UNIVERSITY OF CATANIA

DEPARTMENT OF CHEMICAL SCIENCES

INTERNATIONAL PhD IN CHEMICAL SCIENCES – XXXVI CYCLE

Gabriele Travagliante

Biological role of the conformations of biopolymers. Porphyrins as specific reporters

PhD Thesis

Tutor: Prof. Alessandro D'Urso

PhD Coordinator: Prof. Salvatore Sortino

ACADEMIC CYCLE 2020/2023

Doctor of Philosophy Thesis
International PhD in Chemical Sciences, XXXVI Cycle, 2020-2023
Biological role of the conformations of biopolymers. Porphyrins as specific reporters
GABRIELE TRAVAGLIANTE

Department of Chemical Sciences,
Viale Andrea Doria, 6,
95125, Catania (CT) - Italy
Tel. +39 095 7385125
www.dsc.unict.it

Reviewed and approved by:

Prof. Angela Mammana
University of Dayton (Ohio), Department of Chemistry
amammana1@udayton.edu

Prof. Andrea Romeo
University of Messina (Italy), Department of Chemical, Biological, Pharmaceutical and Environmental Sciences
andrea.romeo@unime.it

page left blank intentionally

Table of Contents

Abstract

Biopolymers, such as polynucleotides and proteins, exhibit a variety of conformations that govern their biological roles. The structure adopted by biomolecule is thermodynamically the most stable, which is allocated in a minimum energy state. The role of the structure is to protect the information that the molecule carries from the attack of other molecules, allowing to the biomolecules to reach the target. Therefore the three-dimensional arrangement of biopolymers directly impacts their biological activity and is essential for their optimal functions.

Hence, studying the conformation of biopolymers is vital for comprehending their structure, function, and interactions within biological systems. Indeed exploring the dynamics of these conformations has yielded insights into fundamental biological processes and contributed to therapeutic advancements. In particular by studying folding patterns, secondary structures, and higher-order arrangements, we have gained knowledge about gene regulation, cellular signaling, enzymatic catalysis, and disease mechanisms.

Moreover, the study of biomolecule conformations has facilitated the design of innovative drugs, therapeutic interventions, and diagnostic tools. Indeed to find ligands that interact specifically with conformation allow to interfere with biological process and/or to obtain more information on that conformation.

Through the examination of three-dimensional structures and dynamic alterations of biopolymers, we can obtain invaluable insights into the fundamental mechanisms governing vital life processes. This knowledge bears immense significance across a range of scientific disciplines, involving molecular biology, biochemistry, drug discovery, and biotechnology.

In this thesis, I explored the secondary structures of various polynucleic acids using biophysical methods and employed porphyrins as specific reporters. We investigated how micro-RNA stability influences biological activity and revealed the role of porphyrins in interacting with miRNA and telomeric RNA sequences. We also studied cytosine-rich variants from the human c-MYC promoter, confirming their ability to adopt i-motif conformations. Our findings provide insights into the roles of nucleic acid structures in biological systems, offering avenues for advancements in disease diagnosis, treatment, and beyond. Future work will focus on molecular interactions, including with porphyrins, with implications for biomedicine, biotechnology, and nanotechnology.

List of abbreviations

A	adenine
AGO	argonaute
a-miR	artificial microRNA
C	cytosine
CD	circular dichroism
D.I.T.	digital integration time
dG4	DNA G-quadruplex
DSC	differential scanning calorimetry
dsRNA	double-stranded RNA
G	guanine
G4	G-quadruplex
H2T4	meso-tetrakis(4-N-methylpyridyl)porphyrin
H2TCPPSpm4	meso-tetrakis(4-carboxyphenylspermine) porphyrin
hsa-miR	human-miR
ICD	induced circular dichroism
IDT	Integrated DNA Technologies
iM	i-motif
miR	microRNA
miRNA	microRNA
oncomiR	oncogenic miRNA
PAGE	polyacrylamide gel electrophoresis
PBS	Phosphate-buffered saline
PDT	photodynamic therapy

pH_T	transition pH
pre-miRNA	precursor miRNA
pri-miRNA	primary miRNA
rG4	RNA G-quadruplex
RISC	RNA-induced silencing complex
RLS	Resonance light scattering
T	thymine
TERRA	telomeric repeat-containing RNA
Tm	melting temperature
TRBP	TAR RNA binding protein
TS	tumor suppressor
U	uracil

1. Introduction

In biochemistry, the structure of biopolymers - proteins, nucleic acids, carbohydrates, lipids - plays an important role in their function. This conformation or three-dimensional structure describes how these molecules interact with their environment and with each other, enabling a complex symphony of biological processes that sustain life.

Molecular conformation is any spatial arrangement of the atoms in a molecule which can be interconverted by rotations about formally single bonds. Biopolymers, such as polynucleotides, polypeptides or polysaccharides, may change conformation in response to changes in their environment.

The precise three-dimensional structure, or native conformation, of the protein is crucial to its function. These conformations are primarily determined by the protein's primary amino acid sequence and the interplay of forces such as hydrophobic interactions, hydrogen bonding, and electrostatic interactions [1,2]. Correctly folded proteins engage in critical physiological processes, including DNA replication, transporting molecules, catalyzing metabolic reactions, and providing structural support to cells [3,4].

Conversely, protein misfolding can lead to harmful consequences. Misfolded proteins can aggregate and form insoluble plaques, causing diseases such as Alzheimer's, Parkinson's, Huntington’s disease and many other degenerative and neurodegenerative disorders [5–7]. The delicate balance between protein folding and misfolding underscores the significance of understanding biopolymer conformations.

Similarly, nucleic acids (DNA and RNA), the biopolymers responsible for storing and transmitting genetic information, adopt specific conformations that influence their function. The iconic double-helix structure of DNA, known as B-DNA, enables efficient storage and replication of genetic information [8,9]. However, DNA can also adopt non-canonical conformations such as A-DNA, Z-DNA, triplex DNA, and quadruplex DNA under specific conditions or sequences [9–14]. These alternative structures have been implicated in various biological processes, including DNA replication, transcription, chromatin organization, and genome stability [15].

RNA, with its greater chemical complexity, can adopt a broader range of structures compared to DNA. These include the well-known A-form helix, but also more complex secondary structures such as hairpins, loops, bulges, and pseudoknots [16,17]. Furthermore, the discovery of G-quadruplexes in RNA has provided a new layer of complexity in our understanding of RNA structure and function [18,19]. These diverse RNA conformations enable its multiple roles in protein synthesis, gene regulation, and as the genetic material in some viruses [20,21]. The folding of RNA into these non-canonical structures can be influenced by environmental conditions, sequence context, and the presence of binding proteins or small molecules [22].

Alterations in the normal conformation of these nucleic acids can cause mutations leading to genetic disorders and cancers. For instance, the formation and stabilization of non-canonical DNA structures can impede the replication machinery, leading to genetic instability and potentially carcinogenesis [15,23]. Similarly, aberrant RNA structures can disrupt normal cellular processes, leading to diseases such as neurodegeneration and cancer [24,25].

Polysaccharides, another category of biopolymers, also exhibit a diverse range of conformations and biological roles [26]. Cellulose, a linear polymer of glucose units, adopts a rigid, crystalline structure and provides structural support in plant cell walls. In contrast, the highly branched structure of glycogen, another glucose polymer, allows it to act as an efficient energy storage molecule in animal cells. It has been reported that polysaccharides participate in biological information carrying and transmission, cell recognition, metabolism and secretion, immune response regulation, and protein transfer process [27,28].

Lipids, characterized by their hydrophobicity, form membrane bilayers due to their amphipathic nature. The conformation of these bilayers enables selective permeability, compartmentalization of cellular processes, and signal transduction [29,30].

At the heart of these biopolymer interactions lies the principle of molecular recognition, an essential concept in biochemistry. Molecular recognition refers to the specific interaction between two or more molecules through noncovalent bonding such as hydrogen bonding, metal coordination, hydrophobic forces, van der Waals forces, π-π interactions, electrostatic effects, or even solvent effects [31,32].

In all of these cases, the precise conformation of the biopolymers allows for highly specific molecular recognition events. For example, the intricate conformation of an enzyme enables it to recognize and bind to its substrate in a lock-and-key manner. Similarly, the specific base-pairing conformation in nucleic acids allows DNA replication and RNA transcription to occur with high fidelity [33].

This intricate interplay between biopolymer conformation, molecular recognition, and biological function forms the cornerstone of biological processes at the molecular level. Further understanding of these relationships can illuminate the fundamental mechanisms of life and guide the design of therapeutic interventions for diseases caused by conformational abnormalities.

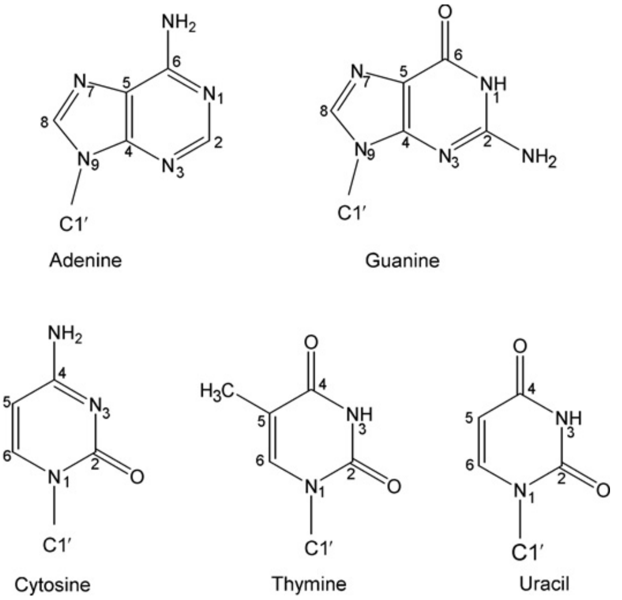
The next paragraph offers an in-depth exploration of the structures of DNA and RNA, which are the principal subjects of this dissertation.

1.1. Nucleic acid structures

Nucleic acids (DNA and RNA) are formed by units called nucleotides. A nucleotide is composed of three components:

1. A phosphate group: It is the part of the nucleotide that allows them to connect together, creating the backbone of the DNA and RNA structures.
2. A sugar: In DNA, the sugar is deoxyribose, while in RNA, it is ribose. The difference lies in the presence of one less oxygen atom in the sugar of DNA as compared to RNA.
3. A nitrogenous base: These bases are planar, aromatic molecules categorized into two groups - pyrimidine bases, thymine (T) and cytosine(C), and purine bases, adenine (A) and guanine (G). In the case of RNA, thymine is substituted by uracil (U) (**Fig. 1**).

A nucleoside, on the other hand, is very similar to a nucleotide, but without the phosphate group. It consists only of a sugar (ribose or deoxyribose) and a nitrogenous base (adenine, guanine, cytosine, thymine, or uracil).



Adenine Guanine

Cytosine Thymine Uracil

***Fig. 1** Chemical structure of the five bases of DNA and RNA*

The connection that links the sugar and base in nucleic acids is termed a glycosidic bond, and its stereochemistry plays a significant role. In natural nucleic acids, this glycosidic bond maintains a beta (β) configuration, signifying that the base, when seen in the plane's orientation, is on the same plane side as the 5' hydroxyl group and is above the sugar plane (**Fig. 2 (a)**).

While α-nucleosides and corresponding α-oligonucleosides, which locate their bases in the "below" orientation (**Fig. 2 (b)**), can be chemically synthesized, they are substantially more resistant to nuclease degradation compared to the conventional natural β-oligomers [34].

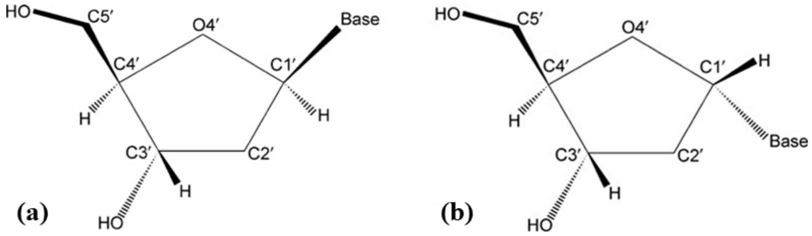


***Fig. 2 (a)** The stereochemistry of a natural β-nucleoside and **(b)** of an α-nucleoside. Solid bonds are extending out of the plane of the page, towards the observer. Conversely, the dashed bonds are retreating from the viewer. Adapted from ref.* [35].

In a nucleic acid, individual nucleoside units are sequentially linked via phosphodiester bonds that connect to the 3′ and 5′ positions of the sugar molecules. As a result, the complete repeating segment in a nucleic acid is a 3′,5′-nucleotide (**Fig. 3**). The polymer as depicted possesses a polarity; one end has a 5′-hydroxyl or phosphate terminal, while the other has a 3′-phosphate or hydroxyl terminal.

It is known that nucleic acids exhibit vast polymorphism, due to the different conformations that the molecules making up DNA and RNA can assume.

Nucleobases for example, despite the individual bases are planar, there can be substantial flexibility in base pairs and successive bases within a single strand. The degree of this flexibility is partly influenced by the properties of the bases and base pairs themselves, but it is more substantially related to the conditions of their base-stacking surroundings [35]. Consequently, the morphological representation of bases has become significant in explaining and comprehending many features and distortions in nucleic acids that are dependent on the sequence. These characteristics are usually considered on a local level, primarily, while longer-range effects, such as helix bending, can also be assessed at a more comprehensive level [36–38].

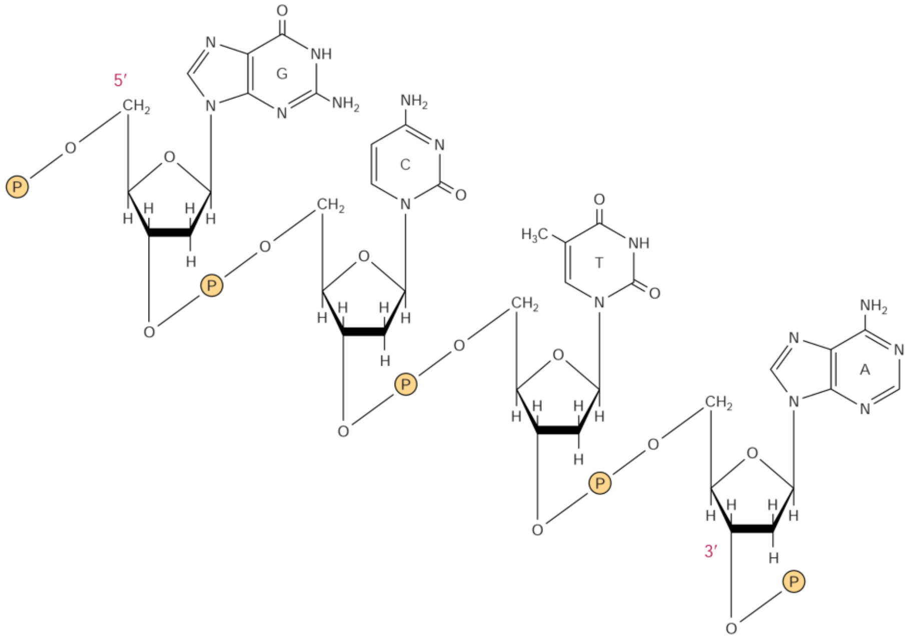


***Fig. 3** A segment of one strand of a DNA molecule in which the bases are held together by a phosphodiester backbone between 2'-deoxyribosyl moieties attached to the nucleobases by an N-glycosidic bond. Reproduced from ref* [39].

Sugar configurations can also influence the conformation of nucleic acids, indeed the pentagonal deoxyribose sugar ring in DNA naturally exhibits a non-flat structure, a feature known as puckering. The exact shape of a deoxyribose ring can be fully defined by the five internal torsion angles it contains. In principle, there is a continuum of interconvertible puckers, separated by energy barriers. These various puckers are produced by systematic changes in the ring torsion angles [40]. Numerous unique geometries of deoxyribose ring pucker have been discovered using X-ray crystallography and NMR methods [41,42]. If one ring atom deviates from the plane of the remaining four, this is categorized as an envelope pucker. More frequently, two atoms diverge from the plane formed by the remaining three, with these two on opposite sides of the plane. Generally, one of these two atoms has a more pronounced deviation from the plane, leading to a twist conformation. The side from which the atomic displacement occurs is crucial. If the major displacement is on the same side as the base and C4′–C5′ bond, the atom involved is termed endo. If it is on the opposite side, it is called exo. The puckers most frequently observed in crystal structures of isolated nucleosides and nucleotides are usually near the C2′-endo or C3′-endo type (**Fig. 4**).

Alterations in sugar pucker play a significant role in the structure of oligo- and polynucleotides because they can shift the positioning of C1′, C3′ and C4′ substituents. This leads to considerable modifications in the conformation of the backbone and the overall structure.

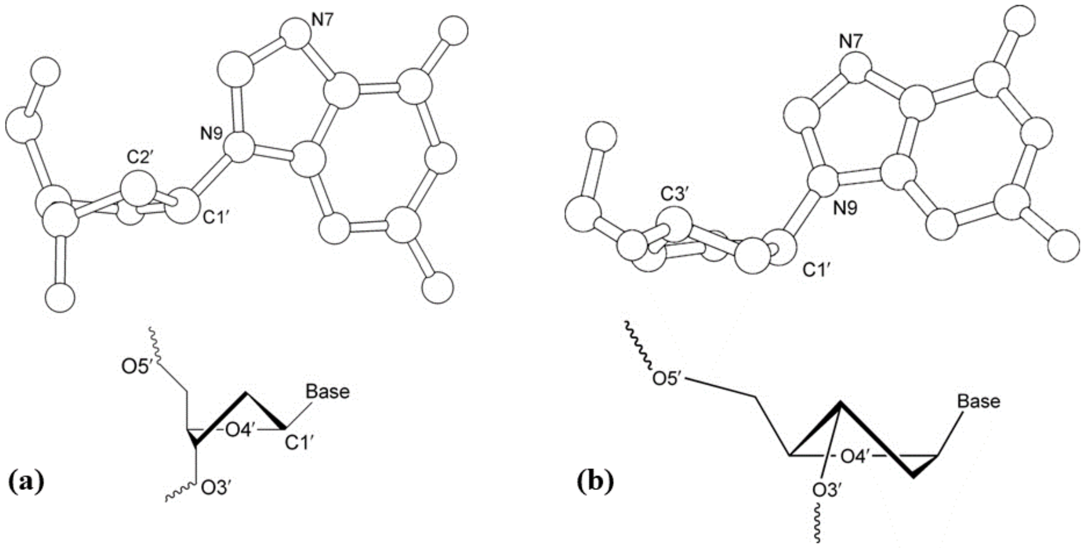
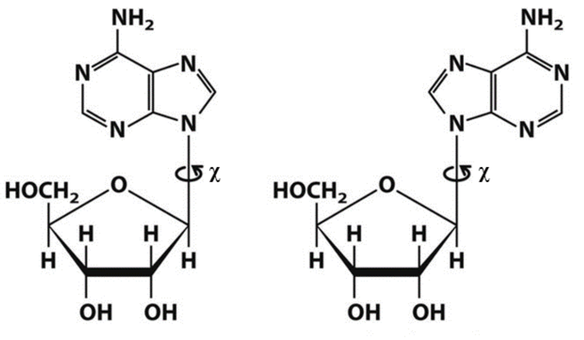


Fig. 4 (a) *Guanosine nucleoside showing C2′-endo sugar puckering and* ***(b)*** *C3′-endo sugar puckering. Both viewed from the perspective of the sugar ring plane. Adapted from ref.* [35].

The glycosidic bond, which connects a deoxyribose sugar and a base (through C1′–N9 bond for purines or C1′–N1 bond for pyrimidines), can technically adopt numerous orientations due to the torsion angle χ. However, structural restrictions commonly lead to specific patterns [43]. This bond's torsion angles are designated by the four atoms: O4′–C1′–N9–C4 for purines, and O4′–C1′–N1–C2 for pyrimidines. According to theory and experimental results, two main low-energy domains exist for the glycosidic angle: *anti* and *syn* conformations [44]. In the *anti* conformation, the Watson–Crick hydrogen-bonding groups are directed away from the sugar ring, positioning hydrogen-bonding groups of the bases outward. In contrast, the *syn* conformation reverses these orientations, bringing these hydrogen-bonding groups closer to the sugar, particularly its O5′ atom (**Fig. 4**). Purine nucleosides are stabilized through hydrogen bonds between the O5′ atom and the N3 base atom. In general, purines slightly favor *anti*-conformation due to lower steric hindrance. However, guanosine-containing nucleotides show a slight preference for the *syn* form because of the positive electrostatic interactions between the N2 exocyclic amine group of guanine and the 5′

phosphate atom. Pyrimidine nucleotides, on the other hand, favour the *anti* conformation over the *syn* because of the negative repulsions between the O2 oxygen atom of the base and the 5′-phosphate group [45].



syn-Adenosine

anti-Adenosine

Fig. 5 *An adenosine nucleoside in syn conformation (on the left) and in anti conformation (on the right).*

The phosphodiester backbone of an oligonucleotide introduces six flexible torsion angles, in addition to the five internal sugar rotations (τ0–τ4) and the glycosidic angle (χ). Physical constraints limit these backbone angles to certain ranges [46,47]. Out of these various possible conformations, only a few have been observed in DNA oligonucleotides and polynucleotides. This is likely due to the constraints imposed by Watson–Crick base pairing on backbone structures when DNA is in double strand.

The convention for defining these backbone angles is as follows: ∼60° is designated as gauche+(g+), −60° as gauche − (g−), and ∼180° as trans (t). For the two torsion angles around the phosphate group (α and ζ), only the g−g− conformation allows successive nucleotide units to arrange their bases for potential hydrogen bonding with a second nucleotide strand. This is the typical phosphate conformation for both DNA and RNA double helices [48].

The torsion angle β, which rotates around the O5′-C5′ bond, is almost always found in the trans conformation. The γ angle, primarily exhibits the g+ conformation in right-handed oligonucleotide and polynucleotide double helices [49].

The torsion angle δ, around the C4′–C3′ bond, varies in relation to the sugar ring's pucker. For C2′-endo and C3′-endo puckers, the internal ring torsion angle τ3, around the same bond, is about 35° and 40°, respectively. δ is approximately 75° for C3′-endo and about 150° for C2′-endo puckers [50].

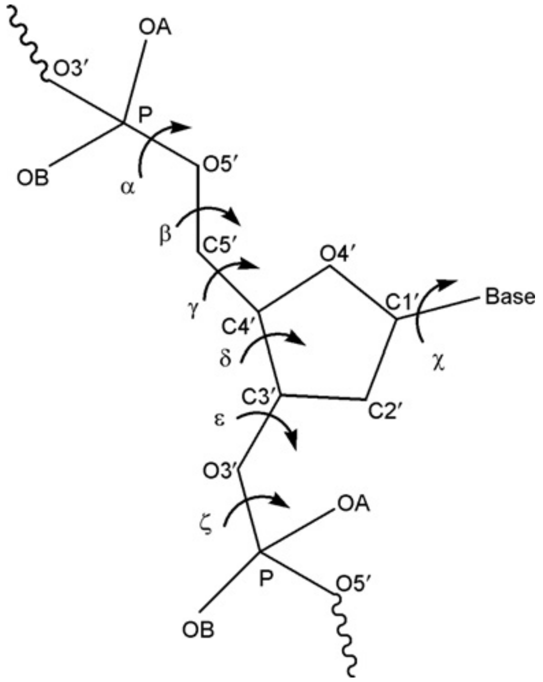


Fig. 6 *Representation of a nucleotide, the curved arrows represent the rotatable bonds of the backbone torsion angles. Adapted from ref.* [35].

1.1.1. DNA structures and its polymorphism

The understanding that planar bases can engage in specific interactions via hydrogen bonds was a pivotal advancement in revealing the structure of DNA.

The significant experimental findings by Chargaff demonstrated that in DNA, the molar ratios of adenine to thymine, and cytosine to guanine, were both equal to one, paired with X-ray diffraction information from the DNA molecule itself, enabled Watson, Crick and Franklin to introduce a double-stranded DNA model in the early 1950s [51,52].

The double helix model of DNA (**Fig. 7**), proposed by James Watson and Francis Crick, is characterized by several key structural features:

1. **Two antiparallel strands**: The DNA molecule consists of two strands that run in opposite directions. Each strand has a 5' end (the end with a free phosphate group) and a 3' end (the end with a free hydroxyl group). In the double helix, one strand runs in the 5' to 3' direction, while the other runs in the 3' to 5' direction.
2. **Base pairing**: The two strands are held together by hydrogen bonds between pairs of nitrogenous bases. Adenine (A) pairs with thymine (T) via two hydrogen bonds, and guanine (G) pairs with cytosine (C) via three hydrogen bonds (**Fig. 7**). This specific pairing is known as Watson-Crick base pairing.
3. **Right-handed helix**: When viewed from above, the DNA strands wind around each other in a clockwise direction, forming a right-handed helix.
4. **Major and minor grooves**: As the strands twist around each other, they create grooves or spaces between the strands. The wider grooves are known as major grooves, and the narrower ones are known as minor grooves.
5. **Backbone and bases**: Each DNA strand has a backbone made of deoxyribose sugar and phosphate groups, with the nitrogenous bases projecting towards the interior of the helix. The backbone is on the exterior of the helix, and it is the phosphodiester bonds in the backbone that give DNA its overall strength and stability.
6. **10 base pairs per turn**: In the Watson-Crick model, there are approximately 10.5 base pairs for every full 360-degree turn of the helix. The distance between each base pair (rise) is about 3.4 Å, leading to a helical repeat (pitch) of approximately 34 Å.

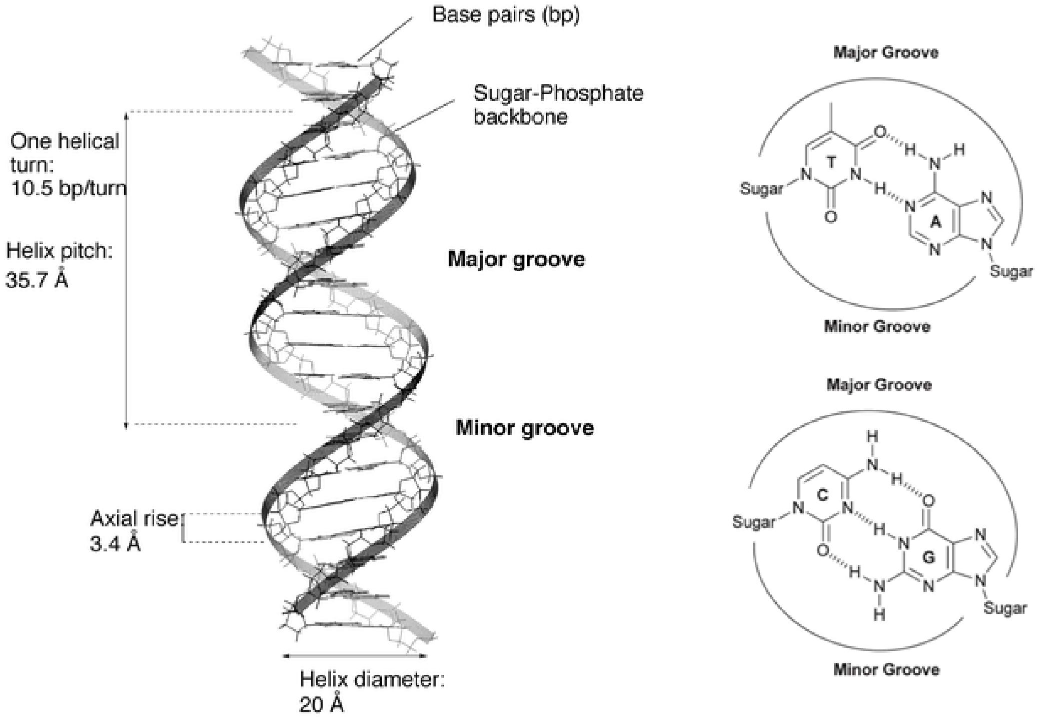


Fig. *7 On the left the double helix model of Watson and Crick. On the right a representation of Watson–Crick hydrogen bonding between adenine and thymine and between guanine and cytosine in the major and minor grooves. Adapted from ref.* [53].

The double helix structure described above, also known as B-DNA, is the most common form for DNA at physiological conditions, it has the glycosidic bond in anti-conformation and the sugar pucker in the C2' conformation. However DNA can adopt other two main conformations with antiparallel strands: A-DNA and Z-DNA (**Fig. 8**).

A-DNA, is another right-handed structure which can be formed from the B-DNA under dehydrating conditions or in the presence of alcohol. This alternative form is more condensed and shorter than the typical B-form. The A-DNA consisting of more than 11 base pairs per turn, which is a larger number than the approximately 10 found in B-DNA. This results in a shorter distance between each base pair, leading to the more compact overall structure. A unique feature of the A-form is its central axial hole, which is not seen in B-DNA. In terms of orientation, the base pairs are inclined with respect to the helical axis, differing from their perpendicular alignment in the B-form. A further distinction of A-DNA lies in the characteristics of its grooves: the major groove is deep and narrow, while the minor groove is wide and shallow. Finally, the sugar moiety in the backbone of A-DNA adopts a C3' endo conformation, but maintaining an anti-glycoside [54,55].

Z-DNA is an atypical lefthanded double helical conformation, can be formed in vitro under specific conditions such as at high salt concentrations, or in vivo in the presence of DNA supercoiling, Z-DNA-binding proteins, and base modifications [56]. The designation 'Z' is derived from the zigzag appearance of the sugar-phosphate backbone when viewed from the side (**Fig. 8**). This is a consequence of a repeating pattern of alternating *syn* and *anti* conformations around the glycosidic bond, corresponding to the alternating purine and pyrimidine bases. In terms of base pairs per helical turn, Z-DNA typically comprises around 12 base pairs per turn, which is more than in B-DNA or A-DNA.

Z-DNA is distinguished by a deep, narrow minor groove and no noticeable major groove due to the irregular alignment of the strands due to the sugar in the C3' endo conformation for purines and C2’ endo for pyrimidines. The overall helix is more slim than B-DNA or A-DNA, with a diameter of approximately 18 Å [57,58].

In **Table 1** are shown the main helix parameters for the three major conformations of double stranded DNA described above.

***Table 1** Main helix parameters. Adapted from ref.* [53].

Parameter	A-DNA	B-DNA	Z-DNA
Helix sense	right-handed	right-handed	left-handed
Residue per turn	11	10.5	11.6
Axial rise	2.55 Å	3.4 Å	3.7 Å
Helix pitch	28 Å	36 Å	45 Å
Rotation per residue	32.7°	36°	-9°, -51°
Diameter of helix	23 Å	20 Å	18 Å
Glycosyl angle	anti	anti	C/G: anti/syn
Sugar pucker	C3’-endo	C2’-endo	C/G: C2’-endo/C3’-endo
Major groove	Narrow, deep	Wide, deep	Flattened
Minor groove	Wide, shallow	Narrow, deep	Narrow, deep

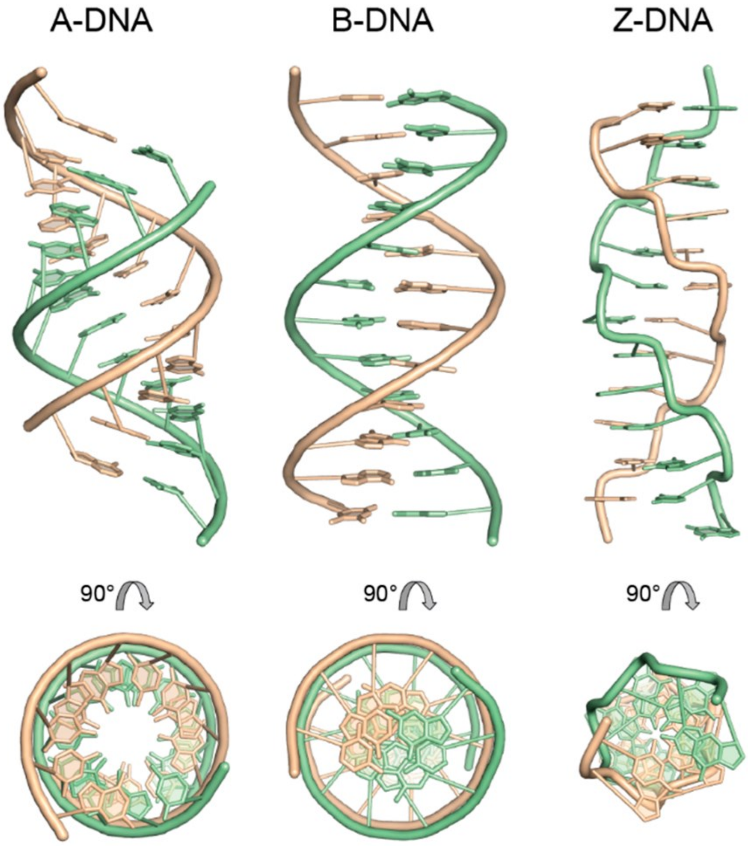


Fig. 8 *Molecular modelling representations of double helix forms of DNA based on X-ray fiber diffraction analysis. The antiparallel strands are represented with different colors, and the sugar-phosphate backbone is drawn as smooth tube. At the bottom the orthogonal view of the respective double helix forms. Adapted from ref* [59].

Over the past few decades, the availability of high-resolution X-ray and precise NMR data has elucidated the concept of DNA polymorphism at a molecular scale, greatly expanding our understanding of DNA architecture. This understanding extends far beyond the traditional representation of DNA structure, with the revelation that DNA polymorphism is heavily dependent on sequence contexts and environmental conditions (ionic strength, pressure, temperature, ligands, etc) giving rise to the major conformations (A, B, Z, etc) and several sub-variants (A, A', C, D, T, BI/BII, ZI/ZII, etc). Additionally, other conformations can be adopted such as cruciforms, slipped loops, three-stranded triple helices (H-DNA), and even unique four-stranded structures known as G-quadruplexes [60–64] (**Fig. 9**). Notably, H-DNA represents an additional structural potential of DNA. The deviation from the traditional double helix is accompanied by the emergence of an alternate type of base-pair in H-DNA, where the third strand of the triple-stranded complex forms

Hoogsteen base pairs with a Watson–Crick paired double strand situated within the major groove of the latter [61]. Double-stranded sequences, capable of forming a G-quadruplex, produce another structure, termed the i-motif, from the C-rich strand [65]. In the quadruplex structure, the four strands are interlinked by Hoogsteen base pairing, whereas the i-motif reports a cytosine–cytosine$^+$ base pairs.

Thus, due to its polymorphic capabilities, DNA can assume a considerable variety of different conformations. Specific structures and biological functionality, such as the i-motif and G-quadruplex, which will be examined in greater detail later in this thesis.

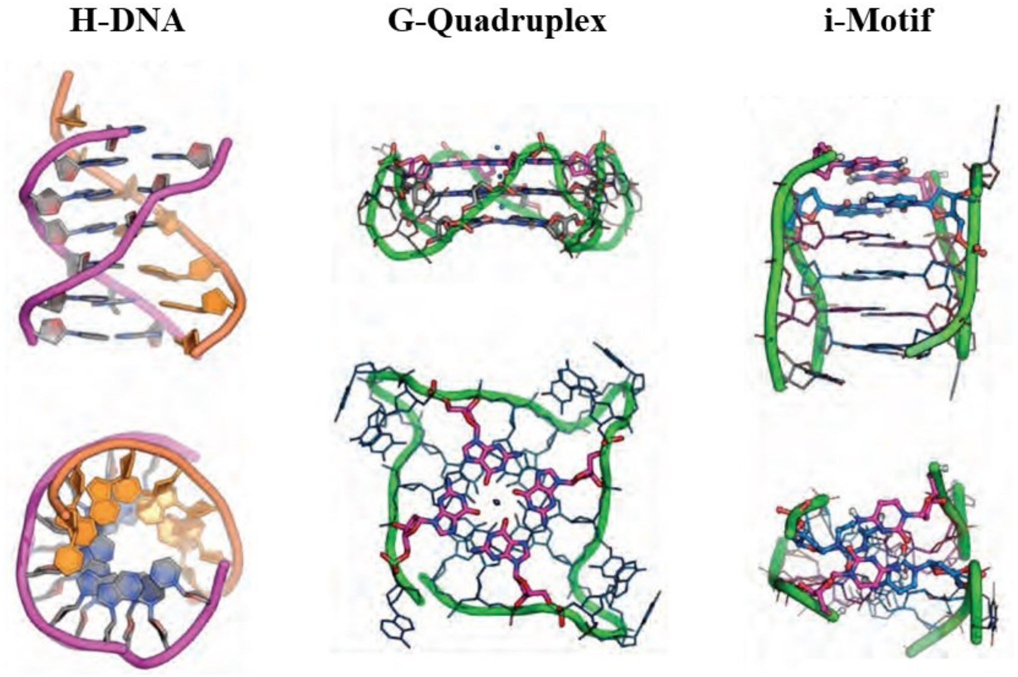


Fig. 9 Three- and four-stranded structures of DNA. Viewed along (top) and down (bottom) their helical axes. Adapted from [66].

1.1.2. RNA structures

In a manner similar to DNA, RNA is a biopolymer that consists of a phosphate backbone. Its fundamental units, termed ribonucleotides, comprise a sugar and a nitrogenous base, but these differ from DNA nucleotides in two critical chemical aspects.

Firstly, the ribose sugar units of RNA possess a hydroxyl group at the 2’ position, a feature absent in the deoxyribose sugars of DNA. This additional hydroxyl group profoundly impacts the chemical structure of RNA, rendering its sugars significantly more rigid than those of DNA and sterically forcing them into a C3’-endo pucker configuration (**Fig. 10**). Furthermore, the 2' hydroxyl group increases RNA's capacity for diverse interactions with itself and a plethora of ligands, as well as facilitating its hydrolysis and cleavage [35].

Secondly, the thymine bases in DNA are substituted by uracils in RNA (**Fig. 1**), which employ identical base-pairing strategies but lack the methyl group at the 5' position.

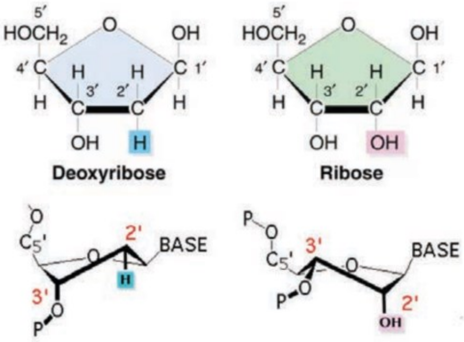


Fig. 10 *Chemical structure differences between deoxyribose and ribose, at the bottom the related sugar configurations.*

Generally, RNA is a single-stranded polymer, but its structure is far from being linear. The ribonucleotides have a tendency to interact amongst themselves in two primary ways: either through stacking via their flat faces or by forming hydrogen bonds along their edges. Isosteric base pairs analogous to the canonical Watson-Crick pairs in DNA can be formed, where A pairs with U, and G with C. However, RNA bases possess an additional isosteric pairing mode, known as the G-U wobble pair. The G-U pair exhibits comparable stability to the A-U pair and is frequently observed in RNA structures [67,68].

These three most prevalent base-pairing modes in RNA maintain the overall dimensions of consecutive base pairs and also locate the ribose group of both bases on the same side of the pair. This alignment allows for the identification of a major and minor groove in the RNA structure. The C3'-endo pucker configuration of the ribose determines that the RNA helix will assemble into the

A-form (**Fig. 11**), with 11 base pairs per turn, a deep narrow major groove, and a relatively shallow minor groove. On an intermediate level, often referred to as its secondary structure, the basic structural element of an RNA sequence is the double helix [69].

Once the helices are determined, the unpaired regions among them can be classified into several types of structural elements, collectively referred to as loops. These loops can be internal, positioned between two helix stems (an internal loop with one side is known as a bulge), or hairpin loops, comprising several unpaired bases that are bordered on each side by the same helix, instead pseudoknot occurs when there is a base-pair interaction among loop. Additionally, there are multi-branched loops located at the intersection of three or more helix stems [70,71] (**Fig. 11**).

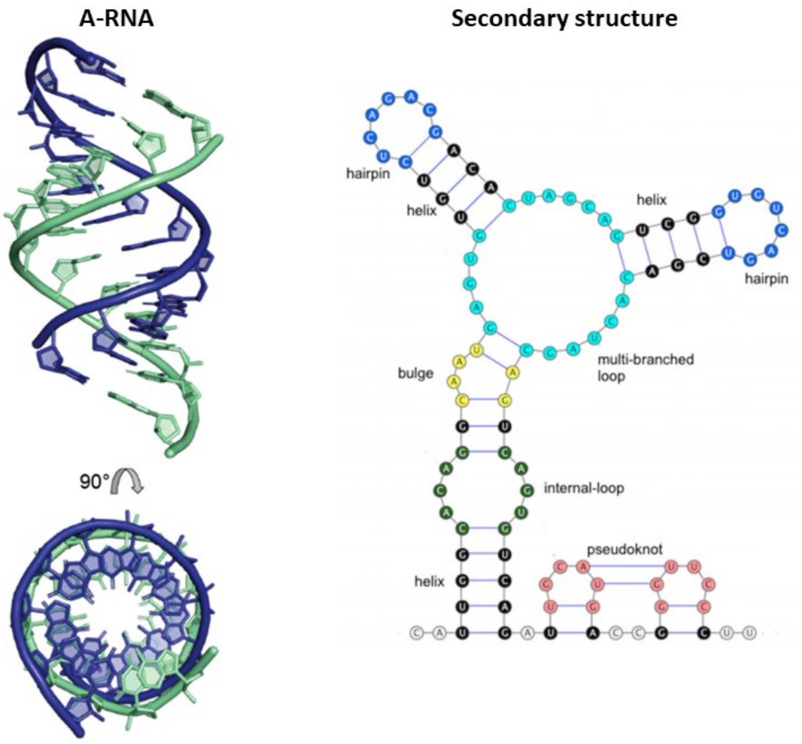


Fig. 11 *On the left: A-RNA structure based on single-crystal X-ray diffraction analysis. At the bottom the orthogonal view (PDB: 1SDR). Adapted from* [59]. *On the right:* examples of RNA secondary structures with 5 different types of loops. Adapted from [70].

Upon binding RNA can assume a variety of tertiary motifs to enhance base stacking, achieved through the co-axial stacking of adjacent helices and other structures, thus stabilizing its 3D motif [72]. Furthermore, RNA 3D motifs are shaped via interactions between secondary structural motifs, including kissing loops, pseudoknots, kink turns, tetraloops, and G-quadruplexes (**Fig. 12**). The K-turn motif causes a sharp kink in the helical axis, which results in the convergence of the minor groove sides of its two supporting helices. Another prevalent RNA motif is the pseudoknot, where a single-stranded region of a hairpin loop base pairs with sequences upstream or downstream within the same RNA strand. The kissing loops motif is formed when the single-stranded loop regions of two hairpins base pair with each other [73].

Of particular importance is a small fold RNA motif known as the three-way junction, the most common type of multi-branched loop. This motif serves as a structural scaffold, converging three helical stems that are connected by a maximum of three single-stranded segments. Although this element can also be represented in the secondary structure where it appears open and unpaired, it is the tertiary structure of this motif, established by non-canonical base interactions, that dictates its distinctive topology [74,75].

Repeated stretches of guanine-rich sequences can form quadruple base pairs that stack on each other to form G-quadruplexes.

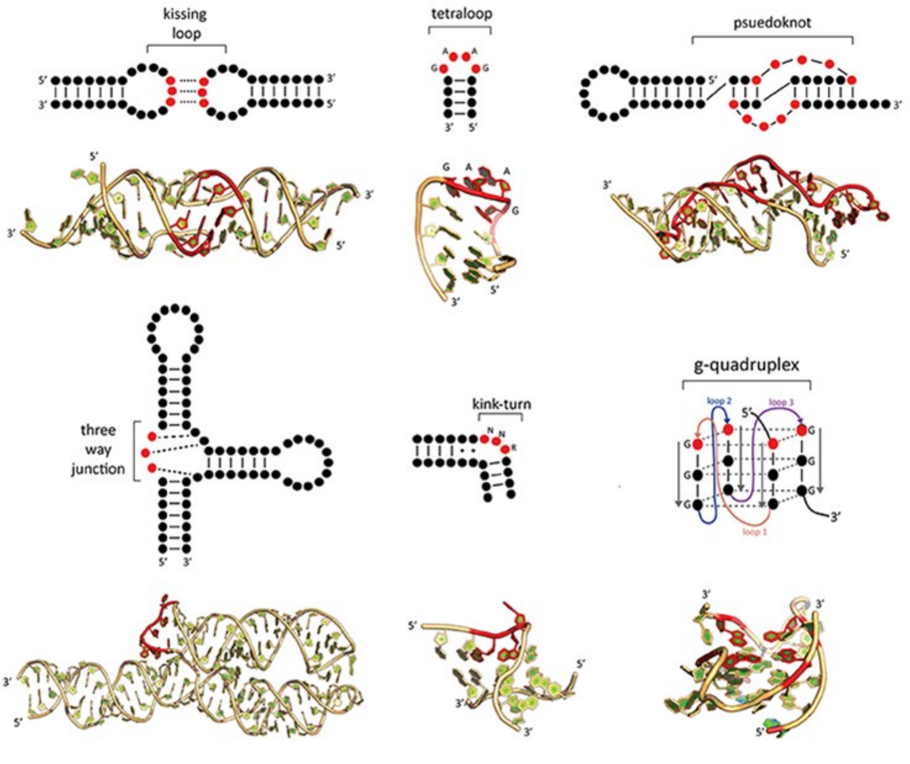


Fig. 12 2D representations of common RNA secondary and tertiary structural motifs with 3D examples of crystal structures. Adapted from [73].

Tertiary folded RNAs interact with proteins, ligands, and other RNA molecules, causing biochemical reactions that influence all aspects of cellular metabolism [76]. Hence, comprehending the molecular characteristics of RNA tertiary structure remains as a fundamental objective in biochemistry. As we enhance our understanding of the intricacies of RNA tertiary structure, it will allow us to predict the structure of RNA molecules based on their sequences. This can offer key insights into their biological roles. Given recent findings that a large portion of our genome is transcribed into RNA with unknown functions, understanding RNA structure becomes all the more crucial [77].

In the pursuit of this thesis, the objective was to illuminate the structural characteristics of select RNA and DNA sequences in solution. To achieve this, a variety of spectroscopic methodologies were employed. Moreover, we utilized a range of achiral porphyrins, with the intent of exploiting them as chiroptical probes. The next paragraphs will explore more comprehensively into the characteristics and function of these porphyrins.

1.2. Porphyrins

Tetrapyrroles denote a significant achievement in chemical evolution. They are similar to other primordial biomolecular units in that the intricacies of their original abiogenesis remain a mystery. They were elegantly realized an estimated four billion years ago. However, it is undeniable that tetrapyrroles such as porphyrins, chlorins, and corroles are ubiquitous and valuable molecular entities.

Porphyrin is derived from the Greek word *πορφύρα* (*porphura*, which means purple) and refers to a fascinating class of macrocycle rings that are engaged in a wide range of biological processes ranging from oxygen transport to photosynthesis and catalysis [78]. Porphyrins and related compounds have accordingly been termed *the pigments of life* [79], chlorophylls and hemes are two significant examples in nature. The former function as both light gathering antennas and charge separation reaction mechanisms in photosynthesis. These are important components of biocatalysts and oxygen transporters in the blood [80].

The porphyrinic macrocyclic structure (**Fig. 13 (a)**) is formed by four pyrrole rings linked by methine units, this macrocycle is a planar and aromatic ring containing 22-π electrons of which only 18 of them are delocalized. As result of this extensive conjugation, porphyrins are intensely coloured and they possess a strong absorption in the visible region of the UV-Vis spectrum. Additionally, the large π-electron system is responsible for numerous properties of porphyrins including electronic [80], optical [81,82], mechanical [83], and chemical [84] properties.

The internal nitrogen atoms produce a core pocket that is perfectly positioned to firmly integrate metal atoms in a tetradentate fashion (Fe, Co, or Mg are common in biological systems), when coordination occurs, two protons are removed from the pyrrole nitrogen atoms, leaving two negative charges [85]. These anionic centres are coordinatively unsaturated units for charge transfer and ligation of adducts, which occurs with reversible changes of electronic configuration, such as oxidation or spin states [78]. Therefore, at the center of the porphyrin core, porphyrins can form several coordinate covalent complexes with transition metals (metalloporphyrins) and some non-metals. Furthermore, the peripheral substituents at α, β, and meso positions (**Fig. 13 (b)**) can be functionalized to produce tuneable molecular and crystal properties like solubility, reactivity, conductivity, and photophysics. Metalloporphyrins can react with ligands to generate a plethora of (porphyrin)metal-ligand complexes [86], which are also useful for many applications [87].

Thanks to their peculiar characteristics porphyrins have been extensively studied and used in a wide range of applications which include photoimmunotherapy, photo diagnosis [88], biosensors [89], cancer therapy, photo-catalysis [90], solar cells [91], chemical sensors [92], optoelectronics [93], chiroptical probes [94] and so forth.

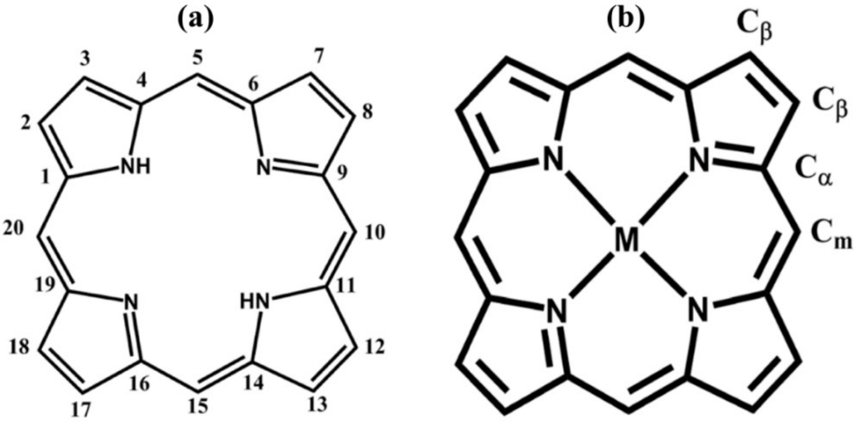


Fig. 13 *Chemical structures of* ***(a)*** *freebase porphyrin and* ***(b)*** *metalloporphyrin. Alpha position (C_α) at 1, 4, 6, 9, 11, 14, 16, 19; beta position (C_β) at 2, 3, 7, 8, 12, 13, 17, 18; and meso position (C_m) at 5, 10, 15, 20 represent three unique peripheral substituent positions. Adapted from* [87].

It is evident that the presence of substituents and/or metals in the porphyrin ring has a significant impact on the spectroscopic behaviour [95,96]. The high aromatic character of porphyrins and their derivatives is responsible for the high absorption properties in the UV-Vis region, due mostly to π-π^* (HOMO - LUMO) electronic transitions [96,97].

Free-base porphyrin UV-Vis spectra (**Fig. 14 (a)**) typically show an intense, narrow absorption band ($\varepsilon > 200000$ L mol^{-1} cm^{-1}) at around 400 nm, known as the Soret or B band, followed by four longer wavelength (450–700 nm), weaker absorptions ($\varepsilon > 20000$ L mol^{-1} cm^{-1}) referred to as the Q bands [96,97].

The intense Soret-band is caused by the second excited transition state ($S_0 \rightarrow S_2$), whereas the weak Q-(*quasi*-allowed) bands are caused by the first excited transition state ($S_0 \rightarrow S_1$). The dissipation of energy via internal conversion (IC) is so rapid that fluorescence is only observed from depopulation of the first excited singlet state to the lower-energy ground state ($S_1 \rightarrow S_0$) (**Fig. 14 (b)**).

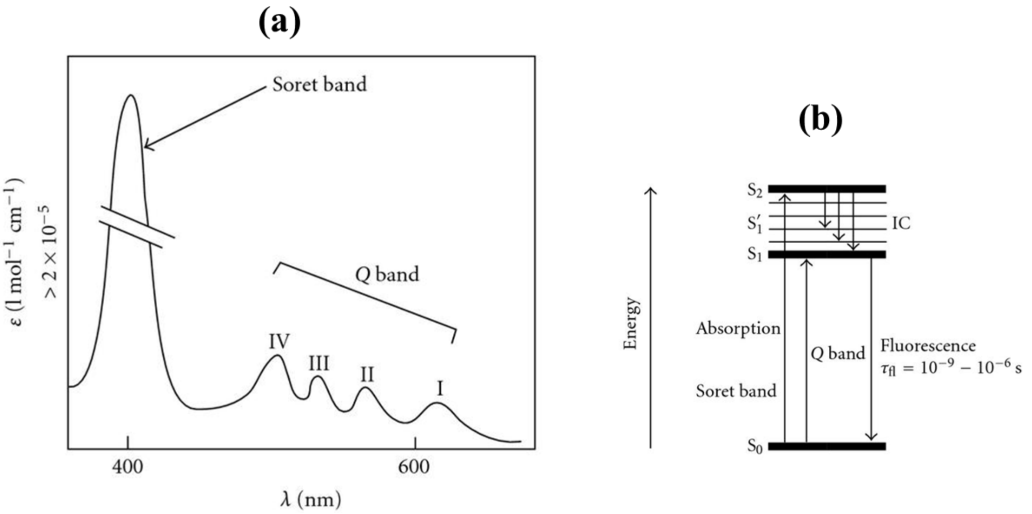


***Fig. 14 (a)** Typical freebase porphyrin absorption spectrum and **(b)** modified Jablonski energy diagram.*

In the 1960s, Martin Gouterman proposed a semi-quantitative Four Orbital Model to explain the spectra of porphyrins [98,99], by this theory the absorption bands are caused by the splitting of the four frontier molecular orbital ($HOMO_{-1}$, HOMO, LUMO, and $LUMO_{+1}$).

His theory states that the absorption bands in a porphyrin system result from a transition between two HOMOs and two LUMOs. The HOMOs were found to be almost degenerate a_{1u} and a_{2u} orbitals, whereas the LUMOs were found to be a degenerate set of e_g orbitals (**Fig. 15 (a)**). Transitions between these orbitals produced two excited states, both of which ^1E_u character.

Orbital mixing splits these two states in energy by a configuration interaction into two pairs of degenerate orbitals, creating a higher energy ^1E_u state with greater oscillator strength, giving rise to the Soret band, and a lower energy ^1E_u state with less oscillator strength, giving rise to the Q-bands (**Fig. 15 (b)**) [100].

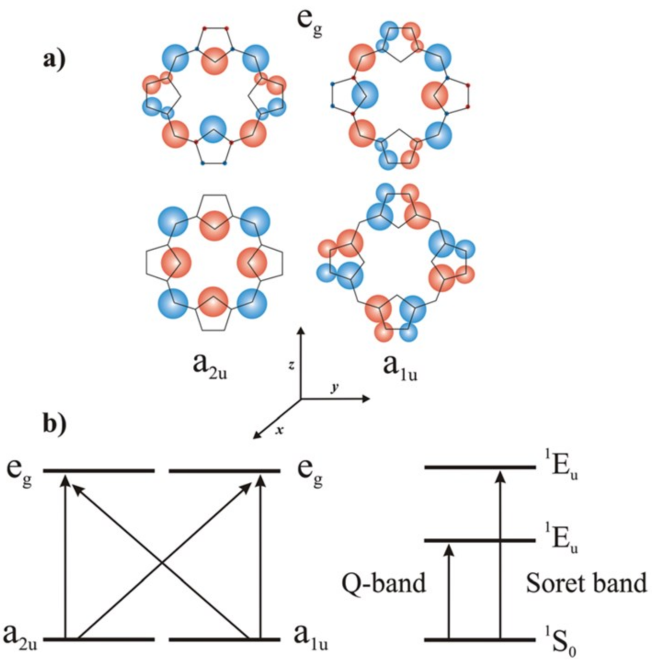


***Fig. 15 (a)** Gouterman's four orbital model and **(b)** energy levels showing the transitions of a porphyrin system.*

The relative energy of these transitions are affected by the identities of the metal core and the substituents on the ring.

For a metalloporphyrin, the d orbitals of a transition metal are not suited to direct interactions with the porphyrin π frontier orbitals, the main effect of the metal on the spectrum results from conjugation of its empty p_z orbital with the π electrons of the rings, which involves the occupied frontier orbital with a_{2u} symmetry. In spite of this modest influence on the spectrum, the inclusion of a metal atom in the porphyrin macrocycle can generate strong electrostatic and charge transfer (CT) interactions [101].

Based on their UV-vis and fluorescence characteristics, metalloporphyrins can be divided into two classes [102]. *Regular* metalloporphyrins comprise closed-shell metal ions (d^0 or d^{10}) such as Zn^{II}, with relatively low energy dπ (d_{xz}, d_{yz}) metal-based orbitals. These have little influence on the porphyrin π to π* energy gap electronic spectra.

Metalloporphyrins with transition metal with partially-filled d orbitals (d^m m= 6-9) metals are known as *hypsoporphyrins*. There is considerable metal dπ to porphyrin π* orbital interaction (metal to ligand backbonding) in *hypsoporphyrins*. As a result, the porphyrin π to π^* energy separation increases, and a blue shift (i.e., hypsochromic effect) in the Soret band is revealed.

Another important difference for a metalloporphyrin is that in general in the visible spectrum only two Q-bands are observed. This is a reflection of the molecular symmetry. The molecular symmetry shifts from rectangular D_{2h} to square D_{4h} in going from freebase-porphyrin to metalloporphyrins [101]. Indeed, in metalloporphyrins (D_{4h}) the two axes defined by opposite pyrrole nitrogen atoms become equivalent, as a result of the increased symmetry both LUMO e_{gx} and e_{gy} levels are degenerated.[103].

Regarding the fluorescence of these macrocycle, free base porphyrins and metalloporphyrins have two emission zones, the first of which is centered between 400 and 500 nm and corresponds to the higher energy $S_2 \rightarrow S_0$ transitions, and the second between 550 and 800 nm and corresponds to the lower energy $S_1 \rightarrow S_0$ transitions. However, many variables make observing the $S_2 \rightarrow S_0$ emission bands challenging (poor quantum relaxation/return rate, internal conversions, spectral overlap, etc.). The majority of the time, excitation of these chromophores results in the observation of only the $S_1 \rightarrow S_0$ emission bands, which are the most commonly exploited in porphyrinic systems [104].

Since these very peculiar spectroscopic characteristics porphyrins have found application in many scientific and technology fields, including molecular recognition [105] and new chiral catalysts for asymmetric synthesis [106], as well as to investigate the secondary structure of different biomolecules in aqueous solution, such as oligopeptides, proteins [107–113] polynucleotides [114–119], and so forth [120,121]. The application of porphyrins as probes for studying polynucleotides is a particular focus and will be discussed in detail in the following paragraph.

1.2.1. Porphyrins as Chiroptical Probes in the Structural Investigation of Polynucleotides

Fiel et al. in 1979 [122] and Fiel and Munson in 1980 [123] published the first extensive description of nucleic acid interactions with tetrakis(4-N-methylpyridyl) porphyrin (TMPyP or H2T4) and several of its metal derivatives (M = Fe, Co, Ni, Zn, Sn, Mn).

In 1983, Pasternack et al. [124] provided a detailed explanation of the structural properties of metalloporphyrins, describing that the binding process with polynucleotides is affected by the type of the metal ion, as well as the size and location of the substituent groups on the porphyrins. Many parameters influence DNA-porphyrin interactions, including the number of charges on the porphyrin and peripheral substituents, solvent medium, ionic strength, and porphyrin/DNA ratio, in particular by changing the metal core and peripheral substituents, the binding modes can be easily modified [125].

The interaction between a porphyrin and DNA can occur via a variety of binding mechanisms (**Fig. 16**), including intercalation, binding to minor and major grooves (outside self-stacking), external complex formation (outside random binding), and binding to DNA G-quadruplexes [126]. Fiel and colleagues [127] refer to the first three as a "three-mode binding model".

The intercalation occurs when a fused-ring aromatic system is introduced into the space between two base pairs, intercalators are able to unwind the supercoiled DNA and lengthen short DNA sections [127].

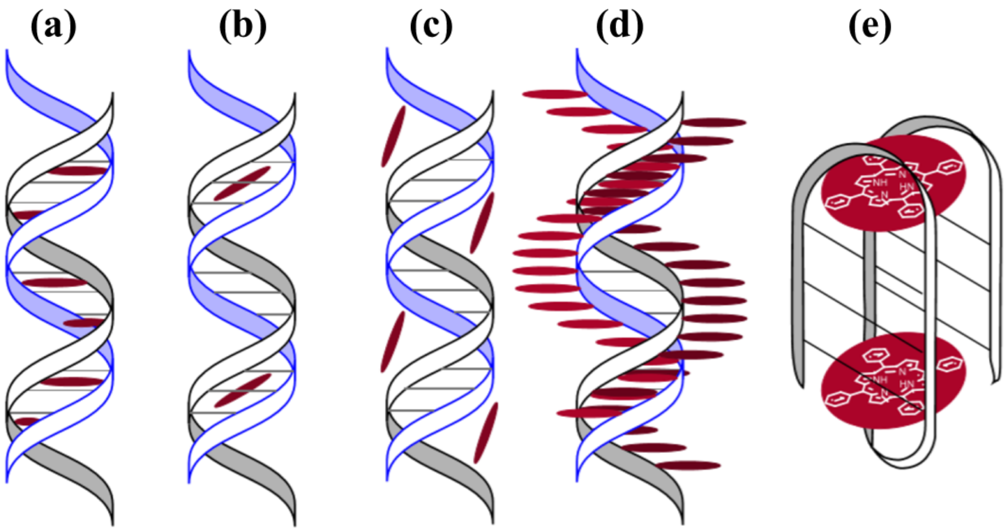


Fig. 16 *Schemes of various binding modes of porphyrin to DNA.* ***(a)*** *Intercalation,* ***(b)*** *binding to a minor groove of DNA,* ***(c)*** *binding to a major groove of DNA,* ***(d)*** *external binding with self-stacking along the DNA surface, and* ***(e)*** *binding to a G-quadruplex. Adapted from* [128].

The groove binding usually involves electrostatic interactions between a cationic ligand and the negatively charged phosphate backbone of the DNA. Generally, adenine-thymine (A-T) base pairs form porphyrin complexes at the minor grooves, and guanine-cytosine (G-C) base pairs may be implicated in the formation of intercalated complexes with porphyrin. It was found that the ionic strength and porphyrin to base-pair ratio affect the competitive binding of intercalation versus outside binding.

Selective binding in major or minor grooves is determined by a number of factors, including the shape of the cation, the DNA's local electric field, van der Waals forces, possible hydrogen-bonding interactions with atoms on the DNA molecule, and hydrophobic effects [129].

The way porphyrin interacts with G-quadruplex (G4) is influenced by numerous elements such as the type of G4, the DNA sequence, and the configuration of strand orientations. Other features, like the groove sizes of different antiparallel G-quadruplexes, the structures of the loops, and nearby non-G-tetrads, may also affect this interaction [130]. Three interaction modes have been suggested: (i) attaching at both the top and bottom of the quadruplex, known as capping [131], (ii) binding externally on the strands of G-quadruplexes [132], or (iii) sliding or intercalating between the guanine tetrads within a quadruplex [133]. Each of these interaction mechanisms has been found to increase the stability of the G4 structure. Particularly, the porphyrin's capping function at the G4's end results in a notable enhancement in the G4's stability.

In relation to targeting, RNA has many attractive properties similar to those of DNA. Both forms of nucleic acids adopt a standard helix shape with major and minor grooves.

However, the A-form of RNA's typical helix is frequently interrupted by regions of mismatched or unpaired bases, which changes the secondary structure of the RNA **(Fig. 11)**. These structural motifs in RNA are substantially smaller than the minor and large grooves of DNA due to the reduced helix pitch and inclination of the nucleotides. According to Hermann and co-workers [134,135] the A-shape helix's major groove only has a diameter of 4 Å, which prevents the binding of porphyrins. Nevertheless, as with DNA, porphyrins can interact with duplex RNA to produce two different types of complexes: intercalation and external binding with self-stacking.

The exploration of cationic H2T4 porphyrin interaction with RNA duplexes was reported for the first time in 1997 [136]. An analogous method, which involves the cationic porphyrins' interaction with RNA G-quadruplexes, has been utilized to evaluate the possibility of using porphyrins in photodynamic therapy [137].

In addition, it was reported [138] that metalloporphyrins form stronger complex with RNA duplexes compared to the naked porphyrins.

Generally, DNA or RNA probes are formed by aromatic chromophores with the potential to intercalate with nucleic bases via π–π interactions, among them porphyrins are an excellent choice because of their unique properties:

i) an highly electronic system, which leads to an intense absorption band in the 380–450 nm region, far from UV-region in which most of biomolecules absorb. Hence the intense absorption greatly enhances the sensitivity of CD;

ii) the possibility to easy functionalize them at the meso-positions with specific groups or charged moieties in order to confer unique features in the whole macrocycle (e.g. to increase the water solubility);

iii) their ability to coordinate a metal in the central core, such as zinc and magnesium porphyrin provide extra stereo-differentiation with their Lewis acid sites for binding OH, NH_2, COO^-, and other functional groups as Lewis bases [94];

iv) when non-chiral substituents are present, porphyrins are achiral molecules therefore they do not display any chiroptical signal. However, after interaction, with chiral molecules, an induced circular dichroism (ICD) signal could be observed in the Soret absorption band region which is very diagnostic for the type of interaction [139];

v) their capacity to behave as photosensitizers in the presence of oxygen in order to employ them in photodynamic therapy (PDT).

Due to their distinct photochemical and photophysical characteristics, a number of research over the past 20 years have used free base or metalated porphyrins to covalently and non-covalently insert them into a DNA scaffold. Thus, the goals of these studies have been to investigate the impact of porphyrins on DNA in order to use them as spectroscopic sensors in DNA conformational studies as well as their stability and electronic properties using circular dichroism (CD), CD-melting, UV-Vis, fluorescence spectroscopy and resonance light scattering (RLS) [140]. Because interactions between these macrocycles and polynucleotides—which are conformationally chiral—induce a dichroic band (ICD) in the absorption region of the achiral ligands—CD measurements are, in general, more precise. A diagnostic of the sort of interactions is provided by the shape and intensity of ICD.

Every spectroscopic technique aforementioned give an important information about the type of binding of a ligand with polynucleotides and like a piece of a puzzle, information has to be put together to get a bigger view.

In general the information obtained by each technique in this specific field can be summarized as follows:

i) UV-vis provides quick, intuitive information regarding the nature of binding, but not often in great depth. Aggregation and intercalation lead to hypochromicity of the Soret band, which is more intense for aggregation than for intercalation, and typically to a red shift of about 10 nm or more.

ii) CD measurements are, in general, more specific because interactions of these macrocycles with polynucleotides (which are conformationally chiral) induce a dichroic band (ICD) in the absorption region of the achiral ligands. The shape and intensity of ICD is diagnostic of the type of interaction (e.g., external binding, intercalation, or aggregation).

iii) Fluorescence can be used to distinguish between intercalation and external binding. In theory, the first binding mode may cause an increase in emission intensity because vibrational freedom is reduced, but the second type typically causes an emission quenching. Similarly, aggregation leads to extensive quenching of the macrocycle emission band [141].

iv) RLS, which is particularly specific in identifying aggregation, can resolve the contrast between external binding and aggregation.

In conclusion we expect for intercalation: a red shift of the Soret band and some hypochromicity, some fluorescence quenching, no RLS signal, and a negative ICD; for external binding: a red shift of the Soret band, a hypochromic effect smaller than that observed for intercalation, no variations of fluorescence (or a small increase in the intensity owing to the loss of vibrational freedom degrees), no RLS signal, and a positive ICD or no signal; for aggregation: a red or blue shift (for J- and H-aggregates, respectively) and broadening of the Soret band accompanied by extended hypochromicity, heavy fluorescence quenching, the presence of RLS signal, whose intensity is proportional to the size of the aggregate, and a bisignate ICD [142].

2. Micro-RNAs

MicroRNAs (miRNAs) are a class of highly conserved, short (18-24 nucleotides) and non-coding RNAs that regulate gene expression by base-pairing to one or more mRNA targets, causing either target degradation or translational repression [143]. The distinctive miRNA sequences are involved in many physiological and pathological processes. MiRNAs play important roles in immunity response [144], hematopoiesis, developmental timing, cell death, cell proliferation and patterning of the nervous system [145]. Dysregulation of miRNAs contributes to the initiation and development of human diseases like cancer, genetic disorders, cardiovascular diseases and altered immune system functions [146–148]. Recent insight into the roles of miRNAs have made them attractive tools and targets for novel therapeutic approaches [149].

miRNA biogenesis

The biogenesis of miRNA (**Fig. 17**) begins in the nucleus with the transcription of primary miRNA (pri-miRNA) by multiprotein complex RNA polymerase II [150]. Pri-miRNAs are 1–3 Kb long RNA hairpin structures possessing a poly-A tail at the 3′ end and a cap at the 5′ end. An enzymatic complex, containing RNAse III Drosha and the double-strand binding protein, DGCR8, specifically recognizes and converts pri-miRNA to 60–70 nucleotides precursor miRNA (pre-miRNA) [151]. Then, the exportin 5–RAN•GTP complex transports pre-miRNA to the cytoplasm. There, another enzyme the RNase III Dicer and TAR RNA binding protein (TRBP) bind to the pre-miRNAs and removes the terminal loop, resulting in a mature miRNA duplex [149]. Dicer action results in the generation of 18–24 base pair duplex RNA with guide and passenger strands. In the next step, the mature miRNA duplex is incorporated into the RNA-induced silencing complex (RISC). The process of unwinding and strand selection of the miRNA duplex is mediated by the argonaute (AGO) family of proteins. Binding of the mature miRNA to RISC leads to the targeting of mRNAs with complementary sites and results in translational repression or mRNA degradation [149].

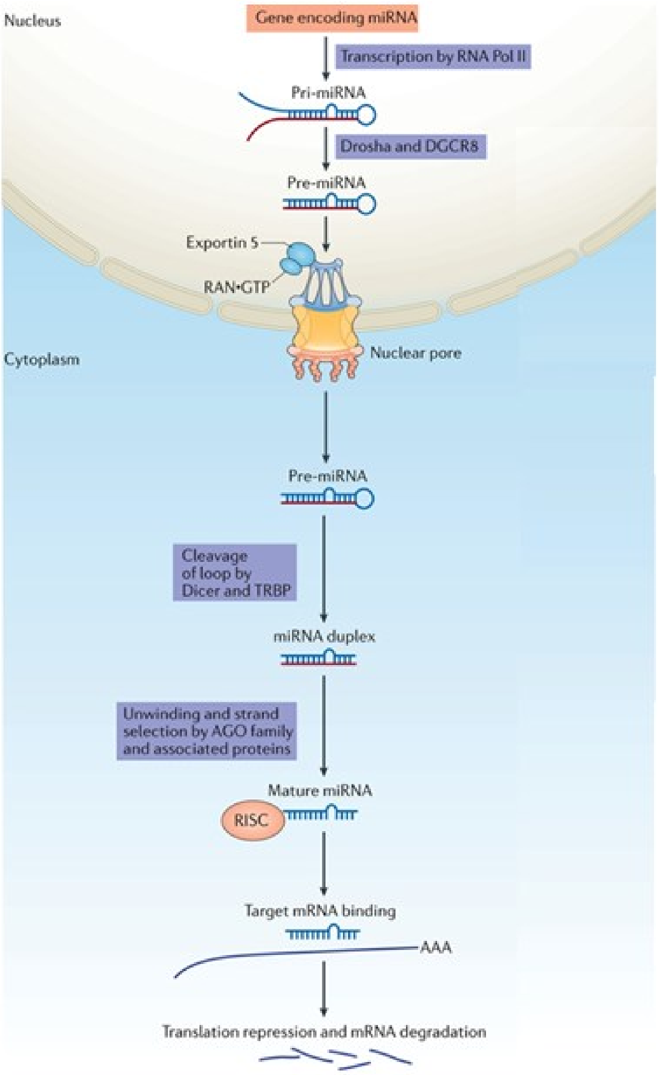


***Fig. 17** Illustration of the main steps in biogenesis of miRNA. Adapted from* [149].

Based on the canonical processing mechanism, both strands derived from the mature miRNA duplex can be loaded into the Argonaute (AGO) family of proteins (AGO1-4 in humans) in an ATP-dependent manner [152]. For any given miRNA, the proportion of AGO-loaded 5p or 3p strand varies depending on the cell type or cellular environment, it is changing from near equal amounts to predominantly one or the other [153]. The selection of the 5p or 3p strand is based in part on the thermodynamic stability at the 5′ ends of the miRNA duplex or a 5′ U at nucleotide position 1 [154]. Generally, the strand with lower 5′ stability or 5′ uracil is preferentially loaded into

AGO and is the guide strand. The unloaded strand is called the passenger strand, which will be unwound from the guide strand through various mechanisms based on the degree of complementarity (**Fig. 18**) [155].

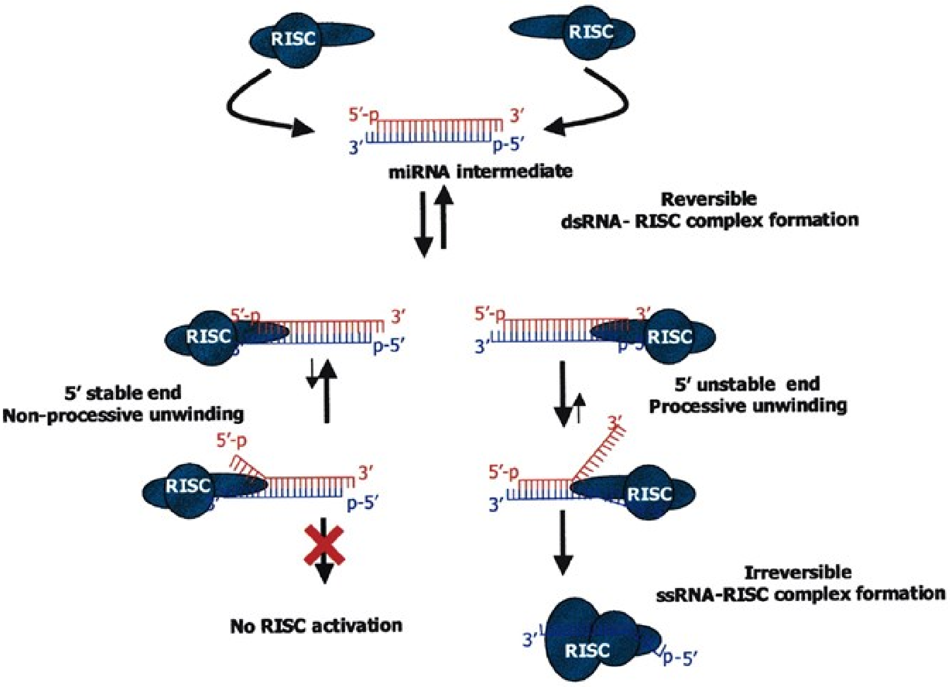


***Fig. 18** Proposed model in which RISC forms reversible complexes with either 5' end of the miRNA duplex. The enhanced flexibility of the 5' terminus facilitates the directional unwinding of the duplex from the 5' end to promote the formation of the more stable activated RISC. In this way, differential stability of the two ends of the miRNA duplex predetermines which strand is selectively loaded into the RISC. Adapted from* [154].

Usually, in plants miRNAs have near-perfect pairing with their mRNA targets, on the contrary in animal cells miRNAs are able to recognize their target mRNAs by using approximately 6–8 nucleotides (the seed region) at the 5' end of the miRNA (**Fig. 19**) [156]. It has been reported that in mammalians the specificity is not limited to the seed region, but the stabilization of miRNA-mRNA interaction can also involve the residues in position 13-16 from the 5’ end of miRNA [157]. Furthermore, a single mature miRNA may have hundreds of different mRNA targets, and a given target might be regulated by multiple miRNAs [158,159].

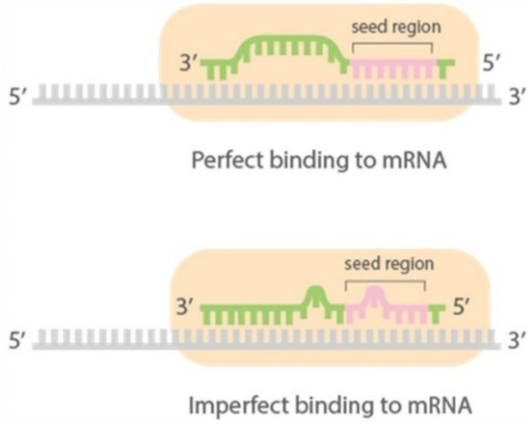


Fig. 19 *Illustration of the perfect binding (top) and the imperfect binding (bottom) between a miRNA and its mRNA target. In gray is represented the mRNA, in green the miRNA and in pink the seed region of the miRNA.*

miRNAs in cancer

The important role of miRNAs in regulating physiological processes and carcinogenesis have represented a revolutionary discovery in the last decade. Dysregulation of miRNA is now thought to affect critical molecular signaling pathways involved in tumor progression, invasion, angiogenesis, and metastasis in a variety of cancers. MiRNAs, based on their altered expression, can be divided into two categories: oncogenic miRNAs (oncomiRs) and Tumor Suppressor (TS) miRNAs. The upregulation of oncomiRs has been observed in several tumor cell lines. These miRNAs promote tumorigenesis by blocking the translation of mRNAs encoding TS proteins [160]. On the contrary, TS miRNAs are down-expressed in many cancer cell lines [161]. These miRNAs generally repress the translation of mRNAs encoding oncoproteins. Dysregulated expression of miRNAs in tumor cells may result from epigenetic alterations, mutations, copy number changes, and/or aberrant miRNA processing. Many in vitro and in vivo studies have shown that defective or suppressed expression of miRNA can promote tumorigenesis [162]. Tumor cells arise from numerous mutations in the genome that interfere with complicated genetic processes involving multiple genes and molecular pathways. Therefore, abnormal expression of miRNAs could be a reliable indicator for early diagnosis and patient prognosis. Moreover, the versatility of miRNAs can be effectively exploited for therapeutic strategies in cancer. For example, reducing oncomiRs, frequently overexpressed in human cancers, allows for the restoration of target tumor-suppressor expression. One of the major advantages of miRNA therapy is its large adaptability. It offers the possibility of silencing multiple mRNA targets with a single miRNA or targeting a single mRNA simultaneously with multiple miRNA sequences [163,164]. Regulating miRNA expression in the tumor microenvironment can be done in a number of ways. OncomiR suppression utilizing synthetic anti-miRNAs or antagonists, or locked nucleic acids (LNAs), such as anti-sense oligonucleotides that completely match mature oncomirs, has been extensively documented [164]. Restoration of decreased TS miRNAs to normal levels, also known as microRNA replacement therapy, is another miRNA therapy regimen that has been well-tested in pre-clinical studies [165].

Circulating miRNAs and their potential as cancer biomarkers

Several studies have revealed that miRNAs can be released into extracellular fluids, such as plasma and serum, cerebrospinal fluid, saliva, breast milk, urine, tears, colostrum, peritoneal fluid, bronchial lavage, seminal fluid, and ovarian follicular fluid [166]. In contrast to cellular RNA, circulating miRNAs (or extracellular miRNAs) showed great stability and resistance to high temperature and extreme variation in pH [167]. Circulating miRNAs can be divided into two groups. One can be found in vesicles such as microvesicles, exosomes, and apoptotic bodies while the other is associated with proteins, generally AGO2. The role of vesicles or proteins is to protect circulating miRNAs to the enzymatic degradation and increase their stability in the extracellular environment. Circulating miRNAs play an important role in intercellular communication and regulate host cell activity. It has been demonstrated that extracellular miRNAs secreted by donor cells can be transferred into other recipient cells and function as endocrine signals [168], furthermore some miRNAs can interact with cell surface receptors, such as Toll-like receptors. Nevertheless, the mechanisms by which miRNAs are secreted and taken up by cells are not well understood and require further investigation.

Circulating miRNAs can be used as biomarkers for a variety of diseases, since it was shown that the miRNA profile is altered in blood and in body fluids of patients compared with that of healthy persons [169]. Circulating miRNAs are becoming candidates of emerging non-invasive cellular and molecular biomarkers especially in cancer, since numerous studies have recognized that dysregulation of extracellular miRNAs is associated with the origin, progression, therapeutic response, and survival of patients [170]. Noteworthy, a study conducted in 2008 by Rosenfeld et al. [171] demonstrated that a microarray of only 48 selected miRNAs could trace and classify 90% of primary tumor in metastatic samples. Nevertheless, there are still some technical challenges to overcome before clinical application, first the isolation and purification of the samples since the integrity and purity of RNA are the basis for detection and quantification. Secondly, the difficulty to accurately measure circulating miRNAs due to their low concentration, currently the main methods to detect the quantity of circulating miRNAs are qPCR, microarray, and next-generation sequencing [170].

2.1. MicroRNA secondary structures: State-of-the-art

In contrast to the canonical process that all miRNAs are associated with Argonaute (Ago) proteins throughout RISC loading, Janas et al. [172] recently discovered that miRNAs are much more expressed in cells than Agos, and some miRNAs are stable without binding to Argonaute proteins, suggesting that miRNAs not involved in the RISC complex may exist as free molecules fulfilling their biological activity.

It is still unknown through which mechanism miRNAs, in the absence of Ago binding, can be stabilized and how they reach their targets, but a great contribution was given by Belter et al. [173]. They demonstrated that mature miRNAs could form highly ordered structures similar to aptamers and concluded that miRNA structural versatility may determine a variety of functions other than the well-known mechanism of mRNA recognition and degradation. Indeed, miRNAs, like RNA aptamers, can be induced to bind with proteins and hence influence their activity directly, suggesting that miRNA secondary structures may have functions outside of the RISC, both in terms of sequence and structure.

Also Purrello et al. [174] explored the link between mature miRNA secondary structures and biological activity using a variety of spectroscopic and non-spectroscopic techniques. This hypothesis was built upon a previous study by Croce and Ferro[175], they exploited the ability of endogenous miRNAs to target multiple gene sites. They aimed to create the bioinformatics tool "miR-Synth" to design artificial microRNAs (a-miR) that could target various genes at multiple sites. The tool ranked the designed miRNAs based on predicted repression efficiency and tested its validity by measuring the silencing efficiency of single and double-stranded miRNAs against c-MET and EGFR, two lung cancer-associated genes.

However, two of the top-six ranked miRNAs displayed by "miR-Synth" (a-miR-23 and a-miR-98) did not show significant inhibition of c-MET and EGFR expression, suggesting that these artificial miRNAs do not exhibit effective anticancer activity. Probably because "miR-Synth" does not consider potential secondary structures adopted by the miRNA sequences, which are increasingly recognized as critical factors influencing miRNA function [176].

Based on this premise, Purrello et al. decided to explore the existence of secondary structures potentially linked to the activity of these sequences using a range of techniques, such as electronic circular dichroism (CD), differential scanning calorimetry (DSC), nuclear magnetic resonance (1H-NMR), and non-denaturing polyacrylamide gel electrophoresis (PAGE).

In particular, they performed a structural investigation on two highly-ranked artificial miRNAs, which significantly inhibited the luciferase activity for both c-MET and EGFR (a-miR-141 and a-miR-196) and they investigated the structures of two a-miRNAs that appeared to be inactive (a-miR-23 and a-miR-98) for comparison. Lastly, they conducted similar experiments with two endogenous miRNA sequences (mir-15a and mir-15b) that have the same seed sequence, but different biological functions.

CD experiments suggested that mature miRNAs investigated fold into ordered secondary structures, like self-dimers or hairpins, and variations in CD spectra among different miRNA sequences proposed that different secondary structures were adopted by the different miRNA sequences. This was further confirmed by CD melting experiments. Other techniques, including ^{1}H NMR and differential scanning calorimetry (DSC), highlighted the presence and stability of different secondary structures among the studied miRNAs. Furthermore, the nucleases resistance PAGE experiment showed that miRNAs arranged in highly ordered secondary structures were more resistant to endo- and exo-nucleases, suggesting that these miRNAs could have a longer half-life and higher likelihood of reaching their target in a biological environment.

Importantly, the study investigated two endogenous miRNAs, miR-15a and miR-15b, which have the same seed sequence but different biological activities. miR-15a is related to the pathogenesis of chronic lymphocytic leukemia, while miR-15b is associated with metastasis progression. The authors found that these two sequences folded into distinct ordered secondary structures, suggesting a link between their structure and biological functions.

These findings indicate that the sequence of natural and artificial miRNA is not the only parameter to consider when predicting their biological function. Indeed, spectroscopic evidence indicates they fold into various three-dimensional structure. These configurations could significantly impact their function and specificity toward targets, especially within a biological context.

2.2. miR-337-3p bat vs human: aim of the work

Building upon the recent insights provided by Purrello et al. [174] about the importance of miRNA secondary and tertiary structures in modulating their biological functions, our project takes a step forward in the exploration of these structure-function relationships. It has been established that miRNAs can adopt different secondary structures when they exist as single-stranded molecules in solution, leading us to speculate that the stability of these conformations may be linked to biological functions. This might include reduced degradation, extended half-life, and more stable interaction with target molecules.

In the first part of my thesis work we focused on miRNA sequences that display only few base differences, but regulate different targets, to show that not only the sequence determines the specific function of miRNAs.

In this context the miR-337-3p sequence lends itself as the perfect case study to investigate the possible secondary structures and the thermodynamic stability of miRNA conformations as this microRNA is variant in bats compared to all other mammals (**Fig. 20**) [177]. Interestingly, the seed region of bat and human miR-337-3p differs by only two nucleotides, but they are predicted to regulate different sets of gene targets. In bats, these targets are associated with developmental, rhythmic, synaptic, and behavioural gene pathways [177], while in humans, miR-337-3p is associated with various cancers, often acting as a tumour suppressor [178,179].

The bat variant of miR-337-3p (bat-miR-337-3p) exhibits both higher and more widespread expression than its human counterpart (hsa-miR-337-3p) and shows stronger repression when binding to a fully complementary target suggesting a prominent biological role [177].

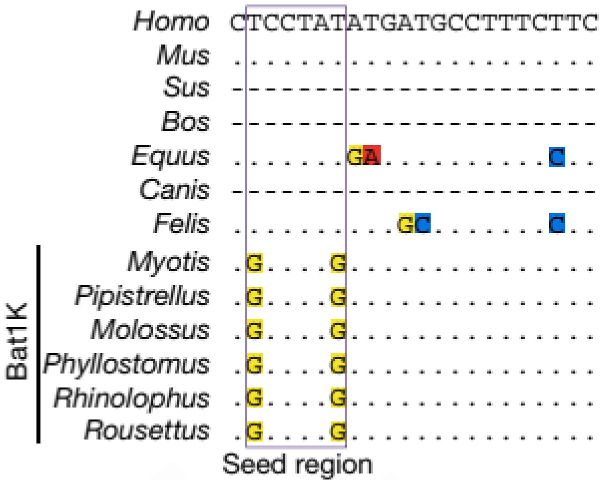


***Fig. 20** Alignment of miR-337-3p among several mammals. Dots represent bases identical to human; dashes represent species in which a functional miR-337-3p could not be identified. Adapted from ref* [177].

The first aim of this work is to characterize by using spectroscopic techniques such as CD and CD-melting the possible secondary structures adopted by hsa-miR-337-3p and bat-miR-337-3p, and their thermodynamic stability. Our hypothesis is that a stronger secondary structure with a higher melting temperature and therefore higher thermodynamic stability can increase the targeting efficiency. In addition, it has been observed that hsa-miR-337-3p has also another isoform which is more expressed (~67% vs ~28%) [180] and differs for only one nucleotide deletion at the 5' end. In this context conformation and stability of the two isoforms were investigated, in order to see if the most expressed isoform (hsa-miR-337-3p_2) is also the most stable in solution.

Secondly, to confirm our first hypothesis we decided to manipulate, by changing as few nucleotides as possible, the hsa-miR-337-3p_1 sequence to increase the thermodynamic stability of its secondary structures without changing the seed region, in order to evaluate if the target efficiency in vitro is increased (e.g. using fully complementary sensors). Conversely, the bat-miR-337-3p sequence was destabilized in the same way, outside of the seed region, in order to reduce the targeting efficiency. The mutated sequences (mut-hsa-miR-337-3p_1 and mut-bat-miR-337-3p) were designed in silico by using the "oligoanalyzer tool" provided by Integrated DNA Technologies IDT®, which permits to predict the secondary structures of a given sequence (e.g. hairpins or self-dimers) and their theoretical melting temperature (Tm). Finally, the mutated sequences designed in silico were purchased from IDT® and characterized by CD and CD-melting experiments and the results were compared with that of the wild-type sequences.

2.2.1. Materials and methods

All the miRNAs (miRs) sequences used were purchased from Integrated DNA Technologies IDT® and used without further purification. Each solid was dissolved in ultra-pure water obtained by Elga Purelab Flex system by Veolia with purity of 18.2 MΩcm, achieving stock solutions with concentration of ~100 µM.

Then, by dilution in PBS buffer 10 mM ([KCl] 2.7 mM; [NaCl] 137 mM; pH 7.4) we prepared work and sample solutions.

Concentration of miRs solutions were checked by UV-Vis measurements using the extinction coefficient for each sequence given by IDT: bat-miR-337-3p $\varepsilon_{260\ \mathrm{nm}}$ = 203,400 L/(mol·cm); hsa-miR-337-3p_1 $\varepsilon_{260\ \mathrm{nm}}$ = 203,200 L/(mol·cm); hsa-miR-337-3p_2 $\varepsilon_{260\ \mathrm{nm}}$ = 196,900 L/(mol·cm); mut-hsa-miR-337-3p_1 $\varepsilon_{260\ \mathrm{nm}}$ = 206,900 L/(mol·cm); mut-bat-miR-337-3p $\varepsilon_{260\ \mathrm{nm}}$ = 213,100 L/(mol·cm).

The sequences of miRNAs used are the following (in **bold** the differences between bat-miR-337-3p and hsa-miR-337-3p_1, in red the difference between bat-miR-337-3p and mut-bat-miR-337-3p and in blue the difference between hsa-miR-337-3p and mut-hsa-miR-337-3p).

bat-miR-337-3p: 5’-C**G**CCUA**G**AUGAUG**C**CUU**U**CUUC-3’
hsa-miR-337-3p_1: 5’-C**U**CCUA**U**AUGAUGCCUUUCUUC-3’
hsa-miR-337-3p_2: 5’-_UCCUAUAUGAUGCCUUUCUUC-3’
mut-hsa-miR-337-3p_1: 5’-CUCCUAUAUGAUGCCUUUCAUC-3’
mut-bat-miR-337-3p: 5’-CGCCUAGAUGAUG**A**CUU**A**CUUC-3’

By using the “oligoanalyzer tool” provided by IDT® we calculated the secondary structures likely adopted by each sequence and their respective theoretical melting temperature.

Electronic Circular Dichroism

miRNA samples were analyzed at 0.5, 2.5, 5 and 10 µM single strand concentration.

CD spectra were recorded at 37°C using a Jasco J-710 spectropolarimeter equipped with a single position Peltier temperature control system. A quartz cuvette with a 1 cm path length was used for all CD experiments. Conditions were as follows: scanning rate 50 nm/min, data pitch 0.5

nm, digital integration time (D.I.T) 2s, band width 2.0 nm. Each CD spectrum was an average of at least five scans.

The CD melting experiments were performed within the temperature range 0–90°C using a temperature heating rate of 1°C/min, monitoring the intensity of the miRNA CD signal at 260 nm.

Melting curves were obtained from the experimental data with the help of a Boltzmann sigmoid equation:

$$Y = \text{bottom} + (\text{top} - \text{bottom})/(1 + e((Tm - x)/\text{slope}))$$

A Boltzmann sigmoid function is characterized by a bottom, a top plateau, and a characteristic value describing the point where the x value is exactly between the top and bottom value, here Tm.

2.2.2. Results and Discussions

Spectroscopic characterization of bat and human miR-337-3p

By using CD and CD-melting we analyze the two sequences hsa-miR-337-3p and bat-miR-337-3p in order to verify if they are arranged in some secondary structures , and in that case to evaluate their thermodynamic stability. The aim was to determine whether the two sequences, differing only in two nucleotides but having distinct gene targets, could assume different structures with different stability.

Firstly, we performed CD measurements of the two miRNA sequences, bat-miR-337-3p and hsa-miR-337-3p_1 at the concentration of 2.5 μM. The CD spectra in **Fig. 21 (a)** shows an intense positive band centered at ~ 270 nm, and a less intense negative band centered near 240 nm, suggesting that the two miRNA sequences are not in random conformations. Indeed, if the sequences were in random conformations the intensity of the signal had to be close to few millidegrees. Furthermore, the intensity of the positive band of the hsa-miR-337-3p_1 results slightly higher and red-shifted than the bat-miR-337-3p. This aspect might suggest that hsa-miR-337-3p_1 organizes in highly ordered secondary structure, differently than bat-miR-337-3p.

Secondly, we performed melting experiments by monitoring the variation of CD signal at 260 nm in the range 0-90 °C (**Fig. 21 (a)**). In general, the loss of ordered structure by heating causes CD changes which, monitored at a fixed wavelength as a function of temperature, allow to estimate the stability of secondary structures.

The melting curves show a clear transition for the two sequences, confirming that these miRNAs adopt well-defined folded structures. In particular, the human sequence has a lower melting temperature (Tm ~ 16°C) than the bat sequence (Tm ~ 30°C). Moreover, if we visualize the whole melting curve (**Fig. 21 (b)**), we can notice that around 37°C (body temperature) the hsa-337-3p_1 sequence is almost completely denatured, whereas the bat-337-3p is almost at the half of the melting event, suggesting that in bats this sequence has a stronger probability to preserve its secondary structures. All these data suggest a possible correlation between the higher stability of secondary structures adopted by the bat sequence and its biological prominent role.

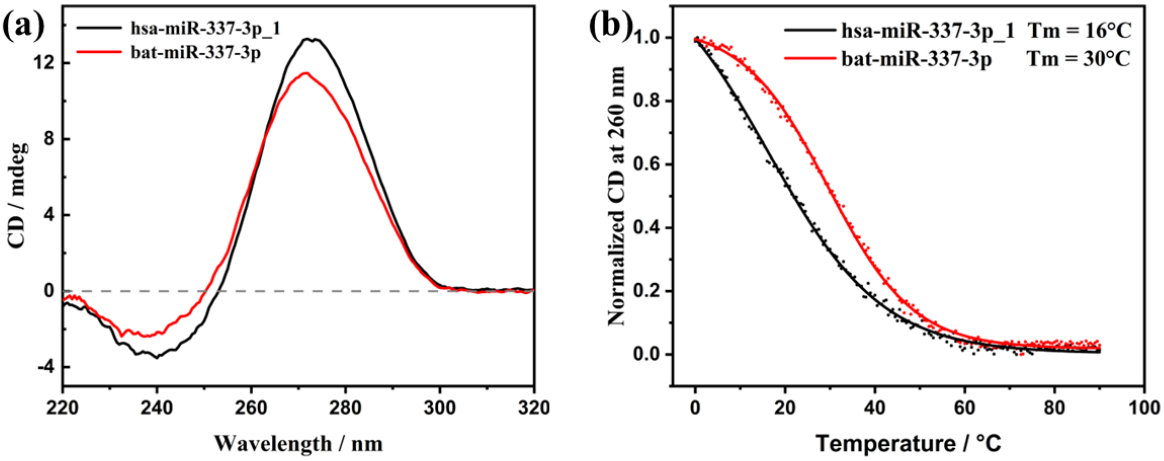


***Fig. 21** CD-spectra **(a)** and normalized CD-melting curves **(b)** of hsa-miR-337-3p_1 [2.5 μM] (black curve) and bat-miR-337-3p [2.5 μM] (red curve) in PBS 10 mM, pH=7.4.*

To further explore the secondary structures assumed by the two sequences, we conducted melting experiments at varying concentrations (0.5 μM, 2.5 μM, 5 μM, and 10 μM) and by Boltzmann equation we obtained the melting temperature (Tm) for each.

It is important to note that monomolecular transitions, such as those linked to the unfolding of hairpins, are virtually independent on strand concentration. In contrast, transitions of higher molecularity, like those related to the dissociation of self-dimers, demonstrate significant dependency on sample concentration.

The Tm values obtained for the human sequence at different concentrations revealed notable differences (**Fig. 22 (a)**). Specifically, the Tm decreased with increasing concentration, suggesting a strong dependence on sample concentration.

In contrast, the CD-melting experiments at various concentrations for the bat sequence (**Fig. 22 (b)**) revealed similar Tm values (~ 30°C) across all tested concentrations, indicating that is not dependent on sample concentration. These observations suggest that the secondary structures of hsa-miR-337-3p_1, due to their concentration dependence, are likely predominantly self-dimers at higher concentrations. Whereas, at lower concentrations, they may exist in multiple conformations (e.g., self-dimers and hairpins). On the other hand, the secondary structures of bat-miR-337-3p, being independent on sample concentration, are likely primarily organized as hairpins, regardless of the concentration examined.

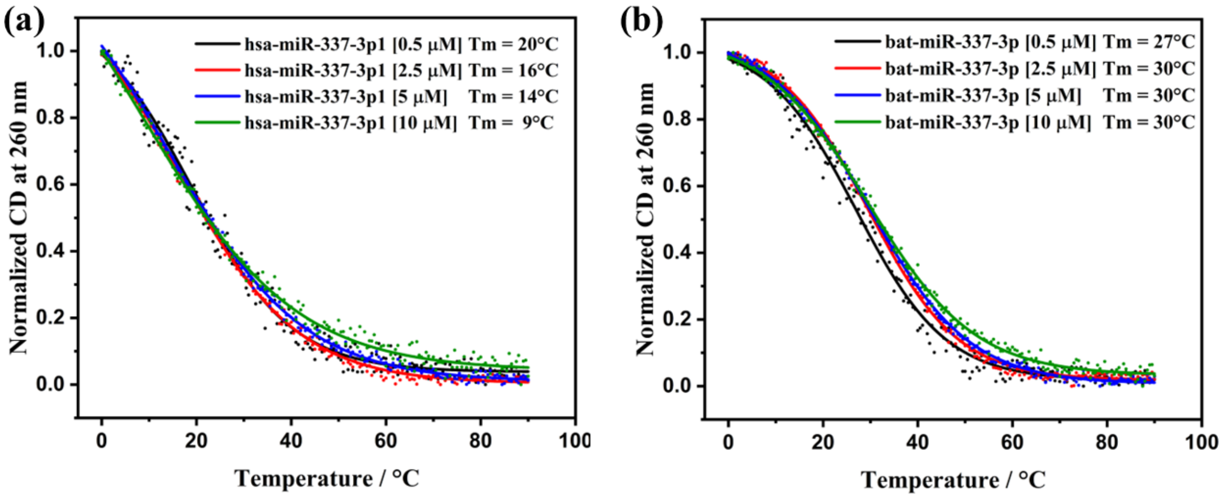


Fig. 22 *Normalized CD-melting curves of hsa-miR-337-3p_1* ***(a)*** *and bat-miR-337-3p* ***(b)*** *at the concentration of 0.5 μM (black curve), 2.5 μM (red curve), 5 μM (blue curve) and 10 μM (green curve) in PBS 10 mM, pH = 7.4. Full circles represent the experimental points.*

Lastly, we used the “oligoanalyzer tool” provided by IDT® which permits to predict the secondary structures (e.g. hairpins or self-dimers) of a given sequence. For hsa-miR-337-3p_1 the IDT tool provided two types of hairpins and two self-dimers with 4 base pairing (**Fig. 23**). Suggesting that for the human sequence by increasing the concentration is possible to promote the formation of self-dimers which are in an equal number of the predicted hairpins. Whereas for the bat sequence the IDT tool provided only one self-dimers with 4 base pairing and three types of hairpins, suggesting that hairpins are more favorite also at higher concentrations.

hsa-miR-337-3p_1

5′-CUCCUAUAUGAUGCCUUUCUUC-3′						
GC %	**Hairpin**	**Self-dimer**				
		6 base	5 base	4 base	3 base	2 base
41%	**2 types**	-	-	**2**	-	8

bat-miR-337-3p

5′-CGCCUAGAUGAUGCCUUUCUUC-3′						
GC %	**Hairpin**	**Self-dimer**				
		6 base	5 base	4 base	3 base	2 base
50%	**3 types**	-	-	**1**	-	11

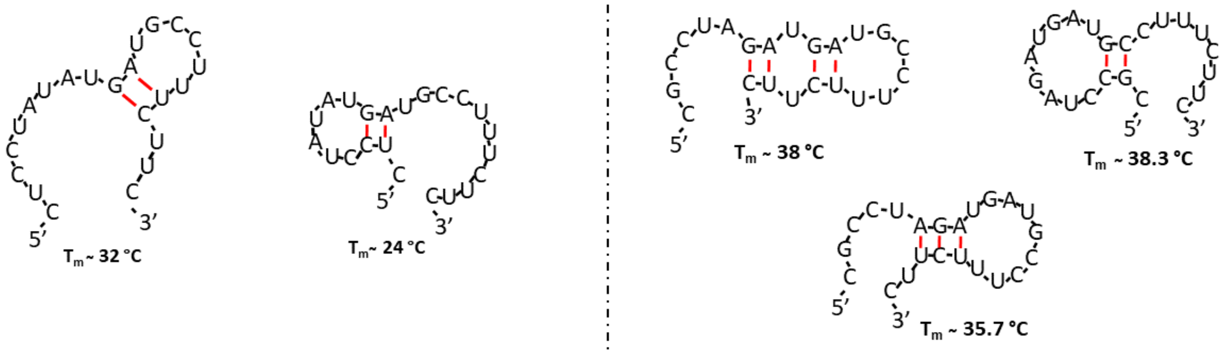


Fig. 23 Tables and cartoon representations of hairpins provided by the IDT® oligoanalyzer tool. Each table includes the percentage of GC content, as well as the number of hairpins and self-dimers, differentiated by base-pair count. On the left, the data refer to hsa-miR-337-3p_1, and on the right, to bat-miR-337-3p. The cartoon representations depict potential hairpin structures along with their respective calculated melting temperatures, red lines represent canonical Watson-Crick base pairings.. Cartoon representations of self-dimers are not depicted.

Isoform of human miR-337-3p

As detailed in **paragraph 2.2**, we conducted CD and CD-melting experiments on the most expressed isoform (hsa-miR-337-3p_2) at a concentration of 2.5 μM to examine if it was also the most stable in solution. The results were compared with those of the least expressed isoform (hsa-miR-337-3p_1) which is slightly longer (22 basis) than hsa-miR-337-3p_2 (21 basis). Indeed hsa-miR-337-3p_2 is misses for the “C” at 5’. The CD spectra (**Fig. 24 (a)**) revealed that the intensity of the positive band around 270 nm for hsa-miR-337-3p_1 was slightly higher than that of hsa-miR-337-3p_2, suggesting a potential variation in the secondary structures assumed by the two isoforms. CD-melting curves (**Fig. 24 (b)**) allowed us to obtain the melting temperatures (Tm) for both sequences. The most expressed isoform showed a higher Tm (~22°C) than the least expressed isoform (~16°C), corroborating our hypothesis that the most expressed sequence tends to be the most stable in solution.

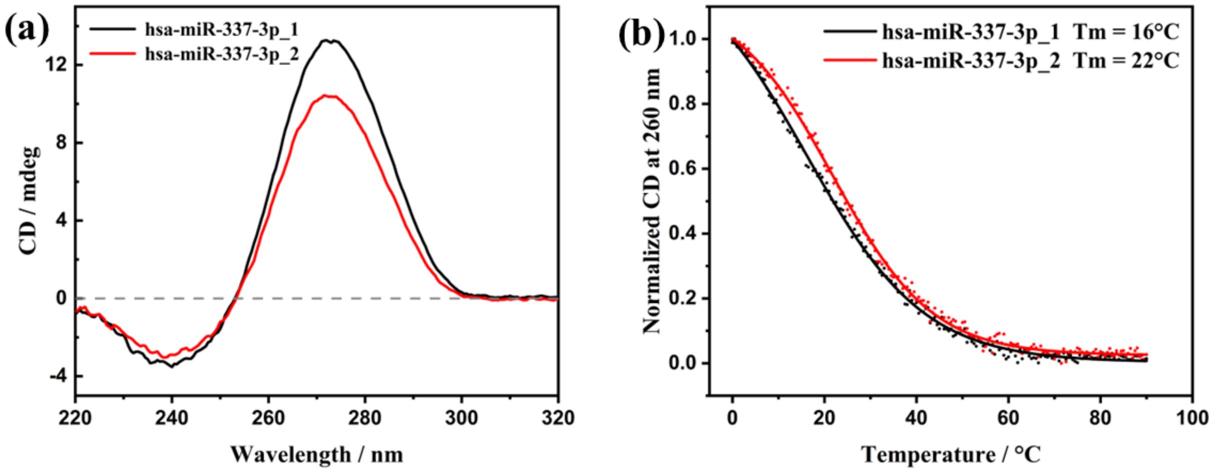


***Fig. 24** CD-spectra **(a)** and normalized CD-melting curves **(b)** of hsa-miR-337-3p_1 [2.5 µM] (black curve) and hsa-miR-337-3p_2 [2.5 µM] (red curve) in PBS 10 mM, pH=7.4.*

We subsequently performed CD-melting experiments at different concentrations for hsa-miR-337-3p_2 and obtained the respective Tm (**Fig. 25**). The Tm remained approximately around 20°C at all concentrations, suggesting that the secondary structures are independent on strand concentration and are likely organized into hairpins, whereas the least expressed miRNA (hsa-miR-337-3p_1), as shown previously, exhibited dependency on strand concentration and at higher concentration it probably folds into self-dimers. Interestingly, although this sequence differs from hsa-miR-337-3p_1 by only a single nucleotide deletion, it exhibits slightly higher thermodynamic stability. These collective observations suggest that the higher stability of the secondary structures adopted by the most expressed sequence could decrease the chances of self-dimer formation at higher concentrations.

A biological explanation for the higher expression of this sequence might be that it also maintains higher stability in biological conditions. This stability could result in a longer half-life and greater resistance to nucleases, enhancing the probability of the sequence reaching its target and successfully achieving its biological activity.

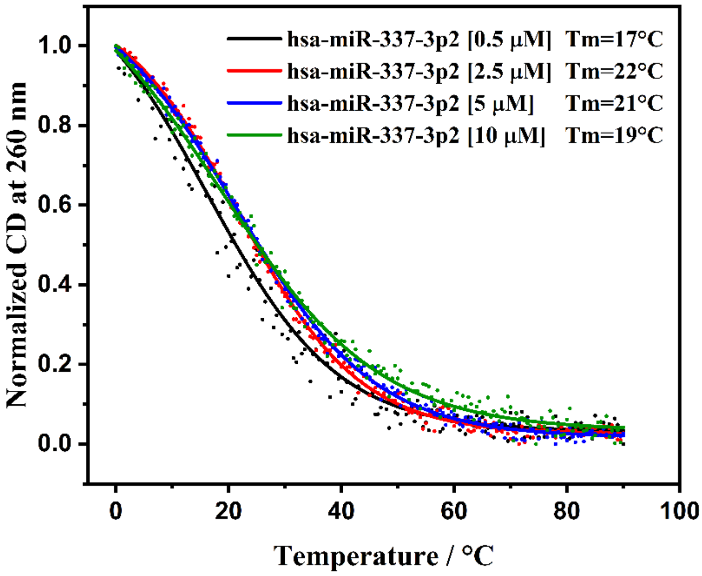


Fig. 25 *Normalized CD-melting curves of hsa-miR-337-3p2 at the concentration of 0.5 μM (black curve), 2.5 μM (red curve), 5 μM (blue curve) and 10 μM (green curve) in PBS 10 mM, pH = 7.4. Full circles represent the experimental points.*

Modification of human miR-337-3p

To verify whether the thermodynamic stability of secondary structures adopted by a miRNA influences its biological activity, we decided to modify both the wild-type sequences (hsa-miR-337-3p_1 and bat-miR-337-3p). Specifically, we aimed to increase the thermodynamic stability of the secondary structures of hsa-miR-337-3p_1 without altering the seed region, by changing as few nucleotides as possible.

We used the “oligoanalyzer tool” provided by IDT®, which allows the prediction of the secondary structures (such as hairpins or self-dimers) of a given sequence and their theoretical Tm. **Fig. 26** displays the most stable secondary structure of hsa-miR-337-3p_1 predicted by the tool. This structure displays two Watson-Crick base-pairings and a theoretical Tm of 32°C.

To stabilize this structure, we planned to replace the uracil (U) at position 20 with adenine (A). This change is projected to yield a new hairpin structure with more base pairings. Specifically, the new adenine at position 20 can interact via hydrogen bonding with the uracil on the opposite side (at position 9), leading the uracil at position 21 to interact with adenine at position 8. This would result in the formation of two new base pairings, A-U and U-A respectively.

For this modified sequence (mut-hsa-miR-337-3p), the IDT tool predicted two new hairpin structures as shown in **Fig. 26**, each displaying four base-pairings and theoretical Tm of 52°C and 47°C, respectively. Thus, in silico analysis suggests that enhancing the stability of the secondary structure of hsa-miR-337-3p can be achieved by changing only a single nucleotide out of the seed region.

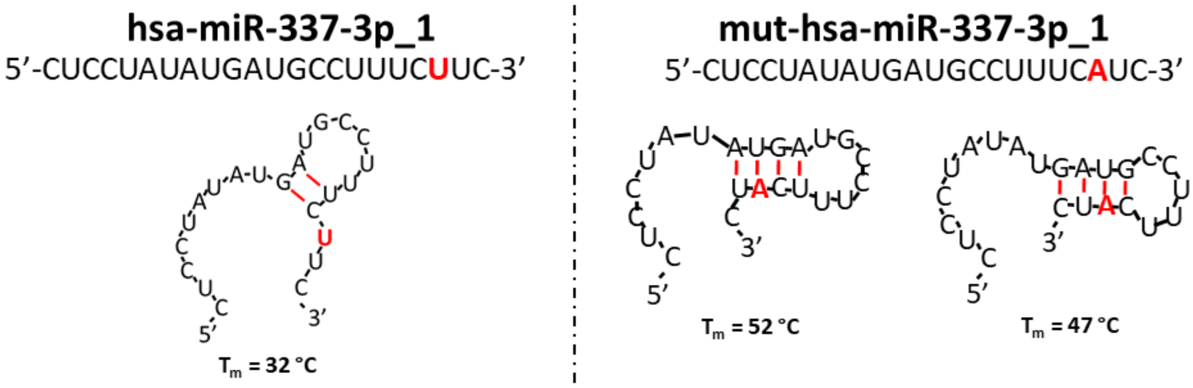


Fig. 26 *The cartoon representations depict potential hairpin structures along with their respective calculated melting temperatures for hsa-miR-337-3p_1 (left) and mut-hsa-miR-337-3p_1 (right), as determined by the "IDT® oligoanalyzer tool". The red highlighted nucleotide indicates the site of modification. Red lines represent canonical Watson-Crick base pairings.*

We subsequently purchased from IDT® the mutated sequence (mut-hsa-miR-337-3p) and characterized by CD and CD-melting experiments to verify if the structures adopted by the sequence were effectively more stable in solution in comparison to the wild-type sequence.

To make a comparison between the wild-type and mutated human sequences, we analysed the CD spectra of both (**Fig. 27 (a)**). The spectra displayed a positive band around 270 nm for hsa-miR-337-3p_1 showing a slightly higher and red-shifted intensity than mut-hsa-miR-337-3p_1. This suggests a potential variation in the organization of the secondary structures adopted by the two miRNAs. The CD-melting curves (**Fig. 27 (b)**) show a markedly higher stability for mut-hsa-miR-337-3p_1 than the corresponding wild-type sequence. Melting temperatures (Tm) extracted from these curves were 16°C for the wild-type sequence and 40°C for the mutated one. These experimental results validate our in silico findings, demonstrating that the secondary structures of hsa-miR-337-3p can be significantly stabilized by altering just one nucleotide outside the seed region.

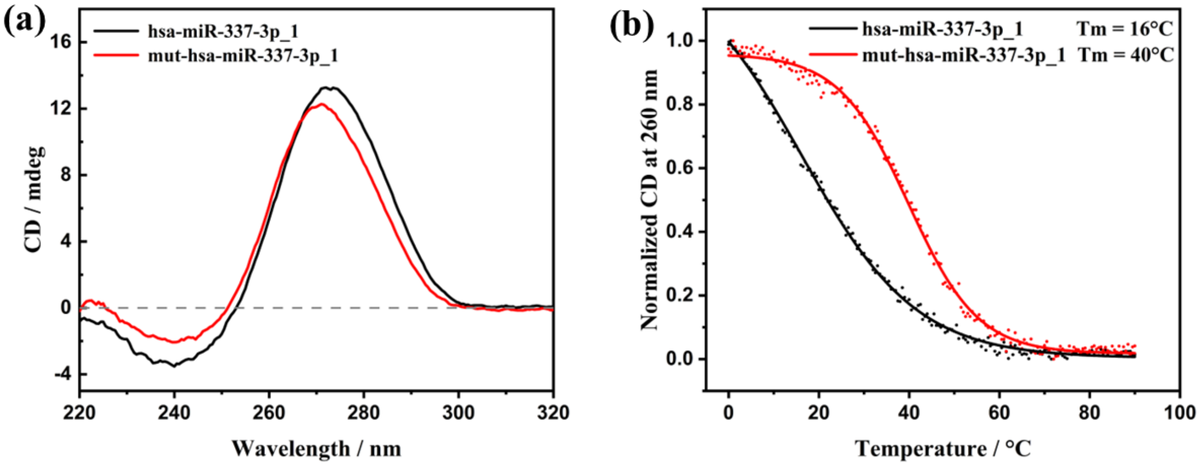


***Fig.** 27 CD-spectra **(a)** and normalized CD-melting curves **(b)** of hsa-miR-337-3p_1 [2.5 μM] (black curve) and mut-hsa-miR-337-3p_1 [2.5 μM] (red curve) in PBS 10 mM, pH=7.4.*

Modification of bat miR-337-3p

In light of the higher thermodynamic stability demonstrated by the bat sequence compared to the human sequence, and to further examine if the thermodynamic stability could influence its biological function, we projected to decrease the stability of the secondary structures adopted by bat-miR-337-3p and then verify in vitro (by luciferase reporter assay) if also the target efficiency can decrease. To achieve this goal, we used again the "IDT® oligoanalyzer tool", which displayed several potential structures (hairpins and self-dimers).

Our design for the new sequence was based on the most stable secondary structure, as shown in **Fig. 28**. Beginning with this hairpin structure, and following a similar logic used to stabilize the human sequence, we initially altered only one nucleotide at a time, outside of the seed region. In particular, we selected those nucleotides involved in base pairing to reduce their number and destabilize the structure.

However, changing just one nucleotide still resulted in structures with a similar number of base pairings and a comparable theoretical Tm, according to the "IDT® oligoanalyzer tool". This result confirmed the high stability of this structure. Consequently, we had to change at least two nucleotides to destabilize this structure. We replaced cytosine (C) at position 14 with adenine (A), and uracil (U) at position 18 with adenine (A). This yielded two new hairpin structures with theoretical Tm values of 16.2 °C and 7.4 °C, respectively, as shown in **Fig. 28**. Thus, at least in silico, it appears possible to destabilize the bat sequence by modifying at least two nucleotides outside the seed region.

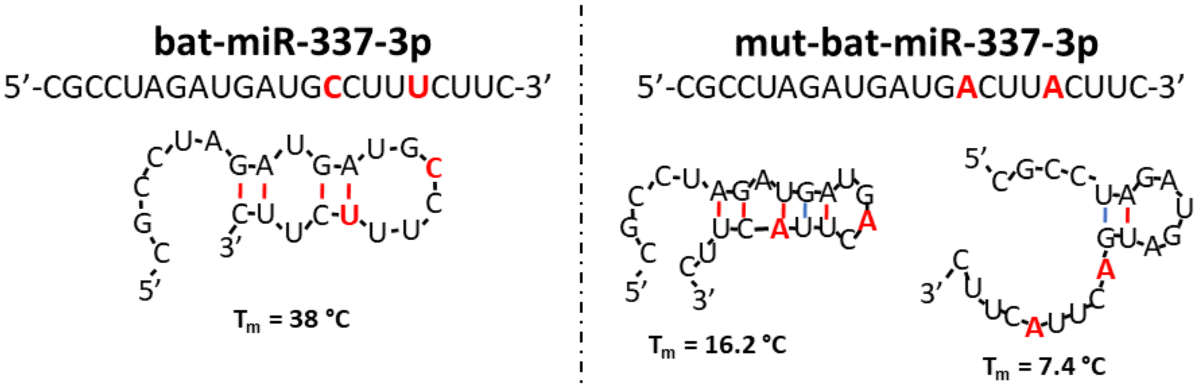


Fig. 28 *The cartoon representations depict potential hairpin structures along with their respective calculated melting temperatures for bat-miR-337-3p (left) and mut-bat-miR-337-3p (right), as determined by the "IDT® oligoanalyzer tool". The red highlighted nucleotide indicates the site of modification. Red lines represent canonical Watson-Crick base pairings, whereas blue lines represent the non-canonical Wobble base pairing.*

We subsequently purchased from IDT® the mutated sequence (mut-bat-miR-337-3p) and characterized by CD and CD-melting experiments to verify if the structures adopted by the sequence were effectively more or less stable in solution in comparison to the wild-type sequence.

We compared the spectroscopic results (wild-type vs mutated) and noted that the CD spectra of the two sequences (**Fig. 29 (a)**) showed remarkably similar strong bands around 270 nm. However, minor differences in the weaker bands around 240 nm and 220 nm were displayed.

We obtained the melting temperatures of the two miRNAs from the CD-melting curves (**Fig. 29 (b)**), which are 30°C for the wild-type sequence and 26°C for the mutated sequence. Interestingly, the mutated bat-miR-337-3p exhibited slightly less stability than its wild-type counterpart. This underscores the intrinsic stability of bat-miR-337-3p, highlighting the fact that even substantial alterations to the sequence resulted in only a minor destabilization. This resilience shows the robustness of bat-miR-337-3p's secondary structures, underlining its considerable thermodynamic stability. It seems that the secondary structure of bat-miR-337-3p is such that it can maintain its stability even when subjected to mutations, contributing to the robust functional role this miRNA likely plays in bats.

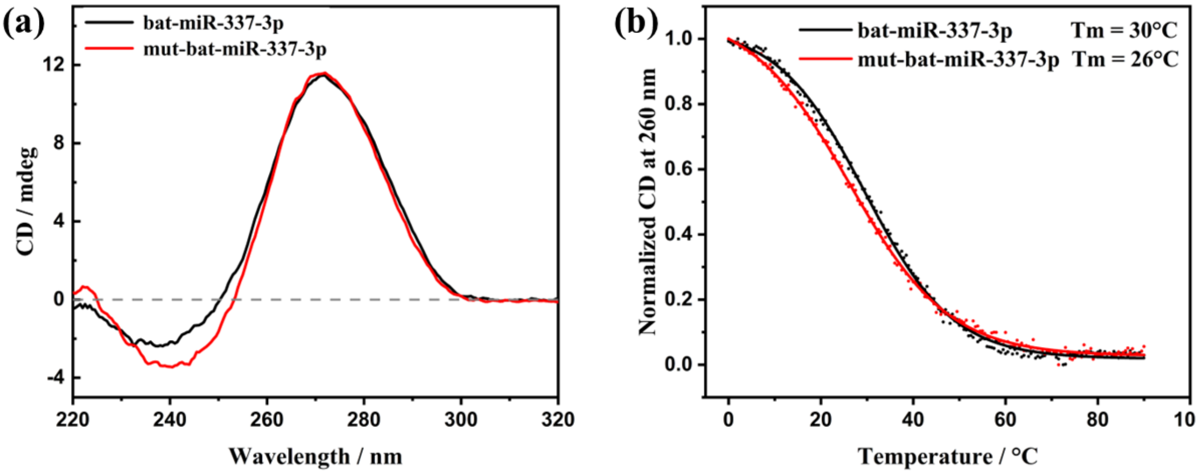


Fig. 29 *CD-spectra* ***(a)*** *and normalized CD-melting curves* ***(b)*** *of bat-miR-337-3p [2.5 μM] (black curve) and mut-bat-miR-337-3p [2.5 μM] (red curve) in PBS 10 mM, pH=7.4.*

Finally, in collaboration with the Max Planck Institute for Psycholinguistics (Nijmegen, Netherlands), a luciferase reporter assay was conducted to provide further confirmation of our hypothesis. From the results (**Fig. S1**), it can be observed that the mut-hsa-miR-337-3p exhibits a stronger capability to suppress the expression of the complementary target compared to the wild-type human miR. Conversely, the mut-bat-337-3p displayed a reduced capacity to inhibit the expression of the complementary target relative to the wild-type bat miR.

2.2.3. Final remarks

In this study, we explored the secondary structures and thermodynamic stability of two mature miRNAs, hsa-miR-337-3p_1 and bat-miR-337-3p, using CD and CD-melting experiments. Notably, despite only a two-nucleotide difference between the sequences, our results indicated distinct secondary structures and significantly different stability. The bat miRNA displayed higher thermodynamic stability, potentially contributing to its prominent biological role and effectiveness in target interaction [177].

We then explored deeper into the relationship between thermodynamic stability and biological activity, examining the most expressed isoform of the human miRNA. Our results showed that this isoform adopted different conformations in solution and exhibited higher thermodynamic stability, supporting the hypothesis that stability can influence a miRNA's biological activity.

To further test this hypothesis, we manipulated the sequences of the miRNAs in silico to modify their stability. The mutated human miRNA showed greater stability than the wild-type sequence, confirming the potential to significantly increase stability by changing a single nucleotide outside the seed region. The bat miRNA, on the other hand, exhibited inherent high stability, which could only be slightly reduced with a double-nucleotide change, underscoring the robust stability of this miRNA.

These findings provide fascinating evidence that a miRNA's secondary structures and their thermodynamic stability can influence its biological function. However, more research is needed to confirm the efficiency of these stabilized miRNAs in repressing targets, and to explore potential changes in target selection. Our study has opened up new pathways for understanding the roles of miRNA structures and stability in target recognition, providing a solid foundation for future research.

2.3. Interactions between miR-26b-5p and four achiral porphyrins: aim of the work

To better investigate the secondary structure of mature miRNAs, in the second part of our study, we explored by circular dichroism (CD) and CD-melting experiments the possible secondary structures adopted by the human miR-26b-5p (UUCAAGUAAUUCAGGAUAGGU). This sequence has been chosen for its biological importance, indeed miR-26b is found in an intron of the Ctdsp2 gene and it acts synergistically with its host gene to regulate neuronal development [181]. It has been revealed that miR-26b is a crucial regulator in carcinogenesis and tumor progression in several forms of cancer by serving as a tumor suppressor gene, indeed, impedes cellular proliferation in breast cancer [182], and inhibits metastasis of osteosarcoma [183]. Furthermore, a recent study found that miR-26b-5p inhibits the proliferation, migration, and invasion of intrahepatic cholangiocarcinoma cells via inhibiting S100A7 [184]. Similarly, downregulation of miR-26b-5p and miR-26a-5p in bladder cancer cells was observed, and both of these miRNAs significantly prevented cancer cell motility and invasion [185,186]. Aside from cancer, altered miR-26b expression and functional abnormalities have been identified in a range of different disorders. These findings show that miR-26b plays a role in cardiac function maintenance [187] and it has been found to be elevated in the human temporal cortex in Alzheimer's disease development [188].

Secondly, we used four achiral porphyrins (**Fig. 30**) with the aim to exploit them as chiroptical probe, as well as to modulate the thermodynamic stability of miR-26b-5p.

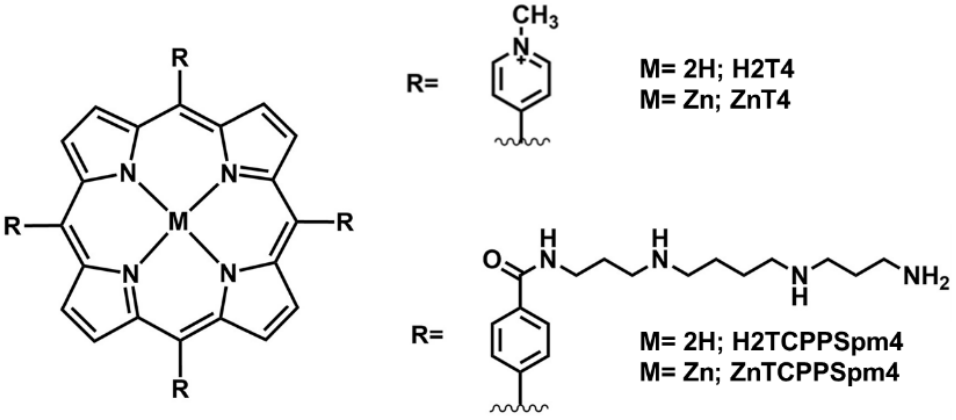


Fig. 30 *Structures of H2T4 and H2TCPPSpm4, with their respective Zn(II) derivatives, ZnT4 and ZnTCPPSpm4. For clarity of the image the axial water molecule attached to Zn(II) of ZnT4 is not depicted.*

By UV-vis, resonance light scattering (RLS), CD and CD-melting, we characterized the interactions between miR-26b-5p and meso-tetrakis(4-N-methylpyridyl)porphyrin (H2T4) and its zinc(II) derivative, ZnT4. In physiological conditions both porphyrins bring in the four *meso*-positions cationic charges. In addition, H2T4 is able to intercalate in DNA GC rich region via stacking interactions with nucleobases [189]. The latter type of interaction is precluded to ZnT4 owing to the penta-coordinated nature of the central Zn, which hinders intercalation [190,191]. Successively, in order to evaluate the role of electrostatic interactions we studied the interactions of the title miRNA with other two porphyrins bearing four multi-cationic spermine arms (meso-tetrakis(4-carboxyphenylspermine) porphyrin, H2TCPPSpm4 [192], and its Zn(II) derivative, ZnTCPPSpm4 (**Fig. 30**). We selected the two sperminated derivatives in order to improve the biocompatibility of porphyrins since it has known the polyamine transport system of the cell affords a selective accumulation of polyamine analogues in neoplastic tissues [193]. In addition, it has been already shown that H2TCPPSpm4 is able to stabilize G-Quadruplex superstructures [194] and it can interact with a variety of DNA structures [118], whereas ZnTCPPSpm4 can detects, catalyzes and stabilizes Z-DNA [195].

2.3.1. Materials and methods

The 5K buffer was prepared by solubilizing the necessary salts in 200mL of ultrapure water. In particular, the 5K buffer consists of cacodylate sodium salt 10 mM, potassium chloride 5mM and lithium chloride 95mM. The final pH was adjusted to 7.3 by concentrated HCl.

The phosphate buffer ([H2PO4−] + [HPO42−] = 10mM, pH = 7.3) was prepared by dissolving the proper amount of corresponding anhydrous sodium salts (NaH2PO4 and Na2HPO4) in 200mL of ultrapure water.

The Phosphate buffered saline (PBS) has been obtained from the corresponding tablets, that is one PBS tablet was dissolved in 200mL of ultrapure water. PBS buffer (pH = 7.3) contains 10mM of phosphate buffer sodium salt ([H2PO4−] + [HPO4−2−] = 10mM), 137mM of sodium chloride, and 2.7mM of potassium chloride.

The miR-26b-5p sequence (5'-UUCAAGUAAUUCAGGAUAGGU-3') used was purchased from Integrated DNA Technologies IDT® and used without further purification. The solid was dissolved in ultra-pure water obtained by Elga Purelab Flex system by Veolia with purity of 18.2 MΩcm, achieving a stock solution with concentration of ~ 100 µM. Annealing is performed increasing the temperature up to 90° for 5 min, then cooling down slowly up to room temperature and finally storing the solution at 4°C for one night. Concentration of miR solutions were checked by UV-vis measurements at 80°C using the extinction coefficient of the sequence given by IDT: ε_{260} nm = 224,200 L/(mol·cm); Then, by dilution in PBS buffer 10 mM ([KCl] 2.7 mM; [NaCl] 137 mM; pH 7.4) we prepared the sample solution.

ZnT4 porphyrin was obtained by metalation of the naked tetracationic porphyrin (purchased from Mid-Century) following the well-known procedure reported in literature [196]. The concentration of the stock solutions of H2T4 and ZnT4 were prepared by dissolving the solid in ultrapure water in order to achieve concentrations ranging from 3×10^{-4} M to 4×10^{-4} M. The concentration of these stock solutions was checked by spectrophotometric experiments using for H2T4 ε_{423} = 224000 $M^{-1}cm^{-1}$ and for ZnT4 ε_{437} = 204000 $M^{-1}cm^{-1}$ in water.

H2TCPPSpm4 and ZnTCPPSpm4 were obtained according to the literature procedure [192,195]. All stock solutions of H2TCPPSpm4 were prepared by dissolving the solid in DMSO in order to avoid aggregation process, whereas stock solutions of ZnTCPPSpm4 were prepared by dissolving the solid in ultrapure water in order to achieve concentrations ranging from 2×10^{-4} M to 3×10^{-4} M. The concentration of these stock solutions was checked by spectrophotometric experiments using for H2TCPPSpm4 ε_{437} = 338000 M^{-1} cm^{-1} at pH = 1 in water, HCl and for ZnTCPPspm4 ε_{423} = 133766 M^{-1} cm^{-1} at pH = 3.5 in water, HCl.

Electronic Circular Dichroism

CD spectra were recorded at 20°C using a Jasco J-715 spectropolarimeter equipped with a single position Peltier temperature control system. A quartz cuvette with a 1 cm path length was used for all CD experiments. The solution of 2.5 μM miR-26b-5p was titrated with increasing amount of porphyrin and CD scans were collected with the following parameters: scanning rate 50 nm/min, data pitch 0.5 nm, digital integration time (D.I.T) 2s, band width 2.0 nm. Each CD spectrum was an average of at least five scans. CD melting was performed at the ratio porphyrin:miR 1:1 and 3:1 respectively, with the following parameters: 260 nm wavelength, 5–90 °C temperature range, 1 °C/min temperature slope and 8 sec response.

CD-melting curve were obtained from the experimental data with the help of a Boltzmann sigmoid equation (as described in **2.2.1 Materials and methods**).

UV–vis absorption spectroscopy

UV–vis absorption spectra were collected at room temperature using a Jasco V-530 UV-vis spectrophotometer equipped with a 1 cm path-length cell. Conditions were as follows: scanning rate 100 nm/min, data pitch 0.5 nm, band width 2 nm.

Titrations of miR-26b-5p [2.5 μM] were performed by adding increasing amount of porphyrin. In general, the porphyrins form with other molecules, complexes which exhibit kinetic inertness; therefore it is possible to perform UV-titration to obtain information on the stoichiometry of these complexes. By plotting of the Soret band absorbance *vs* the [porphyrin]/[miRNA] it is possible to reveal straight lines with different slopes indicating complex stoichiometry. Different slopes indicate the presence in solution of diverse supramolecular species, which are stable, inert and not in equilibrium with each other, characterized by a certain molar extinction coefficient [197]. Conversely, a unique and straight linear plot will appear all over the titration experiment as well as observed in the case of porphyrin alone in buffer solution. The titrations were terminated when the straight-line slope results similar to that obtained by the titrations of the porphyrins alone in buffer solution. Break points are found where the titration line displays at least a 10% slope variation.

Fluorescence spectroscopy and resonance light scattering (RLS)

Fluorescence emission spectra and RLS were recorded using a Fluorolog FL-11 Jobin Yvon Horiba. Each experiment was carried out at room temperature with a 1 cm quartz cuvette. For fluorescence measurements the following parameters were used: emission range of 550–800 nm, increment of 1.0 nm, averaging time of 0.1 sec, one scan, and 2.5 nm slits both for excitation and emission. Excitation wavelength (λ_{ex}) was chosen taking into account the isosbestic points detected in the UV titrations for each porphyrin used: H2T4 (λ_{ex} = 430 nm); ZnT4 (λ_{ex} = 450 nm); H2TCPPSpm4 (λ_{ex} = 420 nm); ZnTCPPSpm4 (λ_{ex} = 450 nm).

2.3.2. Results and discussions

Spectroscopic characterization of miR-26b-5p

Firstly, to evaluate if the miR-26b-5p sequence adopts well-defined secondary structures (*e.g.* hairpins and/or self-dimers) or random conformations, we performed CD measurements at different strand concentrations (0.7 μM, 2.8 μM and 10 μM) in PBS buffer at pH = 7.4 (**Fig. 31 (a)**). For single stranded sequences in random conformations the intensity of the signal is expected to be close to few millidegrees [174] On the contrary, in the UV region between 300 nm and 220 nm, the CD signal of miR-26b-5p (at all the investigated concentrations) is quite structured and shows an intense band around 270 nm, a less intense negative band around 240 nm and a band ~ 220 nm, suggesting that the sequence is not in random conformation. Indeed, the CD spectra is similar to those of A-form RNA, indicating that the sequence is most likely folded in ordered secondary structures (*e.g.* self-dimers or hairpins). Additionally, the intensity of the CD bands increases linearly with the concentration of the sequence (**Fig. 31 (a), inset)** suggesting that the sequence does not change the conformations adopted in the concentration range investigated.

Secondly, we performed CD melting experiments by monitoring the variation of CD signal at 260 nm in the range 5-90 °C at different strand concentrations (0.7 μM, 2.8 μM and 10 μM) in PBS buffer (**Fig. 31 (b)**). In general, for a solution containing nucleic acid sequences, two possible situations are possible. The first one, in which the nucleobases are not involved in hydrogen bonds and a temperature increase causes "only" the breaking of π-π interactions (in the case of highly structured single strands). In this case, the CD intensity will decrease linearly upon increasing the temperature together with the smooth loss of base stacking. The second scenario, in which the nucleobases form either "internal" (e.g. hairpins) and/or "external" (*e.g.* duplexes) base-pairing, in which a temperature increase causes a breaking of hydrogen bonds and a consequent "melting" of the base-base coupling leading to a net transition in the intensity-temperature graph (*i.e.* a sharp decrease of the CD intensity) which is centered at the temperature where 50% of denaturation occurs (Tm) [174]. Altogether, the loss of ordered structure by heating causes CD changes which, monitored at a fixed wavelength as a function of temperature, allow to estimate the stability of secondary structures. The melting curves showed a clear transition for this sequence, confirming that this sequence adopts well-defined folded structures.

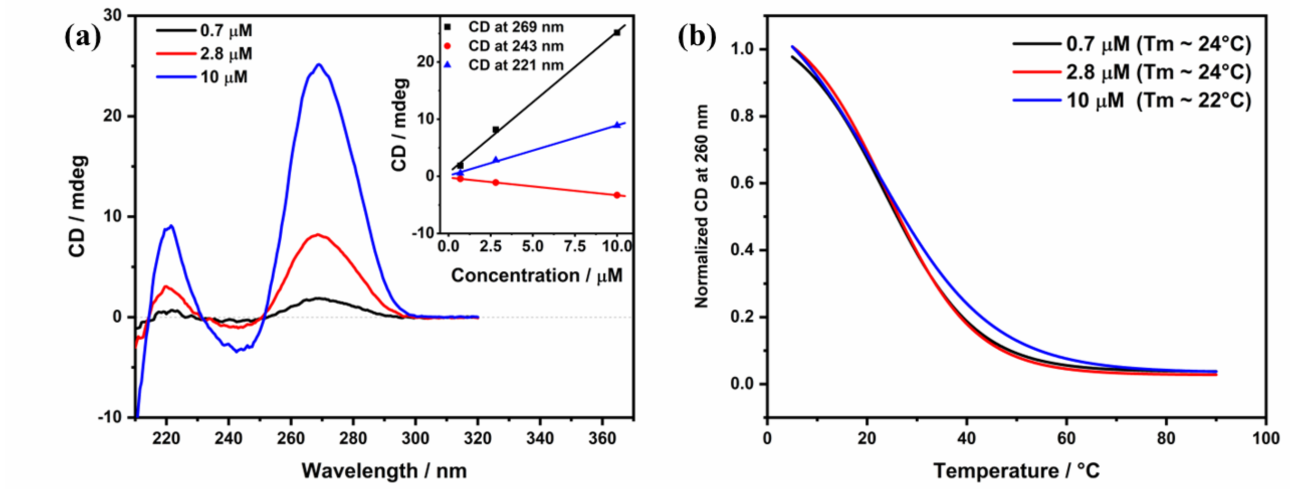


Fig. 31 (a) *CD spectra of miR-26b-5p in PBS 10 mM ([KCl] 2.7 mM; [NaCl] 137 mM; pH 7.4 (0.7 μM black curve; 2.8 μM red curve; 10 μM blue curve). Inset: plot of concentration of miR-26b-5p vs the CD signal at 269 nm (black squares), 243 nm (blue triangles) and 221 nm (red circles).* ***(b)*** *CD-melting curves normalized to [0,1] of miR-26b-5p in PBS 10 mM, pH = 7.4 (0.7 μM black curve; 2.8 μM red curve; 10 μM blue curve).*

To better understand if the miR-26b-5p adopts in solution preferentially one structure (*e.g.* hairpin) rather than another one (*e.g.* self-dimers) we performed melting experiments at different concentrations (0.7 μM, 2.8 μM, and 10 μM) and for each one we extracted the melting temperature (Tm) (**Fig. 31 (b)**). These experiments are very useful because monomolecular transitions (*e.g.* those associated to unfolding of hairpins) are virtually independent on strand concentration, while transitions with higher molecularities, as those associated to dissociation of self-dimers, exhibit a significant dependence on sample concentration. Melting temperatures extracted do not display significant differences among the concentrations investigated and for this reason we speculated that this sequence is mostly folded in hairpins.

Interaction of cationic H2T4 with miR-26b-5p

We firstly studied the interaction of H2T4 with miR-26b-5p by UV-vis absorption spectroscopy by the addition of increasing amounts of H2T4 (in the range 0.5-19 μM) to miR-26b-5p (2.5 μM) in PBS buffer at pH = 7.4 (**Fig. S2 (a)**). In PBS buffer, H2T4 porphyrin displays the Soret band at 421 nm and other four less intense bands at 510, 550, 580 and 650 nm (Q bands) (**Black line, Fig. 32**). After the addition of H2T4 up 2.5 μM to a solution containing 2.5 μM of miR-26b-5p, is possible to observe a large bathochromic shift (Δλ = 17 nm) and an hypochromic effect (~ 40%) of the porphyrin Soret band (**Red line, Fig. 32)** suggesting a significant perturbation in the porphyrin *π* electrons as a result of the binding to the miRNA. In order to verify if the porphyrin binding mode with miR-26b-5p depends on the ratio of the two molecules we have plotted the absorbance at 421 nm versus the [H2T4]/[miR] ratio. The plot reported in the inset of **Fig. 32** shows five distinct break points, indicating the formation of five H2T4:miR complexes (1:1, 2:1, 3:1, 5:1 and 6:1). In particular, for complexes with porphyrin contents higher than 3:1, a strong hypochromic effect is observed, suggesting a stronger interaction among porphyrins with miRNA. Yet, after the formation of the 6:1 complex, further addition of H2T4 has no effect on the slope of the straight line (**inset Fig. 32, orange line**) which became almost identical to that one obtained for H2T4 alone in buffer solution (**inset Fig. 32, black line**). These results indicate that the formation of the H2T4/miR-26b-5p complex take place up to a 6:1 ratio. Lastly, the second derivatives of the absorption spectrum of H2T4/miRNA complexes (**Fig. S2 (b)**) show that from ratio 5:1 the band around 420 nm is predominant, indicating that the further H2T4 added does not interact with H2T4/miRNA complexes.

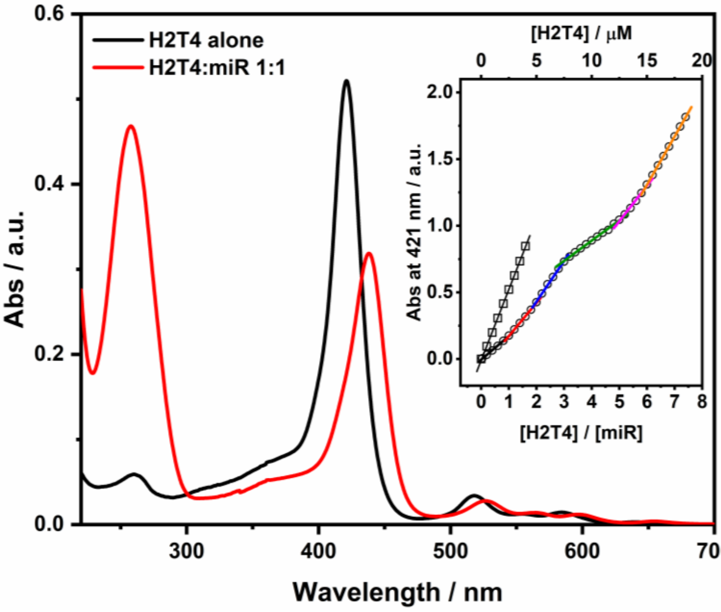


Fig. 32 *Absorption spectra of H2T4 [2.5 µM] in PBS 10 mM ([KCl] 2.7 mM; [NaCl] 137 mM; pH 7.4 (black curve) and in the presence of miR-26b-5p [2.5 µM] (red curve). Inset: plot of absorbance at 421 nm vs. the concentration of H2T4 alone in PBS (squares) and vs. the ratio [H2T4]/[miR-26b-5p] (circles), lines with different colors indicate a different slope.*

As reported in literature, large bathochromic shift (Δλ = 17 nm) and strong hypochromic effect (~ 40%) of the porphyrin Soret band, are ascribed to the intercalation of the ligand between two bases. However as mentioned above, mature miRNA are short sequences and even if arranged in some secondary structure the intercalation of porphyrin could be unlikely. Therefore, weak aggregation of porphyrin onto the miRNA structure could occur. In this contest, the formation of dimers (two porphyrins interacting in a face-to-face mode between two miRNA structures in a fascinating way) cannot be excluded.

The emission spectra of H2T4 (**black line, Fig. 33 (a)**) displays a large broad band from 630 nm to 780 nm which is mainly due to the coupling of the first excited state S1 with a nearby charge-transfer state (CT) from the porphyrin core to the pyridinium group, the coupling is favored in high-polarity solvents and can be assisted by high degree of rotational flexibility of the N-methylpyridinium groups [198]. However, in the presence of miRNA (**red line, Fig. 33 (a)**) it is possible to observe a change both in the fluorescence intensity and shape of the emission spectra, confirming the interaction between the porphyrin and the miRNA. Indeed, the broad emission of H2T4, after the interaction with miRNA, is split into two peaks of Q(0, 0) and Q(0, l) bands with emission maxima at around 660 and 730 nm (**red line, Fig. 33 (a)**). This effect might be caused by the changes in the dielectric of the solvation sphere associated with the porphyrin, in fact generally the complex formation of H2T4 with nucleotides results in dielectric changes by reducing the accessibility of the water molecules to the π orbitals of the porphyrin [199]. In addition, when the pyridinium groups stuck, the charge transfer effect is decreased, increasing the splitting effect of the broad emission band due to reduction of rotational flexibility of the N-methylpyridinium groups. These two effects brought by binding with miRNA render the electronic S1−CT mixing less effective, causing the splitting of Q(0, 0) and Q(0, l) bands. At 3:1 (H2T4:miR-26b-5p) ratio the two emission bands become less resolved (**Fig. S3**) suggesting some changes in the excited state of the porphyrin that could be ascribed to the formation of other H2T4:miR-26b-5p complex, in accordance with the absorbance data in which at this ratio an evident break point is displayed.

The interaction with miR-26b-5p was also confirmed by resonance light scattering measurements (**Fig. S4)**. The RLS intensity of H2T4 alone is low and does not increase with increasing porphyrin concentrations (**Fig. 33 (b)**) indicating an absence of aggregation in buffer solution. Whereas, titration of miR-26b-5p with increasing amount of H2T4 shows a strong enhancement in light scattering intensity when the concentration of H2T4 is 6.5 μM. Then the RLS intensity increases linearly after addition of more H2T4 up to 10 μM (**Fig. 33 (b)**), after that no enhancement is observed. The RLS results suggest that porphyrins are electronically coupled on the miR-26b-5p up to 4:1 (H2T4:miR-26b-5p) ratio. Whereas after the ratio 4:1 the RLS intensity does not increase anymore suggesting that other well-ordered H2T4:miR-26b-5p complexes are not formed. However these data do not exclude the formation of other H2T4:miR-26b-5p complexes at higher ratio as suggesting by UV data. In addition the enhancement of RLS intensity up to H2T4:miR-26b-5p 4:1 ratio can be ascribed to the formation of well-ordered aggregates formed by porphyrins onto the miRNA structure rather than aggregation phenomenon of porphyrins alone. RLS and UV-vis data support our hypothesis of miR-assisted porphyrin self-aggregation which leads to strong electronic communication between individual porphyrins in the assembly.

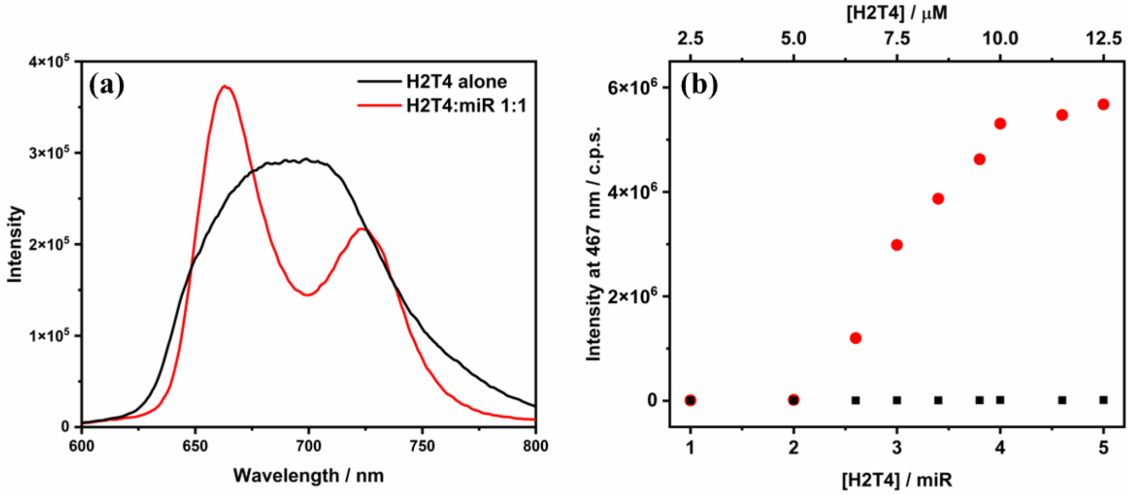


***Fig. 33 (a)** Emission spectra of H2T4 [2.5 μM] in PBS 10 mM ([KCl] 2.7 mM; [NaCl] 137 mM; pH 7.4 (black curve) and in the presence of miR-26b-5p [2.5 μM] (red curve). **(b)** Plot of the intensity of RLS signal at 467 nm vs. the ratio [H2T4]/[miR-26b-5p] (red circles) and vs. the concentration of [H2T4] alone in PBS (black squares).*

The low intensity of the RLS signal at ratio 1:1 and 2:1 H2T4:miR-26b-5p could be ascribed to the low sensibility of the RLS technique to detect short aggregates, like dimers, which cannot be excluded as reported for the UV data.

Circular dichroism titration with increasing amount of H2T4 were performed. The CD signal of the miR-26b-5p alone in solution in the UV region between 300 nm and 220 nm (**black line, Fig. 34**) shows an intense positive band around 270 nm and a less intense negative band around 240 nm, suggesting that the sequence is not in random conformation. Upon increasing addition of H2T4, the intense positive band decreased in intensity proposing a modification of the secondary structures of the miRNA, as if the porphyrin unwinds the oligonucleotide structure. The region between 500 nm and 400 nm is highly diagnostic to differentiate the type of binding of an achiral porphyrin with polynucleotides, despite the fact that many ligands are achiral and optically inactive, they can acquire ICD signals when they interact with nucleic acids [200]. The presence of an ICD signal in the achiral ligand's absorption bands indicates that the ligand interacts with nucleic acid. In particular, very weak bisignate CD signal in the porphyrin Soret band, at the 1:1 and 2:1 (H2T4:miR-26b-5p) ratio is detected (**Fig. 34**). Bisignate CD signal indicate that electronic communication at least between two porphyrins occurs and in general is related to a substantial aggregation of porphyrins, but it could also be linked to an outer stacking of porphyrins onto the miRNA structure [189]. Therefore the bisegnate ICD at 1:1 (H2T4:miR-26b-5p) ratio suggests the formation of porphyrin dimer, with each porphyrin interacting with one miRNA sequence. Whereas after the ratio 3:1 the intensity of bisignate signal drastically increases. It is conceivable to speculate that increasing the concentration of H2T4, especially after the ratio 3:1, triggers porphyrins aggregation onto the miRNA structure, resulting in a stronger bisignate ICD signal, in accordance with the spectroscopic data showed previously.

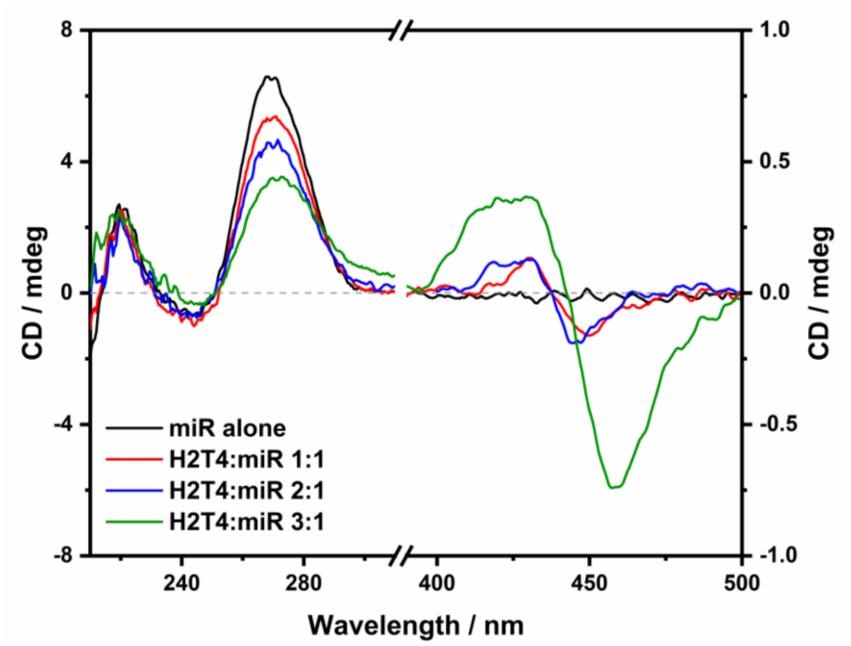


Fig. 34 *CD spectra of miR-26b-5p [2.5 µM] in PBS 10 mM ([KCl] 2.7 mM; [NaCl] 137 mM; pH 7.4 (black curve) and in the presence of increasing amounts of H2T4 (2.5 µM red curve; 5 µM blue curve; 7.5 µM green curve).*

The several spectroscopic techniques used demonstrate that until 3:1 (H2T4:miR) ratio, there is a specific type of interaction involving the formation of dimers in 2:2 and 4:2 (H2T4:miR) ratio. Indeed, the pronounced hypochromic effect in the absorption spectra, accompanied by two resolved bands in the emission spectrum, confirm that porphyrins are involved in a specific interaction. The lack of an increase in RLS intensity also excludes the possibility of strong aggregation, offering further confirmation of weak aggregation or dimers formation between two porphyrins and two miRNA molecules, as evidenced by the weak induced circular dichroism (ICD) signal. From the ratio 3:1, instead, an aggregation of porphyrins onto the miRNA structure becomes predominant. This is indicated by the fluorescence bands becoming less resolved, an increase in RLS, and the enhancement of the bisignate in the ICD spectra, also in this case is not possible to exclude the formation of dimers in 6:2 (H2T4:miR) ratio.

Interaction of cationic ZnT4 with miR-26b-5p

Secondly, we studied the interaction of the zinc(II) derivative of meso-tetrakis(4-N-methylpyridyl)porphyrin (ZnT4) with miR-26b-5p, by using the same spectroscopic techniques used for the interaction of H2T4 with miR-26b-5p. UV-vis titration with increasing amount of ZnT4 (in the range 0.5-12.5 μM) were performed in PBS buffer. The UV–vis absorption spectrum of ZnT4 (2.5 μM) in the absence of miRNA (**black line, Fig. 35 (a)**) showed the Soret band at 436 nm and other two Q bands at 565 and 610 nm, respectively. Whereas the presence of miR-26b-5p (2.5 μM) led to Soret band red shift to 443 nm (Δλ = 7 nm) and ~ 27% hypochromicity (% H) (**red line, Fig. 35 (a)**). We have reported the absorbance variation at 436 nm versus the [ZnT4]/[miR-26b-5p] ratio (**Fig. 35, (a) inset**), conversely to what is observed with H2T4, in this case, except for the first weak break point at the ratio 1:1 (ZnT4:miR) no others evident break points are showed, suggesting that the central metal (Zn) plays a crucial role in the interaction with miRNA avoiding the formation of different types of binding modes. As suggested by the low bathochromicity and the slight hypochromicity a very weak interaction is occurring in comparison to the interaction between H2T4 porphyrin and the miR sequence.

Both emission spectra **(Fig. 35 (b)**) of ZnT4 porphyrin alone in PBS (black line) and in the presence of miRNA-26b-5p (red line) displayed two not well resolved bands around 630 nm and 670 nm, and no significant differences in intensity and in shape are showed, confirming that weak interaction is occurring between ZnT4 and miRNA.

Also RLS titration with increasing amount of ZnT4 confirmed the absence of specific interaction with miRNA, indeed as showed in the RLS spectra (**Fig. S6**) there is not an enhancement in light scattering and this corroborate the hypothesis that porphyrins are not electronically coupled on the miRNA.

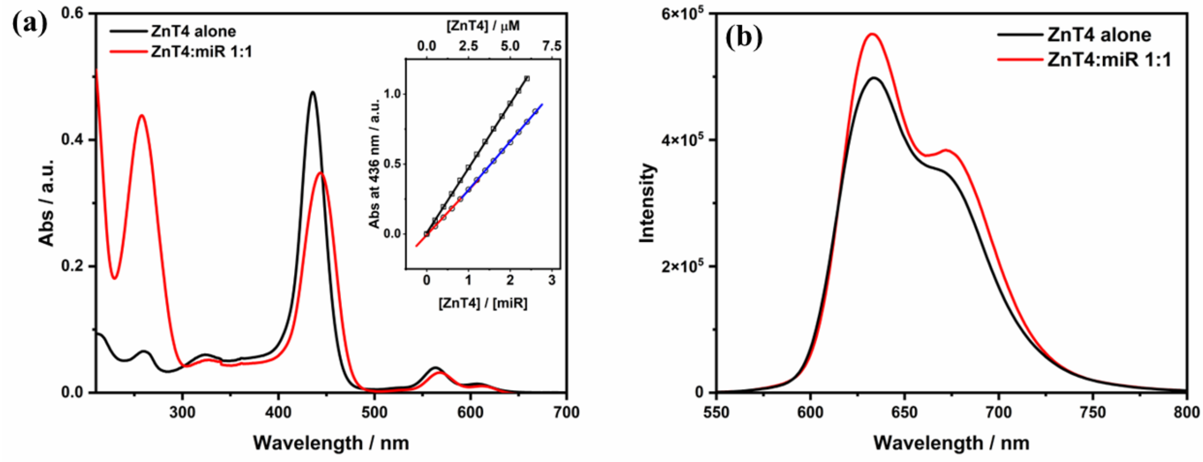


Fig. 35 (a) *Absorption spectra of ZnT4 [2.5 μM] in PBS 10 mM ([KCl] 2.7 mM; [NaCl] 137 mM; pH 7.4 (black curve) and in the presence of miR-26b-5p [2.5 μM] (red curve). Inset: plot of absorbance at 436 nm vs. the concentration of ZnT4 alone in PBS (squares) and vs. the ratio [ZnT4]/[miR-26b-5p] (circles), lines with different colors indicate a different slope.* ***(b)*** *Emission spectra of ZnT4 [2.5 μM] in PBS 10 mM ([KCl] 2.7 mM; [NaCl] 137 mM; pH 7.4 (black curve) and in the presence of miR-26b-5p [2.5 μM] (red curve).*

Circular dichroism titration performed with increasing amounts of ZnT4 (**Fig. 36**) displayed a slight modification of the miR-26b-5p signal around 270 nm, indeed the intense positive band decreased slightly in intensity, proposing a lower destabilization of the miRNA in comparison to the H2T4. Although the previous spectroscopic techniques suggested a very weak interaction between ZnT4 and the miRNA, a weak negative induced CD (ICD) signal appears around 460 nm, and its intensity decreases with increasing the ZnT4 concentration. Normally, a negative induced signal indicates intercalation of the porphyrin within the base-stacking of the oligonucleotide. However, ZnT4 is a pentacoordinate planar porphyrin within a water molecule axially coordinated, which should prevent intercalation between nucleobases, furthermore intercalation generally results in a stronger bathochromic shift ($\Delta\lambda \geq 15$ nm) and more marked hypochromism ($H \geq 35\%$) in the absorption spectra than our results [23]. Therefore it is more plausible that the negative ICD signal is due to a *pseudo*-intercalation of the ZnT4 porphyrin to the final sections of the miRNA sequence close to 5’ and 3’ end respectively. Indeed, the final nucleotides are not involved in a base pairing and consequently they are more available to interact via stacking interactions with ligands having aromatic moieties.

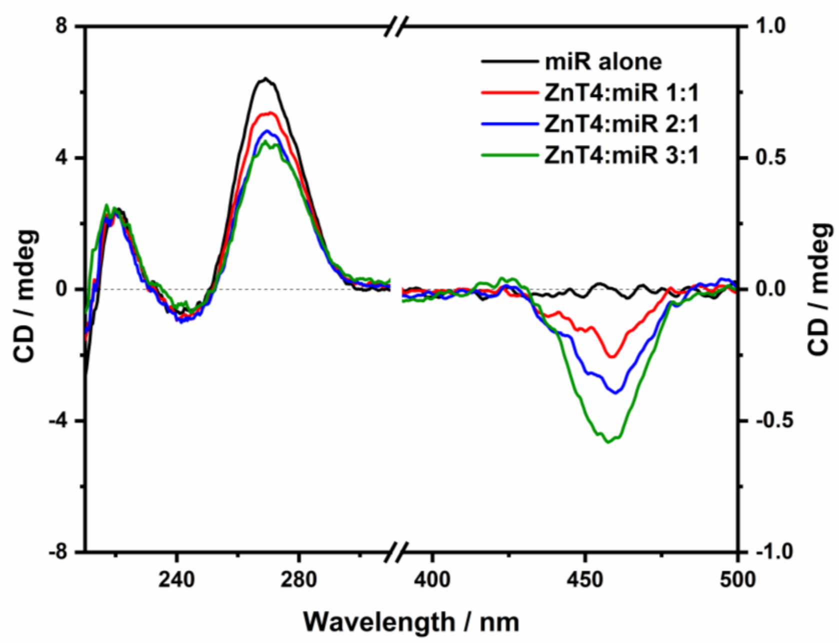


Fig. 36 *CD spectra of miR-26b-5p [2.5 μM] in PBS 10 mM ([KCl] 2.7 mM; [NaCl] 137 mM; pH 7.4 (black curve) and in the presence of increasing amounts of ZnT4 (2.5 μM red curve; 5 μM blue curve; 7.5 μM green curve).*

Interaction of H2TCPPSpm4 with miR-26b-5p

Successively, meso-tetrakis(4-carboxyphenylspermine) porphyrin, H2TCPPSpm4, was used to characterize the interaction with miR-26b-5p, in order to evaluate the contribution of the spermine arms which, in physiological condition, have at least the most external amino groups positively charged [192]. The UV–vis absorption spectrum of H2TCPPSpm4 (2.5 μM) in the absence of miRNA (**black line, Fig. 37**) showed the Soret band at 414 nm and other four less intense Q bands from 520 to 650 nm. While, the presence of miR-26b-5p (2.5 μM) produced a Soret band shift from 414 nm to 426 nm (Δλ = 12 nm) and ~ 40% hypochromicity (**red line, Fig. 37**), indicating strong interactions between the π-system of porphyrin and miRNA sequence, producing a similar behavior to the interaction with H2T4. Also, in this case we plotted the absorbance variation at 414 nm versus the [H2TCPPSpm4]/[miR-26b-5p] ratio (**Fig. 37, inset**). At ratio 1:1 and 2:1 (H2TCPPSpm4:miR) two clear break points are displayed, then the absorbance does not increase until the ratio 3:1 and after that the slope of absorbance at 414 nm remains constant demonstrating that the porphyrin is free in solution and it is no longer interacting with the oligonucleotide structure

(**Fig. S7**). Unlike the interaction with H2T4, in which the complex take place up to a 6:1 ratio (**Fig. 37, inset**), in this case H2TCPPSpm4 interacts with miRNA until the ratio 3:1, this is ascribable to the bigger dimension of the molecule which bears four spermine arms, which hinder the possibility to form a complex with higher binding stoichiometry.

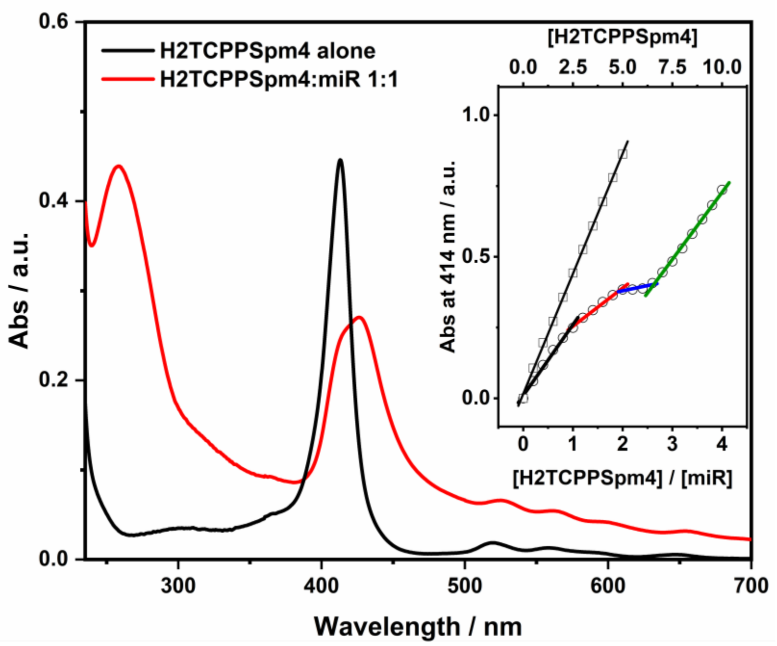


Fig. 37 *Absorption spectra of H2TCPPSpm4 [2.5 μM] in PBS 10 mM ([KCl] 2.7 mM; [NaCl] 137 mM; pH 7.4 (black curve) and in the presence of miR-26b-5p [2.5 μM] (red curve). Inset: plot of absorbance at 414 nm vs. the concentration of H2TCPPSpm4 alone in PBS (squares) and vs. the ratio [H2TCPPSpm4]/[miR-26b-5p] (circles), lines with different colours indicate a different slope.*

For H2TCPPSpm4 porphyrin, as well as H2T4, emission spectra is formed by the initial excited state, S1, and the charge-transfer state (CT) between the porphyrin core and its substituents at meso position, which in this case are formed by the carboxyphenylspermine groups. Polar solvents or restriction of peripheral substituents can cause fluorescence quenching because of the coupling of these two states (S1-CT).

Emission spectra of H2TCPPSpm4 alone in solution (**black curve, Fig. 38 (a)**) produced a peak around 645 nm and a shoulder around 700 nm. Noteworthy from the concentration of 5 μM the peak at 645 nm becomes larger and red shifted (Δλ ~ 5 nm) (**Fig. S8 (a)**). To better understand the origin of this shift we plotted the second derivatives of the absorption spectra at the concentration of H2TCPPSpm4 [1 μM] and [5 μM], respectively (**Fig. S8 (b)**). In the derivatives plot (**Fig. S8 (b), inset**) at the concentration of 5 μM a shoulder around 405 nm appeared, and it became more pronounced at higher concentrations (data not shown), suggesting that at low concentration H2TCPPSpm4 porphyrin is free in solution, while increasing the concentration triggers porphyrins aggregation which can explain the red-shift of the peak around 645 nm in the emission spectra. The presence of miR-26b-5p (**red line, Fig. 38 (a)**) led to a large decrease in fluorescence intensity and a red shift of 10 and 15 nm for the 650 and 700 nm peaks, respectively. The significant quenching of the fluorescence could be attributed to close contacts between porphyrins and miRNA, as well as porphyrin self-association supported by the miRNA backbone. Emission spectra at ratio 1:1 and 2:1 (H2TCPPSpm4:miR) are sharper and better resolved proposing restriction in rotation of the spermine arms upon miRNA structure, whereas at ratio 3:1 and 4:1 (**Fig. S9**) the quenching decreases, and the emission spectra resembled to the spectra of the porphyrin alone in solution, implying that after the ratio 2:1 porphyrins are no longer interacting, in accordance with the absorption data.

H2TCPPSpm4 alone showed a low RLS intensity (**black squares, Fig. 38 (b)**), indicating that strong and ordered aggregation does not occur, whereas titration of miR with increasing amount of H2TCPPSpm4 led to an increase of the RLS intensity until the ratio ~ 2:1 (H2TCPPSpm4:miR) (**red circles, Fig. 38 (b)**), proposing an aggregation phenomenon with strong electronic communications among porphyrins on the miRNA structure. Increasing addition of porphyrin conducted to a slightly decrease of the RLS intensity, indicating that further addition does not form a complex with higher binding stoichiometry, in line with the UV-vis data.

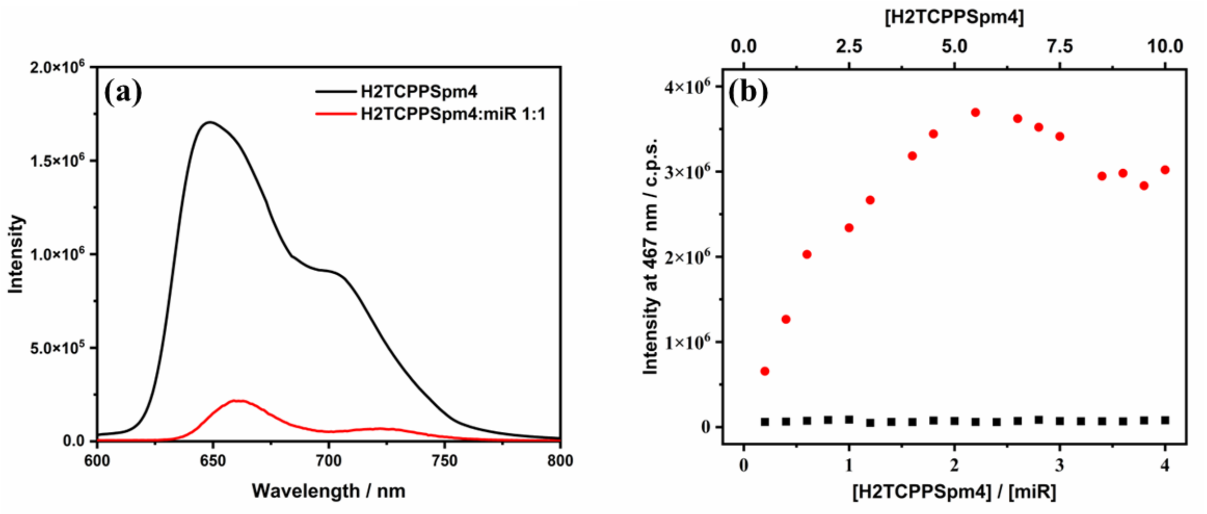


***Fig. 38 (a)** Emission spectra of H2TCPPSpm4 [2.5 μM] in PBS 10 mM ([KCl] 2.7 mM; [NaCl] 137 mM; pH 7.4 (black curve) and in the presence of miR-26b-5p [2.5 μM] (red curve). **(b)** Plot of the intensity of the H2TCPPSpm4 RLS signal at 467 nm vs. the concentration of [H2TCPPSpm4] in the presence of miR-26b-5p (red circles) and alone in PBS (black squares).*

Circular dichroism spectrum of miR-26b-5p with increasing quantity of H2TCPPSpm4 (**Fig. 39**) showed a strong reduction of the band at 270 nm, which could be explained by preferential binding of porphyrins to single-stranded miRNA, which penalize the formation of secondary structures by hydrogen base pairing. In particular, after the ratio 2:1 (H2TCPPSpm4:miR) the CD signal is close to few millidegrees, suggesting that the nucleobases are not involved in hydrogen bonds and the sequence is mainly in random single strands. A weak and negative ICD signal appeared in the Soret region, as also described previously, a negative induced signal suggests intercalation of the porphyrin within the base-stacking of the oligonucleotide. However, it is remarkable to remember that ICD signal arises from chiral distortion of the porphyrin symmetry and/or intermolecular exciton coupling between at least two chiral well-oriented chromophores (porphyrin–porphyrin or porphyrin–nucleobase) and after the ratio 2:1 the ICD signal disappeared, proposing that increasing porphyrin concentration induces the random conformations of oligonucleotide structure, suggesting external binding of the porphyrin onto the miRNA structure.

In conclusion, H2TCPPSpm4 destabilize the secondary structures of miR stronger than H2T4 porphyrin, suggesting that the four spermine pendants have a predominant effect to disrupt the hydrogen base pair interactions of the miR structures causing its conformational change.

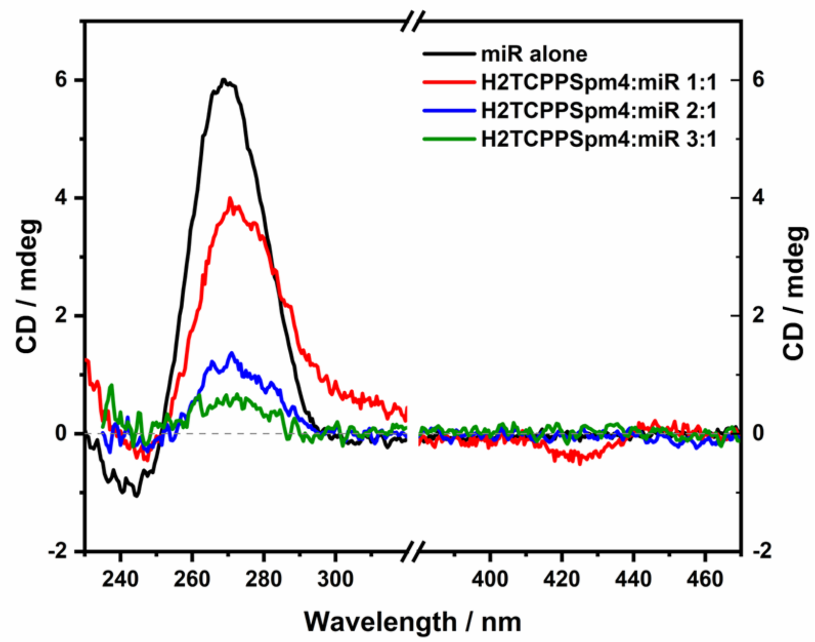


Fig. 39 *CD spectra of miR-26b-5p [2.5 µM] in PBS 10 mM ([KCl] 2.7 mM; [NaCl] 137 mM; pH 7.4 (black curve) and in the presence of increasing amounts of H2TCPPSpm4 (2.5 µM red curve; 5 µM blue curve; 7.5 µM green curve).*

Interaction of ZnTCPPSpm4 with miR-26b-5p

We also examined the interaction between the ZnTCPPSpm4 porphyrin and miR-26b-5p. However, it is crucial to acknowledge, that this porphyrin forms self-aggregates in PBS buffer, a phenomenon that will be discussed in detail in the following paragraph. As a result of the aggregation starting state of ZnTCPPSpm4, it is not possible to perform stoichiometric analysis with miRNA. However, it is possible to observe that in presence of the miRNA, all the absorption bands of ZnTCPPSpm4 are red shifted (**inset, Fig. 40**) suggesting an interaction between porphyrin aggregates and the miRNA structure.

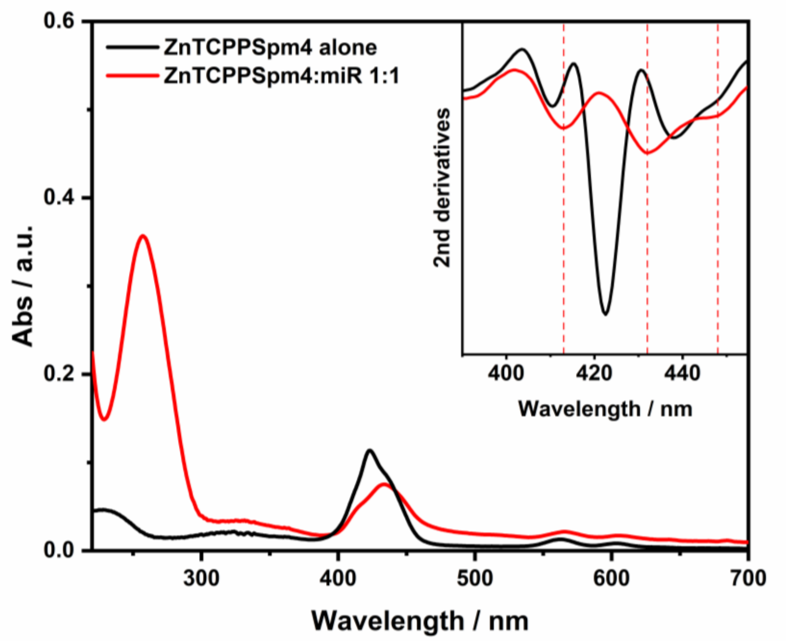


Fig. 40 *Absorption spectra of ZnTCPPSpm4 [2.5 µM] in PBS 10 mM ([KCl] 2.7 mM; [NaCl] 137 mM; pH 7.4 (black curve) and in the presence of miR-26b-5p [2.5 µM] (red curve). Inset: 2nd derivatives of UV-vis spectra.*

Emission measurements of ZnTCPPSpm4 with and without miRNA did not produce detectable fluorescence spectra in both cases, also at higher porphyrin concentration. This is due to the specific assembly of porphyrin aggregates with miRNA's phosphate backbone and phosphate molecules, respectively which cause a quenching of the fluorescent emission.

Also RLS of ZnTCPPSpm4 alone (**black squares, Fig. 41**) confirms self-aggregation phenomenon. Indeed, RLS intensity slightly increases by increasing the porphyrin concentration. However, in the presence of miRNA (**red circles, Fig. 41**) the RLS intensity is enhanced proposing a stronger and well-ordered communication among porphyrins onto the miRNA sequence. These spectroscopic evidences advise the idea that the presence of ordered template, such as miRNA structure, induces better organization of the porphyrin aggregates.

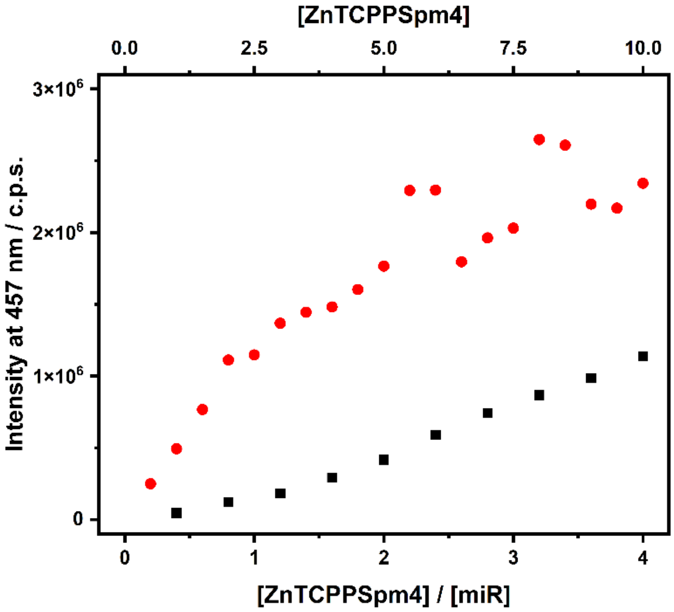


Fig. 41 *Plot of the intensity of RLS signal at 467 nm vs. the ratio [ZnTCPPSpm4]/[miR-26b-5p] (red circles) and vs. the concentration of [ZnTCPPSpm4] alone in PBS (black squares).*

Conversely from what observed with other porphyrins used previously, circular dichroism with increasing amount of ZnTCPPSpm4 (**Fig. 42**) did not produce significantly changes of the miRNA CD signal around 270 nm, indicating that this porphyrin does not destabilize and modify the secondary structures adopted by the miRNA. In the Soret absorption region a bisignate ICD signal is detected (**Fig. 42**). The low intensity of this ICD signal can be attributed to the low absorbance showed in the UV-vis spectra for these aggregates and / or to a very weak interaction between ZnTCPPSpm4 and miRNA.

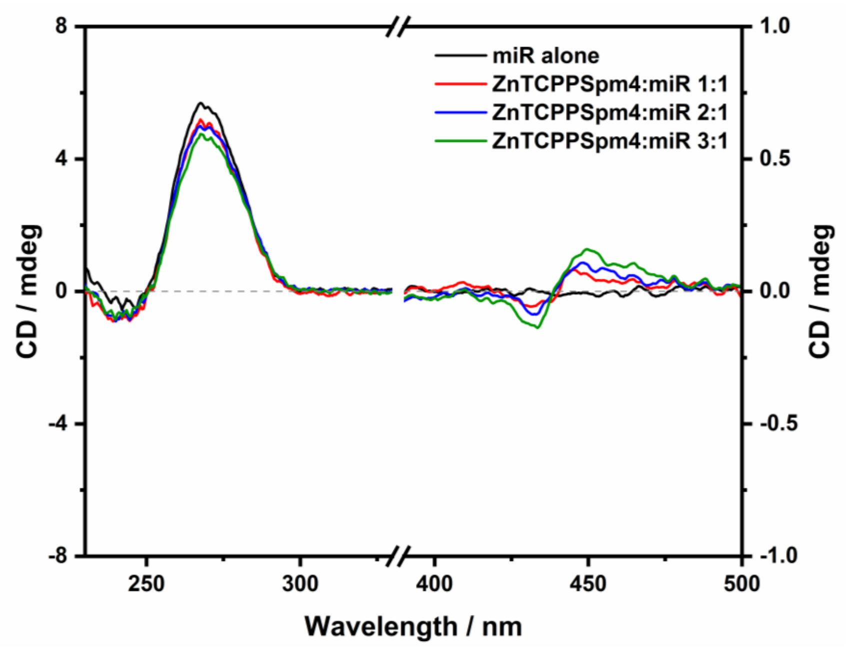


Fig. 42 *CD spectra of miR-26b-5p [2.5 µM] in PBS 10 mM ([KCl] 2.7 mM; [NaCl] 137 mM; pH 7.4 (black curve) and in the presence of increasing amounts of ZnTCPPSpm4 (2.5 µM red curve; 5 µM blue curve; 7.5 µM green curve).*

To support the hypothesis that ZnTCPPSpm4 does not affect the stability of the miRNA and to verify how other porphyrins modulate its stability we performed CD melting experiments at the ratio [porphyrin]/[miR] 1:1 (**Fig. 43 (a)**) and 3:1 (**Fig. 43 (b)**). As we expected miR-26b-5p in presence of 2.5 µM and 7.5 µM of ZnTCPPSpm4 showed a melting temperature (Tm) of ~ 23°C and ~ 26°C, respectively which are similar to those of the miR alone in solution (Tm ~ 24 °C), confirming that this porphyrin does not destabilize the micro-RNA's secondary structures. Other porphyrins used at the ratio [porphyrin]/[miR] 1:1 do not affect the stability of the miRNA, except for the H2TCPPSpm4 which displayed a Tm of ~ 17 °C. Most notably at the ratio 3:1 for H2T4 and H2TCPPSpm4 it was not possible to perform CD melting experiments since at this porphyrin concentration the miRNA structure is almost completely unwind, in line with spectroscopic data described previously, therefore the noise of the CD signal was too high that does not allow to obtain a melting curve.

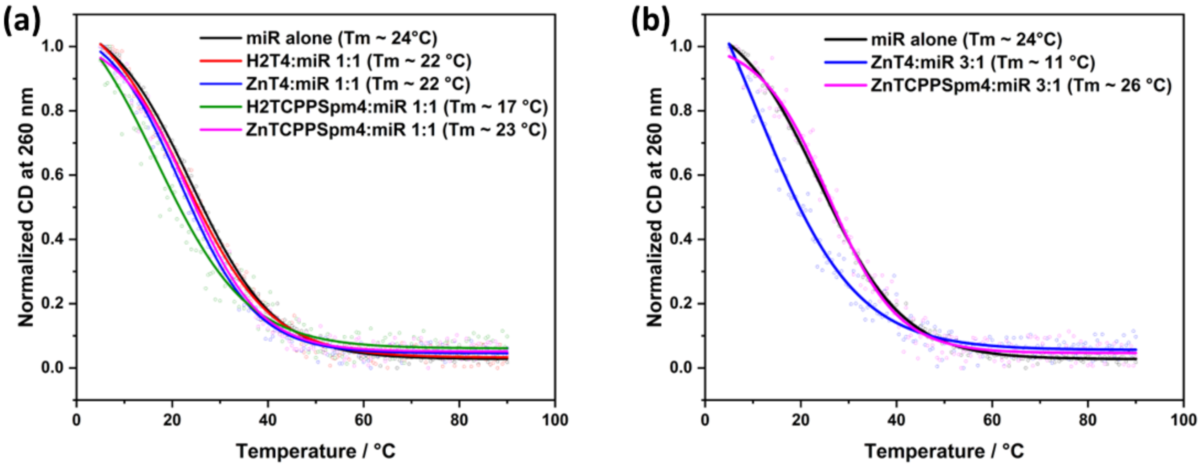


Fig. 43 (a) *Normalized CD melting curves of miR-26b-5p [2.5 μM] alone in PBS buffer (black line) and with the presence of H2T4 [2.5 μM] (red line), ZnT4 [2.5 μM] (blue line), H2TCPPSpm4 [2.5 μM] (green line) and ZnTCPPspm4 [2.5 μM] (magenta line), respectively.* ***(b)*** *Normalized CD melting curves of miR-26b-5p [2.5 μM] in PBS buffer with the presence of H2T4 [7.5 μM] (red circles), ZnT4 [7.5 μM] (blue line) and ZnTCPPspm4 [7.5 μM] (green line), respectively.*

Finally, we estimated the binding constant between miR-26b-5p and each porphyrin based on their absorption data using Eq. 1 [202]. In this equation, 'ν' represents the number of moles of bound porphyrin per mole of total miRNA, 'L' signifies the molar concentration of free porphyrin, 'n' denotes the number of consecutive lattice residues (or miRNA sites) that become inaccessible due to the binding of a single porphyrin molecule, and 'K_{app}' stands for the apparent association constant for porphyrin binding to a site on the nucleic acid. The values of 'ν' and 'L' required for this analysis were determined by using the Peacocke and Sherrett method (see SI for details) [203].

$$\frac{\nu}{L} = K_{app}\,(1 - n\nu)\left(\frac{1-n\nu}{1-(n-1)\nu}\right)^{n-1} \quad \text{(Eq. 1)}$$

In a typical linear Scatchard plot (**Fig. S11 (a), (c) and (d)**), the slope extrapolation corresponds to K_{app} according to Eq. 1. Thus, our estimated K_{app} values are listed in **Table 2**.

As we expected, similar K_{app} values were obtained for H2T4 and H2TCPPSpm4, which is in line with the spectroscopic results we obtained. However, for ZnTCPPSpm4, a higher K_{app} was found, indicating that the pre-organization of ZnTCPPSpm4 can enhance its affinity for miR-26b-6p.

***Table 2** Apparent association constants (K_{app}) for H2T4, H2TCPPSpm4 and ZnTCPPSpm4 with miR-26b-5p obtained from the Scathard plots. It was not possible to estimate Kapp for ZnT4 because the Scatchard plot deviates from linearity (see SI for details).*

miRNA	**Porphyrin**	**K_{app} M^{-1}**
miR-26b-5p	H2T4	4.84×10^5
miR-26b-5p	ZnT4	-
miR-26b-5p	H2TCPPSpm4	3.21×10^5
miR-26b-5p	ZnTCPPSpm4	1.34×10^6

Aggregation State of the ZnTCPPSpm4 in different buffer composition

In a recent work [118] was reported that ZnTCPPSpm4 porphyrin does not aggregate in 10 mM lithium cacodylate buffer (pH 7.2; 5 mM KCl; 95 mM LiCl) up to the porphyrin concentration of 40 μM, showing the Soret band of the monomeric form at 424 nm. However, during the spectroscopic characterization of ZnTCPPSpm4 we noticed that this porphyrin tends to self-aggregate in PBS. For this reason we decide to investigate the aggregation state of ZnTCPPSpm4 in three popular buffers at the same pH value: 5K (cacodylate) buffer, phosphate sodium salt buffer, and PBS. These buffers were chosen as they frequently appear in literature to mimic physiological environments. However, notable differences exist among them in terms of their molecular composition (for instance, cacodylate in 5K, and phosphate in the remaining two) and ionic strength (5K and PBS share similar ionic strength compared to the phosphate buffer). Our spectroscopic studies interestingly showed that the aggregate states of ZnTCPPspm4 can vary based on the particular buffer used to prepare the working solution.

The ZnTCPPspm4 bears four spermine pendants and similar to its naked H2TCPPspm4 counterpart, it is expected to possess tuneable positive charges in an aqueous environment due to the three protonatable nitrogen atoms in each spermine arm. It is reasonable to assume that the outermost primary amino-groups have a pKa_1 around 8, while the two secondary amino-groups have a pKa_2 of approximately 5.8 [192,204]. The final protonation step is not present because the fourth nitrogen atom is actually part of the formation of amide groups with the porphyrinoid's carboxyphenyl moieties [192]. Furthermore, the zinc cation's presence can restrict the protonation of the inner core below pH = 5.0. Based on these observations, we can conclude that the ZnTCPPspm4 is barely positively charged at mild basic conditions (pH range 7.2–7.8).

Hence, the repulsion between the cationic porphyrinoids can effectively hinder the aggregation of ZnTCPPspm4 in ultra-pure water [192]. This characteristic is clearly demonstrated by monitoring the spectroscopic titration of ZnTCPPspm4 porphyrin in water solution at pH = 7.3

(**Fig. 44**, **full triangles and Fig. S12**). In particular, all UV-vis spectra (**Fig. S12**) show the same Soret band (λmax = 423 nm) and the two Q-bands at 558 nm and 600 nm in the investigated concentration range (1 – 10 μM). The plot of the Soret band against the porphyrin concentration (**Fig. 44, full triangle**) indicates a typical Lambert-Beer progression, where the experimental slope of the linear fit (~ $1.32 \cdot 10^5$ M^{-1} cm^{-1}) is very similar to the experimentally obtained molar extinction coefficient for ZnTCPPspm4.

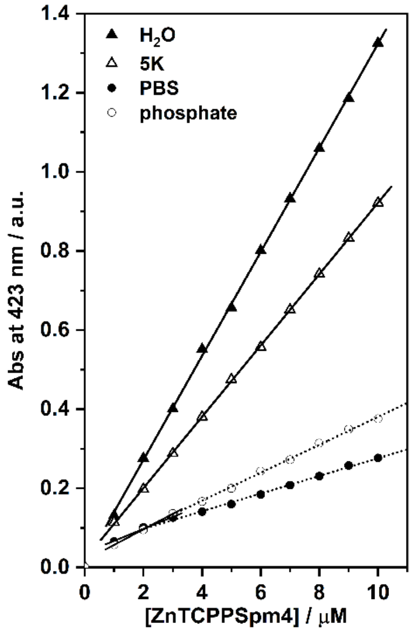


Fig. 44 Absorbance variation of the ZnTCPPspm4 Soret band (λ_{max} = 423 nm) observed upon the increase of the porphyrin concentration (μM) in: (i) ultra-pure water (pH = 7.3; full triangles); (ii) 5K buffer (pH = 7.3; empty triangles); (iii) PBS buffer (pH = 7.3; full dots); (iv) phosphate buffer (pH = 7.3; empty dots). The solid line refers to the linear fit for each titration data set in all buffers. In the case of PBS and phosphate buffer, the dotted line corresponds to the linear fit right after the break-point at 3μM

The titration of ZnTCPPspm4 conducted in the 5K buffer displays a comparable trend (**Fig. 44**, **empty triangles and Fig. S13**). However, the different environment of the porphyrinoid lead to a slightly decreased intensity within its Soret band compared to ultra-pure water. Specifically, as the concentration of porphyrin increases, this reduced intensity becomes more prominent, achieving up to 20-30% reduction at higher concentrations (**Fig. 44**, **empty triangles**). The buffering agent in the 5K is the acid-base pair -cacodylic acid/cacodylate sodium salt- with a pKa value of 6.27. Consequently, we can suppose that under our specific pH conditions (pH = 7.3), the majority of cacodylic acid molecules are deprotonated into cacodylate anions, which could potentially interact with the weakly positive-charged ZnTCPPspm4. However, due to the weak electrostatic interactions, we did not observe significant spectroscopic alterations for the ZnTCPPspm4, as highlighted in the aforementioned titration plot (**Fig. 44**, **empty triangles**). Moreover, the high ionic strength in 5K buffer, owing to the presence of potassium chloride (5mM) and lithium chloride (95mM), can further mitigate any coulombic interaction between the cationic porphyrin and the anionic cacodylate.

A dramatically different behaviour of ZnTCPPspm4 is observed in the context of phosphate-based buffers. For instance, the titration in PBS solution results in a noticeable broadening of the porphyrin's Soret band, accompanied by a pronounced hypochromic effect (**Fig. S14**). Additionally, when we plot the absorbance at 423 nm against the concentration of ZnTCPPpm4, a typical breakpoint is displayed at 3μM (**Fig. 44, full dots**), suggesting that the PBS buffer induces a controlled self-assembly of ZnTCPPSpm4 porphyrin.

Nonetheless, to exclude the influence of ionic strength, we performed the same titration experiment solely in phosphate buffer (pH = 7.3), without other salts (e.g., sodium chloride and/or potassium chloride). Interestingly, the spectroscopic analysis indicates a virtually identical state characterized by both hypochromicity and the expansion of the primary porphyrin visible band (**Fig. S15**), along with a clearly defined breakpoint at 3μM (**Fig. 44**, **empty dots**). Consequently, comparing these two recent titration experiments with earlier ones, allows us to propose that the phosphate anions may promote a form of aggregation in our ZnTCPPSpm4.

In fact, a detailed spectroscopic investigation of our system across all buffers (**Fig. 45**) reveals this complex scenario. As previously discussed, the growing mutual interaction seen in the various buffers results in hypochromism and the broadening of the porphyrin Soret band (**Fig. 45 (a)**), with a singular main component at $\lambda \approx 423$ nm for the system in ultra-pure water and 5K (as observed in the second derivative spectra in **Fig. 45, inset**). However, significant transformations occur in PBS and phosphate, where the second derivative of the Soret band reveals additional components at λ≈409 nm and λ≈440 nm (**Fig. 45, inset**). Therefore, aligning with the literature and the molecular exciton theory [205], we can assert that phosphate-based buffers facilitate the creation of two distinct types of ZnTCPPspm4 aggregates: (i) the absorption band with a blue shift at λ≈409 nm represents face-to-face aggregates, while (ii) the edge-to-edge aggregates are described by a red shifted absorption band at λ≈440 nm.

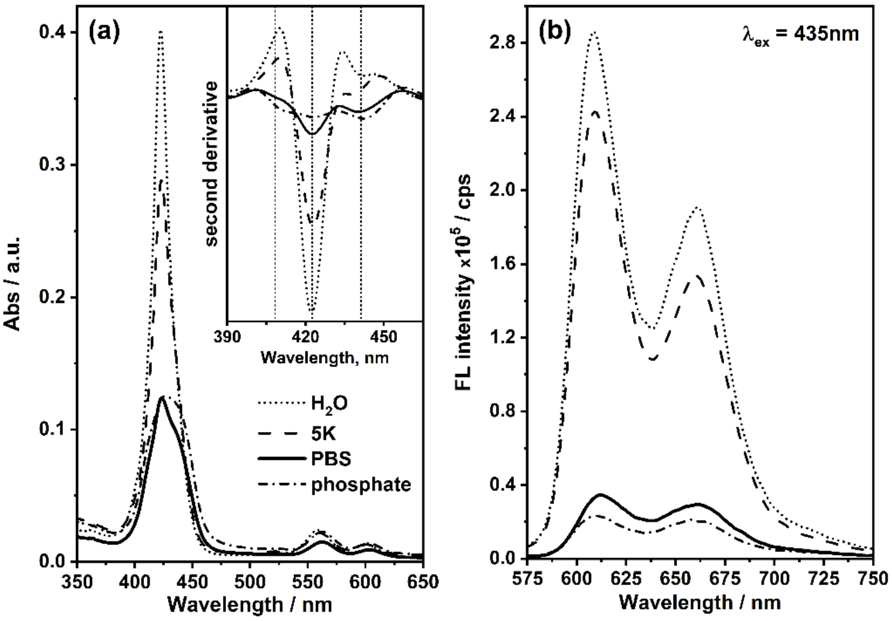


Fig. 45 (a) *UV-Vis (inset: second derivative) and* ***(b)*** *fluorescence (λ_{ex}=435 nm) spectra at pH = 7.3 of ZnTCPPspm4 (3 μM) in various buffer solutions*

Given that most of the spermine pendants of the porphyrin are positively charged at pH = 7.3, the first driving forces for the ZnTCPPspm4 aggregation could be the electrostatic interactions and the selective H-bond interaction with phosphate molecules. Indeed, it has been demonstrated that the supramolecular interaction of cationic porphyrinoids and related derivatives with phosphate or polyphosphate compounds results in a diverse range of adaptable porphyrin aggregates [206–208]. Furthermore, polyamines, such as spermine and spermidine, have a specific affinity for phosphate, facilitated by electrostatic and (N)H···O(P) hydrogen bonding. This results in a stable complex where phosphate ions link different polyamine chains together [209–211].

As already described for H2TCPPSpm4, for ZnTCPPSpm4 porphyrin as well, the emission spectra is formed by the initial excited state, S1, and the charge-transfer state (CT) between the porphyrin core and its substituents at meso position. Polar solvents or restriction of peripheral substituents can cause fluorescence quenching because of the coupling of these two states (S1-CT).

Although we have used highly polar buffers, the rigid rotation of the spermine arms results in a sharp and split peak at 608 and 660 nm in the ZnTCPPsmp4 emission, particularly in water and 5K buffer where aggregation phenomena are absent [118,198]. The fluorescence intensity at λ_{ex} = 435 nm of ZnTCPPSPm4 significantly reduces in phosphate and PBS buffers (**Fig. 45 (b)**). The observed emission quenching in these buffers could be attributable to the particular assembly of porphyrin aggregates with phosphate molecules.

To further corroborate the formation of ZnTCPPspm4 aggregates, resonance light scattering (RLS) measurements were performed.

The RLS signal of ZnTCPPspm4 in water (pH = 7.3) is very low, as shown in **Fig. 46** (inset), pointing to a lack of any form of aggregation. This aligns with the results from UV-Vis (**Fig. 45 (a)**) and emission studies (**Fig. 45 (b)**). A similar pattern is observed throughout the RLS titration in 5K buffer (**Fig. 46**, and **Fig. S16**). On the other hand, a significant increase in RLS intensity in PBS and phosphate solution indicates that the buffer induced ZnTCPPspm4 aggregation (**Fig. 46**), consistent with other spectroscopic techniques. Notably, from the RLS titration in PBS/phosphate buffer (**Fig. 46**), we detected a stoichiometric breakpoint at 3μM, correlating with the UV-vis data. Moreover, the shield-effect produced by high ionic strength on porphyrin-porphyrin interactions may account for the diminished RLS signal intensity in PBS.

Collectively, under our experimental conditions, the RLS data enables us to (i) exclude the possibility of buffer-induced self-aggregation in both ultra-pure water and 5K buffer; and (ii) confirm the formation of substantial ZnTCPPspm4 assemblies initiated by specific interactions with phosphate molecules.

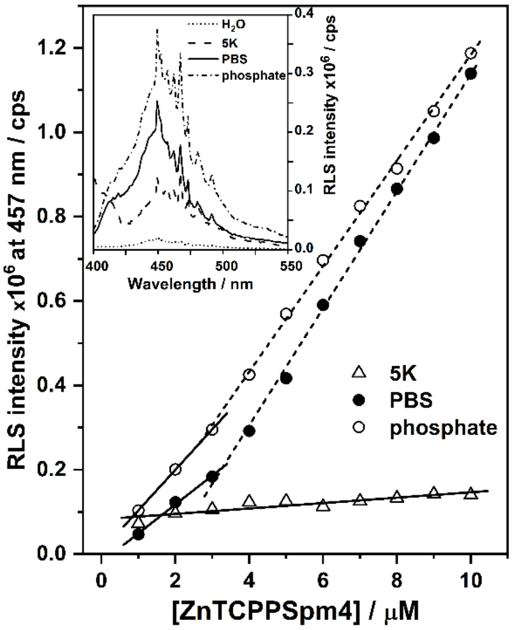


Fig. 46 *RLS intensity at 457 nm of the ZnTCPPspm4 (λ = 457 nm) upon the increase of the porphyrin concentration (μM) in: (i) 5K buffer (pH = 7.3; empty triangles); (ii) PBS buffer (pH = 7.3; full dots); (iii) phosphate buffer (pH = 7.3; empty dots). The solid line refers to the linear fit for each titration data set in all buffers. In the case of PBS and phosphate buffer, the dotted line corresponds to the linear fit right after the break-point at 3μM. Inset: RLS spectra in various buffers are reported ([ZnTCPPspm4] = 3μM, pH = 7.3)*

2.3.3. Final remarks

In conclusion, we reported the supramolecular interaction between several achiral porphyrins and a mature miRNA. It was demonstrated that porphyrins can be used as chiroptical probes for miR-26b-5p and depending on their characteristics, different binding modes were observed, as evidenced by the different ICD signal produced in the Soret region. In particular, H2T4, up to the 2:1 ratio, appears capable to form weak aggregations with the miRNA structure forming dimers between porphyrins and miRNA molecules. After this ratio it aggregates onto the miRNA structure via electrostatic interactions with the phosphate backbone, whereas a pseudo-intercalation occurs in presence of ZnT4. For H2TCPPSpm4 a negative ICD signal is displayed until the ratio 1:1, at higher porphyrin concentration it disappeared suggesting that spermine arms are able to disrupt the miRNA’s hydrogen base pairing, penalizing the formation of well-ordered secondary structures. For ZnTCPPSpm4 a pre-organized multi-porphyrinic system occurs and then it interacts weakly with the miRNA structure.

Lastly, CD-melting experiments demonstrated that porphyrins used in this work does not affect the stability of the miRNA at the ratio 1:1, excluding H2TCPPSpm4 that even at this low concentration showed a strong destabilizing behavior. Intriguing at the ratio 3:1 the unique porphyrin that does not destabilized the miRNA structure was ZnTCPPSpm4, which for this reason makes it the best candidate for further investigation and conformational studies on mature miRNA structures. These porphyrins, especially H2TCPPSpm4, could be used as denaturing agents for mature miRNAs, deactivating in this way their biological function even at micromolar concentrations, making porphyrins potential drugs in connection with pathological contexts partially/totally caused by some miRNA upregulation.

The presence of the Zn (II) in the central core of both porphyrins changes drastically the type of interaction with miRNA, indeed weaker interactions and minor destabilizing effect with the miRNA were observed, suggesting potential use of this porphyrin as chiroptical probes for miRNAs in spatial transcriptomics and single-cell analysis.

Overall, this study could significantly advance our understanding of porphyrin-miRNA interactions, providing potential new tools for biological research, and suggesting possible future therapeutic applications.

3. G-quadruplexes

G-quadruplexes (G4s) are non-canonical nucleic acid secondary structures, primarily composed of guanine-rich sequences. These structures are formed by the stacking of square planar guanine tetrads, or G-tetrads, which are planar arrangements of four guanine bases that are held together by Hoogsteen hydrogen bonding (**Fig. 47 (a)**).

Many studies demonstrated that G-rich DNA or RNA sequences with the ability to form G4 can fold into dG4s and RNA G-quadruplexes (rG4s) (**Fig. 47 (b)**), influencing a variety of cellular functions [212].

G4 formation can be predicted based on primary sequences. Typically, sequences with potential to form G-quadruplexes have four sections of three continuous guanines separated by three loops, with the loop length being between one to seven nucleotides [213].

Monovalent cations, and such metal ions, can intercalate into the central core of a G-tetrad or bind between two G-quartets. This interaction stabilizes hydrogen-bonded tetrads and amplifies base-stacking interactions. Due to the central channel's specific geometry and size, only cations with suitable charge, size, and dehydration energy can coordinate a G4. The G4-stabilizing effects of monovalent ions decrease in the following order: K+ > Na+ > Li+ [214].

G4s can be categorized into two main groups based on the strands number that form them: intermolecular and intramolecular. Intermolecular G4s arise from two or four distinct strands of DNA or RNA, while intramolecular G4s originate from a singular DNA or RNA strand (**Fig. 47 (c)**).

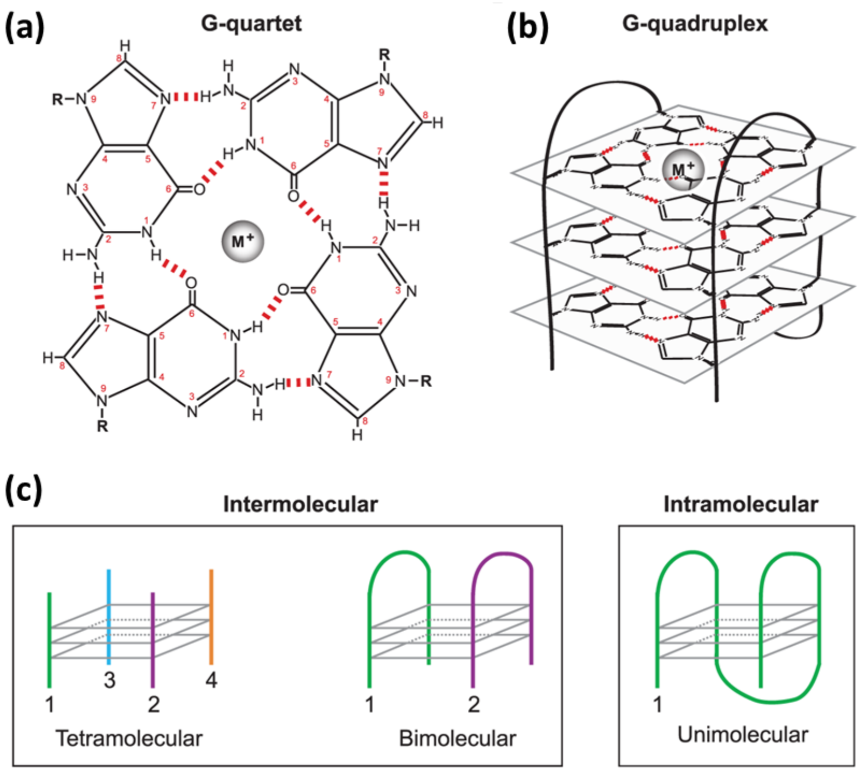


Fig. 47 *Representation of* ***(a)*** *G-tetrad with a monovalent cation, Hydrogen-Hoogsteen bonds are indicated by red dashes* ***(b)*** *An intramolecular G-quadruplex structure, consisting of three planar G-quartets and* ***(c)*** *G-quadruplex classifications, based on the number of strands involved. Adapted from ref* [215]

From a topological perspective, G4s can be parallel with all four G-strands in the same direction, hybrid or mixed with adjacent strands being both co-directional (parallel) and anti-directional (antiparallel), or antiparallel with all adjacent G-strands anti-directional to each other. (**Fig. 48 (b)**). Intramolecular G4s contain different types of loops, such as propeller for connecting parallel strands, lateral for connecting adjacent antiparallel strands, and diagonal for connecting antiparallel strands across the G-tetrad core (**Fig. 48 (a)**).

rG4s typically form a parallel topology, which is due to the anti-conformation of glycosidic bonds in ribonucleosides. Despite a few exceptions [216–218], the majority of documented rG4s so far have a parallel topology.

Contrarily, dG4s display versatility in their topological formation—they can form parallel, anti-parallel, or hybrid structures and may switch between these, depending on the physiochemical conditions and sequence characteristics [219]. Guanines in a G-tetrad may have two conformations:

anti and syn. All parallel G4s have all the guanine glycosidic angles in the *anti* conformation, whereas antiparallel and hybrid G4s have both *syn* and *anti*.

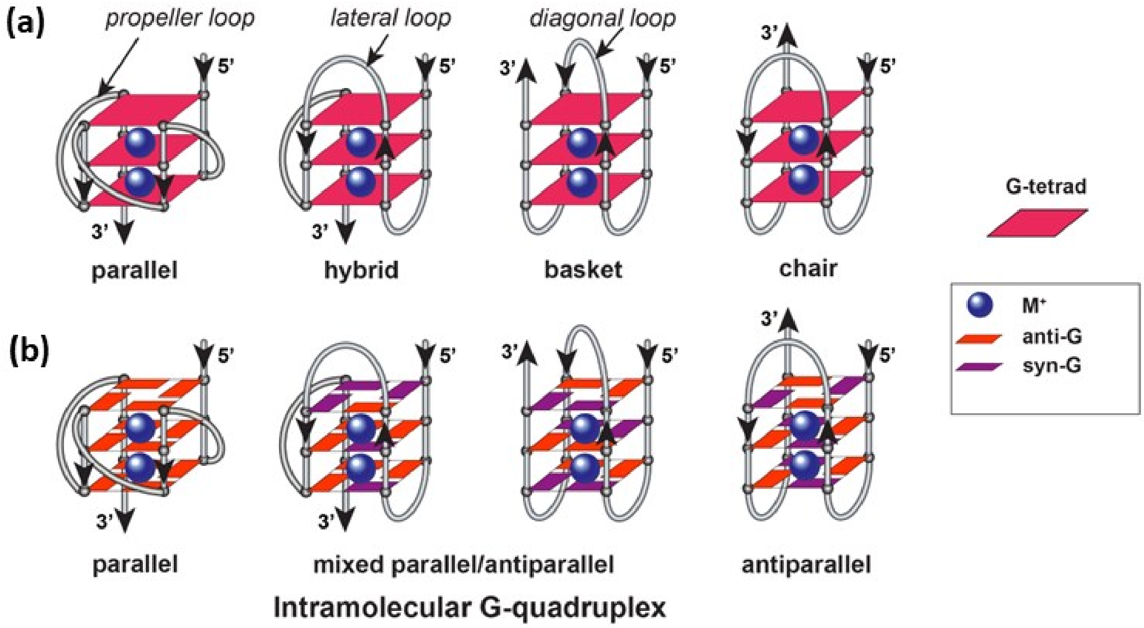


***Fig. 48** Representation of G-quadruplexes with different topologies with a corresponding **(a)** loop and **(b)** syn-anti orientation of guanosine residues. Adapted from ref* [220].

rG4s form a structure that is more thermodynamically stable, compact, and less hydrated compared to their DNA counterparts (**Fig. 49**) [221,222]. The key distinction between RNA and DNA quadruplexes lies in the substitution of deoxyribose sugar with ribose sugar and the replacement of the thymine residue with uracil in RNA quadruplexes.

The 2′ hydroxyl group present in the ribose sugars of rG4s results in more intramolecular interactions within the structure, leading to an increase in its stability. Interestingly, these 2′ hydroxyl groups impose structural limitations on the G-quadruplex topology, preventing it from achieving a syn-conformation, which is necessary for an antiparallel topology. Instead, they limit the orientation of the base around the glycosidic bond to an anti-conformation, leading to C3′ endo puckering and constraining the RNA G-quadruplex to adopt a parallel topology [223]. This suggests that the topology of RNA G-quadruplexes is less influenced by environmental conditions compared to DNA G-quadruplexes [224,225].

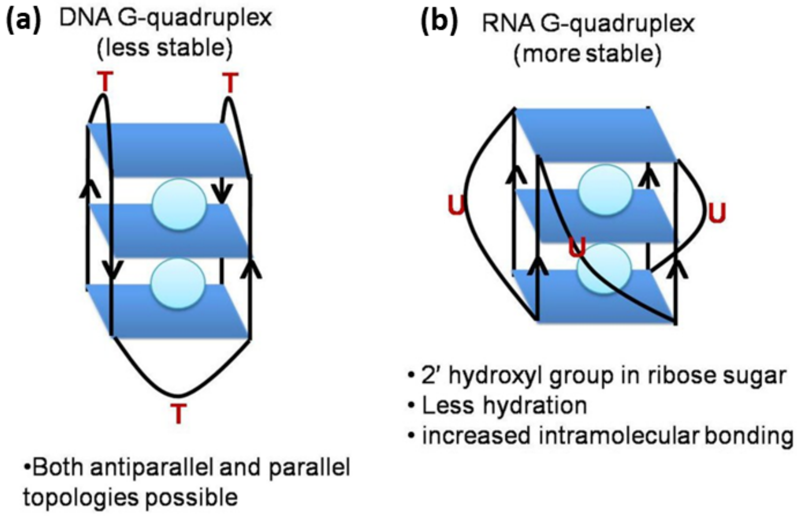


Fig. 49 *Differences between* ***(a)*** *DNA G-quadruplex and* ***(b)*** *RNA G-quadruplex. Adapted from ref.* [226]

Telomeric RNA G-quadruplexes

The functional roles of G-quadruplexes in biological systems were initially discovered in relation to the telomeric regions of chromosomes, where the highest concentration of this secondary structure is found. Telomeric regions of mammalian chromosomes, typically spanning 2-10 kb, comprise repetitive sequences of d(TTAGGG) and a G-rich strand that forms a 100-200 nucleotide long 3′ overhang, known as the G-tail.

Telomeres, along with the enzyme telomerase, are critical for maintaining the integrity of chromosome ends. A protein complex known as shelterin further supports the stability of these telomeric ends. It has been suggested that the continuous formation and resolution of G-quadruplex structures within the repeating sequence d(TTAGGG) contribute to the preservation of chromosome ends [227,228].

Considering that these telomeric G-quadruplexes play a role in cancer biology, significant research efforts are being made towards understanding their therapeutic potential and how to target these structures effectively [229].

Originally, the telomeric region was assumed to be transcriptionally inactive. However, a research study conducted in 2007 revealed that the C-rich strand of telomeric DNA is subject to active transcription by the DNA-dependent RNA polymerase II (Pol II). This process produces transcripts of varying lengths (0.1- 9.0 kb), termed telomeric repeat containing RNA (TERRA) [230].

TERRA was found to be located specifically in the nucleus, in close proximity to the telomeric regions. These long non-coding RNAs participate in several essential cellular processes, such as the regulation of telomere length and inhibition of telomerase activity [231,232].

TERRA, which is formed by UUAGGG repeats, has been extensively studied to verify the formation of G-quadruplexes in these telomeric RNAs. The research indicates that TERRA can form stable G-quadruplexes, and evidence of their in vivo existence has provided insight into their previously unknown roles in telomere biology [233].

A broad spectrum of proteins interacts with TERRA RNA, these proteins are implied to regulate the location and local concentration of TERRA RNA[234]. Notably, the shelterin proteins TRF1 and TRF2 can bind directly to TERRA, as well as to components of the origin recognition complex and H3TK9. Specifically, TRF2 can simultaneously bind to both TERRA RNA G-quadruplex and telomeric DNA G-quadruplex [235].

Therefore, these findings make it plausible that protein binding and associated functions are facilitated via the formation of G-quadruplex structure in TERRA.

TERRA serves as a key example of a non-coding RNA that potentially plays a role in maintaining the integrity of chromosome ends, possibly by acting as a scaffold for protein recruitment. As more literature emerges about long non-coding RNAs with unknown functions, and with bioinformatic predictions suggesting G-quadruplex presence in non-coding RNAs, it is speculated that the occurrence of G-quadruplexes in non-coding RNA might be a widespread phenomenon with potentially significant functions.

3.1. Porphyrins as ligands for G4s: State-of-the-art

Utilizing the secondary structure of nucleic acids to manipulate cellular processes offers a promising approach to regulate gene expression. The employment of small molecules able to bind to G4s allowing the conversion of chemical biology into therapeutic development [236].

Given that G4s at telomeric ends and promoter regions are recognized targets for cancer treatments, there has been substantial research conducted on ligands that target dG4s in recent decades [237]. Various small molecules such as porphyrin (H2T4), acridine, pentacridium, quinacridine, telomestatin, naphthalene diiamide, bisquinolium, and their derivatives have been identified to selectively bind and stabilize DNA G-quadruplexes [238]. Multiple crystal structures of DNA G-quadruplex-small molecule combinations have been reported. Some of these ligands have also been examined for their ability to bind to and modulate the function of DNA G-quadruplexes in cellular environments [220].

Small molecules designed to selectively bind to G-quadruplex generally exhibit certain common characteristics: (i) A large flat aromatic core to optimize π-stacking interactions with G-quartets of the G4s. (ii) A positive charge to counterbalance the negatively charged phosphate groups of the nucleic acid backbone. (iii) Positively charged side chains with functional groups to improve interactions with grooves/loops/inner G-quartet core of the G4s [236].

Cationic porphyrins, which reflect all aforementioned features, were among the initial ligands analyzed for their G4s binding and stabilization capabilities. H2T4, a tetracation metal-free porphyrin, exhibits molecular dimensions similar to G-tetrads and demonstrates efficiency as an inhibitor of human telomerase in vitro [239]. The binding manner of porphyrin–G4 can be influenced by various factors such as the type of G4, DNA sequence, and arrangement of strand orientations. Additionally, the groove sizes of different antiparallel G-quadruplexes, along with the loop structures and adjacent non-G-tetrads, may also impact binding specificity.

Three potential binding modes have been suggested: (i) binding to the quadruplex's top and bottom (capping) [131], (ii) external binding to G4s strands [132], or (iii) intercalation between guanine tetrads [133]. Each of these binding modes has been demonstrated to enhance the stability of the G4 structure.

A study conducted by Sherdy et al. [240] demonstrated through UV-Vis titration that the Soret absorption band of H2T4, when in excess of T4G4 (which forms parallel stranded quadruplexes with K+), is red shifted (10-12 nm) and exhibits a considerable hypochromic effect (35-40%). Notably, the Soret red shift of H24 is somewhere between the shifts observed for intercalation (≥15 nm) and those for external binding (≤8 nm), while the hypochromic effect aligns with the characteristics of an intercalation mechanism. These spectroscopic characteristics and energy transfer studies confirm that H2T4 interacts with T4G4 via intercalation at low concentrations (0.25-0.4 mM DNA and 3-5 µM H2T4). At higher concentrations, a more complex mixed binding mechanism appears.

Intercalation of H2T4 has also been observed with distinct DNA G4s (G2, G3 and G4, **Fig. 50**) [241].

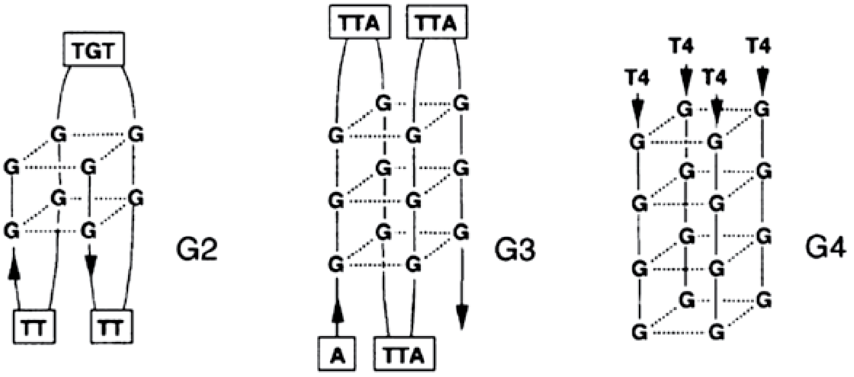


Fig. 50 *Scheme of the G4s studied in reference* [241].

Yamamoto et al. have reported intercalative binding [242], where they examined the interaction of H2T4 with the all-parallel type G-quadruplex d(TTAGGG), which forms under different potassium buffer concentrations. Their findings, based on NMR, UV-Vis absorption, and CD studies, show H2T4's intercalation into the A3pG4 step of the G-quadruplex DNA. Evidence for this includes an 18 nm red shift and 66% hypochromicity of the Soret band in the UV-Vis region, along with a negative induced CD band (**Fig. 51**)

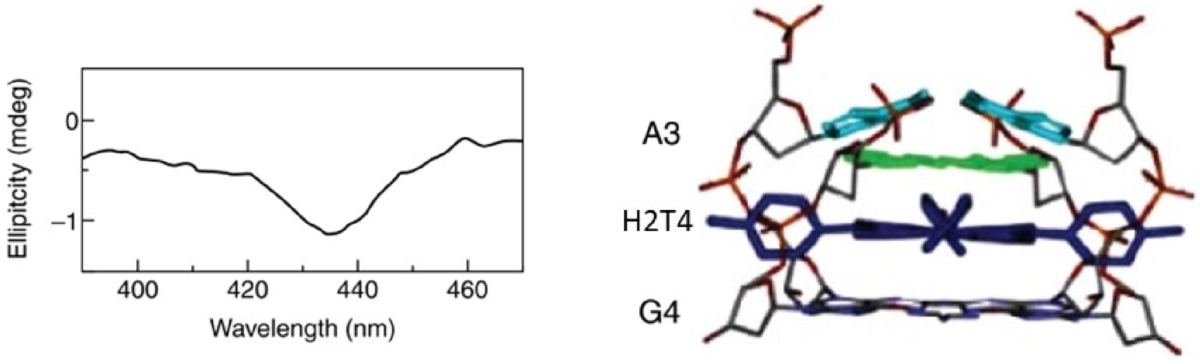


Fig. 51 *CD spectra of H2T4, 6 µM, in the presence of d(TTAGGG). Models of intercalation of H2T4 into the A3pG4 step proposed from the molecular mechanics calculation. Adapted from reference* [242].

Neidle et al. [243] have provided a crystal structure of the bimolecular human telomeric quadruplex composed of the d(TAGGGTTAGGG) sequences interacting with H2T4. Their structural analysis reveals that H2T4 does not intercalate, but one molecule is stacked onto an AT base pair, thereby forming a TT propeller loop. The second H2T4 molecule is externally stacked onto thymine bases at the edges of both TTA and TT loops. There is no direct contact between the ligand and any G-tetrads. The structure demonstrates that the stable parallel quadruplex topology is maintained upon ligand binding. Interestingly, H2T4 induces a significant change in one loop, transforming it from a trinucleotide TTA loop into a dinucleotide TT loop. This change in loop topology provides a more dynamic and complex interface for the binding of H2T4.

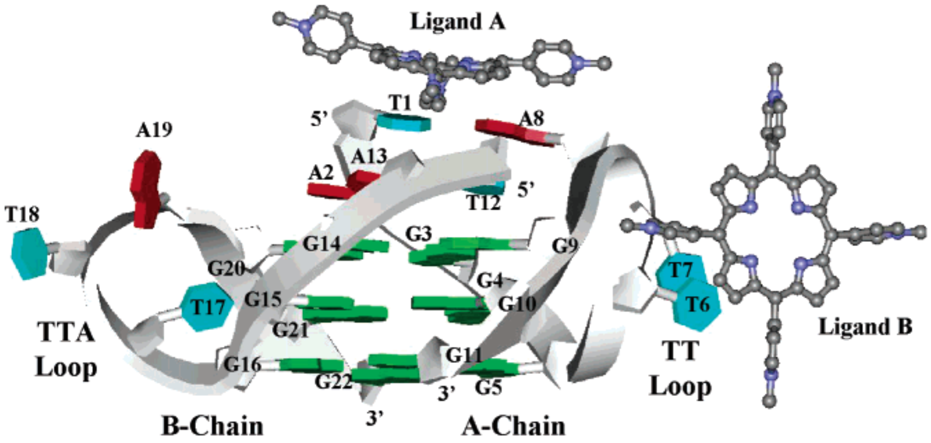


Fig. 52 *Schematic view of the bimolecular quadruplex-H2T4 structure, showing the folding topology, the numbering of nucleotides, the extended TTA loop geometry, and the two H2T4 molecules bound per asymmetric unit. Guanine bases are colored green, adenines red, and thymines cyan. Adapted from ref* [243].

Zhang et al. [244] investigated the impact of the loop region on G4 structures. Circular dichroism (CD) measurements showed that the G-quadruplex structure of the single-strand oligonucleotide d(TTAGGG)2 (12-mer) shifts from a parallel structure to a mixed parallel/antiparallel structure in the presence of K+, upon binding with H2T4.

They performed similar experiments using shorter (6-mer 5′-TTAGGG-3′) and longer (24-mer 5′-TAGGG(TTAGGG)3T-3′) DNA strands. While a 6-mer remains in a random conformation upon porphyrin addition, the 24-mer sequence changes into an antiparallel G4. In the case of 12- and 24-mer DNAs, both the antiparallel G4 structures have end loops for the stacking of H2T4, which will further stabilize the G4 with the help of the stacked porphyrins. On the other hand, the parallel G4s only has external loops and cannot provide an adequate space for accommodating H2T4.

From these findings, it is reasonable to conclude that the end-loops play a pivotal role in the stabilization of G4s induced by H2T4.

Hurley et al. [245].used gel mobility shift experiments and a helicase assay to study the interactions between H2TMPyP2 (ortho), H2TMPyP3 (meta), and H2T4 (para) with both parallel and antiparallel G-quadruplexes.

Even with very similar structures, these porphyrins demonstrated differing binding affinities to antiparallel G4 DNA. H2TMPyP2 did not show significant interaction with either parallel or antiparallel G-quadruplexes. H2T4, however, showed greater specificity to antiparallel G-quadruplexes.

The study suggests that the location of the N-pyridyl group significantly influences the free rotation of the pyridyl groups at the meso positions and the relative dihedral angles between the porphyrin core and pyridyl groups. These factors, in turn, affect the porphyrins' interactions with G-quadruplexes. H2TMPyP3 is better suited to fit the pockets between the end G-tetrads and the loops in a parallel G-quadruplex compared to H2T4, while H2TMPyP2 does not fit any site in the G-quadruplex-parallel.

In summary, this study underscores how small changes in molecular structure can significantly alter interactions with biological targets, with potential implications for drug design and therapeutic applications.

Yatsunyk and coworkers [246] investigated the impact of the central metal on the interaction with G4. They studied the interactions between H2T4 and its Zn(II), Cu(II), and Pt(II) derivatives and G4 DNA oligonucleotide d(TAGGG)2 in the presence of potassium (K+).

They found that CuT4 and PtT4, like H2T4, form square planar complexes. ZnT4, on the other hand, adopts a square pyramidal geometry, which should preclude an intercalative binding mode.

They observed that d(TAGGG)2 forms a propeller-type parallel bimolecular G4 under high K+ concentrations. Conversely, a buffer with a low concentration of K+, or any concentration of lithium (Li+) or sodium (Na+), could not induce G4 structures in d(TAGGG)2. ZnT4 was found to induce a mixture of parallel-antiparallel G4 structures.

Interestingly, the addition of H2T4, CuT4, and PtT4 to d(TAGGG)2 resulted in the complete disintegration of the secondary DNA structure by the end of the titration process. In contrast, ZnT4 was capable of promoting the formation of parallel G4s structures in human telomeric DNA.

Due to its geometry, ZnT4 is not expected to be capable of intercalating between G-tetrads. Therefore, the most probable mode of binding is end-stacking rather than groove or loop binding. These modes do not produce substantial changes in UV-Vis spectra.

Their findings indicate two equivalent and independent ZnT4 binding sites on each end of the G4. They concluded that ZnT4 facilitates the folding of d(TAGGG)2 into parallel G4, shifting the equilibrium between the unfolded and folded G4 structure towards the latter under conditions that would otherwise be unfavourable (like low concentrations of K+).

In a recent study conducted by Gomez et al. [247] they demonstrated that the insertion of a gold(III) ion into the center of the porphyrin ring enhances the porphyrin's binding affinity for the G-quadruplex structure. Their modelling studies suggest that the insertion of the square planar gold(III) ion introduces an additional positive charge to the complex. This not only increases the overall charge of the complex, but also decreases the electron density in the aromatic macrocycle of the porphyrin ring both enhancing the electrostatic and π-stacking interactions between the porphyrin and the G-quadruplex.

However only a limited number of studies have been conducted on ligands targeting RNA G4 and even less is known about the porphyrin binding with TERRA G-quadruplex. Considering that many clinically significant genes are known to contain RNA G-quadruplex, the design of small molecules that either stabilize or destabilize this structure is an interesting strategy for drug discovery against various diseases, including cancer [236]. Particularly, G-quadruplexes in telomeric RNA containing repeat (TERRA), which is transcribed from the subtelomeric loci present at chromosomal ends, and in the untranslated region of mRNA of oncogenes, are considered to be ideal targets for drug discovery.

Leeper et al. [248] demonstrated that the cationic porphyrin H2T4, can unfold an exceptionally stable all-purine RNA G-quadruplex (M3Q), located in the 5′-UTR of MT3-MMP mRNA. The direct interaction between H2T4 and M3Q, monitored by UV spectroscopy, led to a 22-nm bathochromic shift and a 75% decrease in intensity (hypochromicity) of the porphyrin's Soret band, indicative of direct binding. As determined by circular dichroism (CD), 1D 1H NMR, and native gel electrophoresis, H2T4 was found to disrupt the folded M3Q in a concentration-dependent manner. Moreover, the presence of H2T4 during the folding process prevented the M3Q RNA from forming a G4 structure. Through a dual reporter gene construct carrying either the M3Q sequence alone or the complete 5′-UTR of MT3-MMP mRNA, it was demonstrated that H2T4 could reduce the inhibitory effect of the M3Q G-quadruplex. However, the same concentrations of H2T4 were unable to affect the translation of a mutated construct. Thus, it was concluded that H2T4 possesses the capability to unfold an extremely stable RNA G-quadruplex and modulate gene activity, presumably through the disruption of the G-quadruplex structure.

A study conducted by Shao et al. [249] reported that 5,10,15,20-tetrakis(3,5-dihydroxyphenyl)porphyrin (TOHdPP) acts as an efficient fluorescent probe for targeting TERRA. It was observed that only upon binding with TERRA effectively activated the hyperporphyrin spectrum of TOHdPP, resulting in red-shifted spectral bands and enhanced fluorescence emission.

Experimental data suggested that TOHdPP interacts with the 5’ tetrads of two TERRA G4s through a 1:2 sandwich association (**Fig. 53**). The 2’-OH group on the ribose favours the loop adenine residue-extended tetrad G4 plane that is specific for TERRA. Therefore, in addition to π-stacking with the G4 tetrads, it is believed that TOHdPP also interacts with this substructure to enable efficient electron communication between the tetraphenyl substituents and the porphyrin macrocycle, a requirement for the hyperporphyrin effect.

The study also suggests that the hydrogen bonding interactions of the eight hydroxyl substituents in TOHdPP with the backbone phosphate oxygen atoms of TERRA likely further contribute to the binding selectivity. Their findings highlight the potential of TOHdPP as a selective TERRA G4 fluorescent probe and a promising TERRA-based sensor reporter.

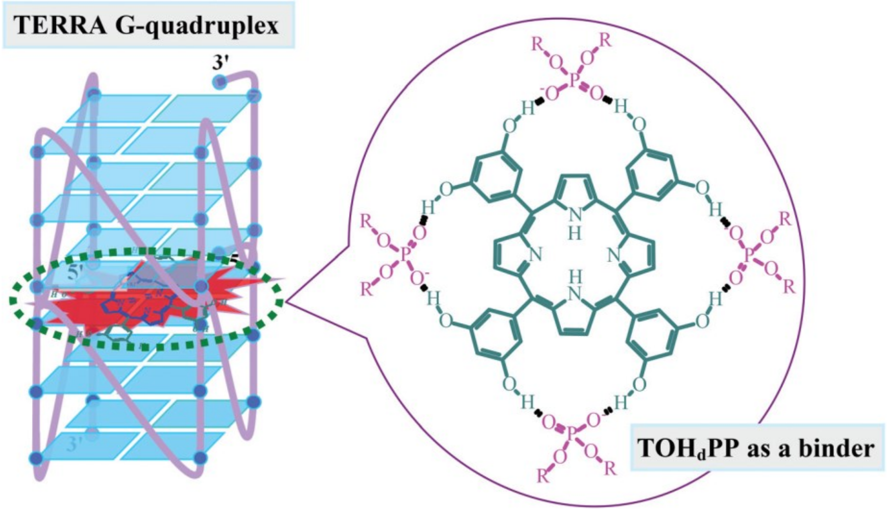


***Fig. 53** The proposed binding model of TOHdPP with TERRA G4. Adapted from ref.* [249].

Su and coworkers [250] have used various techniques such as steady-state and time-resolved spectroscopy analyses along with NMR methods to examine the binding behaviour of H2T4 to a TERRA G4 dimer, which was formed by the 10-nucleotide sequence r(GGGUUAGGGU). They show that H2T4 intercalates into the 5'-5' stacking interface of two G-quadruplex blocks with a binding stoichiometry of 1:1 (**Fig. 54**) and a binding constant of $1.92\ 10^{6}\ M^{-1}$. This finding is in agreement with the unique TERRA structural features, including the enlarged π-π stacking plane of the A·(G·G·G·G)·A hexad at 5'-ends of each G-quadruplex block. They were the first to determine the preferential binding of the π-ligand porphyrin to the 5'-5' stacking interface of the native TERRA G-quadruplex dimer through the integration of dynamic and structural characterization.

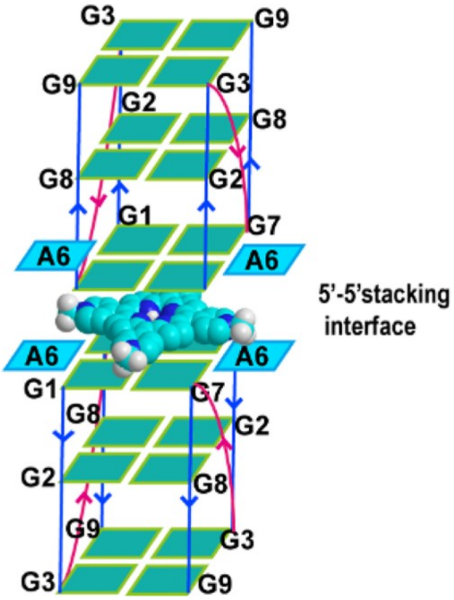


Fig. 54 *Illustration of the Intercalation Binding of H2T4 at the 5′-5′ Stacking Interface of the TERRA G4 Dimer. Adapted from ref.* [250].

In a recent study conducted by Lieberman et al. [251] they performed fluorescence polarization (FP) assays which indicated that N-Methylmesoporphyrin IX (NMM) shows a preference for binding G4 TERRA over G4 Telomeric DNA (TeloDNA). They report that NMM is known to have high selectivity for G4 DNA over other DNA structures, and particularly favours parallel forms of G4 DNA. TERRA has been reported to form a parallel-stranded G-quadruplex and to have a high affinity for related porphyrins, such as H2T4.

These and newer generation G4 interacting molecules may be useful as probes to better explore functions of telomere G4 structures in vivo.

3.2. Interactions between the telomeric RNA (TERRA) and the ZnTCPPSpm4 porphyrin: aim of the work

The human genome is a complex mosaic of sequences with varied functionality, among which telomeric RNA, a guanine (G)-rich sequence, holds a unique place due to its pivotal role in maintaining genomic integrity. Telomeric RNA can adopt a four-stranded G-quadruplex (G4) conformation, a feature that has become increasingly recognized for its profound biological implications. Telomeric RNA G4 structures are not only essential in regulating telomerase activity and telomere homeostasis, but they are also implicated in aging and a range of human diseases, including cancer. Despite their importance, understanding the molecular mechanisms underlying the formation and function of telomeric RNA G4 structures has remained a challenge, largely due to their inherent conformational flexibility and stability.

To address this challenge, it is crucial to adopt a powerful tool that can selectively stabilize and probe these structures. In this context, the ligand ZnTCPPSpm4, a zinc(II) tetracationic porphyrin, emerges as a potential ligand. This compound has demonstrated its ability in stabilizing G-quadruplexes telomeric DNA [118], offering an unprecedented opportunity to study these enigmatic structures in detail. ZnTCPPSpm4 has been found to display a remarkable affinity and selectivity towards G4 structures, thereby enabling us to examine the role of G-quadruplexes.

Our research focuses on the interaction between a telomeric RNA sequence (5’-AGGGUUAGGGUUAGGGUUAGGG-3’) and ZnTCPPSpm4, aiming to explore the biophysical characteristics, thermodynamics, and molecular mechanisms underlying this interaction. We believe that our findings will provide new insights into the regulation of telomeric RNA G4 structures and potentially pave the way for the development of therapeutic strategies targeting diseases linked with telomere dysfunction.

3.2.1. Materials and methods

The telomeric RNA (TERRA) sequence (5’-AGGGUUAGGGUUAGGGUUAGGG-3’) used was purchased from Integrated DNA Technologies IDT® and used without further purification. The solid was dissolved in ultra-pure water obtained by Elga Purelab Flex system by Veolia with purity of 18.2 MΩcm, achieving a stock solution with concentration of ~ 100 µM. Annealing is performed increasing the temperature up to 90° for 5 min, then cooling down slowly up to room temperature and finally storing the solution at 4°C for one night. Concentration of TERRA solution was checked by UV-vis measurements at 80°C using the extinction coefficient of the sequence given by IDT: ε_{260} nm = 236,900 L/(mol·cm); Then, by dilution in 5K buffer ([Sodium cacodylate] 10 mM; [KCl] 5 mM; [LiCl] 95 mM; pH 7.4) we prepared the sample solution.

Methods used such as, UV-Vis, Fluorescence, RLS and Circular Dichroism as well as buffers and ZnTCPPSpm4 preparation are the same described in paragraph **2.3.1 Materials and methods**

Denaturing polyacrylamide gel electrophoresis (PAGE)

PAGE samples contained 40 µM of RNA in 10 µL of ddH$_2$O and 30% v/v formamide was added, samples were annealed at 95°C for 5 minutes followed by fast cool down at RT for other 5 minutes before loading.

Fifteen percent native polyacrylamide gels were prepared with 7 M urea and TBE 1X, and running buffer was made with TBE 1X. The gels were pre-migrated at 300 V for at least 30 minutes, loaded with 10 µL of each sample, and allowed to run for 20 minutes at 300 V and room temperature. An oligothymidylate ladder consisting of dT15, dT24, dT30, and dT57 was used as a length marker, and a tracking dye was used to monitor the gel progress. DNA bands were visualized with Stains-All and the resulting gel was captured using a smartphone camera.

3.2.2. Results and Discussions

First of all, to confirm that the telomeric RNA (TERRA) sequence (5’-AGGGUUAGGGUUAGGGUUAGGG-3’) adopts a G-quadruplex (G4) topology we performed circular dichroism (CD) and CD-melting experiments. At pH 7.4 in 5K buffer the CD spectra of TERRA (**Fig. 55 (a)**) displays a positive band around 265 nm and a negative band around 245 nm, which is the characteristic CD signature of a parallel G4 structure [252,253].

CD-melting experiments (**Fig. 55 (b)**) show two distinct melting events, at ~ 40°C and ~ 66°C respectively.

This phenomenon can be due to several possible circumstances. One explanation could be that the G-quadruplex may exist in multiple conformations in equilibrium. In this case, the first melting event would correspond to the transition of a more stable conformation to a less stable one. The second event would then represent the full denaturation of the structure.

Alternatively, the G-quadruplex could be transitioning through an intermediate state during the melting process. Here, the first melting event would likely be the transition from the G-quadruplex to the intermediate state. The second event would then be the transition from this intermediate state to the fully denatured state.

Another scenario could be that the G-quadruplex structure consists of multiple units, such as dimers or trimers. In this case, the first melting event could correspond to the dissociation of one subunit, while the second melting event would represent the dissociation of the remaining subunits.

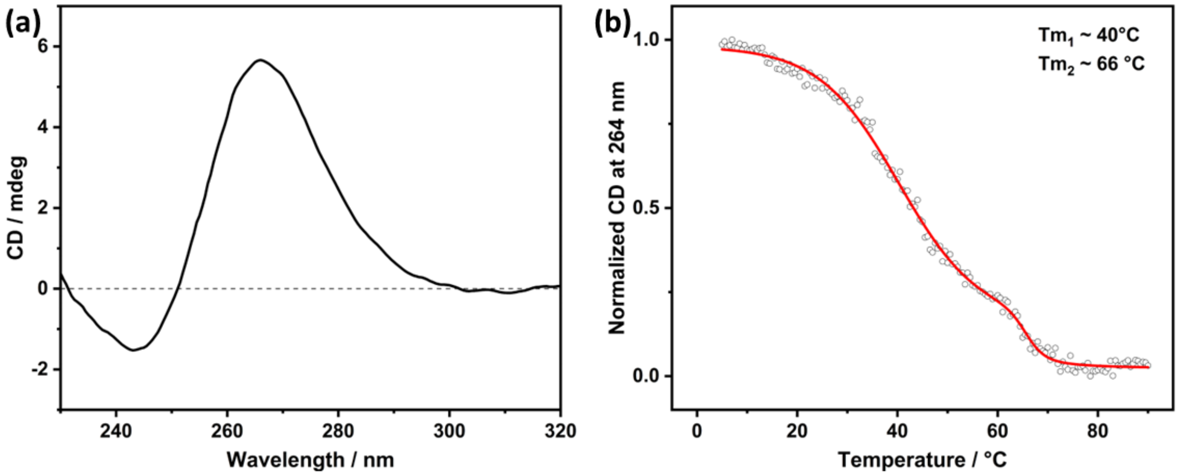


Fig. 55 ***(a)*** *CD spectra and* ***(b)*** *CD-melting curve normalized to [0,1] of TERRA [2 μM] in 5K buffer pH = 7.4. In the CD-melting spectra circles represent experimental points, which were fitted by the BiDoseResp function (red curve).*

To study the interaction of ZnTCPPSpm4 porphyrin with TERRA, we used a variety of spectroscopic techniques such as UV-vis, fluorescence, RLS and CD.

We titrated the TERRA sequence, which was at a concentration of 2 μM, by progressively increasing the porphyrin concentration up to 21 μM (**Fig. S19**).

In the absorption spectra (**Fig. 56**, **red line**), at the ratio ZnTCPPSpm4:miR 2:1, the Soret band of the porphyrin is red shifted to 433 nm (Δλ = 10 nm) and an hypochromic effect is observed (~ 30%) compared to that one of the porphyrin alone in 5K buffer (**Fig. 56**, **black line**). Which suggests a significant perturbation in the porphyrin π electrons as a result of the binding to the TERRA.

To confirm whether the binding interaction between porphyrin and TERRA is dependent on their molecular ratio, we plotted the absorbance at 423 nm against the [ZnTCPPSpm4]/[TERRA] ratio (**Fig. 56**, **inset**). The titration plot reveals four break points, suggesting the formation of four ZnTCPPSpm4:TERRA complexes at ratios of 2:1, 4:1, 6:1, and 8:1. Notably, a significant hypochromic effect, is observed when the porphyrin content in complexes exceeds the ratio of 4:1, implying a enhanced interaction between porphyrins and TERRA. However, once the 8:1 complex is formed, additional porphyrin does not alter the slope of the straight line (**Fig. 56 inset**, **magenta**

line) which became almost identical to that one obtained for ZnTCPPSpm4 alone in 5K buffer. These findings indicate that the formation of the ZnTCPPSpm4/TERRA complex occurs up to a molecular ratio of 8:1. As evidenced by the four break points, complexes with exact stoichiometry are formed at porphyrin:TERRA ratios that are multiples of two. This suggests that the TERRA G4 exists in a dimeric state, further corroborating one of the possible scenarios previously described.

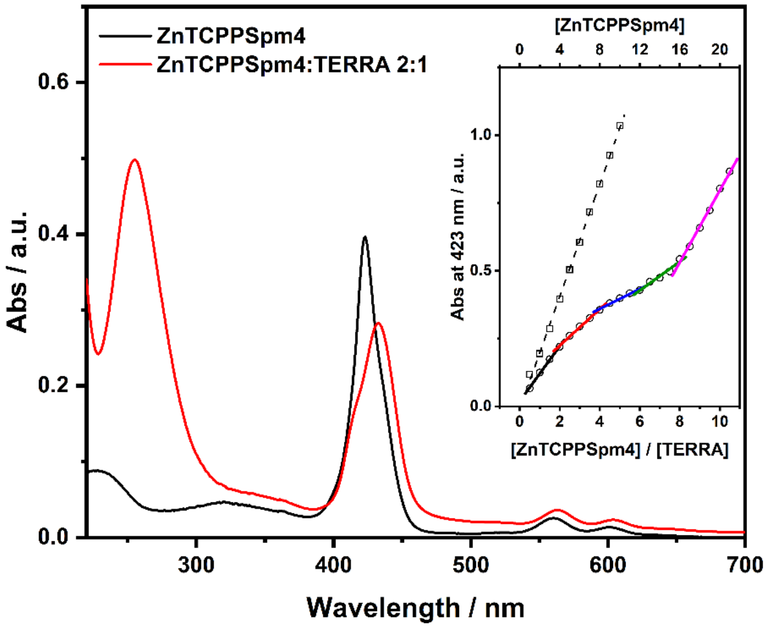


Fig. 56 *Absorption spectra of ZnTCPPSpm4 [4 μM] in 5k buffer (black curve) and in the presence of TERRA [2 μM] (red curve). Inset: plot of absorbance at 423 nm vs. the concentration of ZnTCPPSpm4 alone in 5K buffer (squares) and vs. the ratio [ZnTCPPSpm4]/[TERRA] (circles), lines with different colors indicate a different slope.*

The emission spectra of ZnTCPPSpm4 in 5K buffer have already been well described in this thesis during the characterization study of ZnTCPPSpm4 in different buffers, however it is worth to mention that the emission spectra is formed by the initial excited state, S1, and the charge-transfer state (CT) between the porphyrin core and its substituents at meso position. Polar solvents or restriction of peripheral substituents can cause fluorescence quenching because of the coupling of these two states (S1-CT).

The emission spectra show a pronounced quenching in the fluorescence of ZnTCPPSpm4 upon its interaction with TERRA at 2:1 ratio (**Fig. 57 (a)**). This significant quenching suggests a marked change in ZnTCPPSpm4's photophysical properties and points to a strong interaction with TERRA.

Specifically, the decrease in fluorescence emission continues until the 4:1 ratio (**Fig. S20**, **blue line**), indicating that as the concentration of ZnTCPPSpm4 increases relative to TERRA, the interaction between the two molecules intensifies, leading to fluorescence quenching. This observation aligns with the noticeable break point identified at the 4:1 ratio. However, at higher ratios (6:1 and 8:1), there is a slight increase in fluorescence, pointing to a potentially different type of interaction and/or aggregation of porphyrins onto the TERRA quadruplex.

An increase in Resonance Light Scattering (RLS) intensity was observed upon titrating TERRA with ZnTCPPSpm4 until the ratio ~ 8:1 (ZnTCPPSpm4:TERRA) (**red circles, Fig. 57 (b)**). This suggests an aggregation process involving strong electronic interactions among porphyrins on the TERRA structure. When the addition of porphyrin continued beyond this ratio, a slight decrease in RLS intensity was noted. This decrease indicates that no further complex formation occurs with a higher binding stoichiometry, a conclusion consistent with the UV-Vis spectroscopic data.

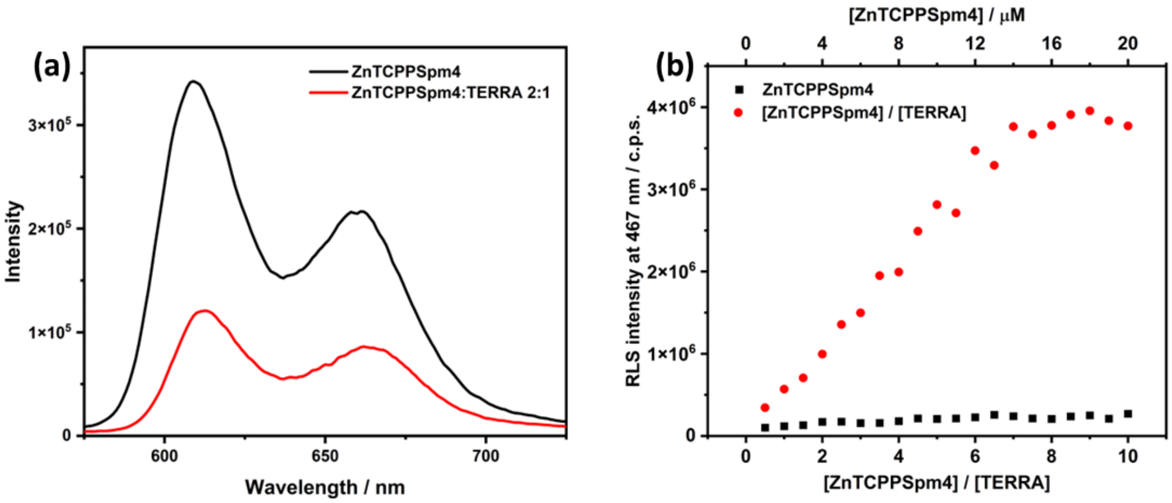


***Fig. 57 (a)** Emission spectra of ZnTCPPSpm4 [4 µM] in 5K buffer pH = 7.4 (black curve) and in the presence of TERRA [2 µM] (red curve). **(b)** Plot of the intensity of RLS signal at 467 nm vs. the ratio [ZnTCPPSpm4]/[TERRA] (red circles) and vs. the concentration of [ZnTCPPSpm4] alone in 5K buffer (black squares).*

Circular dichroism titration with increasing amount of ZnTCPPSpm4 were performed (**Fig. 58**). Upon increasing addition of ZnTCPPSpm4, the intense positive band around 265 nm and the trough around 245 nm decreased in intensity and a slight red-shift is observed, proposing a modification of the G-quadruplex structure. An induced bisignate CD signal (iCD) appeared in the

Soret region, further confirming the interaction between the porphyrin and the G4 TERRA. Bisignate iCD signal indicates that porphyrins are orderly distributed onto the whole G4 structure communicating each other along the assembly.

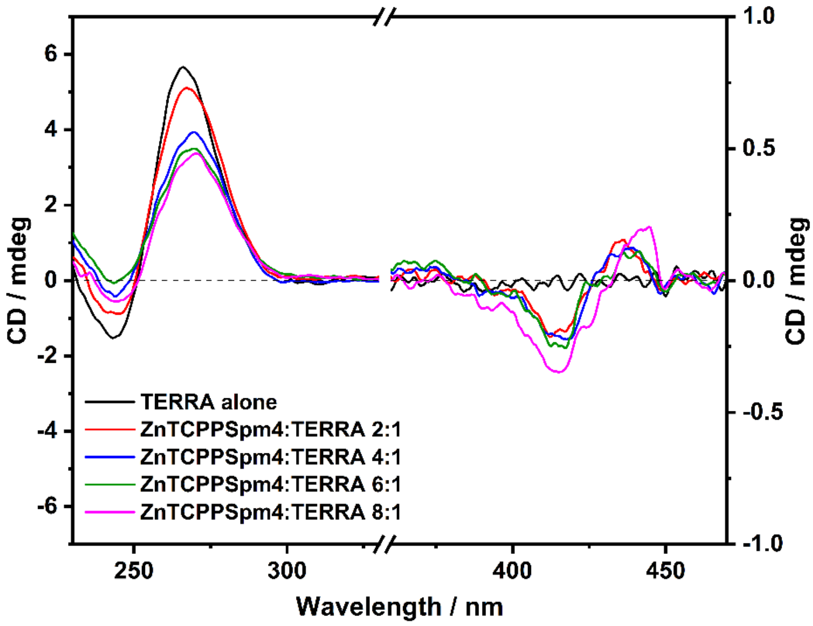


Fig. 58 *CD spectra of TERRA [2 μM] in 5K buffer pH 7.4 (black curve) and in the presence of increasing amounts of ZnTCPPSpm4 (4 μM red curve; 8 μM blue curve; 12 μM green curve; 16 μM magenta curve).*

Finally, in order to verify how the ZnTCPPSpm4 porphyrin can influence the stability of the G4 TERRA we performed CD-melting experiments at the ratio ZnTCPPSpm4:TERRA 2:1 (**Fig. 59**). Even in this case, two melting events occurred, still suggesting the presence of multiple units (e.g. dimers) or multiple conformations in equilibrium. However the presence of the ZnTCPPSpm4 drastically shifts the first melting event from ~ 40 °C (**Fig. 55**) to ~ 15 °C, without changing the second melting event (~ 60°C).

These results suggest that the ZnTCPPSpm4 porphyrin influences the stability of G4 TERRA in a differential manner. It leads to a strong destabilization of the first melting event, likely associated with the unfolding of specific G-quadruplex units or less stable regions. However, the second melting event, representing more stable structural elements in G4 TERRA, remains unaffected as they might be less affected by the presence of ZnTCPPSpm4 or not involved in strong interactions with the porphyrin.

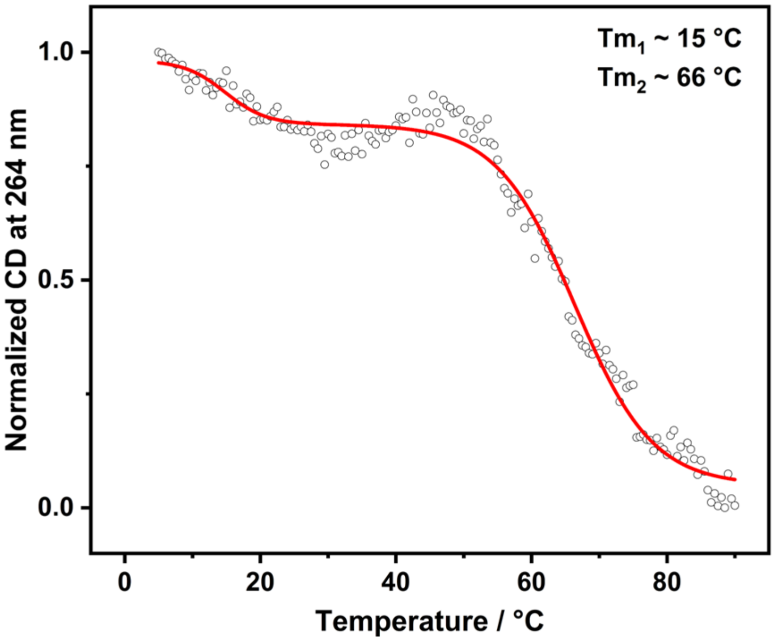


Fig. 59 *CD-melting curve normalized to [0,1] of 2 µM TERRA in presence of 4 µM ZnTCPPSpm4 in 5K buffer pH = 7.4. Circles represent experimental points, which were fitted by the BiDoseResp function (red curve).*

TERRA G-quadruplex can adopt different topologies

During the research period abroad at Swarthmore College, one of the objectives was to further investigate the interaction between the TERRA G4 and the ZnTCPPSpm4 porphyrin. However, the CD spectra of the TERRA sequence in 5K buffer revealed a shoulder at approximately 290 nm (**Fig. 60**, **black line**), which was not present in the CD spectra recorded in Catania (**Fig. 60**, **blue line**). Generally, a positive band in this region can be attributed to the presence of an RNA antiparallel topology. In fact, when the same sequence is in 10 mM Tris buffer with 100 mM NaCl, it exhibits the characteristic spectrum of the antiparallel topology (**Fig. 60**, **red line**) [218], suggesting that the specific sequence depending on the buffer and the nature of the monovalent ion can adopt an antiparallel G4.

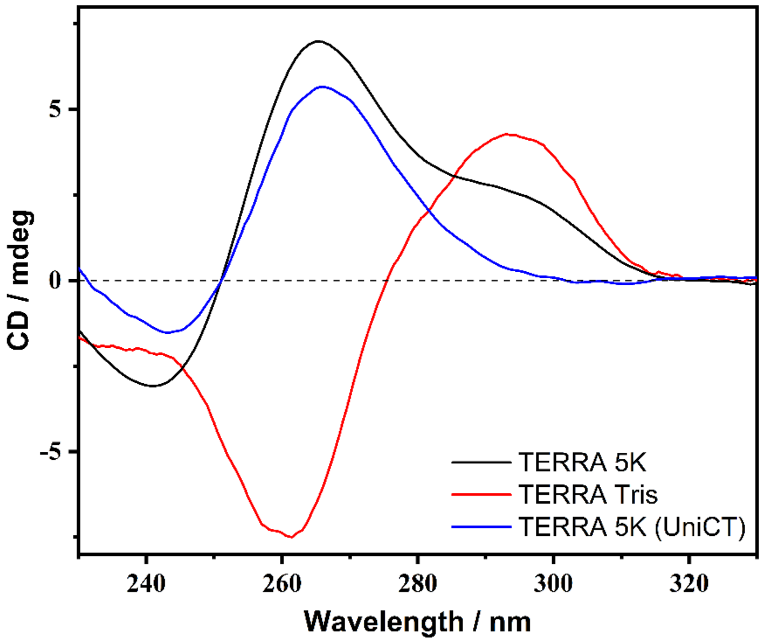


Fig. 60 *CD spectra of TERRA [2 µM] in 10 mM 5K buffer (black line) and in 10 mM Tris buffer implemented with 100 mM of NaCl (red line) both spectra recorded at Swarthmore College. TERRA [2 µM] in 10 mM 5K buffer recorded at University of Catania (blue line).*

Based on this, it was speculated that in the 5K buffer at Swarthmore College, the TERRA sequence adopts both parallel and antiparallel topologies simultaneously in 5K buffer. Upon returning to Catania, the decision was made to investigate the reason for our sequence only adopting the parallel topology.

In order to verify the integrity of the sequence, denaturing PAGE experiments were conducted (**Fig. 61**, **left**). The TERRA sequence shows an intense band around the dT24 marker, which correspond to the intact sequence formed by 22 nucleotides, additionally it is possible to observe other two lower bands, between dT24-dT15 and below dT15 respectively, indicating the presence of two shorter sequences, demonstrating that the original TERRA sequence was in part digested by nucleases probably present in the laboratory environment. We speculate that the presence of these shorter sequences can impede the formation of the antiparallel topology since is the only difference found between the TERRA sequence in Catania and that one in Swarthmore (**Fig. 61**, **right**), which showed only one band around dT24 confirming the integrity of the sequence.

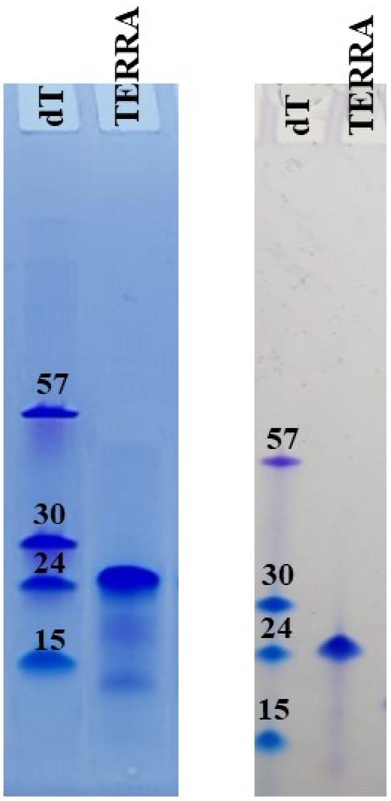


Fig. 61 *Fifteen percent denaturing gel prepared in TBE 1X buffer and 7M urea. Size markers correspond to dTn sequences. The TERRA sample was prepared at the concentration of 40 µM in* ddH_2O *with 30% of formamide. PAGE performed in Catania (left image), and in Swarthmore (right image) using the same experimental conditions.*

3.2.3. Final remarks

In conclusion, the spectroscopic results collectively demonstrate that ZnTCPPSpm4 porphyrin interacts with parallel G4 TERRA, forming different ZnTCPPSpm4:TERRA complexes until the ratio 8:1. The interaction is dependent on their molecular ratio, with multiple complexes formed at specific stoichiometric ratios. The enhanced RLS intensity evidences an aggregation of porphyrins onto the G4 TERRA as also suggested by the bisignate iCD signal.

Modifications in the CD spectra further indicate structural changes in G4 TERRA upon binding to ZnTCPPSpm4. Finally, the CD-melting experiment suggests that the ZnTCPPSpm4 strongly destabilizes specific G-quadruplex units or less stable regions. The findings suggest a complex and intriguing interplay between the porphyrin and the G4 TERRA, potentially influencing the stability and behavior of the G-quadruplex structure.

Furthermore, the TERRA sequence can adopt different G4 topologies (parallel and antiparallel). However, the presence of shorter sequences seems to prevent the formation of the antiparallel topology. Next step will be identify and quantify the shorter sequences formed in the TERRA sample through an intra-laboratory collaboration using high-resolution accurate mass spectrometry (HRAM MS) coupled with data-dependent tandem MS (ddMS2). This approach enables accurate mass measurement, confident identification, and mapping of oligonucleotides.

4. i-motif DNA

i-motif DNA is a non-canonical four-stranded DNA structure that forms in sequences rich in cytosine. The i-motif DNA structure was first characterized by Gehring et al. [254]. It was formed by a hexamer sequence of d(TCCCCC) intercalated in an antiparallel orientation to create a quadruple-helical tetramolecular structure under acidic conditions. This structure consists of two parallel-stranded duplexes held together by hemi-protonated cytosine-cytosine$^+$ (C:C$^+$) base pairs (**Fig. 62 (a)**) [255]. Since the initial report, several i-motif structures have been identified using crystallographic and NMR methods [254,255].

Like G4 structures, i-motifs can fold in an intramolecular manner (monomeric) as a result of the spatial arrangement of four different C-tracts within the same strand, or they can fold in an intermolecular manner using two (dimers) or four independent DNA strands (tetramers). The distance between consecutive base pairs in the i-motif is 3.1 Å, and the right-handed helical twist angle is ∼12–20◦, which is significantly smaller than that of B-DNA [256]. The intercalation of base pairs from two parallel duplexes creates two major grooves and two minor grooves (**Fig. 62 (b)**). The two minor grooves are very narrow, resulting in destabilizing interactions due to close phosphate-phosphate distances. Sugar-sugar contacts partially compensate for these interactions [257].

The i-motif structure can be classified in two different intercalation topologies known as 3'E and 5'E (**Fig. 62 (c)**). The 3'E topology is more stable than the 5'E topology due to the extended sugar-sugar contacts along the narrow grooves [258].

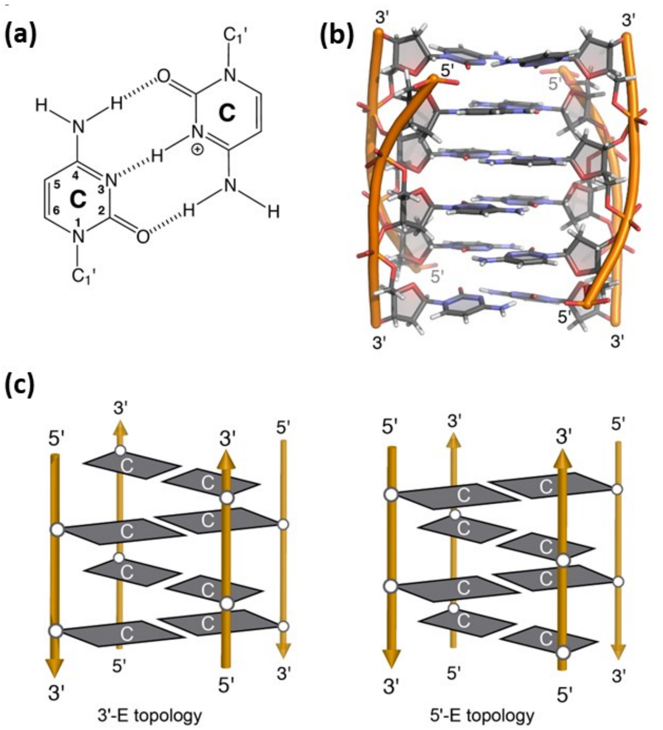


Fig. 62 ***(a)*** *C:C+ base pair.* ***(b)*** *3D structure of a tetramolecular i-motif.* ***(c)*** *Representation of the 3'E (left) and 5'E (right) intercalation topologies of an intermolecular i-motif structure. Adapted from ref.* [259].

It was originally thought that i-motif structures could only form at acidic pH values, but recent studies have shown that they can also form at neutral pH depending on the sequence and environmental conditions [260,261]. i-motif structures have been observed at neutral pH and low temperatures under molecular crowding conditions [262], under negative superhelicity [263], in the presence of certain cations, and inside silica nanochannels [264].

Similar to other structures of nucleic acids, factors like the nature of the sequence, temperature, and ionic strength influence the stability of i-motifs. Contrary to B-DNA or G-quadruplexes where stacking interactions between sequential nucleobases significantly contribute to their stability, the intercalative geometry between subsequent base pairs in i-motif structures leads to minimal overlap among the aromatic pyrimidine bases. Although C:C$^+$ base pairs display favourable stacking of exocyclic carbonyls and amino groups in an antiparallel manner, theoretical computations revealed that these and other advantageous stacking interactions between successive C:C$^+$ base pairs hardly counterbalance the electrostatic repulsion between their charged imino groups [265].

The crucial interactions for the stability of i-motifs are the hemi-protonated C:C+ base pairs (**Fig. 62 (a)**). These base pairs, which are characterized by three hydrogen bonds, offer a high level of stability. The C:C+ base pairs are planar, with a distance between the N3–N3 approximately between 2.6 and 2.8 Angstroms [266].

The i-motif structural formation is destabilized by the close distances between phosphorus atoms (5.9 A) along the minor groove. This instability is balanced by positive interactions between deoxyribose sugar units along the minor groove. Additional stabilization interactions along the minor groove arise from C-H1'···O4' interactions within each pair of antiparallel strands. Furthermore, studies indicate intra-nucleoside hydrogen bonds between O4' and H6, and O4' simultaneously bonding with H1' and H4' [256]. For DNA i-motif structures, the orientation of the glycosidic angles is anti, while the deoxyribose sugars predominantly exhibit a C3'-endo puckering.

Generally, under identical experimental conditions, an i-motif structure with a greater number of C:C+ base pairs tends to be more stable [267]. Recently, how the length of the C-tract impacts the formation of intramolecular i-motif structures under physiological conditions was investigated [260,261]. Their findings generally indicate an increase in transition pH (pH_T) as the number of cytosines per tract rises. This trend was similarly observed in terms of thermal stability (Tm). Moreover, more recent investigations suggest that sequences with shorter C tracts also tend to form stable i-motifs at neutral pH [268,269].

Nevertheless, it is crucial to consider that interactions between loops significantly contribute to the stability of these structures [270,271]. The number and type of residues capping the i-motif core, as well as the length and nature of the connecting loops, are key determinants of i-motif stability [272]. Generally, extremely short loops, like those comprising one nucleotide, tend to promote the formation of mono and bimolecular i-motifs. Conversely, longer loops usually facilitate the formation of intramolecular i-motif structures [273]. Because these extended loops could potentially permit additional stabilizing interactions, it has been hypothesized that they possess greater stability.

Thymidines, capable of forming T:T base pairs analogous to C:C+ base pairs, commonly cap i-motifs and extend the i-motif core [274,275]. T:T base pairs are not only excellent capping pairs, but they can even be accommodated in the middle of the C:C+ base pair stack [276]. Hoogsteen, reverse Watson-Crick A:T, A:A and G:G base pairs have been observed in several i-motif structures [275,277,278], implying that it is not only the loop size, but the exact sequence and resulting interactions of the loop bases that contribute to stabilization.

Unlike G4 structures, where the type of cation significantly influences stability and folding topology, the stability of i-motifs is not impacted by the nature of the cation, but rather by the ionic strength of the solution [279]. Increasing the concentration of NaCl from 0 to 100 mM at a pH near the pKa of cytosine destabilized the i-motif structure. Interestingly, higher NaCl concentrations (to 300 mM) did not induce additional destabilization [273]. Molecular crowding agents, such as high molecular weight polyethylene glycols (PEGs), tend to favour the stability of both i-motif and G4 structures over duplexes and single-stranded DNA [256].

Biological relevance

Bioinformatic analysis showed that C-rich sequences were identified in promoter regions, introns, and both 5'- and 3'-UTRs, while fewer were found in coding and intergenic regions, evidencing that sequences with a higher propensity for i-motif formation are not randomly located [261].

It has been recently demonstrated that i-motif DNA structures are formed in vivo in the nuclei of human cells [280], providing evidence that these structures are formed in regulatory regions of the human genome, including promoters and telomeric regions. These findings support the hypothesis that i-Motif structures may play key roles in the regulation of gene expression.

Poly-C-binding proteins (PCBP) interact with C-rich DNA sequences and are crucial in regulating gene expression [281]. However, it is often unclear whether these proteins bind to a specific i-motif structure or the C-rich strand resulting from i-motif unfolding. Early research by

Marsich et al. [282] found a protein in human HeLa cells that had a specific affinity for the human telomeric sequence, d(CCCTAA)n, containing at least four cytosine tracts.

More recently, the BmILF protein from the Bombyx mori insect was identified as an i-motif binding protein [283].

One of the most thoroughly researched i-motif binding proteins is the BCL2 activating transcription factor hnRNP LL. It was found to bind specifically to i-motif structures and not to BCL2 promoter forming a duplex or to mutated single strand DNA unable to fold into an i-motif. The protein unfolds the i-motif structures into a single-stranded sequence, the more thermodynamically favoured conformation, and activates BCL2 gene transcription [284].
This shows i-motif structures as protein recognition sites that can regulate gene expression.

i-motif structures are proposed to play a role in providing long-range interactions between laterally associated centromeric nucleosomes. The presence of these i-motif forming sequences at the entrance and exit of the nucleosome could facilitate their involvement in forming dimeric i-motifs. This suggests that non-B DNA structures might not only have a role in determining centromere location, but also contribute to the specific architectural features of the centromere [285].

There is a substantial amount of evidence suggesting that i-motif structures are involved in the regulation of transcription. During the G1/S phase of the cell cycle when transcription activity is higher, the number of i-motif foci is found to be greater, suggesting their role in transcription regulation [280].

The BCL2 oncogene, which promotes cell survival and proliferation, can fold into i-motif structures that influence its transcription. Compounds like IMC-48 can stabilize the i-motif structure and lead to significant upregulation of BCL2. Similarly, the hnRNP LL protein influences BCL2 transcription by recognizing and unfolding the BCL2 i-motif. These findings illustrate the effects of molecules and transcription factors on i-motif population and gene expression [286].

Several reports suggest that i-motifs and G4s serve as molecular switches for the regulation of various genes, including MYC, which is regulated by SP1 binding and hnRNP K–i-motif complex interactions. These intricate molecular interactions highlight the complexity of transcription regulation and the potential role of i-motif structures in these processes [287].

Sugimoto and colleagues [288] have recently studied the effect of non-canonical DNA structures, including i-motif-forming sequences, on DNA replication by the Klenow fragment (KF) of DNA polymerase. They found that i-motif-forming sequences in the DNA template strand stall DNA polymerase, thereby inhibiting DNA replication or repair. i-motif, unwinding by KF, demands higher activation energy than other non-canonical structures suggesting that i-motifs may modulate DNA replication in living organisms and may have a greater influence compared to other secondary structures.

Despite significant progress in i-motif structural biology, many aspects still require further study. Current data suggest that i-motifs form transiently within cells, but more in vivo studies are needed to confirm i-motif formation during various stages of the cell cycle. Additional research on how proteins and small ligands recognize i-motifs, both in vitro and in vivo, is essential to understand their roles in biological processes. These studies are complicated by the difficulty of distinguishing between recognition of C-rich sequences and actual i-motif structures. Compared to G4s, only a few i-motif structures have been identified using NMR or crystallographic methods. At present, it is not possible to predict the stability of an i-motif based on its sequence. Hence, more structural information is needed to understand the effects of capping interactions and loops connecting the C-tracts on i-motif stability.

4.1. i-Motif structures formed in the Human c-MYC promoter: State-Of-The-Art

c-MYC is a vital regulator of the cell cycle and is responsible for controlling cell growth, differentiation, apoptosis, and neoplastic transformation [289]. Overexpression of the c-MYC gene is a common occurrence in various forms of human cancer and its alteration has been linked to several diseases [290].

The c-MYC oncogene is a frequent target of genetic alterations in various types of human cancer, with protein overexpression observed in more than 40% of tumors [291]. As a "master regulator," c-MYC drives the expression of multiple pathways associated with cancer hallmarks, including cell cycle signaling, glycolysis, and other metabolic pathways [292–294]. Although c-MYC expression is essential for normal cell growth, tissue development, and apoptosis, it needs to be downregulated after proliferation and growth to enable appropriate differentiation and senescence.

The regulation of c-MYC expression involves multiple promoters, and a conserved 27-base-pair sequence known as NHE III_1 plays a crucial role in c-MYC transcription. NHE III_1 can form two DNA secondary structures, namely G-quadruplex and i-motif, with the former being essential for c-MYC transcriptional silencing [295]. Compounds that stabilize the G-quadruplex can repress c-MYC gene expression. The major c-MYC G-quadruplex structure, formed by consecutive guanine runs, has been characterized by Yang et al. [272]. Additionally, the C-rich strand of NHE III_1 can form i-motif structures with potential interactions with small molecule compounds that regulate gene transcription.

The i-motif formed in the c-Myc NHE III_1 appears to be highly dynamic in its formation. The dynamic equilibrium between multiple structures caused by sequence and structure redundancy appear to provide entropy and stability to the overall i-motif structure, as demonstrated by Dai et al. [272].

Vorlickova et al. investigated a C-rich strand of NHE III_1 formed by 33 nucleotides, which corresponds to bases 2180– 2212 in the sequence of the human c-MYC locus demonstrating that the different sequences investigated adopt an i-motif conformation in vitro, and discuss implications of possible biological roles for i-motif structures in vivo [296].

Very recently Anunay et al. studied a 22-mer c-MYC-promoter-based C-rich sequence (Py22) in the presence of PEG using Förster resonance energy transfer and fluorescence lifetime measurements at the single-molecule level. Molecular crowding agents like PEG facilitate i-Motif formation even at neutral pH, simulating the intracellular environment [297].

Li et al. synthesized and evaluated new bisacridine derivatives for their interactions with c-MYC promoter G-quadruplex and i-motif. Among them, a bisacridine derivative demonstrated the ability to bind to and stabilize both G-quadruplex and i-motif structures, leading to the downregulation of c-MYC gene transcription. Additionally, it showed the potential to inhibit cancer cell proliferation, induce apoptosis, and arrest the cell cycle in SiHa cells. These findings suggest that this derivative, as a dual G-quadruplex/i-motif binder, could effectively target both oncogene replication and transcription [298].

4.2. Biophysical and X-ray study of C-rich DNA variants from c-MYC: aim of the work

The purpose of this chapter is to summarise the research activities which were conducted at Swarthmore College (PA, United States), under the supervision of Prof. Liliya Yatsunyk. Noteworthy, all the following data are currently under further investigations.

We have undertaken an investigation of a cytosine-rich tract formed by 35 nucleotides that is present in the c-MYC gene, which is situated on human chromosome 8 at the 8q24 locus (**Fig. 63**).

This tract consists of seven stretches of 2-4 cytosines Our aim is to understand how this sequence folds into i-motif (iM) and what are the atomic details of this iM structure. Such knowledge will provide the field with the iM coordinates necessary for the screening and design of iM binders as potential anticancer therapeutics.

To avoid potential heterogeneity of the DNA sample we divide the C-rich tract into two shorter variants named PyA and PyB, which contain four C-rich stretches leading to a homogeneous monomolecular iM. These variants are depicted in red and blue, respectively (**Fig. 63**), while the tract in common is shown in purple. To investigate the effect of the overhangs on the fold of the iM and to increase our chances for successful DNA crystallization, we have designed six new variants from each sequence by adding one and two nucleotides to the terminal ends. Our initial goals are to determine if these variants form i-motif structures, as well as to identify their thermodynamic stability, their transition pH and 3D architecture.

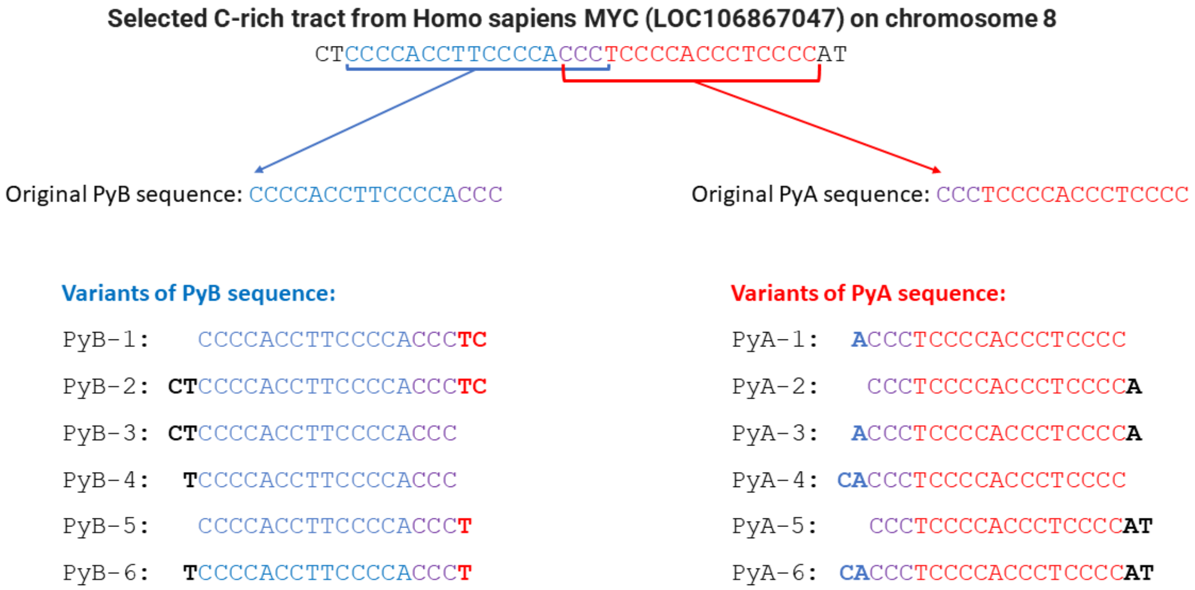


Fig. 63 *Variants design. The common tract in PyA variants is highlighted in red, while the common tract of PyB variants is shown in blue. The three cytosines common to both variant groups are indicated in purple. Any additional nucleotides at the overhang are in bold.*

In the first part of my research work I developed a protocol to obtain the transition pH (pH_T) values for i-motif-forming sequences by using circular dichroism (CD) equipped with an automatic titrator. Then I used this protocol to determine the pH_T values for selected PyA and PyB sequences.

The folding behaviour of i-motifs is pH-dependent. Additionally, it is affected by intrinsic factors such as the length of the oligodeoxynucleotide, the length and number of the poly-dC tracks, and the length and nucleotide composition of non-poly-dC tracks. Extrinsic factors include temperature, salt concentration and identity, and molecular crowding [299].

Transition pH (pH_T) is defined as the pH at a given temperature at which 50% of the oligodeoxynucleotide molecules are in the intramolecularly folded state. Typically, ultraviolet (UV), circular dichroism (CD), or NMR spectroscopy are used to measure the pH_T for i-motifs [300,301].

The transition pH represents an important characteristic of an i-motif sequence because it can provide insight into the conditions under which the i-motif structure is most likely to form in vivo. This information can be useful for understanding how the i-motif structure may contribute to biological processes such as gene regulation and DNA damage repair. Additionally, knowledge of the transition pH can also aid in the design of experiments to study the i-motif structure, as well as in the development of i-motif-based biosensors and therapeutics [259].

Secondly, to gather more information about the topologies and 3D architectures adopted by all Py variants, I conducted investigations using CD, CD-melting, and native PAGE techniques. Additionally, I attempted to crystallize all variants using the hanging-drop vapor diffusion method. Excitingly, I was able to crystallize one variant, PyB-2. Diffraction data were collected at the Advanced Photon Source synchrotron facility. Regrettably, the 3D structure of PyB-2 remains unsolved due to the phase problem and is currently under investigation.

4.2.1. Material and methods

DNA, ligand, and buffers

Lyophilized oligonucleotides were purchased from Integrated DNA Technologies (IDT; Coralville, IA) with standard desalting purification. DNA was hydrated in doubly-distilled water to 1-2 mM and stored at 4 °C. Extinction coefficients for all sequences were obtained using IDT's OligoAnalyzer 3.1 and DNA concentration was determined from UV-vis spectra collected at 90 °C. The full list of DNA sequences used in this work, along with their extinction coefficients and molecular weights, can be found in **Table *3*.**

In my work I used two buffers. The first one was formed by 10 mM MES, pH 5.5; this buffer was used for UV and CD scans, CD-melting experiments, and crystallization trials. pH titrations were performed in 20 mM Britton–Robinson (BR) [302] buffer with 140 mM KCl.

Table 3. DNA sequences studied in this work and their thermodynamic parameters in 10 mM MES buffer pH = 5.5. All sequences are based on PyA and PyB, with additional nucleotides indicated in bold.

Name	*Sequence 5' →* **3'**	ε, $mM^{-1}cm^{-1}$	*MW, g/mol*	*Tm, °C*	*ΔH Fit, kJ/mol*	ΔT_m, °C
PyA-1	**A**CCCTCCCCACCCTCCCC	143.4	5221.4	58.1	170	3.1
PyA-2	CCCTCCCCACCCTCCCC**A**	143.4	5221.4	55.1	158	3.9
PyA-3	**A**CCCTCCCCACCCTCCCC**A**	157.2	5334.6	54	155	0.9
PyA-4	**CA**CCCTCCCCACCCTCCCC	149.2	5510.6	59	161	1.9
PyA-5	CCCTCCCCACCCTCCCC**AT**	150.8	5525.6	54	157	2.9
PyA-6	**CA**CCCTCCCCACCCTCCCC**AT**	170.4	6128.0	50	154	2.9
PyB-1	CCCCACCTTCCCCACCC**TC**	150.1	5525.6	50.5	133	0.5
PyB-2	**CT**CCCCACCTTCCCCACCC**TC**	165.4	6119.0	46.1	141	0.7
PyB-3	**CT**CCCCACCTTCCCCACCC	150.1	5525.6	55	158	0.6
PyB-4	**T**CCCCACCTTCCCCACCC	143.6	5236.4	55	154	0.6
PyB-5	CCCCACCTTCCCCACCC**T**	142.6	5236.4	51	144	0.4
PyB-6	**T**CCCCACCTTCCCCACCC**T**	151.4	5540.6	50	148	0.2

UV-vis spectroscopy

We used an Agilent Cary 3500 UV-vis spectrophotometer equipped with a Peltier block probe temperature controller (± 0.5 °C error) to scan DNA. Data were collected using 1 nm intervals, 0.02 s averaging time, 2 nm spectral bandwidth, automatic baseline correction, and 220–330 nm range for DNA.

Circular dichroism (CD) scans

CD experiments were performed on a Jasco J-1500 circular dichroism spectrophotometer equipped with a Peltier thermocontroller (± 0.5 °C error) in 1 cm quartz cuvettes with Teflon caps at 20 °C. CD scans were collected from 220-330 nm with 1 s averaging time, 200 nm/min speed, 2 nm bandwidth, and 1.00 nm step. Five scans were collected, averaged and processed as described in this work [303].

CD melting experiments

CD melting experiments were performed from 4 to 95 °C with a 1 °C/min temperature rate, 8 s D.I.T., 30 sec equilibration, and 2 nm spectral bandwidth. The experiments were conducted at two different wavelengths (288 and 330 nm) to measure the unfolding of the i-motif and instrument stability, respectively. Two methods were employed to determine the melting temperatures, Tm. The first method involved analyzing the first derivative of the 288 nm CD signal and identifying the temperature at the peak or through visual inspection (with an associated error of ± 0.5 °C). The second method assumed a two-state model for i-motif folding with constant ΔH, which was applied to fully reversible melting transitions where melting and cooling curves were nearly identical [304]. Hysteresis was determined as the difference between Tm determined from the melting and cooling curves. Since hysteresis did not surpass 3°C for all sequences, the systems are considered to be reversible. Thus, the thermodynamic data reported were obtained using the two-state model, which is the second method described above.

Automatic pH titration

CD automatic pH titrations were performed using a Jasco J-1500 equipped with the automated titration unit ATS-530, which is employed with dual syringes of a maximum volume of 1.0 ml each. pH was constantly monitored by the pH-meter FiveEasy F20, which was directly inserted into a quartz cuvette equipped with a stir bar.

The Automatic Titration Scan Measurement was used to perform the pH titration and to automatically correct the sample concentration. The total volume of HCl and LiOH added was less than 10% of the total volume of the DNA sample. After each addition of titrant, the sample was equilibrated for 90 sec after which CD spectra were automatically recorded (with the parameters described above, but with only one scan), keeping track of the pH value. The whole titration was performed under stirring at 120 rpm.

CD spectra were processed using Singular Value Decomposition (SVD) [305,306] to generate a titration plot where the pH has been plotted against the normalized **v1** vector from the SVD matrix **V.**

Native polyacrylamide gel electrophoresis (PAGE)

PAGE samples contained 3 µg DNA in 10 µL of 10 mM MES 5.5 buffer and 7% w/v sucrose was added before loading. To prepare crystal samples, 5–10 crystals were harvested, washed in the crystallization condition, and dissolved in 10–15 µL of 10 mM MES 5.5 buffer.

Fifteen percent native polyacrylamide gels were prepared with 5 mM KCl and 50 mM MES 6.0, and running buffer was also made with 5 mM KCl and 50 mM MES 6.0. The gels were pre-migrated at 150 V for at least 30 minutes, loaded with 10 µL of each sample, and allowed to run for 150 minutes at 150 V and room temperature. An oligothymidylate ladder consisting of dT15, dT24, dT30, and dT57 was used as a length marker, and a tracking dye was used to monitor the gel progress. DNA bands were visualized with Stains-All and the resulting gel was captured using a smartphone camera.

X-ray crystallography

General

The hanging-drop vapor diffusion method was utilized to achieve crystallization at 12 °C. DNA samples at a concentration of 1.5 mM were annealed in 10 mM MES 5.5. To screen for crystallization conditions, a TTP Labtech Mosquito Liquid Handling robot system equipped with a humidity chamber was used with commercial screens, including Natrix (Hampton Research) and Helix (Molecular Dimensions) [307]. 96 well trays were set up with equal volumes of sample and crystallization condition (each 0.1 μL) and 100 μL of well condition. Optimization was conducted manually in 24-well trays by combining 1 μL of DNA sample with 1 μL of crystallization condition and adding 400 μL of crystallization condition into wells. Diffraction data were collected at the Advanced Photon Source synchrotron facility at 24 ID-E beamline at both the native ($\lambda = 0.98$ Å, 12622 eV) and the cobalt wavelength ($\lambda = 1.61$ Å, 7725 eV). Raw diffraction data were processed using RAPD software provided by the beamline. The trials to solve structures were performed through molecular replacement (MR) using PHENIX [308]. The data collected at the cobalt wavelength was also examined for the strength of anomalous signal. The structures were build COOT [309], followed by refinement in PHENIX.

PyB-2

For the original screening, PyB-2 sample was prepared at a DNA concentration of 1.5 mM in a 10 mM MES buffer pH = 5.5. The original crystals grew in one month in condition 1–4 from the HELIX screen which consisted of 0.2 M KCl, 25% PEG400, 5 mM $[Co(NH_3)_6]Cl_3$, 0.05M MES pH = 6.5.

This condition was optimized in 24-well trays screening DNA in the range 1.5-1.7 mM, $[Co(NH_3)_6]Cl_3$ in the range 2-5 mM, PEG400 20-30% and KCl to 0.1-0.3 M. The best crystals were big and show a well-ordered 3D structure, they were obtained after one week in the original home-made 1-4 HELIX condition; no crystals were obtained in 2 mM $[Co(NH_3)_6]Cl_3$, and smaller crystals were obtained in all other optimization instances. Crystals were cryoprotected in the base condition supplemented with either 10% of PEG400 (25% PEG400 total), or 10% of glycerol, or 10% of ethylene glycol before being flash frozen in liquid nitrogen. The latter two conditions produced good quality glass and better quality crystals.

Among 27 different crystals, 10 gave a maximum resolution below 1.90 Å in a space group C121 and I222. The crystal that gave the best maximum resolution of 1.53 Å (named GT1_C1) was one of the first crystals grown in the handmade condition of 1-4 HELIX at concentration of 1.7 mM and included 10% glycerol as cryo condition. Crystals named GT3_C5 and GT3_C6 gave the second and third best results with a maximum resolution of 1.54 and 1.86 Å, respectively. Both were grown in the 1-4 HELIX handmade condition, but with 0.1 M KCl and 1.5 mM DNA. GT3_C5 was protected with the base condition supplemented with 10% ethylene glycol, whereas for GT3_C6 was protected with the base condition and 10% glycerol.

4.2.2. Developing a protocol to obtain pH_T values for i-motif DNA sequences

Aim of the work presented in this Chapter is to develop a protocol to obtain pH_T values for i-motif-forming sequences by using circular dichroism (CD) technique. This protocol will then be used to determine pH_T values for representative PyA and PyB sequences.

We started by literature review to determine what buffer should be used and which one is used more often by the scientific community. Buffers are used to maintain a stable pH environment. Using the correct buffer is essential when studying i-motif sequences because pH greatly affect the stability and formation of the i-motif structure. The pH of the buffer must be carefully controlled to match the transition pH of the i-motif sequence in order to optimize the formation of the i-motif.

We found that there is not a unique buffer. Indeed 25 mM and 10 mM sodium phosphate (NaPi) buffer was used by González et al. and Pagano et al. respectively [310,311]. 0.5× BrittonRobinson (BR) buffer (20 mM H_3BO_3, 20 mM H_3PO_4, 20 mM CH_3COOH, 37 mM KOH) was used by Vorlíčková et al. [312]. 20 mM potassium phosphate (KPi) with 12 mM NaCl, and 140 mM KCl was used by Burrows et al. [299] and 20 mM sodium cacodylate or 20 mM sodium phosphate buffers were used by Trantírek et al. [313]. Smith et al. used 10 mM sodium cacodylate with 100 mM KCl [314].

Nevertheless, all the research groups performed their experiments in a pH range between 4.5 and 8 at pH increments of ~ 0.25 pH by adding aliquots of HCl or NaOH. DNA solution had concentration of iMs between 2 and 4 μM.

pH_T determination via manual and automatic titrations

At the beginning we performed the pH titrations in 20 mM KPi buffer pH = 8.08, with a sample concentration of ~ 2 μM. The CD spectra were recorded at pH increments of ~ 0.2 units, by adding small aliquots of HCl 1M and when the pH reached a value around 4 the titration was reversed to a pH ~ 8 by adding small aliquots of 1 M LiOH. We used LiOH instead of NaOH because we wanted to maintain the physiological concentrations of K^+ or Na^+ as constant and as reported by Burrows et al., 50 mM LiCl does not affect the iM folding behaviour [299]. After each addition of acid/base we waited at least 90 seconds before recording CD spectra to allow equilibration of the sample.

For each sequence investigated, we prepared two identical samples and performed manual and automatic titrations on the same day in order to evaluate possible differences between the two method. The data are reported as plots of x_{IM} vs pH, where x_{IM} represents the fraction of folded iM. x_{IM} was obtained by normalizing the **v1** vector at the given pH. All experiments were conducted at constant temperature of 25 °C.

We first performed pH titration on the PyA-2 (**Fig. 64**). We started at pH = 8 where DNA is unfolded. Gradual addition of HCl resulted in the increase of the 288 nm signal signifying folding of the iM DNA. The original titrations went to near pH of 3 where likely most of cytosines became protonated and electrostatic repulsion lead to unfolding of the DNA and decrease in the CD signal at 288 nm (**Fig. 64 (c) and (d)**). In addition, the dramatic change in the 288 nm signal at low pH could result from reaching outside of the buffering capacity of KPi (4.8-8.0). Therefore, for the subsequent titrations, we limited the lower end of pH values to ~ 4.0, consistent with literature.

The titration was repeated in a similar manner for PyA-3 (**Fig. S21**).

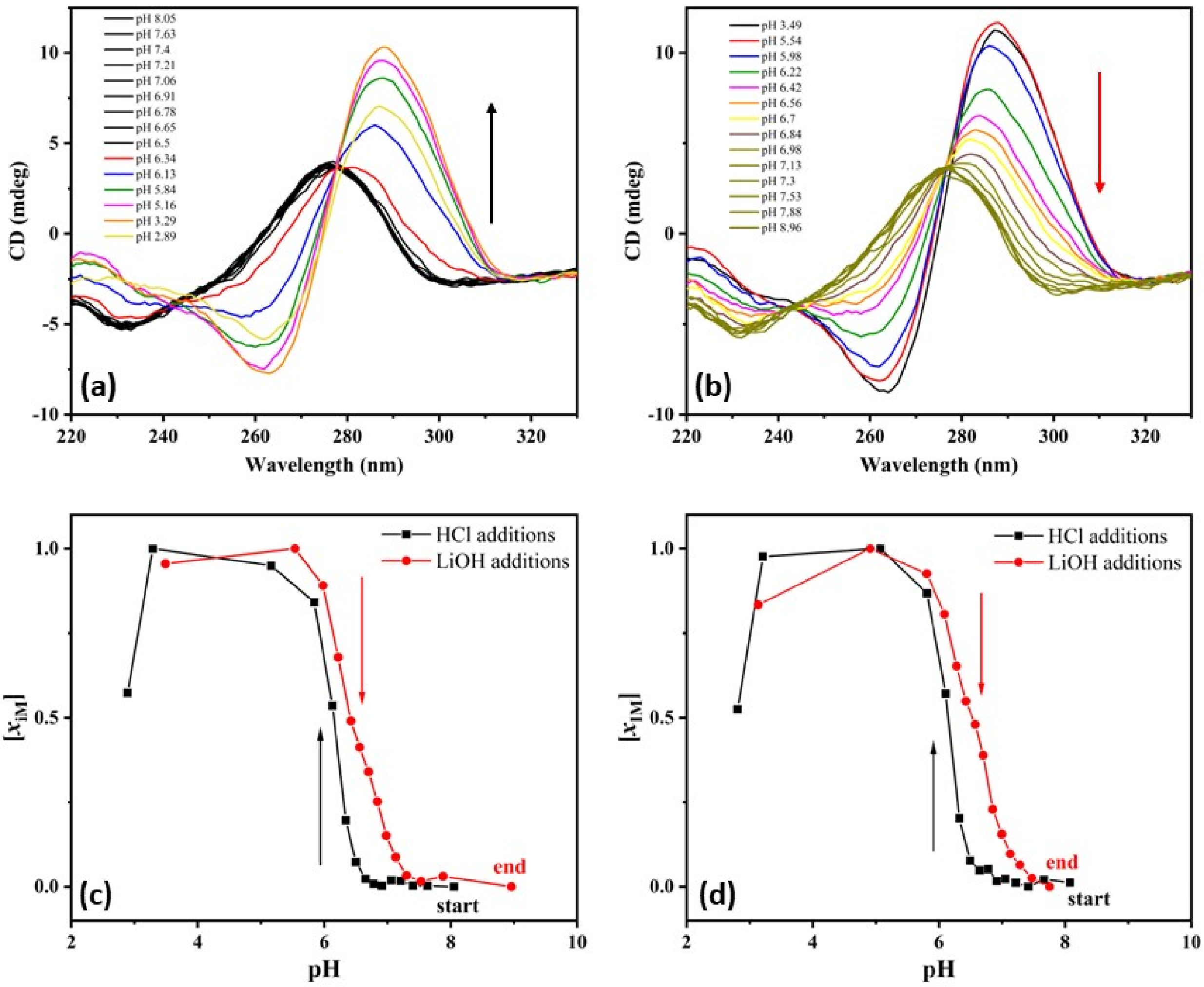


***Fig. 64** The raw CD spectra (top) of PyA-2 in 20 mM KPi buffer at various pH values using manual titration. **(a)** Acid titration with 14 additions of 5 µL of 1.0 M HCl and **(b)** reverse base titration with 14 additions of 5 µL of 1.0 M LiOH. Titration curves for PyA-2 (bottom) in KPi buffer. **(c)** Manual titration and **(d)** automatic titration for PyA-2 each one was performed with 14 additions of 5 µL of 1.0 M HCl (black squares), followed by 14 additions of 5 µL of 1.0 M LiOH (red circles). [PyA-2] ~ 2 µM in 20 mM KPi pH = 8.08.*

The raw data were corrected by the dilution factor and processed using Singular Value Decomposition (SVD) [305,306] to generate a titration plot where the pH has been plotted against the normalized **v1** vector from the SVD matrix **V.**

The titration plots in (**Fig. 64 (c)** and **(d)**) show hysteresis; in addition, the reverse titrations give rise to a multiphasic unfolding behaviour, suggesting a non-reversible process. As reported by Burrows et al. [299] biphasic unfolding for samples subject to a rapid drop in pH is consistent with the formation of a small subpopulation of iM dimers which may or may not be the case here.

In order to obtain the $\mathrm{pH_T}$ value, we fitted the data in the pH range between ~ 4 and ~ 8 in two different ways, using either Boltzmann sigmoid function or the following $\mathrm{pH_T}$ equation [299]:

$$[x] = \frac{-1}{1+e^{((-b1*pH)+b2)}} + 1 \quad \mathrm{pH_T} = \frac{b2}{b1} \qquad \text{(eq. 2)}$$

Where $[x]$ is the folded state population at a given pH value and b1 and b2 are the parameters to be fit via nonlinear regression.

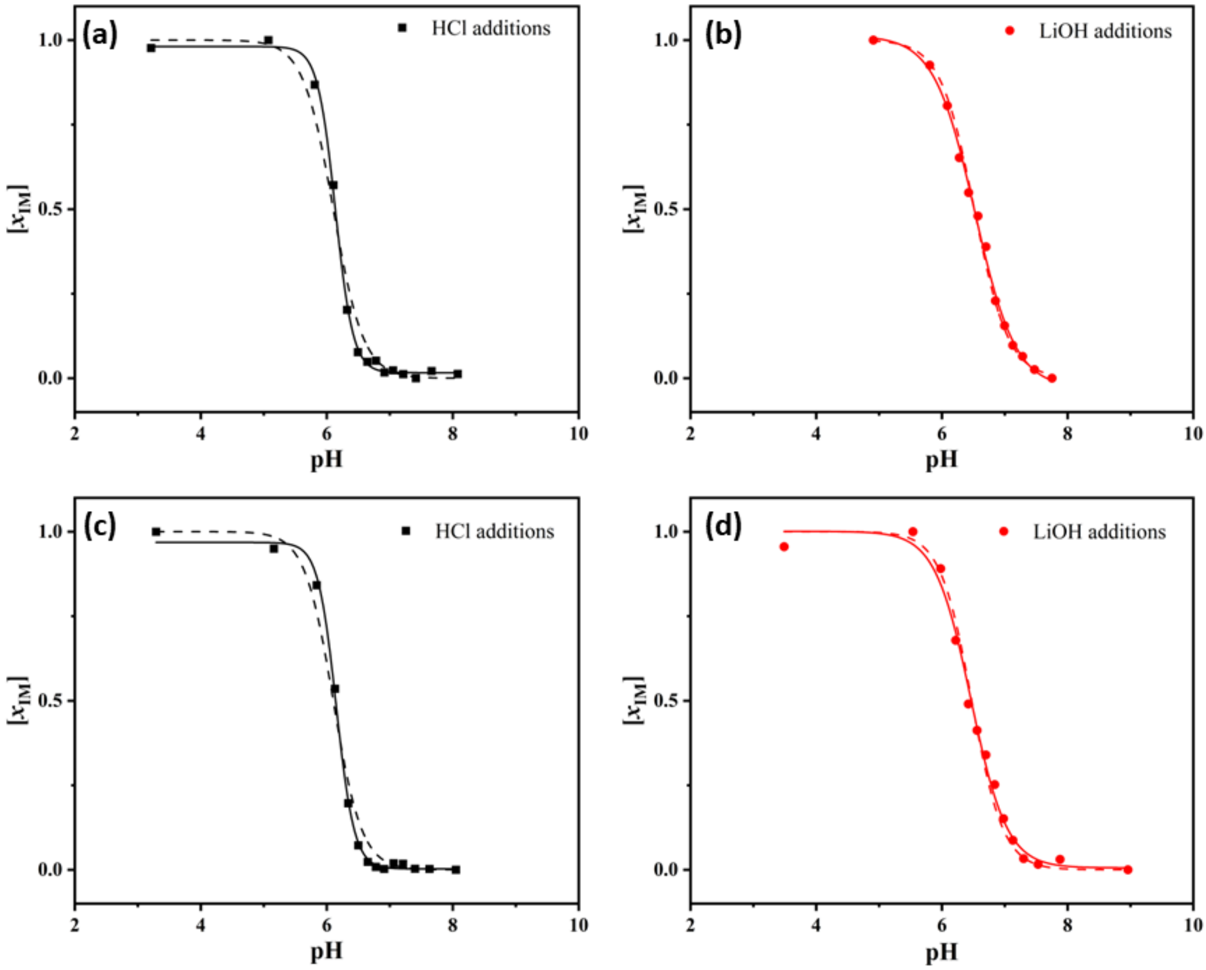


Fig. 65 The fit of the data for pH_T titration on PyA-2 in KPi buffer. ***(a)*** *The fit to the automatic titration with HCl;* ***(b)*** *the fit to the automatic titration with LiOH;* ***(c)*** *the fit to the manual titration with HCl, and* ***(d)*** *the fit to the manual titration with LiOH. Solid lines represent the fitting by using Boltzmann sigmoid equation, whereas dashed lines represent the fitting by using eq. (2).*

The data fits are shown in **Fig. 65** for PyA-2 and in **Fig. S22** for PyA-3; the extracted $\mathrm{pH_T}$ values are shown in **Table 4**. Analysis of the data indicate that both methods give similar $\mathrm{pH_T}$ values. However, Boltzmann sigmoid equation fits better out data.

pH titration in Britton-Robinson (BR) buffer

To linearly cover the range in which titrations were performed (~ 4 to 8) we tested the use of the BR buffer. BR is a "universal" buffer used for the pH range from 2 to 12. It is composed of 0.04 M boric acid, 0.04 M phosphoric acid and 0.04 M acetic acid and titrated to the desired pH with lithium hydroxide [315]. We also added 140 mM of KCl to better mimic the physiological conditions (the buffer name is K-BR).

We collected the CD titration data in the K-BR for PyA-2 (**Fig. 66**) in order to compare its performance relative to the KPi buffer. The titration in K-BR buffer do not show hysteresis suggesting that the process is completely reversible in contrast to what was observed in KPi buffer (**Fig. 64 (c)** and **(d)).**

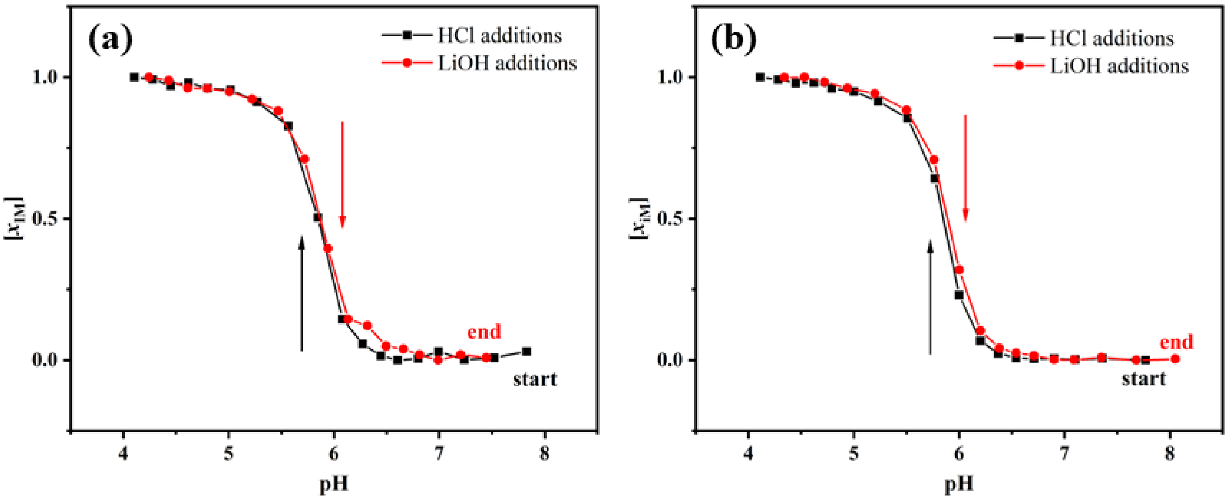


***Fig. 66** Titration curves for PyA-2 in K-BR buffer at 25 ℃. **(a)** Data for automatic titration and **(b)** manual titration for PyA-2, each one was performed with 17 additions of 4 µL of HCl 1 M, followed by 17 additions of 4 µL of LiOH 1 M. [PyA-2] ~ 2 µM.*

We treated the data as described above and display the results in **Table 4**. The pH_T obtained for PyA-2 in BR buffer are ~ 0.3 units lower than that ones in KPi buffer.

Although the aforementioned pH_T values are lower compared to those obtained in KPi buffer, we observed that the initial pH value in BR buffer, as recorded by the pH meter, was approximately 7.82 instead of 8.02. This discrepancy suggests a potential miscalibration of the instrument and would explain the observed difference of 0.2 units. Consequently, the pH meter was recalibrated.

In addition to PyA-2 we determined pH_T for three other representative sequences, PyA-3 (**Fig. S22**), PyB-1 and PyB-2 (**Fig. 67**).

Since no significant differences between automatic and manual titration were observed, and the discrepancy in the pH_T values obtained between the two methods was less than 0.5 units, and automatic titration allows for time savings, the subsequent titrations for PyB-1 and PyB-2 were exclusively conducted using automatic titration (**Fig. 67**).

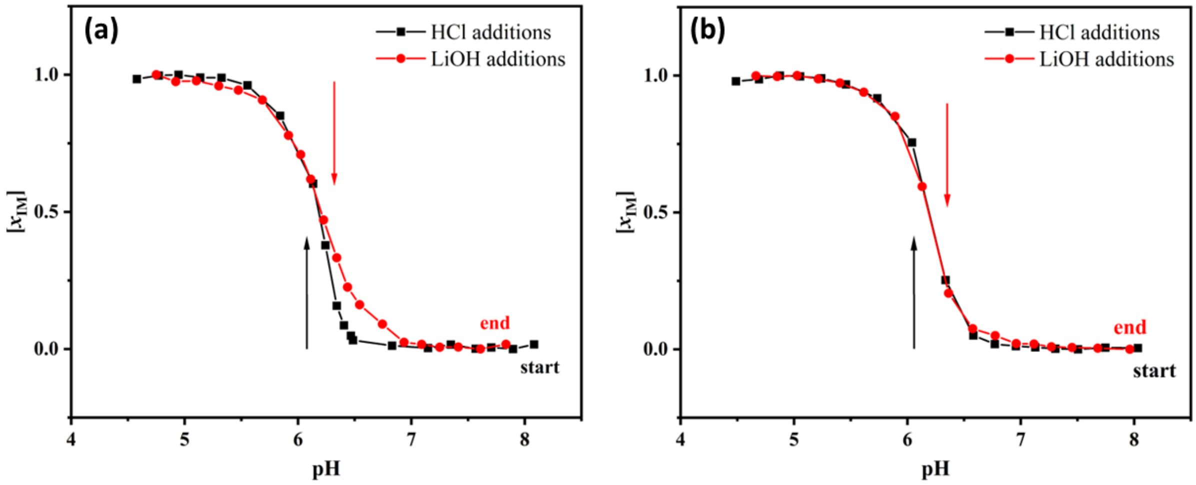


Fig. 67 *Titration curves for PyB-1 (a) and PyB-2 (b) in K-BR buffer at 25 ℃. Each one was performed by automatic titration with 17 additions of 4 μL of HCl 1 M, followed by 17 additions of 4 μL of LiOH 1 M. [PyB-1] ~ 2 μM and [PyB-2] ~ 2 μM in 20 mM K-BR buffer pH = 8.02.*

The data was fitted using the Boltzmann equation and eq (1) (**Fig. S23**) and the corresponding pH_T values are presented in **Table 4**. It is noteworthy that, after the recalibration of the pH meter, the pH_T for PyB-1 was found to be ~ 6.18, indicating an increase of approximately 0.6 units compared to the previous measurements. This observation emphasizes the importance of ensuring that the initial pH value aligns with the pH of the buffer in which the DNA is dissolved prior to initiating the titration process. Moreover, regular calibration of the pH meter and verification of its accuracy using buffer standard solutions with well-established pH values (e.g., 4, 7, and 10) are crucial prerequisites before starting any titration experiment.

Table 4 *pH_T values for PyA-2, PyA-3, PyB-1 and PyB-2. In red the pH_T values obtained when the pH-meter was miscalibrated.*

Sample	**Buffer**	**Titration**	**HCl additions**		**LiOH additions**	
			Boltzmann	**equation (1)**	**Boltzmann**	**equation (1)**
PyA-2	KPi	Automatic	6.14 ± 0.01	6.11 ± 0.02	6.52 ± 0.02	6.52± 0.01
PyA-2	KPi	Manual	6.15 ± 0.01	6.10 ± 0.03	6.46 ± 0.03	6.47 ± 0.02
PyA-3	KPi	Automatic	6.10 ± 0.01	6.07 ± 0.02	6.50 ± 0.03	6.51 ± 0.02
PyA-3	KPi	Manual	6.20 ± 0.01	6.18 ± 0.02	6.41 ± 0.03	6.44 ± 0.02
PyA-2	K-BR	Automatic	5.84 ± 0.01	5.84 ± 0.01	5.87 ± 0.01	5.86 ± 0.02
PyA-2	K-BR	Manual	5.85 ± 0.01	5.81 ± 0.03	5.89 ± 0.01	5.87 ± 0.01
PyB-1	K-BR	Automatic	5.66 ± 0.01	5.63 ± 0.02	5.72 ± 0.01	5.72 ± 0.01
PyB-1	K-BR	Manual	5.61 ± 0.01	5.59 ± 0.02	± 0.01	5.68 ± 0.01
PyB-1	K-BR	Automatic	6.18 ± 0.01	6.11 ± 0.03	6.21± 0.01	6.20 ± 0.01
PyB-2	K-BR	Automatic	6.19 ± 0.01	6.17 ± 0.02	6.17 ± 0.01	6.17 ± 0.02

Furthermore, to verify that once the instrument is calibrated, the obtained pH_T value is repeatable, the titration for PyB-2 was performed 3 times on 3 different days (**Fig. S24** and **Table S1**), Boltzmann equation was used to fit the data. As shown in **Table S1** there is no variability among the measurements, demonstrating that the measurements are perfectly consistent.

To sum up, we develop a method to calculate the pH_T of iM-forming sequences by using CD equipped with automatic titrator.

Automatic pH titration have a series of advantages over manual titrations:

- Faster and more efficient than manual pH titration, reducing the time needed to complete an analysis.
- Reduced operator error

And drawbacks:

- Can be more expensive due to the need for automatic titrator.
- May require more maintenance and calibration.
- Not as flexible as manual titration in terms of adjusting for unexpected changes or variations in samples.
- Surprisingly the data did not look as smooth as some of our manual titrations.

The automatic titration is the best choice when an automatic titrator is available. However, in both cases (automatic and manual) the pH-meter must be calibrated with high periodicity (preferably weekly or daily). To obtain the pH_T we recommend using the Boltzmann sigmoid equation since it fits well the experimental points and often better than to eq (1) although lead to very similar pH_T values. We also recommend using BR buffer (supplemented, if needed with KCl), or in a buffer that has buffering capacity wide enough to cover the pH range between 4 and 8. One should pay close attention to other parameters, such as the thermodynamic stability of DNA under chosen conditions.

4.2.3. Biophysical and X-ray study of C-rich DNA variants from c-MYC: results and discussion

First we performed biophysical studies on PyA and PyB variants (**Table *3***) to characterize them using circular dichroism (CD) and CD melting. CD spectra show that all the variants adopt an i-motif conformation as indicated by the characteristic peak at ~ 288 nm and the trough at ~ 260 nm (**Fig. 68 (a) and (b)**). Suggesting that adding one or two nucleotides to the terminal ends does not prevent the formation of the i-motif structure, which only relies on cytosine base pairing.

CD melting was used to determine the thermodynamic stability of all sequences, and the results were presented in the form of representative melting curves in **Fig. S25** and **Fig. S26**, along with data analysis results shown in **Fig. 68 (c) and (d)** and in **Table *3***.

The melting process exhibited low hysteresis of 3 ± 1 °C in all cases, indicating that it was a reversible process with monomolecular transitions. Thermodynamic stability data suggest that the presence of different nucleotides at the termini of PyA and PyB variants can influence the ΔTm by approximately 9 °C. Specifically, PyA-6 and PyB-2 exhibit the lowest stability, as they are the only two sequences with two additional nucleotides at both termini that probably are able to interfere with the cytosine-cytosine$^+$ base pairs destabilizing the i-motif structure. Conversely, PyA-1, PyA-4, PyB-3, and PyB-4 exhibit the highest stability, which are the only variants with one or two additional nucleotides at the 5'-end. Overall PyA variants shown an higher stability compared to PyB ones, this can probably be attributed to the higher number of tracts formed by 3 cytosine (CCC) in the PyA variants, which may increase the Tm [260].

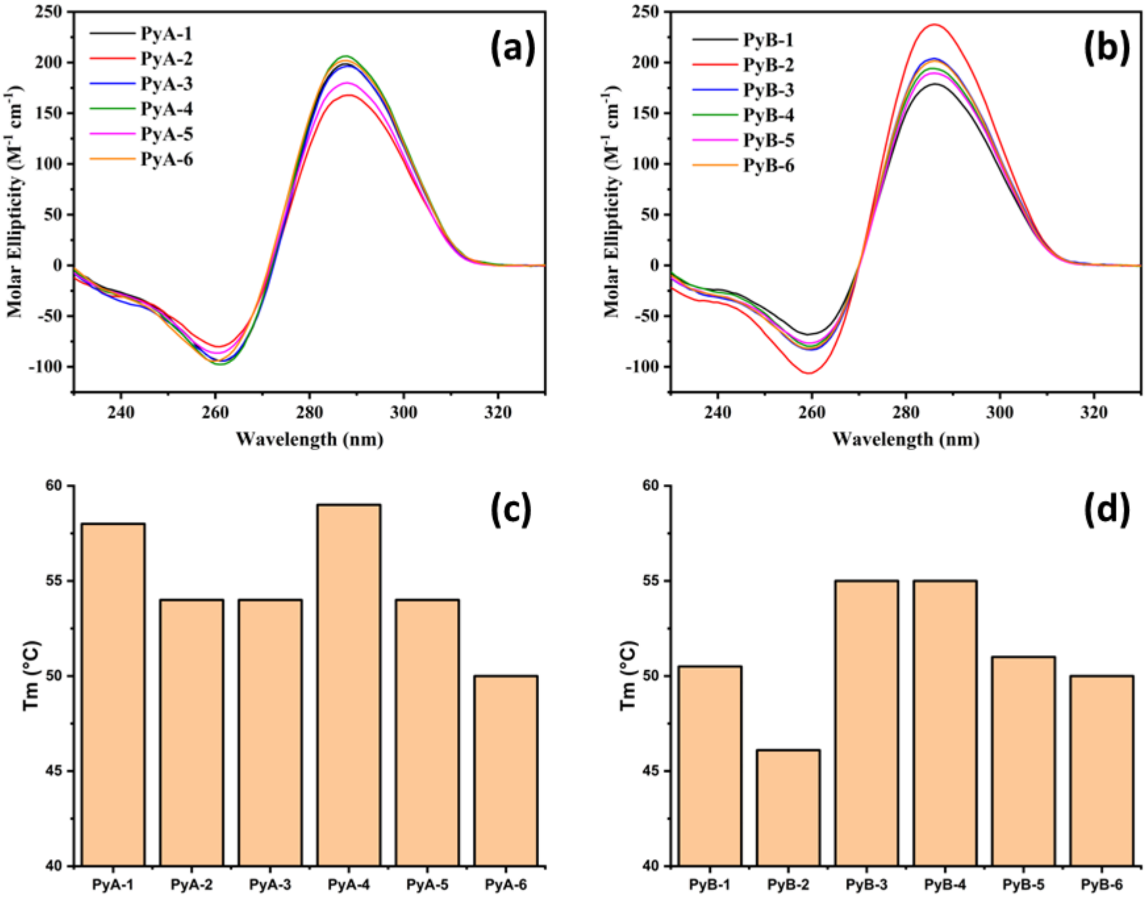


Fig. 68 *CD spectra of PyA* ***(a)*** *and PyB* ***(b)*** *sequences in 10 mM MES, pH = 5.5 at 20 °C. Melting temperatures of PyA* ***(c)*** *and PyB* ***(d)*** *sequences. All samples were prepared at ~ 4 µM DNA concentration.*

Secondly, we performed pH titrations by using circular dichroism in order to obtain the transition pH (pH$_T$) of the variants. We would like to determine the value of pH$_T$ in order to understand whether Py sequences can form iM under pH conditions close to biologically relevant and whether 3' and 5' overhangs influence the pH$_T$ values.

Initially pH titrations were performed in KPi buffer, later to linearly cover the pH range in which titrations were performed (~ 4 to 8) we used the Britton–Robinson (BR) buffer.

Titration spectra were analysed using SVD [305,306]. The data are reported as plots of x_{IM} vs pH, where x_{IM} represents the fraction of folded iM. x_{IM} was obtained by normalizing the **v1** vector at the given pH.

We first performed pH titration on the PyA-2 and PyA-3 sequences. We started at pH = 8 where DNA is unfolded. Gradual addition of HCl resulted in the increase of the 288 nm signal signifying folding of the DNA (**Fig. 64**). The titrations in KPi buffer show hysteresis, indicating a non-reversible process.

Whereas pH titrations performed in BR buffer for PyA-2 (**Fig. 66**), PyB-1 and PyB-2 (**Fig. 67**) sequences, did not show hysteresis, indicating that the process is completely reversible.

We did not performed the pH titrations for all the variants because we assumed that two of PyA and two of PyB variants could be representative for the all group since the number of cytosine pairs do not change among the variants and because we demonstrated that pH_T values for two PyA and PyB variants are very similar.

The pH_T values obtained are summarized in **Table 5**. These values range between 5.8 and 6.2, which are significantly lower than the physiological pH of 7.4. Notably, the value of 5.84 obtained for PyA-2 was determined with a miscalibrated pH meter. However, in certain diseases such as cancer, the pH in specific body locations can become slightly acidic [316], which may facilitate i-motif formation in vivo, additionally it has been recently shown that i-motif formation in the c-MYC promoter region is facilitated by negative superhelicity at physiological pH and salt condition [317].

The variants investigated in this study represent shorter sequences from the selected c-MYC promoter sequence. Indeed, the c-MYC tract selected for this study formed by 35 nucleotides has been investigated by Vorlickova et al. [296] and a pH_T of 6.7 was reported. Generally, under identical experimental conditions, an i-motif structure with a greater number of C:C+ base pairs tends to be more stable [267]. Recently, how the length of the C-tract impacts the formation of intramolecular i-motif structures under physiological conditions was investigated [260,261]. Their findings generally indicate an increase in transition pH (pH_T) as the number of cytosines per tract rises.

The values obtained in KPi buffer for PyA-3 closely resemble those of PyA-2, indicating that the addition of one nucleotide at the 5' end of the PyA sequence does not significantly affect the pH_T. Similar values were also obtained for the PyB-1 and PyB-2 sequences, indicating no significant difference between the two variant groups.

In addition, these values are in line with the pH_T values obtained by other researchers with sequences of similar length [260,296,299].

***Table 5** pH_T values obtained by automatic titration and fitted by using Boltzmann sigmoid equation.*

Sample	**Buffer**	**pH_T ± SD (HCl additions)**	**pH_T ± SD (LiOH additions)**
PyA-2	KPi	6.14 ± 0.01	6.52 ± 0.02
PyA-3	KPi	6.08 ± 0.01	6.50 ± 0.03
PyA-2	BR	5.84 ± 0.01	5.87 ± 0.01
PyB-1	BR	6.18 ± 0.01	6.21 ± 0.01
PyB-2	BR	6.19 ± 0.01	6.17 ± 0.01

In order to solve the 3D structure of the variants we try to crystalize them by screening the crystallization conditions with commercial screens, including Natrix and Helix, by using a TTP Labtech Mosquito Liquid Handling robot system as described in paragraph **4.2.1.**

However, among 12 variants screened, only PyB-2 was successfully crystalized (**Fig. 69 (a)**) in in condition 1–4 from the HELIX screen (Molecular Dimensions): 0.2 M KCl, 25% PEG400, 5 mM $[Co(NH_3)_6]Cl_3$, 0.05M MES pH = 6.5.

This condition was optimized by screening in 24-well trays DNA at 1.5-1.7 mM $[Co(NH_3)_6]Cl_3$ 2-5 mM, PEG400 20-30% and KCl to 0.1-0.3 M. The best crystals were big and show a well-ordered 3D structure, they were obtained within one week in the original home-made 1-4 HELIX condition; no crystals were obtained in 2 mM $[Co(NH_3)_6]Cl_3$, and smaller crystals were obtained in all other optimization instances.

Diffraction data were collected at the Advanced Photon Source synchrotron facility at 24 ID-E beamline at both the native (λ = 0.98 Å, 12622 eV) and the cobalt wavelength (λ = 1.61 Å, 7725 eV). The crystals were analysed in different locations and at different angles (**Fig. 69 (b)**) to acquire as much good quality data as possible. Most of them showed a good diffraction pattern (**Fig. 69 (c)**) without ice rings and smeared spots.

Raw diffraction data were processed using RAPD software provided by the beamline and the three best data collection are reported in **Table 6**.

The quality of all the data were first analysed by Xtriage tool implemented in Phenix [308]. The Xtriage summary reported good quality of the data except for the translational noncrystallographic symmetry (tNCS) (**Fig. S27**), which is higher than 20% and it can dramatically complicate the solution of the crystal structures.

Translational noncrystallographic symmetry (tNCS) refers to a phenomenon observed in crystallography, where multiple copies of a molecule or complex are present in a crystallographic asymmetric unit but are not related by any of the symmetry operations that define the crystal lattice. Instead, they are related by a noncrystallographic symmetry operation [318,319]. It is particularly insidious in causing difficulties in all stages of crystal structure determination, from indexing the diffraction pattern to refining the structure [320,321].

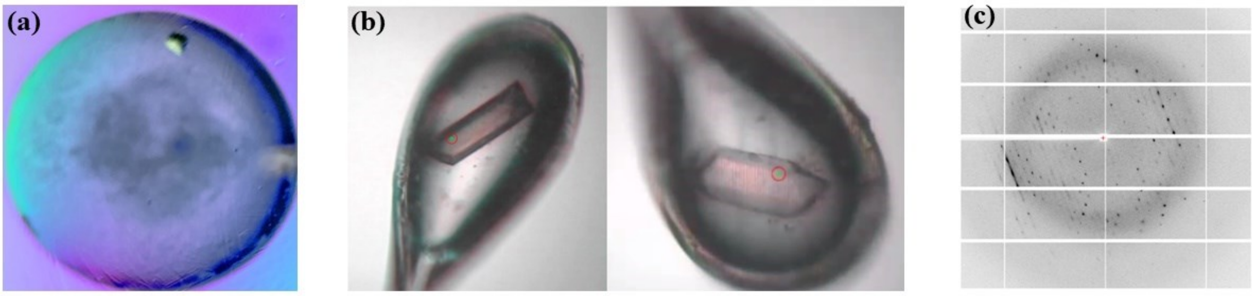


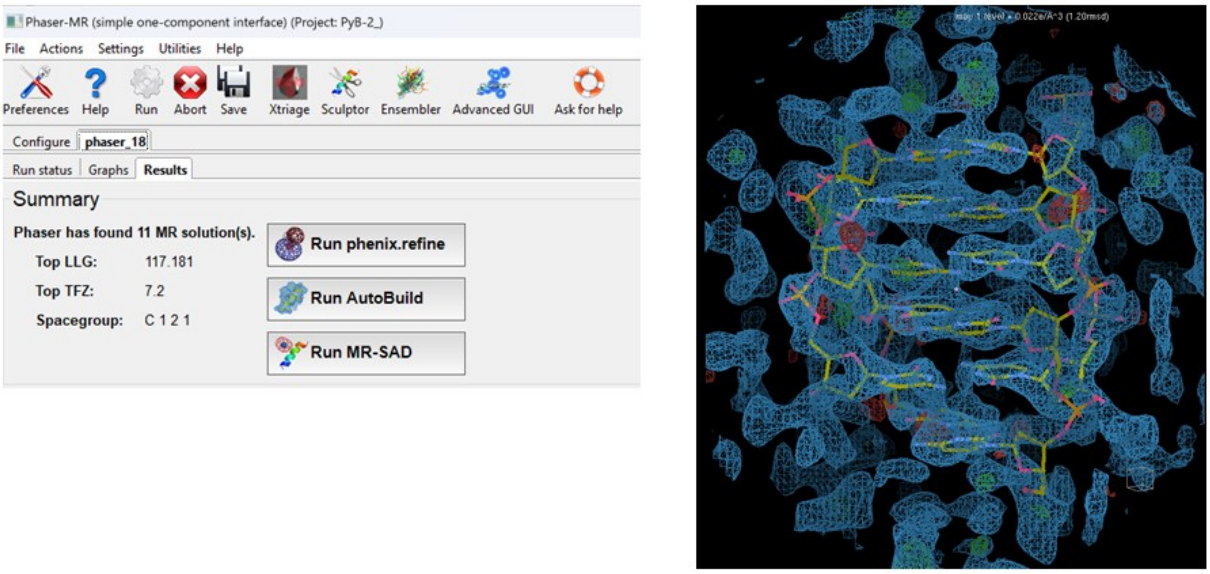
Fig. 69 (a) *PyB-2 crystal grew in a drop containing the 1-4 Helix condition viewed through an optical microscope.* ***(b)*** *Image of the crystals of the PyB-2 in the loop at the 24 ID-E beamline, view from two different angles. The color is due to radiation exposure and does not represent the true color of crystals that were transparent and colorless.* ***(c)*** *X-ray diffraction pattern of a crystal of PyB-2. Obtained by the Advanced Photon Source 24 ID-E synchrotron facility.*

***Table 6** The best three data collections*

	GT1_C1	**GT3_C5**	**GT3_C6**
Space group	C 1 2 1	C 1 2 1	I 2 2 2
Unit cell dimensions			
a, b, c (Å)	49.58, 50.97, 26.91	49.68, 51.06, 27.03	26.98, 41.45, 50.88
α, β, γ (°)	90, 122.75, 90	90, 122.93, 90	90, 90, 90
Resolution (Å)	1.53	1.54	1.86
cc1/2	0.999	0.997	0.99
cc1/2 outer shell	0.64	0.365	0.424
Completeness			
Overall	98.2	96.9	99.8
Outer shell	98.6	96.9	99.6
Multiplicity			
Overall	3.3	3.3	6
Outer shell	3.2	3.2	6.2
I/σ			
Overall	13.6	11.3	9.2
Outer shell	0.6	0.5	0.6
R-merge			
Overall	0.034	0.046	0.098
Outer shell	1.458	2.084	2.231

The next step consisted in solve the crystal structure by starting from a similar existing structure already deposited in the Protein Data Bank (PDB). Different i-Motif models were downloaded from PDB website and were modified in PyMOL in order to have a scaffold with only cytosine base pairs. We produced different models with a different number of cytosine base pairs, ranging from 2 to 6.

After that each modified model was combined with each diffraction data through molecular replacement (MR) in Phenix to output an electron density map. The best results displayed by the software were given combining the modified pdb model (1BQJ) with only six cytosine base pairs and the diffraction data named “GT1_C1” (**Fig. 70**). However the TOP Log Likelihood Gain (LLG), which indicates how well the pdb model resembles the diffraction data and TOP Translation Function Z score **(**TFZ), which indicates the signal-to-noise ratio of the solution, are both lower than the threshold which the large majority of solutions are correct (threshold for TFZ ~ 8 and for LLG ~ 300).


Phaser-MR (simple one-component interface) (Project: PyB-2_)

Summary

Phaser has found 11 MR solution(s).

Top LLG: 117.181
Top TFZ: 7.2
Spacegroup: C 1 2 1

Fig. 70 *On the left the results displayed by MR in Phenix combining 1BQJ pdb model (modified with only Cytosines) and the diffraction data named "GT1_C1". On the right the output MR with the density map opened in Coot.*

Despite the output results not being above the threshold, and our inability to find better results with different models, we decided to refine the structure through extensive manual model building in COOT. Despite our various efforts, we unfortunately have not yet solved the structure.

Finally, We performed native PAGE to examine the homogeneity and oligomeric state of Py variants (**Fig. 71**). All the variants displayed a prominently intense band around the dT15 marker, indicating that they primarily exist in a monomeric form. A second, less intense band around the dT24 marker suggests the presence of dimers, though in low quantities.

Bands displayed a more homogeneous migration on the right side of the gel, while on the left side a 'wave' pattern was observed, with all the bands being more shifted towards the bottom.

However, when the experiment was replicated, a more homogeneous migration was observed across the gel (**Fig. 72**).

Given that the running buffer used has a pH of 6 and from pH titrations we obtained a $pH_T \sim 6$, we can not exclude the simultaneous presence of folded and unfolded structures, in particular the second band around the dT24 marker can be attributed to the unfolded DNA. To determine whether this second higher band is due to the unfolded structure or the formation of dimer, we performed a native PAGE in BR buffer at a pH of 5.5 (**Fig. S28**). The representative PAGE displayed less distinct bands around dT24. However, the smearing suggests the presence of dimers, as at this pH, all the variants are folded into i-Motif structures.

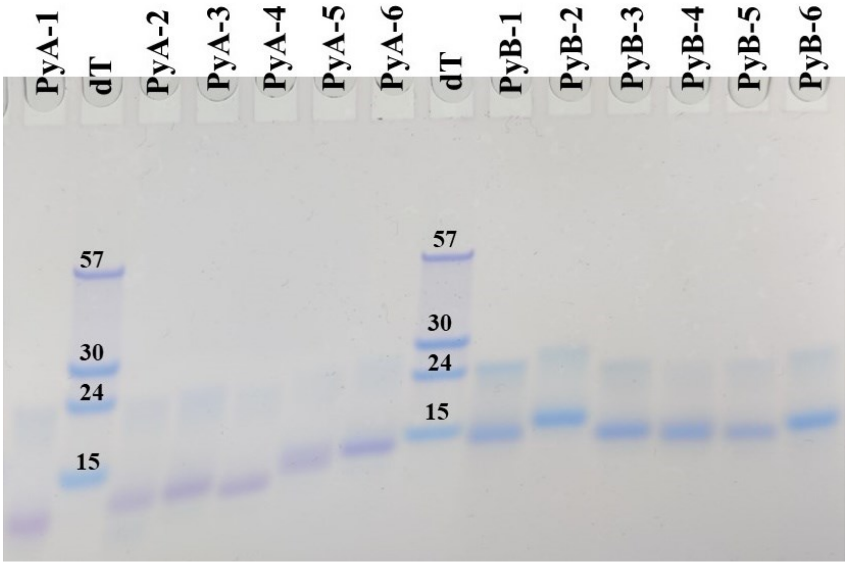


***Fig. 71** Fifteen percent native gel prepared in 10 mM MES buffer pH = 6 supplemented with 5 mM KCl. DNA samples were prepared at ~ 50 µM.*

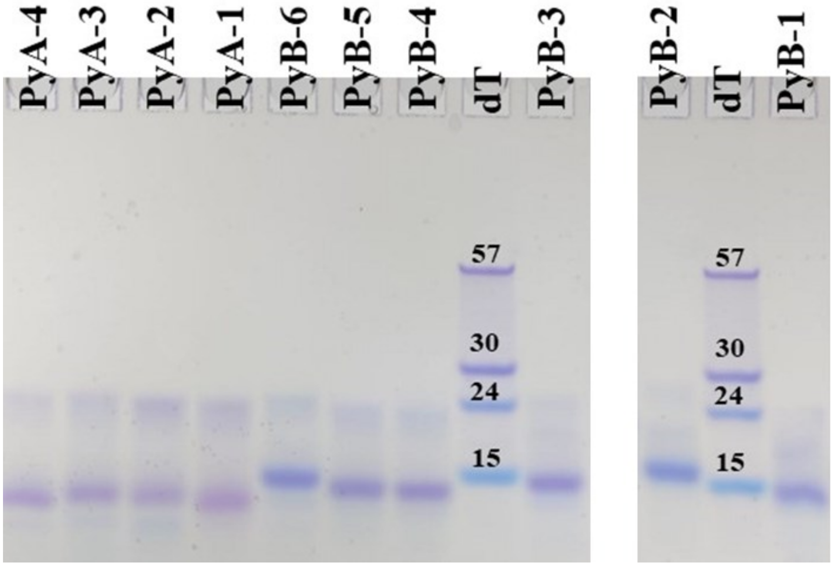


***Fig.** 72 Fifteen percent native gel prepared in 10 mM MES buffer pH = 6 supplemented with 5 mM KCl. DNA samples were prepared at ~ 50 µM. (PyA-5 and PyA-6 are not present because PAGE's lanes were all occupied).*

To gain more insight into the oligomeric state of the PyB-2 crystal, we loaded the sample used to grow crystals to native PAGE (**Fig. 73**, **lane 2**). This sample was annealed at a concentration of 1.5 mM, and on the day of the gel run, it was diluted to 100 µM in MES buffer. Notably, the formation of higher bands was observed, indicating that a higher concentration promotes the formation of higher-order structures. In **Fig. 73, lane 3**, we ran the crystal sample. To prepare it, 5–10 crystals were harvested, washed in the crystallization condition, and then dissolved in 10–15 µL of 10 mM MES buffer. In this case, clear bands are not observable, likely due to incomplete dissolution of the crystals. However, two small spots are visible around dT24 and dT30, and considering the high-ordered species observed in lane 2, we hypothesize that the PyB-2 crystal is not in a monomeric form.

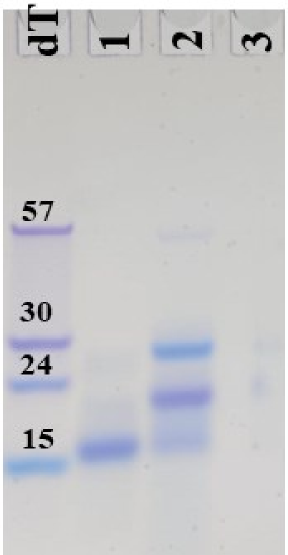


***Fig.** 73 Fifteen percent native gel prepared in 10 mM MES buffer pH = 6 supplemented with 5 mM KCl. Size markers correspond to dTn sequences. The DNA samples were prepared the following way: lane **1** – 50 µM PyB-2 in 10 mM MES buffer; lane **2** – concentrated DNA sample used to grow crystal deposited in lane 3; lane **3** – crystals washed and dissolved in 10 mM MES buffer.*

Given that all biophysical characterizations were performed in 10 mM MES buffer pH = 5.5, we compared the stability of two variants in BR versus MES buffer. Specifically, we performed CD melting experiments for PyA-3 and PyB-1 in 20 mM BR buffer pH = 5.5, supplemented with 140 mM KCl, and compared the results with those obtained using 10 mM MES buffer pH = 5.5 (**Fig. 74**). Surprisingly, we observed a significant difference. In particular, the stability of both variants in BR buffer was significantly lower (by approximately 15 °C), and in the case of PyA-3, two melting events occurred.

To investigate whether KCl was the cause of the difference in Tm between the two buffers, we performed the same CD melting experiments in 20 mM BR buffer without KCl. However, no significant differences were observed. All the melting temperatures as well as the thermodynamic parameters are listed in **Table 7.**

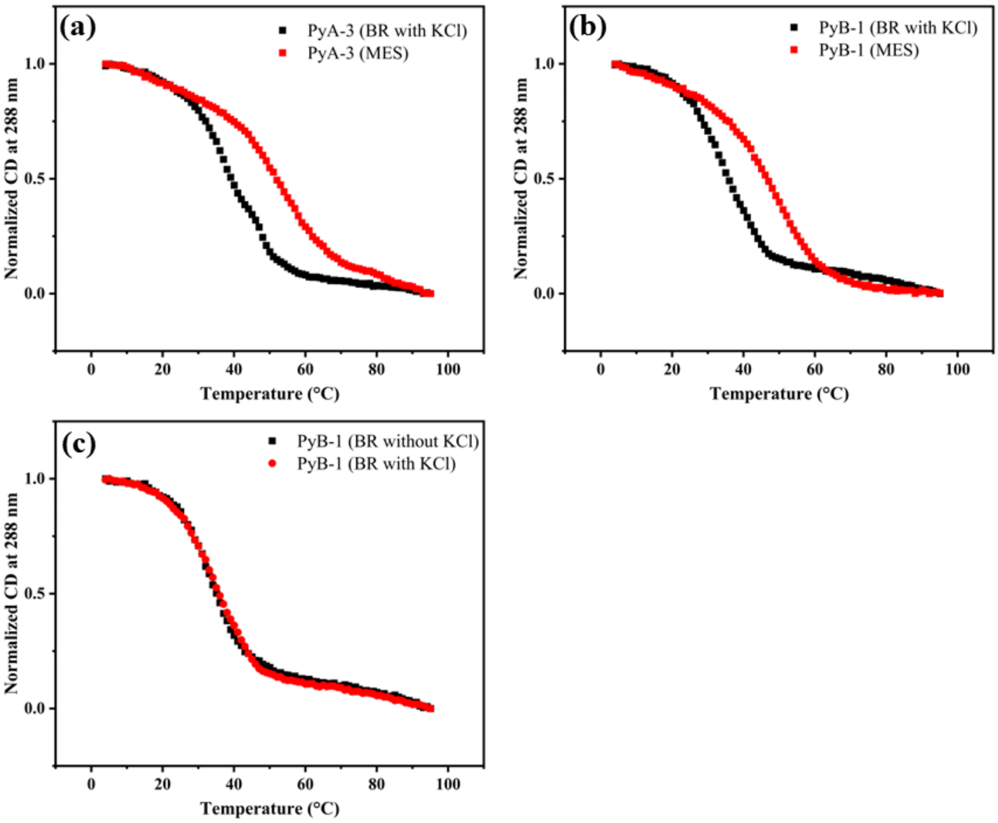


***Fig. 74** CD-melting plots of **(a)** PyA-3 sequence in BR buffer (black squares) and in MES (red squares) and of PyB-1 sequence **(b)** in BR buffer (black squares) and in MES (red squares) **(c)** in BR buffer without KCl (black squares) and in BR buffer with 140 mM KCl (red circles).*

Table 7 Thermodynamic parameters obtained by CD melting experiments in MES and BR buffer.

		HEATING			**COOLING**		
Sequence	Buffer	Tm Derivative, °C; error ± 0.5	T_m Fit, °C	ΔH Fit, kJ/mol	Tm Derivative, °C; error ± 0.5	T_m Fit, °C	ΔH Fit, kJ/mol
PyA-3	MES	54	54	155	53	52	150
PyA-3	BR	37.1 45.1	40.4	138	36	36	223
PyB-1	MES	50.5	50.6	133	49	49.2	119
PyB-1	BR	35.1	35.2	161	30	30.6	204
PyB-1	BR (without KCl)	33	33.6	180	31.1	31.3	172

4.2.4. Final remarks

In this study 12 DNA variants designed from a C-rich tract were investigated by biophysical studies, native PAGE and X-ray crystallography.

CD spectroscopy revealed that all variants adopt an i-Motif conformation. This suggests that even the smaller sequences from c-MYC tract, with additional nucleotides added at the 5’ and 3’ ends, do not impede the formation of i-Motifs at slightly acidic pH. However, the thermodynamic stability is influenced by the different nucleotides present at the terminal ends, with a Tm ranging from 50 °C to 59°C for PyA variants and from 46°C to 55 °C for PyB variants. Notably, for both variant groups, the lowest melting temperature was observed for the sequences with two additional nucleotides present at both ends (PyA-6 and PyB-2). The highest stability was observed for sequences with additional nucleotides only at the 5’ termini (PyA-1, PyA-4, PyB-3, and PyB-4).

From the CD pH titrations, we obtained a pH_T of approximately 6.2. This value was measured for two variants from each group, assuming that this value should be similar for the other variants given the consistent number of C:C+ base pairs among them. This value is lower than the pH_T value reported by Vorlickova et al. [296], where they investigated the full selected c-MYC tract composed of 33 nucleotides (5’-TCCCCACCTTCCCCACCCTCCCCACCCTCCCCA-3’). This suggests that the shorter sequences investigated in this study are less likely to adopt an i-Motif conformation in vivo at physiological pH. It is more probable that they form such conformations when they have more C tracts.

Native PAGE indicated that all the variants are primarily in a monomeric form, even though dimers are also present. Furthermore, it was observed that at higher concentrations, the PyB-2 sequence can form high order species, and it is probable that the PyB-2 crystal exists in an oligomeric state.

Among the 12 variants, only PyB-2 was successfully crystallized. However, despite the good resolution (1.52 Å) of the diffraction data, all the raw data demonstrated high translational noncrystallographic symmetry (~ 21%), which has, so far, prevented the resolution of the structure.

5. Conclusions and perspectives

In this thesis I summarized the results and findings obtained about the secondary structures of various polynucleic acids. These were investigated using biophysical methods, and in some instances, porphyrins were employed as specific reporters.

Herein, I illustrated how the different secondary structures of micro-RNAs and their thermodynamic stability could influence their biological activity (*chap. 2.1*). Evidence shows that a single nucleotide change outside the seed region can significantly enhance the thermodynamic stability. Furthermore, a luciferase reporter assay demonstrates how this increased stability can improve the miRNA's efficiency in repressing its target.

I demonstrated the formation of supramolecular assemblies between various achiral porphyrins and a mature miRNA (*chap. 2.3*). In particular, porphyrins, based on their properties, interact with miRNA differently, as reflected by the unique induced circular dichroism (ICD) signals in the Soret region, moreover at the ratio porphyrin:miR 3:1 porphyrins are able to destabilize the miRNA's secondary structures.

I also explored the interaction between the ZnTCPPSpm4 porphyrin and the telomeric-RNA sequence, which adopts a parallel G-quadruplex topology (*chap. 3.2*). Various spectroscopic techniques revealed that porphyrins are orderly distributed across the entire G4 structure, interacting with each other along the assembly. Furthermore, the presence of shorter sequences, likely derived from nuclease digestion, inhibits the formation of the antiparallel G-quadruplex topology.

Lastly, I examined two groups of cytosine-rich variants with different nucleotides added at the 5' and 3' ends (*chap. 4.2*). These sequences were selected from the human c-MYC promoter. Our biophysical investigations revealed that all these sequences are capable of adopting an i-motif (iM) conformation, exhibiting similar transition pH. We observed minor differences in thermodynamic stability among the sequences. We have successfully crystalized one of the 12 variants, and we are currently working to solve its 3D structure. This will provide deeper insight into the binding mode of existing iM ligands and facilitate the design of new scaffolds that selectively bind to c-MYC and iMs in general.

In the future, we will continue to expand our understanding of nucleic acid structures and their significant roles in biological systems. This comprehensive knowledge could open up new perspectives in disease understanding, diagnosis, and treatment. Additionally, the exploration of molecular interactions, particularly with porphyrins, will offer innovative applications not only in biomedicine but also in areas such as biotechnology and nanotechnology. Overall, this research direction holds the promise of significant scientific advancements and broad implications for various fields.

References

1. Anfinsen, C.B. Principles That Govern the Folding of Protein Chains. *Science (1979)* **1973**, *181*, 223–230, doi:10.1126/SCIENCE.181.4096.223/ASSET/1A648F0F-E322-4108-98DD-B0857103738D/ASSETS/SCIENCE.181.4096.223.FP.PNG.
2. Sanvictores, T.; Farci, F. Biochemistry, Primary Protein Structure. *StatPearls* **2022**.
3. Dobson, C.M. Principles of Protein Folding, Misfolding and Aggregation. *Semin Cell Dev Biol* **2004**, *15*, 3–16, doi:10.1016/J.SEMCDB.2003.12.008.
4. Weikl, T.R.; Paul, F. Conformational Selection in Protein Binding and Function. *Protein Sci* **2014**, *23*, 1508, doi:10.1002/PRO.2539.
5. Ellis, R.J.; Pinheiro, T.J.T. Danger — Misfolding Proteins. *Nature 2002 416:6880* **2002**, *416*, 483–484, doi:10.1038/416483a.
6. Walsh, D.M.; Klyubin, I.; Fadeeva, J. V.; Cullen, W.K.; Anwyl, R.; Wolfe, M.S.; Rowan, M.J.; Selkoe, D.J. Naturally Secreted Oligomers of Amyloid β Protein Potently Inhibit Hippocampal Long-Term Potentiation in Vivo. *Nature 2002 416:6880* **2002**, *416*, 535–539, doi:10.1038/416535a.
7. Chaudhuri, T.K.; Paul, S. Protein-Misfolding Diseases and Chaperone-Based Therapeutic Approaches. *FEBS J* **2006**, *273*, 1331–1349, doi:10.1111/J.1742-4658.2006.05181.X.
8. Biol, F.C.-S.S.E.; 1958, undefined On Protein Synthesis. *academia.edu*.
9. Crick, F.; Barnett, L.; Brenner, S.; Nature, R.W.-T.-; 1961, undefined General Nature of the Genetic Code for Proteins. *Springer* **1961**.

10. Harvey, S.C.; Luo, J.; Lavery, R. DNA Stem-Loop Structures in Oligopurine-Oligopyrimidine Triplexes. *Nucleic Acids Res* **1988**, *16*, 11795, doi:10.1093/NAR/16.24.11795.
11. Sundquist, W.I.; Klug, A. Telomeric DNA Dimerizes by Formation of Guanine Tetrads between Hairpin Loops. *Nature 1989 342:6251* **1989**, *342*, 825–829, doi:10.1038/342825a0.
12. Kang, C.; Zhang, X.; Ratliff, R.; Moyzis, R.; Rich, A. Crystal Structure of Four-Stranded Oxytricha Telomeric DNA. *Nature 1992 356:6365* **1992**, *356*, 126–131, doi:10.1038/356126a0.
13. Crawford, J.L.; Kolpak, F.J.; Wang, A.H.; Quigley, G.J.; van Boom, J.H.; van der Marel, G.; Rich, A. The Tetramer d(CpGpCpG) Crystallizes as a Left-Handed Double Helix. *Proc Natl Acad Sci U S A* **1980**, *77*, 4016–4020, doi:10.1073/PNAS.77.7.4016.
14. Mace, H.A.F.; Pelham, H.R.B.; Travers, A.A. Association of an S1 Nuclease-Sensitive Structure with Short Direct Repeats 5' of Drosophila Heat Shock Genes. *Nature* **1983**, *304*, 555–557, doi:10.1038/304555A0.
15. Wang, G.; Vasquez, K.M. Dynamic Alternative DNA Structures in Biology and Disease. *Nature Reviews Genetics 2022 24:4* **2022**, *24*, 211–234, doi:10.1038/s41576-022-00539-9.
16. Chastain, M.; Tinoco, I. Structural Elements in RNA. *Prog Nucleic Acid Res Mol Biol* **1991**, *41*, 131–177, doi:10.1016/S0079-6603(08)60008-2.
17. Conn, G.L.; Draper, D.E. RNA Structure. *Curr Opin Struct Biol* **1998**, *8*, 278–285, doi:10.1016/S0959-440X(98)80059-6.
18. Kim, J.; Cheong, C.; Moore, P.B. Tetramerization of an RNA Oligonucleotide Containing a GGGG Sequence. *Nature 1991 351:6324* **1991**, *351*, 331–332, doi:10.1038/351331a0.

19. Christiansen, J.; Kofod, M.; Nielsen, F.C. A Guanosine Quadruplex and Two Stable Hairpins Flank a Major Cleavage Site in Insulin-like Growth Factor II MRNA. *Nucleic Acids Res* **1994**, *22*, 5709–5716, doi:10.1093/NAR/22.25.5709.
20. Svoboda, P.; Di Cara, A. Hairpin RNA: A Secondary Structure of Primary Importance. *Cellular and Molecular Life Sciences* **2006**, *63*, 901–918, doi:10.1007/S00018-005-5558-5/METRICS.
21. McManus, C.J.; Graveley, B.R. RNA Structure and the Mechanisms of Alternative Splicing. *Curr Opin Genet Dev* **2011**, *21*, 373–379, doi:10.1016/J.GDE.2011.04.001.
22. Al-Hashimi, H.M.; Walter, N.G. RNA Dynamics: It Is about Time. *Curr Opin Struct Biol* **2008**, *18*, 321–329, doi:10.1016/J.SBI.2008.04.004.
23. Bansal, A.; Kaushik, S.; Kukreti, S. Non-Canonical DNA Structures: Diversity and Disease Association. *Front Genet* **2022**, *13*, doi:10.3389/FGENE.2022.959258.
24. Spitale, R.C.; Incarnato, D. Probing the Dynamic RNA Structurome and Its Functions. *Nature Reviews Genetics 2022 24:3* **2022**, *24*, 178–196, doi:10.1038/s41576-022-00546-w.
25. Childs-Disney, J.L.; Yang, X.; Gibaut, Q.M.R.; Tong, Y.; Batey, R.T.; Disney, M.D. Targeting RNA Structures with Small Molecules. *Nature Reviews Drug Discovery 2022 21:10* **2022**, *21*, 736–762, doi:10.1038/s41573-022-00521-4.
26. Yang, L.; Zhang, L.M. Chemical Structural and Chain Conformational Characterization of Some Bioactive Polysaccharides Isolated from Natural Sources. *Carbohydr Polym* **2009**, *76*, 349–361, doi:10.1016/J.CARBPOL.2008.12.015.
27. Yu, Y.; Shen, M.; Song, Q.; Xie, J. Biological Activities and Pharmaceutical Applications of Polysaccharide from Natural Resources: A Review. *Carbohydr Polym* **2018**, *183*, 91–101, doi:10.1016/J.CARBPOL.2017.12.009.

28. Meng, Y.; Lyu, F.; Xu, X.; Zhang, L. Recent Advances in Chain Conformation and Bioactivities of Triple-Helix Polysaccharides. *Biomacromolecules* **2020**, *21*, 1653–1677, doi:10.1021/ACS.BIOMAC.9B01644/ASSET/IMAGES/LARGE/BM9B01644_0022.JPEG.
29. Seelig, J.; Seelig, A. Lipid Conformation in Model Membranes and Biological Membranes. *Q Rev Biophys* **1980**, *13*, 19–61, doi:10.1017/S0033583500000305.
30. Simons, K.; Ikonen, E. Functional Rafts in Cell Membranes. *Nature 1997 387:6633* **1997**, *387*, 569–572, doi:10.1038/42408.
31. Lehn, J. -M Supramolecular Chemistry—Scope and Perspectives Molecules, Supermolecules, and Molecular Devices (Nobel Lecture). *Angewandte Chemie International Edition in English* **1988**, *27*, 89–112, doi:10.1002/ANIE.198800891.
32. Persch, E.; Dumele, O.; Diederich, F. Molecular Recognition in Chemical and Biological Systems. *Angewandte Chemie International Edition* **2015**, *54*, 3290–3327, doi:10.1002/ANIE.201408487.
33. Pabo, C.; Sauer, R.T. TRANSCRIPTION FACTORS: Structural Families and Principles of DNA Recognition. **1992**.
34. Ni, G.; Du, Y.; Tang, F.; Liu, J.; Zhao, H.; Chen, Q. Review of α-Nucleosides: From Discovery, Synthesis to Properties and Potential Applications. *RSC Adv* **2019**, *9*, 14302–14320, doi:10.1039/C9RA01399G.
35. Neidle, S. Principles of Nucleic Acid Structure. *Principles of Nucleic Acid Structure* **2007**, 1–289, doi:10.1016/B978-0-12-369507-9.X5001-8.
36. Lu, X.J.; Olson, W.K. Resolving the Discrepancies among Nucleic Acid Conformational Analyses. *J Mol Biol* **1999**, *285*, 1563–1575, doi:10.1006/JMBI.1998.2390.

37. Olson, W.K.; Bansal, M.; Burley, S.K.; Dickerson, R.E.; Gerstein, M.; Harvey, S.C.; Heinemann, U.; Lu, X.J.; Neidle, S.; Shakked, Z.; et al. A Standard Reference Frame for the Description of Nucleic Acid Base-Pair Geometry. *J Mol Biol* **2001**, *313*, 229–237, doi:10.1006/JMBI.2001.4987.

38. Lu, X.J.; Olson, W.K. 3DNA: A Software Package for the Analysis, Rebuilding and Visualization of Three-dimensional Nucleic Acid Structures. *Nucleic Acids Res* **2003**, *31*, 5108–5121, doi:10.1093/NAR/GKG680.

39. Murray, R.K. (Robert K. Harper's Illustrated Biochemistry. **2009**, 693.

40. Altona, C.; Sundaralingam, M. Conformational Analysis of the Sugar Ring in Nucleosides and Nucleotides. A New Description Using the Concept of Pseudorotation. *J Am Chem Soc* **1972**, *94*, 8205–8212, doi:10.1021/JA00778A043.

41. Neidle, S.; Berman, H.M. X-Ray Crystallographic Studies of Nucleic Acids and Nucleic Acid-Drug Complexes. *Prog Biophys Mol Biol* **1983**, *41*, 43–66, doi:10.1016/0079-6107(83)90025-1.

42. Van Dam, L.; Ouwerkerk, N.; Brinkmann, A.; Raap, J.; Levitt, M.H. Solid-State NMR Determination of Sugar Ring Pucker in 13C-Labeled 2′-Deoxynucleosides. *Biophys J* **2002**, *83*, 2835–2844, doi:10.1016/S0006-3495(02)75292-2.

43. Rosemeyer, H.; Toth, G.; Golankiewicz, B.; Kazimierczuk, Z.; Bourgeois, W.; Kretschemer, U.; Muth, H.P.; Seela, F. Syn-Anti Conformational Analysis of Regular and Modified Nucleosides by 1D 1H NOE Difference Spectroscopy: A Simple Graphical Method Based on Conformationally Rigid Molecules. *Journal of Organic Chemistry* **1990**, *55*, 5784–5790, doi:10.1021/JO00309A024/SUPPL_FILE/JO00309A024_SI_001.PDF.

44. Foloppe, N.; Hartmann, B.; Nilsson, L.; Mackerell, A.D. Intrinsic Conformational Energetics Associated with the Glycosyl Torsion in DNA: A Quantum Mechanical Study. **2002**.

45. Suck, D.; Saenger, W.; Vorbrüggen, H. Conformation of 6-Methyluridine — a Pyrimidine Nucleoside in the Syn Conformation. *Nature 1972 235:5337* **1972**, *235*, 333–334, doi:10.1038/235333a0.

46. Sundaralingam, M. Stereochemistry of Nucleic Acids and Their Constituents. IV. Allowed and Preferred Conformations of Nucleosides, Nucleoside Mono-, Di-, Tri-, Tetraphosphates, Nucleic Acids and Polynucleotides. *Biopolymers* **1969**, *7*, 821–860, doi:10.1002/BIP.1969.360070602.

47. Neii, S.; Le,); Bermani, H.M. X-RAY CRYSTALLOGRAPHIC STUDIES OF NUCLEIC ACIDS AND NUCLEIC ACID DRUG COMPLEXES. *41*.

48. Kim, S.-H.; Berman, H.M.; Seeman, N.C.; Newton, M.D.; IUCr Seven Basic Conformations of Nucleic Acid Structural Units. *urn:issn:0567-7408* **1973**, *29*, 703–710, doi:10.1107/S0567740873003201.

49. Svozil, D.; Kalina, J.; Omelka, M.; Schneider, B. DNA Conformations and Their Sequence Preferences. *Nucleic Acids Res* **2008**, *36*, 3690, doi:10.1093/NAR/GKN260.

50. Lescrinier, E.; Froeyen, M.; Herdewijn, P. SURVEY AND SUMMARY: Difference in Conformational Diversity between Nucleic Acids with a Six-Membered ‘Sugar’ Unit and Natural ‘Furanose’ Nucleic Acids. *Nucleic Acids Res* **2003**, *31*, 2975, doi:10.1093/NAR/GKG407.

51. Watson, J.D.; Crick, F.H.C. Molecular Structure of Nucleic Acids; a Structure for Deoxyribose Nucleic Acid. *Nature* **1953**, *171*, 737–738, doi:10.1038/171737A0.

52. Klug, A. Rosalind Franklin and the Discovery of the Structure of DNA. *Nature 1968 219:5156* **1968**, *219*, 808–810, doi:10.1038/219808a0.

53. Belmont, P.; Constant, J.F.; Demeunynck, M. Nucleic Acid Conformation Diversity: From Structure to Function and Regulation. *Chem Soc Rev* **2001**, *30*, 70–81, doi:10.1039/A904630E.

54. Wahl, M.C.; Sundaralingam, M. Crystal Structures of A-DNA. **1997**, *44*, 45–63, doi:10.1002/(SICI)1097-0282(1997)44:1.

55. Dickerson, R.E.; Ng, H.L. DNA Structure from A to B. *Proc Natl Acad Sci U S A* **2001**, *98*, 6986–6988, doi:10.1073/PNAS.141238898/ASSET/B95004F3-0D66-431F-9C9B-F0634991EDE5/ASSETS/PNAS.141238898.FP.PNG.

56. Ravichandran, S.; Subramani, V.K.; Kim, K.K. Z-DNA in the Genome: From Structure to Disease. *Biophys Rev* **2019**, *11*, 383–387, doi:10.1007/S12551-019-00534-1/TABLES/1.

57. Herbert, A.; Rich, A. Left-Handed Z-DNA: Structure and Function. *Genetica* **1999**, *106*, 37–47, doi:10.1023/A:1003768526018/METRICS.

58. Rich, A.; Zhang, S. Z-DNA: The Long Road to Biological Function. *Nature Reviews Genetics 2003 4:7* **2003**, *4*, 566–572, doi:10.1038/nrg1115.

59. Heinemann, U.; Roske, Y. Symmetry in Nucleic-Acid Double Helices. *Symmetry (Basel)* **2020**, *12*, 737, doi:10.3390/SYM12050737.

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

61. Mirkin, S.M. H-DNA AND RELATED STRUCTURES. *Annu. Rev. Biophys. Biomol. Struct* **1994**, *23*, 541–576.

62. Schroth, G.P.; Ho, P.S. Occurrence of Potential Cruciform and H-DNA Forming Sequences in Genomic DNA. *Nucleic Acids Res* **1995**, *23*, 1977–1983, doi:10.1093/NAR/23.11.1977.

63. Lipps, H.J.; Rhodes, D. G-Quadruplex Structures: In Vivo Evidence and Function. *Trends Cell Biol* **2009**, *19*, 414–422, doi:10.1016/J.TCB.2009.05.002.

64. Travers, A.; Muskhelishvili, G. DNA Structure and Function. *FEBS J* **2015**, *282*, 2279–2295, doi:10.1111/FEBS.13307.

65. Day, H.A.; Pavlou, P.; Waller, Z.A.E. I-Motif DNA: Structure, Stability and Targeting with Ligands. *Bioorg Med Chem* **2014**, *22*, 4407–4418, doi:10.1016/J.BMC.2014.05.047.

66. Shing, P.; Carter, M. DNA Structure: Alphabet Soup for the Cellular Soul. *DNA Replication-Current Advances* **2011**, doi:10.5772/18536.

67. Giese, M.R.; Betschart, K.; Dale, T.; Riley, C.K.; Rowan, C.; Sprouse, K.J.; Serra, M.J. Stability of RNA Hairpins Closed by Wobble Base Pairs. *Biochemistry* **1998**, *37*, 1094–1100, doi:10.1021/BI972050V/SUPPL_FILE/BI1094.PDF.

68. Varani, G.; McClain, W.H. The G·U Wobble Base Pair. *EMBO Rep* **2000**, *1*, 18–23, doi:10.1093/EMBO-REPORTS/KVD001.

69. Conn, G.L.; Draper, D.E. RNA Structure. *Curr Opin Struct Biol* **1998**, *8*, 278–285, doi:10.1016/S0959-440X(98)80059-6.

70. Mamuye, A.L.; Merelli, E.; Tesei, L. A Graph Grammar for Modelling RNA Folding. *Electronic Proceedings in Theoretical Computer Science, EPTCS* **2016**, *231*, 31–41, doi:10.4204/EPTCS.231.3.

71. Leontis, N.B.; Westhof, E. Geometric Nomenclature and Classification of RNA Base Pairs. *RNA* **2001**, *7*, 499, doi:10.1017/S1355838201002515.

72. Blythe, A.J.; Fox, A.H.; Bond, C.S. The Ins and Outs of LncRNA Structure: How, Why and What Comes Next? *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms* **2016**, *1859*, 46–58, doi:10.1016/J.BBAGRM.2015.08.009.

73. Flores, J.K.; Ataide, S.F. Structural Changes of RNA in Complex with Proteins in the SRP. *Front Mol Biosci* **2018**, *5*, 330379, doi:10.3389/FMOLB.2018.00007/BIBTEX.

74. Lescoute, A.; Leontis, N.B.; Massire, C.; Westhof, E. Recurrent Structural RNA Motifs, Isostericity Matrices and Sequence Alignments. *Nucleic Acids Res* **2005**, *33*, 2395, doi:10.1093/NAR/GKI535.

75. Lescoute, A.; Westhof, E. Topology of Three-Way Junctions in Folded RNAs. *RNA* **2006**, *12*, 83, doi:10.1261/RNA.2208106.

76. Butcher, S.E.; Pyle, A.M. The Molecular Interactions That Stabilize RNA Tertiary Structure: RNA Motifs, Patterns, and Networks. *Acc Chem Res* **2011**, *44*, 1302–1311, doi:10.1021/AR200098T/ASSET/IMAGES/LARGE/AR-2011-00098T_0012.JPEG.

77. Deng, J.; Fang, X.; Huang, L.; Li, S.; Xu, L.; Ye, K.; Zhang, J.; Zhang, K.; Zhang, Q.C. RNA Structure Determination: From 2D to 3D. *Fundamental Research* **2023**, doi:10.1016/J.FMRE.2023.06.001.

78. Auwärter, W.; Écija, D.; Klappenberger, F.; Barth, J. v. Porphyrins at Interfaces. *Nature Chemistry 2015 7:2* **2015**, *7*, 105–120, doi:10.1038/nchem.2159.

79. Battersby, A.R. Tetrapyrroles: The Pigments of Life. *Nat Prod Rep* **2000**, *17*, 507–526, doi:10.1039/B002635M.

80. Saito, S.; Osuka, A.; Vogel, E.; Osuka, A.; Saito, S. Expanded Porphyrins: Intriguing Structures, Electronic Properties, and Reactivities. *Angewandte Chemie International Edition* **2011**, *50*, 4342–4373, doi:10.1002/ANIE.201003909.

81. Tran-Thi, T.H. Assemblies of Phthalocyanines with Porphyrins and Porphyrazines: Ground and Excited State Optical Properties. *Coord Chem Rev* **1997**, *160*, 53–91, doi:10.1016/S0010-8545(96)01341-0.

82. Senge, M.O.; Fazekas, M.; Notaras, E.G.A.; Blau, W.J.; Zawadzka, M.; Locos, O.B.; Mhuircheartaigh, E.M.N. Nonlinear Optical Properties of Porphyrins. *Advanced Materials* **2007**, *19*, 2737–2774, doi:10.1002/ADMA.200601850.

83. Adinehnia, M.; Eskelsen, J.R.; Hipps, K.W.; Mazur, U. Mechanical Behavior of Crystalline Ionic Porphyrins. *https://doi.org/10.1142/S1088424619500147* **2019**, 855–866, doi:10.1142/S1088424619500147.

84. Phillips, J.N. Physico-Chemical Properties of Porphyrins. *Comprehensive Biochemistry* **1963**, *9*, 34–72, doi:10.1016/B978-1-4831-9718-0.50010-2.

85. Biesaga, M.; Pyrzyńska, K.; Trojanowicz, M. Porphyrins in Analytical Chemistry. A Review. *Talanta* **2000**, *51*, 209–224, doi:10.1016/S0039-9140(99)00291-X.

86. Brothers, P.J.; Collman, J.P. The Organometallic Chemistry of Transition-Metal Porphyrin. *Acc. Chem. Res* **1986**, *19*, 209–215.

87. Chilukuri, B.; Mazur, U.; Hipps, K.W. Structure, Properties, and Reactivity of Porphyrins on Surfaces and Nanostructures with Periodic DFT Calculations. *Applied Sciences 2020, Vol. 10, Page 740* **2020**, *10*, 740, doi:10.3390/APP10030740.

88. Huang, H.; Song, W.; Rieffel, J.; Lovell, J.F. Emerging Applications of Porphyrins in Photomedicine. *Front Phys* **2015**, *3*, 23, doi:10.3389/FPHY.2015.00023/BIBTEX.

89. Imran, M.; Ramzan, M.; Qureshi, A.K.; Azhar Khan, M.; Tariq, M. Emerging Applications of Porphyrins and Metalloporphyrins in Biomedicine and Diagnostic Magnetic Resonance Imaging. *Biosensors 2018, Vol. 8, Page 95* **2018**, *8*, 95, doi:10.3390/BIOS8040095.

90. Barona-Castaño, J.C.; Carmona-Vargas, C.C.; Brocksom, T.J.; de Oliveira, K.T.; Graça, M.; Neves, P.M.S.; Amparo, M.; Faustino, F. Porphyrins as Catalysts in Scalable Organic Reactions. *Molecules 2016, Vol. 21, Page 310* **2016**, *21*, 310, doi:10.3390/MOLECULES21030310.

91. Li, L.L.; Diau, E.W.G. Porphyrin -Sensitized Solar Cells. *Chem Soc Rev* **2012**, *42*, 291–304, doi:10.1039/C2CS35257E.

92. Paolesse, R.; Nardis, S.; Monti, D.; Stefanelli, M.; di Natale, C. Porphyrinoids for Chemical Sensor Applications. *Chem Rev* **2017**, *117*, 2517–2583, doi:10.1021/ACS.CHEMREV.6B00361/ASSET/IMAGES/LARGE/CR-2016-003615_0078.JPEG.

93. Jurow, M.; Schuckman, A.E.; Batteas, J.D.; Drain, C.M. Porphyrins as Molecular Electronic Components of Functional Devices. *Coord Chem Rev* **2010**, *254*, 2297–2310, doi:10.1016/J.CCR.2010.05.014.

94. Huang, X.; Nakanishi, K.; Berova, N. Porphyrins and Metalloporphyrins: Versatile Circular Dichroic Reporter Groups for Structural Studies. *Chirality* **2000**, *12*, 237–255, doi:10.1002/(SICI)1520-636X(2000)12:4.

95. Vicente, M.; Smith, K. Syntheses and Functionalizations of Porphyrin Macrocycles. *Curr Org Synth* **2014**, *11*, 3–28, doi:10.2174/15701794113106660083.

96. Karolczak, J.; Kowalska, D.; Lukaszewicz, A.; Maciejewski, A.; Steer, R.P. Photophysical Studies of Porphyrins and Metalloporphyrins: Accurate Measurements of Fluorescence Spectra and Fluorescence Quantum Yields for Soret Band Excitation of Zinc Tetraphenylporphyrin. *Journal of Physical Chemistry A* **2004**, *108*, 4570–4575, doi:10.1021/JP049898V/ASSET/IMAGES/LARGE/JP049898VF00002.JPEG.

97. Senge, M.O.; Ryan, A.A.; Letchford, K.A.; MacGowan, S.A.; Mielke, T. Chlorophylls, Symmetry, Chirality, and Photosynthesis. *Symmetry 2014, Vol. 6, Pages 781-843* **2014**, *6*, 781–843, doi:10.3390/SYM6030781.

98. Gouterman, M. Spectra of Porphyrins. *J Mol Spectrosc* **1961**, *6*, 138–163, doi:10.1016/0022-2852(61)90236-3.

99. Gouterman, M.; Wagnière, G.H.; Snyder, L.C. Spectra of Porphyrins: Part II. Four Orbital Model. *J Mol Spectrosc* **1963**, *11*, 108–127, doi:10.1016/0022-2852(63)90011-0.

100. Namuangruk, S.; Sirithip, K.; Rattanatwan, R.; Keawin, T.; Kungwan, N.; Sudyodsuk, T.; Promarak, V.; Surakhot, Y.; Jungsuttiwong, S. Theoretical Investigation of the Charge-Transfer Properties in Different Meso-Linked Zinc Porphyrins for Highly Efficient Dye-Sensitized Solar Cells. *Dalton Transactions* **2014**, *43*, 9166–9176, doi:10.1039/C4DT00665H.

101. Hirao, K. Theoretical Study of the Q and B Bands of Free-Base, Magnesium, and Zinc Porphyrins, and Their Derivatives. *Journal of Physical Chemistry A* **1999**, *103*, 1894–1904, doi:10.1021/JP984807D/ASSET/IMAGES/LARGE/JP984807DF00005.JPEG.

102. Marsh, D.F.; Mink, L.M. Microscale Synthesis and Electronic Absorption Spectroscopy of Tetraphenylporphyrin H2(TPP) and Metalloporphyrins ZnII(TPP) and NiII(TPP). *J Chem Educ* **1996**, *73*, 1188–1190, doi:10.1021/ED073P1188.

103. Nasri, H. Porphyrins and Metalloporphyrins: An Overview. *DTS 2020 - IEEE International Conference on Design and Test of Integrated Micro and Nano-Systems* **2020**, doi:10.1109/DTS48731.2020.9196129.

104. Baskin, J.S.; Yu, H.Z.; Zewail, A.H. Ultrafast Dynamics of Porphyrins in the Condensed Phase: I. Free Base Tetraphenylporphyrin. *Journal of Physical Chemistry A* **2002**, *106*, 9837–9844, doi:10.1021/JP020398G/ASSET/IMAGES/LARGE/JP020398GF00008.JPEG.

105. Travagliante, G.; Gaeta, M.; Purrello, R.; D'Urso, A. Recognition and Sensing of Chiral Organic Molecules by Chiral Porphyrinoids: A Review. *Chemosensors 2021, Vol. 9, Page 204* **2021**, *9*, 204, doi:10.3390/CHEMOSENSORS9080204.

106. Bergin, E. Asymmetric Catalysis. *Annual Reports on the Progress of Chemistry - Section B* **2012**, *108*, 353–371, doi:10.1039/C2OC90003C.

107. Mihara, H.; Haruta, Y.; Sakamoto, S.; Nishino, N.; Aoyagi, H. Chiral Assembly of Porphyrins Regulated by Amphiphilic α-Helix Peptides. *http://dx.doi.org/10.1246/cl.1996.1* **2006**, 1–2, doi:10.1246/CL.1996.1.

108. Fukushima, Y. Interaction of Porphyrin Derivatives with a β-Sheet Structure of a Zwitterionic Polypeptide in Aqueous Solution. *Polymer Bulletin 2001 45:6* **2001**, *45*, 479–485, doi:10.1007/S002890170101.

109. Urbanova´*, M.; Urbanova´*, U. Bioinspired Interactions Studied by Vibrational Circular Dichroism. *Chirality* **2009**, *21*, E215–E230, doi:10.1002/CHIR.20803.

110. Luk´, L.; Luk´aš Palivec, L.; Urbanov´a, M.; Urbanov´a, U.; Volka, K. Circular Dichroism Spectroscopic Study of Non-Covalent Interactions of Poly-L-Glutamic Acid with a Porphyrin Derivative in Aqueous Solutions. *Journal of Peptide Science* **2005**, *11*, 536–545, doi:10.1002/PSC.672.

111. Occhiuto, I.G.; Samperi, M.; Trapani, M.; de Luca, G.; Romeo, A.; Pasternack, R.F.; Scolaro, L.M. Aggregates of a Cationic Porphyrin as Supramolecular Probes for Biopolymers. *J Inorg Biochem* **2015**, *153*, 361–366, doi:10.1016/J.JINORGBIO.2015.09.013.

112. Thorpe, S.L.; Snyder, G.N.; Mammana, A. Spectroscopic Study of Porphyrin Self-Assembly: Role of PH, Time, and Chiral Template. *Chirality* **2020**, *32*, 5–16, doi:10.1002/CHIR.23140.

113. Albalawi, A.; Omar Castillo Lorenzo Brancaleon Team, Dr.; Albalawi, A.; Omar Castillo Lorenzo Brancaleon Team, Dr. Irradiation of Porphyrins Modulate the Structure of Human Serum Albumin. *APS* **2022**, *2022*, N00.351.

114. Gibbs, E.J.; Tinoco, I.; Maestre, M.F.; Ellinas, P.A.; Pasternack, R.F. Self-Assembly of Porphyrins on Nucleic Acid Templates. *Biochem Biophys Res Commun* **1988**, *157*, 350–358, doi:10.1016/S0006-291X(88)80054-8.

115. Pasternack, R.F. Circular Dichroism and the Interactions of Water Soluble Porphyrins with DNA - A Minireview. *Chirality* **2003**, *15*, 329–332, doi:10.1002/CHIR.10206.

116. Qin, T.; Liu, K.; Song, D.; Yang, C.; Zhao, H.; Su, H. Binding Interactions of Zinc Cationic Porphyrin with Duplex DNA: From B-DNA to Z-DNA. *International Journal of Molecular Sciences 2018, Vol. 19, Page 1071* **2018**, *19*, 1071, doi:10.3390/IJMS19041071.

117. Gangemi, C.M.A.; D’Agostino, B.; Randazzo, R.; Gaeta, M.; Fragalà, M.E.; Purrello, R.; D’Urso, A. Interaction of Spermine Derivative Porphyrin with DNA. *https://doi.org/10.1142/S1088424618500645* **2018**, *22*, 581–587, doi:10.1142/S1088424618500645.

118. Sabharwal, N.C.; Chen, J.; Lee, J.H.J.; Gangemi, C.M.A.; D’urso, A.; Yatsunyk, L.A. Interactions Between Spermine-Derivatized Tentacle Porphyrins and The Human Telomeric DNA G-Quadruplex. *Int J Mol Sci* **2018**, *19*, 3686, doi:10.3390/IJMS19113686.

119. Zhang, L.M.; Cui, Y.X.; Zhu, L.N.; Chu, J.Q.; Kong, D.M. Cationic Porphyrins with Large Side Arm Substituents as Resonance Light Scattering Ratiometric Probes for Specific Recognition of Nucleic Acid G-Quadruplexes. *Nucleic Acids Res* **2019**, *47*, 2727–2738, doi:10.1093/NAR/GKZ064.

120. Occhiuto, I.G.; Samperi, M.; Trapani, M.; De Luca, G.; Romeo, A.; Pasternack, R.F.; Scolaro, L.M. Aggregates of a Cationic Porphyrin as Supramolecular Probes for Biopolymers. *J Inorg Biochem* **2015**, *153*, 361–366, doi:10.1016/j.jinorgbio.2015.09.013.

121. Gaeta, M.; Raciti, D.; Randazzo, R.; Gangemi, C.M.A.; Raudino, A.; D'Urso, A.; Fragalà, M.E.; Purrello, R. Chirality Enhancement of Porphyrin Supramolecular Assembly Driven by a Template Preorganization Effect. *Angewandte Chemie - International Edition* **2018**, *57*, 10656–10660, doi:10.1002/anie.201806192.

122. Fiel, R.J.; Howard, J.C.; Mark, E.H.; Gupta, N.D. Interaction of DNA with a Porphyrin Ligand: Evidence for Intercalation. *Nucleic Acids Res* **1979**, *6*, 3093, doi:10.1093/NAR/6.9.3093.

123. Fiel, R.J.; Munson, B.R. Binding of Meso-Tetra (4-N-Methylpyridyl) Porphine to DNA. *Nucleic Acids Res* **1980**, *8*, 2835, doi:10.1093/NAR/8.12.2835.

124. Pasternack, R.F.; Gibbs, E.J.; Villafranca, J.J. Interactions of Porphyrins with Nucleic Acids. *Biochemistry* **1983**, *22*, 2406–2414, doi:10.1021/BI00279A016.

125. Scolaro, L.M.; Romeo, A.; Pasternack, R.F. Tuning Porphyrin/DNA Supramolecular Assemblies by Competitive Binding. *J Am Chem Soc* **2004**, *126*, 7178–7179, doi:10.1021/JA049669M/SUPPL_FILE/JA049669MSI20040426_065220.PDF.

126. D'Urso, A.; Randazzo, R.; Rizzo, V.; Gangemi, C.M.A.; Romanucci, V.; Zarrelli, A.; Tomaselli, G.A.; Milardi, D.; Borbone, N.; Purrello, R.; et al. Stabilization vs. Destabilization of G-Quadruplex Superstructures: The Role of the Porphyrin Derivative Having Spermine Arms. *Phys Chem Chem Phys* **2017**, *19*, 17404–17410, doi:10.1039/C7CP02816D.

127. Fiel, R.J. Porphyrin- Nucleic Acid Interactions: A Review. *J Biomol Struct Dyn* **1989**, *6*, 1259–1274, doi:10.1080/07391102.1989.10506549.

128. Lebedeva, N.S.; Yurina, E.S.; Gubarev, Y.A.; Syrbu, S.A. Interactions of Tetracationic Porphyrins with DNA and Their Effects on DNA Cleavage. *Spectrochim Acta A Mol Biomol Spectrosc* **2018**, *199*, 235–241, doi:10.1016/J.SAA.2018.03.066.

129. Mathew, D.; Sujatha, S. Interactions of Porphyrins with DNA: A Review Focusing Recent Advances in Chemical Modifications on Porphyrins as Artificial Nucleases. *J Inorg Biochem* **2021**, *219*, 111434, doi:10.1016/J.JINORGBIO.2021.111434.

130. D'Urso, A.; Petrovic, A.G.; Fragalà, M.E.; Tamargo, M.A.; Ellestad, G.A.; Purrello, R.; Berova, N. Exploring Nucleic Acid Conformations by Employment of Porphyrin Non-Covalent and Covalent Probes and Chiroptical Analysis. *DNA in Supramolecular Chemistry and Nanotechnology* **2017**, 172–208, doi:10.1002/9781118696880.CH3.3.

131. Dixon, I.M.; Lopez, F.; Tejera, A.M.; Estève, J.P.; Blasco, M.A.; Pratviel, G.; Meunier, B. A G-Quadruplex Ligand with 10000-Fold Selectivity over Duplex DNA. *J Am Chem Soc* **2007**, *129*, 1502–1503, doi:10.1021/JA065591T/SUPPL_FILE/JA065591TSI20061221_082021.PDF.

132. Nicoludis, J.M.; Barrett, S.P.; Mergny, J.L.; Yatsunyk, L.A. Interaction of Human Telomeric DNA with N-Methyl Mesoporphyrin IX. *Nucleic Acids Res* **2012**, *40*, 5432, doi:10.1093/NAR/GKS152.

133. Lubitz, I.; Borovok, N.; Kotlyar, A. Interaction of Monomolecular G4-DNA Nanowires with TMPyP: Evidence for Intercalation. *Biochemistry* **2007**, *46*, 12925–12929, doi:10.1021/BI701301U/ASSET/IMAGES/LARGE/BI701301UF00004.JPEG.

134. Hermann, T.; Patel, D.J. Stitching Together RNA Tertiary Architectures. *J Mol Biol* **1999**, *294*, 829–849, doi:10.1006/JMBI.1999.3312.

135. Hermann, T.; Patel, D.J. RNA Bulges as Architectural and Recognition Motifs. *Structure* **2000**, *8*, R47–R54, doi:10.1016/S0969-2126(00)00110-6.

136. Ivanov, M.; Sizov, V.; Kudrev, A. Thermal Unwinding of Polyadenylic·Polyuridylic Acid Complex with TMPyP4 Porphyrin in Aqueous Solutions. *J Mol Struct* **2020**, *1202*, 127365, doi:10.1016/J.MOLSTRUC.2019.127365.

137. Ferino, A.; Nicoletto, G.; D'Este, F.; Zorzet, S.; Lago, S.; Richter, S.N.; Tikhomirov, A.; Shchekotikhin, A.; Xodo, L.E. Photodynamic Therapy for Ras-Driven Cancers: Targeting G-Quadruplex RNA Structures with Bifunctional Alkyl-Modified Porphyrins. *J Med Chem* **2020**, *63*, 1245–1260, doi:10.1021/ACS.JMEDCHEM.9B01577/SUPPL_FILE/JM9B01577_SI_002.CSV.

138. Uno, T.; Aoki, K.; Shikimi, T.; Hiranuma, Y.; Tomisugi, Y.; Ishikawa, Y. Copper Insertion Facilitates Water-Soluble Porphyrin Binding to RA·rU and RA·dT Base Pairs in Duplex RNA and RNA·DNA Hybrids. *Biochemistry* **2002**, *41*, 13059–13066, doi:10.1021/BI026139Z/SUPPL_FILE/BI026139Z_S1.PDF.

139. Choi, J.K.; D'Urso, A.; Balaz, M. Chiroptical Properties of Anionic and Cationic Porphyrins and Metalloporphyrins in Complex with Left-Handed Z-DNA and Right-Handed B-DNA. *J Inorg Biochem* **2013**, *127*, 1–6, doi:10.1016/j.jinorgbio.2013.05.018.

140. D'Urso, A.; Mammana, A.; Balaz, M.; Holmes, A.E.; Berova, N.; Lauceri, R.; Purrello, R. Interactions of a Tetraanionic Porphyrin with DNA: From a Z-DNA Sensor to a Versatile Supramolecular Device. *J Am Chem Soc* **2009**, *131*, 2046–2047, doi:10.1021/JA808099U/SUPPL_FILE/JA808099U_SI_001.PDF.

141. Balaz, M.; Holmes, A.E.; Benedetti, M.; Rodriguez, P.C.; Berova, N.; Nakanishi, K.; Proni, G. Synthesis and Circular Dichroism of Tetraarylporphyrin-Oligonucleotide Conjugates. *J Am Chem Soc* **2005**, *127*, 4172–4173, doi:10.1021/JA043373Z/SUPPL_FILE/JA043373ZSI20050207_015933.PDF.

142. D'Urso, A.; Nardis, S.; Pomarico, G.; Fragalà, M.E.; Paolesse, R.; Purrello, R. Interaction of Tricationic Corroles with Single/Double Helix of Homopolymeric Nucleic Acids and DNA. *J Am Chem Soc* **2013**, *135*, 8632–8638, doi:10.1021/JA4023539/SUPPL_FILE/JA4023539_SI_001.PDF.

143. Garzon, R.; Calin, G.A.; Croce, C.M. MicroRNAs in Cancer. *Annu Rev Med* **2009**, *60*, 167–179, doi:10.1146/annurev.med.59.053006.104707.

144. Schickel, R.; Boyerinas, B.; Park, S.M.; Peter, M.E. MicroRNAs: Key Players in the Immune System, Differentiation, Tumorigenesis and Cell Death. *Oncogene* **2008**, *27*, 5959–5974, doi:10.1038/onc.2008.274.

145. Ambros, V. The Functions of Animal MicroRNAs. *Nature* **2004**, *431*, 350–355, doi:10.1038/nature02871.

146. Guay, C.; Roggli, E.; Nesca, V.; Jacovetti, C.; Regazzi, R. Diabetes Mellitus, a MicroRNA-Related Disease? *Translational Research* **2011**, *157*, 253–264, doi:10.1016/j.trsl.2011.01.009.

147. Gupta, S.; Verma, S.; Mantri, S.; Berman, N.E.; Sandhir, R. Targeting MicroRNAs in Prevention and Treatment of Neurodegenerative Disorders. *Drug Dev Res* **2015**, *76*, 397–418, doi:10.1002/ddr.21277.

148. Croce, C.M. MiRNAs in the Spotlight: Understanding Cancer Gene Dependency. *Nature Medicine 2011 17:8* **2011**, *17*, 935–936, doi:10.1038/nm0811-935.

149. Rupaimoole, R.; Slack, F.J. MicroRNA Therapeutics: Towards a New Era for the Management of Cancer and Other Diseases. *Nat Rev Drug Discov* **2017**, *16*, 203–221, doi:10.1038/nrd.2016.246.

150. Lee, Y.; Kim, M.; Han, J.; Yeom, K.H.; Lee, S.; Baek, S.H.; Kim, V.N. MicroRNA Genes Are Transcribed by RNA Polymerase II. *EMBO Journal* **2004**, *23*, 4051–4060, doi:10.1038/sj.emboj.7600385.

151. Lee, Y.; Ahn, C.; Han, J.; Choi, H.; Kim, J.; Yim, J.; Lee, J.; Provost, P.; Rådmark, O.; Kim, S.; et al. The Nuclear RNase III Drosha Initiates MicroRNA Processing. *Nature* **2003**, *425*, 415–419, doi:10.1038/nature01957.

152. Yoda, M.; Kawamata, T.; Paroo, Z.; Ye, X.; Iwasaki, S.; Liu, Q.; Tomari, Y. ATP-Dependent Human RISC Assembly Pathways. *Nat Struct Mol Biol* **2010**, *17*, 17–24, doi:10.1038/nsmb.1733.

153. Meijer, H.A.; Smith, E.M.; Bushell, M. Regulation of MiRNA Strand Selection: Follow the Leader? *Biochem Soc Trans* **2014**, *42*, 1135–1140, doi:10.1042/BST20140142.

154. Khvorova, A.; Reynolds, A.; Jayasena, S.D. Functional SiRNAs and MiRNAs Exhibit Strand Bias. *Cell* **2003**, *115*, 209–216, doi:10.1016/S0092-8674(03)00801-8.

155. Ha, M.; Kim, V.N. Regulation of MicroRNA Biogenesis. *Nat Rev Mol Cell Biol* **2014**, *15*, 509–524, doi:10.1038/nrm3838.

156. Ellwanger, D.C.; Büttner, F.A.; Mewes, H.W.; Stümpflen, V. The Sufficient Minimal Set of MiRNA Seed Types. *Bioinformatics* **2011**, *27*, 1346–1350, doi:10.1093/bioinformatics/btr149.

157. Grimson, A.; Farh, K.K.H.; Johnston, W.K.; Garrett-Engele, P.; Lim, L.P.; Bartel, D.P. MicroRNA Targeting Specificity in Mammals: Determinants beyond Seed Pairing. *Mol Cell* **2007**, *27*, 91–105, doi:10.1016/j.molcel.2007.06.017.

158. Krek, A.; Grün, D.; Poy, M.N.; Wolf, R.; Rosenberg, L.; Epstein, E.J.; MacMenamin, P.; Da Piedade, I.; Gunsalus, K.C.; Stoffel, M.; et al. Combinatorial MicroRNA Target Predictions. *Nat Genet* **2005**, *37*, 495–500, doi:10.1038/ng1536.

159. Friedman, R.C.; Farh, K.K.H.; Burge, C.B.; Bartel, D.P. Most Mammalian MRNAs Are Conserved Targets of MicroRNAs. *Genome Res* **2009**, *19*, 92–105, doi:10.1101/gr.082701.108.

160. Esquela-Kerscher, A.; Slack, F.J. Oncomirs - MicroRNAs with a Role in Cancer. *Nat Rev Cancer* **2006**, *6*, 259–269, doi:10.1038/nrc1840.

161. Melo, S.A.; Sugimoto, H.; O'Connell, J.T.; Kato, N.; Villanueva, A.; Vidal, A.; Qiu, L.; Vitkin, E.; Perelman, L.T.; Melo, C.A.; et al. Cancer Exosomes Perform Cell-Independent MicroRNA Biogenesis and Promote Tumorigenesis. *Cancer Cell* **2014**, *26*, 707–721, doi:10.1016/j.ccell.2014.09.005.

162. Calin, G.A.; Croce, C.M. MicroRNA Signatures in Human Cancers. *Nat Rev Cancer* **2006**, *6*, 857–866, doi:10.1038/nrc1997.

163. Forterre, A.; Komuro, H.; Aminova, S.; Harada, M. A Comprehensive Review of Cancer MicroRNA Therapeutic Delivery Strategies. *Cancers (Basel)* **2020**, *12*, 1–21, doi:10.3390/cancers12071852.

164. Hosseinahli, N.; Aghapour, M.; Duijf, P.H.G.; Baradaran, B. Treating Cancer with MicroRNA Replacement Therapy: A Literature Review. *J Cell Physiol* **2018**, *233*, 5574–5588, doi:10.1002/jcp.26514.

165. Broderick, J.A.; Zamore, P.D. MicroRNA Therapeutics. *Gene Ther* **2011**, *18*, 1104–1110, doi:10.1038/gt.2011.50.

166. Treiber, T.; Treiber, N.; Meister, G. Regulation of MicroRNA Biogenesis and Its Crosstalk with Other Cellular Pathways. *Nat Rev Mol Cell Biol* **2019**, *20*, 5–20, doi:10.1038/s41580-018-0059-1.

167. Duttagupta, R.; Jiang, R.; Gollub, J.; Getts, R.C.; Jones, K.W. Impact of Cellular MiRNAs on Circulating MiRNA Biomarker Signatures. *PLoS One* **2011**, *6*, e20769, doi:10.1371/JOURNAL.PONE.0020769.

168. Zhang, Y.; Liu, D.; Chen, X.; Li, J.; Li, L.; Bian, Z.; Sun, F.; Lu, J.; Yin, Y.; Cai, X.; et al. Secreted Monocytic MiR-150 Enhances Targeted Endothelial Cell Migration. *Mol Cell* **2010**, *39*, 133–144, doi:10.1016/j.molcel.2010.06.010.

169. Kim, Y.-K. Extracellular MicroRNAs as Biomarkers in Human Disease. *Chonnam Med J* **2015**, *51*, 51, doi:10.4068/cmj.2015.51.2.51.

170. Cui, M.; Wang, H.; Yao, X.; Zhang, D.; Xie, Y.; Cui, R.; Zhang, X. Circulating MicroRNAs in Cancer: Potential and Challenge. *Front Genet* **2019**, *10*, 626, doi:10.3389/fgene.2019.00626.

171. Rosenfeld, N.; Aharonov, R.; Meiri, E.; Rosenwald, S.; Spector, Y.; Zepeniuk, M.; Benjamin, H.; Shabes, N.; Tabak, S.; Levy, A.; et al. MicroRNAs Accurately Identify Cancer Tissue Origin. *Nat Biotechnol* **2008**, *26*, 462–469, doi:10.1038/nbt1392.

172. Janas, M.M.; Wang, B.; Harris, A.S.; Aguiar, M.; Shaffer, J.M.; Subrahmanyam, Y.V.B.K.; Behlke, M.A.; Wucherpfennig, K.W.; Gygi, S.P.; Gagnon, E.; et al. Alternative RISC Assembly: Binding and Repression of MicroRNA-MRNA Duplexes by Human Ago Proteins. *Rna* **2012**, *18*, 2041–2055, doi:10.1261/rna.035675.112.

173. Belter, A.; Gudanis, D.; Rolle, K.; Piwecka, M.; Gdaniec, Z.; Naskre¸t-Barciszewska, M.Z.; Barciszewski, J. Mature MiRNAs Form Secondary Structure, Which Suggests Their Function beyond RISC. *PLoS One* **2014**, *9*, e113848, doi:10.1371/journal.pone.0113848.

174. Gangemi, C.M.A.; Alaimo, S.; Pulvirenti, A.; García-Viñuales, S.; Milardi, D.; Falanga, A.P.; Fragalà, M.E.; Oliviero, G.; Piccialli, G.; Borbone, N.; et al. Endogenous and Artificial MiRNAs Explore a Rich Variety of Conformations: A Potential Relationship between Secondary Structure and Biological Functionality. *Sci Rep* **2020**, *10*, 1–11, doi:10.1038/s41598-019-57289-8.

175. Laganà, A.; Acunzo, M.; Romano, G.; Pulvirenti, A.; Veneziano, D.; Cascione, L.; Giugno, R.; Gasparini, P.; Shasha, D.; Ferro, A.; et al. MiR-Synth: A Computational Resource for the Design of Multi-Site Multi-Target Synthetic MiRNAs. *Nucleic Acids Res* **2014**, *42*, 5416–5425, doi:10.1093/nar/gku202.

176. Maiti, M.; Nauwelaerts, K.; Lescrinier, E.; Schuit, F.C.; Herdewijn, P. Self-Complementary Sequence Context in Mature MiRNAs. *Biochem Biophys Res Commun* **2010**, *392*, 572–576, doi:10.1016/J.BBRC.2010.01.072.

177. Jebb, D.; Huang, Z.; Pippel, M.; Hughes, G.M.; Lavrichenko, K.; Devanna, P.; Winkler, S.; Jermiin, L.S.; Skirmuntt, E.C.; Katzourakis, A.; et al. Six Reference-Quality Genomes Reveal Evolution of Bat Adaptations. *Nature* **2020**, *583*, 578–584, doi:10.1038/s41586-020-2486-3.

178. Zhang, Z.; Zhang, L.; Wang, B.; Wei, R.; Wang, Y.; Wan, J.; Zhang, C.; Zhao, L.; Zhu, X.; Zhang, Y.; et al. MiR-337–3p Suppresses Proliferation of Epithelial Ovarian Cancer by Targeting PIK3CA and PIK3CB. *Cancer Lett* **2020**, *469*, 54–67, doi:10.1016/j.canlet.2019.10.021.

179. Wang, Z.; Wang, J.; Yang, Y.; Hao, B.; Wang, R.; Li, Y.; Wu, Q. Loss of Has-MiR-337-3p Expression Is Associated with Lymph Node Metastasis of Human Gastric Cancer. *Journal of Experimental and Clinical Cancer Research* **2013**, *32*, 1–9, doi:10.1186/1756-9966-32-76.

180. Parafioriti, A.; Cifola, I.; Gissi, C.; Pinatel, E.; Vilardo, L.; Armiraglio, E.; Di Bernardo, A.; Daolio, P.A.; Felsani, A.; D'Agnano, I.; et al. Expression Profiling of MicroRNAs and IsomiRs in Conventional Central Chondrosarcoma. *Cell Death Discov* **2020**, *6*, doi:10.1038/S41420-020-0282-3.

181. Dill, H.; Linder, B.; Fehr, A.; Fischer, U. Intronic MiR-26b Controls Neuronal Differentiation by Repressing Its Host Transcript, Ctdsp2. *Genes Dev* **2012**, *26*, 25–30, doi:10.1101/GAD.177774.111.

182. Li, J.; Li, X.; Kong, X.; Luo, Q.; Zhang, J.; Fang, L. MiRNA-26b Inhibits Cellular Proliferation by Targeting CDK8 in Breast Cancer. *Int J Clin Exp Med* **2014**, *7*, 558.

183. Duan, G.; Ren, C.; Zhang, Y.; Feng, S. MicroRNA-26b Inhibits Metastasis of Osteosarcoma via Targeting CTGF and Smad1. *Tumour Biol* **2015**, *36*, 6201–6209, doi:10.1007/S13277-015-3305-6.

184. Fan, F.; Lu, J.; Yu, W.; Zhang, Y.; Xu, S.; Pang, L.; Zhu, B. MicroRNA-26b-5p Regulates Cell Proliferation, Invasion and Metastasis in Human Intrahepatic Cholangiocarcinoma by Targeting S100A7. *Oncol Lett* **2018**, *15*, 386–392, doi:10.3892/OL.2017.7331.

185. Wu, K.; Mu, X.Y.; Jiang, J.T.; Tan, M.Y.; Wang, R.J.; Zhou, W.J.; Wang, X.; He, Y.Y.; Li, M.Q.; Liu, Z.H. MiRNA-26a-5p and MiR-26b-5p Inhibit the Proliferation of Bladder Cancer Cells by Regulating PDCD10. *Oncol Rep* **2018**, *40*, 3523–3532, doi:10.3892/OR.2018.6734.

186. Miyamoto, K.; Seki, N.; Matsushita, R.; Yonemori, M.; Yoshino, H.; Nakagawa, M.; Enokida, H. Tumour-Suppressive MiRNA-26a-5p and MiR-26b-5p Inhibit Cell Aggressiveness by Regulating PLOD2 in Bladder Cancer. *Br J Cancer* **2016**, *115*, 354, doi:10.1038/BJC.2016.179.

187. Wang, D.; Liu, C.; Wang, Y.; Wang, W.; Wang, K.; Wu, X.; Li, Z.; Zhao, C.; Li, L.; Peng, L. Impact of MiR-26b on Cardiomyocyte Differentiation in P19 Cells through Regulating Canonical/Non-Canonical Wnt Signalling. *Cell Prolif* **2017**, *50*, n/a-n/a, doi:10.1111/CPR.12371.

188. Absalon, S.; Kochanek, D.M.; Raghavan, V.; Krichevsky, A.M. MiR-26b, Upregulated in Alzheimer's Disease, Activates Cell Cycle Entry, Tau-Phosphorylation, and Apoptosis in Postmitotic Neurons. *J Neurosci* **2013**, *33*, 14645–14659, doi:10.1523/JNEUROSCI.1327-13.2013.

189. Thomas, K.E.; McMillin, D.R. Competitive Binding Studies of H2T4 with DNA Hairpins (H2T4 = Meso-Tetrakis(4-(N-Methylpyridiniumyl))Porphyrin). *Journal of Physical Chemistry B* **2001**, *105*, 12628–12633, doi:10.1021/jp011860k.

190. Pasternack, R.F.; Gibbs, E.J.; Villafranca, J.J. Interactions of Porphyrins with Nucleic Acids. *Biochemistry* **1983**, *22*, 2406–2414, doi:10.1021/BI00279A016.

191. Sehlstedt, U.; Nordén, B.; Kim, S.K.; Carter, P.; Goodisman, J.; Dabrowiak, J.C.; Vollano, J.F. Interaction of Cationic Porphyrins with DNA. *Biochemistry* **1994**, *33*, 417–426, doi:10.1021/BI00168A005.

192. Gangemi, C.M.A.; Randazzo, R.; Fragalà, M.E.; Tomaselli, G.A.; Ballistreri, F.P.; Pappalardo, A.; Toscano, R.M.; Sfrazzetto, G.T.; Purrello, R.; D'Urso, A. Hierarchically Controlled Protonation/Aggregation of a Porphyrin–Spermine Derivative. *New Journal of Chemistry* **2015**, *39*, 6722–6725, doi:10.1039/C5NJ01264C.

193. Holbert, C.E.; Foley, J.R.; Stewart, T.M.; Casero, R.A. Expanded Potential of the Polyamine Analogue SBP-101 (Diethyl Dihydroxyhomospermine) as a Modulator of Polyamine Metabolism and Cancer Therapeutic. *Int J Mol Sci* **2022**, *23*, 6798, doi:10.3390/IJMS23126798/S1.

194. D'Urso, A.; Randazzo, R.; Rizzo, V.; Gangemi, C.M.A.; Romanucci, V.; Zarrelli, A.; Tomaselli, G.A.; Milardi, D.; Borbone, N.; Purrello, R.; et al. Stabilization vs. Destabilization of G-Quadruplex Superstructures: The Role of the Porphyrin Derivative Having Spermine Arms. *Physical Chemistry Chemical Physics* **2017**, *19*, 17404–17410, doi:10.1039/C7CP02816D.

195. Gangemi, C.M.A.; D'Urso, A.; Tomaselli, G.A.; Berova, N.; Purrello, R. A Novel Porphyrin-Based Molecular Probe ZnTCPPSpm4 with Catalytic, Stabilizing and

Chiroptical Diagnostic Power towards DNA B-Z Transition. *J Inorg Biochem* **2017**, *173*, 141–143, doi:10.1016/J.JINORGBIO.2017.05.008.

196. Herrmann, O.; Mehdi, S.H.; Corsini, D.A. Heterogeneous Metal-Insertion: A Novel Reaction with Porphyrins. *Can J Chem* **1978**, *56*, 1084–1087, doi:10.1139/V78-184.
197. D'urso, A.; Fragalà, M.E.; Purrello, R. From Self-Assembly to Noncovalent Synthesis of Programmable Porphyrins ' Arrays in Aqueous Solution. *Chemical Communications* **2012**, *48*, 8165–8176, doi:10.1039/C2CC31856C.
198. Vergeldt, F.J.; Koehorst, R.B.M.; Van Hoek, A.; Schaafsma, T.J. Intramolecular Interactions in the Ground and Excited State of Tetrakis(N-Methylpyridyl)Porphyrins. *Journal of Physical Chemistry* **1995**, *99*, 4397–4405, doi:10.1021/J100013A007.
199. Hembury, G.A.; Borovkov, V. V.; Inoue, Y. Chirality-Sensing Supramolecular Systems. *Chem Rev* **2008**, *108*, 1–73, doi:10.1021/cr050005k.
200. Pasternack, R.F. Circular Dichroism and the Interactions of Water Soluble Porphyrins with DNA—A Minireview. *Chirality* **2003**, *15*, 329–332, doi:10.1002/CHIR.10206.
201. Choi, J.K.; D'Urso, A.; Balaz, M. Chiroptical Properties of Anionic and Cationic Porphyrins and Metalloporphyrins in Complex with Left-Handed Z-DNA and Right-Handed B-DNA. *J Inorg Biochem* **2013**, *127*, 1–6, doi:10.1016/J.JINORGBIO.2013.05.018.
202. Pasternack, R.F.; Gibbs, E.J.; Villafranca, J.J. Interactions of Porphyrins with Nucleic Acids. *Biochemistry* **1983**, *22*, 2406–2414, doi:10.1021/BI00279A016/ASSET/BI00279A016.FP.PNG_V03.
203. Peacocke, A.R.; Skerrett, J.N.H. The Interaction of Aminoacridines with Nucleic Acids. *Transactions of the Faraday Society* **1956**, *52*, 261–279, doi:10.1039/TF9565200261.

204. Gaeta, M.; Barcellona, M.; Purrello, R.; Fragalà, M.E.; D’Urso, A. Hybrid Porphyrin/DOPA-Melanin Film as Self-Assembled Material and Smart Device for Dye-Pollutant Removal in Water. *Chemical Engineering Journal* **2022**, *433*, 133262, doi:10.1016/J.CEJ.2021.133262.

205. Kasha, M.; Rawls, H.R.; El-Bayoumi, M.A. The Exciton Model in Molecular Spectroscopy. *Pure and Applied Chemistry* **1965**, *11*, 371–392, doi:10.1351/PAC196511030371.

206. Ryazanova, O.; Voloshin, I.; Dubey, I.; Dubey, L.; Zozulya, V. Fluorescent Studies on Cooperative Binding of Cationic Pheophorbide-a Derivative to Polyphosphate. *Ann N Y Acad Sci* **2008**, *1130*, 293–299, doi:10.1196/ANNALS.1430.033.

207. Zozulya, V.N.; Ryazanova, O.A.; Voloshin, I.M.; Ilchenko, M.M.; Dubey, I.Y.; Glamazda, A.Y.; Karachevtsev, V.A. Self-Assemblies of Tricationic Porphyrin on Inorganic Polyphosphate. *Biophys Chem* **2014**, *185*, 39–46, doi:10.1016/J.BPC.2013.11.006.

208. Guo, Y.; Huang, S.; Sun, H.; Wang, Z.; Shao, Y.; Li, L.; Li, Z.; Song, F. Tuning the Aqueous Self-Assembly of Porphyrins by Varying the Number of Cationic Side Chains. *J Mater Chem B* **2022**, *10*, 5968–5975, doi:10.1039/D2TB00720G.

209. Liquori, A.M.; Costantino, L.; Crescenzi, V.; Elia, V.; Giglio, E.; Puliti, R.; De Santis Savino, M.; Vitagliano, V. Complexes between DNA and Polyamines: A Molecular Model. *J Mol Biol* **1967**, *24*, 113–122, doi:10.1016/0022-2836(67)90094-0.

210. Kimura, E.; Yatsunami, T.; Kodama, M. Macromonocyclic Polyamines as Biological Polyanion Complexons. 2. Ion-Pair Association with Phosphate and Nucleotides. *J Am Chem Soc* **1982**, *104*, 3182–3187, doi:10.1021/JA00375A042/SUPPL_FILE/JA00375A042_SI_001.PDF.

211. Wirth, W.; Blotevogel-Baltronat, J.; Kleinkes, U.; Sheldrick, W.S. Interaction of (Amine)M(II) Complexes (Amine=dien, En; M=Pd, Pt) with Purine Nucleoside 2′-, 3′- and 5′-Monophosphates—the Role of the Phosphate Site for Specific Metal Fragment–Nucleotide Recognition by Macrochelation. *Inorganica Chim Acta* **2002**, *339*, 14–26, doi:10.1016/S0020-1693(02)00921-0.

212. Lyu, K.; Chow, E.Y.C.; Mou, X.; Chan, T.F.; Kwok, C.K. RNA G-Quadruplexes (RG4s): Genomics and Biological Functions. *Nucleic Acids Res* **2021**, *49*, 5426–5450, doi:10.1093/NAR/GKAB187.

213. Huppert, J.L.; Balasubramanian, S. Prevalence of Quadruplexes in the Human Genome. *Nucleic Acids Res* **2005**, *33*, 2908–2916, doi:10.1093/NAR/GKI609.

214. Bhattacharyya, D.; Arachchilage, G.M.; Basu, S. Metal Cations in G-Quadruplex Folding and Stability. *Front Chem* **2016**, *4*, 207258, doi:10.3389/FCHEM.2016.00038/BIBTEX.

215. Edwards, D.N.; Machwe, A.; Wang, Z.; Orren, D.K. Intramolecular Telomeric G-Quadruplexes Dramatically Inhibit DNA Synthesis by Replicative and Translesion Polymerases, Revealing Their Potential to Lead to Genetic Change. *PLoS One* **2014**, *9*, e80664, doi:10.1371/JOURNAL.PONE.0080664.

216. Huang, H.; Suslov, N.B.; Li, N.S.; Shelke, S.A.; Evans, M.E.; Koldobskaya, Y.; Rice, P.A.; Piccirilli, J.A. A G-Quadruplex–Containing RNA Activates Fluorescence in a GFP-like Fluorophore. *Nature Chemical Biology 2014 10:8* **2014**, *10*, 686–691, doi:10.1038/nchembio.1561.

217. Warner, K.D.; Chen, M.C.; Song, W.; Strack, R.L.; Thorn, A.; Jaffrey, S.R.; Ferré-D'Amaré, A.R. Structural Basis for Activity of Highly Efficient RNA Mimics of Green Fluorescent Protein. *Nature Structural & Molecular Biology 2014 21:8* **2014**, *21*, 658–663, doi:10.1038/nsmb.2865.

218. Xiao, C. Da; Shibata, T.; Yamamoto, Y.; Xu, Y. An Intramolecular Antiparallel G-Quadruplex Formed by Human Telomere RNA. *Chemical Communications* **2018**, *54*, 3944–3946, doi:10.1039/C8CC01427B.

219. Tang, C.F.; Shafer, R.H. Engineering the Quadruplex Fold: Nucleoside Conformation Determines Both Folding Topology and Molecularity in Guanine Quadruplexes. *J Am Chem Soc* **2006**, *128*, 5966–5973, doi:10.1021/JA0603958/ASSET/IMAGES/LARGE/JA0603958F00005.JPEG.

220. Chen, L.; Dickerhoff, J.; Sakai, S.; Yang, D. DNA G-Quadruplex in Human Telomeres and Oncogene Promoters: Structures, Functions, and Small Molecule Targeting. *Acc Chem Res* **2022**, *55*, 2628–2646, doi:10.1021/ACS.ACCOUNTS.2C00337/ASSET/IMAGES/LARGE/AR2C00337_0012.JPEG.

221. Arora, A.; Maiti, S. Differential Biophysical Behavior of Human Telomeric RNA and DNA Quadruplex. *J Phys Chem B* **2009**, *113*, 10515–10520, doi:10.1021/JP810638N.

222. Joachimi, A.; Benz, A.; Hartig, J.S. A Comparison of DNA and RNA Quadruplex Structures and Stabilities. *Bioorg Med Chem* **2009**, *17*, 6811–6815, doi:10.1016/J.BMC.2009.08.043.

223. Tang, C.F.; Shafer, R.H. Engineering the Quadruplex Fold: Nucleoside Conformation Determines Both Folding Topology and Molecularity in Guanine Quadruplexes. *J Am Chem Soc* **2006**, *128*, 5966–5973, doi:10.1021/JA0603958.

224. Miyoshi, D.; Nakao, A.; Sugimoto, N. Molecular Crowding Regulates the Structural Switch of the DNA G-Quadruplex. *Biochemistry* **2002**, *41*, 15017–15024, doi:10.1021/BI020412F.

225. Miyoshi, D.; Karimata, H.; Sugimoto, N. Hydration Regulates Thermodynamics of G-Quadruplex Formation under Molecular Crowding Conditions. *J Am Chem Soc* **2006**, *128*, 7957–7963, doi:10.1021/JA061267M.

226. Agarwala, P.; Pandey, S.; Maiti, S. The Tale of RNA G-Quadruplex. *Org Biomol Chem* **2015**, *13*, 5570–5585, doi:10.1039/C4OB02681K.

227. Salazar, M.; Thompson, B.D.; Kerwin, S.M.; Hurley, L.H. Thermally Induced DNA·RNA Hybrid to G-Quadruplex Transitions: Possible Implications for Telomere Synthesis by Telomerase†. *Biochemistry* **1996**, *35*, 16110–16115, doi:10.1021/BI961442J.

228. Pinto, A.R.; Li, H.; Nicholls, C.; Liu, J.P. Telomere Protein Complexes and Interactions with Telomerase in Telomere Maintenance. *Frontiers in Bioscience* **2011**, *16*, 187–207, doi:10.2741/3683/PDF.

229. Monchaud, D.; Teulade-Fichou, M.P. A Hitchhiker's Guide to G-Quadruplex Ligands. *Org Biomol Chem* **2008**, *6*, 627–636, doi:10.1039/B714772B.

230. Azzalin, C.M.; Reichenbach, P.; Khoriauli, L.; Giulotto, E.; Lingner, J. Telomeric Repeat Containing RNA and RNA Surveillance Factors at Mammalian Chromosome Ends. *Science* **2007**, *318*, 798–801, doi:10.1126/SCIENCE.1147182.

231. Pan, B.; Xiong, Y.; Shi, K.; Sundaralingam, M. An Eight-Stranded Helical Fragment in RNA Crystal Structure: Implications for Tetraplex Interaction. *Structure* **2003**, *11*, 825–831, doi:10.1016/S0969-2126(03)00108-4.

232. Pan, B.; Xiong, Y.; Shi, K.; Deng, J.; Sundaralingam, M. Crystal Structure of an RNA Purine-Rich Tetraplex Containing Adenine Tetrads: Implications for Specific Binding in RNA Tetraplexes. *Structure* **2003**, *11*, 815–823, doi:10.1016/S0969-2126(03)00107-2.

233. Xu, Y.; Suzuki, Y.; Ito, K.; Komiyama, M. Telomeric Repeat-Containing RNA Structure in Living Cells. *Proc Natl Acad Sci U S A* **2010**, *107*, 14579–14584, doi:10.1073/PNAS.1001177107.

234. De Silanes, I.L.; D'Alcontres, M.S.; Blasco, M.A. TERRA Transcripts Are Bound by a Complex Array of RNA-Binding Proteins. *Nat Commun* **2010**, *1*, doi:10.1038/NCOMMS1032.

235. Biffi, G.; Tannahill, D.; Balasubramanian, S. An Intramolecular G-Quadruplex Structure Is Required for Binding of Telomeric Repeat-Containing RNA to the Telomeric Protein TRF2. *J Am Chem Soc* **2012**, *134*, 11974–11976, doi:10.1021/JA305734X.

236. Collie, G.W.; Parkinson, G.N. The Application of DNA and RNA G-Quadruplexes to Therapeutic Medicines. *Chem Soc Rev* **2011**, *40*, 5867–5892, doi:10.1039/C1CS15067G.

237. Ohnmacht, S.A.; Neidle, S. Small-Molecule Quadruplex-Targeted Drug Discovery. *Bioorg Med Chem Lett* **2014**, *24*, 2602–2612, doi:10.1016/J.BMCL.2014.04.029.

238. Dhamodharan, V.; Pradeepkumar, P.I. Specific Recognition of Promoter G-Quadruplex DNAs by Small Molecule Ligands and Light-up Probes. *ACS Chem Biol* **2019**, *14*, 2102–2114, doi:10.1021/ACSCHEMBIO.9B00475/ASSET/IMAGES/LARGE/CB9B00475_0005.JPEG.

239. Han, F.X.; Wheelhouse, R.T.; Hurley, L.H. Interactions of TMPyP4 and TMPyP2 with Quadruplex DNA. Structural Basis for the Differential Effects on Telomerase Inhibition. *J Am Chem Soc* **1999**, *121*, 3561–3570, doi:10.1021/JA984153M/SUPPL_FILE/JA984153M_S.PDF.

240. Anantha, N. V.; Azam, M.; Sheardy, R.D. Porphyrin Binding to Quadrupled T4G4. *Biochemistry* **1998**, *37*, 2709–2714, doi:10.1021/BI973009V.

241. Haq, I.; Trent, J.O.; Chowdhry, B.Z.; Jenkins, T.C. Intercalative G-Tetraplex Stabilization of Telomeric DNA by a Cationic Porphyrin. *J Am Chem Soc* **1999**, *121*, 1768–1779, doi:10.1021/JA981554T/ASSET/IMAGES/LARGE/JA981554TF00008.JPEG.

242. Mita, H.; Ohyama, T.; Tanaka, Y.; Yamamoto, Y. Formation of a Complex of 5,10,15,20-Tetrakis(N-Methylpyridinium-4-Yl)-21H, 23H-Porphyrin with G-Quadruplex DNA. *Biochemistry* **2006**, *45*, 6765–6772, doi:10.1021/BI052442Z/SUPPL_FILE/BI052442ZSI20060323_042234.PDF.

243. Parkinson, G.N.; Ghosh, R.; Neidle, S. Structural Basis for Binding of Porphyrin to Human Telomeres. *Biochemistry* **2007**, *46*, 2390–2397, doi:10.1021/BI062244N.

244. Zhang, H.; Xiao, X.; Wang, P.; Pang, S.; Qu, F.; Ai, X.; Zhang, J. Conformational Conversion of DNA G-Quadruplex Induced by a Cationic Porphyrin. *Spectrochim Acta A Mol Biomol Spectrosc* **2009**, *74*, 243–247, doi:10.1016/J.SAA.2009.06.018.

245. Han, H.; Langley, D.R.; Rangan, A.; Hurley, L.H. Selective Interactions of Cationic Porphyrins with G-Quadruplex Structures. *J Am Chem Soc* **2001**, *123*, 8902–8913, doi:10.1021/JA002179J/SUPPL_FILE/JA002179J_S.PDF.

246. Bhattacharjee, A.J.; Ahluwalia, K.; Taylor, S.; Jin, O.; Nicoludis, J.M.; Buscaglia, R.; Brad Chaires, J.; Kornfilt, D.J.P.; Marquardt, D.G.S.; Yatsunyk, L.A. Induction of G-Quadruplex DNA Structure by Zn(II) 5,10,15,20-Tetrakis(N-Methyl-4-Pyridyl)Porphyrin. *Biochimie* **2011**, *93*, 1297–1309, doi:10.1016/J.BIOCHI.2011.05.038.

247. Pipier, A.; De Rache, A.; Modeste, C.; Amrane, S.; Mothes-Martin, E.; Stigliani, J.L.; Calsou, P.; Mergny, J.L.; Pratviel, G.; Gomez, D. G-Quadruplex Binding Optimization

by Gold(III) Insertion into the Center of a Porphyrin. *Dalton Transactions* **2019**, *48*, 6091–6099, doi:10.1039/C8DT04703K.

248. Morris, M.J.; Wingate, K.L.; Silwal, J.; Leeper, T.C.; Basu, S. The Porphyrin TmPyP4 Unfolds the Extremely Stable G-Quadruplex in MT3-MMP MRNA and Alleviates Its Repressive Effect to Enhance Translation in Eukaryotic Cells. *Nucleic Acids Res* **2012**, *40*, 4137–4145, doi:10.1093/NAR/GKR1308.

249. Wang, Y.; Hu, Y.; Wu, T.; Liu, H.; Zhang, L.; Zhou, X.; Shao, Y. Specific G-Quadruplex Structure Recognition of Human Telomeric RNA over DNA by a Fluorescently Activated Hyperporphyrin. *Analyst* **2015**, *140*, 5169–5175, doi:10.1039/C5AN00937E.

250. Qi, Q.; Yang, C.; Xia, Y.; Guo, S.; Song, D.; Su, H. Preferential Binding of π-Ligand Porphyrin Targeting 5′-5′ Stacking Interface of Human Telomeric RNA G-Quadruplex Dimer. *Journal of Physical Chemistry Letters* **2019**, *10*, 2143–2150, doi:10.1021/ACS.JPCLETT.9B00637/SUPPL_FILE/JZ9B00637_SI_001.PDF.

251. Mei, Y.; Deng, Z.; Vladimirova, O.; Gulve, N.; Johnson, F.B.; Drosopoulos, W.C.; Schildkraut, C.L.; Lieberman, P.M. TERRA G-Quadruplex RNA Interaction with TRF2 GAR Domain Is Required for Telomere Integrity. *Scientific Reports 2021 11:1* **2021**, *11*, 1–14, doi:10.1038/s41598-021-82406-x.

252. Arora, A.; Dutkiewicz, M.; Scaria, V.; Hariharan, M.; Maiti, S.; Kurreck, J. Inhibition of Translation in Living Eukaryotic Cells by an RNA G-Quadruplex Motif. *RNA* **2008**, *14*, 1290, doi:10.1261/RNA.1001708.

253. Kumari, S.; Bugaut, A.; Huppert, J.L.; Balasubramanian, S. An RNA G-Quadruplex in the 5’ UTR of the NRAS Proto-Oncogene Modulates Translation. *Nat Chem Biol* **2007**, *3*, 218–221, doi:10.1038/NCHEMBIO864.

254. Gehring, K.; Leroy, J.L.; Guéron, M. A Tetrameric DNA Structure with Protonated Cytosine-Cytosine Base Pairs. *Nature 1993 363:6429* **1993**, *363*, 561–565, doi:10.1038/363561a0.

255. Kang, C.H.; Berger, I.; Lockshin, C.; Ratliff, R.; Moyzis, R.; Rich, A. Crystal Structure of Intercalated Four-Stranded d(C3T) at 1.4 A Resolution. *Proc Natl Acad Sci U S A* **1994**, *91*, 11636–11640, doi:10.1073/PNAS.91.24.11636.

256. Berger, I.; Egli, M.; Rich, A. Inter-Strand C-H...O Hydrogen Bonds Stabilizing Four-Stranded Intercalated Molecules: Stereoelectronic Effects of O4' in Cytosine-Rich DNA. *Proc Natl Acad Sci U S A* **1996**, *93*, 12116, doi:10.1073/PNAS.93.22.12116.

257. Phan, A.T.; Leroy, J.L. Intramolecular I-Motif Structures of Telomeric DNA. *J Biomol Struct Dyn* **2000**, *17 Suppl 1*, 245–251, doi:10.1080/07391102.2000.10506628.

258. Guéron, M.; Leroy, J.L. The I-Motif in Nucleic Acids. *Curr Opin Struct Biol* **2000**, *10*, 326–331, doi:10.1016/S0959-440X(00)00091-9.

259. Assi, H.A.; Garavís, M.; González, C.; Damha, M.J. I-Motif DNA: Structural Features and Significance to Cell Biology. *Nucleic Acids Res* **2018**, *46*, 8038–8056, doi:10.1093/NAR/GKY735.

260. Wright, E.P.; Huppert, J.L.; Waller, Z.A.E. Identification of Multiple Genomic DNA Sequences Which Form I-Motif Structures at Neutral PH. *Nucleic Acids Res* **2017**, *45*, 2951–2959, doi:10.1093/NAR/GKX090.

261. Fleming, A.M.; Ding, Y.; Rogers, R.A.; Zhu, J.; Zhu, J.; Burton, A.D.; Carlisle, C.B.; Burrows, C.J. 4n-1 Is a "Sweet Spot" in DNA i-Motif Folding of 2'-Deoxycytidine Homopolymers. *J Am Chem Soc* **2017**, *139*, 4682–4689, doi:10.1021/JACS.6B10117.

262. Zhou, J.; Wei, C.; Jia, G.; Wang, X.; Feng, Z.; Li, C. Formation of I-Motif Structure at Neutral and Slightly Alkaline PH. *Mol Biosyst* **2010**, *6*, 580–586, doi:10.1039/B919600E.

263. Sun, D.; Hurley, L.H. The Importance of Negative Superhelicity in Inducing the Formation of G-Quadruplex and i-Motif Structures in the c-Myc Promoter: Implications for Drug Targeting and Control of Gene Expression. *J Med Chem* **2009**, *52*, 2863–2874, doi:10.1021/JM900055S.

264. Abdelhamid, M.A.S.; Fábián, L.; Macdonald, C.J.; Cheesman, M.R.; Gates, A.J.; Waller, Z.A.E. Redox-Dependent Control of i-Motif DNA Structure Using Copper Cations. *Nucleic Acids Res* **2018**, *46*, 5886–5893, doi:10.1093/NAR/GKY390.

265. Malliavin, T.E.; Gau, J.; Snoussi, K.; Leroy, J.L. Stability of the I-Motif Structure Is Related to the Interactions between Phosphodiester Backbones. *Biophys J* **2003**, *84*, 3838, doi:10.1016/S0006-3495(03)75111-X.

266. Phan, A.T.; Guéron, M.; Leroy, J.L. The Solution Structure and Internal Motions of a Fragment of the Cytidine-Rich Strand of the Human Telomere. *J Mol Biol* **2000**, *299*, 123–144, doi:10.1006/JMBI.2000.3613.

267. Fojtík, P.; Vorlíčková, M. The Fragile X Chromosome (GCC) Repeat Folds into a DNA Tetraplex at Neutral PH. *Nucleic Acids Res* **2001**, *29*, 4684, doi:10.1093/NAR/29.22.4684.

268. Brazier, J.A.; Shah, A.; Brown, G.D. I-Motif Formation in Gene Promoters: Unusually Stable Formation in Sequences Complementary to Known G-Quadruplexes. *Chemical Communications* **2012**, *48*, 10739–10741, doi:10.1039/C2CC30863K.

269. Mir, B.; Serrano, I.; Buitrago, D.; Orozco, M.; Escaja, N.; González, C. Prevalent Sequences in the Human Genome Can Form Mini I-Motif Structures at Physiological PH. *J Am Chem Soc* **2017**, *139*, 13985–13988, doi:10.1021/JACS.7B07383/SUPPL_FILE/JA7B07383_SI_001.PDF.

270. Benabou, S.; Garavís, M.; Lyonnais, S.; Eritja, R.; González, C.; Gargallo, R. Understanding the Effect of the Nature of the Nucleobase in the Loops on the Stability

of the I-Motif Structure. *Phys Chem Chem Phys* **2016**, *18*, 7997–8004, doi:10.1039/C5CP07428B.

271. Fleming, A.M.; Stewart, K.M.; Eyring, G.M.; Ball, T.E.; Burrows, C.J. Unraveling the 4n − 1 Rule for DNA I-Motif Stability: Base Pairs vs. Loop Lengths. *Org Biomol Chem* **2018**, *16*, 4537–4546, doi:10.1039/C8OB01198B.
272. Dai, J.; Hatzakis, E.; Hurley, L.H.; Yang, D. I-Motif Structures Formed in the Human c-MYC Promoter Are Highly Dynamic–Insights into Sequence Redundancy and I-Motif Stability. *PLoS One* **2010**, *5*, e11647, doi:10.1371/JOURNAL.PONE.0011647.
273. Mergny, J.L.; Lacroix, L.; Hélène, C.; Han, X.; Leroy, J.L. Intramolecular Folding of Pyrimidine Oligodeoxynucleotides into an I-DNA Motif. *J Am Chem Soc* **1995**, *117*, 8887–8898, doi:10.1021/JA00140A001.
274. Lieblein, A.L.; Fürtig, B.; Schwalbe, H. Optimizing the Kinetics and Thermodynamics of DNA I-Motif Folding. *Chembiochem* **2013**, *14*, 1226–1230, doi:10.1002/CBIC.201300284.
275. Garavís, M.; Escaja, N.; Gabelica, V.; Villasante, A.; González, C. Centromeric Alpha-Satellite DNA Adopts Dimeric i-Motif Structures Capped by AT Hoogsteen Base Pairs. *Chemistry* **2015**, *21*, 9816–9824, doi:10.1002/CHEM.201500448.
276. Canalia, M.; Leroy, J.L. [5mCCTCTCTCC]4: An i-Motif Tetramer with Intercalated T*T Pairs. *J Am Chem Soc* **2009**, *131*, 12870–12871, doi:10.1021/JA903210T.
277. Esmaili, N.; Leroy, J.L. I-Motif Solution Structure and Dynamics of the d(AACCCC) and d(CCCCAA) Tetrahymena Telomeric Repeats. *Nucleic Acids Res* **2005**, *33*, 213, doi:10.1093/NAR/GKI160.
278. Chen, Y.W.; Jhan, C.R.; Neidle, S.; Hou, M.H. Structural Basis for the Identification of an I-Motif Tetraplex Core with a Parallel-Duplex Junction as a Structural Motif in

CCG Triplet Repeats. *Angew Chem Int Ed Engl* **2014**, *53*, 10682–10686, doi:10.1002/ANIE.201405637.

279. Kim, S.E.; Hong, S.C. Two Opposing Effects of Monovalent Cations on the Stability of I-Motif Structure. *Journal of Physical Chemistry B* **2023**, *127*, 1932–1939, doi:10.1021/ACS.JPCB.2C07069/SUPPL_FILE/JP2C07069_SI_001.PDF.
280. Zeraati, M.; Langley, D.B.; Schofield, P.; Moye, A.L.; Rouet, R.; Hughes, W.E.; Bryan, T.M.; Dinger, M.E.; Christ, D. I-Motif DNA Structures Are Formed in the Nuclei of Human Cells. *Nat Chem* **2018**, *10*, 631–637, doi:10.1038/S41557-018-0046-3.
281. Yoga, Y.M.K.; Traore, D.A.K.; Sidiqi, M.; Szeto, C.; Pendini, N.R.; Barker, A.; Leedman, P.J.; Wilce, J.A.; Wilce, M.C.J. Contribution of the First K-Homology Domain of Poly(C)-Binding Protein 1 to Its Affinity and Specificity for C-Rich Oligonucleotides. *Nucleic Acids Res* **2012**, *40*, 5101–5114, doi:10.1093/NAR/GKS058.
282. Marsich, E.; Xodo, L.E.; Manzini, G. Widespread Presence in Mammals and High Binding Specificity of a Nuclear Protein That Recognises the Single-Stranded Telomeric Motif (CCCTAA)n. *Eur J Biochem* **1998**, *258*, 93–99, doi:10.1046/J.1432-1327.1998.2580093.X.
283. Niu, K.; Zhang, X.; Deng, H.; Wu, F.; Ren, Y.; Xiang, H.; Zheng, S.; Liu, L.; Huang, L.; Zeng, B.; et al. BmILF and I-Motif Structure Are Involved in Transcriptional Regulation of BmPOUM2 in Bombyx Mori. *Nucleic Acids Res* **2018**, *46*, 1710, doi:10.1093/NAR/GKX1207.
284. Kang, H.J.; Kendrick, S.; Hecht, S.M.; Hurley, L.H. The Transcriptional Complex between the BCL2 I-Motif and HnRNP LL Is a Molecular Switch for Control of Gene

Expression That Can Be Modulated by Small Molecules. *J Am Chem Soc* **2014**, *136*, 4172–4185, doi:10.1021/JA4109352.

285. Henikoff, S.; Thakur, J.; Kasinathan, S.; Talbert, P.B. Remarkable Evolutionary Plasticity of Centromeric Chromatin. *Cold Spring Harb Symp Quant Biol* **2017**, *82*, 71–82, doi:10.1101/SQB.2017.82.033605.

286. Kendrick, S.; Kang, H.J.; Alam, M.P.; Madathil, M.M.; Agrawal, P.; Gokhale, V.; Yang, D.; Hecht, S.M.; Hurley, L.H. The Dynamic Character of the BCL2 Promoter I-Motif Provides a Mechanism for Modulation of Gene Expression by Compounds That Bind Selectively to the Alternative DNA Hairpin Structure. *J Am Chem Soc* **2014**, *136*, 4161–4171, doi:10.1021/JA410934B.

287. Sutherland, C.; Cui, Y.; Mao, H.; Hurley, L.H. A Mechanosensor Mechanism Controls the G-Quadruplex/i-Motif Molecular Switch in the MYC Promoter NHE III1. *J Am Chem Soc* **2016**, *138*, 14138–14151, doi:10.1021/JACS.6B09196/SUPPL_FILE/JA6B09196_SI_001.PDF.

288. Takahashi, S.; Brazier, J.A.; Sugimoto, N. Topological Impact of Noncanonical DNA Structures on Klenow Fragment of DNA Polymerase. *Proc Natl Acad Sci U S A* **2017**, *114*, 9605–9610, doi:10.1073/PNAS.1704258114.

289. Dang, C. V. C-Myc Target Genes Involved in Cell Growth, Apoptosis, and Metabolism. *Mol Cell Biol* **1999**, *19*, 1–11, doi:10.1128/MCB.19.1.1.

290. Hsu, T.Y.T.; Simon, L.M.; Neill, N.J.; Marcotte, R.; Sayad, A.; Bland, C.S.; Echeverria, G. V.; Sun, T.; Kurley, S.J.; Tyagi, S.; et al. The Spliceosome Is a Therapeutic Vulnerability in MYC-Driven Cancer. *Nature 2015 525:7569* **2015**, *525*, 384–388, doi:10.1038/nature14985.

291. Dang, C. V. MYC on the Path to Cancer. *Cell* **2012**, *149*, 22–35, doi:10.1016/J.CELL.2012.03.003.

292. Poole, C.J.; van Riggelen, J. MYC—Master Regulator of the Cancer Epigenome and Transcriptome. *Genes (Basel)* **2017**, *8*, doi:10.3390/GENES8050142.

293. Casey, S.C.; Baylot, V.; Felsher, D.W. MYC: Master Regulator of Immune Privilege. *Trends Immunol* **2017**, *38*, 298–305, doi:10.1016/J.IT.2017.01.002.

294. Potts, M.A.; Mizutani, S.; Garnham, A.L.; Li Wai Suen, C.S.N.; Kueh, A.J.; Tai, L.; Pal, M.; Strasser, A.; Herold, M.J. Deletion of the Transcriptional Regulator TFAP4 Accelerates C-MYC-Driven Lymphomagenesis. *Cell Death & Differentiation 2023* **2023**, 1–10, doi:10.1038/s41418-023-01145-w.

295. Yang, D.; Hurley, L. Structure of the Biologically Relevant G-Quadruplex in The c-MYC Promoter. *http://dx.doi.org/10.1080/15257770600809913* **2007**, *25*, 951–968, doi:10.1080/15257770600809913.

296. Simonsson, T.; Pribylova, M.; Vorlickova, M. A Nuclease Hypersensitive Element in the Human C-Myc Promoter Adopts Several Distinct i-Tetraplex Structures. *Biochem Biophys Res Commun* **2000**, *278*, 158–166, doi:10.1006/BBRC.2000.3783.

297. Paul, S.; Hossain, S.S.; Samanta, A. Insights into the Folding Pathway of a C-MYC-Promoter-Based i-Motif DNA in Crowded Environments at the Single-Molecule Level. *Journal of Physical Chemistry B* **2020**, *124*, 763–770, doi:10.1021/ACS.JPCB.9B10633/ASSET/IMAGES/LARGE/JP9B10633_0003.JPEG.

298. Kuang, G.; Zhang, M.; Kang, S.; Hu, D.; Li, X.; Wei, Z.; Gong, X.; An, L.K.; Huang, Z.S.; Shu, B.; et al. Syntheses and Evaluation of New Bisacridine Derivatives for Dual Binding of G-Quadruplex and i-Motif in Regulating Oncogene c-Myc Expression. *J Med Chem* **2020**, *63*, 9136–9153, doi:10.1021/ACS.JMEDCHEM.9B01917/SUPPL_FILE/JM9B01917_SI_004.CSV.

299. Rogers, R.A.; Fleming, A.M.; Burrows, C.J. Unusual Isothermal Hysteresis in DNA I-Motif PH Transitions: A Study of the RAD17 Promoter Sequence. *Biophys J* **2018**, *114*, 1804–1815, doi:10.1016/J.BPJ.2018.03.012.

300. Mathur, V.; Verma, A.; Maiti, S.; Chowdhury, S. Thermodynamics of I-Tetraplex Formation in the Nuclease Hypersensitive Element of Human c-Myc Promoter. *Biochem Biophys Res Commun* **2004**, *320*, 1220–1227, doi:10.1016/J.BBRC.2004.06.074.

301. Pataskar, S.S.; Dash, D.; Brahmachari, S.K. Intramolecular I-Motif Structure at Acidic PH for Progressive Myoclonus Epilepsy (EPM1) Repeat d(CCCCGCCCCGCG)n. *J Biomol Struct Dyn* **2001**, *19*, 307–313, doi:10.1080/07391102.2001.10506741.

302. Britton, H.T.S.; Robinson, R.A. CXCVIII.—Universal Buffer Solutions and the Dissociation Constant of Veronal. *Journal of the Chemical Society (Resumed)* **1931**, 1456–1462, doi:10.1039/JR9310001456.

303. Nicoludis, J.M.; Barrett, S.P.; Mergny, J.L.; Yatsunyk, L.A. Interaction of Human Telomeric DNA with N-Methyl Mesoporphyrin IX. *Nucleic Acids Res* **2012**, *40*, 5432–5447, doi:10.1093/NAR/GKS152.

304. Ramsay, G.D.; Eftink, M.R. Analysis of Multidimensional Spectroscopic Data to Monitor Unfolding of Proteins. *Methods Enzymol* **1994**, *240*, 615–645, doi:10.1016/S0076-6879(94)40066-0.

305. Qu, X.; Chaires, J.B. Analysis of Drug-DNA Binding Data. *Methods Enzymol* **2000**, *321*, 353–369, doi:10.1016/S0076-6879(00)21202-0.

306. Gray, R.D.; Chaires, J.B. Analysis of Multidimensional G-Quadruplex Melting Curves. *Curr Protoc Nucleic Acid Chem* **2011**, *45*, 17.4.1-17.4.16, doi:10.1002/0471142700.NC1704S45.

307. Viladoms, J.; Parkinson, G.N. HELIX: A New Modular Nucleic Acid Crystallization Screen. *J Appl Crystallogr* **2014**, *47*, 948–955, doi:10.1107/S1600576714007407/HE5642SUP1.PDF.

308. Adams, P.D.; Afonine, P. V.; Bunkóczi, G.; Chen, V.B.; Davis, I.W.; Echols, N.; Headd, J.J.; Hung, L.W.; Kapral, G.J.; Grosse-Kunstleve, R.W.; et al. PHENIX: A Comprehensive Python-Based System for Macromolecular Structure Solution. *Acta Crystallogr D Biol Crystallogr* **2010**, *66*, 213–221, doi:10.1107/S0907444909052925.

309. Emsley, P.; Lohkamp, B.; Scott, W.G.; Cowtan, K. Features and Development of Coot. *Acta Crystallogr D Biol Crystallogr* **2010**, *66*, 486–501, doi:10.1107/S0907444910007493.

310. Mir, B.; Serrano, I.; Buitrago, D.; Orozco, M.; Escaja, N.; González, C. Prevalent Sequences in the Human Genome Can Form Mini I-Motif Structures at Physiological PH. *J Am Chem Soc* **2017**, *139*, 13985–13988, doi:10.1021/JACS.7B07383.

311. Amato, J.; Iaccarino, N.; D'Aria, F.; D'Amico, F.; Randazzo, A.; Giancola, C.; Cesàro, A.; di Fonzo, S.; Pagano, B. Conformational Plasticity of DNA Secondary Structures: Probing the Conversion between i-Motif and Hairpin Species by Circular Dichroism and Ultraviolet Resonance Raman Spectroscopies. *Physical Chemistry Chemical Physics* **2022**, *24*, 7028–7044, doi:10.1039/D2CP00058J.

312. Školáková, P.; Renčiuk, D.; Palacký, J.; Krafčík, D.; Dvořáková, Z.; Kejnovská, I.; Bednářová, K.; Vorlíčková, M. Systematic Investigation of Sequence Requirements for DNA I-Motif Formation. *Nucleic Acids Res* **2019**, *47*, 2177–2189, doi:10.1093/NAR/GKZ046.

313. Školáková, P.; Badri, Z.; Foldynová-Trantírková, S.; Ryneš, J.; Šponer, J.; Fojtová, M.; Fajkus, J.; Marek, R.; Vorlíčková, M.; Mergny, J.L.; et al. Composite 5-Methylations of Cytosines Modulate i-Motif Stability in a Sequence-Specific Manner:

Implications for DNA Nanotechnology and Epigenetic Regulation of Plant Telomeric DNA. *Biochim Biophys Acta Gen Subj* **2020**, *1864*, doi:10.1016/J.BBAGEN.2020.129651.

314. Martella, M.; Pichiorri, F.; Chikhale, R. v.; Abdelhamid, M.A.S.; Waller, Z.A.E.; Smith, S.S. I-Motif Formation and Spontaneous Deletions in Human Cells. *Nucleic Acids Res* **2022**, *50*, 3445–3455, doi:10.1093/NAR/GKAC158.

315. Britton, H.T.S.; Robinson, R.A. CXCVIII.—Universal Buffer Solutions and the Dissociation Constant of Veronal. *Journal of the Chemical Society (Resumed)* **1931**, 1456–1462, doi:10.1039/JR9310001456.

316. Webb, B.A.; Chimenti, M.; Jacobson, M.P.; Barber, D.L. Dysregulated PH: A Perfect Storm for Cancer Progression. *Nature Reviews Cancer 2011 11:9* **2011**, *11*, 671–677, doi:10.1038/nrc3110.

317. Sun, D.; Hurley, L.H. The Importance of Negative Superhelicity in Inducing the Formation of G-Quadruplex and i-Motif Structures in the c-Myc Promoter: Implications for Drug Targeting and Control of Gene Expression. *J Med Chem* **2009**, *52*, 2863, doi:10.1021/JM900055S.

318. Caballero, I.; Sammito, M.D.; Afonine, P. V.; Uson, I.; Read, R.J.; McCoy, A.J. Detection of Translational Noncrystallographic Symmetry in Patterson Functions. *Acta Crystallogr D Struct Biol* **2021**, *77*, 131–141, doi:10.1107/S2059798320016836/GM5078SUP1.PDF.

319. Nespolo, M.; Souvignier, B.; Litvin, D.B. About the Concept and Definition of "Noncrystallographic Symmetry." *Zeitschrift fur Kristallographie* **2008**, *223*, 605–606, doi:10.1524/ZKRI.2008.1137/MACHINEREADABLECITATION/RIS.

320. Read, R.J.; Adams, P.D.; McCoy, A.J. Intensity Statistics in the Presence of Translational Noncrystallographic Symmetry. *urn:issn:0907-4449* **2013**, *69*, 176–183, doi:10.1107/S0907444912045374.

321. Jamshidiha, M.; Pérez-Dorado, I.; Murray, J.W.; Tate, E.W.; Cota, E.; Read, R.J. Coping with Strong Translational Noncrystallographic Symmetry and Extreme Anisotropy in Molecular Replacement with Phaser: Human Rab27a. *Acta Crystallogr D Struct Biol* **2019**, *75*, 342, doi:10.1107/S2059798318017825.

322. Keating, L.R.; Szalai, V.A. Parallel-Stranded Guanine Quadruplex Interactions with a Copper Cationic Porphyrin. *Biochemistry* **2004**, *43*, 15891–15900, doi:10.1021/BI0483209/SUPPL_FILE/BI0483209SI20040805_020804.PDF.

323. Guo, L.; Dong, W.; Tong, X.; Dong, C.; Shuang, S. Study on Spectroscopic Characterization of Pd Porphyrin and Its Interaction with CtDNA. *Talanta* **2006**, *70*, 630–636, doi:10.1016/J.TALANTA.2006.01.031.

324. Sabharwal, N.C.; Chen, J.; Lee, J.H.J.; Gangemi, C.M.A.; D'urso, A.; Yatsunyk, L.A. Interactions Between Spermine-Derivatized Tentacle Porphyrins and The Human Telomeric DNA G-Quadruplex. *Int J Mol Sci* **2018**, *19*, 3686, doi:10.3390/IJMS19113686.

325. Ghazaryan, A.A.; Dalyan, Y.B.; Haroutiunian, S.G.; Tikhomirova, A.; Taulier, N.; Wells, J.W.; Chalikian, T. V. Thermodynamics of Interactions of Water-Soluble Porphyrins with RNA Duplexes. *J Am Chem Soc* **2006**, *128*, 1914–1921, doi:10.1021/JA054070N/ASSET/IMAGES/MEDIUM/JA054070NE00006.GIF.

326. Uno, T.; Hamasaki, K.; Tanigawa, M.; Shimabayashi, S. Binding of Meso-Tetrakis(N-Methylpyridinium-4-Yl)Porphyrin to Double Helical RNA and DNA·RNA Hybrids. *Inorg Chem* **1997**, *36*, 1676–1683, doi:10.1021/IC960824A/ASSET/IMAGES/LARGE/IC960824AF00010.JPEG.

327. McGhee, J.D.; von Hippel, P.H. Theoretical Aspects of DNA-Protein Interactions: Co-Operative and Non-Co-Operative Binding of Large Ligands to a One-Dimensional Homogeneous Lattice. *J Mol Biol* **1974**, *86*, 469–489, doi:10.1016/0022-2836(74)90031-X.

Supplementary Information

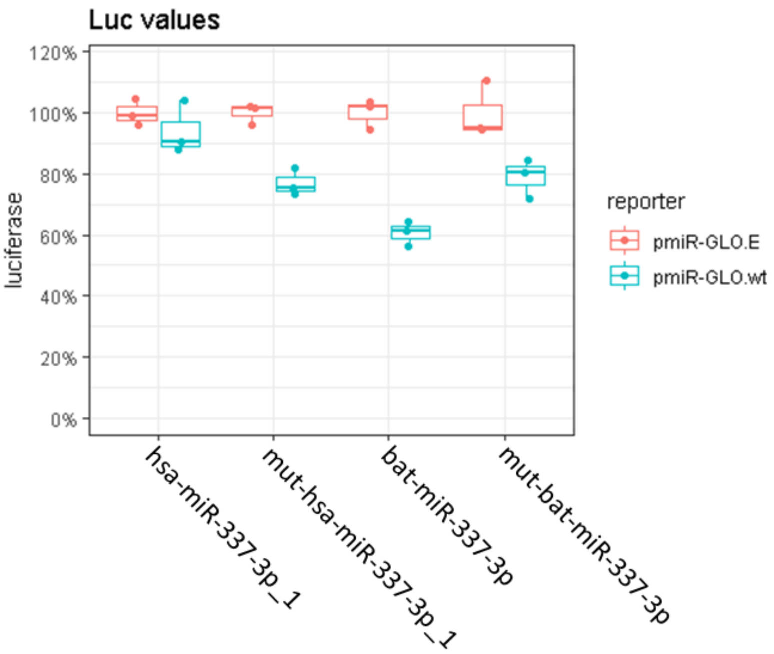


Fig. S1 *Luciferase assay conducted at the Max Planck Institute for Psycholinguistics using complementary targets.*

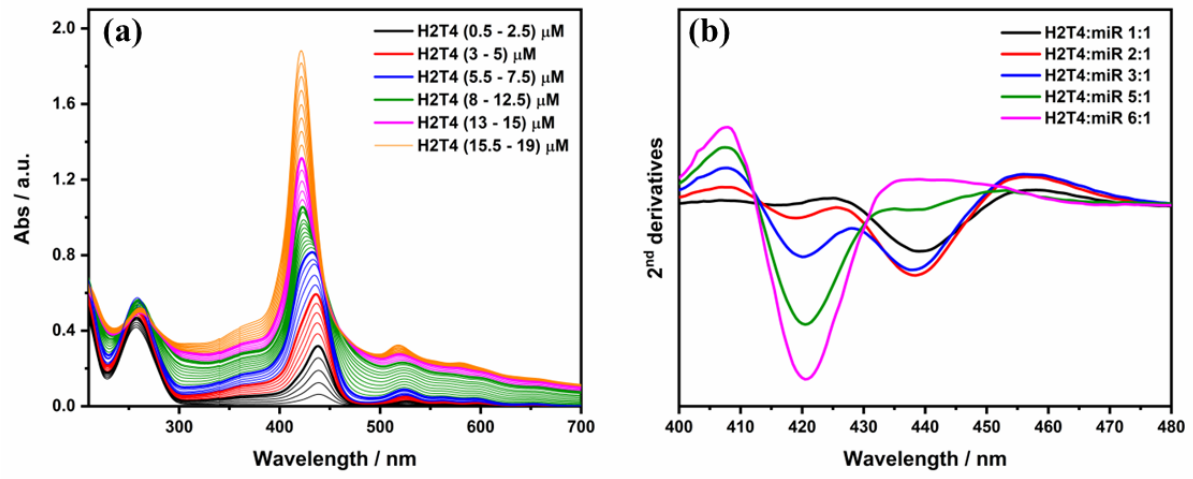


Fig. S2 (a) *UV-vis titration of miR-26b-5p [2.5 μM] in PBS buffer in the presence of increasing amount of H2T4 (from 0.5 μM to 19 μM). Each addition of 0.5 μM of H2T4. Bold lines represent the concentration in which there is a change of the molar extinction coefficient. Displayed colors are the same used in the "plot of absorbance at 421 nm vs. the ratio [H2T4]/[miR-26b-5p]" (**Fig. 32, inset**). **(b)** 2nd derivatives of UV-vis spectra of H2T4 (at the break points) in the presence of the miRNA.*

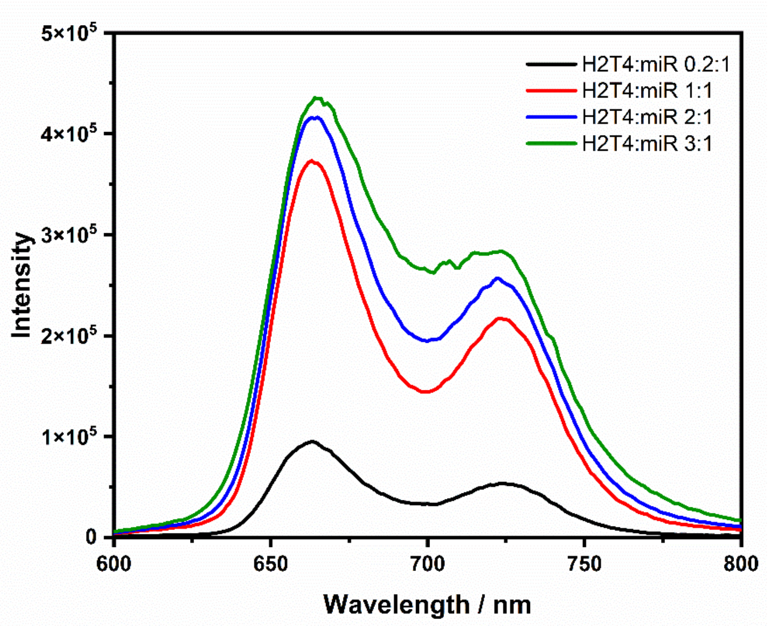


Fig. S3 *Fluorescence titration of miR-26b-5p [2.5 μM] in PBS buffer in the presence of increasing amount of H2T4 (0.5 μM black curve; 2.5 μM red curve; 5 μM blue curve; 7.5 μM green curve).*

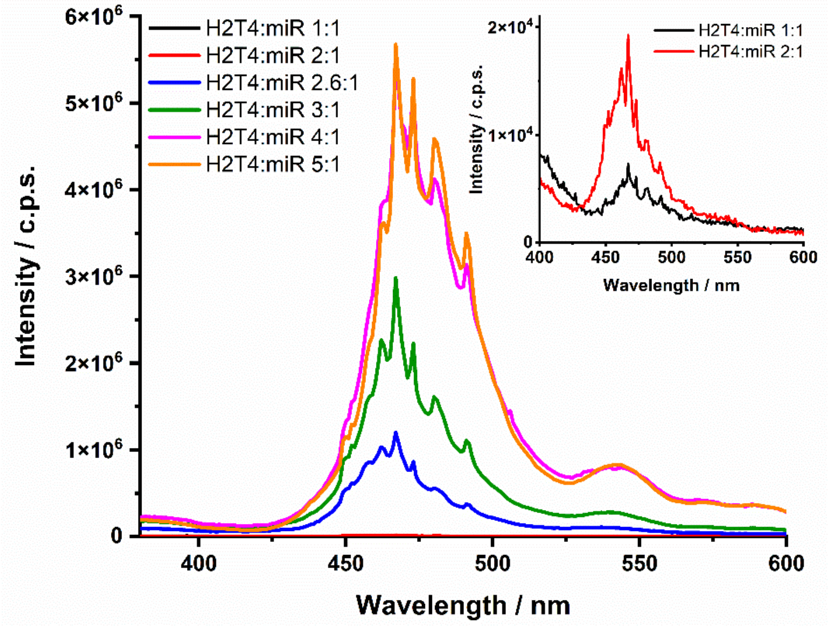


Fig. S4 *RLS titration of miR-26b-5p [2.5 µM] in PBS buffer in the presence of increasing amount of H2T4 (2.5 µM black curve; 5 µM red curve; 6.5 µM blue curve; 7.5 µM green curve; 10 µM magenta curve; 12.5 µM orange curve).*

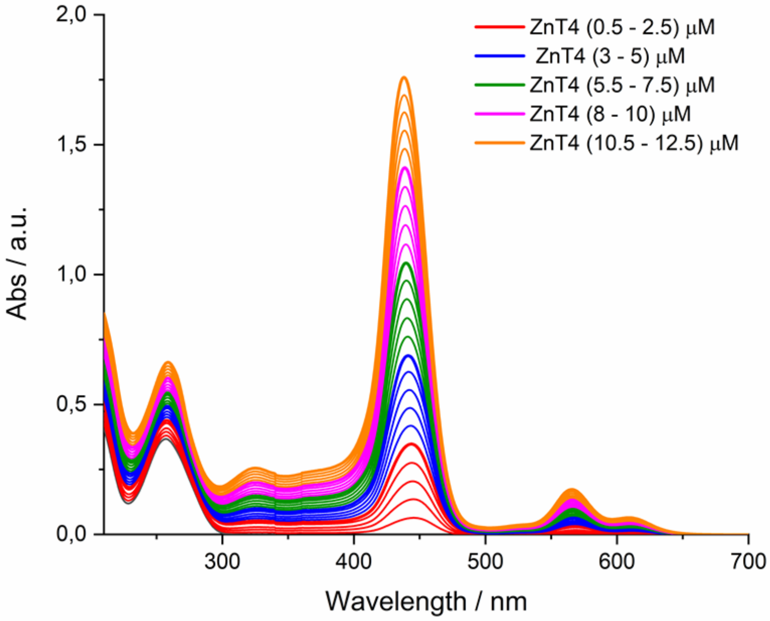


Fig. S5 *UV-vis titration of miR-26b-5p [2.5 µM] in PBS buffer in the presence of increasing amount of Zn2T4 (from 0.5 µM to 12.5 µM). Each addition of ZnT4 of 0.5 µM.*

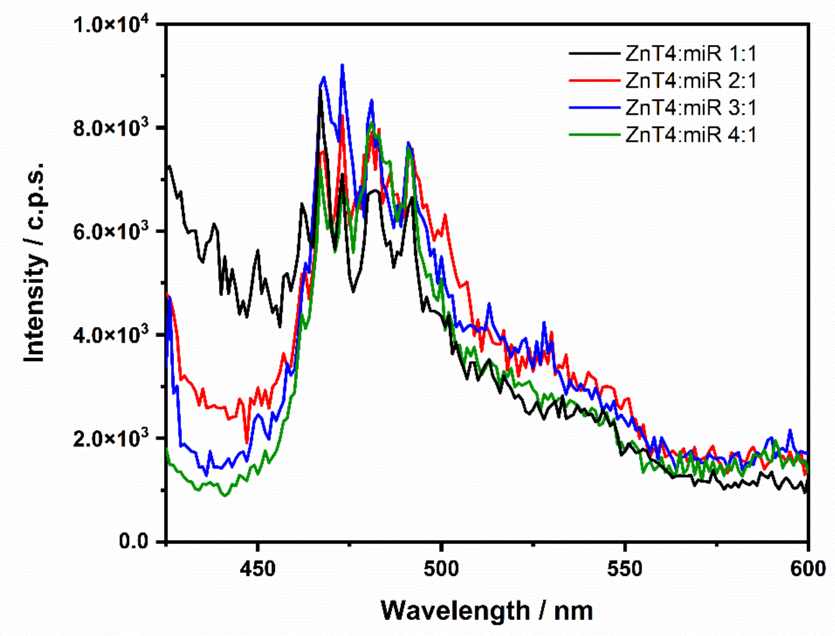


Fig. S6 *RLS titration of miR-26b-5p [2.5 µM] in PBS buffer in the presence of increasing amount of ZnT4 (2.5 µM black curve; 5 µM red curve; 7.5 µM blue curve; 10 µM green curve.*

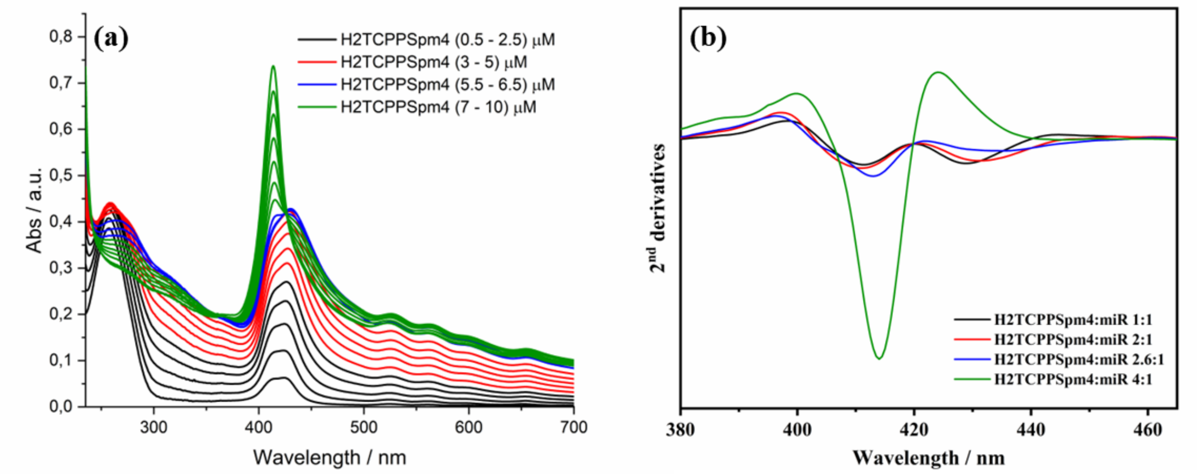


Fig. S7 ***(a)*** *UV-vis titration of miR-26b-5p [2.5 µM] in PBS buffer in the presence of increasing amount of H2TCPPSpm4 (from 0.5 µM to 10 µM). Each addition of 0.5 µM of H2TCPPspm4. Displayed colors are the same used in the "plot of absorbance at 414 nm vs. the ratio [H2TCPPSpm4] / [miR-26b-5p]" (**Fig. 37**, inset).* ***(b)*** *2nd derivatives of UV-vis spectra of H2TCPPSpm4 (at the break points) in the presence of the miRNA.*

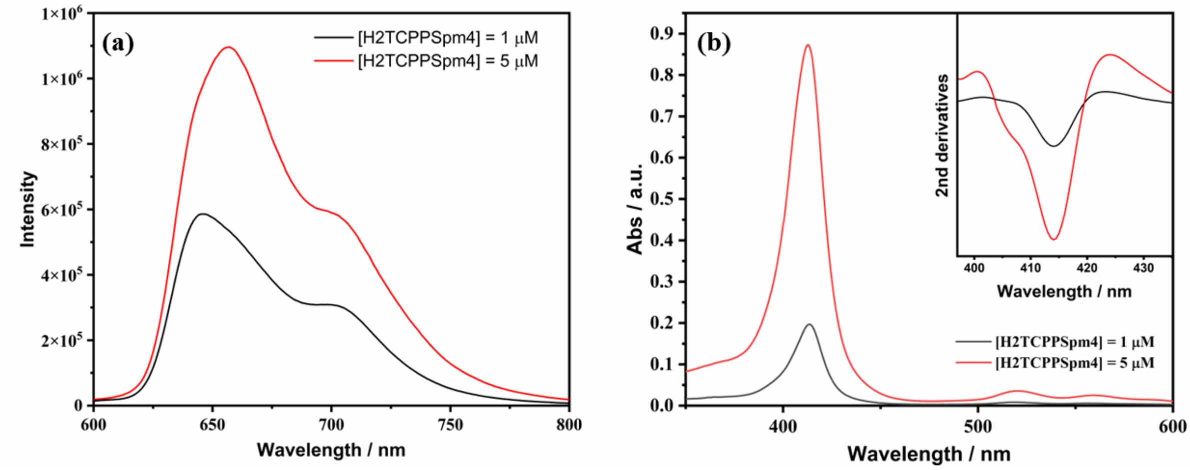


Fig. S8 (a) *Emission spectra of H2TCPPSpm4 [1 µM] (black curve) and [5 µM] (red curve) in PBS buffer.* ***(b)*** *Absorption spectra of H2TCPPSpm4 [1 µM] (black curve) and [5 µM] (red curve) in PBS buffer. Inset: second derivatives of H2TCPPSpm4 [1 µM] (black curve) and [5 µM] (red curve).*

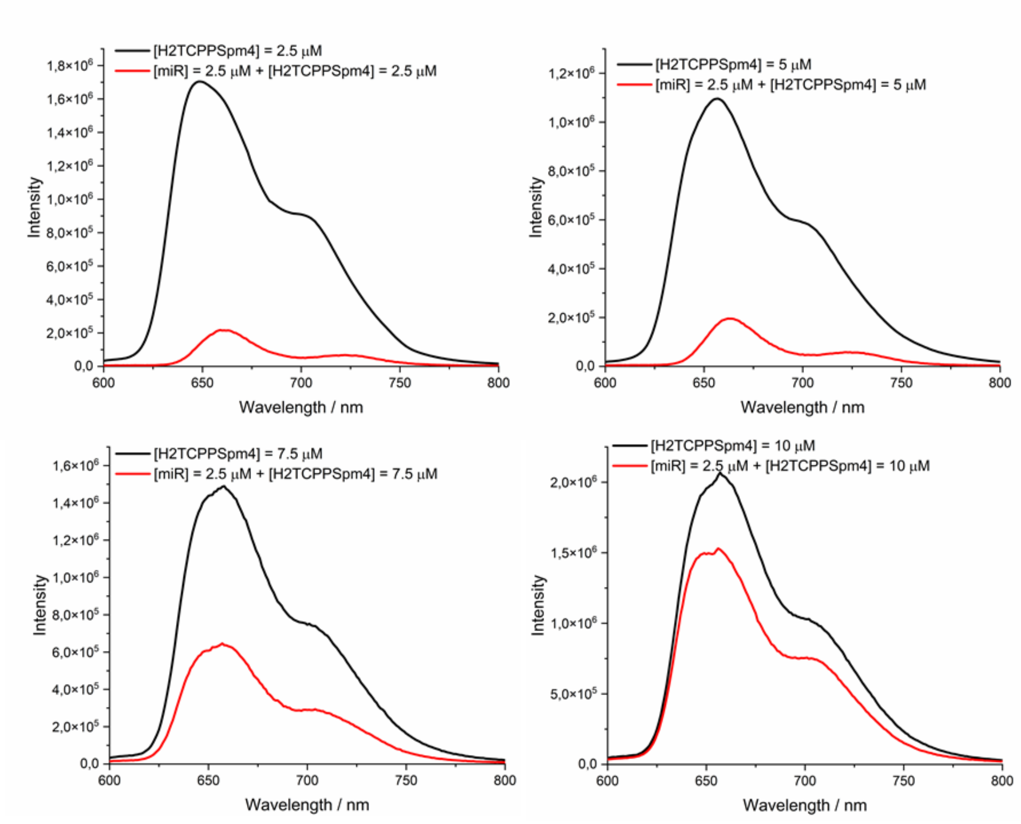


Fig. S9 *Emission spectra of H2TCPPSpm4 (black curve) at different concentrations and in the presence of miR-26b-5p [2.5 µM] (red curve) in PBS buffer.*

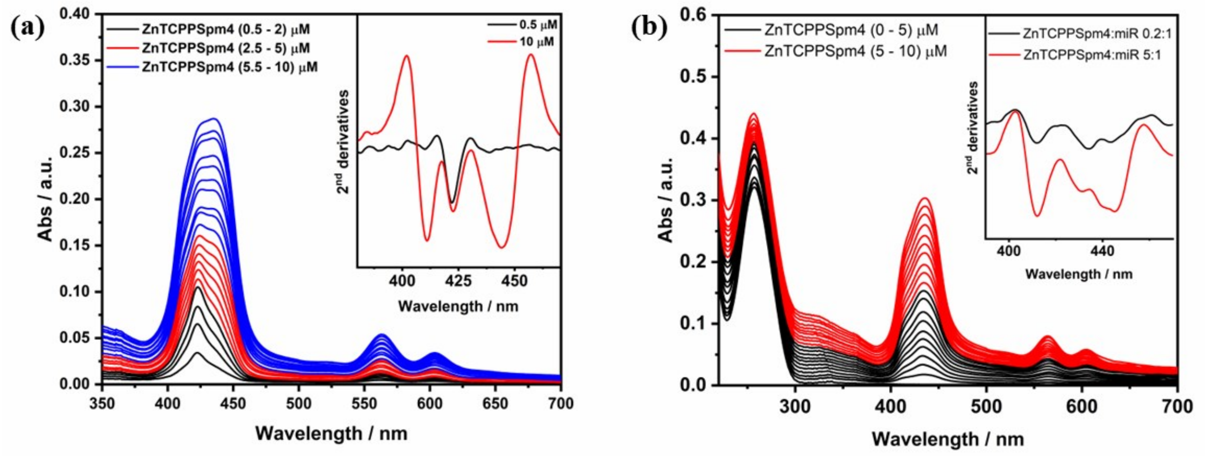


Fig. S10 (a) *UV-vis titration of increasing amount of ZnTCPPSpm4 (from 0.5 μM to 10 μM) in PBS buffer. Each addition of 0.5 μM of ZnTCPPspm4. Inset: 2*nd *derivatives of UV-vis spectra of ZnTCPPSpm4 at the concentration of 0.5 μM (black line) and 10 μM (red line) .* ***(b)*** *UV-vis titration of miR-26b-5p [2.5 μM] in PBS buffer in the presence of increasing amount of ZnTCPPSpm4 (from 0.5 μM to 10 μM). Each addition of 0.5 μM of ZnTCPPspm4. Inset: 2*nd *derivatives of UV-vis spectra of ZnTCPPSpm4 at the concentration of 0.5 μM (black line) and 10 μM (red line) in presence of miR.*

Scatchard plots and apparent association constants

The titration data were treated with Peacocke-Sherrett method's [203] in order to calculate the values of ν (number of moles of bound porphyrin per mole of total miRNA) and L (molar concentration of free porphyrin at equilibrium) needed to graph the Scatchard plot (**Fig. S11**).

At each intermediate titration point, the fraction of bound porphyrin was calculated by the use of the expression $\alpha = (A_f - A)/(A_f - A_b)$, where A_f is the absorption at the Soret maximum for free porphyrin (421 nm for H2T4, 436 nm for ZnT4, 414 nm for H2TCPPSpm4 and 423 nm for ZnTCPPSpm4), A_b is the absorbance at same wavelength of A_f but of a solution in which the miRNA concentration was 25 times greater than the initial concentration of porphyrin (conditions for which complete binding may be assumed) and A is the absorbance at the Soret maximum of porphyrin at any given point during the titration.

In a typical linear Scatchard plot (**Fig. S11 (a), (c) and (d)**), the slope extrapolation corresponds to K_{app} according to Eq. 1. Thus, our estimated K_{app} values are listed in **Table 2**. Such values align with previously reported data on the interaction between porphyrins and various oligonucleotide structures Specifically, for DNA G-quadruplex structures [322–324] the binding constants typically fall within the range of 10^5 to 10^7. In the case of RNA G-quadruplexes [250] the binding constant is generally around 10^5, while for RNA duplex structures [325,326] it typically ranges from 10^4 to 10^5. For ZnT4 was not possible to obtain the K_{app} because the Scatchard plot significantly deviates from linearity with an upward curvature (**Fig. S11 (b)**). The latter arises as a consequence of cooperativity and neighbour exclusion effects [327] suggesting a negative co-operativity binding of ZnT4 with miR-26b-5p. This negative cooperative binding behaviour and curvature in the Scatchard plot could be attributed to very weak binding interactions between the porphyrin and miRNA, consistent with the spectroscopic results. In fact, the weaker the interactions, the more difficult it is for multiple porphyrins to bind to the miRNA, leading to non-linearity in the Scatchard plot.

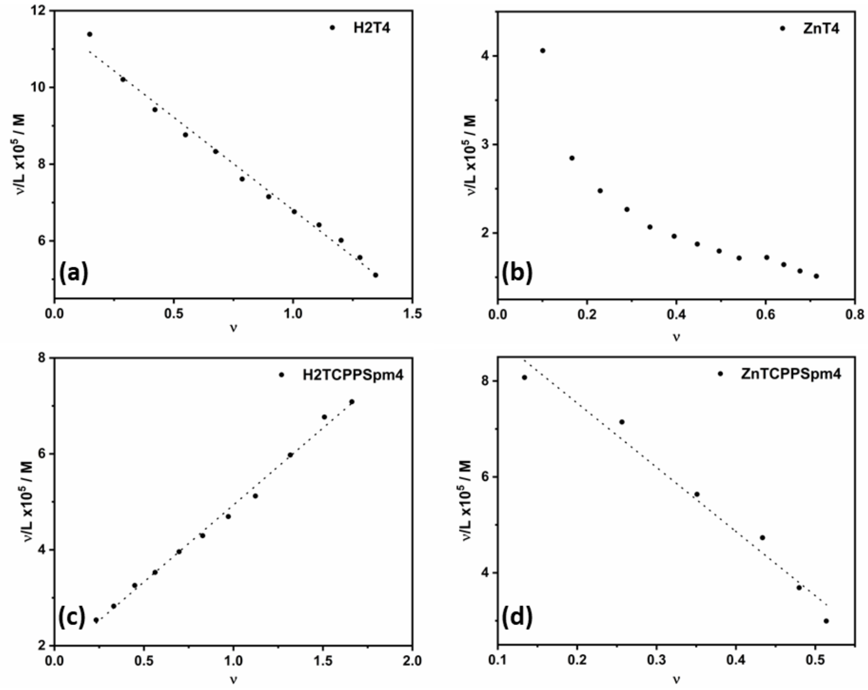


***Fig. S11** Scatchard plots for titration of miR-26b-5p with H2T4 **(a)**, ZnT4 **(b)**, H2TCPPSpm4 **(c)** and ZnTCPPSpm4 **(d)**.*

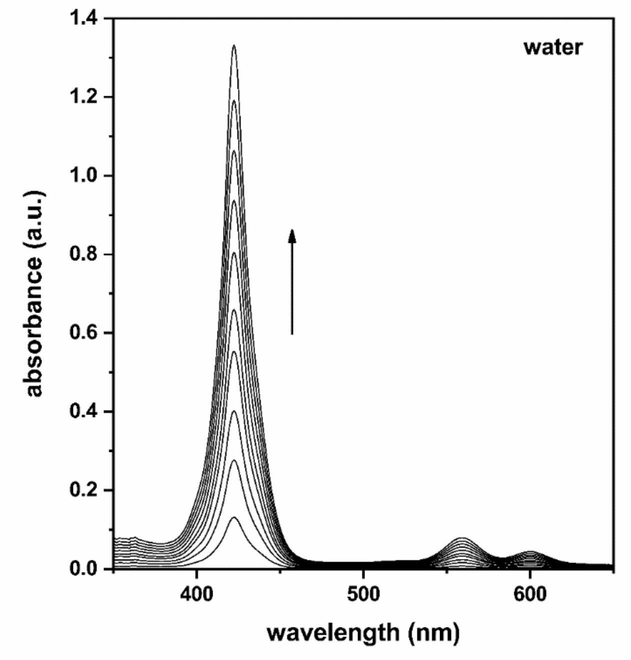


Fig. S12 *UV-Vis titration of ZnTCPPspm4 in ultra-pure water (pH = 7.3) from 1 µM to 10 µM.*

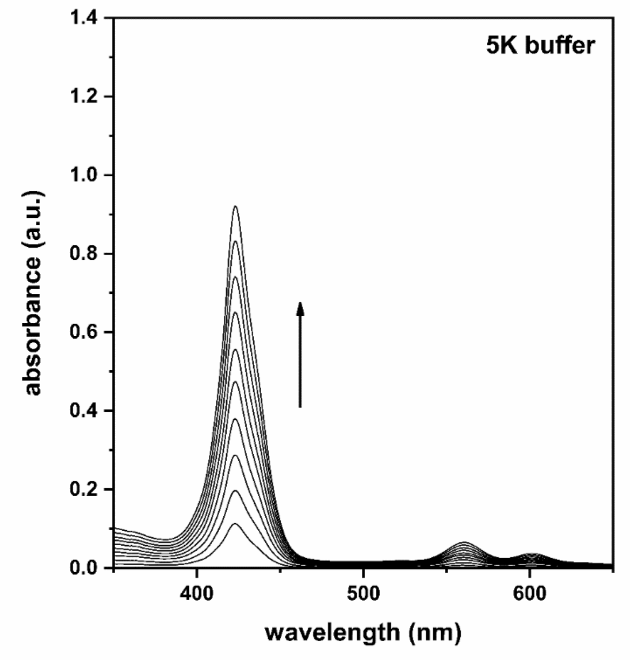


Fig. S13 *UV-Vis titration of ZnTCPPspm4 in 5K buffer (pH = 7.3) from 1 µM to 10 µM.*

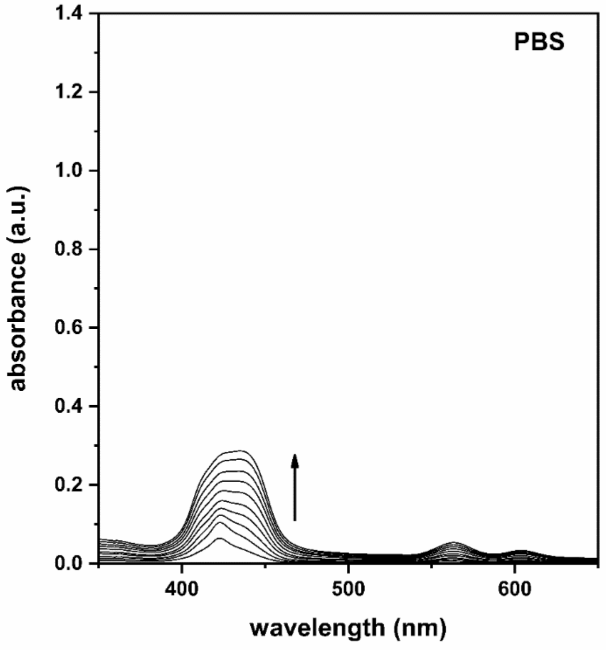


Fig. S14 *UV-Vis titration of ZnTCPPspm4 in PBS buffer (pH = 7.3) from 1 µM to 10 µM.*

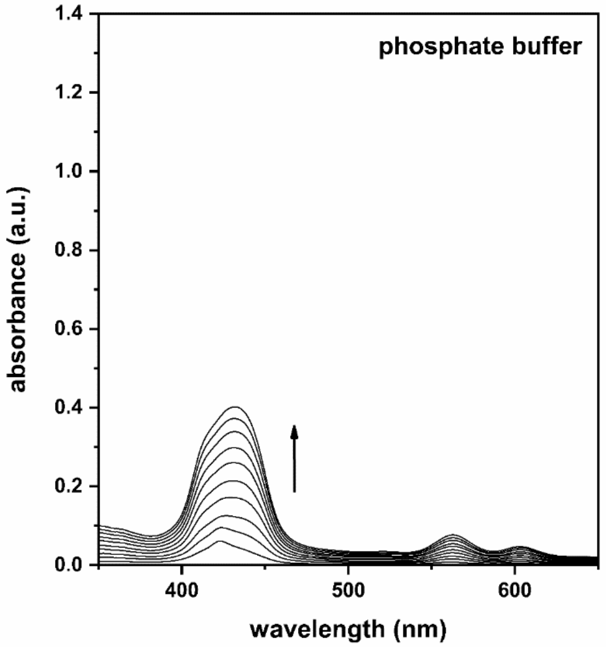


Fig. S15 *UV/Vis titration of ZnTCPPspm4 in phosphate buffer (pH = 7.3) from 1 µM to 10 µM.*

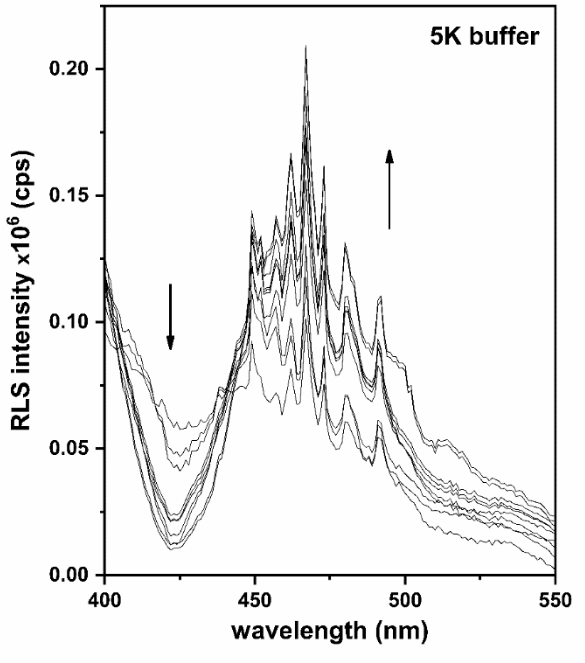


Fig. S16 *RLS titration of ZnTCPPspm4 in 5K buffer (pH = 7.3) from 1 µM to 10 µM.*

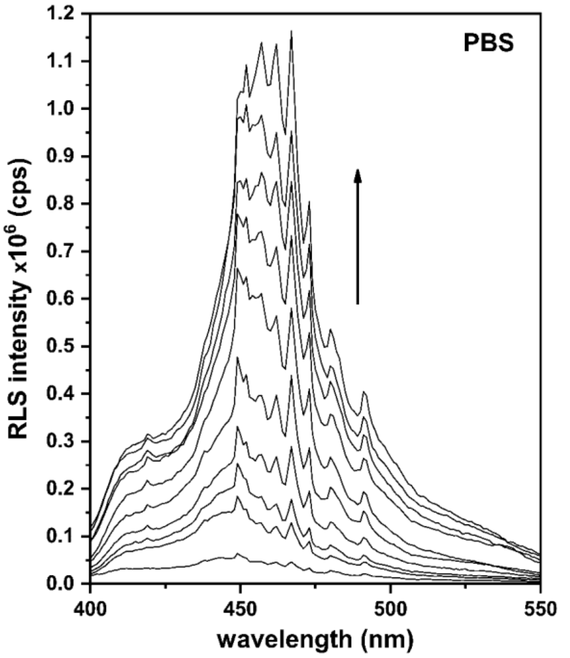


Fig. S17 *RLS titration of ZnTCPPspm4 in PBS buffer (pH = 7.3) from 1 µM to 10 µM.*

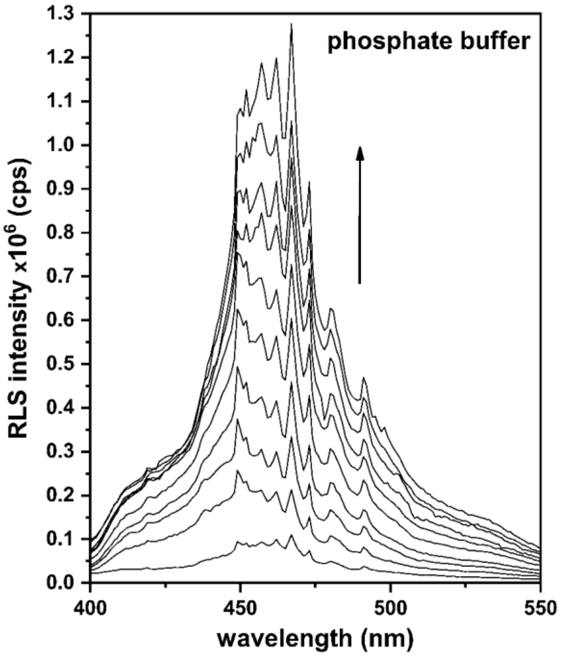


Fig. S18 *RLS titration of ZnTCPPspm4 in phosphate buffer (pH = 7.3) from 1 µM to 10 µM.*

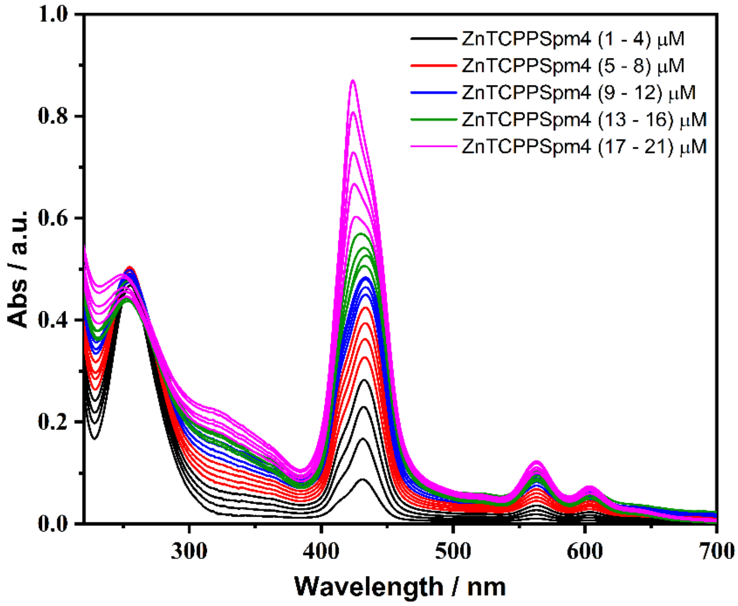


Fig. S19 *UV-vis titration of TERRA [2 μM] in 5K buffer in the presence of increasing amount of ZnTCPPSpm4 (from 1 μM to 21 μM). Each addition of 1 μM of ZnTCPPSpm4. Displayed colors are the same used in the "plot of absorbance at 423 nm vs. the ratio [ZnTCPPSpm4]/[TERRA]"* ***(Fig. 56, inset)****.*

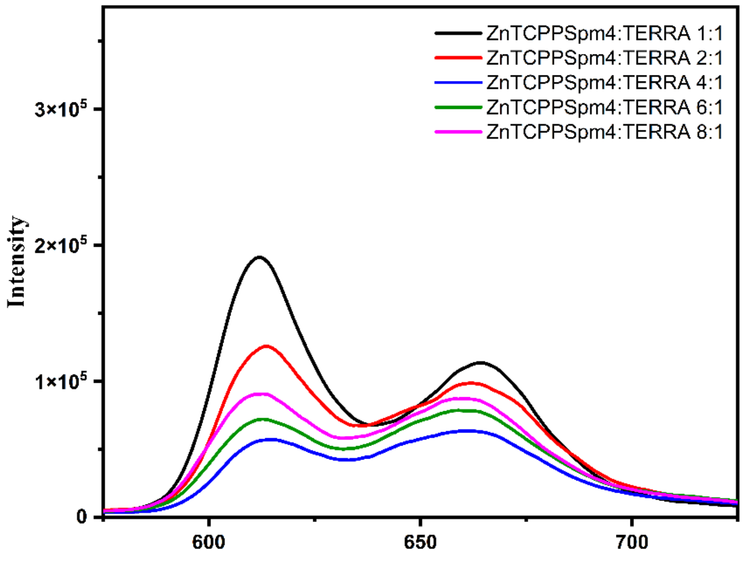


Fig. S20 *Fluorescence titration of TERRA [2 μM] in 5K buffer in the presence of increasing amount of ZnTCPPSpm4(2 μM black curve; 4 μM red curve; 8 μM blue curve; 12 μM green curve; 16 μM magenta curve).*

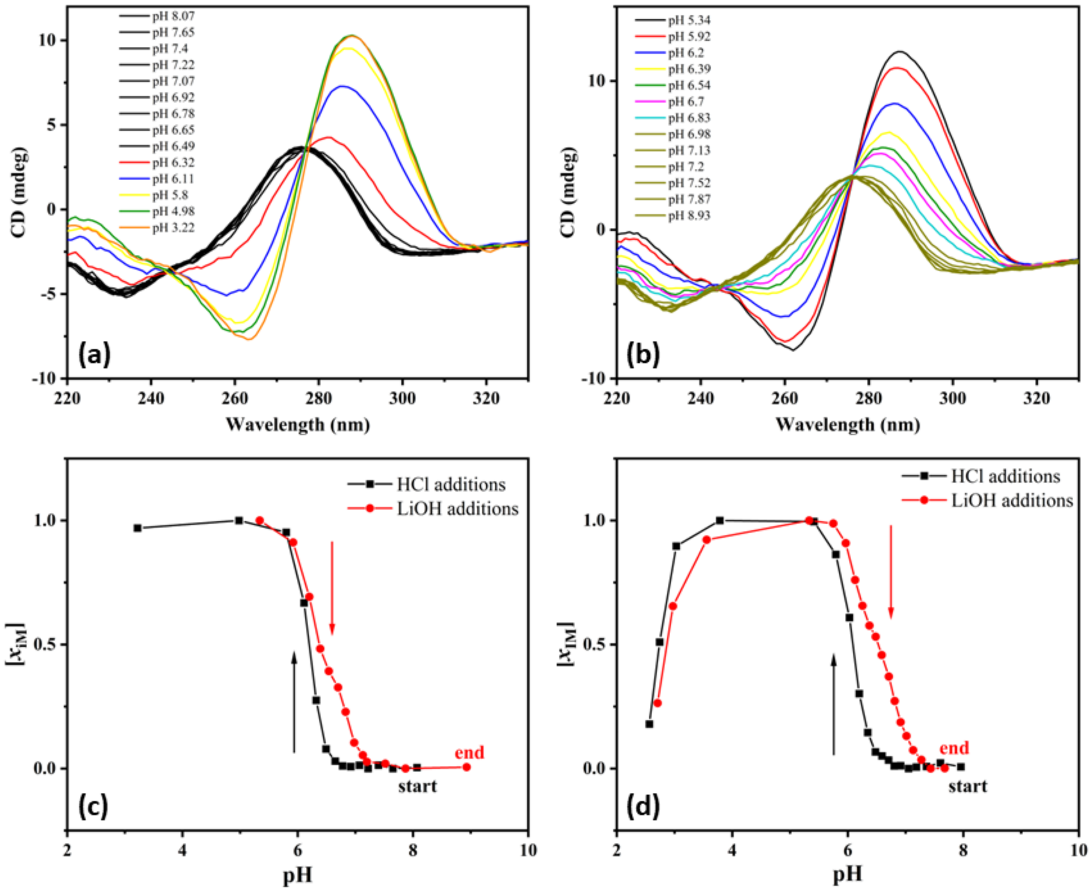


Fig. S21 *The raw CD spectra (top) of PyA-3 in 20 mM KPi buffer at various pH values using manual titration.* ***(a)*** *Acid titration with 13 additions of 5 µL of 1.0 M HCl and* ***(b)*** *reverse base titration with 13 additions of 5 µL of 1.0 M LiOH. Titration plots for PyA-3 (bottom) in KPi buffer.* ***(c)*** *Manual titration and* ***(d)*** *automatic titration for PyA-2 each one was performed with 13 additions of 5 µL of 1.0 M HCl (black squares), followed by 13 additions of 5 µL of 1.0 M LiOH (red circles). [PyA-3] ~ 2 µM in 20 mM KPi pH = 8.08.*

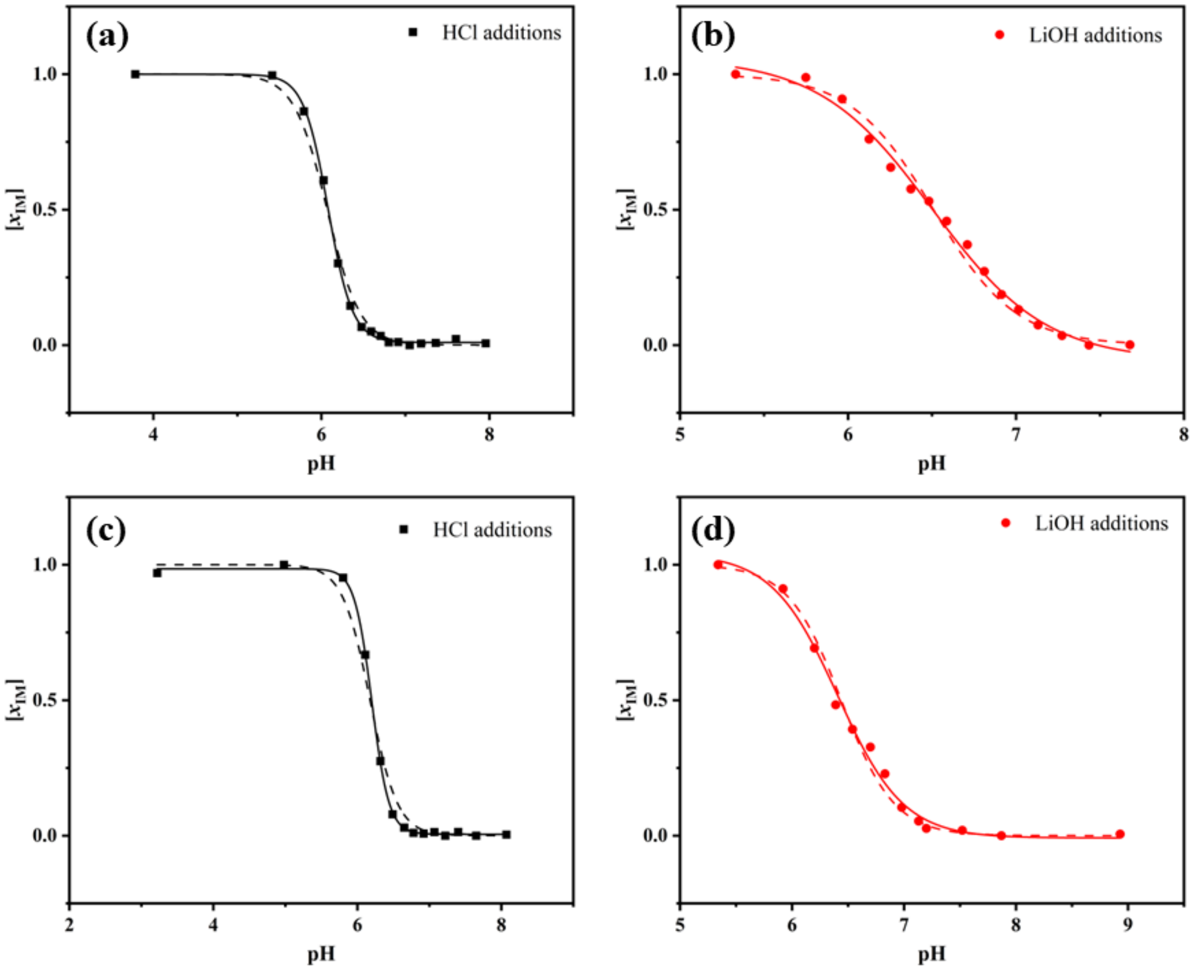


Fig. S22 *The fit of the data for pH_T titration on PyA-3 in KPi buffer. The fit of the automatic titration with HCl* ***(a)*** *and with LiOH* ***(b)****. The fit of the manual titration with HCl* ***(c)*** *and with LiOH* ***(d)****. Solid lines represent the fitting by using Boltzmann sigmoid equation, whereas dashed lines by using eq. (2).*

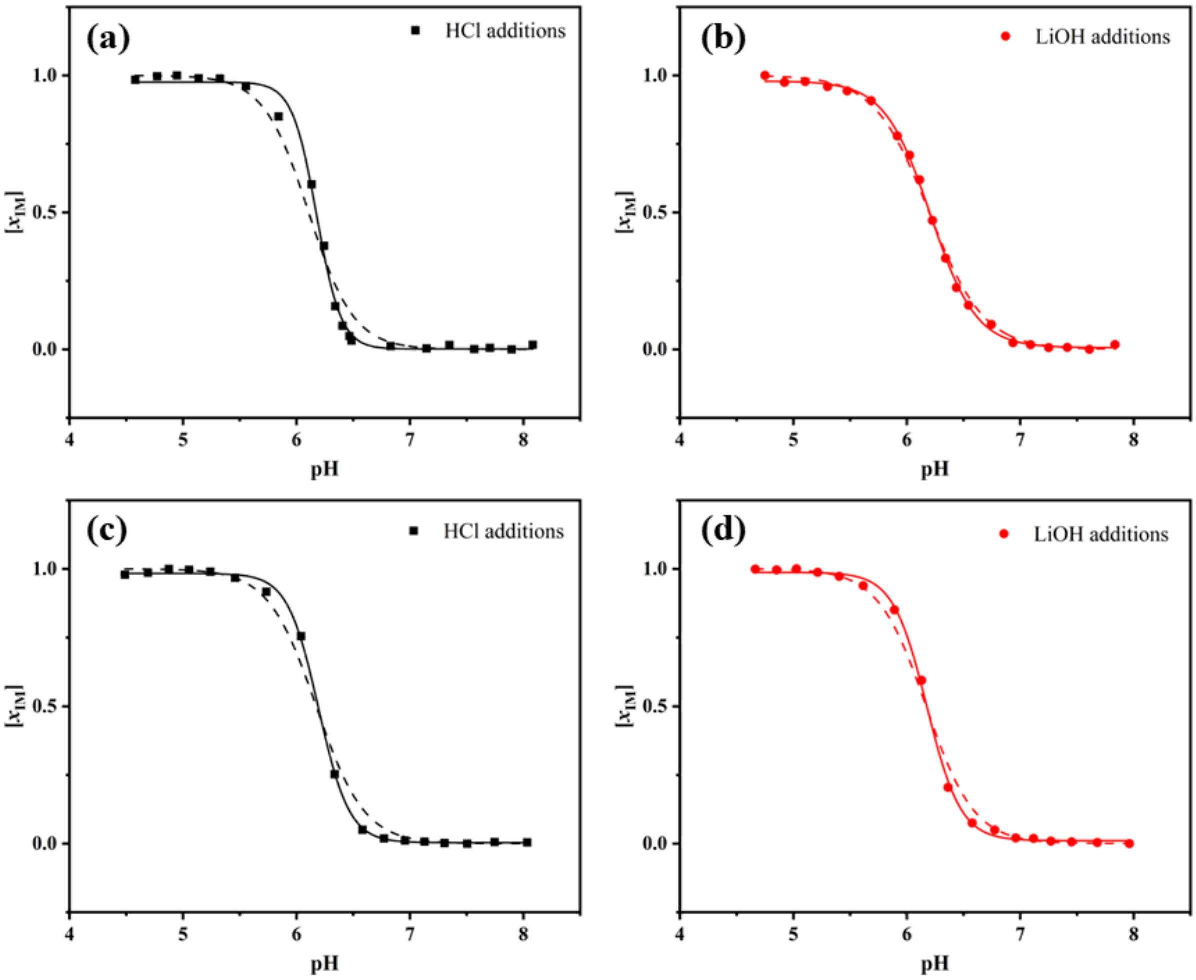


Fig. S23 *The fit of the data for pH_T titration on PyB-1 (top) and PyB-2 (bottom) in BR buffer. The fit of the automatic titration on PyB-1 with HCl* ***(a)*** *and with LiOH* ***(b)****.The fit of the automatic titration on PyB-2 with HCl* ***(c)*** *and with LiOH* ***(d)****. Solid lines represent the fitting by using Boltzmann sigmoid equation, whereas dashed lines by using eq. (2).*

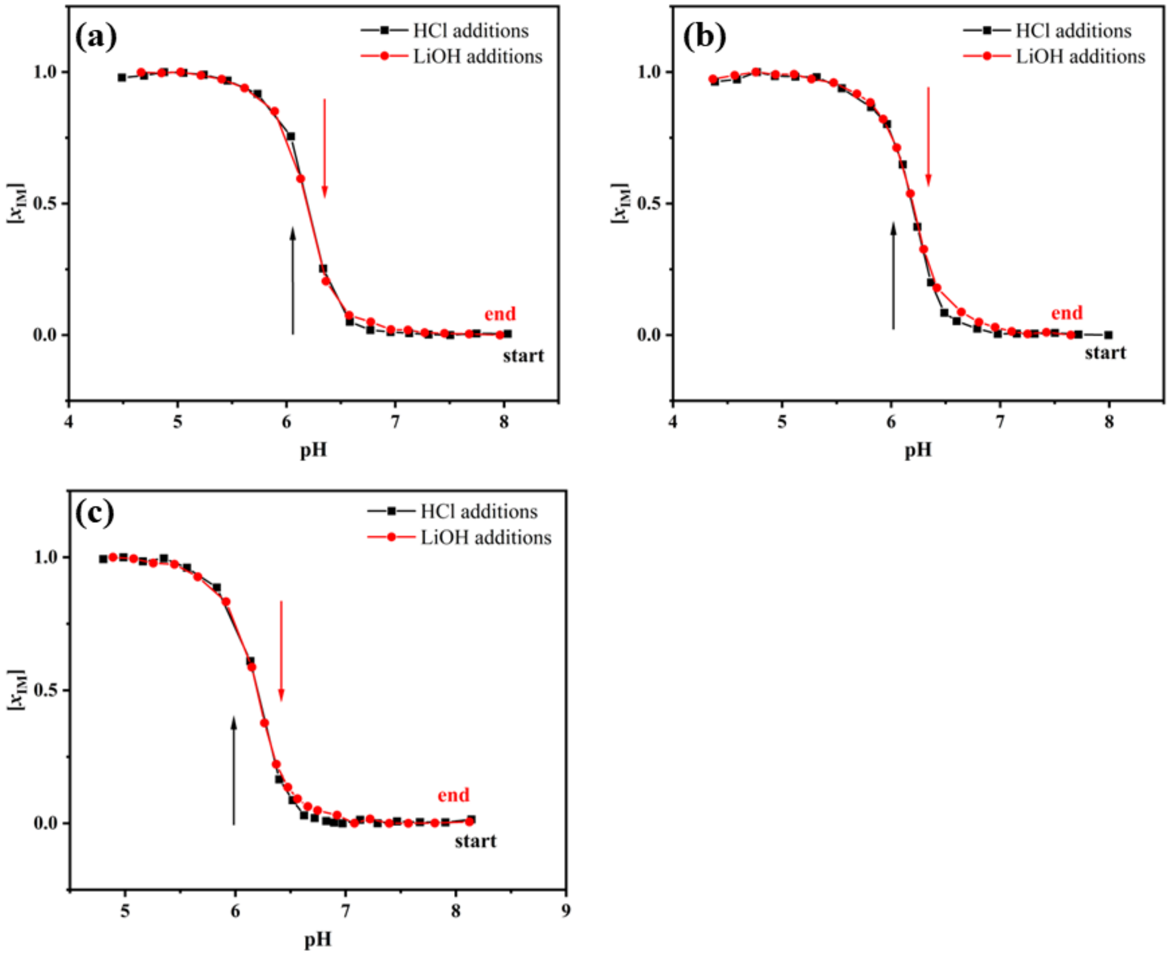


***Fig. S24** Titration plots for PyB-2 in BR buffer at 25 °C performed on three different days. First titration **(a)** was performed by automatic titration with 17 additions of 4 μL of HCl 1 M, followed by 17 additions of 4 μL of LiOH 1 M. Second **(b)** and third **(c)** titrations were performed by automatic titration with HCl 1 M (7 additions of 4 μL, 6 additions of 2 μL, 7 additions of 4 μL), followed by LiOH 1 M additions (7 additions of 4 μL, 6 additions of 2 μL, 7 additions of 4 μL). [PyB-2] ~ 2 μM in 20 mM BR buffer [KCl] = 140 mM pH = 8.02.*

***Table S1** pH_T values for PyB-2 obtained in 20 mM BR buffer [KCl] = 140 mM on 3 different days.*

Sample	Buffer	Titration	HCl additions (Boltzmann)	LiOH additions (Boltzmann)
PyB-2 (first)	BR	Automatic	6.19 ± 0.01	6.17 ± 0.01
PyB-2 (second)	BR	Automatic	6.19 ± 0.01	6.19 ± 0.01
PyB-2 (third)	BR	Automatic	6.19 ± 0.01	6.19 ± 0.01

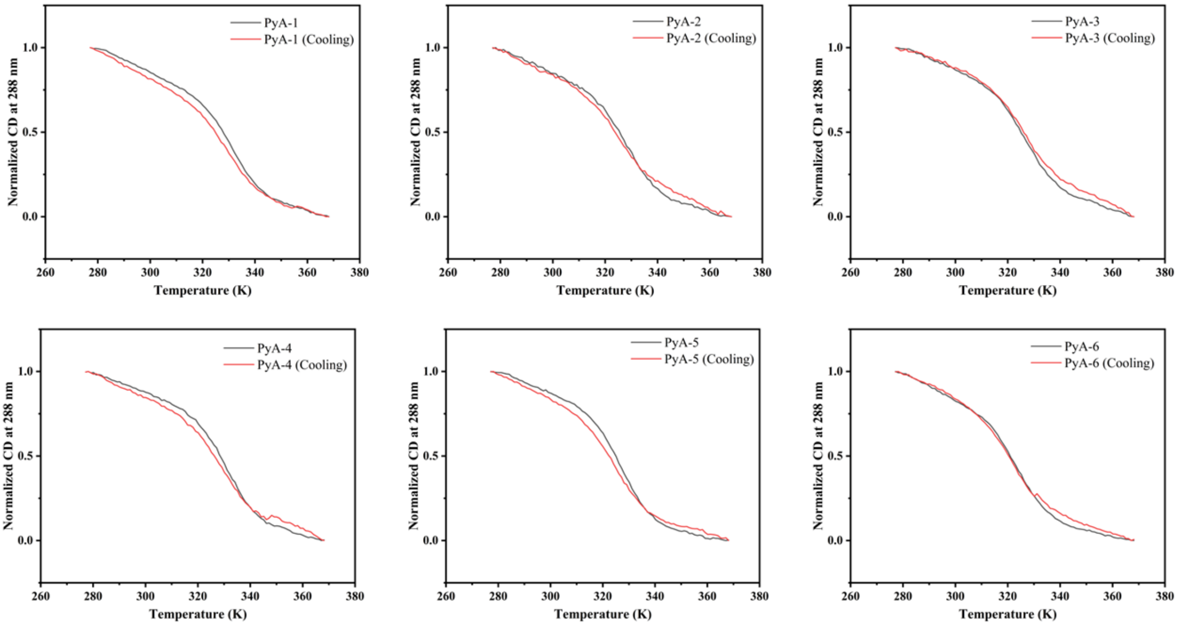


Fig. S25 *CD melting and cooling curves of PyA variants. Experiments were collected from 4 °C (277 K) to 95 °C (368 K) at 288 nm. DNA was prepared at ~4 µM in 10 mM MES buffer pH = 5.5.*

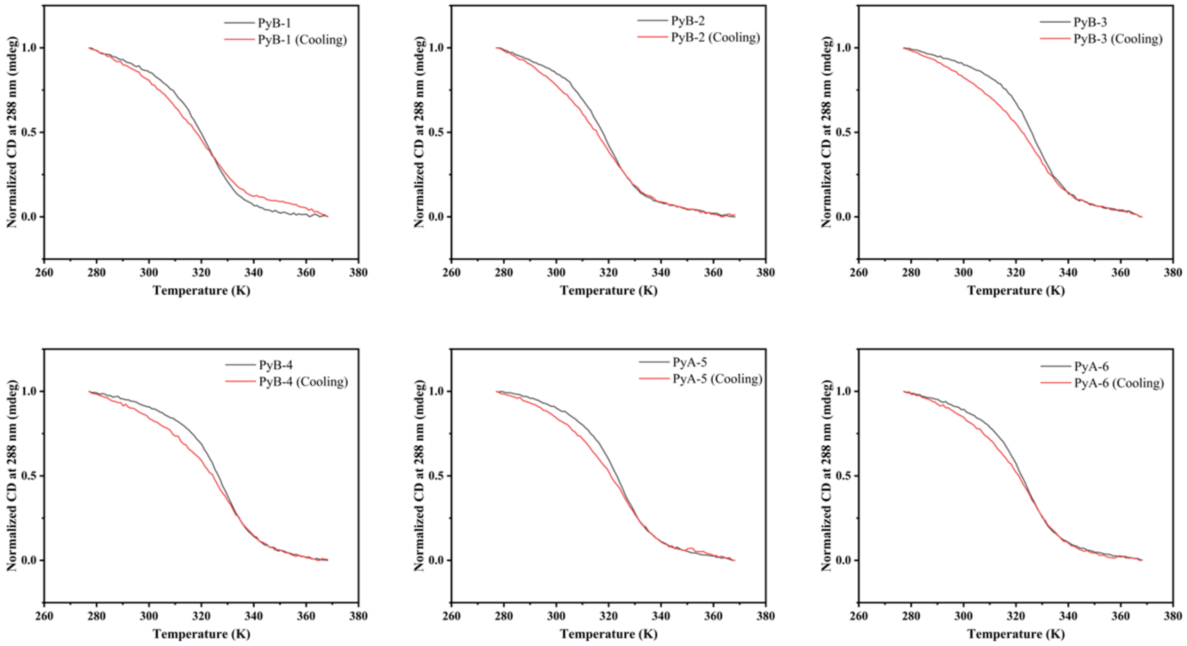
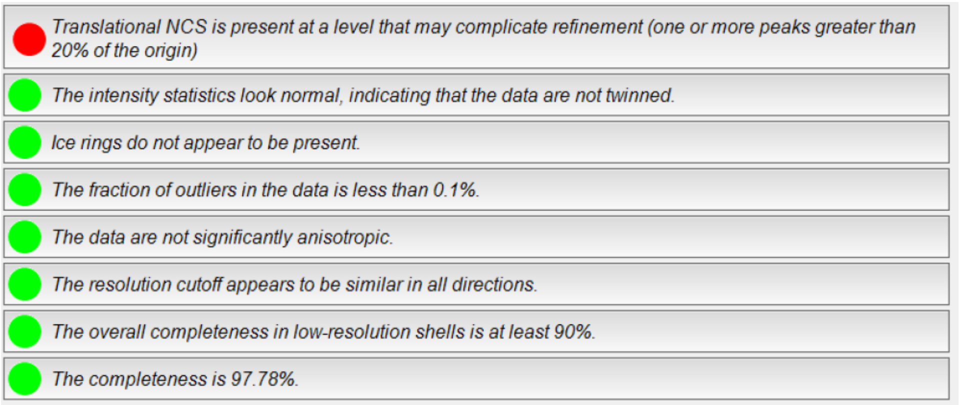


Fig. S26 *CD melting and cooling curves of PyB variants. Experiments were collected from 4 °C (277 K) to 95 °C (368 K) at 288 nm. DNA was prepared at ~4 µM in 10 mM MES buffer pH = 5.5.*



🔴 *Translational NCS is present at a level that may complicate refinement (one or more peaks greater than 20% of the origin)*

🟢 *The intensity statistics look normal, indicating that the data are not twinned.*

🟢 *Ice rings do not appear to be present.*

🟢 *The fraction of outliers in the data is less than 0.1%.*

🟢 *The data are not significantly anisotropic.*

🟢 *The resolution cutoff appears to be similar in all directions.*

🟢 *The overall completeness in low-resolution shells is at least 90%.*

🟢 *The completeness is 97.78%.*

Fig. S27 *Xtriage summary obtained analysing the raw data (free.mtz) of “GT1_C1”.*

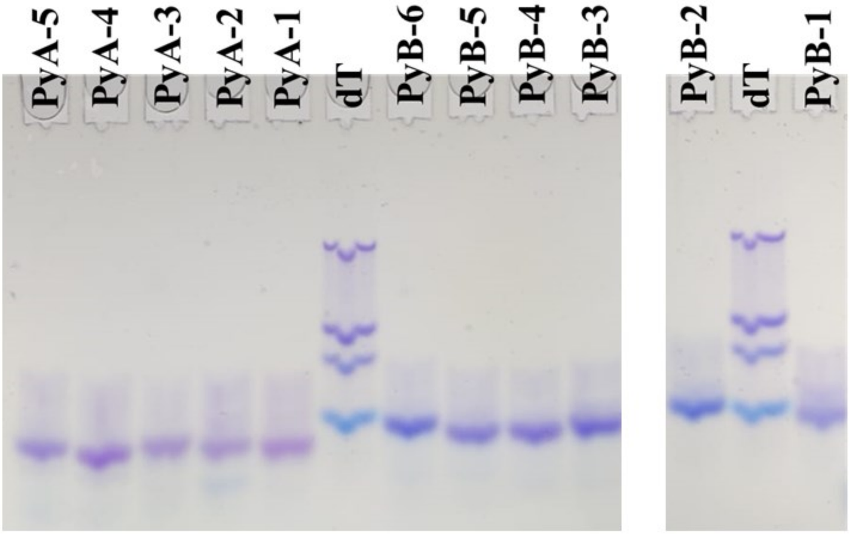


Fig. S28 *Fifteen percent native gel prepared in 10 mM BR buffer pH = 5.5. DNA samples were prepared at ~ 50 µM. The electrophoresis system was connected to a water bath, and the migration was conducted at 10 °C to avoid overheating. Bubbles were formed during the GEL polymerization which affected the shape of the bands.*

Appendix

Instrumental techniques' overview

Uv-Vis Spectroscopy

UV-visible spectrophotometry primarily serves as a quantitative analysis method that focuses on the absorption of near-ultraviolet (180–390 nm) or visible (390–780 nm) radiation by chemical substances either in solution or gas phase. The energy provided by these portions of the electromagnetic spectrum facilitates electronic transitions. However, due to the overlapping of vibrational and rotational transitions, UV-visible spectra of solution-based analytes usually lack detailed structure.

Typically, organic compounds absorb energy in the near-ultraviolet zone owing to transitions like $\sigma \rightarrow \sigma^*$, $n \rightarrow \sigma^*$, $n \rightarrow \pi^*$, and $\pi \rightarrow \pi^*$. Most interesting transitions encompass the promotion of n or π electrons to the π^* excited state, making molecules with delocalized π electrons (such as aromatic and conjugated species) more susceptible to absorption. The section of the molecule responsible for this absorption is referred to as a chromophore. The maximum absorption wavelength of a chemical species can be influenced by the nature of the substituents attached to the aromatic ring or the conjugated system. Substituents that contain lone pairs of electrons, such as –OH and –NH2, often lead to an extended wavelength of maximum absorption and an increased molar absorptivity. These groups are also identified as auxochromes.

Furthermore, specific inorganic compounds can be detected in the visible region due to their natural absorption of radiation. As an example, transition metal ions and their water complexes, which are often coloured, exhibit a decent molar absorptivity thanks to the electronic transitions involving the 3d and 4d orbitals.

With controlled experimental conditions, there is a direct correlation between the absorbed radiation amount and the analyte concentration in solution. This connection is summarized in Beer-Lambert's law and is represented by the formula:

$$A = \varepsilon\, b\, c.$$

In this formula, A stands for the solution's absorbance (no units), ε represents the molar absorptivity (L $\mathrm{mol^{-1}cm^{-1}}$), b signifies the path length of radiation through the absorbing medium (cm), and c indicates the concentration (mol $\mathrm{L^{-1}}$).

Nonetheless, there are limitations, known as deviations, to Beer-Lambert's law. For instance, the law is not applicable to high analyte concentrations (higher than 0.01 mol $\mathrm{L^{-1}}$) due to intermolecular interactions among the analytes. Deviations can be divided into two main categories: instrumental and chemical.

Instrumental deviations occur because the electromagnetic radiation falling on the sample is polychromatic rather than monochromatic (unless laser light is used). This can cause negative deviations from Beer-Lambert's law at higher analyte concentrations and reduced sensitivity as the averaged absorptivity value (ε) is lower than the theoretical one. Other sources of deviation include stray radiation, which arises from scattering, reflection, and refraction effects caused by mirrors, lenses, filters, and other optical components. Moreover, radiation loss can occur due to reflection or refraction at the interfaces between the environment, cuvette, and sample, and from concentration gradients within the sample solution.

Chemical deviations occur due to changes in absorbance resulting from alterations in the sample's chemical environment, such as pH. Absorbance can also be affected by acid-base equilibria, redox reactions, and complexation reactions. At a particular wavelength known as the isosbestic point, the absorbances of all species in equilibrium are equal because they share the same ε value. This point can be used as a reference for kinetic studies, to verify the wavelength accuracy of a spectrophotometer, or as the reference wavelength in dual-wavelength spectrophotometry as it is unaffected by the development of any chemical reaction.

Additionally, positive deviations can be caused by changes in the solvent's nature and refractive index. These should be checked as part of the blank measurement protocol. Positive interferences may also occur due to absorption by other components in the sample matrix or excess reagents, necessitating separation prior to detection, masking by the addition of an appropriate reagent, or minimization through sample dilution and/or lowering the reagent concentration.

Finally, it's crucial to consider the presence of potential interfering species that might react with the primary reagent to produce similarly colored species or lower the effective reagent concentration. Analytical selectivity remains a primary consideration for all applications.

Furthermore, derivative spectrophotometry can be used to process data mathematically after its collection, thereby improving spectral resolution in systems with multiple components. Here, the zero-order absorbance spectrum is derived with respect to the wavelength to yield first-order ($dA/d\lambda$) or higher order plots of absorbance rate change against absorbance. The key benefits of this approach include the amplification of any differences between (similar) spectra and the minimization of any broadband matrix interferences.

The fundamental components of the instrument include a radiation source, a wavelength-selecting device, a sample compartment, a detector, and an output device. For visible radiation, a tungsten filament is the most common source, while a deuterium lamp is typically used for near-ultraviolet radiation. A grating monochromator, used for improved resolution, serves as the wavelength-selecting device and can also be employed for scanning the wavelength range of interest. The sample compartment is typically a cuvette with a 1.0 cm square cross-section, made of quartz (for the near-ultraviolet region), or glass or plastic (for the visible region). Photomultiplier tubes are most commonly used for detection, although phototubes, and increasingly, photodiodes and diode arrays, are also utilized.

The two most frequent optical configurations for UV-Vis spectrophotometers are single-beam and double-beam. Single-beam instruments are typically less expensive and require simpler optics, while double-beam instruments are more suitable for spectral scanning. Additional benefits of the double-beam configuration include the ability to compensate for stray light, fluctuations in the intensity of source radiation and detector response, and the presence of solid particles in the sample solution.

An alternative optical configuration is used in diode array spectrophotometry, where polychromatic light from the source passes through the sample cuvette and is then dispersed by a grating onto a linear diode array detector. Advantages include speed, negligible stray light effects due to the reverse optical configuration, wavelength setting reproducibility because of the absence of moving optical components, and simultaneous multi-wavelength data acquisition. This feature is particularly useful for post-acquisition digital data processing techniques like dual-wavelength spectrophotometry, derivative spectrophotometry, and multivariate calibration.

Fluorescence Spectroscopy

Fluorescence spectroscopy refers to an emission process in which a transition from a high energy state to a lower energy state produces radiation. Fluorescence can only be emitted by molecules in their excited states, which implies they must be excited to a higher state before the emission can occur.

A more comprehensive understanding of this phenomenon can be gained by referring to the diagram (**Fig. 75**), also known as Jablonski's diagram.

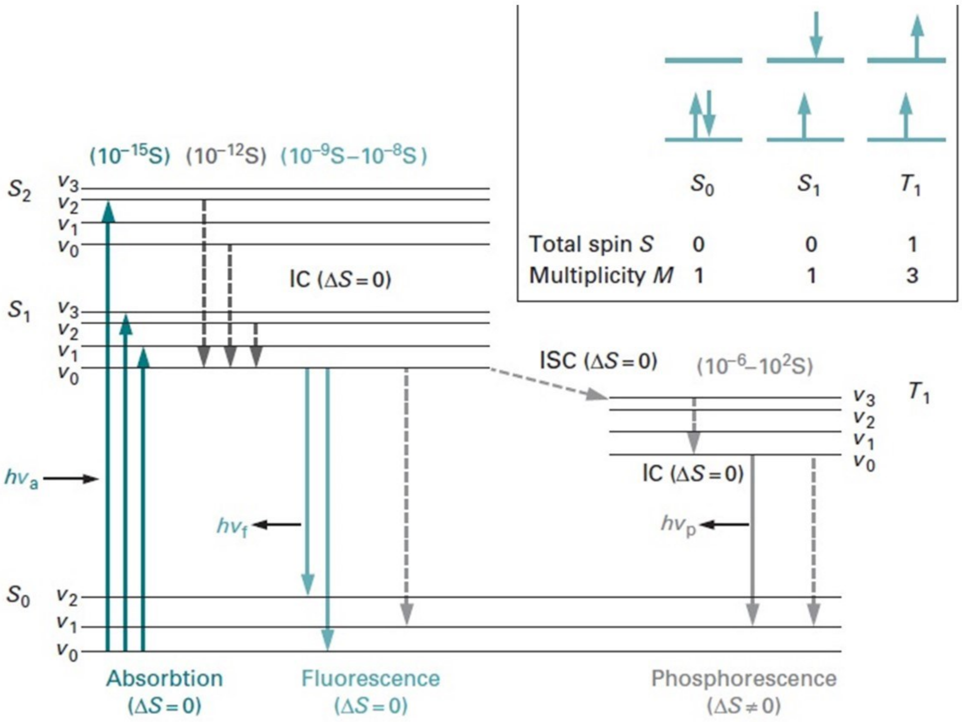


***Fig.** 75 Jablonski's Diagram- In the diagram are shown: the electronic ground state (S0), two excited singlet states (S1, S2) and a triplet state (T1). Vibrational level (V) are only illustrated exemplarily. Solid vertical lines indicate radiative transitions, dotted-lines show non-radiative transitions (i.e. Internal Conversion, IC). The inset shows the relationship between electron configurations, total spin number S and multiplicity M.*

These diagrams depict energy transitions via vertical lines. However, not all transitions are permissible; the selection rules of quantum mechanics dictate which ones are allowed.

A molecule in its electronic and vibrational ground state (S_0V_0) can absorb photons that match the energy difference of its various discrete states. The photon's energy must exceed the energy needed to reach the vibrational base state of the first electronically excited state (S_1V_0). Excess energy is consumed as vibrational energy ($V > 0$) and is rapidly shed as heat through interactions with solvent molecules. Thus, the molecule returns to the vibrational ground state of the first electronic excited state (S_1V_0). These relaxation processes are non-radiative transitions from one energetic state to another with lower energy, referred to as internal conversion (IC). From the lowest level of the first electronically excited state, the molecule reverts to the ground state (S_0), either by emitting light (fluorescence) or through a non-radiative transition.

When a radiative transition occurs, the molecule can land in any of the vibrational states of the electronic ground state (as dictated by quantum mechanical rules). If the vibrational levels of the ground state overlap with those of the electronically excited state, the molecule will not emit fluorescence but will instead revert to the ground state through non-radiative internal conversion. This is the most common way for excitation energy to be dissipated, which is why fluorescent molecules are rather uncommon. Most molecules are flexible, thus having high vibrational levels in the ground state. In fact, most fluorescent molecules possess fairly rigid aromatic rings or ring systems. The fluorescent group in a molecule is known as a fluorophore.

Given that radiative energy is lost in fluorescence compared to absorption, the fluorescent light is always at a longer wavelength than the exciting light (referred to as the Stokes shift). The emitted radiation presents as a band spectrum because there are many closely associated wavelength values dependent on the attained vibrational and rotational energy levels. However, a molecule's fluorescence spectrum is independent of the wavelength of the exciting radiation and exhibits a mirror-image relationship with the absorption spectrum. Additionally, the transition probability from the electronically excited state to the ground state is proportional to the intensity of the emitted light.

In this context, phosphorescence is a related phenomenon that results from the transition from a triplet state (T_1) to the electronic singlet ground state (S_0). The molecule enters the triplet state from an electronically excited singlet state (S_1) via a process known as intersystem crossing (ISC). Notably, the transition from singlet to triplet state is not permitted by quantum mechanics (i.e., occurs with very low probability) and thus only happens in certain molecules where the electronic structure allows it, such as in heavy atoms. As the rate constants for phosphorescence are much longer, phosphorescence occurs with a significant delay and can persist even when the excitation energy is no longer being applied.

A molecule's fluorescence properties are determined by the molecule's own properties (internal factors), as well as environmental influences (external factors). The fluorescence intensity emitted by a molecule depends on the lifetime of the excited state. The transition from the excited state to the ground state can be treated like a first-order decay process, i.e., the number of molecules in the excited state reduces exponentially over time.

The quantum yield Φ is defined as the ratio of photons emitted to photons absorbed by a fluorophore. It is a dimensionless quantity and is the only absolute measure of a molecule's fluorescence. Determining the quantum yield can be challenging and generally involves comparing the fluorescence of the molecule in question with that of a fluorophore with a known quantum yield. This usually involves comparing the fluorescence emissions of two or more similar samples and analyzing their relative differences.

Fluorescence spectroscopy is particularly effective at very low concentrations of emitting fluorophores, a range where UV/Vis spectroscopy is less precise. One reason for the high sensitivity of fluorescence applications is the spectral selectivity provided by the Stokes shift, where the wavelength of the emitted light differs from that of the exciting light. Another advantage is that fluorescence is emitted in all directions. By positioning the detector perpendicular to the excitation pathway, the background of the incident beam is minimized.

In a standard spectrofluorimeter, two monochromators are used. The first monochromator is used to adjust the wavelength of the exciting beam, while the second is used to analyse the fluorescence emission. Since the emitted light always has a lower energy than the exciting light, the wavelength of the excitation monochromator is set at a lower wavelength than the emission monochromator. The more sophisticated fluorescence spectrometers found in laboratories often feature a photon-counting detector for high sensitivity. Temperature control is essential for precise measurements because the emission intensity of a fluorophore is temperature-dependent. The optical geometries with a 90° arrangement are commonly used.

There can also be pre- and post-filter effects that arise due to absorption of light before it reaches the fluorophore and reduction of emitted radiation. These phenomena are known as inner filter effects and are more pronounced in solutions with high concentrations.

Resonance Light Scattering Spectroscopy

Resonance Light Scattering (RLS) spectroscopy is a sensitive and selective method used to extract size and shape information from supramolecular assemblies of chromophores. While traditional light scattering experiments use wavelengths away from absorption bands, wavelengths within the absorption band can provide extremely informative results when the absorption is not too high and the aggregate is large enough.

In RLS spectroscopy, an increased scattering intensity is observed at wavelengths very close to the absorption of a molecular aggregate. This effect can be enhanced by several orders of magnitude when there is strong electronic coupling among the chromophores. Additionally, the wavelength dependence of RLS enables selective observation of aggregates, even in multicomponent systems that include a large fraction of monomers or other aggregates.

Two processes generally occur when light passes through a solution of aggregates: absorption and scattering by the aggregates, assuming the solvent is non-absorbing. The light scattering component arises due to differences in polarizability between the aggregates and the solvent. The incident electromagnetic wave induces an oscillating dipole in the assembly, which then radiates light in all directions.

The *absorption cross section* is the ratio of the rate of energy absorption from the incident beam to the intensity of the incident beam. The *scattering cross section*, on the other hand, is the ratio of the rate of energy scattering out of the incident beam (in all directions) to the intensity of the incident beam. When the induced dipole is ideal (which is often the case if the size of the aggregate is small compared to the wavelength of the light in the solvent), both cross sections are related to the polarizability α of the aggregates. This polarizability can be considered as the sum of its real part α_r (related to light-scattering phenomena) and its imaginary part α_i (related to absorption processes).

Therefore, absorption by a solution of aggregates at a certain wavelength leads to a peak in the imaginary part of the polarizability α_i, causing α_r to behave anomalously, resulting in increased scattering. This increased scattering can be hard to detect under normal conditions due to increased absorption and the weakness of the enhanced scattering effect. However, when investigated for aggregates of chromophores, the enhanced RLS scattering can be high.

The absorption depends on the first power of the polarizability, which in turn depends linearly on the volume of the aggregate. While a solution with a fixed concentration of monomers will not show changes in absorbance, the aggregation process could lead to larger volumes, triggering enhanced scattering. This makes RLS extremely sensitive to even low concentrations of extended aggregates.

Finally, it is worth noting that RLS experiments can often be conducted with conventional equipment, such as a standard double-monochromator fluorimeter in synchronous scan.

Circular Dichroism Spectroscopy

Circular Dichroism (CD) spectroscopy is a powerful tool in stereochemical analysis, with strong connections to chirality, as it is a manifestation of diastereomer discrimination. This technique is sensitive not only to the absolute configuration of a molecule but also to its conformational features, which are often obscured in regular absorption spectroscopy.

CD spectroscopy is based on the principle that plane-polarized light can be conceived as the superposition of a left circularly polarized light wave and a right circularly polarized light wave, both of which have identical amplitudes and wavelengths. This concept is illustrated in the schematic below (**Fig. 76**).

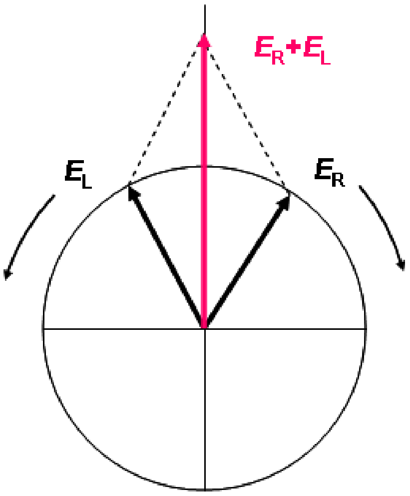


***Fig.** 76 The superposition between the electric field vector's left circularly polarized light E_L, with the electric field vector's right circularly polarized light E_R results in a plane-polarized electric field E.*

Chromophores that are naturally asymmetric, though rare, or symmetric chromophores within asymmetric environments exhibit distinct interactions with right- and left-circularly polarized light, resulting in a pair of interrelated phenomena. Right- and left-circularly polarized light will move at variable speeds through an optically active medium due to the differing refraction indices for the two types of light. This occurrence is known as optical rotation or circular birefringence.

Additionally, right- and left-circularly polarized light will experience varying levels of absorption at certain wavelengths because of the difference in extinction coefficients for the two polarized rays, a phenomenon termed circular dichroism (CD). Formally, this can be expressed as:

$$CD = A_L - A_R$$

where A_L and A_R denote the absorbance of left- and right-polarized light, respectively.

The Lambert-Beer law provides the absorbance, A, which is calculated as:

$$A = \varepsilon\, l\, c$$

where ε represents the molar absorptivity coefficient (dependent on the wavelength), l signifies the optical path length, and c refers to the sample concentration.

In a CD measurement, both l and c remain constant, allowing the CD signal to be defined as a difference in molar absorptivity coefficients:

$$CD = \varepsilon_L - \varepsilon_R.$$

An optically active sample, with different extinction coefficients for right- and left-polarized light, results in a discrepancy in the magnitude of the two components. Consequently, the resultant electric field vector (E-vector) outlines an ellipse over time, not a circle, meaning the light is elliptically polarized.

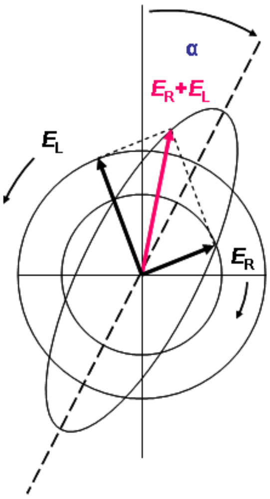


Fig. *77 Illustration of the elliptically polarized light.*

Ellipticity is the unit of CD and is defined as the tangent of the ratio of the minor to major elliptical axis. It is directly proportional to the difference in molar absorptivity coefficients of the two light components, given by the equation:

$$\theta = \frac{\pi\lambda(\varepsilon L - \varepsilon R)}{l}$$

where l is the optical path and λ is the wavelength of the incident radiation.

A CD spectrum, which is a measure of the ellipticity versus the wavelength, can be expressed in molar ellipticity ([θ]) to facilitate the comparison of spectra obtained at different concentrations:

$$[\theta] = \frac{\theta}{[c]\cdot l} = 32.98\ \Delta\varepsilon \text{ degrees } M^{-1}\ cm^{-1}$$

This spectrum is recorded in the UV-Vis wavelength region and bears resemblance to the absorption spectrum, but it can have both positive and negative bands since it represents a difference between molar absorptivity coefficients.

Chromophores sensitive to CD are inherently dissymmetric molecules. This dissymmetry can be structural or a result of the presence of a chiral center. CD is applicable only to samples that absorb in the UV-Vis region.

The standard equipment for measuring Circular Dichroism (CD) is a traditional spectropolarimeter. In the most frequent arrangement, a photoelastic modulator (PEM) is situated between the monochromator and the sample compartment, supplying alternating left and right circularly polarized light.

This modulator operates on the basis of a piezoelectric quartz crystal, typically oscillating at a frequency around 50 kHz. Throughout each cycle, the polarization of the light changes, while the intensity remains unchanged. When this light passes through a chiral, non-racemic sample, the two circularly polarized light components are absorbed to different degrees, resulting in a time-modulated light reaching the detector at the same frequency as the PEM. This time-modulated signal can be amplified using phase-locked amplification to simultaneously provide absorbance and circular dichroism measurements.

In theory, CD should not need a baseline correction as the baseline is expected to be 0. However, in practice, it is highly recommended to perform a baseline subtraction correction on any CD spectrum. This correction is ideally obtained from a measurement of the racemic compound under the same conditions (concentration and cell). Alternatively, using a blank of the same solvent typically yields satisfactory results.

Electrophoresis

Electrophoresis is a universal term for the movement and separation of charged particles under the influence of an electric field. Many biological molecules (such as peptides, proteins, and nucleic acids) have ionizable groups and thus, they can exist as charged species in a solution, depending on the pH.

Gel electrophoresis is a frequently utilized method for the separation, purification, and identification of ionic or ionizable macromolecules in a solution. An electrophoretic system consists of two electrodes - an anode and a cathode - which are linked by a conductive medium known as an electrolyte. The electrophoresis gel is submerged within this electrolyte. The sample is loaded into specific wells, and the species move towards either electrode based on their charge.

The electric force (F_e) can be calculated with the equation:

$$F_e = E\,q$$

where E represents the electric field and q denotes the charge.

There is a frictional force (F_f) that counteracts the motion created by the electric field. This force can be calculated using the formula:

$$F_f = \gamma\, v$$

where γ is the friction coefficient and v is the velocity of migration.

At a constant speed, the electric force (F_e) equals the friction force (F_f), which leads us to the equation:

$$v = q / \gamma$$

Here, v is the migration velocity. The electrophoretic mobility (m) is defined as the ratio of the velocity of a molecule to the electric field (E), which gives us:

$$m = v / E = q / (\gamma E)$$

Thus, the electrophoretic mobility only depends on the properties of the particle.

In the case of a spherical molecule, the friction coefficient (γ) can be given by the Stokes equation:

$$\gamma = 6\pi\eta r$$

Here, η represents the viscosity of the medium, and r stands for the radius of the molecule. Therefore, the electrophoretic mobility (m) can be expressed as:

$$m = q / (6\pi\eta r E)$$

In conclusion, the electrophoretic mobility of a molecule is influenced by its charge, size, and shape, as well as the viscosity of the medium and the size of the pores in the gel. Electrophoresis allows for the separation of molecules in a mixture based on their charge and size.

X-Ray Crystallography

X-ray Crystallography is a scientific method used to determine the arrangement of atoms of a crystalline solid in three dimensional space. This technique takes advantage of the interatomic spacing of most crystalline solids by employing them as a diffraction gradient for x-ray light, which has wavelengths on the order of 1 angstrom.

The arrangement of the atoms needs to be in an ordered, periodic structure in order for them to diffract the x-ray beams. A series of mathematical calculations is then used to produce a diffraction pattern that is characteristic to the particular arrangement of atoms in that crystal. The scattered monochromatic x-rays that are in phase give constructive interference, according to Bragg's law:

$$n\lambda = 2d \sin \theta$$

where n is an integer called the order of reflection, λ is the wavelength of x-rays, d is the characteristic spacing between the crystal planes of a given specimen and θ is the angle between the incident beam and the normal to the reflecting lattice plane. This law relates the wavelength of electromagnetic radiation to the diffraction angle and the lattice spacing in a crystalline sample.

In a single-crystal X-ray diffraction measurement, a crystal is mounted on a goniometer, which is used to position the crystal at selected orientations. The crystal is illuminated with a finely focused monochromatic beam of X-rays, leading to a diffraction pattern of regularly spaced spots known as reflections. X-ray crystallography works in a manner of elastic scattering with the outgoing X-rays having the same energy and wavelength as the incoming X-rays, which get an altered direction after diffraction. A crystallographer can then produce a three-dimensional picture of the density of electrons within the crystal by measuring the angles and intensities of these diffracted beams under the assistance of the mathematical method Fourier transforms. From this electron density, the mean positions of the atoms, chemical bonds, crystallographic disorder, and some other information in the crystal can be determined. Poor resolution or even errors may occur if the crystals are too small, or not uniform enough in their internal makeup.

The technique of single crystal X-ray crystallography has three basic steps. The first and usually most difficult step is to produce an adequate crystal of the studied material. The crystal should be sufficiently large with all dimensions larger than 0.1 mm, pure in composition and regular in structure, and have no significant internal imperfections such as cracks or twinning. The crystal is subsequently placed in an intense beam of X-rays, usually of a single wavelength, to produce regular reflection pattern. The angles and intensities of diffracted X-rays are measured with each compound having a unique diffraction pattern. Previous reflections disappear and new ones appear along with the gradual rotation of the crystal, and the intensity of every spot is recorded at every orientation of the crystal. Multiple data sets may have to be collected since each set covers slightly more than half a full rotation of the crystal and typically contains tens of thousands of reflections. Ultimately, these collected data are combined computationally with complementary chemical information to obtain and refine a model from the arrangement of atoms within the crystal. The final refined model of the atomic arrangement is called a crystal structure and usually stored in a public database.

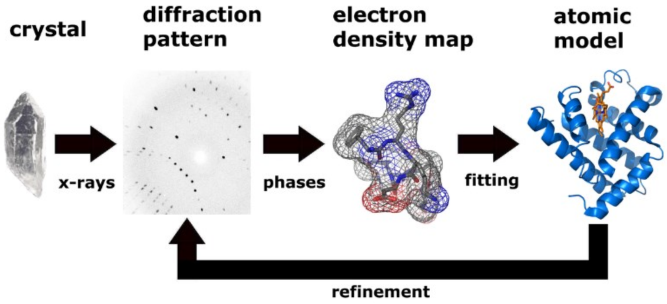


***Fig.* 78** *Workflow for solving the molecular structure by X-ray crystallography.*

List of Publications and Proceedings

-Publications

1. Interactions between achiral porphyrins and a mature miRNA.

Travagliante, G.; Gaeta, M.; Gangemi, C.M.A.; Purrello, R.; D'Urso, A.

Nanoscale. Submitted

2. Higher-Order G-Quadruplex Structures and Porphyrin Ligands: Towards a Non-Ambiguous Relationship.

Falanga, A.P.; D'Urso, A.; **Travagliante, G.**; Gangemi, C.M.A.; Marzano, M.; D'Errico, S.; Terracciano, M.; Greco, F.; De Stefano, L.; Dardano, P.; Rea, I.; Piccialli, G.; Oliviero, G.; Borbone, N.

Anal. Chim. Acta. Submitted

3. Chiral Multi-walled Nanotubes of H4TPPS4 J-aggregates are induced by Amino Acid templates.

Gaeta, M.; Kosta, A.; Randazzo, R.; **Travagliante, G**.; Berriman, J.A.; Purrello, R.; D'Urso, A. J. Am. Chem. Soc. Submitted

4. Effect of the Buffer Composition on the Aggregation State of the Zinc(II) Tetra-Spermine Porphyrin Derivative.

Travagliante, G.; Gaeta, M.; Gangemi, C.M.A.; Purrello, R.; D'Urso, A.

J Porphyr Phthalocyanines **2023**, 27, 509. https://doi.org/10.1142/S108842462350027X

5. Supramolecular Chirality in Porphyrin Self-Assembly Systems in Aqueous Solution.

Travagliante, G.; Gaeta, M.; Purrello, R.; D'Urso, A.

Curr Org Chem **2022**, *26*, 6. http://dx.doi.org/10.2174/1385272826666220330112648

6. Recognition and Sensing of Chiral Organic Molecules by Chiral Porphyrinoids: A Review.

Travagliante, G.; Gaeta, M.; Purrello, R.; D'Urso, A.

Chemosensors **2021**, *9*, 204. https://doi.org/10.3390/chemosensors9080204

Proceedings – oral communications

1. Interactions between achiral porphyrins and a mature miRNA.

Convengo Nazionale della Divisione di Chimica dei Sistemi Biologici 2022 (DCSB), Napoli, 20th-22nd June, 2022.

***Gabriele Travagliante**, Massimiliano Gaeta, Roberto Purrello, and Alessandro D'Urso*

2. Different strategies to modulate the thermodynamic stability of micro-RNAs.

Società Chimica Italiana Convegno Regionale della Sezione Sicilia 2021, Catania, December 2nd, 2021.

***Gabriele Travagliante**, Massimiliano Gaeta, Salvatore Alaimo, Alfredo Ferro, Roberto Purrello, and Alessandro D'Urso*

3. The thermodynamic stability of miRNAs might influence their biological activity. Merck Young Chemists' Symposium 2021, Rimini, November 22nd-24th, 2021.

***Gabriele Travagliante**, Massimiliano Gaeta, Salvatore Alaimo, Alfredo Ferro, Roberto Purrello, and Alessandro D'Urso*

4. Spectroscopic study on interactions of porphyrins and micro-RNA. XXVII Congresso Nazionale della Società Chimica Italiana - Divisione di Chimica dei Sistemi Biologici, September 14th-23rd, 2021.

***Gabriele Travagliante**, Massimiliano Gaeta, Roberto Purrello, and Alessandro D'Urso*

Acknowledgements

I would like to express my heartfelt gratitude to several people who have played an important role in my academic journey.

First of all, I extend my sincere appreciation to my supervisor, Prof. Alessandro D'Urso, for his solid support, valuable guidance and unwavering belief in my abilities. His guidance has been instrumental in shaping my research and personal growth.

I am deeply grateful to the entire LSSC group, especially Prof. Roberto Purrello and PhD Massimiliano Gaeta, for their collaborative spirit, valuable insights and the lively intellectual environment they provided me. Your contributions have enriched my research experience.

During my time abroad in the United States, I had the privilege of working under the guidance of Professor Liliya Yatsunyk. I would like to express my sincere appreciation for her guidance, which broadened my horizons and added an international perspective to my research.

I would also like to acknowledge the support and encouragement of my family, friends and colleagues during this challenging journey. Your trust in me has been a constant source of motivation.

Finally, I would like to express my gratitude to all those who have supported me in both visible and invisible ways, and to the wider academic community for fostering an environment of learning and discovery.

Thank you all for being an integral part of my academic and personal growth. Your contributions have been invaluable and I am deeply grateful for your support.