocation of the ligands in the binding sites, as observed in particular for the Tel12-Coptisine and the CGATCG-Coptisine adducts where such moieties are involved in stabilizing interactions.

4.2 PHENYLALKYL AND DIPHENYLALKYL BERBERINE DERIVATIVES

The promising properties of natural alkaloid Berberine to bind human telomeric G-quadruplex and exert good anticancer activity [266, 267, 276, 278, 283] have led to the investigation of several Berberine semi-synthetic derivatives functionalized at the 13-position. The addition of aromatic moieties through alkyl linkers might increase the G4 binding strength with respect to the natural precursor and result in enhanced telomerase inhibition activity. In the first place, phenylalkyl and diphenylalkyl derivatives (figure 49) were studied by means of spectrophotometry and X-ray crystallography. Also, evaluations on the cytotoxicity of the compounds against cancer cells were performed by Doctor Ivana Scovassi from the CNR in Pavia, thanks to a scientific collaboration.

4.2.1 Solution studies

The interaction of 13-phenylalkyl and 13-diphenylalkyl derivatives of Berberine with the human telomeric sequence sequence Tel24 d[$(TTAGGG)_4$] was evaluated by UV-Vis absorption and CD spectroscopy. UV-Vis absorption titrations were performed in the first place for the derivatives using Berberine as reference under the current experimental conditions. Upon addition of Tel24 to solutions of the ligands, a clear hypochromic and bathochromic shift was observed for the band centred at about 340 nm together with the appearance of an isosbestic point in the range 350-355 nm (figure 50). The spectra demonstrated the interaction of the Berberine analogs with human telomeric G-quadruplex which presumably take to the formation of single adduct species over the concentration range investigated. Since the band shifts observed can be

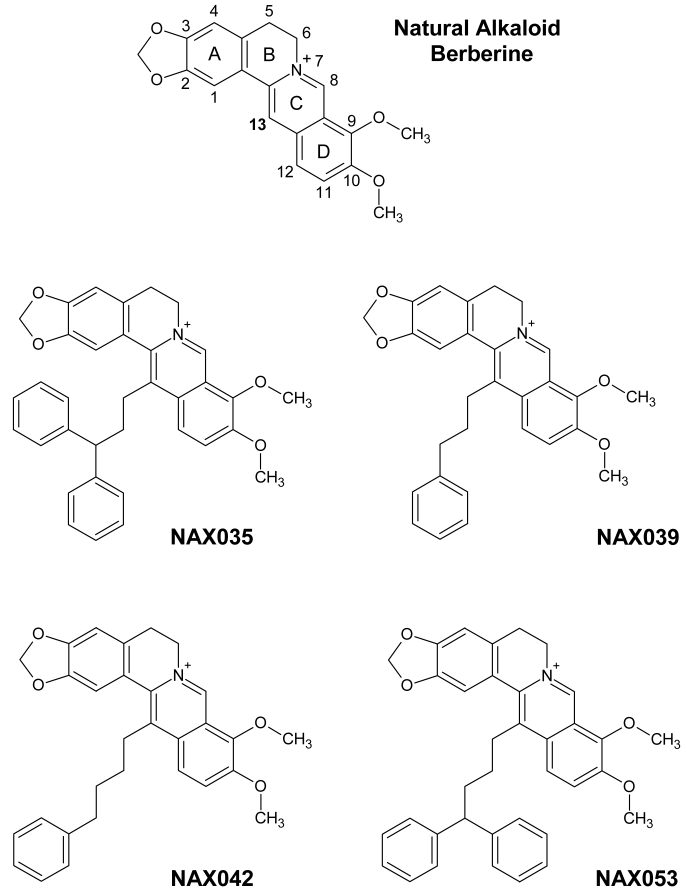


Figure 49: Natural alkaloid Berberine and the phenylalkyl and diphenylalkyl derivatives studied as G4 binding ligands.

explained in terms of the stabilization of π^* orbitals [280, 282], it is supposed that the derivatives bind through π-π stacking interactions at the external G-quartets, as already reported for Berberine [276, 278, 283, 288]. The variations in the intensity at about 340 nm were analysed in order to estimate the dissociation constants for the adducts K_d and approximated G4:ligand stoichiometries (table 17), following methods previously described [280, 282, 339, 340].

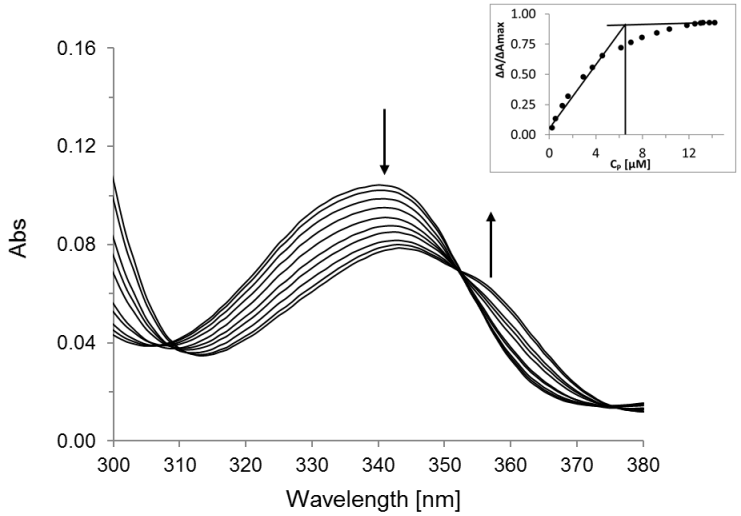


Figure 50: Representative UV-Vis titration performed. Spectra of 14.2 μM NAX039 in the presence of increasing concentrations of Tel24 in 10 mM potassium phosphate buffer containing 150 mM KCl at pH 7 at 25°C. Curves shown in the plot are relative to the addition of DNA between 0 and 0.5 equivalents with respect to the ligand. Inset: binding stoichiometry obtained from the intersection of the straight lines representing the linear fit of the initial and final data points of the binding isotherm.

Table 17: Analysis of UV-Vis absorption spectra for the interaction of Tel24 with Berberine and its 13-monophenyl and 13-diphenylalkyl derivatives. Dissociation constants K_ds (e.s.d. in parentheses) and G4:ligand stoichiometry ratios are reported. Experiments carried out in 10 mM potassium phosphate buffer pH 7, 150 mM KCl (temperature 25°C).

Ligand	Carbons in the alkyl chain	Phenyl groups	K_d (μM)	G4:ligand stoichiometry
Berberine	/	/	9(1)	1:1
NAX035	3	2	3.9(6)	2:1
NAX039	3	1	3.2(5)	1:2
NAX042	4	1	1.3(3)	2:3
NAX053	4	2	1.1(3)	2:3

The dissociation constants are in the μM range, thus suggesting a strong interaction between Tel24 and the ligands. In according to published papers [276, 278], in spite of slight differences associated to different experimental conditions, Berberine strongly bind to human telomeric G-quadruplex as in 1:1 G4:ligand stoichiometry. The addition of phenylalkyl and diphenylalkyl pendant groups take to lower dissociation constants and also to alternative stoichiometries for the adducts. In more detail, the length of the alkyl linker plays an important role in the formation of the adducts as suggested by the comparison of NAX035, NAX039 vs NAX042, NAX053 K_ds. Surprisingly, the number of phenyl groups does not significantly affect the binding strength (NAX035 vs NAX039, NAX042 vs NAX053). The most interesting consequence of the modifications on the Berberine skeleton is the observed ability of the derivatives to bind multiple G-quadruplex in solution, with the only exception of NAX039. Comparable behaviours have been observed for the natural compounds Ellipticine and Chelerythrine characterized by similar isoquinoline scaffolds [282, 340] which can bind to the extremity and to the interface of two or more contiguous G-quadruplexes and have

shown 2:3 G4:ligand stoichiometry, that is the same displayed by NAX042 and NAX053.
CD spectroscopy was then used to get information about the topology and the thermal stability of Tel24 G-quadruplex upon ligand binding (figure 51). The human telomeric sequence Tel24 folds in hybrid structures in potassium solutions with the hybrid-1 as the predominant configurations [184, 187, 193, 401]. CD spectra recorded at room temperature show the characteristic features of mixed parallel/antiparallel tetraplexes that are a positive peak at 290 nm, a shoulder at 270 nm and a negative peak at 235 nm, in agreement with previous studies [280, 282, 401]. The binding of Berberine takes to decreased intensities at 290 nm and 270 nm while the band at 235 nm becomes less negative. A ligand-induced band is also observed around 360 nm, similarly to other G4 binders [280, 282]. As regard the derivatives, spectra of NAX039 resemble those of Berberine, excepting for a slight increase in the 290 nm band intensity. On the other hand, a different profile is displayed by the NAX035 spectra where a decrease of the 290 nm band and a less evident lowering of the shoulder at 268 nm can be appreciated. NAX042 and NAX053 show, conversely, quite comparable spectra in which the appearance of a broad negative band at around 340 nm, the increase in the 290 nm band intensity, the decrease in the shoulder at 268 nm and the contemporary rise of a positive band around 245 nm are recognized. On the basis of these findings, NAX042 and NAX053 are the derivatives which more strongly bind to Tel24 and they induce even a topological conversion from hybrid to antiparallel G-quadruplex, as already evidenced for strong G4 binders as TMPyP4 [342–346, 402].

Thermal melting experiments were carried out in order to supply further indications of the stabilization of Tel24 upon ligand binding (table 18). The melting transitions were monitored at 290 nm through CD spectroscopy, using analogous experimental con-

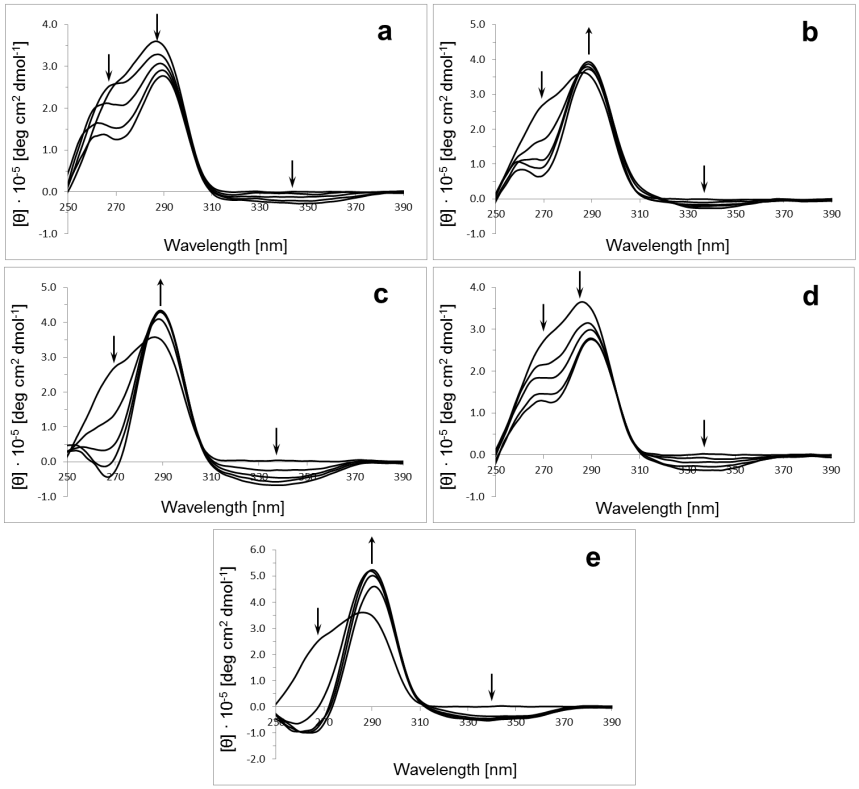


Figure 51: CD spectra at 25°C of 56.3 μM Tel24 in the presence of 1-4 equivalents of Berberine (a), NAX039 (b), NAX042 (c), NAX035 (d) and NAX053 (e). Solutions prepared in 90 mM LiCl, 10 mM KCl and 10 mM lithium cacodylate (pH 7). The concentration of DMSO, due to the additions of ligand stock solutions, was kept constant at 3.75% v/v.

ditions as those employed for the room temperature experiments. Tel24 showed a single thermal transition at about 50°C, in agreement with previous calorimetric studies which have pointed out also the presence of a second transition at about 66°C [280, 282, 340, 403]. Tel24 is thermally stabilized upon Berberine binding and $\Delta T_m \approx 11°C$ was found. Similar results have been obtained in Prof. Sissi laboratory in Padua (figure 34) whereas different values ($3°C \leqslant \Delta T_m \leqslant 8°C$) have been reported in other works, most likely because of the alternative experimental conditions used [278, 294–296]. Anyway, the semi-synthetic derivatives, with the exception of NAX035, induce thermal stabilizations comparable to Berberine, ΔT_m = 11-12°C. The Berberine derivatives do not improve significantly the thermal stability of the telomeric G-quadruplex with respect to the natural precursor thus suggesting an analogous binding mode for the ligands, at least for NAX039, NAX042 and NAX053. The relatively low stabilization of Tel24 by NAX035 in the conditions explored is still matter of investigation.

Table 18: Analysis of CD thermal melting curves of Tel24 (56.3 μM) in the presence of saturating concentration of Berberine and its 13-phenylalkyl and 13-diphenylalkyl derivatives at constant quadruplex:ligand stoichiometric ratio of 1:4. Solutions were prepared in 10 mM lithium cacodylate pH 7, 90 mM LiCl and 10 mM KCl. The concentration of DMSO, due to the additions of ligand stock solutions, was kept constant at 3.75% v/v.

Ligand	T_m [°C]	ΔT_m
Tel24 alone	49.9	/
Berberine	61.2	11.3
NAX035	56.2	6.3
NAX039	61.8	11.9
NAX042	61.2	11.3
NAX053	61.4	11.5

4.2.2 X-ray crystallography

Crystallization screenings were performed for the adducts formed by the 13-monophenyl and 13-diphenylalkyl Berberine derivatives and the human telomeric sequences Tel12 d[$TAG_3TTAGGGT$] and Tel23 d[$TAG_3(TTAGGG)_3$] with the purpose to shed light on the interactions of the ligands and telomeric G-quadruplexes. Suitable crystal for X-ray diffraction have been obtained for the Tel12-NAX053 (figure 49) adduct whose crystal structure was successfully solved (table 19).

Table 19: Crystal data and refinement statistics for the Tel12-NAX053 structure.

Space group	P 4_12_12
Crystal system	tetragonal
Cell dimension (Å)	a=b=41.13, c=71.20
Cell angles (°)	$\alpha=\beta=\gamma=90$
Resolution range (Å)	35.62-1.70
Rfactor	0.231
Rfree	0.284
RMSD bonds (Å)	0.006
RMSD angles (°)	1.813
PDB code	5CDB

The structure shows bimolecular parallel G-quadruplex units and ligand molecules in 1:1 stoichiometric ratio (figure 52). Analogously to the Tel12-Coptisine structure, the G4 units are head-to-tail disposed and form columns growing along the 4-fold crystallographic axis. The tetraplexes are originated by two symmetry independent strands and feature three planar G-tetrads stacked at 3.4 Å apart from one another as well as TTA propeller loops which can be described as type 5 and type 1 (subtype 2), according to the analysis reported by Neidle and coworkers [379]. Potassium ions are found in the internal channel at 2.6-2.9 Å from the O6 atoms of guanines in an overall anti-prismatic coordination, in according with previously reported structures [380].

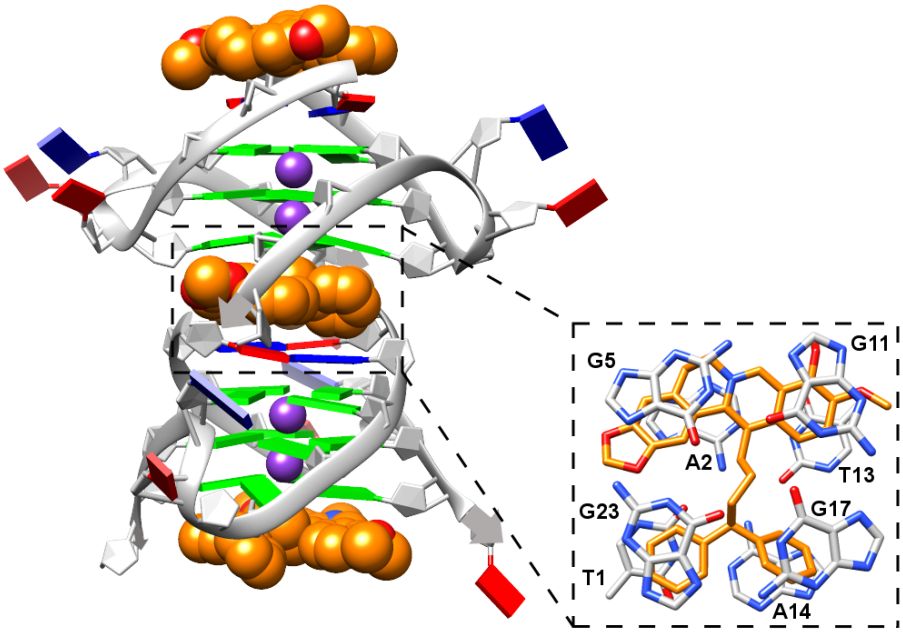


Figure 52: Crystal structure of the Tel12-NAX053 adduct. The ligand is found between the 3′ end G-tetrad and the TATA tetrad originated at the 5′ end. Guanine, thymine and adenine coloured in green, red and blue, respectively.

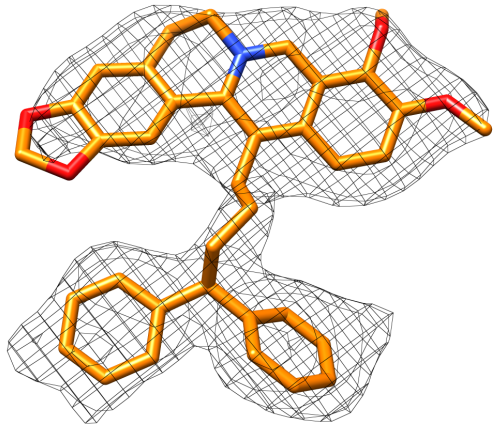


Figure 53: NAX053 shown within the electron density map [F_o-F_c] at 2.0σ level.

As shown in figure 52, NAX053 (omit map, figure 53) is located between the 3′ end G-tetrad and the TATA tetrad that is formed by flanking residues at the 5′ end. In more detail, the ligand is sandwiched between the quartets at about 3.9 Å on average and it is involved in π-π stacking interactions with all the bases belonging to the quartets (figure 54). The conformation assumed by Berberine (quinoline and benzodioxolo groups form a dihedral angle of about 24°) together with the significant deviation from planarity displayed by the phenyl pendant groups (dihedral angle of about 46° between the rings) are responsible for the non-optimal distances of stacking. The conformation assumed by the derivative also affects the geometry of the tetrads, especially the 3′ end G-tetrad towards which the alkyl chain is directed. In fact, such tetrads display non-negligible deviations from planarity most likely due to steric hindrance produced by the derivative (dihedral angles for TATA and GGGG tetrads in the range 7-14° and 14-19°, respectively).

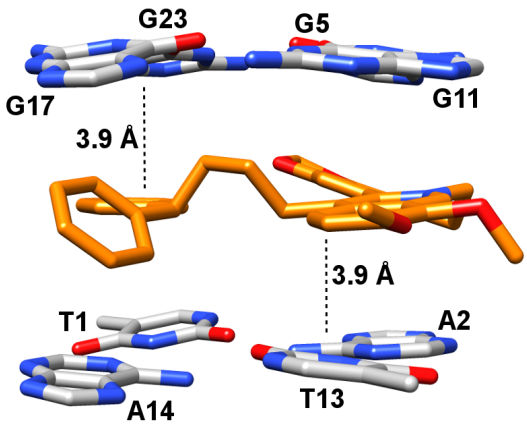


Figure 54: Binding site of NAX053 defined by the 3′ end G-tetrad and the 5′ end TATA quartet.

The comparison of the NAX053-Tel12 structure with the crystal structure previously reported for Berberine bounded to the human telomeric sequence Tel23 [288] supplies further interesting information (figure 55). Berberine was found to stack at two different binding sites, one defined by two 5′ end G-tetrads and the other defined by the 3′ end G-tetrad and a TA base pair from a symmetry related G-quadruplex. Both binding sites host two coplanar Berberine molecules, each interacting with a couple of guanines. It is noteworthy how the Berberine core is located in a strictly analogous manner on the 3′ end G-tetrad in both the crystal structures. The alkaloid skeleton is able indeed to interact via π-π stacking with two guanines and its positively charged nitrogen atom is oriented toward the TTA loops. The stacking distances observed for Berberine are however different in the structures (3.4-3.5 Å in the Berberine-Tel23 adduct vs 3.9 Å in the NAX053-Tel12 adduct) thus suggesting that the addition of the pendant group is determinant for increasing the aromatic surface available for the stacking, even though it induces a certain degree of steric hin-

drance and the binding is optimized only at the expense of the target's external tetrads planarity.

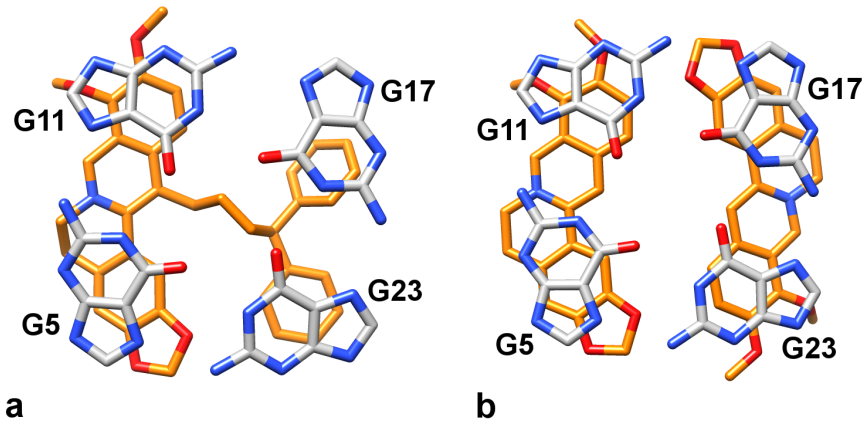


Figure 55: Binding site of NAX053 (a) and Berberine (b) at the 3′ end G-tetrad in their respective crystal structures with Tel12 and Tel23.

4.2.3 Cytotoxic activity on cancer cells

MTT assays were performed by Doctor Ivana Scovassi from the CNR in Pavia to assess the cytotoxicity of the promising NAX042 and NAX053 Berberine derivatives (figure 49) on selected cancer cell lines characterized by telomerase reactivation [303, 304] (figure 56). The derivatives have shown the most interesting improvements in terms of G4 binding ability with respect to the natural precursor and not coincidentally NAX053 has been recently reported to have a strong impact on the viability of human colon cancer cells with significant better properties than Berberine [303, 304]. The investigation on the biological activity of NAX053 was extended to HeLa and MCF-7 cancer cells lines (respectively de-

rived from uterine cervix and breast carcinomas) which have been previously used by other authors to test the effects of Berberine [404, 405]. The results shown in figure 56 highlight the greater efficiency of NAX053 in affecting the viability of both cancer cells lines in a dose-dependent manner in comparison with Berberine (not shown). The IC_{50} values for NAX053 are 2.56±0.10 μM and 2.27±0.02 μM for HeLa and MCF-7 cells, respectively, compared to 18.82±1.27 μM and 11.75±1.14 μM of Berberine. On the other hand, the cytotoxicity of NAX042 was tested on a panel of human cancer cell lines, including MCF-7, melanoma A375 and two colon carcinoma cells (SW613-B3 and HCT116). NAX042 exerts a dose-dependent effect on the viability of all the tested cancer cell lines, which are characterized by telomerase reactivation, having an epithelial origin like HeLa cells. Overall, NAX042 and NAX053 display anticancer activities which are effectively more potent than Berberine.

In summary, the addition of phenylalkyl and diphenylalkyl groups at the 13-position of the Berberine skeleton takes to improvements in telomeric G-quadruplex binding ability of the natural compound. NAX042 and NAX053 have shown the greatest improvements, as highlighted by UV-Vis absorption and CD experiments. The effect of ligand binding on G4 thermal melting however pointed out that the derivatives do not stabilize Tel24 to a greater extent with respect to Berberine, although the correlation between increase in melting temperature and ligand binding strength is not completely correct [406]. Anyway, X-ray crystallography has confirmed the analogous interaction of the Berberine core of NAX053 on the 3′ end G-tetrad of telomeric G-quadruplexes with respect to the natural precursor and also suggested that the increase in binding strength observed by spectrophotometry is associated to the greater aromatic surface of the derivative, despite the evident steric hindrance related to the alkyl moiety.

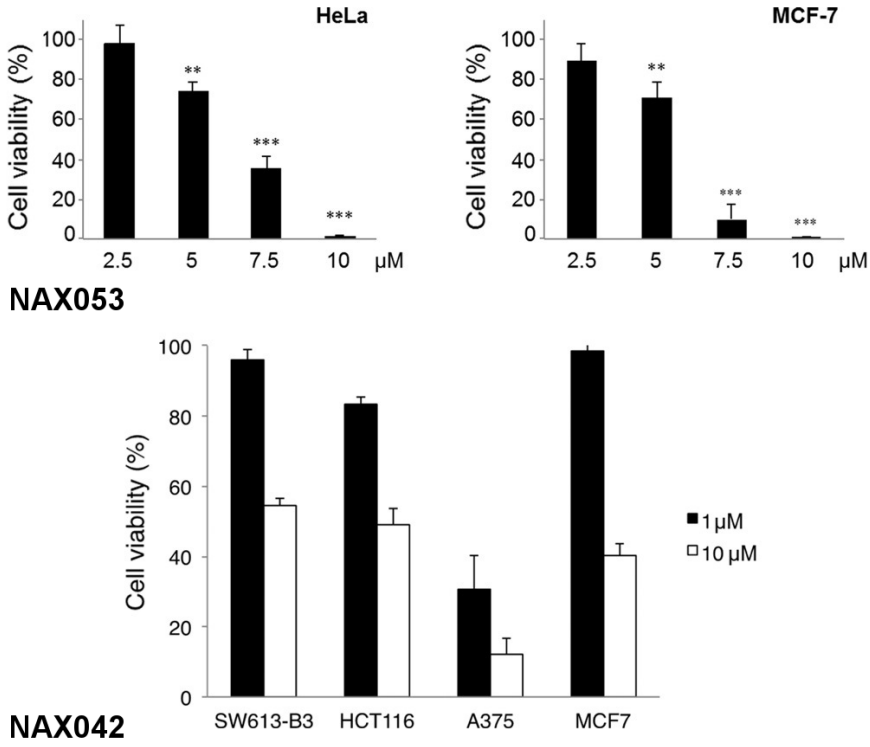


Figure 56: Effect of NAX053 and NAX042 on cancer cell viability. Four increasing concentrations of NAX053 (upper panel) and 1μm/10μm of NAX042 (lower panel) were used for 24 h treatments in quadruplicate (three independent MTT experiments were carried out). Data obtained from untreated control cells were considered as 100% to normalize the absorbance of treated samples and are expressed as mean±SD. **P<0.01; ***P<0.001. Adapted from reference [351].

It is likely that the alkyl groups adopt different conformations in solutions where the pendant chains could be able to bind further G-quadruplexes, as indicated by the G4:ligand stoichiometries calculated for the titrations (table 17). The ability to bind multiple G4 structures is probably the most interesting consequence of the modifications on the Berberine skeleton as telomeres are actually composed of dozens of TTAGGG repeats and so can form multiple tetraplexes [166, 203, 204]. The stimulating cytotoxic activity shown by NAX042 and NAX053 confirms the idea that these modifications are promising for the identification of potential anticancer drugs able to target human telomeres. Nevertheless, the compounds should be further modified in order to improve the overall lipophilicity, which could prevent significant bioavailability, and also in order to increase their selectivity over alternative nucleic acids foldings that appears to be not too great [374–376, 407].

4.3 PYRIDIL BERBERINE DERIVATIVES

In a second phase, different functionalizations at the 13-position of the Berberine skeleton were considered. Pyridine groups were added through alkyl linkers of different length in order to investigate how groups carrying hydrogen bond acceptors and possibly positive charges could interact with human telomeric G-quadruplex foldings (figure 57). Up to now, only crystallographic experiments have shed light on the interaction of the derivatives with telomeric G-quadruplex, although solutions studies are currently being performed in the research group of Prof. Gopinatha Suresh Kumar from the Indian Institute of Chemical Biology, thanks to the established collaboration.

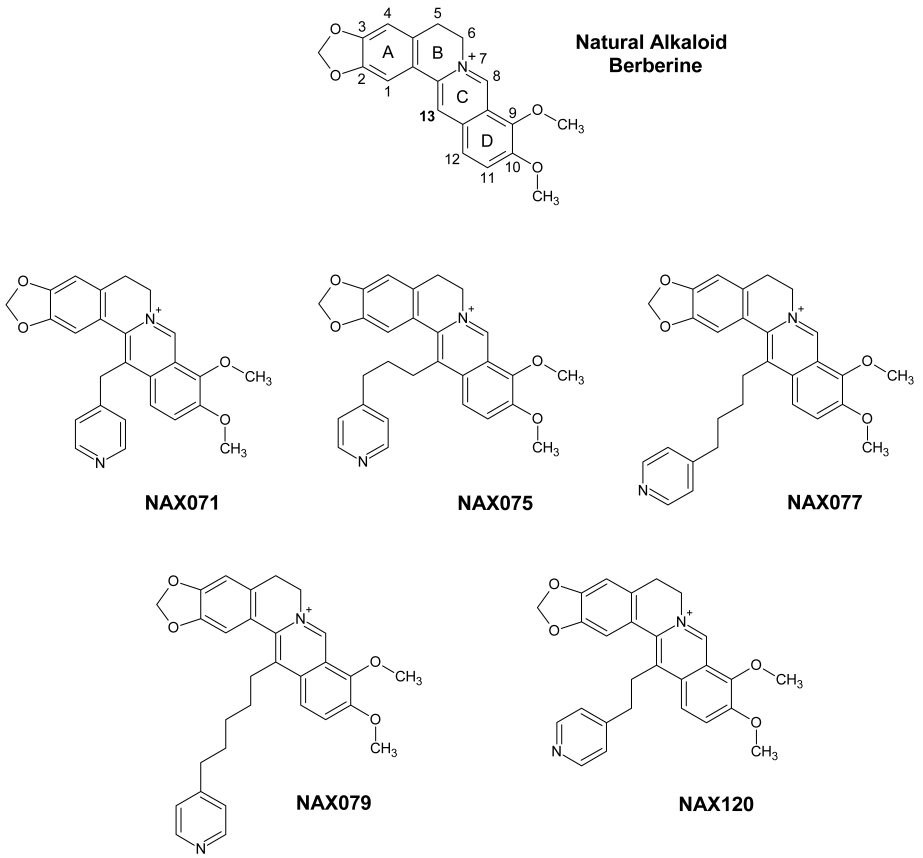


Figure 57: Natural alkaloid Berberine and the pyridil derivatives studied as G4 binding ligands.

4.3.1 X-ray crystallography

Crystallization screenings were performed for the adducts formed by the derivatives and the human telomeric sequences Tel12 d[TA-G_3TTAGGGT] and Tel23 d[$TAG_3(TTAGGG)_3$]. Recently, the crystal structures of Tel12 bounded to NAX075, NAX077 and NAX120 have been successfully solved allowing a more extensive analysis of the interaction of Berberine 13-derivatives (table 20).

Table 20: Crystal data and refinement statistics for the structures of Tel12 bounded to NAX075, NAX077 and NAX120.

	Tel12-NAX075	Tel12-NAX077	Tel12-NAX120
Space group	P 4_12_12	P 4_22_12	P 4_22_12
Crystal system	tetragonal	tetragonal	tetragonal
Cell dimension (Å)	a=b=41.11, c=68.39	a=b=42.10, c=34.15	a=b=41.68, c=34.31
Cell angles (°)	$\alpha=\beta=\gamma=90$	$\alpha=\beta=\gamma=90$	$\alpha=\beta=\gamma=90$
Resolution range (Å)	35.23-1.60	29.77-2.04	29.47-2.04
Rfactor	0.256	0.269	0.249
Rfree	0.290	0.312	0.295
RMSD bonds (Å)	0.015	0.015	0.009
RMSD angles (°)	2.582	2.805	1.898

The crystal structures of the adducts are strictly similar so they are discussed collectively. Bimolecular parallel G-quadruplexes and ligand molecules are found in 1:1 stoichiometric ratio and give rise to columns growing along the four-fold rotation axis (figure 58). As a consequence of the different screw symmetries related to such axes, that are 4_1 for NAX075 while 4_2 for NAX077 and NAX120, the quadruplex units are either originated by symmetry independent (NAX075) or symmetry related strands (NAX077 and NAX120). The tetraplexes feature three G-tetrads stacked at 3.4 Å apart from one another. On the basis of the analysis performed by Neidle and coworkers [379], the TTA propeller loops can be described as type 1 (subtype 2), albeit NAX075 adduct shows also a type-5 loop. Potassium ions are found in the internal channels

at 2.6-3.0 Å apart from the guanine O6 atoms leading to an overall antiprismatic coordination sphere.
The binding sites of the ligands are found between the 3′ end G-tetrad and the TATA tetrad formed by flanking residues at the 5′ end (figure 58, omit maps figure 59). The derivatives are involved in π-π stacking interactions with three bases per quartet by means of the Berberine core and the pyridine pendant group (inter-planar distance around 3.6 Å). No hydrogen bond interactions are observed between the pyridine groups and DNA residues. It is noteworthy how NAX077 and NAX120 are disordered and for both two symmetry related molecules (4_2 screw symmetry) have been found in the binding site sharing the same position. The Berberine core is bent in all the analogs, however the dihedral angle between quinoline and benzodioxolo groups is notably greater in the case of NAX075 (about 25° in NAX075 vs 9° in NAX077 and NAX120). In all the structures, the pyridine groups are well disposed for the stacking interactions and they are quite in-plane with the Berberine core. Similarly, the alkyl linkers stand on the plane of the molecule, in particular in the case of NAX120 which features the shortest linker, and are oriented toward the end G-tetrad, as clearly observed for NAX075.

The comparison of the crystal structures obtained suggests that the length of the alkyl linker is determinant for improving the Berberine binding properties and short and even-numbered linkers appear to be valuable for the purpose. Among the derivatives crystallized with the human telomeric G-quadruplex, NAX120 seems to be the most suitable for the stacking on the 3′ end G-tetrad. The ligand displays indeed the shortest alkyl linker, composed of two carbon atoms, which appear to favour the planarity of the Berberine core as well as the planarization of the entire molecole and so the overall strength of π-π stacking interactions. It is no coincidence that the smaller deviations from planarity for

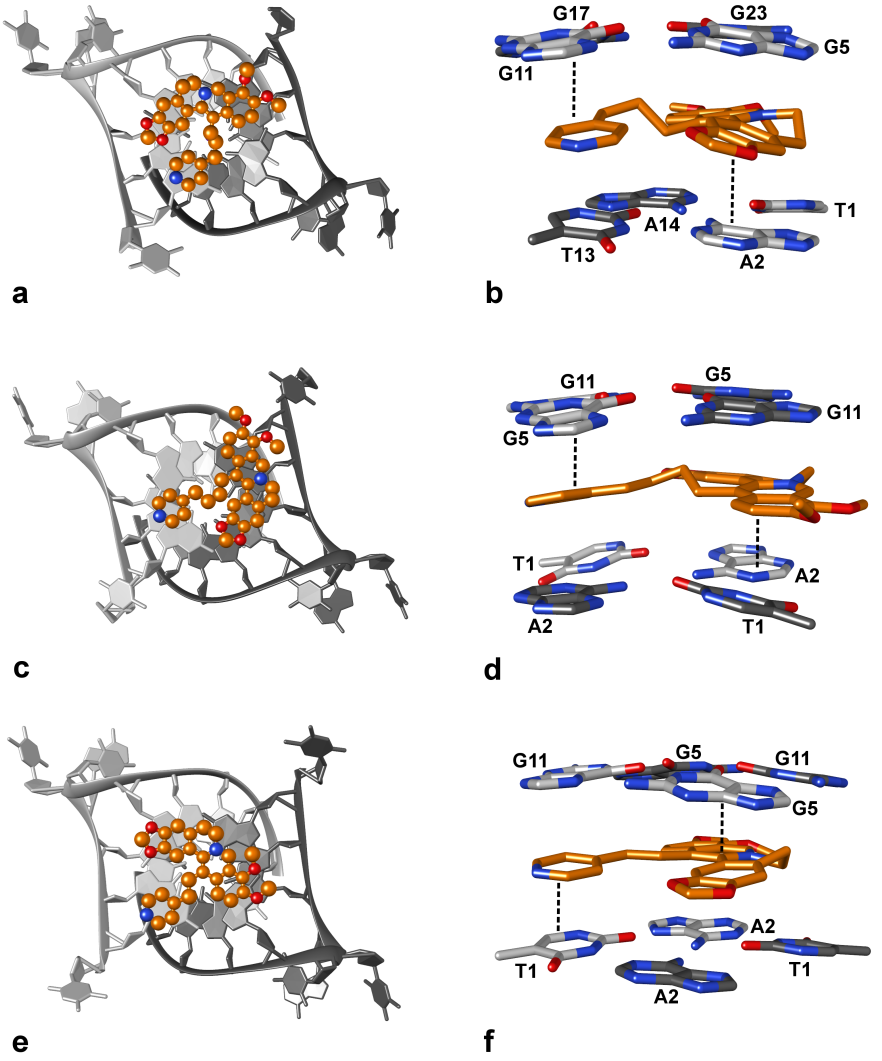


Figure 58: Crystal structures of the Tel12 human telomeric G-quadruplex bounded to Berberine derivatives NAX075 (a), NAX077 (c) and NAX120 (e) with a focus on the binding site of the ligands (b,d,f, respectively).

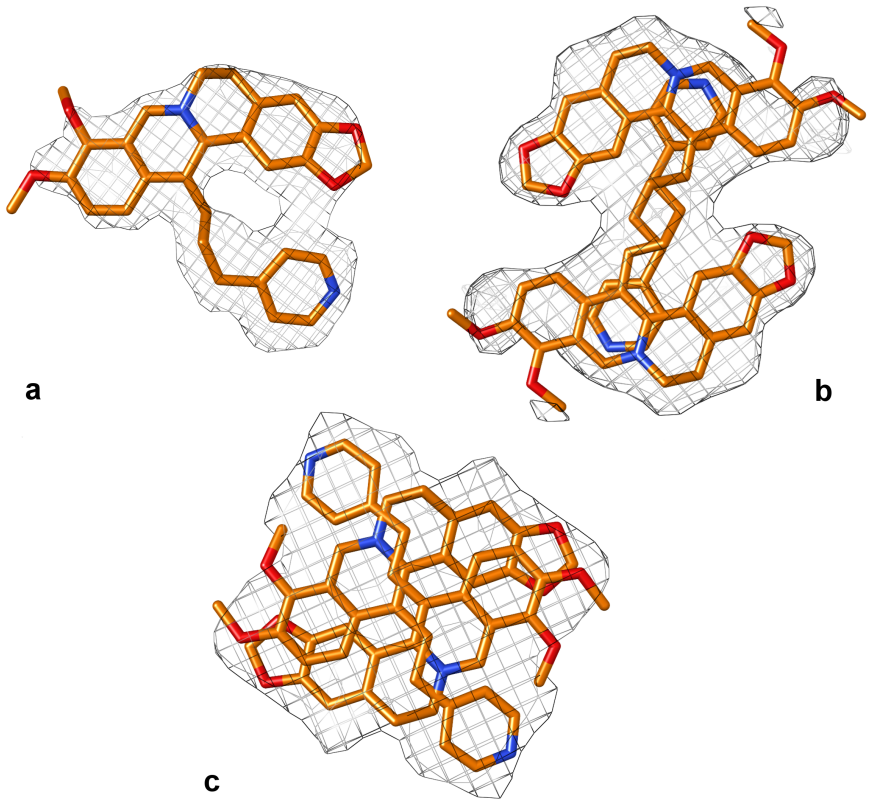


Figure 59: NAX075 (a), NAX077 (b) and NAX120 (c) shown within the electron density map [F_o-F_c] at 3.0σ, 2.0σ, 2.5σ level, respectively.

the tetrads have been observed in the NAX120 crystal. On the other hand, the derivative NAX075, featuring three carbon atoms in the alkyl chain, and the derivative NAX077, characterized by a linker of four carbon atoms (figure 57), exhibit less planar conformations which are associated to higher deviations from planarity for the 3′ end G-tetrad due to steric hindrances, especially in the case of NAX075. These observations suggest that the interaction trend of the derivatives should follow the order NAX120 > NAX077 > NAX075, in agreement with preliminary spectroscopic and calorimetric titrations performed in the research group of Prof. Gopinatha Suresh Kumar from the Indian Institute of Chemical Biology (table 21), thanks to the established scientific collaboration. Further studies will be however carried out in the next months in order to fully confirm the hypotheses formulated. Anyway, the comparison of the dissociation constants reported with those previously evaluated for the phenylalkyl and diphenylalkyl Berberine derivatives studied (table 17) shows how the binding strength of these ligands is largely determined by the length of the alkyl linker on the 13-position (NAX035, NAX039 vs NAX075 and NAX042, NAX053 vs NAX077), despite some discrepancies likely due to the different human telomeric sequences taken into account for the experiments.

Table 21: Spectroscopic and calorimetric titrations carried out for the interaction of the human telomeric sequence Tel12 with the Berberine derivatives NAX075 and NAX077. K_d dissociation constant and *n* ligand:G4 stoichiometry.

Ligand	UV-Vis absorption spectroscopy		Fluorescence spectroscopy		Isothermal titration calorimetry	
	K_d (μM)	*n*	K_d (μM)	*n*	K_d (μM)	*n*
NAX075	3.13	1:1	3.06	1:1	3.16	1:1
NAX077	1.92	1:1	1.98	1:1	1.99	1:1

The crystal structures obtained for the Berberine analogs featuring pyridinealkyl groups have been compared with those obtained for the derivative NAX053 and the natural precursor Berberine (figure 60) [288]. In particular, the binding of the compounds at the 3′ end G-tetrad has been examined. In all the structures, the Berberine core binding is largely driven by π-π stacking interactions involving two guanines. The addition of aromatic groups result in a greater extent of ligand-DNA interactions although the presence of non-planar moieties, as a diphenyl group or alkyl chains of a certain length, induces significant deviations from planarity in the interacting quartets along with increased stacking distances. This scenario has been clearly observed in the Tel12-NAX053 structure. On the other hand, pyridinealkyl modifications lead to ligands which better allocate in the binding sites because of the more planar character of the molecules. The Berberine core interacts in a similar fashion in the structures with the positively charged nitrogen atom oriented outward from the channel of the quadruplex, where a negative electron density is present. Actually, the Berberine cores of NAX053 and NAX077 are nearly superimposable and also the NAX075 structure displays good similarities, although the skeleton of the natural alkaloid is found to stack on different guanines of the tetrad. Conversely, the Berberine core of NAX120 is disposed along the diagonal of the quartet, most likely in order to favour the stacking of the pyridine moiety on a guanine. Overall, these findings might be of great help in the design of Berberine derivatives with enhanced telomeric G4 binding properties in comparison to the natural precursor.

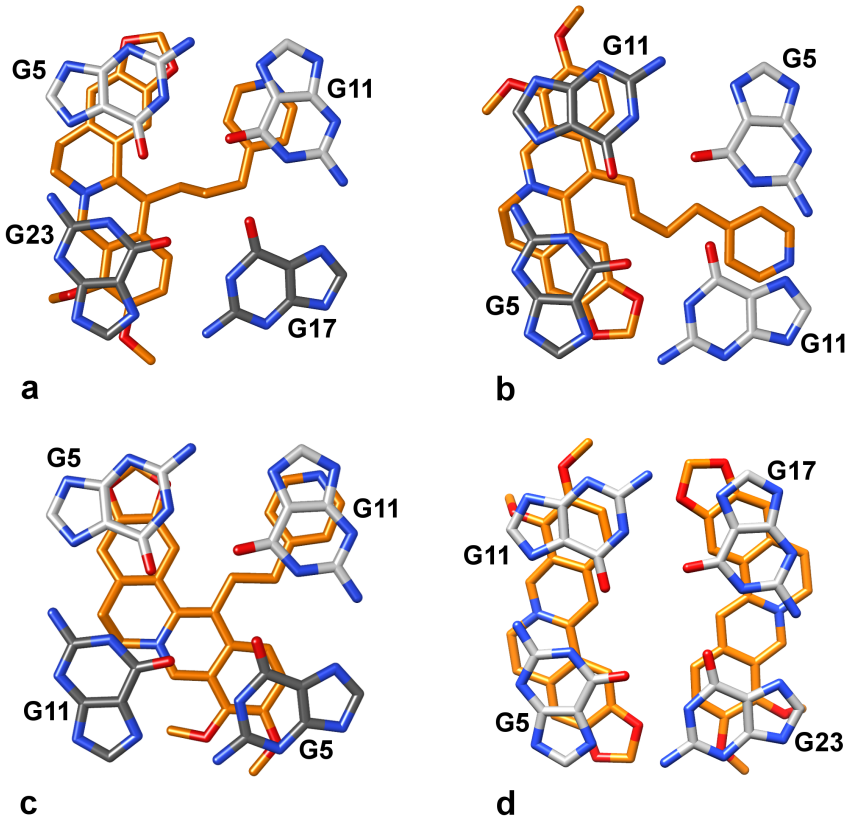


Figure 60: Comparison of the binding of Berberine derivatives NAX075 (a), NAX077 (b) and NAX120 (c) at the 3′ end G-tetrad of Tel12 tetraplex. The localization of natural precursor Berberine at the 3′ end G-tetrad of Tel23 is reported as well (d).

4.4 Au^{I} N-HETEROCYCLIC CARBENE COMPLEXES, AUNHC **I**

Gold complexes are attractive compounds in anticancer drug discovery. From the times when gold was just an alternative to platin for the covalent binding of DNA, great advancements have been made for the identification of complexes able to interact with different biological targets and exert interesting anticancer activity [308–310, 313]. The promising ability of some compounds to interact and stabilize the human telomeric G-quadruplex folding caught the interest of the research activity and led to the study of selected Au^{I} and Au^{III} complexes. The present section concerns Au^{I} complexes, while the following deals with an Au^{III} compound.

N-heterocyclic carbene (NHC) groups are promising ligands for the stabilization of Au^{I} in a linear bicoordinated geometry that is ideal for the synthesis of planar, aromatic and positively charged G4 binders [314]. The AuNHC **I** complex (figure 61) has indeed shown promising telomeric G4 binding properties with a very good selectivity with respect to the interaction with double-stranded DNA [314–316]. These valuable features raised the interest for a more detailed investigation of the interaction of the compound with human telomeric G-quadruplex. At first stage, crystallographic and ESI-MS experiments were carried out. Lately, spectrophotometric characterization and *in silico* calculations supplied further information on the behaviour of ligand-G4 adducts in solution. The inhibition activity of the compound towards telomerase was also evaluated by Prof. Donato Colangelo from the Università del Piemonte Orientale, thanks to the established collaboration.

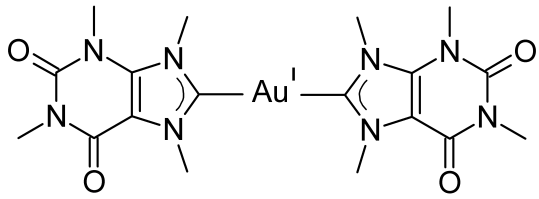


Figure 61: The AuNHC **I** complex studied as G4 binding ligand.

4.4.1 X-ray crystallography

Crystallization screenings were performed for the adducts formed by AuNHC **I** and the human telomeric sequences Tel12 d[TAG_3-TTAGGGT] and Tel23 d[$TAG_3(TTAGGG)_3$]. Suitable crystals for X-ray diffraction analysis were obtained for the Tel23 adduct whose crystal structure was then successfully solved (table 22).

Table 22: Crystal data and refinement statistics for the Tel23-AuNHC **I** structure.

Space group	I $2_1 2_1 2_1$
Crystal system	orthorhombic
Cell dimension (Å)	a=47.13, b=51.29, c=58.77
Cell angles (°)	$\alpha = \beta = \gamma = 90$
Resolution range (Å)	30.00-1.89
Rfactor	0.225
Rfree	0.260
RMSD bonds (Å)	0.006
RMSD angles (°)	1.133
PDB code	5CCW

The asymmetric unit contains a monomolecular parallel G-quadruplex and ligand molecules. The overall G4:ligand stoichiometric ratio is 1:3 (figure 62). G-quadruplexes feature three stacked G-tetrads (inter-planar distances 3.4 Å) and potassium ions in the internal channel (2.7-3.1 Å apart from the guanine O6 atoms). TTA propeller loops can be described as type 9 on the basis of the clas-

sification made by Neidle and coworkers [379], as the thymine residues are directed outward from the guanine core almost lying on a same plane while the adenines are protruded towards the guanines. The G4 units, which are symmetry related by two-fold rotation axes, form columns growing along the [010] direction. AuNHC **I** molecules are found at the interface between adjacent tetraplexes in two different binding sites defined by symmetry related 3′ end and 5′ end G-tetrads, respectively. Interestingly, the binding sites host a different number of ligand molecules. Two coplanar complexes lie in the 5′-5′ binding site, defined by almost perfectly planar G-tetrads, while tilted 3′ end G-tetrads make the 3′-3′ binding site to accommodate just a ligand molecule (dihedral angle of about 38° between symmetry related G17-G23 base pairs, figure 62).

The complex stacks on the external 3′-3′ and 5′-5′ tetrads at about 3.4 Å and it is involved in π-π interactions with bases from these quartets. AuNHC **I** adopts a linear geometry in according to the crystal structure of its BF_4^- salt [315]. The presence of two-fold symmetry also applies to the ligand which displays in fact double positions for the caffeine moieties at the 3′ end G-tetrad, the gold atom being located on the two-fold rotation axis, whereas each complex is found in a pair of disordered positions at the 5′ end G-tetrad (figure 63, omit maps figure 64). The disorder shown by the gold complexes is actually a quite common feature of small molecules crystallized in ligand-quadruplex adducts [288, 408–410]. This observation supports the idea that the binding process is largely driven by non-directional forces, such as electrostatic, Van der Waals and π-π stacking interactions.

Nevertheless, a closer look at the binding sites draws the attention on the reduced overlap between the caffeine moieties of the metal complex and guanines belonging to the quartets. In fact, only one caffeine out of two is involved in significant stacking

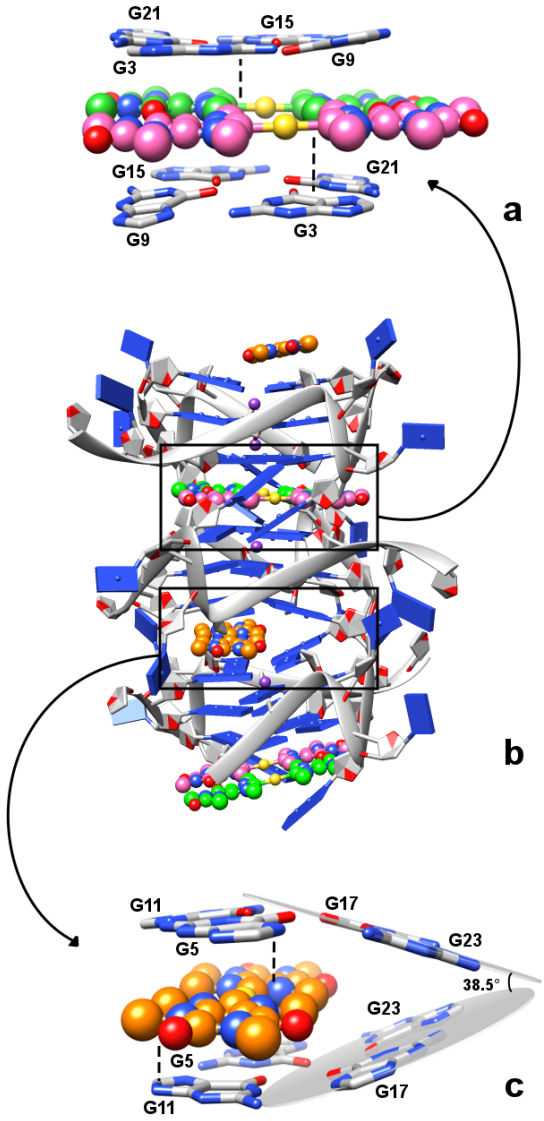


Figure 62: Crystal structure of the Tel23-AuNHC **I** adduct. Binding sites of the ligand at the 5′ end (a) and at the 3′ end G-tetrads (c). Columnar disposition of the adducts along the [010] direction is shown in (b).

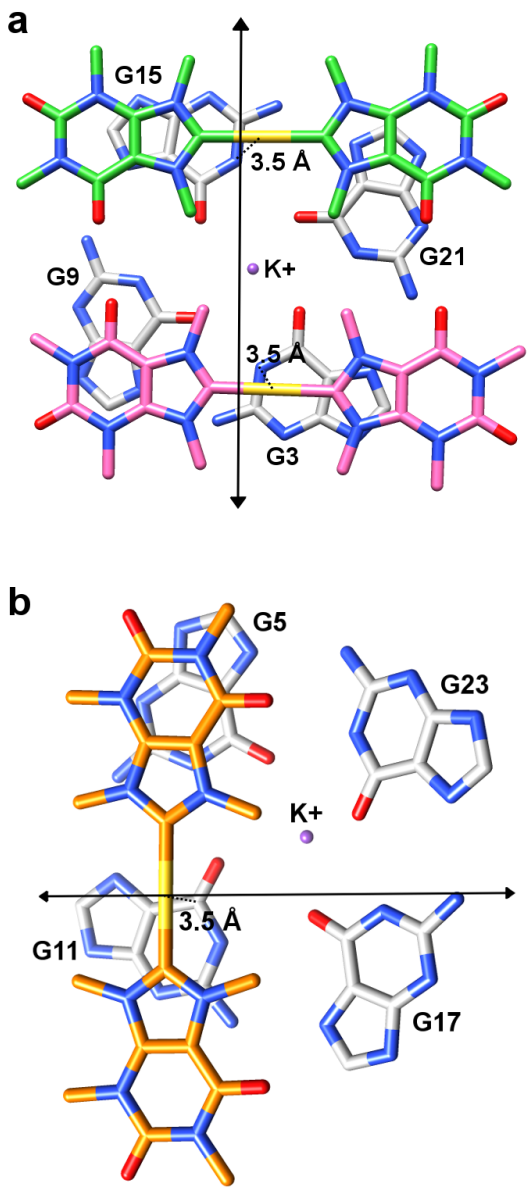


Figure 63: Localization of AuNHC **I** molecules at the 5′-5′ (a) and 3′-3′ (b) binding sites. The black straight lines represent two-fold symmetry axes.

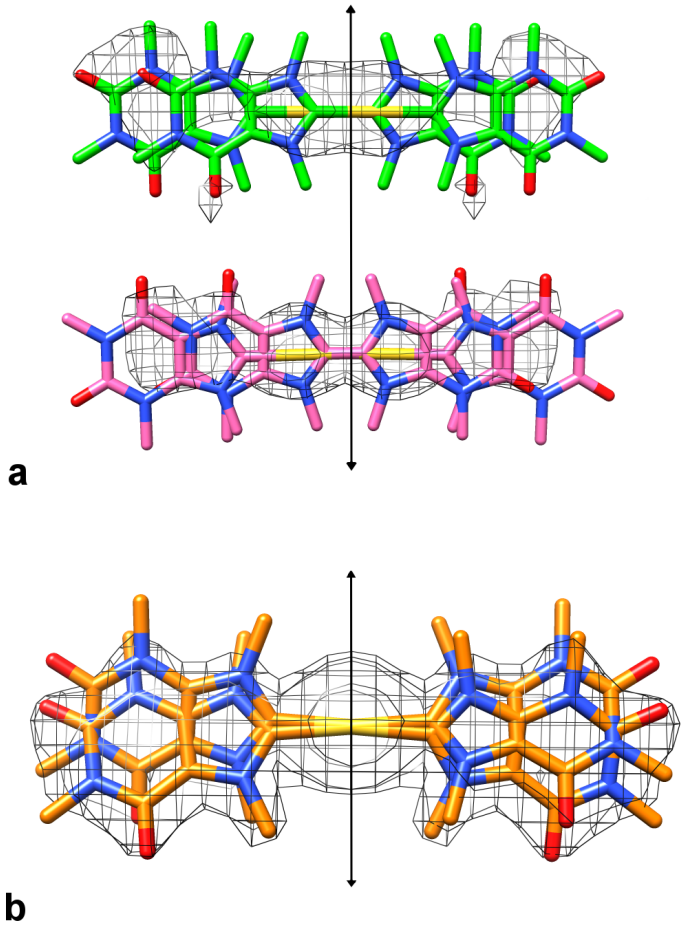


Figure 64: AuNHC **I** shown within the electron density map [$2F_o$-F_c] at 1.5σ level.

interactions as the other one protrudes outward from the G-core. The AuNHC **I** length is too greater with respect to the G-tetrad side (16 Å vs 11 Å) although a good dimensional matching could be achieved with the tetrad diagonal. Presumably, the positioning of the metal complex on the G-tetrad diagonal is prevented by the bent 3′ end quartet (most likely caused by crystal packing forces), while it can be supposed that the overall extent of stabilization achieved by the binding of two ligand molecules at the 5′ end quartet is greater than that possibly achieved by the binding of a single molecule diagonally placed with respect to the tetrad. Hypothetically, the diagonal binding mode could be present in solution for lower G4:ligand stoichiometric ratios. The elongated shape of AuNHC **I** is the most probable cause of the low degree of interaction displayed towards double-stranded DNA [314]. Despite having a diameter of similar length with respect to the G-tetrad side, duplex DNA features indeed a reduced aromatic surface accompanied by a more hindering DNA backbone, which presumably prevent the ligand from intercalation.

The crystal structure obtained for AuNHC **I** is the first one ever reported for a gold compound interacting with a G-quadruplex folding and one of the few with evidences of G4:ligand molar ratios higher than 1:1. Among the reported structures, the adducts of copper(II) and nickel(II) salphen complexes bounded to human telomeric DNA have drawn the attention for a comparison regarding where heavy metal atoms are found with respect to the G-tetrads. Copper and nickel atoms were localized above the channel of tetraplex, in according with studies which have highlighted the presence of a negative electron density in there [57–59]. Conversely, gold atoms are found in proximity of pyrimidine rings at about 3.5 Å from N1 atoms suggesting the presence of favourable interactions between the soft Au^{I} atoms and the π orbitals of guanines (figure 63).

4.4.2 ESI-MS spectra

The crystallographic results prompted us to further investigate the Tel23-AuNHC **I** adducts in solution. Electrospray ionization mass spectrometry (ESI-MS) experiments were carried out by Doctor Lara Massai and Professor Luigi Messori from the University of Florence (figure 65), thanks to a scientific collaboration. Upon annealing in presence of NH_4^+ ions, Tel23 was found to form adducts at saturating concentration of the ligand with different stoichiometries, namely 1:1, 1:2 and 1:3 Tel23:ligand molar ratios (the metal complex was not broken by the ionization). Thus, AuNHC **I** is able to bind human telomeric G-quadruplex up to the stoichiometry observed at the solid state.

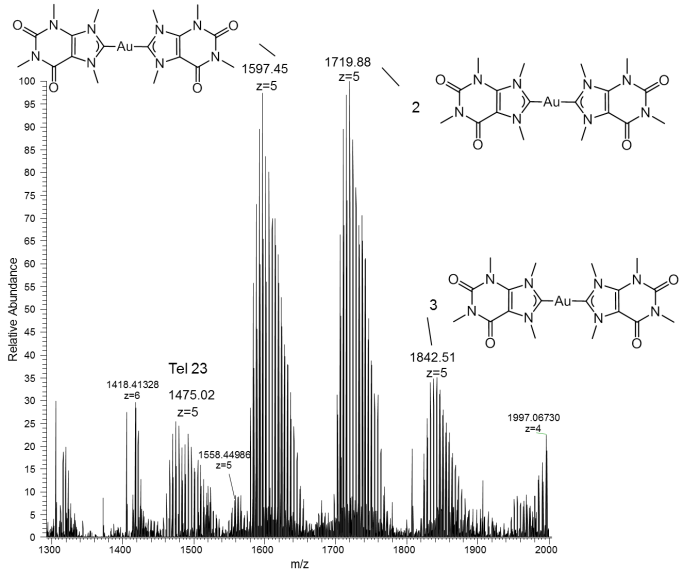


Figure 65: Section of the ESI-MS spectrum of Tel23-AuNHC **I** (25 μM in water in the presence of 60% EtOH, 1:3 G4:metal complex molar ratio). Adapted from reference [353].

4.4.3 Solution studies, CD spectroscopy

The formation of adducts with different stoichiometries emerged from ESI-MS spectra was quite interesting, hence solution studies were then performed to better characterize the Tel23-AuNHC **I** interaction in solution (figure 66). UV-Vis absorption spectroscopy was, unfortunately, not suited for the purpose because of the extensive overlap of DNA and metal complex bands. CD spectroscopy was so used for the investigations.
The CD spectrum of Tel23 shows a positive band at 286 nm with a shoulder at 268 nm and a negative band around 236 nm, hence indicating the formation of hybrid G-quadruplex structures [184, 187, 411–413]. Upon addition of a saturating amount of AuNHC **I**, the CD profile displays progressive changes which lead to a final spectrum characterized by two positive bands centred at 245 and 290 nm and a negative band at 260 nm. On the basis of current knowledge [342–346], the final profile was assigned to an antiparallel topology. The binding of the ligand hence induces a transition from hybrid to antiparallel topology, as already observed for other strong-interacting G4 ligands as TMPyP4 [402]. Interestingly, the binding process was associated to a slow kinetics and the completion was attained only after 72 hours of incubation at 25°C, although significant modifications in the overall spectral appearance were already evident after one hour. Two isoelliptic points were observed at about 280 nm and 250 nm suggesting the presence of multiple adduct species in equilibrium.

A second set of experiments was subsequently carried out to shed light on the interaction between Tel23 and AuNHC **I** at various ligand concentrations. In more detail, a CD titration was performed using a number of independent samples, prepared with a fixed concentration of Tel23 (60 μM oligonucleotide) and increasing amounts of AuNHC **I** (figure 67). CD spectra were recorded after 72 hours of incubation at 25°C, after a careful check of the

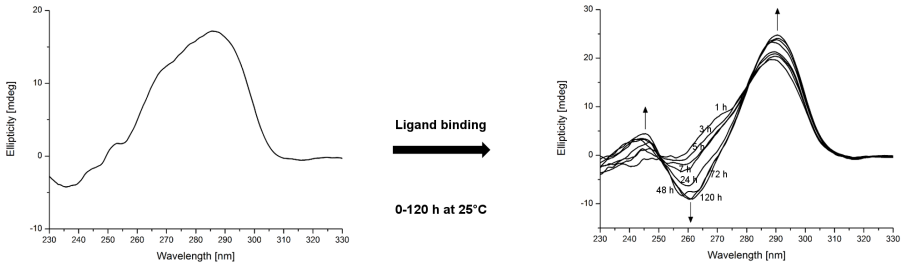


Figure 66: CD spectra of 60 μM Tel23 upon addition of AuNHC **I** (1:4 G4:ligand stoichiometric ratio) after incubation at 25°C for 1, 3, 5, 7, 24, 48, 72 and 120 hours. Experiments carried out in 10 mM lithium cacodylate buffer pH 7.2, 90 mM LiCl, 10 mM KCl (temperature 25°C).

actual achievement of reaction equilibrium for all the samples. Upon inspection of spectra, it emerges how the ligand produces distinct spectral changes in dependence of the applied ligand:G4 stoichiometry. In particular, at 1:1 molar ratio the CD spectrum suggests the conservation of hybrid G4 topology, as two positive bands of similar intensity, centred at 268 and 286 nm, are easily recognized. It is noteworthy how the band at 268 nm is the most affected upon ligand binding, showing a far greater intensity with respect to the native form. Increasing the ligand concentration up to 1:5 molar ratio results in a progressive dramatic change in the spectral shape as the bands at 245 and 290 nm become more intense while the positive band at 268 nm disappears in favour of a negative band with maximum at 260 nm. As shown in figure 67, plots of the data at 286 nm and 236 nm display monotonic increasing trends in this regard. Conversely, data at 268 nm show an increased intensity up to 1:1 Tel23:AuNHC **I** stoichiometric ratio, followed by a decrease that leads to the appearance of the negative CD band characteristic of the antiparallel G4 topology.

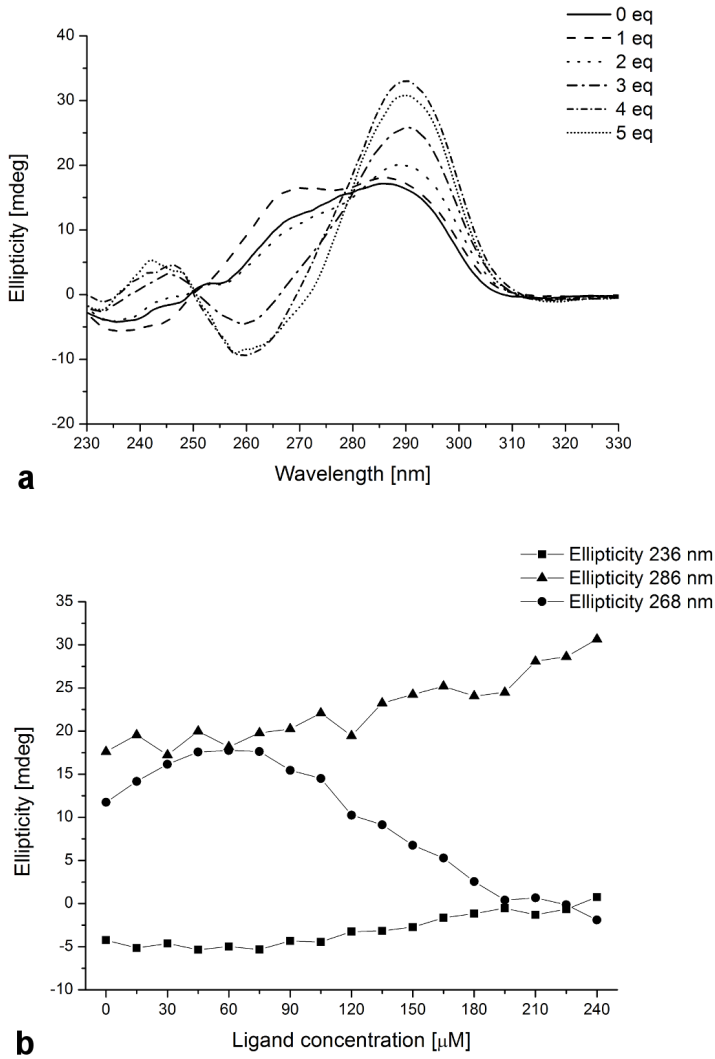


Figure 67: Representative CD profiles of 60 μM Tel23 in presence of different equivalents of the gold metal complex AuNHC **I** after 72 hours of incubation at 25°C (a); changes in intensity at selected wavelengths upon increasing ligand concentration (b). All the experiments were carried out in 10 mM lithium cacodylate buffer pH 7.2, 90 mM LiCl, 10 mM KCl (temperature 25°C).

On the basis of the spectra, it can be speculated that the ligand bind Tel23, as in hybrid structure, with a stoichiometry of 1:1. As the concentration of the metal complex is increased, adducts characterized by higher Tel23:AuNHC **I** stoichiometries emerge in solution. Additionally, the conversion from hybrid to antiparallel G4 topology is triggered, as highlighted by the discontinuity visible in the plot of 268 nm data at the overcoming of 1:1 molar ratio. The Tel23 conversion is quantitatively attained for molar ratios greater than 1:3. Hybrid G4 structures interacting with two molecules are likely to be found at the intermediate 1:2 Tel23:AuNHC **I** stoichiometry. Thus, it can be hypothesized that AuNHC **I** binding takes place through several reaction steps involving the formation of different species in solution, in according to ESI-MS spectra which were performed on solutions where the completion of reaction was most likely not achieved. This hypothesis can also explain why a definite isoelliptic point at about 280 nm is lacking in the spectra shown in figure 66. Probably, in the first set of experiments the saturating amount of the ligand quickly induced the conversion from hybrid to antiparallel topology and, as a consequence, the isoelliptic point at 280 nm was not clearly observed. Spectral changes at 268 nm were subsequently analysed according to the formalism of the Benesi-Hildebrand (BH) plot to gain information on the associated binding constants [347–350] (figure 68). The BH plot was chosen due to the peculiar changes of the signal at 268 nm which prevented us from treating data by other established methods [339, 414, 415]. Although not extremely accurate, the BH plot is a robust method that has allowed the estimation of dissociation constants for the three distinct binding events associated to the formation of 1:1, 1:2 and 1:3 Tel23:AuNHC **I** adducts (table 23). From inspection of these values, it clearly emerges that the binding process, although not significantly strong, is highly cooperative.

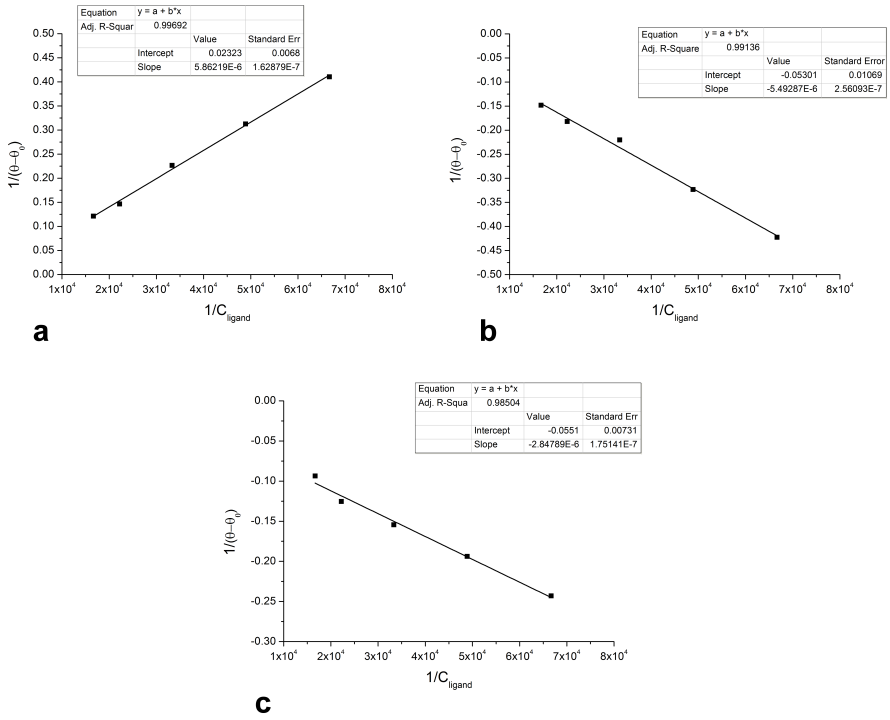


Figure 68: Benesi-Hildebrand plots of CD data at 268 nm for the estimation of dissociation constants for the adducts with G4:ligand molar ratio 1:1 (a) 1:2 (b) and 1:3 (c).

Table 23: Dissociation constants for the binding steps of gold metal complex AuNHC **I** to the human telomeric sequence Tel23. Results obtained by fitting CD data at 268 nm in the Benesi-Hildebrand plot. Estimated fitting errors are reported as well.

Tel23:AuNHC **I** molar ratio	Dissociation constant
1:1	$(2.5 \pm 0.8) \cdot 10^{-4}$ M
1:2	$(1.0 \pm 0.3) \cdot 10^{-4}$ M
1:3	$(5.2 \pm 1.0) \cdot 10^{-5}$ M

In order to verify if the ligand stabilizes the G-quadruplex structure, thermal melting curves were recorded by means of CD spectroscopy (figure 69). The experiments highlighted a modest thermal stabilization, but they confirmed the final 1:3 Tel23:AuNHC **I** stoichiometric ratio for the adducts, in agreement with data already found both at the solid state and through ESI-MS. The thermal stabilization observed was not in good agreement with previously reported data which indicated a more pronounced stabilization, most likely because of the different telomeric sequence and experimental conditions adopted by other authors [314].

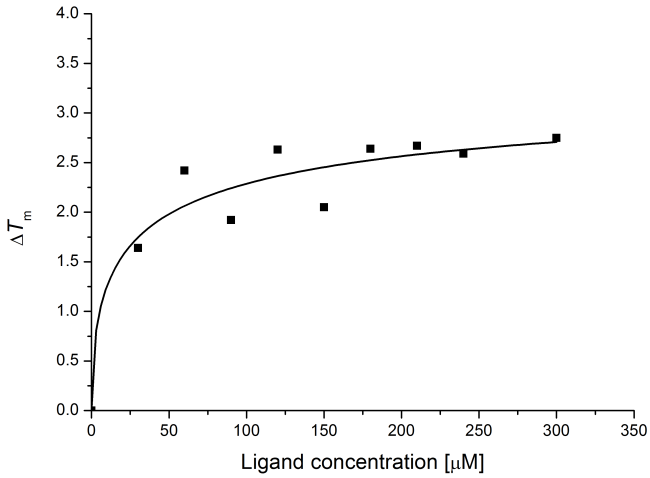


Figure 69: Increase in melting temperature for Tel23 G-quadruplex (60 μM) upon raising AuNHC **I** concentration. Melting observed by CD signal at 290 nm. All the experiments were carried out in 10 mM lithium cacodylate buffer pH 7.2, 90 mM LiCl, 10 mM KCl.

4.4.4 *In silico* calculations

Taking into account the results of the spectrophotometric characterization, *in silico* calculations were performed in order to supply structural models of the Tel23-AuNHC **I** interaction in solution. The starting coordinates for the metal complex were obtained from the crystal structure of [Au(9-methylcaffein-8-ylidene)$_2$] BF_4 (CCDC code 885458) [315]. On the other hand, the tetraplex models were selected as for hybrid-1 (Tel26 d[$A_3G_3(TTAGGG)_3A_2$], PDB code 2HY9 [185]), hybrid-2 (wtTel26 d[$(TTAGGG)_4T_2$], PDB code 2JPZ [190]) and antiparallel-1 basket (Tel22 d[$AG_3(TTAGGG)_3$], PDB code 2MCC [372]) topologies which were modified according to the Tel23 human telomeric sequence, d[$TAG_3(TTAGGG)_3$]. While the chosen hybrid structures concerned native telomeric G-quadruplexes, the antiparallel structure used for the simulations was a ligand-G4 adduct where, similarly to what observed for AuNHC **I**, both external tetrads were able to accommodate ligand molecules.

Firstly, the ligand was QM optimized at the B3LYP/LACV3P$**+$ level of theory and the atomic electrostatic charges were calculated and used in the following steps. Rigid docking calculations were carried out using grids centred on external tetrads and grooves of the targets. Subsequently, molecular dynamics with frozen ligands and QM/MM optimizations (ligand only as the QM region) were recursively applied in order to favour a better disposition of DNA residues in the proximity of the binding sites. In fact, while the force field used can well simulate the behaviour of DNA, it is not suited for the treatment of heavy metals as it induces large distortions in the coordination geometry. Quantum mechanical approaches are more appropriate in this regard. In the end, the final coordinates were optimized by pure QM calculations. Clearly, the procedure has been repeated for the addition of further AuNHC **I** molecules on the targets.

The interaction of three AuNHC **I** molecules and the Tel23 antiparallel G-quadruplex was investigated in the first place. The calculations led to an antiparallel tetraplex displaying two binding sites, one on the external quartet in the proximity of the lateral loops where two molecules are present, and a second one on the G-tetrad close to the diagonal loop where just a ligand molecule is located (figure 70).
Looking at the former, AuNHC **I** adopts a similar configuration to that observed in the crystal structure previously solved and it is involved in π-π stacking interactions with two guanines (interplanar distance around 4.0 Å). Additionally, the T7 residue from a TTA lateral loop stacks on the closest caffeine moiety at about 4.0 Å apart. In according to the binding mode evidenced in the crystal structure, only one caffeine moiety per metal complex interacts efficiently with the aromatic surface of the tetrad as the other one partially protrudes outside from the guanine platform. The gold atoms are placed just above pyrimidine rings of guanines in an analogous manner with respect to that observed at the solid state. Despite the overall resemblance with the crystallographic binding site, the presence of lateral instead of propeller loops takes to a different localization of AuNHC **I** molecules in the binding site, as they are more distant from each other (difference in Au-Au distance 1.0 Å) and one molecule is translated of 2.3 Å along its axis due to the steric hindrance determined by A8 and T18 residues. Furthermore, ligand molecules experience a slight deviation from planarity (dihedral angles between the planes defined by the caffeine moieties lower than 15°).
The other binding site confirms the hypothesis that, at least in solution, AuNHC **I** can stack along the diagonal of a G-quartet, as suggested by the inspection of the crystallographic structure. Actually, the NMR structure chosen for the calculations displays a binding site not hindered by loops and flanking residues and

allowed to successfully test this supposition. AuNHC **I** is sandwiched between the G-tetrad and the TTA diagonal loop where, thanks to its high degree of planarity, it is involved in π-π stacking interactions especially with G3 and G23 residues (inter-planar distances about 3.8 Å). Additional CH$\cdots\pi$ interactions are observed for a methyl group belonging to a caffeine moiety with the A14 residue (carbon-centroid distance 3.5 Å). In spite of the different positioning, the gold atom is still not placed above the G4 channel and a displacement of 2.2 Å from the K$\cdots$K line is observed.

Analogous calculations have been performed for hybrid-1 and hybrid-2 G4 structures bounded to two AuNHC **I** molecules (figure 71). The final structures pointed out the binding of a metal complex at the 3′ end and a second complex at the 5′ end G-tetrad. The molecules are almost perfectly planar and interact through π-π stacking with bases belonging to the quartets (inter-planar distances around 3.8 Å). The ligand binding is strengthen by interactions in which flanking residues and TTA loops are involved (propeller loop in the hybrid-1 adduct whereas lateral loop in the hybrid-2 adduct). Ligand molecules are planar and involved in π-π stacking interactions with bases from the quartets and from the close lateral loops.

Overall, AuNHC **I** molecules can interact with two guanines per G-quartet, although the stacking is significant for just a caffeine moiety, as already observed in the X-ray crystal structure. In all the calculated structures, the gold atoms are likewise not found above the channel of the G-quadruplex. The disposition of ligand molecules along the diagonal of the tetrads has been recognized, hence supporting the reliability of such binding mode for the ligand, at least in solution and when additional ligand molecules are not present. The inspection of the calculated structures suggests that the organization of the loops is determinant for the formation of ligand-G4 adducts. The G-quadruplex topology transition from

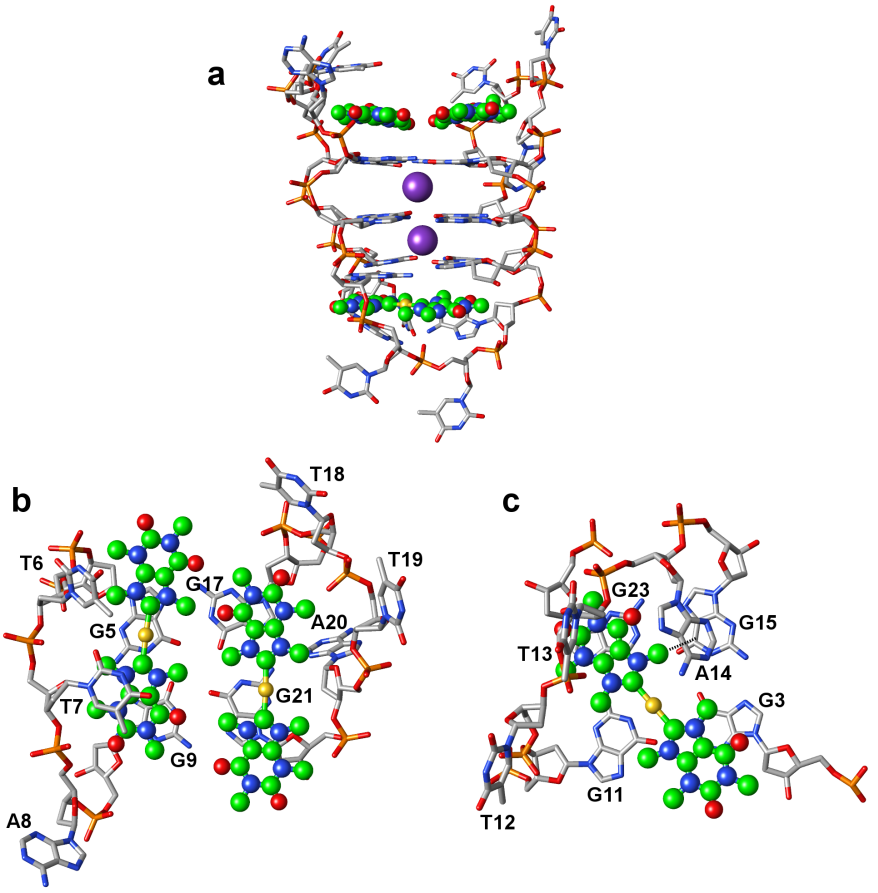


Figure 70: Calculated structure of the antiparallel Tel23 G-quadruplex bounded to three AuNHC **I** molecules. Overall structure (a) together with focuses on the lateral (b) and diagonal loops (c) end binding sites.

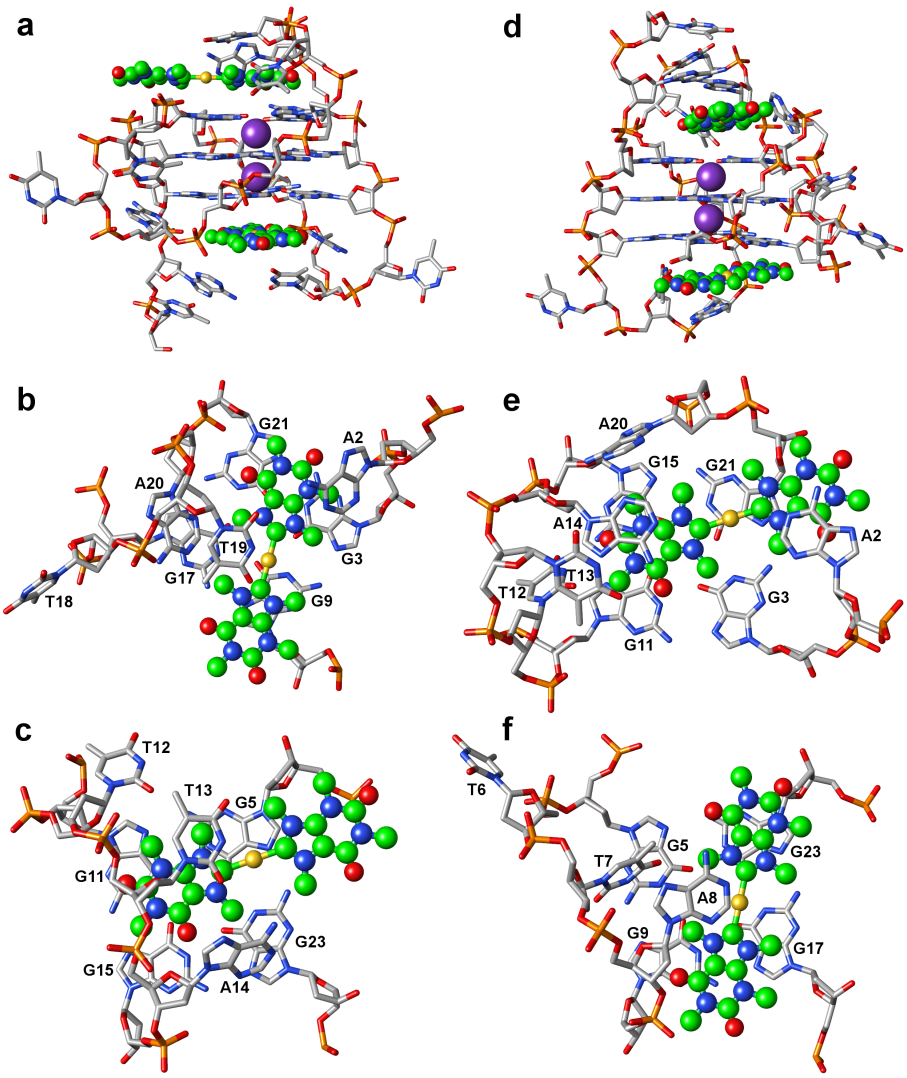


Figure 71: Calculated structures of hybrid-1 (a) and hybrid-2 (d) Tel23 G-quadruplex bounded to two AuNHC **I** molecules with focuses on the 5′ end G-quartet (b,e) and on the 3′ end G-quartet (c,f) binding sites.

hybrid to antiparallel is most likely caused by the rearrangement of loops which allows a better allocation of the ligand on the external G-tetrads. In more detail, the lateral loops of the hybrids, which protrude on the tetrad platform, act as G4 caps and hinder the approach of ligand molecules to the external G-quartets to a greater extent than the loops of the antiparallel structure. Thus, the binding of a third ligand molecule on the 5′ end G-tetrad is energetically favoured and compensates the energetic cost of topological rearrangement. The lack of a clear isoelliptic point at 280 nm in the first set of CD experiments (figure 66) also suggests that the topological conversion might start occurring even at 1:2 Tel23:AuNHC **I** molar ratio in order to favour a better allocation of the ligands on the guanine tetrads.

4.4.5 Telomerase inhibition assays

The interesting binding ability displayed by AuNHC **I** towards Tel23 human telomeric G-quadruplex suggested the potential telomerase inhibition activity of the compound. Telomerase inhibition assays were then performed by Prof. Donato Colangelo from the Università del Piemonte Orientale on Tel23, AuNHC **I** and the Tel23-AuNHC **I** adduct with the purpose to evaluate the effect of the metal complex with respect to the native telomeric tetraplex (figure 72). Cell-free tests were carried out including a true positive control, where telomerase exerted its normal elongation activity on a synthetic telomere (TS), in order to set the reference control activity. The active enzyme was extracted and assayed as previously described, with minor modifications [416]. The experiments demonstrated that telomerase is not able to recognize Tel23 as a natural substrate to be elongated, although some competition between TS and Tel23 could be observed. This may be due to the fact that TS contains a consensus terminal sequence

(GAGTT) that is specifically recognized and elongated by telomerase. Furthermore, AuNHC **I** is not able to form adducts on the TS sequence, since normal TS extension was observed in the presence of the compound during the elongation step of the reaction. Interestingly, the Tel23-AuNHC **I** adduct interferes with this elongation step and this interference is additive to the interference shown for Tel23 alone. It is noteworthy that telomerase is not able to elongate the adduct. These data are interesting, but the experiments can not give insight into the mechanism of action responsible for the lack of elongation. Some hypotheses can be theorized which include the stuck of the enzyme, the lack of recognition of the sequence by telomerase or the partial recognition not sufficient to induce the elongation. Long term treatments of cells with the metal complex will be able to show if telomerase can be inhibited by this compound and the telomere length measurements will give conclusive data on the potential of AuNHC **I** on modulating telomere functions and cell proliferation.

In summary, AuNHC **I** has shown promising G4 binding properties associated to the formation of adducts with different stoichiometries and a slow kinetics. In spite of a relatively not strong, yet cooperative, binding (K_ds of 10^{-4}-10^{-5} M magnitude), the ligand is able to induce a topological transition on Tel23 in solution. The ligand can bind to Tel23 up to 1:3 Tel23:AuNHC **I** molar ratio, as observed by means of X-ray crystallography, CD spectroscopy and ESI-MS spectra. A more extensive investigation on the binding of the ligand to different G-quadruplex topologies has clarified the behaviour observed in solution. Actually, the interaction of the compound at the external G-tetrads is quite hindered by flanking residues and TTA loops and it is optimized when the G-quadruplex assumes the antiparallel-1 or parallel topology. Besides the possible dispositions of the ligand on the tetrads, AuNHC **I** binding is mainly driven by non-directional forces as π-π stack-

ing, electrostatic and Van der Waals interactions. In the end, telomerase inhibition assays have further supported the idea that AuNHC **I** induce a relevant interference on the enzyme activity thus indicating the compound as a promising potential anticancer drug.

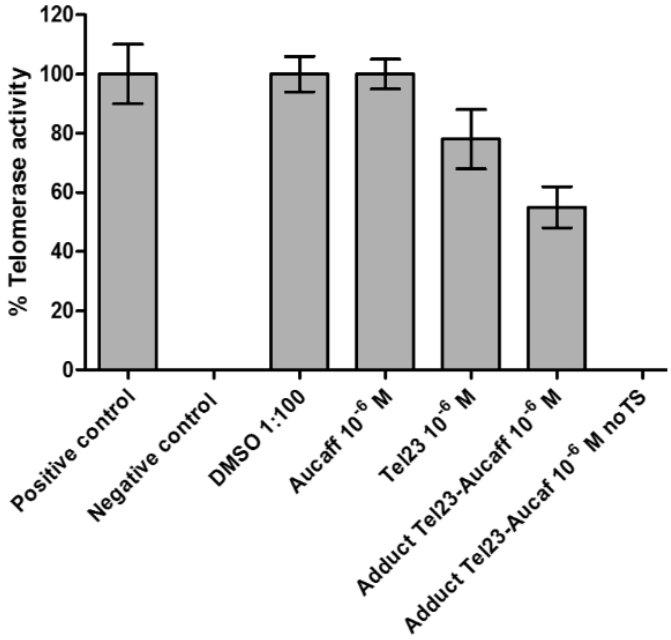


Figure 72: Telomerase activity measured as the ability to elongate a synthetic telomere TS in presence of AuNHC **I**, Tel23 or the adduct Tel23-AuNHC **I** (bar 4, 5, 6, respectively). Data are shown as the mean±SD activity respect to the untreated control (bar 1). Bar 2, 3 and 7 represent the true negative control (no TS), the interference control (vehicle) and telomerase in presence of the adduct Tel23-AuNHC **I** alone, respectively. Bar 5 $p<0.05$ and bar 6 $p<0.01$ versus untreated control. Adapted from reference [353].

4.5 Au^{I} N-HETEROCYCLIC CARBENE COMPLEXES, AUNHC II

The investigation of Au^{I} metal complexes able to interact and stabilize human telomeric G-quadruplex foldings was later extended to AuNHC **II** (figure 73) which features NHC groups with reduced aromatic surfaces and bulky alkyl side chains. The compound has shown valuable anticancer properties [316, 317] and possesses structural similarity to AuNHC **I**, although it is significantly shorter.

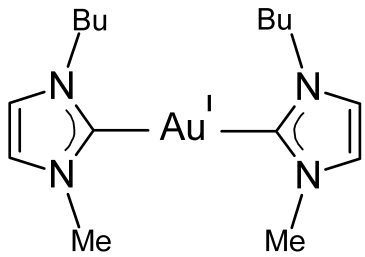


Figure 73: The AuNHC **II** complex studied as G_4 binding ligand.

4.5.1 X-ray crystallography

Crystallization screenings were performed for the adducts formed by the ligand and the human telomeric sequences Tel12 d[TAG_3-TTAGGGT] and Tel24′ d[$TAG_3(TTAGGG)_3T$] which yielded Tel24′-AuNHC **II** crystals suitable for X-ray diffraction analysis (table 24).

The crystal structure of the adduct was solved and it shows monomolecular parallel G-quadruplexes that give rise to columns growing along the two-fold rotation axis (figure 74). The G_4 struc-

Table 24: Crystal data and refinement statistics for the Tel24/-AuNHC **II** structure. As regards the ligand, only gold atoms have been added to the structure.

Space group	C 2
Crystal system	monoclinic
Cell dimension (Å)	a=36.60, b=71.37, c=27.05
Cell angles (°)	$\alpha = \gamma = 90,\ \beta = 92.42$
Resolution range (Å)	35.69-1.80
Rfactor	0.231
Rfree	0.284
RMSD bonds (Å)	0.008
RMSD angles (°)	1.650

tures are characterized by three stacked G-tetrads (inter-planar distances 3.4 Å) and potassium ions in the internal channel (2.6-3.1 Å apart from the guanine O6 atoms). The TTA propeller loops appear to belong to an unclassified loop topology, in spite of a similar configuration with respect to the type-1 loops, in according to the analysis carried out by Neidle and coworkers [379]. In fact, while the first thymine (TTA) stacks on the adenine, the second thymine (TTA) is directed outwards from the G-tetrads in a different fashion with respect those evidenced for example in the Tel12-Coptisine and Tel12-NAX053 crystal structures. Symmetry related G_4 units stack on one another and form pairs related by the two-fold rotation axis. The formation of such dimers is favoured by the presence of potassium ions at the interface between adjacent quadruplexes (figure 74).

The binding site of the ligand is defined by symmetry related end G-tetrads apart 11 Å from one another, contrary to what observed in previously described structures, all displaying an ≈ 7 Å gap. Unfortunately, the electron density maps in the binding site could not be clearly interpreted. Nevertheless, eight significantly high peaks in the density maps confirmed the presence of heavy atoms, as also suggested by the diffraction data featuring anoma-

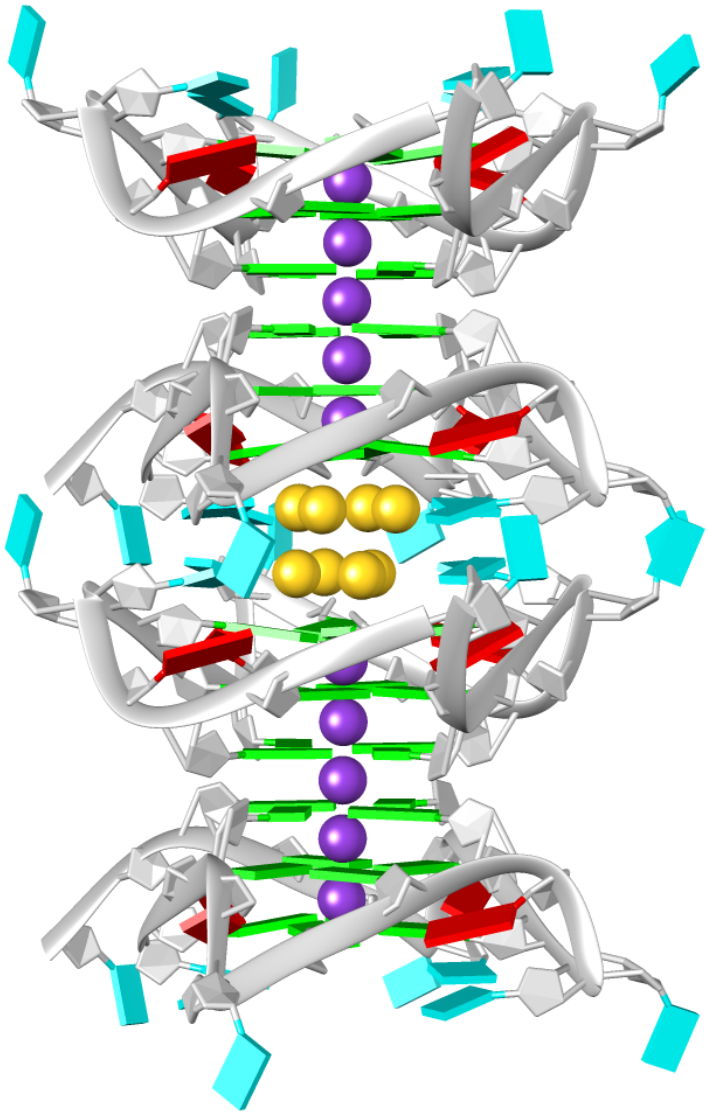


Figure 74: Columnar disposition of stacked G-quadruplexes in the crystal packing of the Tel24/-AuNHC **II** crystal structure. Guanine, thymine and adenine residues coloured in green, cyan and red, respectively. Gold atoms depicted in yellow.

lous signal. In particular, two planes parallel to each other and to the tetrads were observed (inter-planar distances 3.7 Å), each containing four electron density peaks. Every peak was consistent with a quarter of the electron density of a gold atom. This led to hypothesize the presence of a metal complex per plane, spread over four different positions (figure 75). While the electron density of gold atoms was clearly observed, the same was not true for the lighter carbon, nitrogen and oxygen atoms belonging to the N-heterocyclic carbene moieties. Anyway, the overall G4:ligand stoichiometry appears to be 1:1.

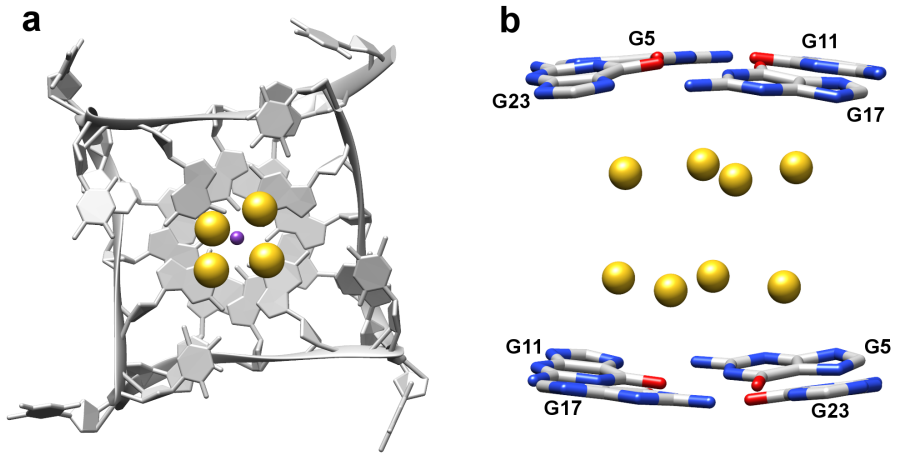


Figure 75: Crystal structure of the Tel24/-AuNHC **II** adduct (a) with a focus on the binding site of the ligand (b).

Despite the significant disorder affecting the ligand, the X-ray crystal structure proved that AuNHC **II** interacts with the guanines tetrads and its binding might be similar to that observed for AuNHC **I**. Notably, the gold atoms of both complexes in their respective crystal structures are localized above the pyrimidine rings of guanines (Au-N1 distances 3.5 Å on average). The disorder ob-

served at the solid state for AuNHC **II** is probably caused by the reduced aromatic surface of the ligand, which, together with the non-directional character of the ligand-G4 interactions, determine an overall modest binding strength. The alkyl side groups most likely exacerbate the disorder.

4.5.2 Corroborative *in silico* calculations

Taking into account the structural information achieved through X-ray crystallography, *in silico* calculations were carried out with the purpose to find the most probable ligand positions in the Tel24/-AuNHC **II** adduct. The starting coordinates for AuNHC **II** were obtained from the crystal structure of [Au(bis(1-butyl-3-methyl-imidazole-2-ylidene))]PF_6 (CCDC code 936581) [317]. The crystallographic coordinates of the Tel24/ G-quadruplex were used as a frozen target in every step. The ligand was quantum mechanically optimized as in both *cis* and *trans* conformations at the B3LYP/LACV3P$**+$ level of theory and the atomic electrostatic charges were calculated and used in the following steps. The ligand was flexibly docked on the target restraining the gold atom positions to those evidenced by crystallography. Flexible docking was necessary to let the alkyl chains move. One by one, all the four gold positions close to the same G-tetrad were considered for the calculations. The resulting adducts were then submitted to molecular dynamics simulations with suitable constraints for the ligand geometry (N-C-Au-C dihedral angles set to zero) and the gold atom positions, and then optimized at QM/MM level of theory (ligand only as the QM region).
The results of the simulations essentially indicate that the *cis* conformation is the most probable for the ligand in the adduct. In each position, linear AuNHC **II** molecules are found to stack on the external G-tetrads being involved in π-π interactions with two

guanines (inter-planar distances 3.5 Å, figure 76). It is noteworthy to highlight that the aromatic moieties of the ligand are both involved in such interactions, contrary to what observed in the AuNHC **I** crystal structure where the size of the ligand did not allow a good matching between the caffeine moieties and the quartets. The positions of gold atoms, which were not restrained in the final QM/MM optimization, are slightly displaced from those refined from crystallographic data (displacements in the range 0.3-0.4 Å) and are similarly located above pyrimidine rings of guanines. The alkyl chains of the molecules appear to interact with the guanines so reinforcing the overall binding. Anyway, the simulations have supplied a reasonable binding mode of the ligand to telomeric G-quadruplexes. Further efforts will be made in the future in the attempt to study the interaction of the ligand with human telomeric sequences in solution.

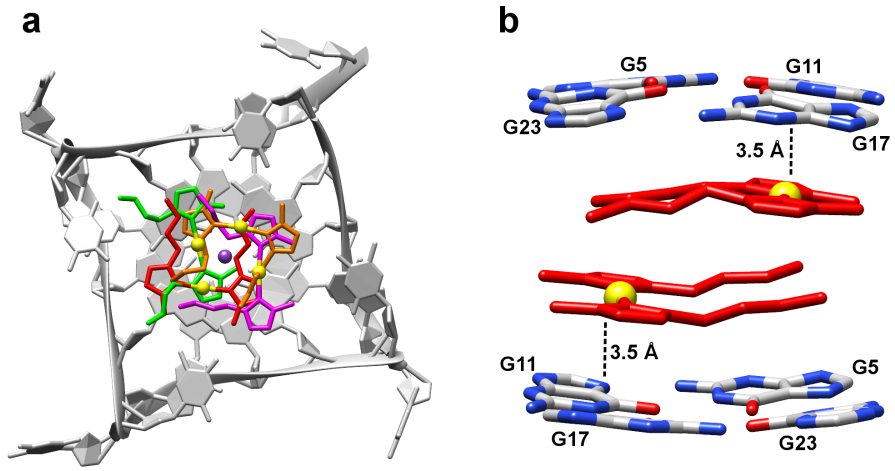


Figure 76: Poses calculated for the ligand in the Tel24/-AuNHC **II** crystal structure (a). The binding site can be filled by the presence of two AuNHC **II** molecules at stacking distance from one another (b).

4.6 Au^{III} COMPLEX AUOXO6

The Au^{III} complex Auoxo6 (figure 77) has displayed valuable telomeric G4 binding properties along with poor interaction with duplex DNA [322]. The anticancer activity of the compound might be indeed connected to G-quadruplex stabilization and consequent telomerase inhibition. Structural detailed data about the ligand-G4 adducts could be useful for a better understanding of binding mode of such compound and could offer useful suggestions for the development of further anticancer drug candidates. For this reason, crystallization screenings of the adducts formed with the human telomeric sequences Tel12 d[$TAG_3TTAGGGT$] and Tel23 d[$TAG_3(TTAGGG)_3$] were performed, but unfortunately no crystals suitable for X-ray diffraction analysis were obtained. The request of structural data about the ligand-G4 interaction was then moved to NMR spectroscopy. Thanks to a scientific collaboration, NMR experiments have been carried out in the research group of Prof. Harald Schwalbe from the University of Frankfurt on the adducts formed by the ligand and the human telomeric sequences Tel26 d[$A_3G_3(TTAGGG)_3A_2$] and wtTel26 d[$(TTAGGG)_4T_2$], able to fold into hybrid-1 and hybrid-2 G-quadruplex, respectively [185, 190]. The personal research activity within the project has been focused in the simulations of Auoxo6 bounded to such sequences in 1:1 stoichiometric ratio, taking into account the information obtained from 1D-NMR and 2D-NMR experiments. The coordinates of hybrid-1 Tel26 (PDB code 2HY9 [185]) and hybrid-2 wtTel26 (PDB code 2JPZ [190]) were used for the simulations. The combined efforts have led to the solution of the wtTel26-Auoxo6 solution NMR structure which has been recently published [417], while the investigation of the binding of Auoxo6 to the Tel26 sequence is still ongoing and not presented here.

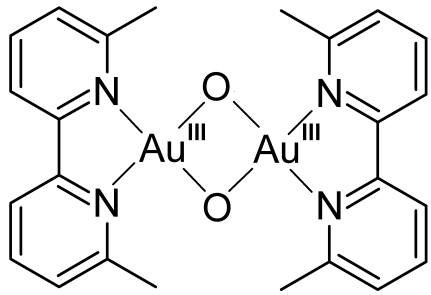


Figure 77: The Auoxo6 complex studied as G4 binding ligand.

4.6.1 *In silico* calculations and wtTel26-Auoxo6 NMR structure

NMR experiments have pointed out the binding of Auoxo6 on the 5′ end G-tetrads of the hybrid structures, especially with the G10 and G22 residues of Tel26 whereas the G12 and G16 residues of wtTel26. Thus, the ligand seems to stack on the 5′ end G-quartet by means of its bipyridine groups having the gold atoms positioned near the internal channel.
The starting coordinates for the metal complex were obtained from the crystal structure of $[(bipy^{2Me})_2Au_2(\mu\text{-}O)_2](PF_6)_2$ (CSD entry QIYZII) [418] and were quantum mechanically optimized at the B3LYP/LACV3P$**+$ level of theory. Atomic electrostatic charges were also calculated and used in the subsequent steps. The ligand was rigidly docked using grids centred on the 5′ end G-tetrads. The poses were then scored on the basis of the minimum distance between the gold atoms of the ligand and the potassium ions of the G4 internal channel. Molecular dynamics simulations were then carried out by using ligand restraints in order to retain the gold coordination geometries (N-Au-O-Au dihedral angles set to zero). In spite of the restraints used, the MM dynamics took to a large alteration of the ligand's geometry, hence the QM optimized metal complex was docked again on the targets centring

the grids where the distorted molecules were found at the end of the simulations. QM/MM calculations were then performed on the adducts (ligand and tetrads' guanines at stacking distance as the QM region).
Hybrid-2 topology shows parallel/antiparallel strands along with a lateral and two propeller loops [190]. The binding of the ligand is coincidentally observed at the 5′ end G-tetrad, in the proximity of the lateral loop, and it takes to a reduced mobility of the G-quartets contrary to the high flexibility of loops. Thanks to the good degree of planarity, the ligand interacts with guanines of the tetrad, especially with G4, G12 and G16 residues (figure 78). Additional π-π stacking interactions involve also Auoxo6 and the A3 residue from the 5′ flank. It is noteworthy how the $[Au_2O_2]^{2+}$ group is positioned above the channel of the quadruplex, where a negative electron density is present [57–59], contrary to what observed for the Au^{I} ligands studied. It can be hypothesized that the alternative geometry and size of Auoxo6 with respect to AuNHC **I** and AuNHC **II** (figure 61,73) have determined a different localization of gold atoms on the G-tetrad, probably also favoured by the hard character of Au^{III} ions.

4.7 PALLADIUM AND PLATINUM COMPLEXES

The palladium and platinum complexes $[Pd(bapbpy)]^{2+}$ and $[Pt(bapbpy)]^{2+}$ (figure 79) have shown remarkable anticancer properties possibly associated to DNA binding. As they display good features for the interaction with non-canonical DNA foldings, structural studies have been carried in order to supply detailed information regarding their binding properties. The information achieved could be useful for a better understanding of the biological activity of the complexes. Up to now, the efforts have been

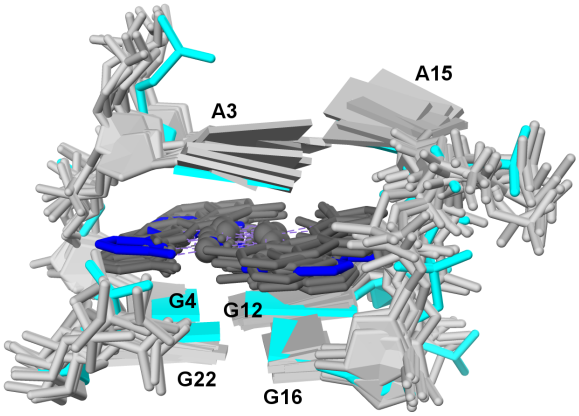


Figure 78: Superposition of *in silico* QM/MM calculated and NMR ensemble structures of the wtTel26-Auoxo6 adduct, PDB code 5MVB [417] (cyan, blue and light grey, dark grey coloured, respectively).

focused on crystallographic experiments taking into account different DNA sequences. In addition, *in silico* calculations have provided complementary data.

4.7.1 X-ray crystallography

Crystallization screenings were performed for the adducts formed by the metal complexes with the human telomeric sequences Tel12 d[$TAG_3TTAGGGT$] and Tel23 d[$TAG_3(TTAGGG)_3$] as well as the oligonucleotides d[CGTACG] and d[CGATCG] able to give rise to local non-B DNA arrangements. Up to now, suitable crystals for X-ray diffraction analysis have been obtained only for the CGTACG-$[Pt(bapbpy)]^{2+}$ adduct whose crystal structure has been recently solved (table 25).

The crystal structure displays double-stranded DNA units and ligand molecules in 2:1 stoichiometric ratio. The duplex units

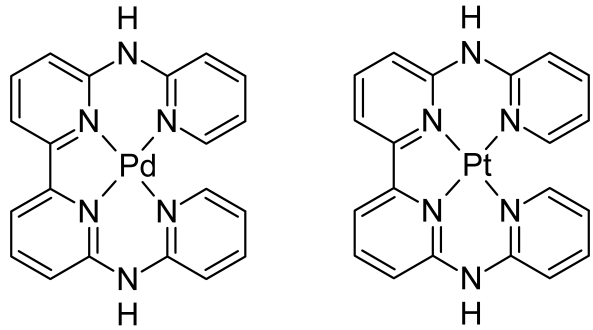


Figure 79: The palladium and platinum complexes studied as binders of non-canonical DNA structures.

Table 25: Crystal data and refinement statistics for the CGTACG-$[Pt(bapbpy)]^{2+}$ structure.

Space group	P $3_2$21
Crystal system	trigonal
Cell dimension (Å)	a=b=30.21, c=117.44
Cell angles (°)	$\alpha = \beta = 90,\ \gamma = 120$
Resolution range (Å)	26.16-2.30
Rfactor	0.243
Rfree	0.326
RMSD bonds (Å)	0.010
RMSD angles (°)	1.580
Twin law	-h -k l
Twin fraction	0.26

stack on one another and form columns growing along the [110], [010] and [100] crystallographic directions (figure 80). The binding site of the ligand is observed at the interface between adjacent B-type helices where a non-canonical arrangement can be recognized. The residues involved in the binding site show significant distortion from the classical B-DNA organization as well as the bases located at the junction between contiguous DNA units.
The binding site of the ligand is defined by an eight-bases arrange-

ment composed of the 3′ end C5 and G6 residues (chain A,C), the 5′ end G8 (chain B,D) and additional C7 residues (chain B,D) from symmetry related units (figure 80). The bases give rise to quite planar CGGC quartets characterized by C5-G8 and G6-C7 base pairs (dihedral angle between the base pairs about 23°) together with hydrogen bonds formed by G6 and G8 residues interacting through their minor groove faces (guanines form a dihedral angle of 15-18°, inter-atomic distances 2.3-2.7 Å). Unfortunately, the disorder observed for $[Pt(bapbpy)]^{2+}$ together with the high twin fraction of the crystal analysed (≈25%) make the electron density maps difficult to be interpreted. Anyway, some attempts have been made in order to identify the most probable conformations of the metal complex at the binding site.

The binding site hosts two platinum atoms about 2.5 Å apart from each other, which were clearly localized by analysing the anomalous signal from diffraction data. As a consequence, two different positions of $[Pt(bapbpy)]^{2+}$ share the site. The best results in terms of residual densities in the $[F_o\text{-}F_c]$ maps and atomic thermal factors pointed out two ligand configurations, indicated in the following as complex **F** and complex **G**, almost coplanar and approximately related by a rotation of 90° along the helical axis (omit map figure 81). The ligand molecules mainly interact by π-π stacking with DNA bases of the binding site (mean inter-planar distances 3.4 Å, figure 80). The binding is further strengthen by electrostatic attraction between the positively charged platinum ions and the negative electron density of the DNA bases.

The 5′ end C7 residue of each CGTACG strand protrudes outside the helix and gives rise to a connection with an adjacent, symmetry related, column in the crystal. Interestingly, this kind of connection between adjacent columns reveals a structural motif which resembles that assumed by the Holliday junction of stacked-X type (figure 82) [34, 35, 37–39]. Despite the obvious diversity between

the present structure and Holliday junctions (HJ), $[Pt(bapbpy)]^{2+}$ is found at such connection similarly to what observed in the crystal structures of previously reported adducts formed by HJ-DNA (PDB codes 1FHY, 2GWA) [419, 420]. This finding might indicate that the metal complex could be able to interact with this kind of secondary DNA structure.

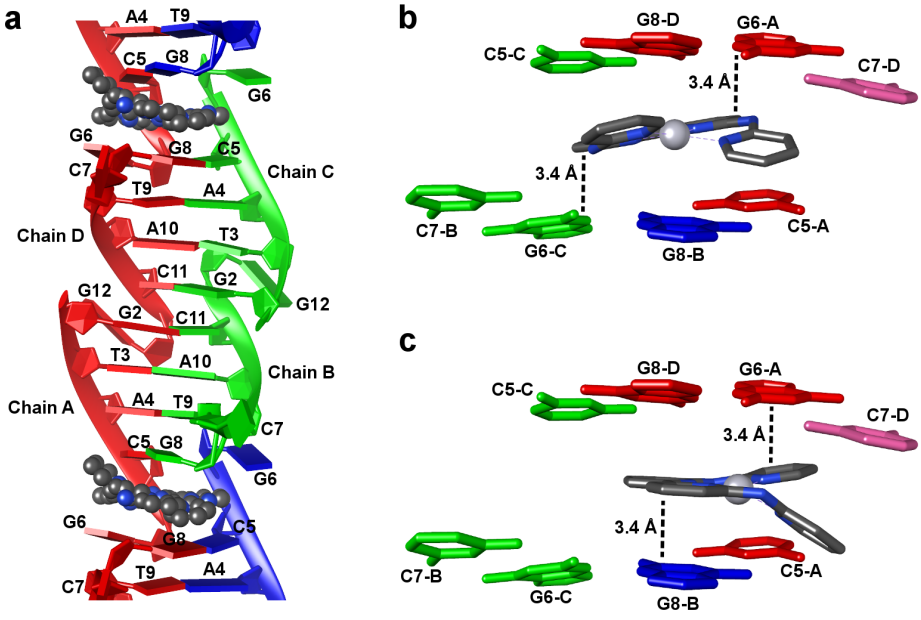


Figure 80: Crystal structure of the $[Pt(bapbpy)]^{2+}$-CGTACG adduct. Formation of columns of double-stranded DNA along the [110] direction (a). Focuses on the binding site filled by complex **F** (b) and complex **G** (c).

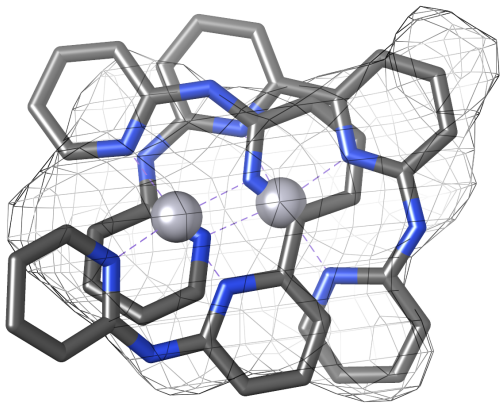


Figure 81: The platinum complex molecules shown within the electron density map [F_o-F_c] at 1.5σ level.

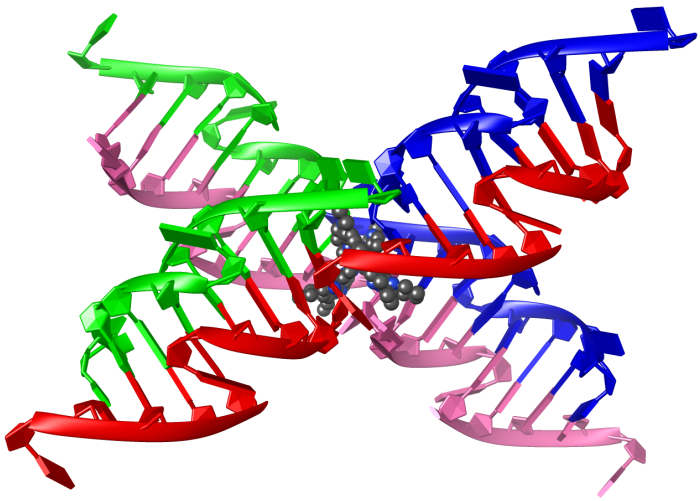


Figure 82: The crystal packing of the [Pt(bapbpy)]$^{2+}$-CGTACG structure defines a motif which is similar to a Holliday junction organization of strands.

4.7.2 Corroborative *in silico* calculations

In order to obtain a more detailed view of the overall interactions, QM/MM calculations were performed on the coordinates of each ligand-DNA adduct obtained by X-ray crystallography. On the whole, the simulations confirmed the crystallographic position for each ligand (figure 83). Metal complex **F** is involved in π-π stacking interactions with C5 and G6 residues from chain A, G8 residue from chain B, C5 residue from chain C and G8 residue from chain D (inter-planar distances 3.2-3.3 Å). Conversely, complex **G** interacts with C5 and G6 residues from chain A, G8 residue from chain B and D (inter-planar distances 3.3-3.6 Å).

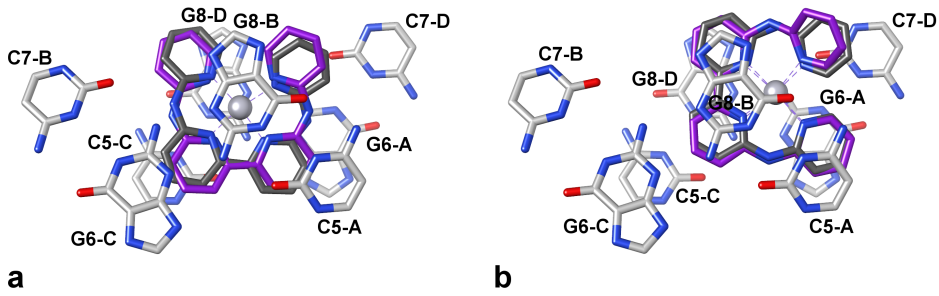


Figure 83: Superposition of the crystallographic (dark grey coloured) and *in silico* calculated (purple coloured) positions of complex **F** (a) and complex **G** (b) in the binding site of the $[Pt(bapbpy)]^{2+}$-CGTACG structure.

5 CONCLUSIONS

The research project has focused its attention on ligands able to interact and stabilize non-canonical DNA foldings. As these secondary DNA structures show high biological activity and are probably involved in various processes, their targeting might open stimulating perspectives for the treatment of various diseases, including cancer. Human telomeric DNA has emerged in particular as a promising target in anticancer drug discovery because it is able to fold into G-quadruplex structures exerting inhibition activity on telomerase, which is an enzyme implicated in the proliferation of about 85% of cancer forms. During the three years of PhD, the interaction of non-canonical foldings with natural and synthetic compounds has been investigated in order to select valuable anticancer drug candidates.

Natural alkaloids Berberine, Coptisine, Sanguinarine and Chelerythrine have shown notable binding properties towards telomeric tetraplexes, with particular mention of Coptisine and Chelerythrine that poorly stabilize duplex DNA. The detailed analyses of their interaction with such DNA structures have allowed to shed light on the importance of the dioxolo groups and the overall planarity of the molecules in the formation of stable adducts.

The research activity has also concerned derivatives of Berberine featuring functionalizations at the 13-position. The analogs have shown improved G4 binding properties accompanied by better cytotoxic activity towards cancer cells in comparison to the natural precursor. Structural investigations have pointed out the determinant role of alkyl chain length and kind of pendant groups in the

stability of resulting adducts. These factors shape the planarity of the ligands that seems crucial for the improvement of the G4 binding properties. Interestingly, some derivatives have shown the ability to bind multiple G-quadruplexes in solution which is a intriguing scenario for the targeting of human telomeres.
As regards synthetic compounds, gold metal complexes have covered an important section of the research. Au_I complexes have displayed quite singular behaviour for the interaction with human telomeric G-quadruplexes. The binding of AuNHC **I** has been analysed in great detail as it is characterized by multiple reaction steps along with a slow kinetics. The formation of multiple adducts in solution has been experimentally observed and theoretically simulated revealing the important dimensional matching between the ligand and the binding site on the G-tetrads. As a consequence, the binding of AuNHC **I** appears to be greatly influenced by the quadruplex topology and in particular by the disposition of loops and flanking residues in the proximity of guanine tetrads. Interestingly, G4-AuNHC **I** adducts have shown valuable telomerase inhibition activity. On the basis of such results, the structural investigation has been extended to AuNHC **II** which however has displayed significant disorder when bounded to human telomeric G-quadruplexes at the solid state. In according to the X-ray crystal structure, the binding of the ligand has been theoretically predicted by *in silico* calculations. Solution studies will be anyway helpful to draw more complete conclusions about the binding properties of such compound.
The interaction of the Au^{III} complex Auoxo6 with telomeric G-quadruplex has been studied by a combined *in silico* and NMR investigation which has successfully led to the NMR structure of the adduct formed by the ligand and the wtTel26 sequence. The structure has clarified the binding mode of the ligand and also the consequent rearrangements of loops and flanking bases involved

in the capping of the external G-tetrads.
Furthermore, $[Pd(bapbpy)]^{2+}$ and $[Pt(bapbpy)]^{2+}$ complexes were studied as ligands able to bind non-canonical DNA structures. The platinum complex was found to interact with local non-canonical arrangements of double helix DNA by both crystallography and *in silico* calculations. The packing of the crystal structure suggests also the potential binding properties of the compounds towards Holliday junctions which will be the object of future research.

BIBLIOGRAPHY

[1] E. P. Consortium et al., "An integrated encyclopedia of DNA elements in the human genome", *Nature* **2012**, *489*, 57–74.

[2] S. P. Jackson, J. Bartek, "The DNA-damage response in human biology and disease", *Nature* **2009**, *461*, 1071–1078.

[3] B. Stewart, C. P. Wild, et al., "World cancer report 2014", *Health* **2017**.

[4] T. M. Allen, "Ligand-targeted therapeutics in anticancer therapy", *Nature reviews. Cancer* **2002**, *2*, 750–763.

[5] J. Zhao, A. Bacolla, G. Wang, K. M. Vasquez, "Non-B DNA structure-induced genetic instability and evolution", *Cellular and Molecular Life Sciences* **2010**, *67*, 43–62.

[6] R. Dahm, "Friedrich Miescher and the discovery of DNA", *Developmental biology* **2005**, *278*, 274–288.

[7] C. L. Woodcock, S. Dimitrov, "Higher-order structure of chromatin and chromosomes", *Current opinion in genetics & development* **2001**, *11*, 130–135.

[8] J. P. Reddington, S. Pennings, R. R. Meehan, "Non-canonical functions of the DNA methylome in gene regulation", *Biochemical Journal* **2013**, *451*, 13–23.

[9] T. Carell, M. Q. Kurz, M. Müller, M. Rossa, F. Spada, "Non-canonical bases in the genome: The regulatory information layer in DNA", *Angewandte Chemie* **2017**.

[10] S. Neidle, *Principles of nucleic acid structure*, Academic Press, **2010**.

[11] M. Egli, W. Saenger, *Principles of nucleic acid structure*, Springer Science & Business Media, **2013**.

[12] J. D. Watson, F. H. Crick, et al., "Molecular structure of nucleic acids", *Nature* **1953**, *171*, 737–738.

[13] D. W. Ussery, "DNA Structure: A-, B- and Z-DNA Helix Families", *eLS* **2002**.

[14] L. Pray, "Discovery of DNA structure and function: Watson and Crick", *Nature Education* **2008**, *1*, 100.

[15] P. Jurečka, P. Hobza, "True Stabilization Energies for the Optimal Planar Hydrogen-Bonded and Stacked Structures of Guanine-Cytosine, Adenine-Thymine, and Their 9- and 1-Methyl Derivatives: Complete Basis Set Calculations at the MP2 and CCSD(T) Levels and Comparison with Experiment", *Journal of the American Chemical Society* **2003**, *125*, 15608–15613.

[16] J. Šponer, P. Jurečka, P. Hobza, "Accurate interaction energies of hydrogen-bonded nucleic acid base pairs", *Journal of the American Chemical Society* **2004**, *126*, 10142–10151.

[17] Y. Mo, "Probing the nature of hydrogen bonds in DNA base pairs", *Journal of molecular modeling* **2006**, *12*, 665–672.

[18] R. Sedlák, P. Jurečka, P. Hobza, "Density functional theory-symmetry adapted perturbation treatment energy decomposition of nucleic acid base pairs taken from DNA crystal geometry", *The Journal of chemical physics* **2007**, *127*, 08B610.

[19] J. Černý, M. Kabeláč, P. Hobza, "Double-helical-ladder structural transition in the B-DNA is induced by a loss of dispersion energy", *Journal of the American Chemical Society* **2008**, *130*, 16055–16059.

[20] P. Yakovchuk, E. Protozanova, M. D. Frank-Kamenetskii, "Base-stacking and base-pairing contributions into thermal stability of the DNA double helix", *Nucleic acids research* **2006**, *34*, 564–574.

[21] M. Kolář, T. Kubař, P. Hobza, "On the role of London dispersion forces in biomolecular structure determination", *The Journal of Physical Chemistry B* **2011**, *115*, 8038–8046.

[22] J. Lipfert, S. Doniach, R. Das, D. Herschlag, "Understanding nucleic acid-ion interactions", *Annual review of biochemistry* **2014**, *83*, 813–841.

[23] A. Vologodskii, M. D. Frank-Kamenetskii, "Strong bending of the DNA double helix", *Nucleic acids research* **2013**, *41*, 6785–6792.

[24] A. Ghosh, M. Bansal, "A glossary of DNA structures from A to Z", *Acta Crystallographica Section D: Biological Crystallography* **2003**, *59*, 620–626.

[25] D. Svozil, J. Kalina, M. Omelka, B. Schneider, "DNA conformations and their sequence preferences", *Nucleic acids research* **2008**, *36*, 3690–3706.

[26] F. DiMaio, X. Yu, E. Rensen, M. Krupovic, D. Prangishvili, E. H. Egelman, "A virus that infects a hyperthermophile encapsidates A-form DNA", *Science* **2015**, *348*, 914–917.

[27] A. Rich, S. Zhang, "Z-DNA: the long road to biological function", *Nature reviews. Genetics* **2003**, *4*, 566–572.

[28] G. Wang, K. M. Vasquez, "Z-DNA, an active element in the genome.", *Frontiers in bioscience: a journal and virtual library* **2007**, *12*, 4424–4438.

[29] N. Saini, Y. Zhang, K. Usdin, K. S. Lobachev, "When secondary comes first-the importance of non-canonical DNA structures", *Biochimie* **2013**, *95*, 117–123.

[30] M. Kaushik, S. Kaushik, K. Roy, A. Singh, S. Mahendru, M. Kumar, S. Chaudhary, S. Ahmed, S. Kukreti, "A bouquet of DNA structures: Emerging diversity", *Biochemistry and Biophysics Reports* **2016**, *5*, 388–395.

[31] Y. Du, X. Zhou, "Targeting Non-B-Form DNA in Living Cells", *The Chemical Record* **2013**, *13*, 371–384.

[32] G. R. Smith, "Meeting DNA palindromes head-to-head", *Genes & development* **2008**, *22*, 2612–2620.

[33] V. Brázda, R. C. Laister, E. B. Jagelská, C. Arrowsmith, "Cruciform structures are a common DNA feature important for regulating biological processes", *BMC molecular biology* **2011**, *12*, 33.

[34] D. M. Lilley, "Structures of helical junctions in nucleic acids", *Quarterly reviews of biophysics* **2000**, *33*, 109–159.

[35] Y. Liu, S. C. West, "Happy Hollidays: 40th anniversary of the Holliday junction", *Nature Reviews Molecular Cell Biology* **2004**, *5*, 937–944.

[36] W. Wang, L. M. Nocka, B. Z. Wiemann, D. M. Hinckley, I. Mukerji, F. W. Starr, "Holliday Junction Thermodynamics and Structure: Coarse-Grained Simulations and Experiments", *Scientific reports* **2016**, *6*, 22863.

[37] J. Matos, S. C. West, "Holliday junction resolution: regulation in space and time", *DNA repair* **2014**, *19*, 176–181.

[38] L. A. Howell, M. Searcey, "Targeting Higher-Order DNA: Beyond the G-Quadruplex", *ChemBioChem* **2009**, *10*, 2139–2143.

[39] P. S. Ho, "Structure of the Holliday junction: applications beyond recombination", *Biochemical Society Transactions* **2017**, BST20170048.

[40] K. Hoogsteen, "The structure of crystals containing a hydrogen-bonded complex of 1-methylthymine and 9-methyladenine", *Acta crystallographica* **1959**, *12*, 822–823.

[41] K. Hoogsteen, "The crystal and molecular structure of a hydrogen-bonded complex between 1-methylthymine and 9-methyladenine", *Acta Crystallographica* **1963**, *16*, 907–916.

[42] H. Zhou, B. J. Hintze, I. J. Kimsey, B. Sathyamoorthy, S. Yang, J. S. Richardson, H. M. Al-Hashimi, "New insights into Hoogsteen base pairs in DNA duplexes from a structure-based survey", *Nucleic acids research* **2015**, *43*, 3420–3433.

[43] E. C. Friedberg, G. C. Walker, W. Siede, R. D. Wood, *DNA repair and mutagenesis*, American Society for Microbiology Press, **2005**.

[44] M. Duca, P. Vekhoff, K. Oussedik, L. Halby, P. B. Arimondo, "The triple helix: 50 years later, the outcome", *Nucleic acids research* **2008**, *36*, 5123–5138.

[45] A. Jain, G. Wang, K. M. Vasquez, "DNA triple helices: biological consequences and therapeutic potential", *Biochimie* **2008**, *90*, 1117–1130.

[46] A. Bacolla, G. Wang, K. M. Vasquez, "New perspectives on DNA and RNA triplexes as effectors of biological activity", *PLoS genetics* **2015**, *11*, e1005696.

[47] M. Brázdová, V. Tichỳ, R. Helma, P. Bažantová, A. Polášková, A. Krejčí, M. Petr, L. Navrátilová, O. Tichá, K. Nejedlỳ, et al., "p53 specifically binds triplex DNA in vitro and in cells", *PloS one* **2016**, *11*, e0167439.

[48] R. Besch, C. Giovannangeli, K. Degitz, "Triplex-forming oligonucleotides-sequence-specific DNA ligands as tools for gene inhibition and for modulation of DNA-associated functions", *Current drug targets* **2004**, *5*, 691–703.

[49] H. A. Day, P. Pavlou, Z. A. Waller, "i-Motif DNA: structure, stability and targeting with ligands", *Bioorganic & medicinal chemistry* **2014**, *22*, 4407–4418.

[50] I. Bang, "Untersuchungen über die Guanylsäure", *Biochemische Zeitschrift* **1910**, *26*, 293–311.

[51] M. Gellert, M. N. Lipsett, D. R. Davies, "Helix formation by guanylic acid", *Proceedings of the National Academy of Sciences* **1962**, *48*, 2013–2018.

[52] S. Arnott, R. Chandrasekaran, C. M. Marttila, "Structures for polyinosinic acid and polyguanylic acid", *Biochemical Journal* **1974**, *141*, 537–543.

[53] S. B. Zimmerman, G. H. Cohen, D. R. Davies, "X-ray fiber diffraction and model-building study of polyguanylic acid and polyinosinic acid", *Journal of molecular biology* **1975**, *92*, 181–192.

[54] E. Henderson, C. C. Hardin, S. K. Walk, I. Tinoco, E. H. Blackburn, "Telomeric DNA oligonucleotides form novel intramolecular structures containing guanine-guanine base pairs", *Cell* **1987**, *51*, 899–908.

[55] S. Burge, G. N. Parkinson, P. Hazel, A. K. Todd, S. Neidle, "Quadruplex DNA: sequence, topology and structure", *Nucleic acids research* **2006**, *34*, 5402–5415.

[56] A. T. Phan, "Human telomeric G-quadruplex: structures of DNA and RNA sequences", *The FEBS journal* **2010**, *277*, 1107–1117.

[57] J. Gu, J. Leszczynski, M. Bansal, "A new insight into the structure and stability of Hoogsteen hydrogen-bonded G-tetrad: an ab initio SCF study", *Chemical physics letters* **1999**, *311*, 209–214.

[58] Y. P. Yurenko, J. Novotnỳ, V. Sklenář, R. Marek, "Exploring non-covalent interactions in guanine-and xanthine-based model DNA quadruplex structures: a comprehensive quantum chemical approach", *Physical Chemistry Chemical Physics* **2014**, *16*, 2072–2084.

[59] K. Gkionis, H. Kruse, J. A. Platts, A. Mládek, J. Koča, J. Šponer, "Ion binding to quadruplex DNA stems. Comparison of MM and QM descriptions reveals sizable polarization effects not included in contemporary simulations", *Journal of chemical theory and computation* **2014**, *10*, 1326–1340.

[60] F. B. Howard, H. T. Miles, "Poly(inosinic acid) helixes: essential chelation of alkali metal ions in the axial channel", *Biochemistry* **1982**, *21*, 6736–6745.

[61] F. Zaccaria, G. Paragi, C. F. Guerra, "The role of alkali metal cations in the stabilization of guanine quadruplexes: why K+ is the best", *Physical Chemistry Chemical Physics* **2016**, *18*, 20895–20904.

[62] D. Bhattacharyya, G. M. Arachchilage, S. Basu, "Metal cations in G-quadruplex folding and stability", *Frontiers in chemistry* **2016**, *4*.

[63] S. Nagatoishi, Y. Tanaka, K. Tsumoto, "Circular dichroism spectra demonstrate formation of the thrombin-binding DNA aptamer G-quadruplex under stabilizing-cation-deficient conditions", *Biochemical and biophysical research communications* **2007**, *352*, 812–817.

[64] Z.-y. Kan, Y. Yao, P. Wang, X.-h. Li, Y.-h. Hao, Z. Tan, "Molecular crowding induces telomere G-quadruplex formation under salt-deficient conditions and enhances its competition with duplex formation", *Angewandte Chemie International Edition* **2006**, *45*, 1629–1632.

[65] C. Saintomé, S. Amrane, J.-L. Mergny, P. Alberti, "The exception that confirms the rule: a higher-order telomeric G-quadruplex structure more stable in sodium than in potassium", *Nucleic acids research* **2016**, *44*, 2926–2935.

[66] N. Zhang, A. Gorin, A. Majumdar, A. Kettani, N. Chernichenko, E. Skripkin, D. J. Patel, "V-shaped scaffold: a new architectural motif identified in an A-G-G-G-G pentad-containing dimeric DNA quadruplex involving stacked G(anti)-G (anti)-G (anti)-G (syn) tetrads", *Journal of molecular biology* **2001**, *311*, 1063–1079.

[67] R. Tippana, W. Xiao, S. Myong, "G-quadruplex conformation and dynamics are determined by loop length and sequence", *Nucleic acids research* **2014**, *42*, 8106–8114.

[68] A. Risitano, K. R. Fox, "Influence of loop size on the stability of intramolecular DNA quadruplexes", *Nucleic acids research* **2004**, *32*, 2598–2606.

[69] P. Hazel, J. Huppert, S. Balasubramanian, S. Neidle, "Loop-length-dependent folding of G-quadruplexes", *Journal of the American Chemical Society* **2004**, *126*, 16405–16415.

[70] A. Bugaut, S. Balasubramanian, "A sequence-independent study of the influence of short loop lengths on the stability and topology of intramolecular DNA G-quadruplexes", *Biochemistry* **2008**, *47*, 689–697.

[71] A. Guédin, J. Gros, P. Alberti, J.-L. Mergny, "How long is too long? Effects of loop size on G-quadruplex stability", *Nucleic acids research* **2010**, *38*, 7858–7868.

[72] I. Kejnovská, M. Vorlíčková, M. Brázdová, J. Sagi, "Stability of human telomere quadruplexes at high DNA concentrations", *Biopolymers* **2014**, *101*, 428–438.

[73] A. Arora, D. R. Nair, S. Maiti, "Effect of flanking bases on quadruplex stability and Watson-Crick duplex competition", *The FEBS journal* **2009**, *276*, 3628–3640.

[74] L. Petraccone, B. Pagano, C. Giancola, "Studying the effect of crowding and dehydration on DNA G-quadruplexes", *Methods* **2012**, *57*, 76–83.

[75] R. D. Gray, J. Li, J. B. Chaires, "Energetics and kinetics of a conformational switch in G-quadruplex DNA", *The Journal of Physical Chemistry B* **2009**, *113*, 2676–2683.

[76] L. Petraccone, A. Malafronte, J. Amato, C. Giancola, "G-quadruplexes from human telomeric DNA: how many conformations in PEG containing solutions?", *The Journal of Physical Chemistry B* **2012**, *116*, 2294–2305.

[77] A. N. Lane, J. B. Chaires, R. D. Gray, J. O. Trent, "Stability and kinetics of G-quadruplex structures", *Nucleic acids research* **2008**, *36*, 5482–5515.

[78] A. N. Lane, "The stability of intramolecular DNA G-quadruplexes compared with other macromolecules", *Biochimie* **2012**, *94*, 277–286.

[79] D. Miyoshi, H. Karimata, N. Sugimoto, "Hydration regulates thermodynamics of G-quadruplex formation under molecular crowding conditions", *Journal of the American Chemical Society* **2006**, *128*, 7957–7963.

[80] M. C. Miller, R. Buscaglia, J. B. Chaires, A. N. Lane, J. O. Trent, "Hydration is a major determinant of the G-quadruplex stability and conformation of the human telomere 3′ sequence of d(AG3(TTAG3)3)", *Journal of the American Chemical Society* **2010**, *132*, 17105–17107.

[81] J. B. Chaires, "Human telomeric G-quadruplex: thermodynamic and kinetic studies of telomeric quadruplex stability", *The FEBS journal* **2010**, *277*, 1098–1106.

[82] A. K. Todd, M. Johnston, S. Neidle, "Highly prevalent putative quadruplex sequence motifs in human DNA", *Nucleic acids research* **2005**, *33*, 2901–2907.

[83] J. L. Huppert, S. Balasubramanian, "Prevalence of quadruplexes in the human genome", *Nucleic acids research* **2005**, *33*, 2908–2916.

[84] K. W. Lim, Z. J. Khong, A. T. Phan, "Thermal stability of DNA quadruplex–duplex hybrids", *Biochemistry* **2013**, *53*, 247–257.

[85] A. Bedrat, L. Lacroix, J.-L. Mergny, "Re-evaluation of G-quadruplex propensity with G4Hunter", *Nucleic acids research* **2016**, *44*, 1746–1759.

[86] S. Neidle, "Quadruplex nucleic acids as novel therapeutic targets", *Journal of medicinal chemistry* **2016**, *59*, 5987–6011.

[87] V. S. Chambers, G. Marsico, J. M. Boutell, M. Di Antonio, G. P. Smith, S. Balasubramanian, "High-throughput sequencing of DNA G-quadruplex structures in the human genome", *Nature biotechnology* **2015**, *33*, 877–881.

[88] K. W. Lim, P. Jenjaroenpun, Z. J. Low, Z. J. Khong, Y. S. Ng, V. A. Kuznetsov, A. T. Phan, "Duplex stem-loop-containing quadruplex motifs in the human genome: a combined genomic and structural study", *Nucleic acids research* **2015**, *43*, 5630–5646.

[89] J. Eddy, N. Maizels, "Gene function correlates with potential for G4 DNA formation in the human genome", *Nucleic acids research* **2006**, *34*, 3887–3896.

[90] J. Eddy, N. Maizels, "Conserved elements with potential to form polymorphic G-quadruplex structures in the first intron of human genes", *Nucleic acids research* **2008**, *36*, 1321–1333.

[91] A. Verma, K. Halder, R. Halder, V. K. Yadav, P. Rawal, R. K. Thakur, F. Mohd, A. Sharma, S. Chowdhury, "Genome-wide computational and expression analyses reveal G-quadruplex DNA motifs as conserved cis-regulatory elements in human and related species", *Journal of medicinal chemistry* **2008**, *51*, 5641–5649.

[92] N. Maizels, L. T. Gray, "The G4 genome", *PLoS genetics* **2013**, *9*, e1003468.

[93] W.-C. Huang, T.-Y. Tseng, Y.-T. Chen, C.-C. Chang, Z.-F. Wang, C.-L. Wang, T.-N. Hsu, P.-T. Li, C.-T. Chen, J.-J. Lin, et al., "Direct evidence of mitochondrial G-quadruplex DNA by using fluorescent anti-cancer agents", *Nucleic acids research* **2015**, *43*, 10102–10113.

[94] D. Rhodes, H. J. Lipps, "G-quadruplexes and their regulatory roles in biology", *Nucleic acids research* **2015**, *43*, 8627–8637.

[95] Y. Qin, L. H. Hurley, "Structures, folding patterns, and functions of intramolecular DNA G-quadruplexes found in eukaryotic promoter regions", *Biochimie* **2008**, *90*, 1149–1171.

[96] P. A. Jones, J.-P. J. Issa, S. Baylin, "Targeting the cancer epigenome for therapy", *Nature reviews Genetics* **2016**, *17*, 630–641.

[97] T. A. Brooks, S. Kendrick, L. Hurley, "Making sense of G-quadruplex and i-motif functions in oncogene promoters", *The FEBS journal* **2010**, *277*, 3459–3469.

[98] L. Yuan, T. Tian, Y. Chen, S. Yan, X. Xing, Z. Zhang, Q. Zhai, L. Xu, S. Wang, X. Weng, et al., "Existence of G-quadruplex structures in promoter region of oncogenes confirmed by G-quadruplex DNA cross-linking strategy", *Scientific reports* **2013**, *3*, 1811.

[99] A. Ambrus, D. Chen, J. Dai, R. A. Jones, D. Yang, "Solution structure of the biologically relevant G-quadruplex element in the human c-MYC promoter. Implications for G-quadruplex stabilization", *Biochemistry* **2005**, *44*, 2048–2058.

[100] A. T. Phan, V. Kuryavyi, H. Y. Gaw, D. J. Patel, "Small-molecule interaction with a five-guanine-tract G-quadruplex structure from the human MYC promoter", *Nature chemical biology* **2005**, *1*, 167–173.

[101] R. I. Mathad, E. Hatzakis, J. Dai, D. Yang, "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 research* **2011**, *39*, 9023–9033.

[102] A. T. Phan, V. Kuryavyi, S. Burge, S. Neidle, D. J. Patel, "Structure of an unprecedented G-quadruplex scaffold in the human c-kit promoter", *Journal of the American Chemical Society* **2007**, *129*, 4386–4392.

[103] D. Wei, G. N. Parkinson, A. P. Reszka, S. Neidle, "Crystal structure of a c-kit promoter quadruplex reveals the structural role of metal ions and water molecules in maintaining loop conformation", *Nucleic acids research* **2012**, *40*, 4691–4700.

[104] D. Wei, J. Husby, S. Neidle, "Flexibility and structural conservation in a c-KIT G-quadruplex", *Nucleic acids research* **2014**, *43*, 629–644.

[105] S.-T. D. Hsu, P. Varnai, A. Bugaut, A. P. Reszka, S. Neidle, S. Balasubramanian, "A G-rich sequence within the c-kit oncogene promoter forms a parallel G-quadruplex having asymmetric G-tetrad dynamics", *Journal of the American Chemical Society* **2009**, *131*, 13399–13409.

[106] V. Kuryavyi, A. T. Phan, D. J. Patel, "Solution structures of all parallel-stranded monomeric and dimeric G-quadruplex scaffolds of the human c-kit2 promoter", *Nucleic acids research* **2010**, *38*, 6757–6773.

[107] J. Dai, D. Chen, R. A. Jones, L. H. Hurley, D. Yang, "NMR solution structure of the major G-quadruplex structure formed in the human BCL2 promoter region", *Nucleic acids research* **2006**, *34*, 5133–5144.

[108] A. Kerkour, J. Marquevielle, S. Ivashchenko, L. A. Yatsunyk, J.-L. Mergny, G. F. Salgado, "High-resolution three-dimensional NMR structure of the KRAS proto-oncogene promoter reveals key features of a G-quadruplex involved in transcriptional regulation", *Journal of Biological Chemistry* **2017**, *292*, 8082–8091.

[109] D. Wei, A. K. Todd, M. Zloh, M. Gunaratnam, G. N. Parkinson, S. Neidle, "Crystal structure of a promoter sequence in the B-raf gene reveals an intertwined dimer quadruplex", *Journal of the American Chemical Society* **2013**, *135*, 19319–19329.

[110] K. W. Lim, L. Lacroix, D. J. E. Yue, J. K. C. Lim, J. M. W. Lim, A. T. Phan, "Coexistence of two distinct G-quadruplex conformations in the hTERT promoter", *Journal of the American Chemical Society* **2010**, *132*, 12331–12342.

[111] P. Agrawal, E. Hatzakis, K. Guo, M. Carver, D. Yang, "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 research* **2013**, *41*, 10584–10592.

[112] M. Marušič, R. N. Veedu, J. Wengel, J. Plavec, "G-rich VEGF aptamer with locked and unlocked nucleic acid modifications exhibits a unique G-quadruplex fold", *Nucleic acids research* **2013**, *41*, 9524–9536.

[113] S. Balasubramanian, S. Neidle, "G-quadruplex nucleic acids as therapeutic targets", *Current opinion in chemical biology* **2009**, *13*, 345–353.

[114] V. González, L. H. Hurley, "The c-MYC NHE III1: function and regulation", *Annual review of pharmacology and toxicology* **2010**, *50*, 111–129.

[115] M. Z. Wojtukiewicz, E. Sierko, D. Hempel, S. C. Tucker, K. V. Honn, "Platelets and cancer angiogenesis nexus", *Cancer and Metastasis Reviews* **2017**, *36*, 249–262.

[116] V. Viglasky, T. Hianik, "Potential uses of G-quadruplex-forming aptamers", *General physiology and biophysics* **2013**, *32*, 149–172.

[117] R. F. Macaya, P. Schultze, F. W. Smith, J. A. Roe, J. Feigon, "Thrombin-binding DNA aptamer forms a unimolecular quadruplex structure in solution.", *Proceedings of the National Academy of Sciences* **1993**, *90*, 3745–3749.

[118] P. Schultze, R. F. Macaya, J. Feigon, "Three-dimensional solution structure of the thrombin-binding DNA aptamer d(GGTTGGTGTGGTTGG)", *Journal of molecular biology* **1994**, *235*, 1532–1547.

[119] V. M. Marathias, K. Y. Wang, S. Kumar, T. Q. Pham, S. Swaminathan, P. H. Bolton, "Determination of the number and location of the manganese binding sites of DNA quadruplexes in solution by EPR and NMR in the presence and absence of thrombin", *Journal of molecular biology* **1996**, *260*, 378–394.

[120] V. M. Marathias, P. H. Bolton, "Structures of the potassium-saturated, 2:1, and intermediate, 1:1, forms of a quadruplex DNA", *Nucleic acids research* **2000**, *28*, 1969–1977.

[121] X.-a. Mao, L. A. Marky, W. H. Gmeiner, "NMR structure of the thrombin-binding DNA aptamer stabilized by Sr2+", *Journal of Biomolecular Structure and Dynamics* **2004**, *22*, 25–33.

[122] L. Martino, A. Virno, A. Randazzo, A. Virgilio, V. Esposito, C. Giancola, M. Bucci, G. Cirino, L. Mayol, "A new modified thrombin binding aptamer containing a 5'–5' inversion of polarity site", *Nucleic acids research* **2006**, *34*, 6653–6662.

[123] C. Platella, C. Riccardi, D. Montesarchio, G. N. Roviello, D. Musumeci, "G-quadruplex-based aptamers against protein targets in therapy and diagnostics", *Biochimica et Biophysica Acta (BBA) - General Subjects* **2017**, *1861*, 1429–1447.

[124] J. Zhou, J. Rossi, "Aptamers as targeted therapeutics: Current potential and challenges", *Nature Reviews Drug Discovery* **2017**, *16*, 181–202.

[125] D. M. Tasset, M. F. Kubik, W. Steiner, "Oligonucleotide inhibitors of human thrombin that bind distinct epitopes", *Journal of molecular biology* **1997**, *272*, 688–698.

[126] I. Russo Krauss, A. Pica, A. Merlino, L. Mazzarella, F. Sica, "Duplex–quadruplex motifs in a peculiar structural organization cooperatively contribute to thrombin binding of a DNA aptamer", *Acta Crystallographica Section D: Biological Crystallography* **2013**, *69*, 2403–2411.

[127] W. Zhou, P.-J. J. Huang, J. Ding, J. Liu, "Aptamer-based biosensors for biomedical diagnostics", *Analyst* **2014**, *139*, 2627–2640.

[128] A. Kettani, A. R. Kumar, D. J. Patel, "Solution structure of a DNA quadruplex containing the fragile X syndrome triplet repeat", *Journal of molecular biology* **1995**, *254*, 638–656.

[129] S. Stefanovic, B. A. DeMarco, A. Underwood, K. R. Williams, G. J. Bassell, M. R. Mihailescu, "Fragile X mental retardation protein interactions with a G quadruplex structure in the 3'-untranslated region of NR2B mRNA", *Molecular bioSystems* **2015**, *11*, 3222–3230.

[130] R. Perrone, M. Nadai, I. Frasson, J. A. Poe, E. Butovskaya, T. E. Smithgall, M. Palumbo, G. Palù, S. N. Richter, "A dynamic G-quadruplex region regulates the HIV-1 long terminal repeat promoter", *Journal of medicinal chemistry* **2013**, *56*, 6521–6530.

[131] E. Tosoni, I. Frasson, M. Scalabrin, R. Perrone, E. Butovskaya, M. Nadai, G. Palù, D. Fabris, S. N. Richter, "Nucleolin stabilizes G-quadruplex structures folded by the LTR promoter and silences HIV-1 viral transcription", *Nucleic acids research* **2015**, *43*, 8884–8897.

[132] E. Y. N. Lam, D. Beraldi, D. Tannahill, S. Balasubramanian, "G-quadruplex structures are stable and detectable in human genomic DNA", *Nature communications* **2013**, *4*, 1796.

[133] G. Biffi, D. Tannahill, J. McCafferty, S. Balasubramanian, "Quantitative visualization of DNA G-quadruplex structures in human cells", *Nature chemistry* **2013**, *5*, 182–186.

[134] A. Henderson, Y. Wu, Y. C. Huang, E. A. Chavez, J. Platt, F. B. Johnson, R. M. Brosh, D. Sen, P. M. Lansdorp, "Detection of G-quadruplex DNA in mammalian cells", *Nucleic acids research* **2013**, *42*, 860–869.

[135] G. Biffi, D. Tannahill, J. Miller, W. J. Howat, S. Balasubramanian, "Elevated levels of G-quadruplex formation in human stomach and liver cancer tissues", *PloS one* **2014**, *9*, e102711.

[136] G. Fang, T. R. Cech, "The beta subunit of Oxytricha telomere-binding protein promotes G-quartet formation by telomeric DNA", *Cell* **1993**, *74*, 875–885.

[137] R. Hänsel-Hertsch, D. Beraldi, S. V. Lensing, G. Marsico, K. Zyner, A. Parry, M. Di Antonio, J. Pike, H. Kimura, M. Narita, et al., "G-quadruplex structures mark human regulatory chromatin", *Nature genetics* **2016**, *48*, 1267–1272.

[138] P. Murat, S. Balasubramanian, "Existence and consequences of G-quadruplex structures in DNA", *Current opinion in genetics & development* **2014**, *25*, 22–29.

[139] V. Brázda, L. Hároníková, J. C. Liao, M. Fojta, "DNA and RNA quadruplex-binding proteins", *International journal of molecular sciences* **2014**, *15*, 17493–17517.

[140] O. Mendoza, A. Bourdoncle, J.-B. Boulé, R. M. Brosh, J.-L. Mergny, "G-quadruplexes and helicases", *Nucleic acids research* **2016**, *44*, 1989–2006.

[141] C. G. Wu, M. Spies, "G-quadruplex recognition and remodeling by the FANCJ helicase", *Nucleic acids research* **2016**, *44*, 8742–8753.

[142] G. Biffi, M. Di Antonio, D. Tannahill, S. Balasubramanian, "Visualization and selective chemical targeting of RNA G-quadruplex structures in the cytoplasm of human cells", *Nature chemistry* **2014**, *6*, 75–80.

[143] A. Laguerre, K. Hukezalie, P. Winckler, F. Katranji, G. Chanteloup, M. Pirrotta, J.-M. Perrier-Cornet, J. M. Wong, D. Monchaud, "Visualization of RNA-quadruplexes in live cells", *Journal of the American Chemical Society* **2015**, *137*, 8521–8525.

[144] J. U. Guo, D. P. Bartel, "RNA G-quadruplexes are globally unfolded in eukaryotic cells and depleted in bacteria", *Science* **2016**, *353*, aaf5371.

[145] A. Cammas, S. Millevoi, "RNA G-quadruplexes: emerging mechanisms in disease", *Nucleic acids research* **2017**, *45*, 1584–1595.

[146] P. Agarwala, S. Pandey, S. Maiti, "The tale of RNA G-quadruplex", *Organic & biomolecular chemistry* **2015**, *13*, 5570–5585.

[147] B. Balk, A. Maicher, M. Dees, J. Klermund, S. Luke-Glaser, K. Bender, B. Luke, "Telomeric RNA-DNA hybrids affect telomere-length dynamics and senescence", *Nature structural & molecular biology* **2013**, *20*, 1199–1205.

[148] E. H. Blackburn, "Telomeres and telomerase: their mechanisms of action and the effects of altering their functions", *FEBS letters* **2005**, *579*, 859–862.

[149] J. W. Shay, "Role of telomeres and telomerase in aging and cancer", *Cancer discovery* **2016**, *6*, 584–593.

[150] E. Lazzerini-Denchi, A. Sfeir, "Stop pulling my strings-what telomeres taught us about the DNA damage response", *Nature reviews. Molecular cell biology* **2016**, *17*, 364–378.

[151] J. Maciejowski, T. de Lange, "Telomeres in cancer: tumour suppression and genome instability", *Nature Reviews Molecular Cell Biology* **2017**, *18*, 175–186.

[152] C. Wang, L. Zhao, S. Lu, "Role of TERRA in the regulation of telomere length", *International journal of biological sciences* **2015**, *11*, 316–323.

[153] K. Rippe, B. Luke, "TERRA and the state of the telomere", *Nature structural & molecular biology* **2015**, *22*, 853–858.

[154] J. W. Shay, W. E. Wright, "Telomeres and telomerase in normal and cancer stem cells", *FEBS letters* **2010**, *584*, 3819–3825.

[155] O. Samassekou, M. Gadji, R. Drouin, J. Yan, "Sizing the ends: normal length of human telomeres", *Annals of Anatomy-Anatomischer Anzeiger* **2010**, *192*, 284–291.

[156] N. W. Kim, M. A. Piatyszek, K. R. Prowse, C. B. Harley, M. D. West, P. L. Ho, G. M. Coviello, W. E. Wright, S. L. Weinrich, J. W. Shay, "Specific association of human telomerase activity with immortal cells and cancer", *Science* **1994**, *266*, 2011–2015.

[157] H. Vaziri, S. Benchimol, "Alternative pathways for the extension of cellular life span: inactivation of p53/pRb and expression of telomerase", *Oncogene* **1999**, *18*, 7676–7680.

[158] E. S. Hickman, M. C. Moroni, K. Helin, "The role of p53 and pRB in apoptosis and cancer", *Current opinion in genetics & development* **2002**, *12*, 60–66.

[159] Z. Feng, W. Hu, G. Rajagopal, A. J. Levine, "The tumor suppressor p53: cancer and aging", *Cell Cycle* **2008**, *7*, 842–847.

[160] K. C. Low, V. Tergaonkar, "Telomerase: central regulator of all of the hallmarks of cancer", *Trends in biochemical sciences* **2013**, *38*, 426–434.

[161] C. M. Counter, M. Meyerson, E. N. Eaton, L. W. Ellisen, S. D. Caddle, D. A. Haber, R. A. Weinberg, "Telomerase activity is restored in human cells by ectopic expression of hTERT (hEST2), the catalytic subunit of telomerase", *Oncogene* **1998**, *16*, 1217–1222.

[162] H. D. Wyatt, S. C. West, T. L. Beattie, "In TERT preting telomerase structure and function", *Nucleic acids research* **2010**, *38*, 5609–5622.

[163] M. Peifer, F. Hertwig, F. Roels, D. Dreidax, M. Gartlgruber, R. Menon, A. Krämer, J. L. Roncaioli, F. Sand, J. M. Heuckmann, et al., "Telomerase activation by genomic rearrangements in high-risk neuroblastoma", *Nature* **2015**, *526*, 700–704.

[164] A. J. Cesare, R. R. Reddel, "Alternative lengthening of telomeres: models, mechanisms and implications", *Nature reviews. Genetics* **2010**, *11*, 319–330.

[165] H. A. Pickett, R. R. Reddel, "Molecular mechanisms of activity and derepression of alternative lengthening of telomeres", *Nature structural & molecular biology* **2015**, *22*, 875–880.

[166] H.-Q. Yu, D. Miyoshi, N. Sugimoto, "Characterization of structure and stability of long telomeric DNA G-quadruplexes", *Journal of the American Chemical Society* **2006**, *128*, 15461–15468.

[167] A. Bugaut, P. Alberti, "Understanding the stability of DNA G-quadruplex units in long human telomeric strands", *Biochimie* **2015**, *113*, 125–133.

[168] A. J. Zaug, E. R. Podell, T. R. Cech, "Human POT1 disrupts telomeric G-quadruplexes allowing telomerase extension in vitro", *Proceedings of the National Academy of Sciences of the United States of America* **2005**, *102*, 10864–10869.

[169] A. L. Moye, K. C. Porter, S. B. Cohen, T. Phan, K. G. Zyner, N. Sasaki, G. O. Lovrecz, J. L. Beck, T. M. Bryan, "Telomeric G-quadruplexes are a substrate and site of localization for human telomerase", *Nature communications* **2015**, *6*, 7643.

[170] A. M. Zahler, J. R. Williamson, et al., "Inhibition of telomerase by G-quartet DNA structure", *Nature* **1991**, *350*, 718–720.

[171] W. C. Hahn, S. A. Stewart, M. W. Brooks, S. G. York, E. Eaton, A. Kurachi, R. L. Beijersbergen, J. H. Knoll, M. Meyerson, R. A. Weinberg, "Inhibition of telomerase limits the growth of human cancer cells.", *Nature medicine* **1999**, *5*, 1164–1170.

[172] C. B. Harley, "Telomerase and cancer therapeutics", *Nature reviews. Cancer* **2008**, *8*, 167–179.

[173] R. Rodriguez, K. M. Miller, J. V. Forment, C. R. Bradshaw, M. Nikan, S. Britton, T. Oelschlaegel, B. Xhemalce, S. Balasubramanian, S. P. Jackson, "Small-molecule-induced DNA damage identifies alternative DNA structures in human genes", *Nature chemical biology* **2012**, *8*, 301–310.

[174] C. Buseman, W. Wright, J. Shay, "Is telomerase a viable target in cancer?", *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis* **2012**, *730*, 90–97.

[175] S. Mocellin, K. A. Pooley, D. Nitti, "Telomerase and the search for the end of cancer", *Trends in molecular medicine* **2013**, *19*, 125–133.

[176] M. Ruden, N. Puri, "Novel anticancer therapeutics targeting telomerase", *Cancer treatment reviews* **2013**, *39*, 444–456.

[177] R. Hänsel, F. Löhr, S. Foldynová-Trantírková, E. Bamberg, L. Trantírek, V. Dötsch, "The parallel G-quadruplex structure of vertebrate telomeric repeat sequences is not the preferred folding topology under physiological conditions", *Nucleic acids research* **2011**, *39*, 5768–5775.

[178] Y. Wang, D. J. Patel, "Solution structure of the human telomeric repeat d[AG3(T2AG3)3] G-tetraplex", *Structure* **1993**, *1*, 263–282.

[179] K. W. Lim, S. Amrane, S. Bouaziz, W. Xu, Y. Mu, D. J. Patel, K. N. Luu, A. T. Phan, "Structure of the human telomere in K+ solution: a stable basket-type G-quadruplex with only two G-tetrad layers", *Journal of the American Chemical Society* **2009**, *131*, 4301–4309.

[180] Z. Zhang, J. Dai, E. Veliath, R. A. Jones, D. Yang, "Structure of a two-G-tetrad intramolecular G-quadruplex formed by a variant human telomeric sequence in K+ solution: insights into the interconversion of human telomeric G-quadruplex structures", *Nucleic acids research* **2009**, *38*, 1009–1021.

[181] K. W. Lim, V. C. M. Ng, N. Martín-Pintado, B. Heddi, A. T. Phan, "Structure of the human telomere in Na+ solution: an antiparallel (2+2) G-quadruplex scaffold reveals additional diversity", *Nucleic acids research* **2013**, *41*, 10556–10562.

[182] G. N. Parkinson, M. P. Lee, S. Neidle, "Crystal structure of parallel quadruplexes from human telomeric DNA", *Nature* **2002**, *417*, 876–880.

[183] B. Heddi, A. T. Phan, "Structure of human telomeric DNA in crowded solution", *Journal of the American Chemical Society* **2011**, *133*, 9824–9833.

[184] K. N. Luu, A. T. Phan, V. Kuryavyi, L. Lacroix, D. J. Patel, "Structure of the human telomere in K+ solution: an intramolecular (3+1) G-quadruplex scaffold", *Journal of the American Chemical Society* **2006**, *128*, 9963–9970.

[185] J. Dai, C. Punchihewa, A. Ambrus, D. Chen, R. A. Jones, D. Yang, "Structure of the intramolecular human telomeric G-quadruplex in potassium solution: a novel adenine triple formation", *Nucleic acids research* **2007**, *35*, 2440–2450.

[186] A. Matsugami, Y. Xu, Y. Noguchi, H. Sugiyama, M. Katahira, "Structure of a human telomeric DNA sequence stabilized by 8-bromoguanosine substitutions, as determined by NMR in a K+ solution", *The FEBS journal* **2007**, *274*, 3545–3556.

[187] A. T. Phan, V. Kuryavyi, K. N. Luu, D. J. Patel, "Structure of two intramolecular G-quadruplexes formed by natural human telomere sequences in K+ solution", *Nucleic acids research* **2007**, *35*, 6517–6525.

[188] V. V. Cheong, B. Heddi, C. J. Lech, A. T. Phan, "Xanthine and 8-oxoguanine in G-quadruplexes: formation of a G· G· X· O tetrad", *Nucleic acids research* **2015**, *43*, 10506–10514.

[189] J. Dickerhoff, L. Haase, W. Langel, K. Weisz, "Tracing Effects of Fluorine Substitutions on G-Quadruplex Conformational Changes", *ACS Chemical Biology* **2017**, *12*, 1308–1315.

[190] J. Dai, M. Carver, C. Punchihewa, R. A. Jones, D. Yang, "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 research* **2007**, *35*, 4927–4940.

[191] A. Ambrus, D. Chen, J. Dai, T. Bialis, R. A. Jones, D. Yang, "Human telomeric sequence forms a hybrid-type intramolecular G-quadruplex structure with mixed parallel/antiparallel strands in potassium solution", *Nucleic acids research* **2006**, *34*, 2723–2735.

[192] D. Renčiuk, I. Kejnovska, P. Školáková, K. Bednářová, J. Motlova, M. Vorlíčková, "Arrangements of human telomere DNA quadruplex in physiologically relevant K+ solutions", *Nucleic acids research* **2009**, *37*, 6625–6634.

[193] A. T. Phan, K. N. Luu, D. J. Patel, "Different loop arrangements of intramolecular human telomeric (3+1) G-quadruplexes in K+ solution", *Nucleic acids research* **2006**, *34*, 5715–5719.

[194] J. Li, J. J. Correia, L. Wang, J. O. Trent, J. B. Chaires, "Not so crystal clear: the structure of the human telomere G-quadruplex in solution differs from that present in a crystal", *Nucleic acids research* **2005**, *33*, 4649–4659.

[195] M. Bončina, G. Vesnaver, J. B. Chaires, J. Lah, "Unraveling the Thermodynamics of the Folding and Interconversion of Human Telomere G-Quadruplexes", *Angewandte Chemie* **2016**, *128*, 10496–10500.

[196] Y. Xu, Y. Noguchi, H. Sugiyama, "The new models of the human telomere d[AGGG-(TTAGGG)3] in K+ solution", *Bioorganic & medicinal chemistry* **2006**, *14*, 5584–5591.

[197] Y. Xue, Z.-y. Kan, Q. Wang, Y. Yao, J. Liu, Y.-h. Hao, Z. Tan, "Human telomeric DNA forms parallel-stranded intramolecular G-quadruplex in K+ solution under molecular crowding condition", *Journal of the American Chemical Society* **2007**, *129*, 11185–11191.

[198] R. Buscaglia, M. C. Miller, W. L. Dean, R. D. Gray, A. N. Lane, J. O. Trent, J. B. Chaires, "Polyethylene glycol binding alters human telomere G-quadruplex structure by conformational selection", *Nucleic acids research* **2013**, *41*, 7934–7946.

[199] T. I. Gaynutdinov, R. D. Neumann, I. G. Panyutin, "Structural polymorphism of intramolecular quadruplex of human telomeric DNA: effect of cations, quadruplex-binding drugs and flanking sequences", *Nucleic acids research* **2008**, *36*, 4079–4087.

[200] J. Sagi, "G-quadruplexes incorporating modified constituents: a review", *Journal of Biomolecular Structure and Dynamics* **2014**, *32*, 477–511.

[201] R. Hänsel, S. Foldynová-Trantírková, F. Löhr, J. Buck, E. Bongartz, E. Bamberg, H. Schwalbe, V. Dötsch, L. Trantírek, "Evaluation of parameters critical for observing nucleic acids inside living Xenopus laevis oocytes by in-cell NMR spectroscopy", *Journal of the American Chemical Society* **2009**, *131*, 15761–15768.

[202] R. Hänsel, F. Löhr, L. Trantírek, V. Dotsch, "High-resolution insight into G-overhang architecture", *Journal of the American Chemical Society* **2013**, *135*, 2816–2824.

[203] Y. Xu, T. Ishizuka, K. Kurabayashi, M. Komiyama, "Consecutive Formation of G-Quadruplexes in Human Telomeric-Overhang DNA: A Protective Capping Structure for Telomere Ends", *Angewandte Chemie International Edition* **2009**, *48*, 7833–7836.

[204] L. Petraccone in *Quadruplex Nucleic Acids*, Springer, **2012**, pp. 23–46.

[205] S. Haider, G. N. Parkinson, S. Neidle, "Molecular dynamics and principal components analysis of human telomeric quadruplex multimers", *Biophysical journal* **2008**, *95*, 296–311.

[206] I. Fotticchia, C. Giancola, L. Petraccone, "G-quadruplex unfolding in higher-order DNA structures", *Chemical Communications* **2013**, *49*, 9488–9490.

[207] T. Biver, "Stabilisation of non-canonical structures of nucleic acids by metal ions and small molecules", *Coordination Chemistry Reviews* **2013**, *257*, 2765–2783.

[208] B. Rosenberg, L. Van Camp, T. Krigas, "Inhibition of cell division in Escherichia coli by electrolysis products from a platinum electrode", *Nature* **1965**, *205*, 698–699.

[209] C. A. Rabik, M. E. Dolan, "Molecular mechanisms of resistance and toxicity associated with platinating agents", *Cancer treatment reviews* **2007**, *33*, 9–23.

[210] S. M. Attia, "Deleterious effects of reactive metabolites", *Oxidative medicine and cellular longevity* **2010**, *3*, 238–253.

[211] L. H. Hurley, "DNA and its associated processes as targets for cancer therapy", *Nature reviews. Cancer* **2002**, *2*, 188–200.

[212] S. Neidle, R. J. Harrison, A. P. Reszka, M. A. Read, "Structure-activity relationships among guanine-quadruplex telomerase inhibitors", *Pharmacology & therapeutics* **2000**, *85*, 133–139.

[213] D. Monchaud, M.-P. Teulade-Fichou, "A hitchhiker's guide to G-quadruplex ligands", *Organic & biomolecular chemistry* **2008**, *6*, 627–636.

[214] J.-H. Tan, L.-Q. Gu, J.-Y. Wu, "Design of selective G-quadruplex ligands as potential anticancer agents", *Mini reviews in medicinal chemistry* **2008**, *8*, 1163–1178.

[215] R. Haudecoeur, L. Stefan, F. Denat, D. Monchaud, "A model of smart G-quadruplex ligand", *Journal of the American Chemical Society* **2013**, *135*, 550–553.

[216] A. Paul, P. Sengupta, Y. Krishnan, S. Ladame, "Combining G-Quadruplex Targeting Motifs on a Single Peptide Nucleic Acid Scaffold: A Hybrid (3+1) PNA-DNA Bimolecular Quadruplex", *Chemistry–A European Journal* **2008**, *14*, 8682–8689.

[217] N. Zhang, T. Bing, X. Liu, C. Qi, L. Shen, L. Wang, D. Shangguan, "Cytotoxicity of guanine-based degradation products contributes to the antiproliferative activity of guanine-rich oligonucleotides", *Chemical Science* **2015**, *6*, 3831–3838.

[218] M. Ivancich, Z. Schrank, L. Wojdyla, B. Leviskas, A. Kuckovic, A. Sanjali, N. Puri, "Treating Cancer by Targeting Telomeres and Telomerase", *Antioxidants* **2017**, *6*, 15.

[219] S. Alcaro, A. Artese, G. Costa, S. Distinto, F. Ortuso, L. Parrotta, "Conformational studies and solvent-accessible surface area analysis of known selective DNA G-Quadruplex binders", *Biochimie* **2011**, *93*, 1267–1274.

[220] S. Alcaro, C. Musetti, S. Distinto, M. Casatti, G. Zagotto, A. Artese, L. Parrotta, F. Moraca, G. Costa, F. Ortuso, et al., "Identification and characterization of new DNA G-quadruplex binders selected by a combination of ligand and structure-based virtual screening approaches", *Journal of medicinal chemistry* **2013**, *56*, 843–855.

[221] S. Müller, K. Laxmi-Reddy, P. V. Jena, B. Baptiste, Z. Dong, F. Godde, T. Ha, R. Rodriguez, S. Balasubramanian, I. Huc, "Targeting DNA G-Quadruplexes with Helical Small Molecules", *ChemBioChem* **2014**, *15*, 2563–2570.

[222] A. M. Burger, F. Dai, C. M. Schultes, A. P. Reszka, M. J. Moore, J. A. Double, S. Neidle, "The G-quadruplex-interactive molecule BRACO-19 inhibits tumor growth, consistent with telomere targeting and interference with telomerase function", *Cancer Research* **2005**, *65*, 1489–1496.

[223] L. Rossetti, M. Franceschin, S. Schirripa, A. Bianco, G. Ortaggi, M. Savino, "Selective interactions of perylene derivatives having different side chains with inter- and intramolecular G-quadruplex DNA structures. A correlation with telomerase inhibition", *Bioorganic & medicinal chemistry letters* **2005**, *15*, 413–420.

[224] S. Cosconati, L. Marinelli, R. Trotta, A. Virno, L. Mayol, E. Novellino, A. J. Olson, A. Randazzo, "Tandem application of virtual screening and NMR experiments in the discovery of brand new DNA quadruplex groove binders", *Journal of the American Chemical Society* **2009**, *131*, 16336–16337.

[225] S. Cosconati, L. Marinelli, R. Trotta, A. Virno, S. De Tito, R. Romagnoli, B. Pagano, V. Limongelli, C. Giancola, P. G. Baraldi, et al., "Structural and conformational requisites in DNA quadruplex groove binding: another piece to the puzzle", *Journal of the American Chemical Society* **2010**, *132*, 6425–6433.

[226] F. Hamon, E. Largy, A. Guédin-Beaurepaire, M. Rouchon-Dagois, A. Sidibe, D. Monchaud, J.-L. Mergny, J.-F. Riou, C.-H. Nguyen, M.-P. Teulade-Fichou, "An Acyclic Oligoheteroaryle That Discriminates Strongly between Diverse G-Quadruplex Topologies", *Angewandte Chemie International Edition* **2011**, *50*, 8745–8749.

[227] A. De Rache, J.-L. Mergny, "Assessment of selectivity of G-quadruplex ligands via an optimised FRET melting assay", *Biochimie* **2015**, *115*, 194–202.

[228] M. Porru, S. Artuso, E. Salvati, A. Bianco, M. Franceschin, M. G. Diodoro, D. Passeri, A. Orlandi, F. Savorani, M. D'Incalci, et al., "Targeting G-quadruplex DNA structures by EMICORON has a strong antitumor efficacy against advanced models of human colon cancer", *Molecular cancer therapeutics* **2015**, *14*, 2541–2551.

[229] A. Cummaro, I. Fotticchia, M. Franceschin, C. Giancola, L. Petraccone, "Binding properties of human telomeric quadruplex multimers: a new route for drug design", *Biochimie* **2011**, *93*, 1392–1400.

[230] L.-P. Bai, M. Hagihara, Z.-H. Jiang, K. Nakatani, "Ligand binding to tandem G quadruplexes from human telomeric DNA", *ChemBioChem* **2008**, *9*, 2583–2587.

[231] K.-i. Shinohara, Y. Sannohe, S. Kaieda, K.-i. Tanaka, H. Osuga, H. Tahara, Y. Xu, T. Kawase, T. Bando, H. Sugiyama, "A chiral wedge molecule inhibits telomerase activity", *Journal of the American Chemical Society* **2010**, *132*, 3778–3782.

[232] C. Zhao, L. Wu, J. Ren, Y. Xu, X. Qu, "Targeting human telomeric higher-order DNA: dimeric G-quadruplex units serve as preferred binding site", *Journal of the American Chemical Society* **2013**, *135*, 18786–18789.

[233] A. R. Cousins, D. Ritson, P. Sharma, M. F. Stevens, J. E. Moses, M. S. Searle, "Ligand selectivity in stabilising tandem parallel folded G-quadruplex motifs in human telomeric DNA sequences", *Chemical Communications* **2014**, *50*, 15202–15205.

[234] X.-X. Huang, L.-N. Zhu, B. Wu, Y.-F. Huo, N.-N. Duan, D.-M. Kong, "Two cationic porphyrin isomers showing different multimeric G-quadruplex recognition specificity against monomeric G-quadruplexes", *Nucleic acids research* **2014**, *42*, 8719–8731.

[235] C.-Q. Zhou, T.-C. Liao, Z.-Q. Li, J. Gonzalez-Garcia, M. Reynolds, M. Zou, R. Vilar, "Dinickel-Salphen Complexes as Binders of Human Telomeric Dimeric G-Quadruplexes", *Chemistry–A European Journal* **2017**, *23*, 4713–4722.

[236] S. Neidle, "Human telomeric G-quadruplex: The current status of telomeric G-quadruplexes as therapeutic targets in human cancer", *The FEBS journal* **2010**, *277*, 1118–1125.

[237] S. A. Ohnmacht, S. Neidle, "Small-molecule quadruplex-targeted drug discovery", *Bioorganic & medicinal chemistry letters* **2014**, *24*, 2602–2612.

[238] E. Largy, A. Granzhan, F. Hamon, D. Verga, M.-P. Teulade-Fichou in *Quadruplex Nucleic Acids*, Springer, **2012**, pp. 111–177.

[239] B. R. Vummidi, J. Alzeer, N. W. Luedtke, "Fluorescent probes for G-quadruplex structures", *ChemBioChem* **2013**, *14*, 540–558.

[240] A. C. Bhasikuttan, J. Mohanty, "Targeting G-quadruplex structures with extrinsic fluorogenic dyes: promising fluorescence sensors", *Chemical Communications* **2015**, *51*, 7581–7597.

[241] S. N. Georgiades, N. H. Abd Karim, K. Suntharalingam, R. Vilar, "Interaction of metal complexes with G-quadruplex DNA", *Angewandte Chemie International Edition* **2010**, *49*, 4020–4034.

[242] S. F Ralph, "Quadruplex DNA: a promising drug target for the medicinal inorganic chemist", *Current topics in medicinal chemistry* **2011**, *11*, 572–590.

[243] M. Wang, Z. Mao, T.-S. Kang, C.-Y. Wong, J.-L. Mergny, C.-H. Leung, D.-L. Ma, "Conjugating a groove-binding motif to an Ir(III) complex for the enhancement of G-quadruplex probe behavior", *Chemical Science* **2016**, *7*, 2516–2523.

[244] B. Shen, "A new golden age of natural products drug discovery", *Cell* **2015**, *163*, 1297–1300.

[245] A. L. Harvey, R. Edrada-Ebel, R. J. Quinn, "The re-emergence of natural products for drug discovery in the genomics era", *Nature reviews. Drug discovery* **2015**, *14*, 111–129.

[246] D. J. Newman, G. M. Cragg, "Natural products as sources of new drugs from 1981 to 2014", *Journal of Natural Products* **2016**, *79*, 629–661.

[247] T. Efferth, P. C. Li, V. S. B. Konkimalla, B. Kaina, "From traditional Chinese medicine to rational cancer therapy", *Trends in molecular medicine* **2007**, *13*, 353–361.

[248] J. M. Hagel, P. J. Facchini, "Benzylisoquinoline alkaloid metabolism: a century of discovery and a brave new world", *Plant and Cell Physiology* **2013**, *54*, 647–672.

[249] M. Iranshahy, R. Quinn, M. Iranshahi, "Biologically active isoquinoline alkaloids with drug-like properties from the genus Corydalis", *RSC Advances* **2014**, *4*, 15900–15913.

[250] M. Tillhon, L. M. G. Ortiz, P. Lombardi, A. I. Scovassi, "Berberine: new perspectives for old remedies", *Biochemical pharmacology* **2012**, *84*, 1260–1267.

[251] A. Kumar, K. Chopra, M. Mukherjee, R. Pottabathini, D. K. Dhull, et al., "Current knowledge and pharmacological profile of berberine: an update", *European journal of pharmacology* **2015**, *761*, 288–297.

[252] S. Kulkarni, A. Dhir, "Berberine: a plant alkaloid with therapeutic potential for central nervous system disorders", *Phytotherapy Research* **2010**, *24*, 317–324.

[253] T. Ahmed, M. Abdollahi, M. Daglia, S. F. Nabavi, S. M. Nabavi, et al., "Berberine and neurodegeneration: A review of literature", *Pharmacological Reports* **2015**, *67*, 970–979.

[254] W.-J. Kong, Y.-L. Zhao, X.-H. Xiao, Z.-L. Li, C. Jin, H.-B. Li, "Investigation of the antifungal activity of coptisine on Candida albicans growth by microcalorimetry combined with principal component analysis", *Journal of applied microbiology* **2009**, *107*, 1072–1080.

[255] Z. Zhang, A. Deng, J. Yu, Z. Li, L. Wu, W. Wang, H. Qin, "Advance in studies on pharmacological activity of coptisine hydrochloride", *Zhongguo Zhong yao za zhi= Zhongguo zhongyao zazhi= China journal of Chinese materia medica* **2013**, *38*, 2750–2754.

[256] J. Guo, S.-B. Wang, T.-Y. Yuan, Y.-J. Wu, Y. Yan, L. Li, X.-N. Xu, L.-l. Gong, H.-l. Qin, L.-H. Fang, et al., "Coptisine protects rat heart against myocardial ischemia/reperfusion injury by suppressing myocardial apoptosis and inflammation", *Atherosclerosis* **2013**, *231*, 384–391.

[257] T. Friedemann, U. Schumacher, Y. Tao, A. K.-M. Leung, S. Schröder, "Neuroprotective activity of coptisine from Coptis chinensis", *Evidence-Based Complementary and Alternative Medicine* **2015**, *2015*.

[258] J. Herbert, J. Augereau, J. Gleye, J. Maffrand, "Chelerythrine is a potent and specific inhibitor of protein kinase C", *Biochemical and biophysical research communications* **1990**, *172*, 993–999.

[259] P. Kosina, D. Walterová, J. Ulrichová, V. Lichnovskỳ, M. Stiborová, H. Rỳdlová, J. Vičar, V. Krečman, M. J. Brabec, V. Šimánek, "Sanguinarine and chelerythrine: assessment of safety on pigs in ninety days feeding experiment", *Food and Chemical Toxicology* **2004**, *42*, 85–91.

[260] V. Dvorak, Zdenek Kuban, J. Klejdus, Borivoj Hlavac, J. Vicar, Jaroslav Ulrichova, V. Simanek, "Quaternary Benzo[c]phenanthridines Sanguinarine and Chelerythrine: A Review of Investigations from Chemical and Biological Studies", *Heterocycles* **2006**, *68*, 2403–2422.

[261] X. Niu, T. Fan, W. Li, W. Xing, H. Huang, "The anti-inflammatory effects of sanguinarine and its modulation of inflammatory mediators from peritoneal macrophages", *European journal of pharmacology* **2012**, *689*, 262–269.

[262] S. J. Chmura, M. E. Dolan, A. Cha, H. J. Mauceri, D. W. Kufe, R. R. Weichselbaum, "In vitro and in vivo activity of protein kinase C inhibitor chelerythrine chloride induces tumor cell toxicity and growth delay in vivo", *Clinical Cancer Research* **2000**, *6*, 737–742.

[263] C.-C. Lin, L. T. Ng, F.-F. Hsu, D.-E. Shieh, L.-C. Chiang, "Cytotoxic effects of Coptis chinensis and Epimedium sagittatum extracts and their major constituents (berberine, coptisine and icariin) on hepatoma and leukaemia cell growth", *Clinical and Experimental Pharmacology and Physiology* **2004**, *31*, 65–69.

[264] J. Malikova, A. Zdařilová, A. Hlobilkova, J. Ulrichova, "The effect of chelerythrine on cell growth, apoptosis, and cell cycle in human normal and cancer cells in comparison with sanguinarine", *Cell biology and toxicology* **2006**, *22*, 439–453.

[265] I. De Stefano, G. Raspaglio, G. F. Zannoni, D. Travaglia, M. G. Prisco, M. Mosca, C. Ferlini, G. Scambia, D. Gallo, "Antiproliferative and antiangiogenic effects of the benzophenanthridine alkaloid sanguinarine in melanoma", *Biochemical pharmacology* **2009**, *78*, 1374–1381.

[266] Y. Sun, K. Xun, Y. Wang, X. Chen, "A systematic review of the anticancer properties of berberine, a natural product from Chinese herbs", *Anti-cancer drugs* **2009**, *20*, 757–769.

[267] L. M. Guamán Ortiz, P. Lombardi, M. Tillhon, A. I. Scovassi, "Berberine, an epiphany against cancer", *Molecules* **2014**, *19*, 12349–12367.

[268] I. Slaninová, K. Pěnčıǩová, J. Urbanová, J. Slanina, E. Táborská, "Antitumour activities of sanguinarine and related alkaloids", *Phytochemistry reviews* **2014**, *13*, 51–68.

[269] R. Gaziano, G. Moroni, C. Buè, M. T. Miele, P. Sinibaldi-Vallebona, F. Pica, "Antitumor effects of the benzophenanthridine alkaloid sanguinarine: Evidence and perspectives", *World journal of gastrointestinal oncology* **2016**, *8*, 30–39.

[270] T.-K. Li, E. Bathory, E. J. LaVoie, A. Srinivasan, W. K. Olson, R. R. Sauers, L. F. Liu, D. S. Pilch, "Human Topoisomerase I Poisoning by Protoberberines: Potential Roles for Both Drug-DNA and Drug-Enzyme Interactions", *Biochemistry* **2000**, *39*, 7107–7116.

[271] J. Holy, G. Lamont, E. Perkins, "Disruption of nucleocytoplasmic trafficking of cyclin D1 and topoisomerase II by sanguinarine", *BMC cell biology* **2006**, *7*, 13.

[272] S. S. Matkar, L. A. Wrischnik, U. Hellmann-Blumberg, "Sanguinarine causes DNA damage and p53-independent cell death in human colon cancer cell lines", *Chemico-biological interactions* **2008**, *172*, 63–71.

[273] G. S. Kumar, S. Hazra, "Sanguinarine, a promising anticancer therapeutic: photochemical and nucleic acid binding properties", *Rsc Advances* **2014**, *4*, 56518–56531.

[274] K.-H. Lee, H.-L. Lo, W.-C. Tang, H. H.-y. Hsiao, P.-M. Yang, "A gene expression signature-based approach reveals the mechanisms of action of the Chinese herbal medicine berberine", *Scientific reports* **2014**, *4*, 6394.

[275] Y. Wang, M. M. Kheir, Y. Chai, J. Hu, D. Xing, F. Lei, L. Du, "Comprehensive study in the inhibitory effect of berberine on gene transcription, including TATA box", *PLoS One* **2011**, *6*, e23495.

[276] A. Arora, C. Balasubramanian, N. Kumar, S. Agrawal, R. P. Ojha, S. Maiti, "Binding of berberine to human telomeric quadruplex-spectroscopic, calorimetric and molecular modeling studies", *The FEBS journal* **2008**, *275*, 3971–3983.

[277] K. Bhadra, G. S. Kumar, "Therapeutic potential of nucleic acid-binding isoquinoline alkaloids: Binding aspects and implications for drug design", *Medicinal research reviews* **2011**, *31*, 821–862.

[278] K. Bhadra, G. S. Kumar, "Interaction of berberine, palmatine, coralyne, and sanguinarine to quadruplex DNA: a comparative spectroscopic and calorimetric study", *Biochimica et Biophysica Acta (BBA) - General Subjects* **2011**, *1810*, 485–496.

[279] X. Ji, H. Sun, H. Zhou, J. Xiang, Y. Tang, C. Zhao, "The interaction of telomeric DNA and C-myc22 G-quadruplex with 11 natural alkaloids", *Nucleic acid therapeutics* **2012**, *22*, 127–136.

[280] S. Ghosh, S. K. Pradhan, A. Kar, S. Chowdhury, D. Dasgupta, "Molecular basis of recognition of quadruplexes human telomere and c-myc promoter by the putative anticancer agent sanguinarine", *Biochimica et Biophysica Acta (BBA) - General Subjects* **2013**, *1830*, 4189–4201.

[281] L.-P. Bai, M. Hagihara, K. Nakatani, Z.-H. Jiang, "Recognition of chelerythrine to human telomeric DNA and RNA G-quadruplexes", *Scientific reports* **2014**, *4*, 6767.

[282] S. Ghosh, J. Jana, R. K. Kar, S. Chatterjee, D. Dasgupta, "Plant Alkaloid Chelerythrine Induced Aggregation of Human Telomere Sequence - A Unique Mode of Association between a Small Molecule and a Quadruplex", *Biochemistry* **2015**, *54*, 974–986.

[283] L.-N. Wen, M.-X. Xie, "Spectroscopic investigation of the interaction between G-quadruplex of KRAS promoter sequence and three isoquinoline alkaloids", *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy* **2017**, *171*, 287–296.

[284] S. Yang, J. Xiang, Q. Yang, Q. Zhou, X. Zhang, Q. Li, Y. Tang, G. Xu, "An important functional group, benzo[1,3]dioxole, of alkaloids induces the formation of the human telomeric DNA G-quadruplex", *Chinese Science Bulletin* **2011**, *56*, 613–617.

[285] L. Zhang, H. Liu, Y. Shao, C. Lin, H. Jia, G. Chen, D. Yang, Y. Wang, "Selective lighting up of epiberberine alkaloid fluorescence by fluorophore-switching aptamer and stoichiometric targeting of human telomeric DNA G-Quadruplex multimer", *Analytical chemistry* **2014**, *87*, 730–737.

[286] S. K. Noureini, H. Esmaeili, F. Abachi, S. Khiali, B. Islam, M. Kuta, A. A. Saboury, M. Hoffmann, J. Sponer, G. Parkinson, et al., "Selectivity of major isoquinoline alkaloids from Chelidonium majus towards telomeric G-quadruplex: A study using a transition-FRET (t-FRET) assay", *Biochimica et Biophysica Acta (BBA) - General Subjects* **2017**, *1861*, 2020–2030.

[287] I. Bessi, C. Bazzicalupi, C. Richter, H. R. Jonker, K. Saxena, C. Sissi, M. Chioccioli, S. Bianco, A. R. Bilia, H. Schwalbe, et al., "Spectroscopic, molecular modeling, and NMR-spectroscopic investigation of the binding mode of the natural alkaloids berberine and sanguinarine to human telomeric G-quadruplex DNA", *ACS chemical biology* **2012**, *7*, 1109–1119.

[288] C. Bazzicalupi, M. Ferraroni, A. R. Bilia, F. Scheggi, P. Gratteri, "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 research* **2013**, *41*, 632–638.

[289] H. L. Wu, C. Y. Hsu, W. H. Liu, B. Y. Yung, "Berberine-induced apoptosis of human leukemia HL-60 cells is associated with down-regulation of nucleophosmin/B23 and telomerase activity", *International journal of cancer* **1999**, *81*, 923–929.

[290] J. Li, L. Gu, H. Zhang, T. Liu, D. Tian, M. Zhou, S. Zhou, "Berberine represses DAXX gene transcription and induces cancer cell apoptosis", *Laboratory investigation; a journal of technical methods and pathology* **2013**, *93*, 354–364.

[291] V. M. Kumarasamy, Y.-J. Shin, J. White, D. Sun, "Selective repression of RET proto-oncogene in medullary thyroid carcinoma by a natural alkaloid berberine", *BMC cancer* **2015**, *15*, 599.

[292] J. Jana, S. Mondal, P. Bhattacharjee, P. Sengupta, T. Roychowdhury, P. Saha, P. Kundu, S. Chatterjee, "Chelerythrine down regulates expression of VEGFA, BCL2 and KRAS by arresting G-quadruplex structures at their promoter regions", *Scientific reports* **2017**, *7*, 40706.

[293] D. Bhowmik, G. Suresh Kumar, "Recent Advances in Nucleic Acid Binding Aspects of Berberine Analogs and Implications for Drug Design", *Mini reviews in medicinal chemistry* **2016**, *16*, 104–109.

[294] M. Tera, T. Hirokawa, S. Okabe, K. Sugahara, H. Seimiya, K. Shimamoto, "Design and Synthesis of a Berberine Dimer: A Fluorescent Ligand with High Affinity towards G-Quadruplexes", *Chemistry–A European Journal* **2015**, *21*, 14519–14528.

[295] M. Franceschin, L. Rossetti, A. D'Ambrosio, S. Schirripa, A. Bianco, G. Ortaggi, M. Savino, C. Schultes, S. Neidle, "Natural and synthetic G-quadruplex interactive berberine derivatives", *Bioorganic & Medicinal Chemistry Letters* **2006**, *16*, 1707–1711.

[296] Y. Ma, T.-M. Ou, J.-H. Tan, J.-Q. Hou, S.-L. Huang, L.-Q. Gu, Z.-S. Huang, "Synthesis and evaluation of 9-O-substituted berberine derivatives containing aza-aromatic terminal group as highly selective telomeric G-quadruplex stabilizing ligands", *Bioorganic & medicinal chemistry letters* **2009**, *19*, 3414–3417.

[297] W.-J. Zhang, T.-M. Ou, Y.-J. Lu, Y.-Y. Huang, W.-B. Wu, Z.-S. Huang, J.-L. Zhou, K.-Y. Wong, L.-Q. Gu, "9-Substituted berberine derivatives as G-quadruplex stabilizing ligands in telomeric DNA", *Bioorganic & medicinal chemistry* **2007**, *15*, 5493–5501.

[298] K. C. Gornall, S. Samosorn, B. Tanwirat, A. Suksamrarn, J. B. Bremner, M. J. Kelso, J. L. Beck, "A mass spectrometric investigation of novel quadruplex DNA-selective berberine derivatives", *Chemical Communications* **2010**, *46*, 6602–6604.

[299] Y.-X. Xiong, H.-F. Su, P. Lv, Y. Ma, S.-K. Wang, H. Miao, H.-Y. Liu, J.-H. Tan, T.-M. Ou, L.-Q. Gu, et al., "A newly identified berberine derivative induces cancer cell senescence by stabilizing endogenous G-quadruplexes and sparking a DNA damage response at the telomere region", *Oncotarget* **2015**, *6*, 35625–35635.

[300] P. Lombardi, F. Buzzetti, A. G. Arcamone, Benzoquinolizinium salt derivatives as anticancer agents, US Patent 8, 188, 109, **2012**.

[301] E. Pierpaoli, A. G. Arcamone, F. Buzzetti, P. Lombardi, C. Salvatore, M. Provinciali, "Antitumor effect of novel berberine derivatives in breast cancer cells", *Biofactors* **2013**, *39*, 672–679.

[302] K. F. Albring, J. Weidemüller, S. Mittag, J. Weiske, K. Friedrich, M. C. Geroni, P. Lombardi, O. Huber, "Berberine acts as a natural inhibitor of Wnt/beta-catenin signaling-Identification of more active 13-arylalkyl derivatives", *Biofactors* **2013**, *39*, 652–662.

[303] L. M. Guamán Ortiz, M. Tillhon, M. Parks, I. Dutto, E. Prosperi, M. Savio, A. G. Arcamone, F. Buzzetti, P. Lombardi, A. I. Scovassi, "Multiple effects of berberine derivatives on colon cancer cells", *BioMed research international* **2014**, *2014*.

[304] L. M. Guamán Ortiz, A. L. Croce, F. Aredia, S. Sapienza, G. Fiorillo, T. M. Syeda, F. Buzzetti, P. Lombardi, A. I. Scovassi, "Effect of new berberine derivatives on colon cancer cells", *Acta biochimica et biophysica Sinica* **2015**, *47*, 824–833.

[305] E. Pierpaoli, E. Damiani, F. Orlando, G. Lucarini, B. Bartozzi, P. Lombardi, C. Salvatore, C. Geroni, A. Donati, M. Provinciali, "Antiangiogenic and antitumor activities of berberine derivative NAX014 compound in a transgenic murine model of HER2/neu-positive mammary carcinoma", *Carcinogenesis* **2015**, *36*, 1169–1179.

[306] S. Chatterjee, S. Mallick, F. Buzzetti, G. Fiorillo, T. M. Syeda, P. Lombardi, K. D. Saha, G. S. Kumar, "New 13-pyridinealkyl berberine analogues intercalate to DNA and induce apoptosis in HepG2 and MCF-7 cells through ROS mediated p53 dependent pathway: biophysical, biochemical and molecular modeling studies", *RSC Advances* **2015**, *5*, 90632–90644.

[307] S. Vieira, S. Castelli, M. Falconi, J. Takarada, G. Fiorillo, F. Buzzetti, P. Lombardi, A. Desideri, "Role of 13-(di)phenylalkyl berberine derivatives in the modulation of the activity of human topoisomerase IB", *International journal of biological macromolecules* **2015**, *77*, 68–75.

[308] I. Ott, "On the medicinal chemistry of gold complexes as anticancer drugs", *Coordination Chemistry Reviews* **2009**, *253*, 1670–1681.

[309] B. Bertrand, A. Casini, "A golden future in medicinal inorganic chemistry: the promise of anticancer gold organometallic compounds", *Dalton Transactions* **2014**, *43*, 4209–4219.

[310] T. Zou, C. T. Lum, C.-N. Lok, J.-J. Zhang, C.-M. Che, "Chemical biology of anticancer gold(III) and gold(I) complexes", *Chemical Society Reviews* **2015**, *44*, 8786–8801.

[311] J. C. Lin, R. T. Huang, C. S. Lee, A. Bhattacharyya, W. S. Hwang, I. J. Lin, "Coinage metal-N-heterocyclic carbene complexes", *Chemical reviews* **2009**, *109*, 3561–3598.

[312] L. Oehninger, R. Rubbiani, I. Ott, "N-Heterocyclic carbene metal complexes in medicinal chemistry", *Dalton Transactions* **2013**, *42*, 3269–3284.

[313] W. Liu, R. Gust, "Update on metal N-heterocyclic carbene complexes as potential antitumor metallodrugs", *Coordination Chemistry Reviews* **2016**, *329*, 191–213.

[314] L. Stefan, B. Bertrand, P. Richard, P. Le Gendre, F. Denat, M. Picquet, D. Monchaud, "Assessing the differential affinity of small molecules for noncanonical DNA structures", *ChemBioChem* **2012**, *13*, 1905–1912.

[315] B. Bertrand, L. Stefan, M. Pirrotta, D. Monchaud, E. Bodio, P. Richard, P. Le Gendre, E. Warmerdam, M. H. de Jager, G. M. Groothuis, et al., "Caffeine-based gold(I) N-heterocyclic carbenes as possible anticancer agents: synthesis and biological properties", *Inorganic chemistry* **2014**, *53*, 2296–2303.

[316] A. Casini, N. Estrada-Ortiz, C. Gabbiani, F. Guarra, G. Groothuis, I. de Graaf, M. de Jager, L. Marchetti, "Anticancer gold N-heterocyclic carbene complexes: a comparative in vitro and ex vivo study", *ChemMedChem* **2017**, *12*, 1429–1435.

[317] L. Messori, L. Marchetti, L. Massai, F. Scaletti, A. Guerri, I. Landini, S. Nobili, G. Perrone, E. Mini, P. Leoni, et al., "Chemistry and biology of two novel gold(I) carbene complexes as prospective anticancer agents", *Inorganic chemistry* **2014**, *53*, 2396–2403.

[318] S. Nobili, E. Mini, I. Landini, C. Gabbiani, A. Casini, L. Messori, "Gold compounds as anticancer agents: chemistry, cellular pharmacology, and preclinical studies", *Medicinal research reviews* **2010**, *30*, 550–580.

[319] C. Gabbiani, M. A. Cinellu, L. Maiore, L. Massai, F. Scaletti, L. Messori, "Chemistry and biology of three representative gold(III) compounds as prospective anticancer agents", *Inorganica Chimica Acta* **2012**, *393*, 115–124.

[320] A. Casini, M. A. Cinellu, G. Minghetti, C. Gabbiani, M. Coronnello, E. Mini, L. Messori, "Structural and solution chemistry, antiproliferative effects, and DNA and protein binding properties of a series of dinuclear gold(III) compounds with bipyridyl ligands", *Journal of medicinal chemistry* **2006**, *49*, 5524–5531.

[321] C. Gabbiani, F. Scaletti, L. Massai, E. Michelucci, M. A. Cinellu, L. Messori, "Medicinal gold compounds form tight adducts with the copper chaperone Atox-1: biological and pharmacological implications", *Chemical Communications* **2012**, *48*, 11623–11625.

[322] P. Gratteri, L. Massai, E. Michelucci, R. Rigo, L. Messori, M. A. Cinellu, C. Musetti, C. Sissi, C. Bazzicalupi, "Interactions of selected gold(III) complexes with DNA G quadruplexes", *Dalton Transactions* **2015**, *44*, 3633–3639.

[323] J. Drenth, *Principles of protein X-ray crystallography*, Springer Science & Business Media, **2007**.

[324] I. Russo Krauss, A. Merlino, A. Vergara, F. Sica, "An overview of biological macromolecule crystallization", *International journal of molecular sciences* **2013**, *14*, 11643–11691.

[325] R. Henderson, "Cryo-protection of protein crystals against radiation damage in electron and X-ray diffraction", *Proceedings: Biological Sciences* **1990**, *241*, 6–8.

[326] C. Giacovazzo, *Fundamentals of crystallography*, Oxford university press, USA, **2002**.

[327] J. R. Deschamps, "X-ray crystallography of chemical compounds", *Life sciences* **2010**, *86*, 585–589.

[328] W. A. Hendrickson, "Anomalous diffraction in crystallographic phase evaluation", *Quarterly reviews of biophysics* **2014**, *47*, 49–93.

[329] A. J. Cohen, P. Mori-Sánchez, W. Yang, "Insights into current limitations of density functional theory", *Science* **2008**, *321*, 792–794.

[330] F. Jensen, *Introduction to computational chemistry*, John wiley & sons, **2017**.

[331] P. J. Stephens, F. J. Devlin, C. F. Chabalowski, M. J. Frisch, "Ab Initio Calculation of Vibrational Absorption and Circular Dichroism Spectra Using Density Functional Force Fields", *The Journal of Physical Chemistry* **1994**, *98*, 11623–11627.

[332] A. R. Leach, *Molecular modelling: principles and applications*, Pearson education, **2001**.

[333] G. A. Kaminski, R. A. Friesner, J. Tirado-Rives, W. L. Jorgensen, "Evaluation and reparametrization of the OPLS-AA force field for proteins via comparison with accurate quantum chemical calculations on peptides", *The Journal of Physical Chemistry B* **2001**, *105*, 6474–6487.

[334] J. L. Banks, H. S. Beard, Y. Cao, A. E. Cho, W. Damm, R. Farid, A. K. Felts, T. A. Halgren, D. T. Mainz, J. R. Maple, et al., "Integrated modeling program, applied chemical theory (IMPACT)", *Journal of computational chemistry* **2005**, *26*, 1752–1780.

[335] V. Hornak, R. Abel, A. Okur, B. Strockbine, A. Roitberg, C. Simmerling, "Comparison of multiple Amber force fields and development of improved protein backbone parameters", *Proteins: Structure Function and Bioinformatics* **2006**, *65*, 712–725.

[336] A. K. Rappé, C. J. Casewit, K. Colwell, W. Goddard Iii, W. Skiff, "UFF, a full periodic table force field for molecular mechanics and molecular dynamics simulations", *Journal of the American chemical society* **1992**, *114*, 10024–10035.

[337] A. Warshel, M. Levitt, "Theoretical studies of enzymic reactions: dielectric, electrostatic and steric stabilization of the carbonium ion in the reaction of lysozyme", *Journal of molecular biology* **1976**, *103*, 227–249.

[338] H. M. Senn, W. Thiel, "QM/MM methods for biomolecular systems", *Angewandte Chemie International Edition* **2009**, *48*, 1198–1229.

[339] S. Chakrabarti, D. Bhattacharyya, D. Dasgupta, "Structural basis of DNA recognition by anticancer antibiotics, chromomycin A3, and mithramycin: Roles of minor groove width and ligand flexibility", *Biopolymers* **2000**, *56*, 85–95.

[340] S. Ghosh, A. Kar, S. Chowdhury, D. Dasgupta, "Ellipticine binds to a human telomere sequence: an additional mode of action as a putative anticancer agent?", *Biochemistry* **2013**, *52*, 4127–4137.

[341] J. Kypr, I. Kejnovská, D. Renčiuk, M. Vorlíčková, "Circular dichroism and conformational polymorphism of DNA", *Nucleic acids research* **2009**, *37*, 1713–1725.

[342] A. Włodarczyk, P. Grzybowski, A. Patkowski, A. Dobek, "Effect of ions on the polymorphism, effective charge, and stability of human telomeric DNA. Photon correlation spectroscopy and circular dichroism studies", *The Journal of Physical Chemistry B* **2005**, *109*, 3594–3605.

[343] S. Paramasivan, I. Rujan, P. H. Bolton, "Circular dichroism of quadruplex DNAs: applications to structure, cation effects and ligand binding", *Methods* **2007**, *43*, 324–331.

[344] S. Masiero, R. Trotta, S. Pieraccini, S. De Tito, R. Perone, A. Randazzo, G. P. Spada, "A non-empirical chromophoric interpretation of CD spectra of DNA G-quadruplex structures", *Organic & biomolecular chemistry* **2010**, *8*, 2683–2692.

[345] A. Randazzo, G. P. Spada, M. W. da Silva in *Quadruplex nucleic acids*, Springer, **2012**, pp. 67–86.

[346] M. Vorlíčková, I. Kejnovská, J. Sagi, D. Renčiuk, K. Bednářová, J. Motlová, J. Kypr, "Circular dichroism and guanine quadruplexes", *Methods* **2012**, *57*, 64–75.

[347] H. A. Benesi, J. Hildebrand, "A spectrophotometric investigation of the interaction of iodine with aromatic hydrocarbons", *Journal of the American Chemical Society* **1949**, *71*, 2703–2707.

[348] D. Lin, X. Fei, Y. Gu, C. Wang, Y. Tang, R. Li, J. Zhou, "A benzindole substituted carbazole cyanine dye: a novel targeting fluorescent probe for parallel c-myc G-quadruplexes", *Analyst* **2015**, *140*, 5772–5780.

[349] M. M. Islam, S. Fujii, S. Sato, T. Okauchi, S. Takenaka, "A selective G-quadruplex DNA-stabilizing ligand based on a cyclic naphthalene diimide derivative", *Molecules* **2015**, *20*, 10963–10979.

[350] M.-H. Hu, S.-B. Chen, B. Wang, T.-M. Ou, L.-Q. Gu, J.-H. Tan, Z.-S. Huang, "Specific targeting of telomeric multimeric G-quadruplexes by a new triaryl-substituted imidazole", *Nucleic acids research* **2017**, *45*, 1606–1618.

[351] M. Ferraroni, C. Bazzicalupi, F. Papi, G. Fiorillo, L. M. Guamán-Ortiz, A. Nocentini, A. I. Scovassi, P. Lombardi, P. Gratteri, "Solution and Solid-State Analysis of Binding of 13-Substituted Berberine Analogues to Human Telomeric G-quadruplexes", *Chemistry–An Asian Journal* **2016**, *11*, 1107–1115.

[352] C. Bazzicalupi, M. Ferraroni, F. Papi, L. Massai, B. Bertrand, L. Messori, P. Gratteri, A. Casini, "Determinants for Tight and Selective Binding of a Medicinal Dicarbene Gold(I) Complex to a Telomeric DNA G-Quadruplex: a Joint ESI MS and XRD Investigation", *Angewandte Chemie* **2016**, *128*, 4328–4331.

[353] F. Papi, C. Bazzicalupi, M. Ferraroni, L. Massai, B. Bertrand, P. Gratteri, D. Colangelo, L. Messori, "The [Au(9-methylcaffein-8-ylidene)2]+/DNA Tel23 System: Solution, Computational and Biological Studies", *Chemistry–A European Journal* **2017**, *23*, 13784–13791.

[354] I. Berger, C. Kang, N. Sinha, M. Wolters, A. Rich, "A highly efficient 24-condition matrix for the crystallization of nucleic acid fragments", *Acta Crystallographica Section D: Biological Crystallography* **1996**, *52*, 465–468.

[355] N. H. Campbell, G. N. Parkinson, "Crystallographic studies of quadruplex nucleic acids", *Methods* **2007**, *43*, 252–263.

[356] W. Kabsch, "Xds", *Acta Crystallographica Section D: Biological Crystallography* **2010**, *66*, 125–132.

[357] G. M. Sheldrick, "A short history of SHELX", *Acta Crystallographica Section A: Foundations of Crystallography* **2008**, *64*, 112–122.

[358] A. J. McCoy, R. W. Grosse-Kunstleve, P. D. Adams, M. D. Winn, L. C. Storoni, R. J. Read, "Phaser crystallographic software", *Journal of applied crystallography* **2007**, *40*, 658–674.

[359] A. Vagin, A. Teplyakov, "MOLREP: an automated program for molecular replacement", *Journal of applied crystallography* **1997**, *30*, 1022–1025.

[360] G. N. Murshudov, A. A. Vagin, E. J. Dodson, "Refinement of macromolecular structures by the maximum-likelihood method", *Acta Crystallographica Section D: Biological Crystallography* **1997**, *53*, 240–255.

[361] M. D. Winn, C. C. Ballard, K. D. Cowtan, E. J. Dodson, P. Emsley, P. R. Evans, R. M. Keegan, E. B. Krissinel, A. G. Leslie, A. McCoy, et al., "Overview of the CCP4 suite and current developments", *Acta Crystallographica Section D: Biological Crystallography* **2011**, *67*, 235–242.

[362] P. D. Adams, P. V. Afonine, G. Bunkóczi, V. B. Chen, I. W. Davis, N. Echols, J. J. Headd, L.-W. Hung, G. J. Kapral, R. W. Grosse-Kunstleve, et al., "PHENIX: a comprehensive Python-based system for macromolecular structure solution", *Acta Crystallographica Section D: Biological Crystallography* **2010**, *66*, 213–221.

[363] P. Emsley, B. Lohkamp, W. G. Scott, K. Cowtan, "Features and development of Coot", *Acta Crystallographica Section D: Biological Crystallography* **2010**, *66*, 486–501.

[364] Schrödinger Release 2016: Maestro, Schrödinger, LLC, New York, NY, **2016**.

[365] N. W. Moriarty, R. W. Grosse-Kunstleve, P. D. Adams, "electronic Ligand Builder and Optimization Workbench (eLBOW): a tool for ligand coordinate and restraint generation", *Acta Crystallographica Section D: Biological Crystallography* **2009**, *65*, 1074–1080.

[366] Schrödinger Suite 2016, Schrödinger, LLC, New York, NY, **2016**.

[367] M. Ferraroni, C. Bazzicalupi, F. Papi, P. Gratteri in *G-Quadruplex Structures, Formation and Role in Biology*, Nova Science Publishers, **2016**, pp. 81–102.

[368] Schrödinger Release 2016: Jaguar, Schrödinger, LLC, New York, NY, **2016**.

[369] Schrödinger Release 2016: Glide, Schrödinger, LLC, New York, NY, **2016**.

[370] Schrödinger Release 2016: MacroModel, Schrödinger, LLC, New York, NY, **2016**.

[371] Schrödinger Release 2016: Impact, Schrödinger, LLC, New York, NY, **2016**.

[372] T. Wilson, P. J. Costa, V. Félix, M. P. Williamson, J. A. Thomas, "Structural studies on dinuclear ruthenium(II) complexes that bind diastereoselectively to an antiparallel folded human telomere sequence", *Journal of medicinal chemistry* **2013**, *56*, 8674–8683.

[373] Schrödinger Release 2016: QSite, Schrödinger, LLC, New York, NY, **2016**.

[374] D. Bhowmik, M. Hossain, F. Buzzetti, R. D'Auria, P. Lombardi, G. S. Kumar, "Biophysical studies on the effect of the 13 position substitution of the anticancer alkaloid berberine on its DNA binding", *The Journal of Physical Chemistry B* **2012**, *116*, 2314–2324.

[375] D. Bhowmik, F. Buzzetti, G. Fiorillo, P. Lombardi, G. S. Kumar, "Spectroscopic studies on the binding interaction of novel 13-phenylalkyl analogs of the natural alkaloid berberine to nucleic acid triplexes", *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy* **2014**, *120*, 257–264.

[376] D. Bhowmik, F. Buzzetti, G. Fiorillo, L. Franchini, T. M. Syeda, P. Lombardi, G. S. Kumar, "Calorimetry and thermal analysis studies on the binding of 13-phenylalkyl and 13-diphenylalkyl berberine analogs to tRNAphe", *Journal of Thermal Analysis and Calorimetry* **2014**, *118*, 461–473.

[377] F. Papi, M. Ferraroni, R. Rigo, S. Da Ros, C. Bazzicalupi, C. Sissi, P. Gratteri, "Role of the benzodioxole group in the interactions between the natural alkaloids Chelerythrine and Coptisine and the human telomeric G-quadruplex DNA. A multiapproach investigation.", *Journal of Natural Products* **2017**, accepted for publication.

[378] M. Maiti, G. S. Kumar, "Polymorphic nucleic acid binding of bioactive isoquinoline alkaloids and their role in cancer", *Journal of nucleic acids* **2010**, *2010*.

[379] G. W. Collie, N. H. Campbell, S. Neidle, "Loop flexibility in human telomeric quadruplex small-molecule complexes", *Nucleic acids research* **2015**, *43*, 4785–4799.

[380] S. Neidle, "The structures of quadruplex nucleic acids and their drug complexes", *Current opinion in structural biology* **2009**, *19*, 239–250.

[381] N. H. Campbell, G. N. Parkinson, A. P. Reszka, S. Neidle, "Structural basis of DNA quadruplex recognition by an acridine drug", *Journal of the American Chemical Society* **2008**, *130*, 6722–6724.

[382] B. M. Kariuki, W. Jones, "Five salts of berberine", *Acta Crystallographica Section C: Crystal Structure Communications* **1995**, *51*, 1234–1240.

[383] S. Man, M. Potáček, M. Nečas, Z. Žák, J. Dostál, "Molecular and crystal structures of three berberine derivatives", *Molecules* **2001**, *6*, 433–441.

[384] J. Marek, D. Hulová, J. Dostál, R. Marek, "Berberine formate-succinic acid (1/1)", *Acta Crystallographica Section C: Crystal Structure Communications* **2003**, *59*, 0583–0585.

[385] I. Russo Krauss, S. Ramaswamy, S. Neidle, S. Haider, G. N. Parkinson, "Structural Insights into the Quadruplex-Duplex 3' Interface Formed from a Telomeric Repeat: A Potential Molecular Target", *Journal of the American Chemical Society* **2016**, *138*, 1226–1233.

[386] A. Kettani, S. Bouaziz, A. Gorin, H. Zhao, R. A. Jones, D. J. Patel, "Solution structure of a Na cation stabilized DNA quadruplex containing G·G·G·G and G·C·G·C tetrads formed by GGGC repeats observed in adeno-associated viral DNA", *Journal of molecular biology* **1998**, *282*, 619–636.

[387] N. B. Leontis, J. Stombaugh, E. Westhof, "The non-Watson–Crick base pairs and their associated isostericity matrices", *Nucleic acids research* **2002**, *30*, 3497–3531.

[388] K. W. Lim, P. Alberti, A. Guédin, L. Lacroix, J.-F. Riou, N. J. Royle, J.-L. Mergny, A. T. Phan, "Sequence variant (CTAGGG)n in the human telomere favors a G-quadruplex structure containing a G·C·G·C tetrad", *Nucleic acids research* **2009**, *37*, 6239–6248.

[389] N. Borbone, J. Amato, G. Oliviero, V. D'Atri, V. Gabelica, E. De Pauw, G. Piccialli, L. Mayol, "d(CGGTGGT) forms an octameric parallel G-quadruplex via stacking of unusual G(:C):G(:C):G(:C):G(:C) octads", *Nucleic acids research* **2011**, *39*, 7848–7857.

[390] P. L. T. Tran, A. De Cian, J. Gros, R. Moriyama, J.-L. Mergny in *Quadruplex Nucleic Acids*, Springer, **2012**, pp. 243–273.

[391] E. N. Nikolova, H. Zhou, F. L. Gottardo, H. S. Alvey, I. J. Kimsey, H. M. Al-Hashimi, "A historical account of hoogsteen base-pairs in duplex DNA", *Biopolymers* **2013**, *99*, 955–968.

[392] J. Sheng, J. Gan, Z. Huang, "Structure-Based DNA-Targeting Strategies with Small Molecule Ligands for Drug Discovery", *Medicinal research reviews* **2013**, *33*, 1119–1173.

[393] A. Granzhan, N. Kotera, M.-P. Teulade-Fichou, "Finding needles in a basestack: recognition of mismatched base pairs in DNA by small molecules", *Chemical Society Reviews* **2014**, *43*, 3630–3665.

[394] N. Valls, R. A. Steiner, G. Wright, G. N. Murshudov, J. A. Subirana, "Variable role of ions in two drug intercalation complexes of DNA", *JBIC Journal of Biological Inorganic Chemistry* **2005**, *10*, 476–482.

[395] M. Ferraroni, C. Bazzicalupi, A. R. Bilia, P. Gratteri, "X-Ray diffraction analyses of the natural isoquinoline alkaloids Berberine and Sanguinarine complexed with double helix DNA d(CGTACG)", *Chemical Communications* **2011**, *47*, 4917–4919.

[396] X.-l. Yang, H. Robinson, Y.-G. Gao, A. H.-J. Wang, "Binding of a macrocyclic bisacridine and ametantrone to CGTACG involves similar unusual intercalation platforms", *Biochemistry* **2000**, *39*, 10950–10957.

[397] S.-H. Chou, K.-H. Chin, A. H.-J. Wang, "Unusual DNA duplex and hairpin motifs", *Nucleic acids research* **2003**, *31*, 2461–2474.

[398] T. Sunami, J. Kondo, T. Kobuna, I. Hirao, K. Watanabe, K.-i. Miura, A. Takénaka, "Crystal structure of d(GCGAAAGCT) containing a parallel-stranded duplex with homo base pairs and an anti-parallel duplex with Watson–Crick base pairs", *Nucleic acids research* **2002**, *30*, 5253–5260.

[399] S. E. Muser, P. J. Paukstelis, "Three-dimensional DNA crystals with pH-responsive non-canonical junctions", *Journal of the American Chemical Society* **2012**, *134*, 12557–12564.

[400] J. N. Lisgarten, M. Coll, J. Portugal, C. W. Wright, J. Aymami, "The antimalarial and cytotoxic drug cryptolepine intercalates into DNA at cytosine-cytosine sites", *Nature Structural & Molecular Biology* **2002**, *9*, 57–60.

[401] S. K. Pradhan, D. Dasgupta, G. Basu, "Human telomere d[(TTAGGG)4] undergoes a conformational transition to the Na+-form upon binding with sanguinarine in presence of K+", *Biochemical and biophysical research communications* **2011**, *404*, 139–142.

[402] L. Martino, B. Pagano, I. Fotticchia, S. Neidle, C. Giancola, "Shedding light on the interaction between TMPyP4 and human telomeric quadruplexes", *The journal of physical chemistry B* **2009**, *113*, 14779–14786.

[403] C. Antonacci, J. B. Chaires, R. D. Sheardy, "Biophysical characterization of the human telomeric (TTAGGG)4 repeat in a potassium solution", *Biochemistry* **2007**, *46*, 4654–4660.

[404] J. B. Patil, J. Kim, G. Jayaprakasha, "Berberine induces apoptosis in breast cancer cells (MCF-7) through mitochondrial-dependent pathway", *European journal of pharmacology* **2010**, *645*, 70–78.

[405] B. Lu, M. Hu, J. Peng, et al., "Cytotoxicity of berberine on human cervical carcinoma HeLa cells through mitochondria, death receptor and MAPK pathways, and in-silico drug-target prediction", *Toxicology in vitro* **2010**, *24*, 1482–1490.

[406] J.-L. Mergny, L. Lacroix, "Analysis of thermal melting curves", *Oligonucleotides* **2003**, *13*, 515–537.

[407] D. Bhowmik, F. Buzzetti, G. Fiorillo, F. Orzi, T. M. Syeda, P. Lombardi, G. S. Kumar, "Synthesis of new 13-diphenylalkyl analogues of berberine and elucidation of their base pair specificity and energetics of DNA binding", *MedChemComm* **2014**, *5*, 226–231.

[408] G. N. Parkinson, R. Ghosh, S. Neidle, "Structural basis for binding of porphyrin to human telomeres", *Biochemistry* **2007**, *46*, 2390–2397.

[409] G. N. Parkinson, F. Cuenca, S. Neidle, "Topology conservation and loop flexibility in quadruplex-drug recognition: crystal structures of inter and intramolecular telomeric DNA quadruplex-drug complexes", *Journal of molecular biology* **2008**, *381*, 1145–1156.

[410] M. Micco, G. W. Collie, A. G. Dale, S. A. Ohnmacht, I. Pazitna, M. Gunaratnam, A. P. Reszka, S. Neidle, "Structure-based design and evaluation of naphthalene diimide G-quadruplex ligands as telomere targeting agents in pancreatic cancer cells", *Journal of medicinal chemistry* **2013**, *56*, 2959–2974.

[411] Z.-F. Wang, M.-H. Li, S.-T. D. Hsu, T.-C. Chang, "Structural basis of sodium-potassium exchange of a human telomeric DNA quadruplex without topological conversion", *Nucleic acids research* **2014**, *42*, 4723–4733.

[412] K. L. Hayden, D. E. Graves, "Addition of bases to the 5'-end of human telomeric DNA: influences on thermal stability and energetics of unfolding", *Molecules* **2014**, *19*, 2286–2298.

[413] J. Amato, R. Morigi, B. Pagano, A. Pagano, S. Ohnmacht, A. De Magis, Y.-P. Tiang, G. Capranico, A. Locatelli, A. Graziadio, et al., "Toward the development of specific G-quadruplex binders: Synthesis, biophysical, and biological studies of new hydrazone derivatives", *Journal of medicinal chemistry* **2016**, *59*, 5706–5720.

[414] G. Scatchard, "The attractions of proteins for small molecules and ions", *Annals of the New York Academy of Sciences* **1949**, *51*, 660–672.

[415] J. D. McGhee, P. H. von Hippel, "Theoretical aspects of DNA-protein interactions: co-operative and non-co-operative binding of large ligands to a one-dimensional homogeneous lattice", *Journal of molecular biology* **1974**, *86*, 469–489.

[416] D. Colangelo, A. Ghiglia, I. Viano, H. Mahboobi, A. Ghezzi, C. Cassino, D. Osella, "Might telomerase enzyme be a possible target for trans-Pt(II) complexes?", *Journal of inorganic biochemistry* **2004**, *98*, 61–67.

[417] J. Wirmer-Bartoschek, L. E. Bendel, H. R. Jonker, J. T. Grün, F. Papi, C. Bazzicalupi, L. Messori, P. Gratteri, H. Schwalbe, "Solution NMR Structure of a Ligand/Hybrid-2-G-Quadruplex Complex Reveals Rearrangements that Affect Ligand Binding", *Angewandte Chemie* **2017**, *56*, 7102–7106.

[418] C. Gabbiani, A. Casini, L. Messori, A. Guerri, M. A. Cinellu, G. Minghetti, M. Corsini, C. Rosani, P. Zanello, M. Arca, "Structural characterization, solution studies, and DFT calculations on a series of binuclear gold(III) oxo complexes: relationships to biological properties", *Inorganic chemistry* **2008**, *47*, 2368–2379.

[419] B. F. Eichman, B. H. Mooers, M. Alberti, J. E. Hearst, P. S. Ho, "The crystal structures of psoralen cross-linked DNAs: drug-dependent formation of Holliday junctions", *Journal of molecular biology* **2001**, *308*, 15–26.

[420] A. L. Brogden, N. H. Hopcroft, M. Searcey, C. J. Cardin, "Ligand Bridging of the DNA Holliday Junction: Molecular Recognition of a Stacked-X Four-Way Junction by a Small Molecule", *Angewandte Chemie International Edition* **2007**, *46*, 3850–3854.

RINGRAZIAMENTI

La scrittura di questi tesi è il compimento di un progetto durato tre anni che mi ha permesso di inserirmi nel contesto accademico e della ricerca. Sono stati tre anni ricchi di esperienze, speranze e delusioni, successi e risultati inattesi, i quali senza dubbio mi hanno arricchito come persona e come scienziato. Vorrei ringraziare in primo luogo la Prof.ssa Carla Bazzicalupi per l'opportunità offertami, per il sostegno fornito in questi anni e per la costante e scrupolosa attenzione al lavoro svolto. Vorrei inoltre ringraziare la Dott.ssa Marta Ferraroni e la Prof.ssa Paola Gratteri per il supporto e l'impegno nella conduzione dell'attività di ricerca, in particolar modo, rispettivamente, per la parte cristallografica e di simulazione *in silico*. Un ringraziamento sentito al Prof. Luigi Messori ed al suo gruppo di ricerca (Lara, Tiziano, Damiano) con cui è stata avviata una stimolante e produttiva collaborazione scientifica nel campo dei complessi metallici con possibili applicazioni farmaceutiche. Vorrei ringraziare infine la Prof.ssa Claudia Sissi, la Dott.ssa Paola Ivana Scovassi, il Dott. Paolo Lombardi e la ditta Naxospharma srl, il Prof. Donato Colangelo, il Prof. Harald Schwalbe, il Prof. Sylvestre Bonnet, e loro collaboratori per la disponibilità nell'avviare un'attività di ricerca comune che ha portato a positivi risultati nel corso dei tre anni di dottorato.

Vorrei qui ringraziare le persone che mi hanno accompagnato, o per meglio dire sopportato, durante questa affascinante ed intensa avventura. In primis babbo e mamma, per la pazienza e il continuo supporto. Un grazie alla mia cara sorellina, alla cara nonna Aldina, ai nonni Sergio e Giuliana, alla zia Annina, agli zii Alessandro e Rossana, ai cugini Lorenzo e Valentina. Un grazie ai tanti amici che mi hanno accompagnato in questi tre anni e che spero possa ancora vedere in futuro. Amici, amici da sempre: Matteo, Niccolò, Deni, Pax-Pacis, Cloud, Simo, Goozer, Pelli, Ippiero, Tommy, Cami. Amici dell'unifi: Damiano, Alessio, Mattia, Tiziana, Maher, Laurina, Tesino, Nao & Franci, Chiarina, Iddonno, Iccionco, Riccardino, Luca, Matteo, la Sarina, Marta, la Mery, Giacche, Enrico, Stefano, Fabio, Lorenzo C., Mr. Marzi e tanti altri. Un ringraziamento speciale alle laureande con cui ho avuto il piacere di ~~bestemm~~ lavorare insieme: Claudina, Giulia O., Giulia C. Direi possa bastare, grazie a tutti e W Rocket Man!

♡