Synthesis and coordination chemistry of tetradentate chelators based on ligand-appended G-quadruplex structures

Dissertation
zur Erlangung des mathematisch-naturwissenschaftlichen Doktorgrades
"Doctor rerum naturalium"
der Georg-August-Universität Göttingen
im Promotionsprogram IRTG 1422
Metal Sites in Biomolecules – Structures, Regulation and Mechanisms
der Georg-August University School of Science (GAUSS)

vorgelegt von

David Maximilian Engelhard

aus Essen

Göttingen 2016

Thesis committee

Prof. Dr. Guido H. Clever (supervisor),
Institute for inorganic chemistry, Georg-August University Göttingen, now TU Dortmund University

Prof. Dr. Ulf Diederichsen,
Institute for organic and biomolecular chemistry, Georg-August University Göttingen

Prof. Dr. Derek T. Logan,
Department of Biochemistry and Structural Biology, Lund University

Members of the examination board

Prof. Dr. Guido H. Clever (supervisor and assessor),
Institute for inorganic chemistry, Georg-August University Göttingen, now TU Dortmund University

Prof. Dr. Claudia Höbartner (assessor),
Institute for organic and biomolecular chemistry, Georg-August University Göttingen

Prof. Dr. Franc Meyer,
Institute for inorganic chemistry, Georg-August University Göttingen

Dr. Inke Siewert,
Institute for inorganic chemistry, Georg-August University Göttingen

Dr. Franziska Thomas,
Institute for organic and biomolecular chemistry, Georg-August University Göttingen

Prof. Dr. Kai Tittmann,
Department for Bioanalytics, Albrecht-von-Haller-Institute for Plant Sciences, Georg-August University Göttingen

Date of the oral examination

14.01.2016

Affidavit

I hereby declare that my doctoral thesis entitled “Synthesis and coordination chemistry of tetradentate chelators based on ligand-appended G-quadruplex structures” has been written independently and with no other sources and aids then quoted. I have indicated the parts which were performed by project collaborators.

David Engelhard
Göttingen, December 2016

"*Ja, ja, die Musik*"
(Loriot, frei nach T. Adorno)

Acknowledgements

Mein allererster Dank geht an Prof. Dr. Guido Clever für die Möglichkeit in seiner Arbeitsgruppe an einem spannenden Promotionsprojekt zu arbeiten und so vielfältige Einblicke in die verschiedensten Themenfelder zu bekommen. Auch wenn das hieß die kleinen Kämpfe mit Linux, Mac, und Windows führen zu müssen. Vielen Dank auch für eine gute und intensive Betreuung, die dennoch auch Forschungsspielraum zuließ.

Ein besonderer Dank geht an Prof. Dr. Claudia Höbartner für die Übernahme des Zweitgutachtens, sowie den weiteren Mitgliedern der Prüfungskommission Prof. Dr. Franc Meyer, Dr. Franziska Thomas, Dr. Inke Siewert, und Prof. Dr. Kai Tittmann. Ebenso sei den weiteren Mitgliedern des Betreuungsausschusses Prof. Dr. Ulf Diederichsen und Prof. Dr. Derek Logan gedankt.

Bei Prof. Dr. Ulf Diederichsen möchte ich mich zusätzlich für die Benutzung des UV-VIS und CD Spektrometers, DNA Synthesizers, und der HPLC bedanken. Ohne Zugang dazu hätte ich mein Promotionsprojekt nicht durchführen können.

Eine gute wissenschaftliche Arbeit kommt nicht ohne kompetente Kooperationspartner aus. Deshalb vielen Dank an Dr. Roberta Pievo und Prof. Dr. Marina Bennati (MPI Göttingen) für die fruchtbare Zusammenarbeit bei EPR Messungen. Auch hier schon ein Dank für die fortlaufende Kooperation bezüglich EPR mit der Gruppe von Prof. Dr. Olav Schiemann (Bonn).

Für die Unterstützung bei den kleinen und großen Aufgaben im täglichen Laborbetrieb möchte ich mich bei den verschiedenen Serviceabteilungen und Werkstätten bedanken, insbesondere dort Dr. Michael John und Ralf Schöne (NMR) und Dr. Holm Frauendorf (Massenspetrometrie), sowie Matthias Hesse (Chemikalienausgabe). Auch die Hausmeister standen jederzeit mit Rat und Tat, inklusive Fahrradreparatur, zur Seite.

Vielen Dank an alle vom Clever Lab, insbesondere natürlich die Crew des Dungeon Labs: Fernanda, Marcel, und Susanne, durch die die Arbeit im Labor nicht nur angenehm, sondern auch abwechslungsreich

war. Ein herzlicher Dank für die schöne gemeinsame Zeit geht auch an die anderen Mitstreiterinnen und Mitstreiter Marina, Muxin, Thorben, und all die anderen Jetzigen und Ehemaligen des Clever-Labs 1.0, 2.0, und 3.0.

Ich möchte mich auch ganz herzlich bei meiner gesamten Familie und meinen Freunden bedanken, ihr habt mich unterstützt, mit mir Musik gemacht, mit mir spannende Momente in und außerhalb der Uni erlebt und mir dadurch insgesamt eine schöne Zeit in Göttingen beschert. Ein großer Dank gebührt Anne und Johanna für die extrem schnelle Korrektur meiner Arbeit, trotz eurer vielen anderweitigen Verpflichtungen.

Contents

1. Introduction

Since the identification of deoxyribonucleic acid (DNA) as the source of genetic information storage, scientists have strived to understand the close relationship between the structure and exact function of DNA, as well as the various interactions with other biomolecules, like proteins, within the cell's nucleus. An initial hallmark along this pathway was the discovery of the right-handed, double-helical nature of DNA as the predominant DNA secondary structure in living organisms, evidenced by James Watson, Francis Crick, Maurice Wilkens and Rosalind Franklin. [1] As only four different nucleobases are used to program the genetic code for the plethora of proteins which are necessary for the cellular machinery, each genome of an organism consists of a tremendous amount of these repeating units. So, not surprisingly, DNA does not exist as long rod-like wires but is packed instead into more condensed structures: nucleosomes, chromatin and ultimately the chromosomes. [1] Since the first discovery of DNA's native secondary structure, it could be demonstrated that DNA is neither simply a static storage of information nor is it invariable, which in fact would prevent evolution. The alterability of DNA was demonstrated by the 2015 noble prize laureates in chemistry Thomas Lindahl, Paul Modrich, Aziz Sancar, who were recognized for their mechanistic studies of DNA repair. [2]

Apart from the canonical duplex B-DNA, additional biologically relevant secondary structures have been identified in the last decades, ranging from alterations of double-helical DNA (hairpin, cruciform, A-, and Z-DNA), to DNA with more than two interacting strands, or strand sections, like triplex DNA, three- and four-way junctions, i-motif, and G-quadruplexes. It should be noted that the usually single-stranded RNA is known to be structurally even more variable and is able to

adopt a great number of complex secondary structures. Among the DNA topologies, especially G-quadruplexes have gained an increasing interest, as their formation *in vivo* is believed to play a crucial role for example in (onco)gene expression, through regulation of promoter regions, and the elongation of the telomers, DNA's end marker and safety region against loss of genetic information. The common structural motif of G-quadruplexes is formed from guanine-rich regions of DNA strands, resulting in guanine-quartets, which stack on top of each other (see chapter 4.1 for more details).[3] It is easily understandable that the formation of this secondary structure will influence numerous DNA related biological mechanisms, e. g. DNA–protein interactions. As an example may serve the single-stranded overhangs of DNA, which are recognized by the enzyme telomerase only when the DNA binds to an RNA template within the enzyme's active site (see figure 1.1). G-quadruplex formation prevents these specific interactions, thereby inhibiting the enzyme.[4]

Despite, or indeed for the very reason, of all the scientific efforts made in the field of DNA–protein/enzyme recognition and structure-activity correlation, controlling the exact 3D folding of DNA and also RNA into specific secondary structures, as well as the identification and characterization of these topologies remains a challenge.

Parallel to the research on the purely biological role, the unique folding behaviour and intrinsic properties of DNA have sparked interest in the field of DNA nanotechnology, being a part of supramolecular chemistry. As scientists discovered that biological activity arises mainly from non-covalent interactions, like dative bonds, hydrogen bonds, ion-dipole interactions, and π–π stacking, supramolecular scientists not only try to understand the nature of these interactions but use them to mimic biochemical processes or to design (supra)molecules with completely new properties. Spanning multiple disciplines, applications for such systems range from sensors, catalytically active species, switches, to compounds for clinical therapies.[6]

Concerning the study of DNA's biological and also its supramolecular role, the chemistry of metal ions takes an important portion of this

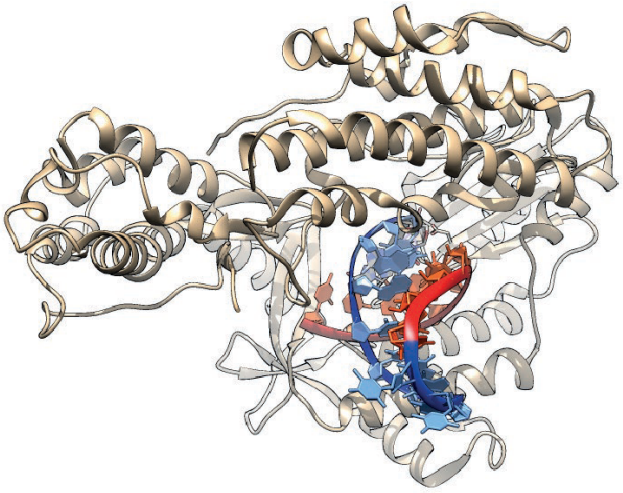


Figure 1.1.: Crystal structure (PDB 3KYL) of the active *Tribolium castaneum* telomerase catalytic subunit, TERT, bound to an RNA-DNA hairpin designed to resemble the putative RNA-templating region and telomeric DNA.[5] The RNA part is highlighted in red, while the DNA part is coloured blue.

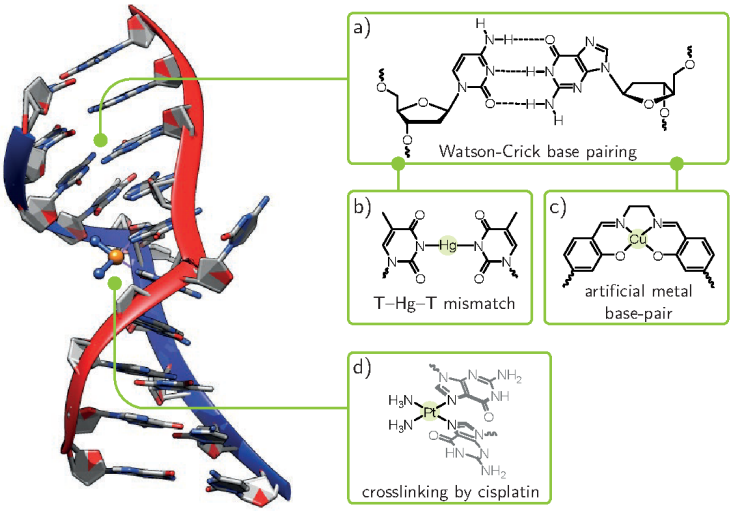


Figure 1.2.: Illustrative examples for a) native Watson-Crick base-pairing (here G≡C, A=T not shown), b) mercury induced thymine-thymine mismatch base-pair, c) artificial metal base-pair composed of a salen ligand covalently bridging between the two DNA strands and Cu^{2+} coordination [7], d) intra-strand guanine-guanine cross-linking induced by the chemotherapy drug cisplatin. The shown crystal structure was taken from PDB entry 1A2E ref. [8].

research, due to the fact that they play a crucial role in many biological processes. This role ranges from alkali metals as ubiquitous ions producing membrane potentials essential for fuelling the cell's machinery to transition metal ions as essential ingredients for the activity and function of numerous proteins, e. g. iron in haemoglobin.[9] In terms of the interaction with DNA (but not solely limited to), metal ions can also have a detrimental role, which is yet sometimes intended in medical treatments. A famous example for this is the chemotherapy drug *cis*-diamminedichloroplatinum(II) (cisplatin), which cross-links two guanines by binding to the *N*7 positions, thereby leading to a kink in the DNA (see figure 1.2d). As a result, cancer cell growth and division is inhibited.[9]

Moreover, extensive research has been spent on modifying the chemical composition of DNA, either by changes of the phosphate backbone and the sugar moieties, or the nucleobases themselves, in order to gain control over stability and functionality.[10] This led to the invention of the so called metal base-pairing concept, in which the hydrogen bond interplay between the natural nucleobases is exchanged for the coordination to transition metal ions (see figure 1.2b and c; for more details, see chapter 4.2.3).[11] The focus of researchers until now, however, has been on alterations of duplex DNA, namely exchanging one or several nucleobases on one DNA strand and its counterpart on the complementary strand. This somewhat limits the coordination environments obtainable to linear arrangements or systems with multidentate ligands.

Based on these considerations, the goal of this thesis is to combine the field of metal base-pairing to the DNA secondary structure of G-quadruplexes and its unique features and topologies. More precisely, the aim is to exchange several natural nucleobases in the strand(s) of a given G-quadruplex topology for a ligand, the latter designed to bind to a transition metal ion like Cu^{2+} or Ni^{2+}. For simplification the whole sugar moiety of the nucleotide is omitted, so the ligand's donor functionality is appended, via a covalent linker, directly to the rest of the G-quadruplex forming oligonucleotide strand. Only the natural phosphate backbone is retained. The task is then to find a

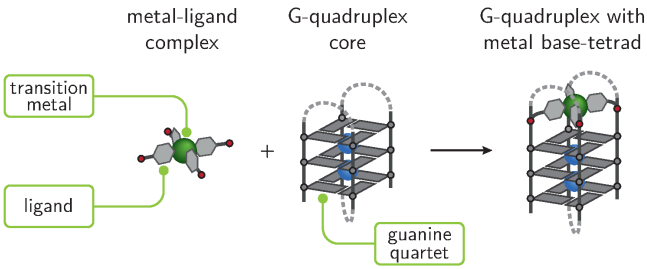


Figure 1.3.: Schematic construction of the G-quadruplex/metal-complex hybrid systems anticipated in this thesis. The dotted lines indicate exemplary strand connecting loops.

suitable coordination environment (metal and ligand) which fits into the fourfold symmetry of the G-quadruplex core, e.g. square-planar, octahedral, or similar coordination environments [12]. Based on this a "proof-of-principle" system should be constructed, demonstrating the successful metal ion incorporation into the G-quadruplex scaffold by full characterization based on ultraviolet absorption spectroscopy, circular dichroism spectroscopy, gel electrophoresis, electron paramagnetic resonance and molecular modelling. Once established, tolerance of the assembly towards ligand modifications should be tested and more complicated topologies investigated.

2. Objectives

The goal of this thesis is to construct DNA G-quadruplexes carrying a "metal base-tetrad", analogous to the metal base-pairing concept that was developed previously for duplex DNA. To achieve this, natural nucleobases in a G-quadruplex forming sequence are exchanged for artificial ligands. These ligands are composed of a donor functionality that is capable of binding to a transition metal ion, and a linker, which provides the necessary spatial separation between the G-quadruplex backbone and the metal coordination site. Based on this overall goal, three main objectives can be outlined for this thesis:

1. **Proof-of principle**: Pyridine ligand synthesis and incorporation into G-rich DNA, and subsequent monitoring of tetramolecular G-quadruplex formation and metal binding.

2. **Ligand variations**: Synthesis of alternative ligands, including variation of linker lengths and ligand backbone, and testing of their implications on formation and metal binding abilities of tetramolecular G-quadruplexes. Usage of the metal base-tetrads as spin labels.

3. **Unimolecular G-quadruplexes**: Control over stability and topology of unimolecular G-quadruplexes carrying the metal base-tetrad.

The first objective will serve as a "proof-of-principle case", with which the general feasibility of the metal base-tetrad concept can be established. The initial simple ligand design should comprise a pyridine donor moiety, an alkyl linker, and a $3'$ phosphate linkage to the next DNA nucleotide. The ligand will be introduced into guanine-rich

DNA oligonucleotide strands via the established phosphoramidite-based solid-state DNA synthesis. Initially, the natural DNA part of the oligonucleotide will just comprise a short guanine stretch $d(G_n)$, so that the resulting oligonucleotide $\mathbf{L}d(G_n)$ will be able to form tetramolecular G-quadruplexes $[\mathbf{L}d(G_n)]_4$. Once suitable conditions for G-quadruplex formation are found, the construct should be tested for the ability to bind transition metal ions like Cu^{2+} or Ni^{2+}. Methods for characterisation will include firstly UV-VIS based thermal denaturation experiments, to assess the thermal stabilities of the investigated G-quadruplexes with and without bound transition metal, and secondly circular dichroism spectroscopy, to probe the topologies of the folded oligonucleotides.

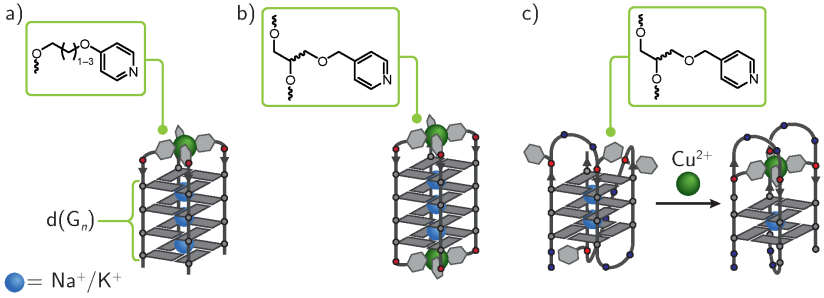


Figure 2.1.: Three illustrative G-quadruplex/metal-complex hybrid systems anticipated in this thesis, corresponding to the three objectives outlined in the text. a) proof-of-principle system with a metal base-tetrad attached at the 5′ end of a tetramolecular G-quadruplex and linker length variation; b) 2nd-generation ligand in a tetramolecular G-quadruplex with two metal base-tetrads; c) unimolecular G-quadruplex with topological switch upon formation of the the metal base-tetrad.

Once the proof-of-principle system is established, the second objective will be to test, if the metal base-tetrad based on pyridine ligands tolerates alterations in the linker length. A comparison of varying alkyl chains as linkers regarding thermodynamic stability of the respective tetramolecular G-quadruplexes should be performed, including the effect of transition metal ion binding. Furthermore, a ligand with a branched

linker should be constructed, which is able to bridge between two nucleotides. The so designed ligand can then be inserted at any position (5′, 3′, or internal) within the DNA oligonucleotide strand. This will enable the construction of G-quadruplexes with more than one instance of the metal base-tetrad, e. g. at opposite or same sides of the tetramolecular G-quadruplex. The two metal base-tetrads could then also be used as spin labels to measure their spatial separation. Moreover, metal-related functionalities could be addressed, like magnetism, redox or catalytic abilities. Thus, this objective will demonstrate the versatility of the metal base-tetrad and the ability to easily adapt this concept to a specific ligand design and application.

The third objective is to use the second-generation ligand for the synthesis of unimolecular G-quadruplexes with one metal base-tetrad, the latter substituting either a whole G-quartet or some of the loop nucleotides, which are necessary for connecting the four strand sections involved in the G-quadruplex assembly. A further advancement will be to use the metal base-tetrad motif to invoke a topological change in a unimolecular G-quadruplex. These results will show that the metal base-tetrad can be used to not only introduce structure stabilisation but also functionality into more complicated G-quadruplex topologies.

3. Summary

In this thesis, the concept of metal base-pairing was adapted for usage with the DNA secondary structure of G-quadruplexes. A "metal base-tetrad" motif was established and incorporated into different tetramolecular and unimolecular G-quadruplex topologies. The metal base-tetrad consists of four identical monodentate pyridine ligands, each covalently appended via an alkyl linker to a guanine-rich oligonucleotide strand, and a transition metal ion coordinating the ligands. Formation of the metal base-tetrad stabilises the respective G-quadruplex topology and is also able to trigger G-quadruplex formation.

Three main results have been achieved. First, a simple but easily adaptable ligand synthesis was developed and the successful ligand incorporation into guanine-rich oligonucleotides was demonstrated. Second, stabilisation of tetramolecular G-quadruplexes, with even more than one instance of the metal base-tetrad could be achieved. Third, stabilisation and topology switching induced by transition metal ion coordination was shown for unimolecular G-quadruplexes.

The first-generation ligand was synthesised in two steps only before usage in standard solid-phase oligonucleotide synthesis. Despite the simple design, the ligand nevertheless allowed for subtle variations, which was demonstrated for different linker lengths between the pyridine donor functionality and the connection point to the DNA backbone. A second-generation ligand with a glycol backbone was successfully obtained, allowing the incorporation at any possible position and of up to four instances of the ligand inside the given guanine-rich oligonucleotide. The ligand has an intrinsic chiral nature, so the effect of the two ligand enantiomers on the stability and folding of the respective ligand-appended G-quadruplex structures was tested.

Appending one ligand to each of the four oligonucleotide strands in tetramolecular G-quadruplex assemblies resulted in a preorganised ligand environment, suitable for square-planar, or octahedral coordination of a Cu^{2+} or Ni^{2+} ion. The transition metal ion coordination not only stabilised a preformed G-quadruplex, but was also able to trigger G-quadruplex assembly from the DNA single-strands, although the renaturation rates were still found to be very slow. Furthermore, the stabilising effect could be reversed by removal of the Cu^{2+} ion with a competitive chelating agent. Variation of the ligand's linker length was well tolerated by the G-quadruplex stem and did not impair formation of the metal-base tetrad. The second-generation ligand enabled the construction of G-quadruplex topologies with two metal base-tetrads. Depending on their position at the same or at different sides of the G-quadruplex core positions, the binding event of a second Cu^{2+} ion was shown to be either detrimental or beneficial for the G-quadruplex stability.

The robustness of the metal base-tetrad, even when inserted into more complicated G-quadruplex topologies, could be demonstrated by using unimolecular G-quadruplexes with four covalently attached ligands and copper as the coordinating transition metal ion. Formation of the metal base-tetrad not only highly stabilised the unimolecular human telomeric G-quadruplex, but was also able to trigger its formation directly from the single-strand. The G-quadruplex folding process was shown to be remarkably fast, and, by removing the metal ion with a chelating agent, to be easily reversible. Moreover, it could be shown that the metal base-tetrad not only stabilised a unimolecular G-quadruplex topology but was also able to induce a topology switching from a mixture of conformers to a stable topology in a G-quadruplex based on the *tetrahymena* telomeric repeat.

All in all, the results depicted in this thesis represent major achievements in the field of G-quadruplex (bio-) nanotechnology, by introducing a novel metal base-tetrad, which could be used to increase the stability, trigger the formation, and switch the topology of tetramolecular or unimolecular G-quadruplexes.

4. State of the art and methods

In this chapter the different research fields important for the results depicted in this thesis will be introduced. This includes the state of the art in DNA secondary structure investigations (section 4.1), DNA nanotechnology with metal-base pairing (section 4.2), as well as short theoretical backgrounds on the most important experimental methods used in this work (section 4.3). Considering the scope of this thesis, a focus on G-quadruplexes will be made.

4.1. DNA secondary structures

Naturally occurring DNA single strands, in their primary structure, are constructed from four different nucleotides, which themselves are composed of three main building blocks. Namely a sugar (ribose) moiety, connected to the next one via a negatively charged phosphate group, and to each ribose one of four possible nucleobases (adenine, guanine, cytosine, and thymine) is attached (see figure 4.1). Combining billions of nucleotides into a single DNA strand then leads to the wealth of DNA sequences found in living organisms. [1]

4.1.1. Structures derived from canonical base-pairing

Under physiological conditions, the most stable DNA secondary structure is B-DNA, i. e. the right-handed double-helical motif formed by two complementary DNA strands connected via Watson-Crick base pairing. The formation is greatly favoured by the hydrogen bonding between adjacent nucleobases of the two opposite strands and π-π-stacking between neighbouring nucleobases on the same strand. [1] In cellular environments, the unfavourable accumulation of negative charges from the phosphate

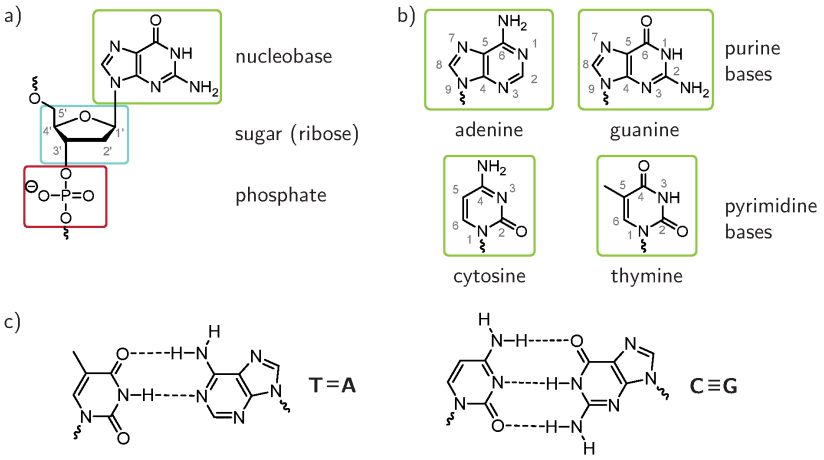


Figure 4.1.: DNA primary structure: a) structure of DNA nucleosides with b) four different nucleobases. c) canonical Watson-Crick T=A and C≡G base-pairing

backbones is compensated by surrounding (counter) cations like K^+, Na^+, Ca^{2+}, or Mg^{2+}. For a comprehensive discussion of nucleic acid metal ion interactions see the references.[13–15]

In the last decades, however, it has been found that several additional, different DNA secondary structures may form *in vivo* and *in vitro*. Other forms of DNA double helices are the more compact, right-handed A-DNA, which is adopted upon dehydration for example during crystallization, or left-handed Z-DNA, which can occur at high electrolyte strength or in DNA hairpins.[3] DNA does not always remain in its canonical double-stranded form but becomes transiently single-stranded during several cellular processes, like transcription and replication. This unwinding can lead to other secondary structures, including hairpins, cruciform, three- and four-way junctions (Holliday junction), triple stranded DNA (triplex), and quadruple stranded DNA (with G-quadruplexes and the i-motif).[1,3]

DNA hairpins can form in regions of the same DNA strand which are self-complementary to each other, e. g. palindromic sequences or

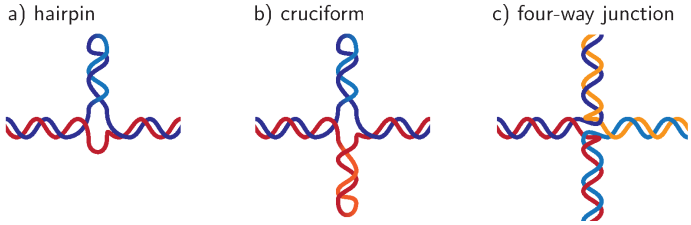


Figure 4.2.: Exemplary double-helical DNA secondary structures.

inverted repeats, thereby interrupting the normal Watson-Crick base-pairing in the double helix. When the palindromic sequence is copied into the complementary strand then two opposite hairpins are formed, resembling the shape of a cross and therefore named cruciform (see figure 4.2a and b).[1]

The Holliday junction is transiently formed during genetic recombination and is an important feature of DNA repair mechanisms of double-strand or single-strand breaks. Its structure is formed by a crossover between two DNA double-stranded sections, which results in a DNA four-way junction (see figure 4.2c). A variant is the three-way junction, which forms analogously, and is found during DNA replication and recombination.[1,16,17]

4.1.2. Structures with Hoogsteen base-pairing

The standard Watson-Crick base-pairing between adjacent nucleobases is not the only possible hydrogen bond pattern between nucleobases. In triplex DNA a third strand binds e. g. in the major groove of the double-helix by additional Hoogsteen base-pairing, resulting in base-triplets.[18] Hoogsteen base pairing involves $N7$ and $C6\text{-}NH_2/O$ of the purine base forming hydrogen bonds to the $N3$ and $C4\text{-}NH_2/O$ Watson-Crick side of the pyrimidine counter part. One example of a base triplet (T=A · T) is depicted in figure 4.3a. Among the other possible triplets, C=G · C^+ involves an unusual hemiprotonated cytosine.[18]

Hemiprotonated cytosine is also the basis for a quadruple stranded DNA secondary structure, namely the i-motif. Here, protonation of the

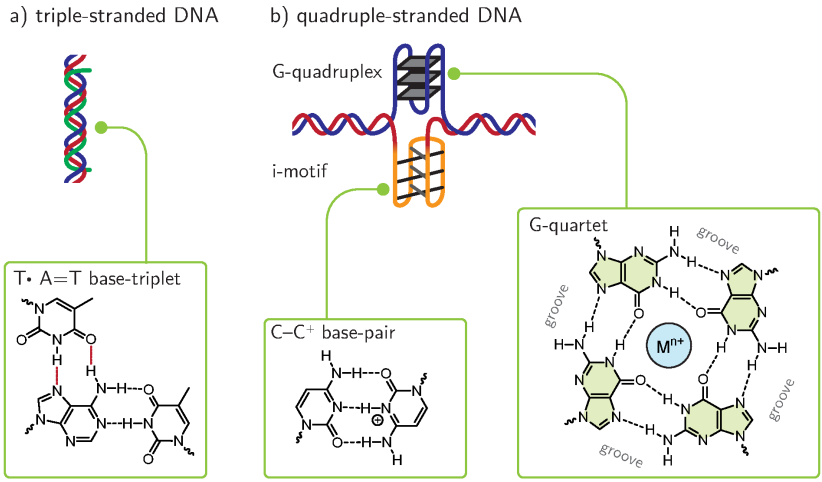


Figure 4.3.: Exemplary triple (a) and quadruple (b) stranded DNA secondary structures. The depicted T • A=T triplet is only one of several possible nucleobase combinations. Hoogsteen hydrogen bonds are depicted with red bullets, Watson-Crick hydrogen bonds with black dashes.

$N3$ position of cytosine under acidic conditions enables the formation of a $C \cdot C^+$ base pair. Two parallel double helices are then connected in an antiparallel fashion via intercalated $C \cdot C^+$ base pairs (see figure 4.3b). *In vitro* studies make use of acidic conditions to trigger formation of the i-motif, so its occurrence *in vivo* is debatable. However, the concerted formation of G-quadruplexes in the complementary strand could shift the equilibrium towards the i-motif. In duplex DNA, a given C-rich DNA strand will always have a G-rich complementary strand, so local unwinding of the double-helix should favour formation of both the G-quadruplex and the i-motif.[19]

G-quadruplexes, as already stated above, assemble when guanine-rich DNA and RNA folds into quadruply stranded DNA, by formation of stacked, planar G-quartets held together by Hoogsteen base pairing (see figure 4.3b). A more detailed description will follow further below.

It should be noted that nucleosides can be involved in two other types of interaction apart from Watson-Crick and Hoogsteen base-pairing, namely the sugar-edge or CH interaction. Sugar-edge hydrogen bonds can involve the C2′-OH found in RNA oligonucleotides, whereas for CH–O interactions adenine $H2$, purine $H8$, and pyrimidine $H5$ and $H6$ may be involved.[20]

4.1.3. G-quadruplex formation and topology

Among the non-canonical DNA secondary structures, G-quadruplexes have attracted special interest of researchers, as putative G-quadruplex forming sequences have been located throughout the whole human genome.[21]

In order to understand the biological role of G-quadruplexes it is a prudent necessity to understand the underlying principles which govern G-quadruplex formation, topology, and reactivity towards other molecules. Tremendous amounts of effort have been put into these questions, the results highlighted in various scientific reports.[22–33] It was found that G-quadruplexes are able to assemble intermolecularly from four separate guanine-rich oligonucleotide strands (tetramolecular G-quadruplexes), two strands (bimolecular), and intramolecularly from

one strand (unimolecular). Even three-stranded G-quadruplexes can be achieved.[34] However, due to the fact that the *in vivo* DNA concentration is rather low, formation of bimolecular or tetramolecular G-quadruplexes should be strongly disfavoured.[35,36] As mentioned above, four guanine nucleobases form a planar, four-fold pattern with Hoogsteen base-pairing. These G-quartets, also called tetrads, stack on top of each other, thus leading to a helical assembly.

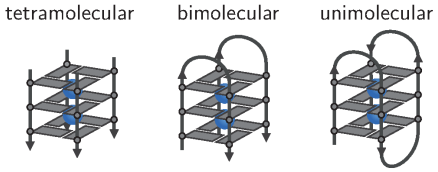


Figure 4.4.: Schematic illustrations for inter- and intramolecular G-quadruplexes. The arrows indicate 5′→3′ strand directionality. Note that the depicted topologies are only illustrative examples.

Cation coordination in the central G-quadruplex channel is a prerequisite for their formation. A detailed description can be found in the literature.[37] The disfavoured accumulation of negative (partial) charges is balanced by the incorporation of mono- or divalent metal cations, typically K^+ or Na^+, which coordinate in the central channel of the G-quadruplex helix to the *O*6 oxygen atoms of the adjacent guanines. Cation coordination is governed by attractive ion-dipole interactions, ion size, and dehydration energy. It has been verified by solution NMR and crystallographic studies that K^+ ions with their ionic radius of 1.33 Å cannot fit into the cavity of one G-quartet and are therefore found between the planes of two adjacent G-quartets, with an octadentate coordination environment.[38] Crystallographic experiments revealed that Na^+ ions can adopt coordination both within one plane or between two planes, and intermediate positions, which can be explained by the smaller ionic radius (0.95 Å) which is exactly the size of the G-quartet cavity.[39] The different coordination sites found in crystal structures are highlighted in figure 4.5, which also demonstrates that the G-quartet does not necessarily have to be completely planar.[37,39] Especially the

G-quartets at the 5′ or 3′ ends tend to bend inwards, thereby maximizing the attractive ion–dipole interactions. The in-plane coordination of Na^+ ions, however, seems to be an artefact of crystallization, as it is not observed in solution or unordered solid state, as evidenced by solid-state NMR experiments.[37]

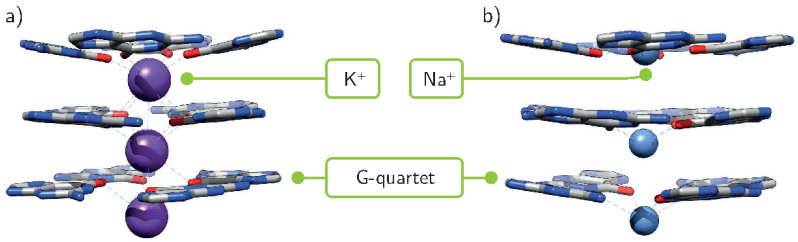


Figure 4.5.: Coordination of stabilizing cations in the central channel of the G-quadruplex structure. a) Crystal structure of a unimolecular, parallel G-quadruplex (PDB 1KF1)[38] with K^+ located between the planes of the G-quartets; b) crystal structure of a tetramolecular, parallel G-quadruplex (PDB 244D)[39] with Na^+ ions coordinating in the plane and intermediate positions. Hydrogen atoms, sugar and phosphate backbone are omitted for clarity.

Among the alkali metal cations, Li^+ and Cs^+ normally do not trigger G-quadruplex formation, which can be explained by the high penalty of dehydration for Li^+, which is not compensated by the attractive electrostatic interactions; and the size of the Cs^+ ion, which is too big to fit into the central G-quadruplex channel.[37] In terms of thermodynamic and kinetic stabilization, experimental evidence like melting temperature analysis (see chapter 4.3.1) revealed the following trend: $K^+ > NH_4^+ > Rb^+ > Na^+ > Cs^+ > Li^+$.[37,40,41] It should be noted that under physiological conditions, K^+ ions have the highest concentration in the cell's nucleus, so G-quadruplexes will predominantly form with this cation *in vivo*.[6] For the divalent metal cations, stabilization occurs for Mg^{2+} (at low concentrations only; most likely due to interactions with the phosphate backbone), Ca^{2+}, Sr^{2+}, Ba^{2+}, and Pb^{2+}, whereas

transition metal ions like Mn^{2+}, Co^{2+}, Ni^{2+} tend to disfavour formation at higher concentrations.[23,37]

As stated above, tetramolecular G-quadruplexes assemble from four separate strands, whereas bimolecular and unimolecular G-quadruplexes are formed in an intramolecular manner. The latter requires that two or all four strand sections are connected via so called loops, which can be as short as one single nucleotide. Four different loop arrangements have been identified (see figure 4.6), although a prediction which kind will be present in a given sequence is not possible until now, as it depends on too many different factors like loop nucleobases and length, electrolyte, solvent, and additives. A lateral (or edge-wise) loop connects adjacent guanines, whereas a diagonal loop bridges guanines on opposite corners. Both loop types result in a reversal of the strand polarities and so in an antiparallel arrangement, whereas the same polarities are found for the propeller loop and the more uncommon V-shape loop, the latter running along the side of the G-quadruplex stem (see figure 4.6).[27] As a consequence of the different possible loop types and strand polarities, bimolecular and especially unimolecular G-quadruplexes exhibit a plethora of typologies.[26] Sometimes, the G-quadruplex assembly invokes or is stabilized by triads, triplets, or quartets (or even higher), which are composed of varying combinations of the other nucleotides (e.g. mixed quartets like G–C–G–C).[42]

Even more complexity arises from the glycosidic bond angle of the guanosines, *syn* or *anti* conformation (see figure 4.7), which not only steers loop arrangement and strand polarity, but also the shape and size of the four helix grooves. In case of tetramolecular G-quadruplexes, the guanosines nearly exclusively adopt all-*anti* conformations within the G-quartets, which in turn invokes the predominantly found parallel strand orientation of the four DNA strands. The relationship between glycosidic bond angle and strand orientation within a G-quartet is nicely phrased in a recent work by Cang et al. (X. Cang, J. Šponer, T. E. Cheatham III, 2011)[44]: "For a given G-quartet, if any two guanines belong to the two strands oriented in the same direction, the two guanines display identical glycosidic conformation. If their strand orientations are opposite, they

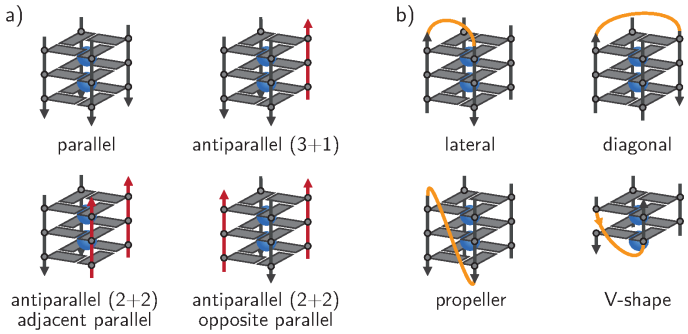


Figure 4.6.: G-quadruplex loops and strand orientations. a) parallel and antiparallel strand orientations; b) lateral, diagonal, propeller, and V-shape loop connections. The arrows indicate 5′→3′ strand directionality.

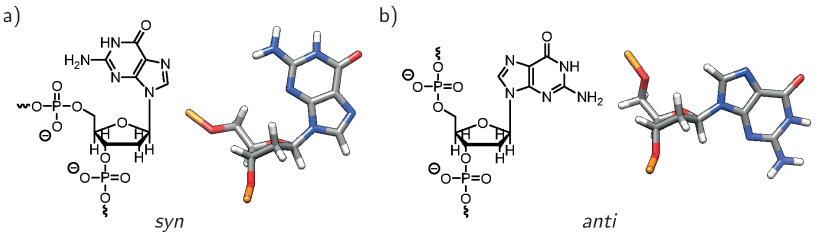


Figure 4.7.: Glycosidic bond angle differences in *syn* and *anti* conformations of guanosine. The molecular models have been taken from the crystal structure of a bimolecular G-quadruplex.[43] Note that in both cases, the sugar is in the C2′-endo conformation.

show opposite glycosidic conformations." In the same publication, the authors derived simple rules to assess the most stable glycosidic bond patterns along a given G-tract. Their findings are based on molecular dynamics simulations for two-quartet models $[dGG]_4$ and comparison with NMR and crystal structure data. The order of stability is 5′–*syn-anti*>*anti-anti*>*anti-syn*>*syn-syn*. This explains why G-quadruplexes with two stacked G-quartets nearly always show *syn-anti* steps, whereas those with three have to combine *syn-anti* steps with more unfavourable arrangements.

4.1.4. Biological role of G-quadruplexes

Understanding the nature of G-quadruplex folding and topology is one important aspect, the effect on biological processes another. As already said above in the previous section, G-quadruplex forming sequences are common in the human genome. Many show the consensus DNA sequence $d(G_{3+}N_{1-7}G_{3+}N_{1-7}G_{3+}N_{1-7}G_{3+}N_{1-7})$, although G-quadruplex formation may not strictly be limited to this sequence pattern due to over- or underestimation of the loop or G-tract length. As these sequences are located non-randomly in the genome, a regulatory role of G-quadruplexes is proposed.[21] In this regard, formation of G-quadruplexes may influence first: DNA transcription (figure 4.8a), by promoting or down-regulating gene expression, e.g. in promoter regions of oncogenes; second: DNA replication (figure 4.8b), as specialized helicases are apparently needed to unwind not only the double helical B-DNA but also G-quadruplex structures; third: translation (figure 4.8c), as a G-quadruplex forming sequence in a sense DNA strand is transferred to its mRNA part, the single-stranded mRNA being more prone to higher-order structure formation due to the lack of a counter strand, so G-quadruplex formation becomes very likely; fourth: telomere maintenance (figure 4.8d), where the repetitive, single stranded overhang at the 3′-end of the chromosomes DNA is elongated by the enzyme telomerase, which is inhibited by G-quadruplex formation.[21,35,36,45,46]

Numerous examples of proteins interacting with G-quadruplex structures have been identified,[47–49] whereby all the above mentioned find-

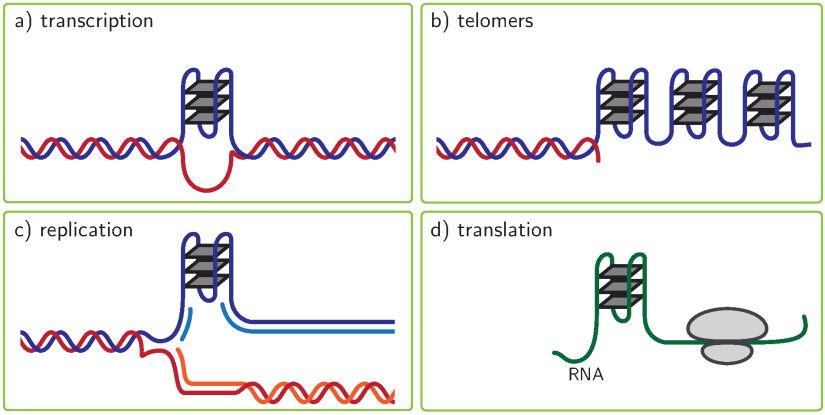


Figure 4.8.: Schematic illustration of putative G-quadruplex formation throughout the genome with implications on a) transcription, b) telomere maintenance, c) replication, and d) translation. Adapted from ref. [21]

ings are based on *in vitro* experiments, supported by computational investigations. However, increasing *in vivo* evidence for the formation of G-quadruplexes was only found recently by quantitative staining of DNA and RNA G-quadruplexes structures in living cells with a fluorescent dye (Figure 4.9). [50,51]

In conjunction with the biological role of G-quadruplexes, investigations have also aimed at understanding medical related implications of G-quadruplex formation, especially concerning the treatment of cancer and neurodegenerative diseases, as G-quadruplexes could either interfere with or restore normal cell functionality. Research was initially fuelled by the finding that the enzyme telomerase, which is overexpressed in cancer cells and partly responsible for their longevity, can be inhibited by the stabilization of G-quadruplex structures in the single-stranded telomeres (see figure 1.1 in chapter 1). [52–56] A whole branch of G-quadruplex related research is now dedicated to the identification of G-quadruplex stabilizing drugs, also called G-quadruplex ligands, some of which are already in the later stages of clinical trials. [57–63] The three main targets

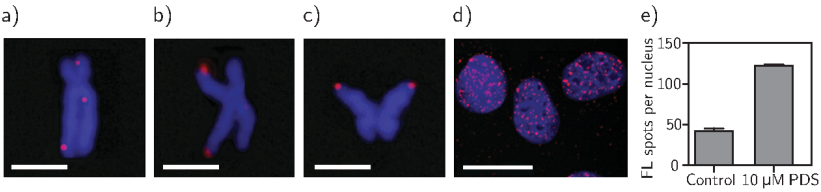


Figure 4.9.: a)–c) Immunofluorescence based visualization of G-quadruplex structures in chromosomes (HeLa cells, G-quadruplex structures highlight in red). Scale bars, 2.5 µm. d) and e) Stabilization of endogenous G-quadruplex structures by a small-molecule ligand. e) Increase in G-quadruplex structures after treatment with the G-quadruplex binding ligand pyridostatin (PDS). Scale bar, 20 µm. Reprinted by permission from Macmillan Publishers Ltd: Nature Chemistry G. Biffi et al.[50] Copyright ©2013.

for anticancer treatments comprise (1) gene promoter G-quadruplexes, e. g. in order to downregulate oncogene expression; (2) protein/enzyme targets as G-quadruplex aptamers can influence interactions between genomic DNA and proteins; and (3) telomeric G-quadruplexes which can help controlling telomerase activity.[53]

These findings not only show the structural diversity and polymorphism of G-quadruplex structures, but also their implications on various cellular processes. Both aspects render G-quadruplexes as ideal candidates not only in chemical biology but also as tools in DNA nanotechnology. This concept and the examples already presented in the current research will be introduced in the next chapter.

4.2. DNA nanotechnology and supramolecular chemistry

Parallel to the aim of understanding the importance and role of DNA secondary structures in living organisms, another approach focuses on using the unique folding behaviour and intrinsic properties of these assemblies for artificial systems. Here, the DNA is assembled with

a specific functional or structural design. This field of research was coined DNA nanotechnology and is part of the supramolecular chemistry research field. This chapter very briefly introduces some important concepts of supramolecular chemistry by means of a simple example, which is based on the target G-quadruplex metal-complex hybrid structure introduced in chapter 1. This is followed by a more detailed description of DNA nanotechnology (subsection 4.2.2) including DNA modifications and the metal base-pairing concept (subsection 4.2.3). The last part focuses on G-quadruplex based DNA nanotechnology and the state of research regarding G-quadruplex metal ion interactions, including G-quadruplexes modified with transition metal complexes (subsection 4.2.4).

4.2.1. Supramolecular chemistry

Supramolecular chemistry research is dedicated to understand non-covalent interactions between molecular assemblies and to use them in the construction of higher-order aggregates (supramolecules) with designed properties and functionalities. The interest in non-covalent interactions dates back several decades and is closely related to the growing knowledge about the structure and interactions of biomolecules, including DNA. An early example of a supramolecular interaction is the acetic acid dimer, which forms through two hydrogen bonds between the single molecules. According to the current IUPAC definition, the hydrogen bond is: "[...] an attractive interaction between a hydrogen atom from a molecule or a molecular fragment X–H in which X is more electronegative than H, and an atom or a group of atoms in the same or a different molecule, in which there is evidence of bond formation." (see also figure 4.10)[64] As seen in the previous chapter, hydrogen bond formation is essential for all the secondary structures of DNA, although the individual interaction can be quite weak in strength ($<$12–120 kJ mol^{-1}).[6] An additional interaction found between the DNA strands in a given secondary structure is the π–π interaction (50–500 kJ mol^{-1})[65] between adjacent nucleobases, resulting in the ladder-like stacking on top of each other. When looking at the more

general electromagnetic interactions between particles, non-covalent interactions comprise mainly ion–ion (100–350 kJ mol^{-1}), ion–dipole (50–200 kJ $mol^{-}1$), , dipole–dipole (5–50 kJ mol^{-1}),[6] and van-der-Waals or hydrophobic interactions (0.05–40 kJ mol^{-1}).[65]†

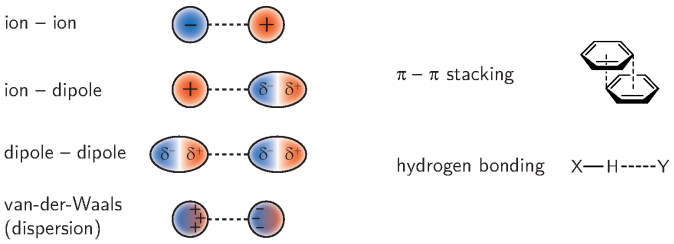


Figure 4.10.: Illustrations of important non-covalent interactions. The depicted examples are only one of sometimes many possible variants.

It becomes immediately evident that, depending on the kind of non-covalent interaction, even a single occurrence of these interactions can be on the same order of magnitude in strength as compared to a covalent bond, the latter ranging from around 144 kJ mol^{-1} (O–O) to 646 kJ mol^{-1} (B–F) and 1077 kJ mol^{-1} (C≡O).[12] Additionally, in biomolecules these interactions normally do not occur solitary but as a combination, so for a given system several contributions to the overall stability and reactivity have to be considered.

Many supramolecular concepts are important for DNA related chemistry and therefore, some of them will be very shortly introduced here. A key concept is the so called *host-guest* chemistry, which describes the interaction between two or more interacting molecules.[6] Commonly, the host is a larger assembly with a hole or cavity, offering the binding sites to which a guest molecule is bound. For the structural goal of this thesis, that is the construction of a G-quadruplex modified with a "metal base-tetrad" as explained earlier in the introductory chapter, the assembled G-quadruplex DNA part can be regarded as the "host", offer-

†The literature values for the mentioned interaction energies differ quite considerably and should be seen only as a rough guide.

ing the binding pocket via the attached donor functionalities to a "guest" ion, namely a transition metal ion. The transition from unordered G-rich DNA single-strands to the assembled G-quadruplex structure is a process of *self-assembly*, as a complex and aggregated DNA structure is spontaneously, but still reversibly, formed starting from simple building blocks. The assembly process not only includes attractive hydrogen bonds but also π–π interactions, which comprise in addition to attractive van-der-Waals forces also *hydrophobic* and *solvation effects*, as the interaction between the hydrophobic, aromatic nucleobases excludes the water molecule solvation shell around the nucleobases. This is also seen for G-quadruplexes, as normally there are no solvent molecules found in the central cation channel of the G-quadruplex structure. Another feature of the assembled donor-functionalised G-quadruplex is that the metal ion binding site is *preorganised*, i. e. it is positioned on one side of the G-quadruplex stem and is therefore readily binding the metal cation. Furthermore, the assembled G-quadruplex can also be seen as a *chelate* ligand, as the donor sites are linked together by the G-quadruplex core, thereby favouring intramolecular instead of intermolecular coordination. If more than one binding site is present in a given supramolecular structure, than the effect of *cooperativity*, or sometimes also *allostery*, has to be considered. The first binding event can alter the *binding constant* for the next. This constant is a numerical value used to measure the thermodynamic strength, that is the position of the equilibrium between bound and unbound state of the binding event. For a more thorough discussion of all these terms and their application in supramolecular chemistry, the reader is referred to some recently published books in this field. [6,66,67]

4.2.2. DNA nanotechnology

Simultaneously to the deeper understanding of the underlying principles of supramolecular chemistry, the idea arouse to use DNA as a structural tool on the nanometre level, something initially proposed by Nadrian Seeman. [68] Since then, the research field called DNA nanotechnology has progressed tremendously. It can be roughly divided into two main

parts: structural and functional DNA nanotechnology. In structural DNA nanotechnology two-dimensional and three-dimensional sheets or lattices, and objects are constructed,[69–71] while the functional DNA nanotechnology branch explores the area of DNA nanodevices, like switches and motors.[72–79]

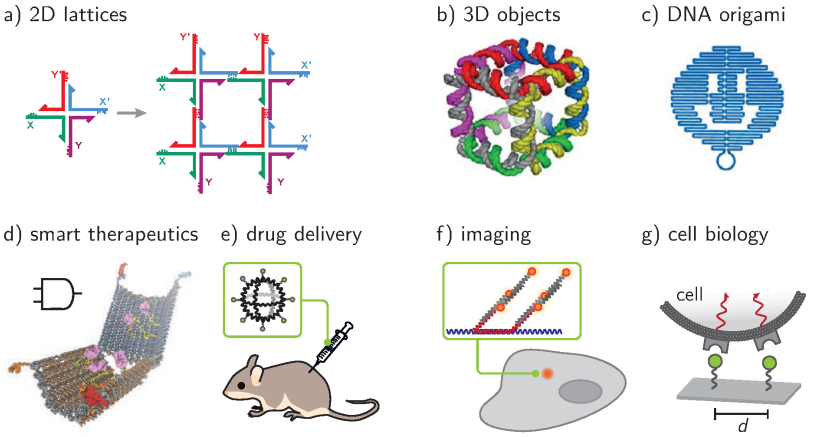


Figure 4.11.: Examples for DNA nanotechnoloy. a) Self-assembly of DNA to larger 2D sheets (tiles) starting from four-way junctions with sticky ends;[70] b) hexacatenane, cube-like 3D DNA assembly;[80] c) DNA origami leading to defined shapes (here: smiley);[81] d) smart therapeutics: an aptamer DNA nanorobot[82]; e) drug delivery;[83] f) imaging;[84] g) spatial organisation of functional groups for quantitative measurements in cells.[85] For more details see the text. Figures reproduced with permission from a) ref.[70], b) and c) ref.[86] ACS, d) ref.[87] AAAS. Figures d)–g) redrawn and adapted from ref.[88].

Structural DNA nanotechnology makes use of DNA's ability to break up the canonical wire like duplex structure to form branched assemblies. Indeed, a simple and stable 2D (or 3D) lattice can be constructed from a four-way junction with sticky ends. The four-way junction was already introduced as an alternative DNA secondary structure and serves

here as a scaffold for the larger DNA assembly. Sticky ends are small single stranded DNA overhangs at the DNA duplex ends. Hybridization between two sticky ends joins two DNA fragments together, holding them in place. Figure 4.11a shows such a four-way junction with sticky ends. The overhangs are designed to be complementary to each other, resulting in the depicted 2x2 DNA tile assembly. As it again contains open sticky ends, it will assemble into even larger structures. [70] Another approach in structural DNA nanotechnology is to use long (viral) single-stranded DNA and to let it self-assemble with numerous small DNA single-strands, called staple strands, by a careful annealing process. Depending on the sequence design this large assembly can fold into nearly any imaginable shape, including more proof-of-concept like structures (e. g. a DNA smiley, see figure 4.11c). The so called DNA origami was a major breakthrough in the field, as it enables to easily introduce diverse functionalities in a spatial oriented way. [70,81,89] This way, DNA nanotechnology becomes interesting in the field of chemical biology, where the focus is on analysing and manipulating molecular information, for example of a living cell. Applications could be found (1) in smart therapeutics, which can react to the local cellular conditions (figure 4.11d); (2) in drug delivery systems, where a DNA cage mimics a virus' protein capsid (figure 4.11e); (3) imaging, e. g. mRNA imaging in living cells by duplex formation with a chromophore-bound DNA strands (figure 4.11f); and (4) distant dependant examination of e. g. protein–DNA interactions (figure 4.11g). [88,90] These examples already combine structural DNA nanotechnology with the functional side. Concerning the latter, so far molecular motors like walking motors, molecular circuits, and catalytic amplifiers have been realised. [88] Also systems showing DNAzyme activity or DNA–enzyme coupled structures have been investigated. [91] This demonstrates that DNA nanotechnology is at the brink of stepping up from a mere peculiarity to a field competing with alternative established techniques like antibodies or polymeric materials and even offering completely new applications. [88]

Using just the natural nucleotide library offers in itself a plethora of structural and functional applications but does not include the possi-

bilities which arise from introducing other non-canonical functionalities. This includes for example drug delivery systems where the DNA acts as a platform or template to offer conjugated peptides, proteins, small molecules, or carbohydrates to an intracellular target.[92] Furthermore, the stability of artificial DNA constructs needs to be finely tuned at the bridge between stability in the cell and impairment of regular cellular processes.[88] As a consequence, scientists strive to extend the natural canon of nucleotides to design artificial DNA hybrids with a specific functionality. This led to the development of a variety of DNA modifications, ranging from (1) backbone alterations, (2) sugar analogues, to (3) nucleobase derivatives.

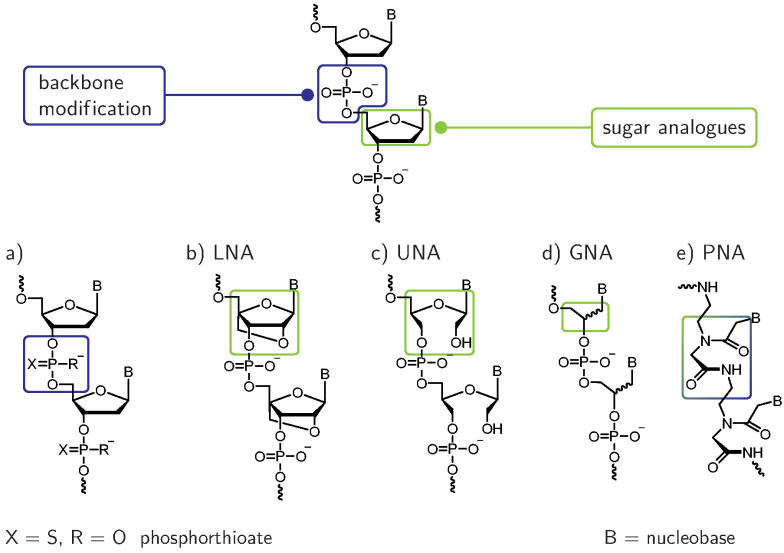


Figure 4.12.: Examples for DNA nucleotide modifications. a) phosphodiester backbone variants, b) locked nucleic acid (LNA), c) unlocked nucleic acid (UNA), d) glycol nucleic acid (GNA) and e) peptide nucleic acid (PNA).

Variations in the backbone comprise subtle alterations of the phosphodiester linkage leading for example to phosphorthioates (see figure 4.12a), or conformational locking using a dioxaphosphorinane ring, but

also major modifications, where a completely new chemical functionality is incorporated.[10] Here, examples include a methylene/formacetal bridge between the nucleotide's 5′ and 3′ oxygen, a triazole linker formed via a 1,3-dipolar cycloaddition reaction, or diphosphodiester moieties. An important aspect of all these modifications is the already above mentioned fine-tuning of the nucleic acid stability. Two typical examples for this idea are made use of in the branch of sugar modifications: either the degree of rigidity of the sugar is increased by conformational locking, as implemented in the locked nucleic acid (LNA), or the conformational freedom is enhanced in case of unlocked nucleic acids, named UNA (see figure 4.12b and c).[93] Both cases are structurally derived from RNA with its additional ribose 2′-OH group. Another approach was introduced by N. Ueda et al.,[94] T. Seita et al.,[95] and further advanced by the group of E. Meggers,[96–99] who exchanged the sugar moieties for a glycol based backbone (GNA) to investigate duplex formation in a simplified nucleic acid (see figure 4.12d).[10] The glycol backbone is not completely achiral: one stereocentre is retained and enables studying effects on stability and reactivity depending on the different diastereomers. This feature of GNA will be important for the results discussed in later chapters. Apart from its acyclic nature, one other important difference of the GNA backbone to its natural ribose counterpart is that it also has one carbon less in the linkage between two phosphodiester moieties. As a consequence a GNA:DNA hybrid structure does not show perfect conformational complementarity and therefore has a significant lower stability when compared to its natural DNA:DNA duplex counterpart. The opposite is observed, however, when a GNA:GNA assembly is used.[99]

One of the most intensively studied derivatives is the peptide nucleic acid (PNA), where the complete sugar-phosphodiester backbone is exchanged for a peptide sequence (see figure 4.12e). In contrast to the natural DNA with its anionic phosphate groups, PNA is intrinsically uncharged, unless additional charged groups are introduced.[100–103]

An important part of backbone or sugar modifications is represented by bioconjugation to other biomolecules like peptides, proteins, pro-

tein ligands, carbohydrates, and lipophilic compounds, using one of the numerous available ligation techniques, for example the copper(I) catalysed alkyne azide 1,3-dipolar cycloaddition (CuAAC) or Staudinger reaction.[92,104–107] The modification is often introduced at the 5′ or 3′ end of the synthesised oligonucleotide or can be inserted as part of the phosphoramidite-based solid-state DNA synthesis as an artificial backbone or at the $C2'$ position. Conjugation also involves attachment of reporter functionalities, which can react to or sense an external stimulus, for example fluorescent dyes or chromophores detecting trace metal ions (see figure 4.13).[92,108]

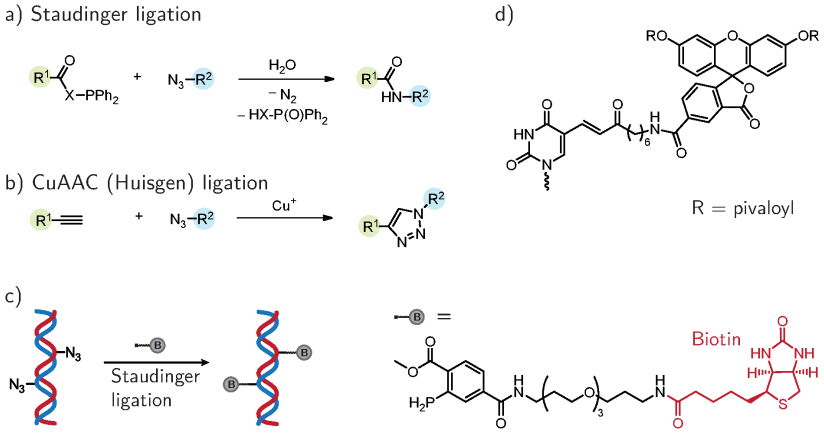


Figure 4.13.: Illustrative examples for DNA (bio-) conjugation.[104,105] a) General bioconjugation via a Staudinger ligation reaction, or b) copper(I) catalysed alkyne azide 1,3-dipolar cycloaddition (CuAAC); c) DNA labelling with biotin via a Staudinger reaction;[109,110] d) fluorescent dye attached to thymine.[92]

Next to backbone alterations and sugar analogues, the third kind of alteration is the possibility to change the nature of the nucleobases themselves, which is an even vaster field of research (see figure 4.14 for illustrative examples). Firstly, hydrogen bonding can be retained, but with a different bonding scheme, or extended with an additional shape complementarity of the opposite nucleobases, which prevents formation

of unwanted base-pairs by steric hindrance, or by just extending the size of the purine or pyrimidine moieties to affect DNA duplex shape and stability (see figure 4.14a–c). Secondly, hydrogen bonding can be discarded completely and exchanged for base-pairs which are recognized by hydrophobicity and shape complementarity, or interstrand stacking.[111] This class also encompasses covalent bridges between the two interacting strands and the metal base-pairing concept, which will be discussed in more detail below. Thirdly, base-pairing can be excluded and instead other functionalities be introduced into the oligonucleotide strand.[111] Here the metal base-pairing concept briefly described in the introductory chapter 1 comes into play.

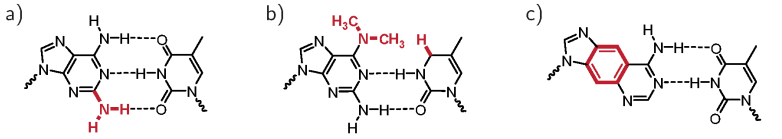


Figure 4.14.: Illustrative examples for DNA nucleobase modifications.[111] a) Hydrogen bonding variation; b) hydrogen bonding and shape complementarity; c) extended nucleobases.

4.2.3. The metal base-pairing concept

Metal-ion interactions with unmodified DNA strands can occur at many different locations, e. g. coordination to the phosphate backbone or the ribose units, or by cross-links via nucleobase donor groups not involved in canonical Watson-Crick base pairing, as seen for cisplatin Furthermore, the metal ion can either stabilize non-standard nucleobase pairs, or be inserted into Watson-Crick base pairs by replacing hydrogen atoms of the involved hydrogen bonds.[7] Two examples which make use of canonical nucleotides comprise the T–Hg–T and C–Ag–C metal base-pair, in which the Hg^{2+} or Ag^{+} ions bridge between the deprotonated nucleobases.[11] The nucleobase mismatch is held together only by the dative bonds between the transition metal ion and the pyrimidine $N3$ atoms. This observation encouraged scientists to introduce other metal base-pairs

into DNA. The metal binding ligand can be a natural nucleobase, like in the just mentioned cases, or an artificial donor functionality attached to the canonical ribose. The latter design of an artificial ligand attached to the ribose was coined "ligandoside" and many examples of these have been reported in the literature, together with their metal binding abilities.[7,112–115]

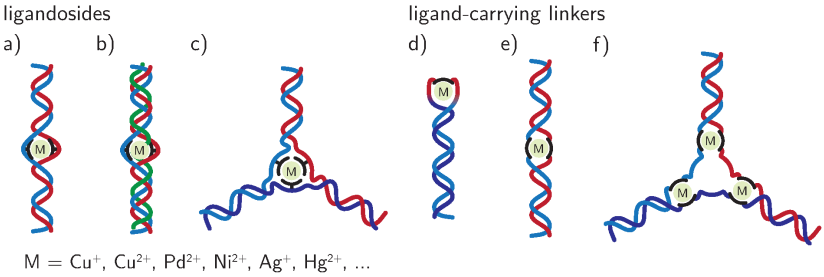


Figure 4.15.: Illustrative examples for simple DNA structures with ligandoside-based metal base-pairs and base-triplets, or metal complex linkers. a) Duplex DNA with one ligandoside metal base-pair; b) triplex DNA with a ligandoside base-triplet; c) three-way junction with one metal coordinated by three ligandosides; d) DNA hairpin mimic; e) two duplex DNA sections bridged by a metal complex linker; f) three way junction with three metal complex joints and three bridging DNA single-strands.

In most cases, the ligand-metal coordination environment is either linear, square-planar, when neglecting loosely bound solvent molecules or ions, or (distorted) octahedral. Planar ligand systems help to maintain the favourable stacking of nucleobases within the DNA duplex structure. In analogy to the natural nucleobases, ligand donor functionalities are mainly composed of mono-, di-, or tridentate nitrogen heterocycles, or aromatic compounds with mostly nitrogen, oxygen, sulphur or mixed *N*, *O*, *S* donor groups.[11] The first example of a metal bound ligandoside was introduced by the group of M. Shionoya with ligands based on 1,2-disubstituted phenylene, although they initially did not report the incorporation into a DNA duplex (see figure 4.16a).[116] A first

metal base-pair inside a DNA duplex was achieved by E. Meggers et al. from the group of P. G. Schultz for a mixed-ligand system comprising a pyridine-2,6-dicarboxylate and pyridine coordinated to Cu^{2+} (figure 4.16b).[117] As already seen above, metal base-pairing can also occur between mismatches of natural nucleobases, giving rise to e. g. the T–Hg^{2+}–T or the C–Ag^{+}–C base-pair (figure 4.16c).[7,115] Linear coordination of a Ag^{+} ion was also used by the group of J. Müller to construct consecutive imidazole–Ag^{+}–imidazole metal base-pairs (figure 4.16d), for which they were also able to obtain the first NMR-derived solution state structure of metal base-pairing inside a DNA duplex.[118] The same group also complemented the findings of H. Weizman and Y. Tor on ligandosides based on bipyridine that the metal base-pairing concept can tolerate spatially more demanding ligand systems (figure 4.16e).[119,120] A remarkable metal base-pair was introduced by G. H. Clever et al., who combined the metal base-pairing concept with a covalent linkage in a copper-bound salen moiety (figure 4.16f).

Only very few examples for metal base-pairing in non-duplex DNA structures have been reported. From the group of M. Shionoya came the report of triplex DNA with a metal base-triplet based on the coordination of Ag^{+} by three pyridine ligandosides (figure 4.16g).[121] Later, the same group was able to construct an artificial DNA triplex held together by octahedral coordination of Fe^{3+} to hydroxy-pyridone modified nucleosides (figure 4.16h).[122] Attachment of a bipyridine functionality at the 2′-*O* position of an uracil nucleotide enabled M. Shionoya and coworkers to stabilise a DNA three-way junction by coordination to a divalent transition metal ion (see figure 4.16i).[123] In the case that instead of a ligandoside a completely artificial ligand linker is introduced into DNA other DNA structures like hairpins[124,125], triplexes with metal base pairs[125] and three-way junctions can be obtained (see also figure 4.15d–f).[125] With the same approach and when using the metal-coordination environments as vertices to connect duplex strands, even higher-order assemblies can be constructed.[126,127]

It is important to notice that the metal base-pairing concept based on ligandosides tries to mimic the natural Watson-Crick base-pairing, whereas the metal complexes based on artificial organic linkers just serve as connection joints for the rest of the DNA assembly.

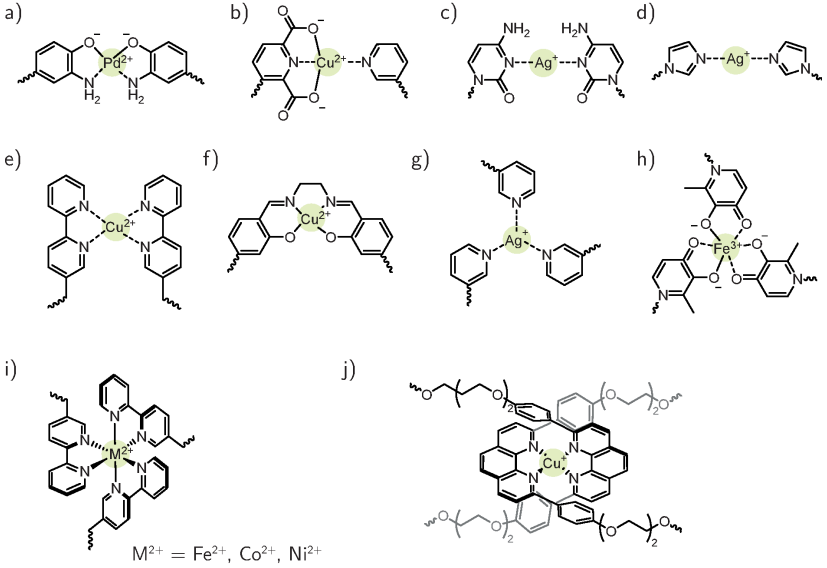


Figure 4.16.: Illustrative examples for metal base-pairs and base-triplets. Proposed metal ion coordination and ligandosides a) 2-aminophenol ap–Pd^{2+}–ap; b) pyridine-2,6-dicarboxylate and pyridine dipic–Cu^{2+}–py; c) cytosine C–Ag^{+}–C; d) imidazole im–Ag^{+}–im; e) bipyridine bipy–Cu^{2+}–bipy; f) Cu^{2+}–salen; g) pyridine py–Ag^{+}–$(py)_2$; h) hydroxypyridone H–Fe^{3+}–H_2; i) bipyridine bipy–M^{2+}–$(bipy)_2$, M^{2+} = Fe, Co, Ni); j) diphenylphenanthroline dpp–Cu^{+}–(dpp).

Metal coordination in conjunction with G-quadruplex structures will be discussed at the end of the next section. Before, the state of research regarding G-quadruplex based DNA nanotechnology will be highlighted.

4.2.4. G-quadruplex based DNA nanotechnology

The discovery that guanylic acid forms a gel at higher concentrations and the subsequent elucidation of the G-quartet structure stacked into helical columns (figure 4.17a) is the stepping stone into the world of G-quadruplexes and simultaneously the first example in the field of supramolecular chemistry and DNA nanotechnology.[128] Since then, this phenomenon has been exploited in quite a few examples of mostly lipophilic guanosine analogues assembling into higher-order structures. Studies include G-quadruplex ionophores, enantiomeric self-association, G-quadruplexes acting as a template e. g. to enhance the reaction between a groove binding peptide and an acridone ligand stacked on the G-quartet faces, and solitary G-quartets without a central cation.[24,129,130]

An interesting case is the pH controlled sol–gel equilibrium between a monomer, comprising two guanines linked together, and an oligomer of covalently linked G-quartets forming a gel (figure 4.17b).[132] The G-quartet stabilising potassium cations are either bound in the G-quartet core or by a [2.2.2]cryptand scavenger. In another example, the central cation cavity in the G-quadruplex structure is used to construct an ion channel transporting Na^+ ions across a phospholipid bilayer membrane (figure 4.17c).[133] Furthermore, the concept can be brought to completely new systems, e. g. by combining the hydrogen bonding abilities of cytosine and guanine in one molecule (figure 4.17d). The resulting assembly are six-membered rings, shaped like a rosette, which stack on top of each other. Attachment of outside pointing crown ether moieties enable the binding of chiral amino acids, resulting in the formation of homochiral helices instead of racemic mixtures via a self-sorting process.[134]

Coming back to G-quartets assembling from oligonucleotide strands, another important field is the designed arrangement of G-quadruplexes topologies. A prime example for this is the construction of tetramolecular and bimolecular G-quadruplexes by means of a cyclic, peptidic frame (see figure 4.18a).[135,136,141,142] Guanine-rich oligonucleotide strands are connected via oxime or 1,3-dipolar alkyne azide (CuAAC) ligation to the peptide scaffold. In case of bimolecular G-quadruplexes, the

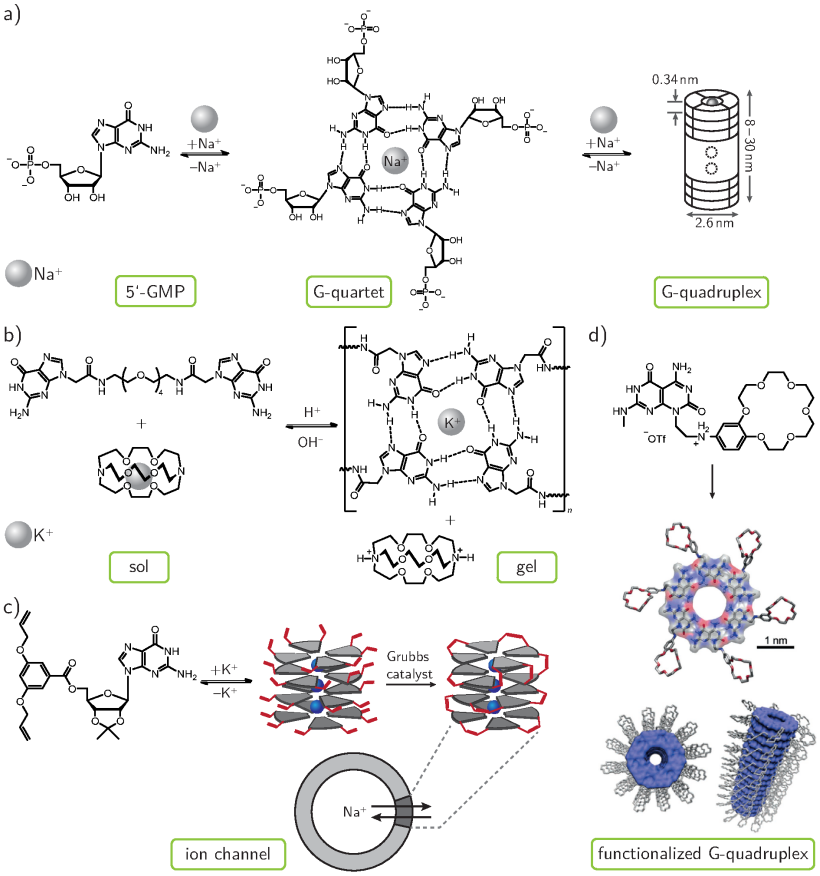


Figure 4.17.: Examples of the supramolecular assembly of guanosine derivatives. a) Na^+ mediated formation of G-quartets and long wire-like G-quadruplexes from 5′-GMP;[131] b) pH induced sol–gel switching of a linked guanine dimer and its corresponding polymeric G-quadruplex;[132] c) formation of a columnar G-quadruplex structure from a modified guanosine, fixation of the topology by cross-linking, incorporation into lipid bilayers, and ion channel activity.[133]; d) assembly of crown-ether functionalized guanine derivatives into rosette like hexagons and stacking into long G-quadruplex wires. Figures reprinted with permission from a) ref.[129] RSC, c) ref.[133] ©2006 ACS (scheme redrawn), d) ref.[134] ©2002 ACS.

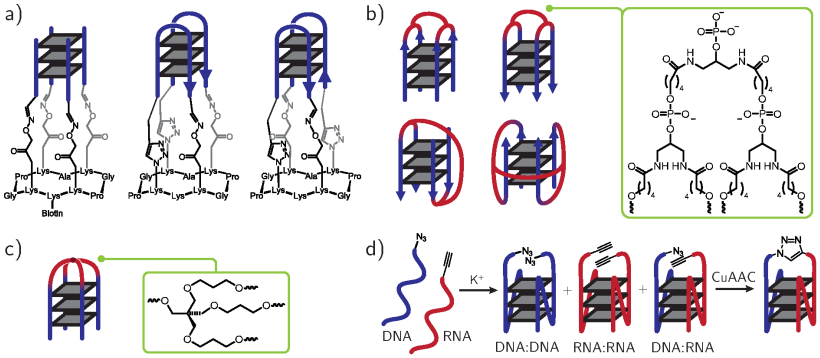


Figure 4.18.: Steering G-quadruplex molecularity and strand orientations. a) Tetramolecular and bimolecular G-quadruplex formation promoted by a cyclic peptide scaffold;[135,136] b) tetra end-linked G-quadruplexes;[137,138] c) branched G-quadruplexes;[139] d) Covalent fixation of a DNA:RNA hybrid G-quadruplex via CuAAC reaction;[140]

strand connection to the peptide template steers the folding into a topology where the 3′-ends of the DNA strands are located either on adjacent (parallel arrangement) or opposite sides (antiparallel). A similar approach was reported by G. Oliviero et al.[137,138,143] (figure 4.18b) and H. Ferreira et al.,[139] (4.18c) who used a tetra end-linkage to covalently connect four guanine-rich oligonucleotides. Alterations of the linker chemistry enables the formation of e. g. G-quadruplexes with partially inverted strand polarities.

Attaching reactive functionalities to G-rich strands can also be used to get insights into DNA:RNA G-quadruplex hybrids. Normally, when using G-rich DNA and RNA strands simultaneously, a mixture of all strand combinations is formed. The group around M. Komiyama used the CuAAC reaction to covalently connect a DNA with an RNA strand.[140] The reaction only proceeds in the hybrid DNA:RNA G-quadruplex assembly, and not in the homostrand assemblies (see figure 4.18d).

Another approach towards the specific design of tetramolecular G-quadruplexes came from the group of J.-L. Mergny, namely using

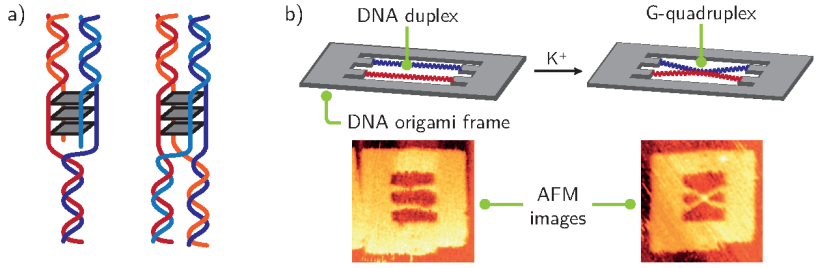


Figure 4.19.: Steering G-quadruplex molecularity and strand orientations. a) Guided formation of tetramolecular DNA and RNA G-quadruplexes by double strand formation;[144,145] b) fixation of DNA duplex strands with guanine-rich parts in a DNA origami frame, G-quadruplex formation induced by K^+ ions and visualisation by AFM imaging;[146,147] AFM images reproduced with permission from ref.[146] ©2013 Oxford University Press.

the hybridisation of duplex DNA to construct a tetramolecular G-quadruplex from four different oligonucleotide strands (figure 4.19a). Without the duplex formation a mixture of all possible G-quadruplex structures would be obtained.[144] The same group also tried to use this technique to trigger the formation of an antiparallel DNA as well as an antiparallel RNA G-quadruplex, respectively. Although they were successful in case of the DNA G-quadruplex, however, they were only able to obtain the canonical parallel RNA G-quadruplex.[145,148] Monitoring formation of G-quadruplexes with different molecularities and strand orientations was nicely achieved by fastening the relevant oligonucleotide strands into an origami framework with subsequent analysis by atomic force microscopy (AFM).[146] One frame contained two adjacent DNA duplexes, with a G-rich section in the middle of each strand. The AFM images clearly show two non-interacting DNA duplex structures in the absence of K^+ ions, whereas an X-shape due to G-quadruplex formation is observed after K^+ ions are added (see figure 4.19b). This technique was later also extended to hairpins and triplex structures.[147]

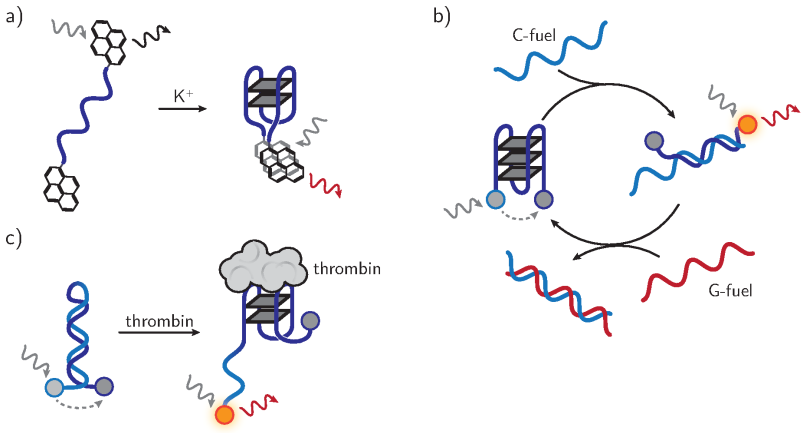


Figure 4.20.: G-quadruplex sensors based on chromophore fluorescence. a) G-quadruplex dependant formation of a pyrene excimer;[149] b) G-quadruplex–duplex motor;[150] c) binding of thrombin to a fluorophore–quencher labelled aptamer;[151]

A major branch of G-quadruplex related research in DNA nanotechnology is dedicated to biosensors, and nanodevices like switches and motors. Regarding the former approach, one common technique already introduced earlier is to covalently attach chromophores to the G-quadruplex. For example, chromophore fluorescence can then be indicative of G-quadruplex formation or denaturation.[149] More precisely, a fluorophore and quencher pair can be attached to either ends of a unimolecular G-quadruplex forming oligonucleotide strand. The DNA denatured state corresponds to a large distance between the two, giving rise to a strong fluorescence signal, whereas the G-quadruplex assembly brings both chromophores together resulting in a quenched fluorescence signal. Folding and unfolding of a G-quadruplex as part of a duplex–G-quadruplex transition can also be used to construct a motor with DNA single strands as a fuel (see figure 4.20b).[150] Starting point is again a unimolecular G-quadruplex with a fluorophore–quencher pair at the ends. Hybridisation with a long complementary strand opens the G-quadruplex structure and leads to duplex formation. This “fuel”

strand is detached by a scavenger strand, the latter making a stronger duplex with the fuel strand, due to a longer complementary nucleotide tract. The shorter G-rich strand is thereby released, again forming a G-quadruplex and thus closing the cycle. [150]

This technique of folding and unfolding a G-quadruplex structure can also be used in the recognition of proteins. A prime example is the thrombin binding aptamer sequence (TBA), which folds into a G-quadruplex structure and is thereby able to bind the protein thrombin. When the TBA sequence is incorporated into a hairpin structure which carries a fluorophore–quencher pair at the ends, the resulting fluorescence signal is indicative of thrombin binding (see figure 4.20c). The interaction of the enzyme shifts the equilibrium towards the G-quadruplex state, accompanied by an increase in fluorescence.

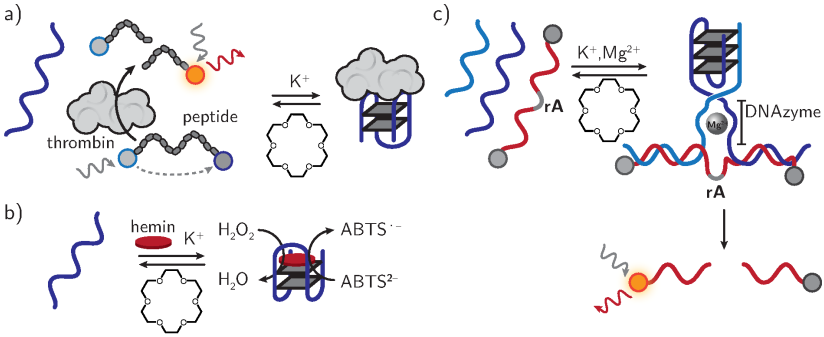


Figure 4.21.: G-quadruplex structures used to inhibit or amplify enzyme activity. a) aptamer inhibition of thrombin; b) G-quadruplex induced amplification of hemin catalysed hydrogen peroxide reduction; c) G-quadruplex based control over re- and denaturation of a Mg^{2+} dependant DNAzyme. [152]

Thrombin and its binding DNA aptamer have been used quite extensively in G-quadruplex related research. [153] An illustrative example is the controlled thrombin inhibition (figure 4.21a) of a G-quadruplex system devised by the group of I. Willner. [152] In the presence of K^+ ions the TBA aptamer sequence folds into a G-quadruplex structure which then binds to thrombin, thereby inhibiting its serine protease

activity. The inhibition or activity of thrombin is monitored by a peptide substrate carrying a fluorophore–quencher pair. When the G-quadruplex structure is disrupted by removal of the K^+ ions with a crown ether, the serine protease activity of thrombin is restored. G-quadruplex structures can also amplify the activity of an artificial enzyme mimic. It has been found out that G-quadruplexes can stimulate the peroxidase activity of hemin, an iron containing porphyrin, when the planar hemin binds to the G-quadruplex structure (figure 4.21b).[154–156] Binding of hemin to the assembled G-quadruplex structure promotes its peroxidase activity, as monitored by the catalytic oxidation of a reporter molecule (ABTS) in the presence of H_2O_2. The group of I. Willner constructed a system which is able to exhibit both the thrombin inhibition and the peroxidase amplification.[152] Switching the activity of an enzyme by re- or denaturation of a G-quadruplex structure was also implemented by the same group, by using DNAzymes[91,152,157], e. g. Mg^{2+} dependant DNAzymes showing phosphodiesterase activity (figure 4.21c).[152]

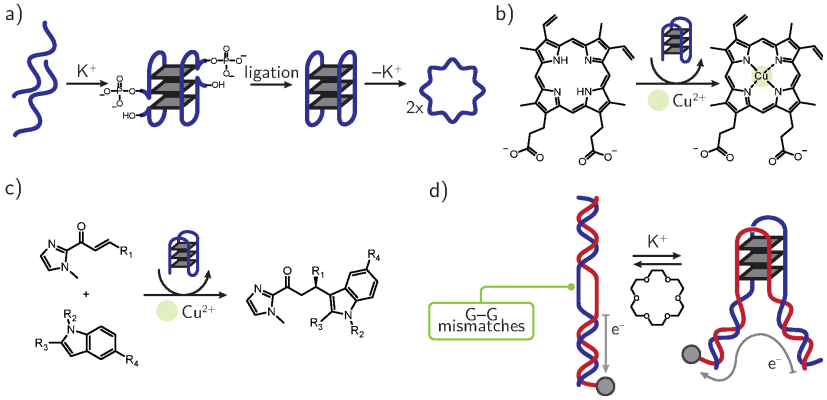


Figure 4.22.: Examples for G-quadruplexes as structural and functional tools. a) G-quadruplex templated formation of circular oligonucleotides;[158] b) G-quadruplex catalysed metallation of porphyrins;[154] c) Cu^{2+} and G-quadruplex catalysed enantioselective Friedel-Crafts reaction;[159] d) control over electron- or hole-transfer by G-quadruplex formation or denaturation.[160]

Folding of G-quadruplexes can also be used to bring reactive functionalities into close proximity. This was achieved in the synthesis of circular DNA oligonucleotides. Here, the G-quadruplex strand acts both as a structural template as well as the enzyme target substrate (figure 4.22a).[158] The ability of G-quadruplexes to bind planar molecules also enables their use as catalysts, e. g. in the metallation of porphyrin derivatives (figure 4.22b).[154] In fact, binding of the planar porphyrin is suggested to lead to a slight bending in its structure, which then facilitates deprotonation and the subsequent metal ion coordination.[161] A unimolecular G-quadruplex structure was also reported to promote a Cu^{2+} catalysed enantioselective Friedel-Crafts reaction (figure 4.22c).[159]

DNA duplex structures can show electron or hole transfer along the double helix and are therefore potential nanometre "cables" for artificial circuits and electronic devices. The ability of a G-rich sequence within a DNA duplex to fold into a G-quadruplex structure can then be used to regulate conductance (see figure 4.22d). In the G-quadruplex denatured state, electron transfer is inhibited by the guanine mismatch region in the DNA duplex, whereas renaturation of the G-quadruplex core restores efficient transfer.[160] Indeed, direct conductance of the G-quadruplex structure was reported by A. Erbe and coworkers, who could also show, by stretching the G-quadruplex molecule, that the transport of current is independent of the molecule elongation.[162]

Although the metal base pairing concept has been implemented in numerous examples for duplex DNA as seen above and despite the fact that research on metal complexes binding to G-quadruplexes is a focus of many scientists, the direct binding of transition or main group metal ions is only scarcely investigated, aside binding of alkali metal ions in the central G-quadruplex channel.

Systems bridging into this area are metal ion triggered G-quadruplex–duplex transitions. In the absence of e. g. Ag^{+} or Hg^{2+} ions, a G-rich strand folds into a G-quadruplex. When the transition metal ions are added, the G-quadruplex structure is denatured and metal base-pairs are formed between C–C or T–T mismatches, thereby stabilizing the

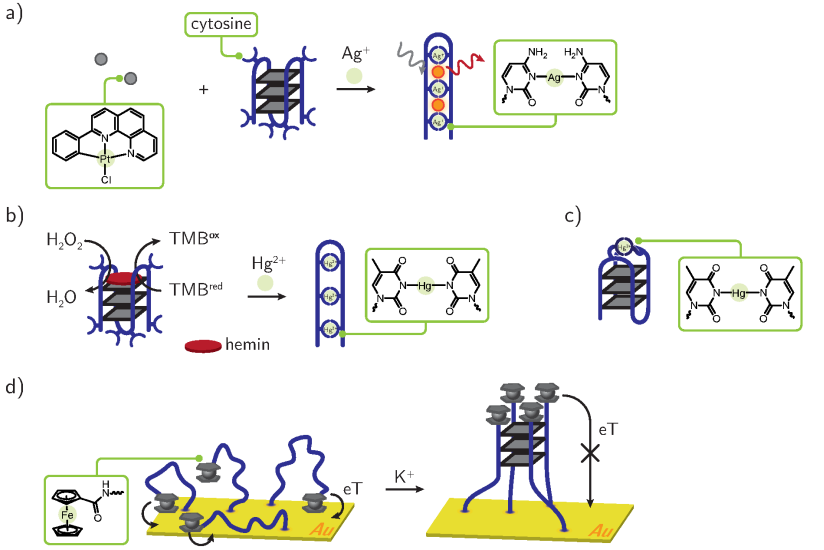


Figure 4.23.: Examples for G-quadruplex metal interactions. a) G-quadruplex–duplex transition induced by Ag^+ mediated C–C mismatch formation and detection via intercalation of a fluorescent Pt complex;[163] b) peroxidase-like activity of a G-quadruplex switched by a G-quadruplex–duplex transition via T–Hg–T mismatch formation;[164] c) T–Hg–T metal base-pair in the loops of a unimolecular G-quadruplex;[165] d) control over electron transfer in ferrocene modified oligonucleotide strands attached to a gold surface.[166]

duplex structure (see figure 4.23a and b).[164] One example combines this even with the duplex intercalator abilities of a planar Pt complex (see figure 4.23a).[163] Another approach is the combination of G-quadruplex structures with metal nanoparticles, which can be used e. g. to transfer electrons, catalyse reactions, or as the binding site for other biomolecules.[72,76]

One example of a metal base-pair incorporated into an intact G-quadruplex structure was accomplished by the group of J.-L. Mergny, who exchanged loop nucleotides in a "chair-type" unimolecular G-quadruplex for a T–Hg–T base-pair (figure 4.23c). Binding of the Hg^{2+} ion was nicely evidenced by ^{1}H-NMR and mass spectrometric analyses.[165]

Attachment of ferrocene units at the 5′ ends of a tetramolecular G-quadruplex and fixation of the G-quadruplex assembly onto a Au surface led to a system in which conductance between the ferrocene units and the Au surface could be controlled by the G-quadruplex formation (see figure 4.23d). In the absence of stabilizing K^+ ions, no G-quadruplex formation is observed, so the flexible oligonucleotide strands can bend and bring the ferrocene units into close proximity to the Au surface. Upon K^+ addition, G-quadruplex formation is triggered, and as a result, the ferrocene units are too far away for electron transfer.[166]

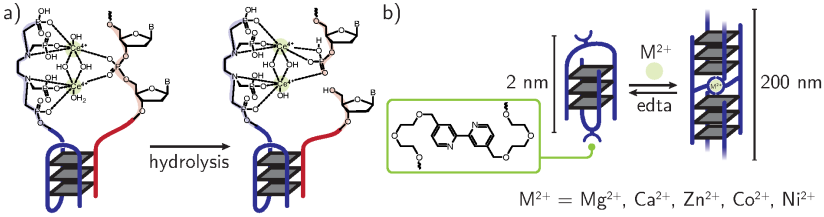


Figure 4.24.: G-quadruplexes with metal interactions. a) cerium(IV) complex catalytically hydrolysing the phosphodiester backbone of a target DNA strand.[167] b) transition between a discrete antiparallel-stranded G-quadruplex and long G-wires with parallel strand polarities induced by coordination of transition metal ions to a bipyridine ligand;[168]

One intriguing case of direct binding of a metal ion is a system capable of catalytically hydrolysing a phosphodiester bond in a G-rich target (see figure 4.24a). The system consists of a G-quadruplex with a covalently attached ligand, comprising an ethylenediaminetetraphosphoric acid (EDTP) unit.[167] The ligand binds two Ce^{4+} ions, which are assumed to be bridged by two hydroxide functions. One of the four G-quadruplex strand sections is taken by the target oligonucleotide, in this case a human telomeric sequence. This prearrangement of both the cerium catalytic centre and the target phosphodiester moiety explains the efficient phosphodiesterase activity of this complex.

Another system with direct metal coordination is a bimolecular G-quadruplex in which bipyridine units are inserted into the loop nucleotides (figure 4.24b).[168] In the absence of metal ions, apart from the necessary Na^+ ions, the modified G-rich sequence folds into discrete, antiparallel G-quadruplexes. Upon addition of either Mg^{2+}, Ca^{2+}, Zn^{2+}, Co^{2+}, Ni^{2+} ions, the G-quadruplex refolds into long G-wires, held together by the coordination complexes between the metal ions and the bipyridine units. The transformation was monitored by circular dichroism spectroscopy, gel electrophoresis, and AFM imaging. The process can be reversed by treatment with Na_2H_2EDTA and even repeatedly cycled between the G-wire and the discrete G-quadruplex state.

In summary, this chapter highlighted the state of knowledge in the field of DNA nanotechnology, thereby showing that designed DNA constructs with a structural and functional role have found many applications in past years and even slowly crossing into the field of understanding and controlling biological processes, also in terms of therapeutic targets. The same holds true for G-quadruplex structures, which become even more popular due to their unique folding abilities. The metal base-pairing concept well introduced for DNA duplex structures, is however still scarcely implemented for other DNA secondary structures, including G-quadruplexes. The many possibilities concerning incorporation of transition metal related structure and function have yet to be exploited for G-quadruplex structures. As a consequence, one major task of this thesis is to successfully transfer the metal base-pairing concept into G-

quadruplex assemblies. Before the results for this task will be presented and discussed, necessary experimental methodologies will be introduced in the next chapter.

4.3. Spectroscopic and computational methods

In this section, the basic principles of relevant spectroscopic and other methods and their application in studying G-quadruplex structures will be discussed. This includes UV-VIS spectroscopy (with thermal difference spectra, see subsection 4.3.1), circular dichroism (see subsection 4.3.2) and molecular dynamics simulations (see subsection 4.3.3). Other methods used within this work will be explained as part of the chapters 6 and 7.

4.3.1. UV-VIS spectroscopy and thermal denaturation

One of the most important features of a given supramolecular assembly is its thermodynamic stability. The same holds true for the DNA secondary structures investigated in this work, both native and modified/artificial systems, whereby one normally is interested in the equilibrium between the folded (annealed) structure and the unfolded (denatured) compound, i. e. the stability against thermal denaturation. In order to investigate the thermal stability of a given DNA secondary structure several methods can be applied. A DNA sample is prepared in a buffered solution of a certain concentration and with an appropriate electrolyte. Usually, the temperature of the sample is swept recursively from low to high temperatures (or vice versa) and an experimental value for the transition from folded to unfolded state (denaturation) is measured. This experimental value typically is for example UV-VIS absorption, a CD signal (see subsection 4.3.2), or one measured with NMR, fluorescence, or Raman spectroscopy.[169]

In the UV-VIS absorption range, the DNA nucleobases absorb from 200 nm to 300 nm, whereas the contributions from the sugar and phos-

phate backbone can be neglected (absorption ≤190 nm). Upon folding of the random coil single strand, present at high temperatures, into the DNA secondary structure, the absorption changes, due to the hydrogen bonding and π-π stacking between the nucleobases (see figure 4.25a). The difference between the high-temperature and low-temperature absorption spectrum (thermal difference spectrum or TDS) reveals those wavelengths with the highest sensitivity towards renaturation and denaturation. These TDS profiles differ considerably between each of the possible secondary structures and can therefore be used as fingerprints. For double stranded DNA a distinct hyperchromicity upon denaturation is observed at ∼260 nm. For G-quadruplexes hyperchromicity is observed at ∼243 nm and ∼273 nm (positive peaks in the TDS spectrum), whereas hypochromicity is detected at ∼295 nm (negative peak in the TDS profile, see for example figure 4.25b). The increase in absorption at lower wavelengths, however, is dominated by the temperature dependence of the single-strand absorption and cannot be used to monitor the folded–unfolded transition.[170]

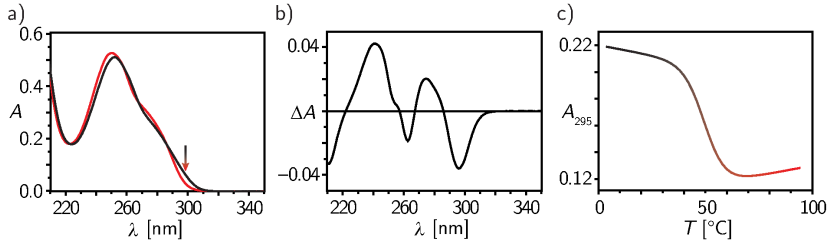


Figure 4.25.: Exemplary UV-VIS absorption spectra of a G-quadruplex (a) in the folded (black line) and unfolded state (red curve); b) thermal difference spectrum between high (unfolded) and low temperature (folded) spectra; c) thermal denaturation profile monitored at 295 nm.

Due to cooperativity, the UV-VIS signal transition follows a sigmoidal pattern, as depicted in figure 4.25c. The midpoint of the transition, called the melting temperature T_m, can be used as a measure for the thermodynamic stability of the DNA secondary structure. A system

with a higher T_m value is assumed to have also a higher thermodynamic stability in comparison to one with a lower T_m. Thus, the melting temperature gives insights into relative stabilities of different secondary structures and also ligand binding. T_m values of different samples can only be compared, if the sample conditions are kept constant, including the concentration of DNA, buffer and electrolyte, or pH of the buffered solution. Furthermore, the mere comparison of melting temperature values neglects differences in the absolute signal gap between high and low temperature absorption, plus the shape of the melting curve. The absolute values can be checked against each other using the TDS profiles. The influence of the curve's shape will be discussed shortly further below. Nevertheless, when these limitations are considered, the melting temperature comparison is a valid method when determining relative thermal stabilities.[169] Two methods of melting temperature determination will be presented in the next paragraph.

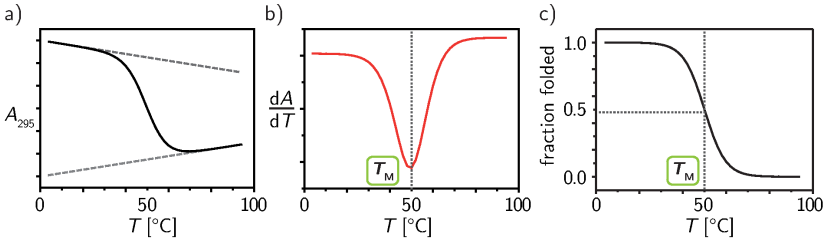


Figure 4.26.: Melting temperature T_m determination. a) original absorption denaturation profile with fitted linear low and high temperature baselines; b) 1st derivative curve with T_m as local minimum; c) fraction folded curve as calculated from the original curve and the fitted linear baselines. T_m is at $T(\alpha = 0.5)$.

The predominant method throughout the literature is to calculate the first derivative of the melting curve (see figure 4.26b) and then to determine the maximum or minimum of the resulting curve, either graphically or by fitting the curve and extracting the value from the corresponding regression function.[169] It has the advantage that it is very easy to perform and that no other data processing is necessary. T_m

calculation by this method may be biased, however, due to low signal to noise ratios or deviations from a unimolecular denaturation/renaturation process. High noise in the melting curve becomes even more dominant in the derivative curve, giving rise to high error margins for the T_m calculation, as the position of the maximum/minimum may not be so obvious. Molecularity has an even greater impact, as the maximum/minimum in the derivative curve is only equal to T_m when the thermal denaturation is an intramolecular unimolecular process and only when plotting $d\alpha/d(T^{-1})$, with α the fraction folded, see below.[169]

The second method established is an approach based on the average between two baselines.[169] In the simple picture of a transition from a completely folded sample to an unfolded one, T_m is the point where 50 % of the sample is in the folded state. Below and above the melting transition, the melting profile will show a linear dependency on the temperature. This is attributed to the temperature dependent absorption of the single-strands or the folded oligonucleotide. By fitting two linear baselines, this dependency can be accounted for. The melting temperature is then the crossing point between the average or bisecting line of the two baselines and the melting curve. Alternatively, the melting curve based on absorption can be converted to a fraction folded curve, ranging from completely unfolded $\alpha = 0$ to completely folded $\alpha = 1$ (see figure 4.26c)†, where T_m is then read out at $\alpha = 0.5$. This method normally gives more accurate results than the derivative method, although the fitting of the two baselines is a subjective approach. Equation 4.1 depicts the calculation of fraction folded α values at a given temperature, by using the original absorption values A_{295}, the high-temperature baseline absorption values A_{high} and those from the low-temperature baseline A_{low}.

$$\alpha = \frac{A_{295} - A_{high}}{A_{low} - A_{high}} \tag{4.1}$$

In principle, thermal denaturation curves can be used to extract quantitative thermodynamic parameters for the equilibrium between the

†The fraction folded is sometimes also referred to as θ. To avoid confusion with the ellipticity θ it will not be used in this thesis.

folded and unfolded oligonucleotide.[171,172] This is, however, beyond the scope of this work and will therefore not be discussed in more detail.

A few implications have to be considered when analysing thermal denaturation curves: kinetic aspects like reversibility and shape of the transition profile. All of the above mentioned methods assume that the thermal denaturation and renaturation profiles are superimposable, i. e. no hysteresis occurs, meaning that the system is in equilibrium at any temperature. Small hysteresis is commonly observed, but normally this does not impair with the conclusions drawn from the melting curves. For some systems, hysteresis will be extremely pronounced, leading to melting profiles which do not show any renaturation in the temperature range examined. This is the case for tetramolecular G-quadruplexes, as the renaturation from four single-strands is extremely slow, as it is kinetically disfavoured, and is normally not observed within the time range of the melting experiment. As the system is not in equilibrium, the apparent melting temperature is not equal to T_m, and named $T_{1/2}$ instead. Strictly speaking, an increase in $T_{1/2}$ does not imply an increase in thermodynamic stability, but a decrease of the denaturation rate (k_off) at a given temperature. It should be noted here that due to the extremely slow association process, denaturation of tetramolecular G-quadruplexes is independent from strand concentration and electrolyte strength, but is very sensitive to the temperature gradient applied during the melting curve analysis. Care must therefore be taken to maintain the same gradient throughout all experiments.[173]

Unimolecular assembly:

$$\mathrm{S}_\mathrm{unfolded} \underset{k_\mathrm{off}}{\overset{k_\mathrm{on}}{\rightleftharpoons}} \mathrm{S}_\mathrm{folded} \tag{4.2}$$

Tetramolecular assembly:

$$4 \times \mathrm{S}_\mathrm{unfolded} \underset{k_\mathrm{off}}{\overset{k_\mathrm{on}}{\rightleftharpoons}} \mathrm{S}_\mathrm{4,folded} \tag{4.3}$$

Isothermal renaturation experiments allow the determination of association rate constants, which are strongly dependent on electrolyte strength and strand concentration. Concerning the nature of the electrolyte itself, the strong thermodynamically stabilizing effect of K^+

compared to Na^+ is mainly due to a large increase in the association rate constants and only with a minor contribution from the decrease in dissociation rate. The influence of the transition metal ions used in this work will be discussed in chapter 6.

A second problem in determining relative thermodynamic stabilities may arise if the slopes of the investigated melting transitions differ greatly. A sharp transition (transition occurs over a small temperature range) implies a strong dependency of the reaction free enthalpy on the temperature, which means a high affinity constant.[169] On the opposite, a broad transition indicates a weaker dependency and therefore a smaller affinity constant. A compound with a higher melting temperature but a broader transition can thus have the lower thermodynamic stability as compared to another compound with lower melting temperature but a sharper transition.[169]

4.3.2. Circular Dichroism Spectroscopy

As discussed in the previous chapter, the chromophores in the purine and pyrimidine bases of oligonucleotides exhibit electric dipole allowed transitions, which can be analysed by standard optical UV-VIS absorption spectroscopy. In addition to these electric dipole allowed transitions, chiral molecules also show magnetic dipole transition moments. These electromagnetic transitions can be analysed by circular dichroism (CD) spectroscopy, in which the difference in absorption of left- and right-circularly polarized light is measured.

Circularly polarized light can be generated by phase shifting ($\pi/2$) two linearly polarized light beams, which propagate in a 90 ° angle to each other (see figure 4.27a). Depending on the algebraic sign of the phase difference, the circular light is either right- or left-circularly polarized (*P*- or *M*- helicity). Optically active media absorb these two oppositely polarized light beams to a different extent. Classically, this difference is expressed in degrees of ellipticity (θ, see figure 4.27c/d) . In contrast to a normal optical absorption spectrum, a CD spectrum can therefore show positive and negative bands (positive or negative Cotton effects).[174] The difference $\Delta\epsilon$, also called molar circular dichroism,

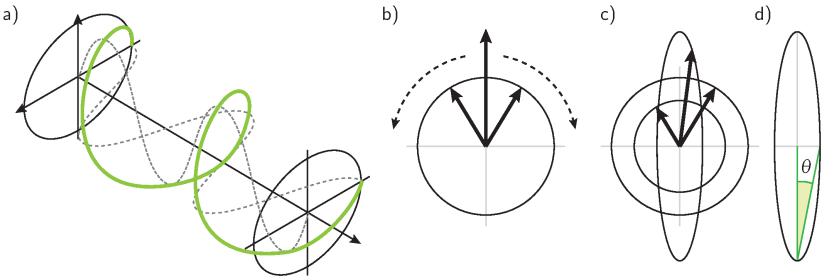


Figure 4.27.: a) Schematic representation of circular-polarized light as the combination of two linear polarized light beams with a $\pi/2$ phase difference, as is it produced in a modern CD spectrometer; b) an optically inactive medium absorbs left- and right-circularly polarized light equally, which results in no change of the polarization; c) in the case of an optically active substance, one of the two circularly polarized light beams is absorbed more strongly as compared to the other. This leads to an elliptically polarized light wave; d) the amount of ellipticity can be given as the angle θ. The optical rotation α is omitted. Adapted from ref. [174,175]

between the molar extinction coefficients of left- (ϵ_l) and right-circularly polarized light (ϵ_r) can be easily calculated from the molar ellipticity values (θ_{M}) according to equation 4.4.

$$\Delta\epsilon = \epsilon_l - \epsilon_r = \frac{4 \times \pi}{100 \times \ln(10) \times 180^\circ} \times \theta_M = \frac{1}{3298} \times \theta_M \tag{4.4}$$

The molar ellipticity θ_M is defined as:

$$\theta_M = \frac{100}{c \times d} \times \theta \tag{4.5}$$

Often, the molar circular dichroism values are divided by the number of residues in the biomolecule (e.g. number of peptides in proteins, nucleobases in oligonucleotides), giving the mean residue molar circular dichroism.

$$\Delta\epsilon_{\mathrm{mean}} = \frac{1}{n_{\mathrm{residues}}} \times \Delta\epsilon \tag{4.6}$$

CD spectroscopy is a powerful tool to investigate the secondary and tertiary structure of biomolecules, as the interaction of multiple chromophores gives rise to unique and intense CD profiles.[176] Here, a focus on the CD characteristics of G-quadruplex structures will be made.

As stated in the previous chapter, only the purine and pyrimidine moieties of the DNA contribute to the electronic absorption spectra in the UV-VIS range, whereas absorption due to the sugar or phosphate backbone can be neglected. In case of guanine, two major absorptions arise, both π-π^* transitions, with one being long axis polarized (absorption maximum at 249 nm), while the other is short axis polarized (absorption maximum at 279 nm). As the π-π^* transition is forbidden according to Laporte's rule, a single guanine nucleobase will only show a very weak electronic absorption band. The corresponding CD band is proportional to the rotational strength R, which in turn depends on both the electric dipole and the magnetic dipole moments. As a consequence the CD band is even more weak than the electronic absorption band, as the magnetic dipole transition is four orders of magnitude

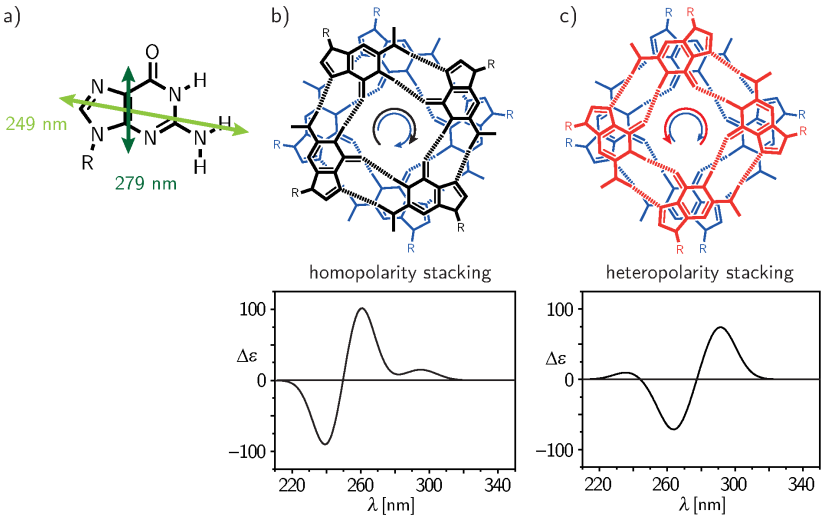


Figure 4.28.: a) Polarization axes of the two main π-π^* transitions of guanine in the UV-VIS range; Stacking arrangement of two G-tetrads and corresponding calculated CD spectrum for b) a homopolar and c) a heteropolar coupling of two G-tetrads. Figures reproduced in part with permission from a)–c) (top) ref. [178] the Royal Society of Chemistry and b)–c) (CD spectra) ref. [179] Copyright ©2008 John Wiley & Sons, Inc.

smaller than the electric dipole transition. When two or more guanosine chromophore transitions are strongly coupled, however, like in the case of a G-quadruplex, the CD effect is greatly enhanced. Depending on the distance and alignment of the stacked guanine moieties, the CD spectrum will show different Cotton effects. [174,177]

Two stacked guanine moieties can either have identical glycosidic bond angles (*syn-syn* or *anti-anti*), or different ones (*syn-anti* or *anti-syn*). Because the electric dipole transition moments are axis polarized (see figure 4.28a), two different coupling arrangements occur, giving rise to oppositely signed CD effects. The CD profile around a given transition is composed of a negative band shifted to higher or lower wavenumbers and a positive band shifted to the opposite direction,

whereas at maximum absorption (in the UV-VIS absorption spectrum) the CD effect is zero. So, a negative band around 240 nm and a positive at 260 nm is attributed to the case of a homopolarity stacking between two stacked G-tetrads (see figure 4.28b) and identical glycosidic bond angles of the stacking guanosines. These type of spectra are signed as 'type I' (or 'group I'). Likewise, a positive band at 240 nm and one at 290 nm, and a negative at 260 nm is indicative of a heteropolarity stacking (see figure 4.28c), with guanosines of different glycosidic bond angles ('type III' spectra). When both contributions are present (e.g. in three stacked G-tetrads), the resulting spectrum is a mixture of type I and III and is referred to as 'type II'. If several topologies are present simultaneously, the average CD spectrum often resembles that of type II. [177]

Type I spectra are commonly found for parallel G-quadruplexes (all strand sections pointing in the same direction), whereas type III is attributed to antiparallel strand orientation. Strictly speaking, however, the CD profiles do not allow an unambiguous assignment of the G-quadruplex topology. Firstly, the nucleobases not involved in G-tetrad formation can still stack onto the G-tetrads and will therefore contribute to the CD spectrum, secondly, aggregation of G-quadruplexes to higher order structures can alter the CD spectrum significantly, and thirdly (as mentioned above), several equally stable confirmations may be present simultaneously, giving rise to an average CD spectrum. For a correct assignment other techniques like NMR structure elucidation or X-ray analysis might be necessary. Still, CD spectroscopy gives valuable insights and indications into G-quadruplex stability and topology. [177,180]

4.3.3. Molecular Dynamics Simulations

Probing the topology, stability and interactions of biomolecules by experimental methods like spectroscopy often requires the generation of molecular models to visualize and test the assumptions and conclusions drawn from the experimental data. With computers becoming more and more powerful, the branch of computational chemistry and molecular modelling has advanced significantly as well. Although quantum

chemical calculations can be done in a reasonable amount of time for small molecules, quantum chemical description of large biomolecules like nucleic acids and especially proteins is still not feasible. Efforts have been made to tackle this problem and have lead to the implementation of molecular dynamics simulation and energy minimization.

In classical molecular dynamics simulations, all interactions between atoms are treated with classical mechanics, by solving Newton's second law of motion (equation 4.7). [181]

$$m_i \frac{\partial^2 r_i}{\partial t^2} = F_i, \quad \text{with } i = 1 \ldots N. \tag{4.7}$$

Here, m_i is the mass of the particle/atom i, r_i its positional vector, t the time, and the force F_i. When both the forces acting on an atom and its mass are known, then the time-dependant movement, including positions, velocities and accelerations can be derived. Using initial coordinates and velocities for each atom of a molecule at a time t, new positions are calculated for a small time interval Δt by integration of equation 4.7. After that the forces arising between the atoms are recalculated. [182] The forces F_i can be described by the derivative of a potential function:

$$F_i = -\frac{\partial V}{\partial r_i} \tag{4.8}$$

Several interactions between atoms contribute to this potential function: Non-bonded forces like Coulomb or dipole interactions (with dispersion), and bonded interactions like covalent and dative bonds. [181,182] The potential functions need to be parametrized by fitting the energies obtained from molecular dynamics (Newton mechanics) to those derived from high level quantum chemical calculations on small reference systems. This set of potential functions and its parameters is called a force field. Numerous force fields are available, the application of a given force field depends on the investigated system, e. g. protein vs. nucleic acid. Although very simple in terms of the theoretical basis, molecular dynamics simulations can nevertheless describe large biomolecules with good accuracy, also in comparison to solution state NMR and solid state crystallographic data, as the underlying force fields are designed to reproduce high level quantum chemical energies and conformations. [181]

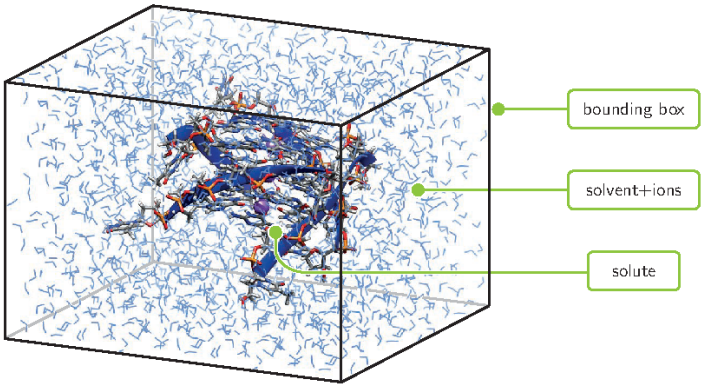


Figure 4.29.: A G-quadruplex structure (solute) in an explicit solvent box as used in MD simulations. The water solvent molecules are depicted as blue wires.

At the start of the typical MD procedure, an initial structure is constructed, e.g. from analogues, already known structures, and placed into a symmetric box, the simplest one being a cube or cuboid (see figure 4.29), which is then virtually filled with solvent molecules in the case of an explicit solvent approach and other compounds like ions for charge neutralization. Normally, the contents of the box (solvent and solute) are brought to a constant pressure and temperature (NpT ensemble). By introducing periodic boundary conditions (PBC) infinite copies of the generated box are produced, mimicking a macroscopic sample.[182] This has the advantage of low computational costs as all of the copies will behave exactly the same, and avoids problems of solvent–vacuum boundaries. It comes at the price that long range interactions have to be cut off, as the individual box size has to have rather small extensions.[181] Starting from the given structure and based on the potential function(s) (with their respective parameters), Newton's equations are solved and new atom positions are generated. At this point, the forces are updated and the cycle is repeated. Using small time steps, the atoms' positions are tracked over a certain time period,

called trajectory. Scheme 4.1 depicts a simple outline of the steps made in a MD simulation as described in the manual for the MD software GROMACS.[181]

Simultaneously to the growing experimental details on canonical and non-canonical DNA structures, molecular dynamics simulations of these assemblies have progressed as well.[183–186] This is especially true for four-stranded DNA structures.[187–189] Not only have reference simulations of non-canonical nucleic acids been developed,[190] but it was also found that G-quadruplexes can be used as force field benchmark systems for canonical DNA as well, due to the complexity arising from the different torsional angles (especially glycosidic) present in these topologies.[191] Additional effort has been made to understand the preference of guanine nucleosides to certain glycosidic torsion angle patterns,[44] the folding process from single strands to the G-quadruplex as the thermodynamic product,[192,193] as well as the formation of higher order structures[194] and G-quadruplex/ligand interactions.[187]

All these investigations, however, are focused on unmodified DNA, i. e. oligonucleotide strand sequences which can be found in the genome of living organisms. As this work will bring the natural occurring DNA strands together with metal-ligand complexes, thus forming artificial hybrid systems, it is necessary to incorporate these new features into the existing MD protocols and force fields. The dynamic bond properties of the metal–ligand dative bond, in comparison to a covalent bond, are quite difficult to parametrize, so the need for a correct structural and electronic description of the metal coordination environment[195,196] has to be balanced with that for allowing dative bond breaking and (re-) formation.[197,198]

Table 4.1.: Simplified MD algorithm. Adapted from ref. [181]

Global MD algorithm

1. Input initial conditions
Potential interaction V as a function of atom positions
Positions r of all atoms in the system
Velocities v of all atoms in the system

Repeat 2, 3, 4 for the required number of steps:

2. Compute forces
The force on any atom is computed by calculating the force between non-bonded atom pairs plus the forces due to bonded interactions (which may depend on 1, 2, 3, or 4 atoms), plus restraining and/or external forces.
The potential and kinetic energies and the pressure tensor may be computed.

3. Update configuration
The movement of the atoms is simulated by numerically solving Newton's equations of motion.

4. If required: Output step
write positions, velocities, energies, temperature, pressure, etc.

5. Ligand synthesis

In this chapter the organic syntheses of the ligands used in this thesis will be discussed, based on the task set in the introductory chapter 1, specifically the design of monodentate ligands which can be covalently appended to DNA oligonucleotide strands. Based on the successful synthesis, the ligands are then incorporated into guanine-rich oligonucleotides, which will be discussed in the following chapters.

5.1. Synthesis of first-generation ligand $\mathbf{L^1}$ and modifications

As outlined in the introductory chapter, the first task towards a metal-base tetrad in G-quadruplexes is the appropriate design of the metal ions and its ligand. Based on the nature of both the metal coordination site and the G-quadruplex part, the following requirements could be established:

- simple and fast ligand synthesis, in order to obtain a proof-of-principle case
- compatibility with an aqueous, buffered solution containing high salt/electrolyte concentrations, to guarantee formation of the desired complex
- tolerance of coordinated metal and ligands by the G-quadruplex structure and *vice versa*
- preferentially fast metal–ligand complexation kinetics and reversibility of the metal–ligand bond, to avoid kinetic traps

- steric hindrance as small as possible, preferentially planar metal–ligand environment

As discussed in the previous chapter, all these requirements can be met in case of DNA duplex structures. Unfortunately they are difficult to be directly transferred to quadruplex structures, especially the compatibility with the planar G-tetrad moiety.

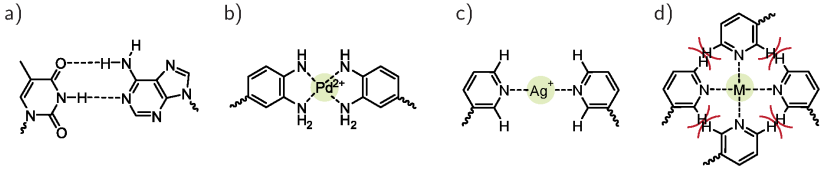


Figure 5.1.: Comparison between ligand/nucleobase orientations in natural and artificial metal base-pairs, as well as the metal base-tetrad. a) planar, canonical T=A Watson-Crick base-pair; b) planar, *o*-phenylenediamine Pd^{2+} metal base-pair; c) planar, pyridine Ag^{+} metal base-pair; d) hypothetical planar, pyridine metal base-tetrad, which is not formed due to steric repulsion between the hydrogen atoms in *ortho* position to the pyridine nitrogen.

In figure 5.1, the base-pairing found between natural base-pairs and artificial metal base-pairs is compared to that of a pyridine-based metal base-tetrad. In the exemplary case of a metal base-pair composed of the square-planar coordination of *ortho*-phenylenediamine to a metal ion (figure 5.1b), the aromatic moieties can adopt a co-planar arrangement with respect to the metal coordination environment. The metal base-pairing thus mimics the same co-planar arrangement found for the natural nucleobases in Watson-Crick base pairing. The co-planarity in the metal base-pair can be explained by the fact that the amine donor functionalities prevent a close proximity between the aromatic units. This kind of ligand is unfortunately unsuitable for the construction of a metal base-tetrad, as the chelate ligand has already two donor functionality. In the pyridine–Ag^{+}–pyridine metal base-pair (figure 5.1c) these spacers are not present and the nitrogen donors are directly part of the aromatic rings. Here, however, the linear coordination environment

prevents any steric hindrance between the pyridines. In contrast, the construction of a metal base-tetrad based on the coordination of four pyridine (figure 5.1d) is not possible, as the hydrogen atoms in *ortho*-position to the pyridine nitrogen atoms come into close proximity. As a result, the pyridine rings rotate out of the metal coordination plane.

Except for small-molecule ligands, nearly all cases of planar ligand environments around a square-planar coordinated metal centre are constructed from chelating or macrocyclic ligands like porphyrins, therefore unsuitable for the attempted scaffold. An exception are the 12-metallacrown-4 systems first introduced by Gibney et al.[199] from the group of V. Pecoraro, which are composed of four triple donor hydroxamate ligands coordinating to four or five metal ions (see figure 5.2a). Four of the metal cations are circularly arranged between the coplanar ligands, while the fifth metal cation is positioned in the centre of the assembly. On the one hand these systems can be truly planar, but on the other hand their synthesis is not straightforward and easily adapted for the aqueous environment. Instead, the profound knowledge in the group about the coordination behaviour of metal–pyridine assemblies was used as the basis of a first, simple ligand.[200] The square-planar metal–pyridine coordination environment is not planar as discussed above, but is easily synthesized and offers direct modifications at the pyridine ring (see figure 5.2b). It is compatible with an aqueous medium as pyridine is a stronger ligand than the water solvent molecules, and offers fast or moderately fast kinetics in case of the first-row transition metals. This is especially true for the Jahn-Teller distorted metal coordination environments like in Cu^{2+}.[12] Although transition metal ions are known to coordinate to DNA phosphate groups and purine and pyrimidine nucleobases[201,202], it was anticipated that the preorganized chelating pyridine ligands should bind more strongly to the respective metal ion, thus this arrangement should be favoured in relation to all other possible interactions.

The remaining question was how to set up the chemical linkage between the DNA and the pyridine unit. As the ligand should be incorporated into the oligonucleotides via the phosphoramidite method

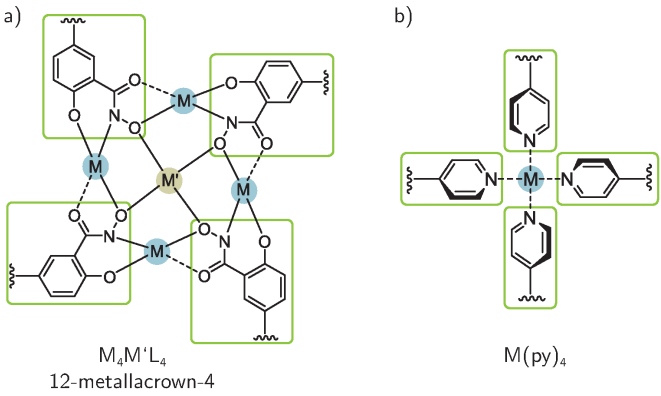


Figure 5.2.: a) Example of a 12-M-4 metallacrown based on hydroxamate ligands[199]; b) schematic illustration of the metal–pyridine square-planar coordination environment. M = transition metal ion.

(see the next chapter 6), a free hydroxyl function opposite to the pyridine donor is needed. As a straightforward synthesis was anticipated, the easiest linker is an alkyl chain, connected to the pyridine via an ether bridge. Direct attachment of the alkyl chain is in principle possible, e. g. by a Suzuki cross-coupling reaction,[203] but involves an additional step and has the risk of carrying over Pd^{2+} traces.

For the question of the appropriate linker length, molecular modelling was used for evaluation. A first, very simplistic model of the desired G-quadruplex structure (see figure 5.3) was created in the following way: All manipulations were carried out in the software Spartan '08.[204] For the G-quadruplex stem the single crystal X-ray structure of a parallel G-quadruplex drug complex ($[d(G_4)]_4$ co-crystallized with daunomycin in the presence of Na^+ and Mg^{2+}; resolution 1.08 Å)[205], was used. Daunomycin, water molecules and cations, except for the three central Na^+ ions, were removed and one propoxypyridyl ligand attached via a phosphate ester at each 5′ end. A square-planar coordinated Cu^{2+} ion was placed in between the donor atoms of the ligands. All atom positions were frozen, with the exception of the 5′-CH_2–O–phosphate

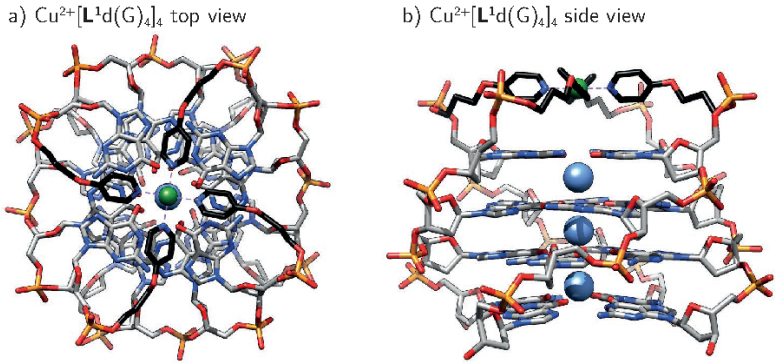


Figure 5.3.: Molecular model for $Cu^{2+}[\mathbf{L^1}d(G_4)]_4$ from a) top view and b) side view. The G-quadruplex stem is taken from the crystal structure of a $[d(G_4)]_4$ drug complex, PDB entry 3TVB. The ligand and Cu^{2+} part were optimized on the MMFF level. Hydrogen atoms have been omitted for clarity. Carbon atoms are coloured grey (black for ligand $\mathbf{L^1}$), nitrogen blue, oxygen red, phosphorus yellow, sodium blue, and copper green, respectively.

linkages, as well as the ligands and the Cu^{2+} ion. The coordination environment of the Cu^{2+} ion was constrained to a square-planar geometry, while another constraint was set for its position on the central axis between the central Na^+ ions. The so obtained model with the active constraints was then geometry optimized at the MMFF level (build in force field of the Spartan software). When the Cu^{2+} ion is fixed in position too far away from the top G-quartet, the flexible alkyl linkers of the ligand $\mathbf{L^1}$ have to adopt a conformationally disadvantageous, linear and stretched arrangement. Once the distance constraint is removed, the structure relaxes, moving the copper ion and its pyridine ligands closer to the G-quadruplex stem, with a final Cu^{2+}–Na^+ distance of 5.8 Å. As no unusual bond lengths or angles were observed, and the structure still adopted its native four-fold symmetry, it was deduced that this linker length and ligand design is suitable for the tasks set in the introductory chapter.

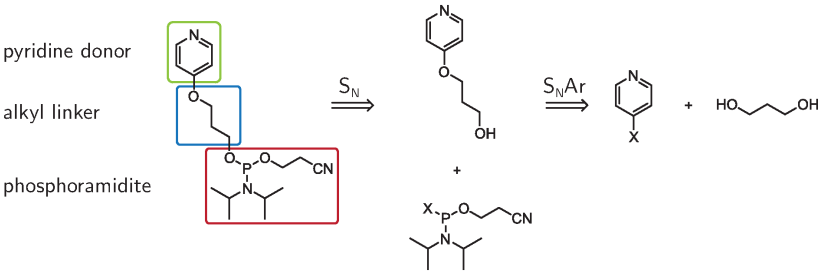


Figure 5.4.: Retrosynthesis of ligand $\mathbf{L^1}$ phosphoramidite via two nucleophilic substitution reactions (S_N and S_NAr).

Using this design, the ligand synthesis could be established by a retrosynthetical dissection. The most straightforward way proceeds via two nucleophilic substitution reactions, one being an aromatic nucleophilic substitution reaction at the pyridine ring (see figure 5.4). The synthesis of the pyridyloxy substituted propanol is already described in the literature, so this procedure could be directly used.[206] It commences from 4-chloro-pyridine hydrochloride **1**, which is treated with 1,3-propanediol **2** in the presence of a base. The hydrochloride salt of the pyridine (pyridinium cation) is used instead of the free base to prevent decomposition due to **1** reacting with itself, leading to oligo- and polymerization. An excess of the diol minimizes double substitution, whereas the base deprotonates the pyridinium cation and acts as a scavenger of the HCl byproduct formed during the reaction. The product **3** is obtained in moderate yields, consistent with the literature values, and is then transformed in the next step to the phosphoramidite **5**. As both the chlorophosphitylating agent **4** and the phosphoramidite product are moisture and air-sensitive, the reaction is carried out under an inert nitrogen atmosphere and with anhydrous solvents.

Hünig's base NEt*i*Pr$_2$ (DIPEA) traps the formed HCl and precipitates as the corresponding ammonium salt in the solvent THF, thus shifting the reaction equilibrium further to the phosphoramidite side. A slight excess of phosphitylating agent is used to fully convert the alcohol **3**, as the latter would impair the DNA solid-phase synthesis due to its

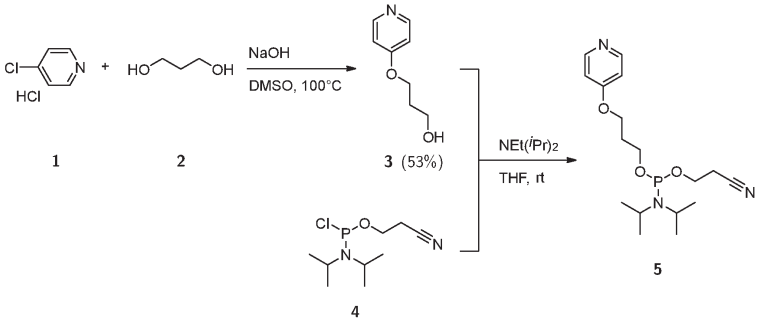


Figure 5.5.: Synthesis of ligand $\mathbf{L^1}$ phosphoramidite **5**.

nucleophilic reactivity. Normally, an aqueous reaction work-up is done to remove the ammonium salt byproduct. In this case, however, it was found that the phosphoramidite **5** is too unstable and readily hydrolysed when exposed to moisture. This step was therefore omitted and the crude product purified only by column chromatography on deactivated silica, under an inert atmosphere and using degassed solvents. The so obtained phosphoramidite **5** was then either used directly on the same day or frozen at –32 °C and then used as fast as possible to avoid further decomposition. The compound is readily characterized by standard ^{1}H-NMR, although incompletely removed NEt*i*Pr$_2$ leads to a strong signal overlap in the aliphatic region of the spectrum. Nevertheless, the observed $J_{\text{P-H}}$ couplings show the bonding between the phosphoramidite part and the linker moiety of the ligand. Unambiguous proof of the successful synthesis is achieved with ^{31}P-NMR, where the phosphoramidite shows a characteristic shift at around $\delta = +148$–149 ppm. Hydrolysed or oxidized phosphoramidite or phosphitylating agent, e. g. the phosphonate, are detected at lower chemical shift values, between $\delta = \sim$2–10 ppm.[207–209] The ESI mass spectrum additionally shows the correct composition, here in the form of the $[M+Na]^+$ adduct (see figure 5.6).

In order to detect if the pyridine moiety is indeed able to take part in metal ion coordination, reference oligonucleotide strands lacking this ability needed to be synthesized. Ideally, this control should be

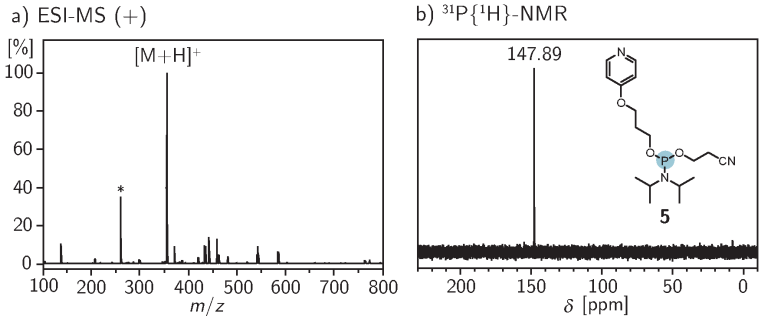


Figure 5.6.: ESI-MS(+) and $^{31}P\{^{1}H\}$-NMR of phosphoramidite **5**. *hydrolysed phosphitylating agent.

as similar as possible to the original strands. Accordingly, the ligand $\mathbf{L^1}$ is altered only in the position of the pyridine unit. Based on a similar procedure as the one described above, reference ligand ($\mathbf{L^{1*}}$) phosphoramidite **9** is obtained (see figure 5.7), which carries a phenyl ring instead of the pyridine group. As the phenyl group does not contain a donor functionality like nitrogen, it should be unable to take part in metal ion coordination. The synthesis is achieved according to a literature procedure[210] and includes the linkage between phenol **6** and 3-chloro-propanol **7** by a nucleophilic substitution reaction to yield the alcohol **8**. This precursor is then transferred to the phosphoramidite **9** as described above, with successful characterization again by ^{1}H-, ^{13}C-, and ^{31}P-NMR.

Both ligands ($\mathbf{L^1}$ and $\mathbf{L^{1*}}$) now enable to conduct proof-of-principle experiments on the metal ion coordination to donor-functionalized tetramolecular G-quadruplexes. The results for these ligands will be described and discussed in chapter 6.

In the beginning, the correct linker length was determined with basic molecular modelling using a crystal structure as a reference. To test if the assumptions made there about the length of the alkyl linker are valid, additional derivatives of ligand $\mathbf{L^1}$ were constructed by varying exactly this length. For this purpose, ligands $\mathbf{L^{1b}}$–$\mathbf{L^{1d}}$ with one ($\mathbf{L^{1d}}$), two ($\mathbf{L^{1b}}$), and four ($\mathbf{L^{1c}}$) methylene groups were synthesized using the same procedure as for $\mathbf{L^1}$ (figure 5.9).

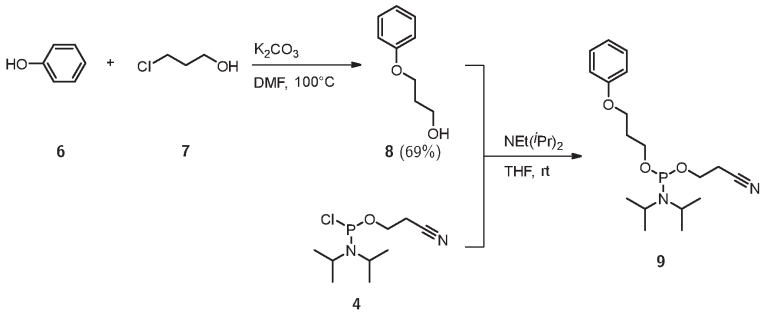


Figure 5.7.: Synthesis of ligand $\mathbf{L^{1*}}$ phosphoramidite **9**.

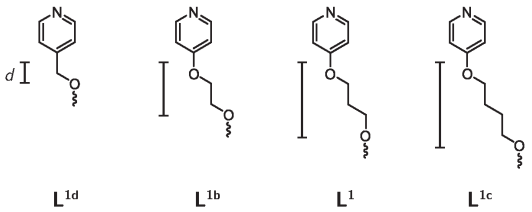


Figure 5.8.: Comparison of linker length variation for ligand $\mathbf{L^{1d}}$, $\mathbf{L^{1b}}$, $\mathbf{L^{1}}$, $\mathbf{L^{1d}}$.

For convenience, the shortest linker length, ligand $\mathbf{L^{1d}}$, with one methylene group was constructed from (4-pyridyl)methanol, instead of (4-pyridyloxy)methanol. The latter would be too unstable, most likely immediately decomposing to (4-hydroxy)pyridine and formaldehyde. It should be noted that as **13** ($\mathbf{L^{1d}}$) lacks this additional oxygen atom, the linker is significantly shorter than that of the next ligand in the series (**11a**,$\mathbf{L^{1b}}$), in comparison to the linker length variation between **11a** and **3** ($\mathbf{L^{1}}$). Ligand $\mathbf{L^{1d}}$ thus poses as an extremely short ligand with nearly no conformational flexibility, whereas the longest ligand $\mathbf{L^{1c}}$ exhibits a large degree of conformational freedom. The different type of substituents attached at the *C*4 position of the pyridine ring, namely –OR versus –CH_2R may also influence the stability of the pyridine–metal bond, as the –CH_2R substituent exhibits no mesomeric effect in contrast to the –OR substituent.

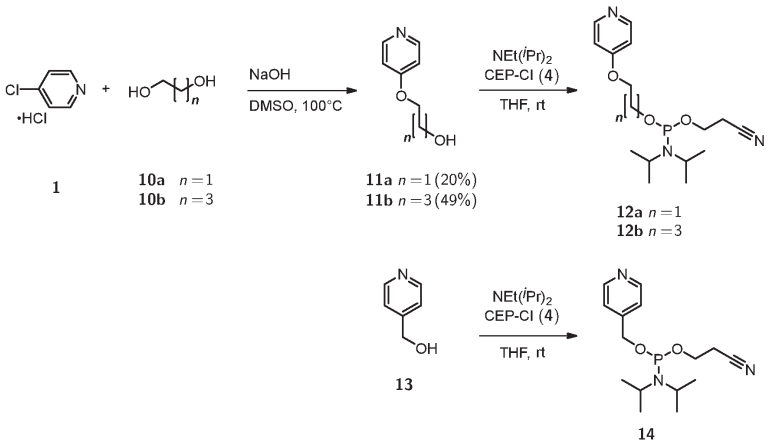


Figure 5.9.: Synthesis of ligand $\mathbf{L^{1b}}$, $\mathbf{L^{1c}}$, $\mathbf{L^{1d}}$ phosphoramidites **14**, **12a**, **12b**.

Concerning the synthesis of (4-pyridyloxy)ethanol **11a** a significantly lower yield (20 %) than for the analogues compounds with longer alkyl chains (53 % and 49 % for **3** and **11b**, respectively) was obtained. Although the original literature procedure by Katrizky et al.[211] from

1987 states a 72 % obtained yield and recently published patents (see for example ref.[206]) also reference this routine, the same experimental recipe did not yield the clean product, even with variations in reaction solvent, temperature, duration, and work-up. This was partly attributed to the fact that the product and the in excess used ethylene glycol behave very similar in terms of polarity and are therefore difficult to separate (during the work-up and column chromatography). Lowering the amount of ethylene glycol led to a significant build-up of disubstituted glycol. Known literature procedures based on the reaction of 4-hydroxypyridine with 2-chloroethanol[212] or 4-chloropyridine with ethylene glycol in the presence of sodium as a base[213] also state low yields (both 19 %) in accordance with the value found in this work.

For the last step, i. e. reaction with the phosphitylating agent to the corresponding phosphoramidite, the same procedure as described above was employed, however with one modification. As it turned out that in most cases the column chromatography step did not greatly purify the phosphoramidite product and even led to a significant hydrolysis or oxidation, this step was omitted in all reactions. To reduce the amount of byproducts, e. g. unreacted phosphitylating agent, the latter amount was reduced to ∼1.1 equivalents, while still obtaining full conversion of the respective alcohol. It should be noted that a common purification technique, namely precipitation of the phosphoramidite in hexane, failed for the compounds described here, so this procedure could not be used. The work-up thus consisted solely of precipitating the Hünig's base ammonium salt with THF, filtration of the reaction solution with a syringe filter and evaporation of the solvents. Just prior usage, the phosphoramidite was redissolved in anhydrous acetonitrile, followed by an additional filtration step. The phosphoramidite was then sufficiently pure for the subsequent DNA solid phase synthesis.

5.2. Synthesis of second-generation chiral ligand L^2

All of the above mentioned ligands have a common major shortcoming, as the incorporation into DNA oligonucleotide strands is solely limited to the attachment at the 5′-end. Although suitable for a proof-of-concept for the metal-base tetrad motif and the testing of linker length variations, this ligand design is only valid for tetramolecular G-quadruplexes, preventing the study on unimolecular assemblies, with the metal-base tetrad in internal positions (or the 3′-ends).

As a consequence, a second-generation ligand design needed to be established, which allows for all these requirements. The alcohol functionality of compound **3** already served as the 3′ connection to the 5′ end of the next nucleotide via the phosphate linker. Thus, an additional alcohol function, serving as the 5′-end attachment position to the next nucleotide, had to be introduced.

In the chapter about DNA backbone modifications (chapter 4.2, page 24), several candidates were already discussed. In order to be as similar as possible to the first-generation ligands and to be as simple as possible, the glycol-based backbone introduced by the group of Eric Meggers [96,99] was chosen. This ligand design has several benefits: first, it is nearly identical to the established ligand $\mathbf{L^1}$ apart from the additional –CH_2–O– functionality, second, the GNA backbone allows for the simple alteration of the donor functionality, thereby easily accommodating possible future modifications, and third, the possibility to introduce a chiral centre into the ligand backbone (see also figure 5.10 for a direct visual comparison between the structures of ligand $\mathbf{L^1}$ and $\mathbf{L^2}$).

The synthesis of the ligand $\mathbf{L^2}$ phosphoramidite was not based on the one introduced by Zhang et al., [96] but instead based initially on the synthesis of a more similar pyridine compound reported by Ikegashira et al. [214], where the pyridine is substituted in the 2-position with a azetidine-1-carbonyl based functionality. The corresponding retrosynthetic dissection is depicted in figure 5.11.

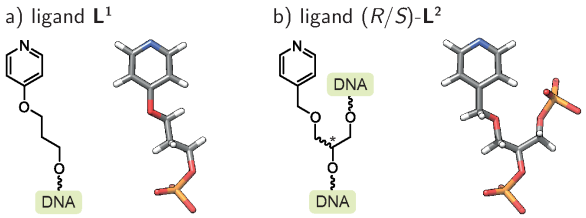


Figure 5.10.: Comparison between a) first-generation ligand $\mathbf{L^1}$ and b) second-generation ligand (R/S)-$\mathbf{L^2}$.

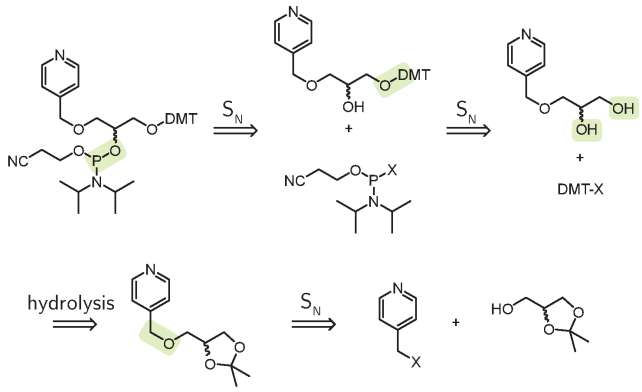


Figure 5.11.: Retrosynthesis of ligand (R/S)-$\mathbf{L^2}$ phosphoramidite via three nucleophilic substitutions and a hydrolysis reaction.

Here, commercially available solketal is used, either in the enantiomerically pure (*R*)- or (*S*)-form, which is coupled to the pyridine unit by a nucleophilic substitution reaction. For the 4-pyridylmethyl synthon a good leaving group has to be present, i. e. a mesylate functionality in case of the original procedure.[214] The 4-pyridylmethyl mesylate **13a**, however readily obtained in good yields from 4-pyridylmethanol (**13**), was found to be prone to decomposition even at lower temperatures. This is likely due to the nucleophilic reactivity of the pyridine nitrogen towards the mesylate carbon, leading to oligo- or polymerization. All attempts to yield the coupling product **16** were therefore unsuccessful. The reverse reaction, i. e. using the 4-pyridylmethanol as the nucleophile and the solketal mesylate **15a**[215–217] as the electrophile, failed likewise, yielding only trace amounts of the desired product (see figure 5.12).

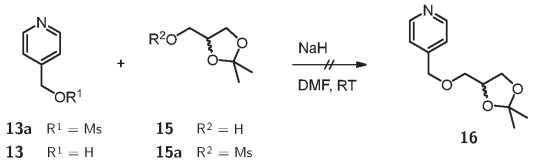


Figure 5.12.: Anticipated synthesis of pyridine solketal **16** via the mesylate route.

To circumvent the high reactivity of the free pyridine base towards itself, the corresponding protonated form, namely the hydrobromide salt, was used. Protonation should lower the nucleophilic ability of the nitrogen atom dramatically. Simultaneously, the leaving group was changed to the more unreactive analogues bromide. The 4-bromomethyl pyridine hydrobromide salt **17** can be either easily synthesized from 4-pyridylmethanol **13**[218] or purchased directly from commercial sources. With these modifications of the synthesis, the desired product could be obtained in excellent yields. The full synthetic route is depicted in figure 5.13.

The next step requires the release of the alcohol functionalities by the acid catalysed hydrolysis of the isopropylidene (acetonide) protection group in solketal **16**, yielding the free diol **18**. Numerous examples

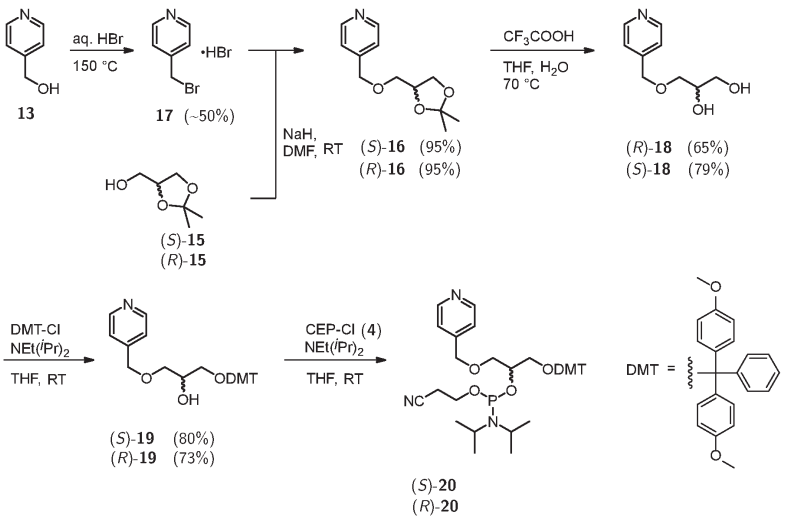


Figure 5.13.: Synthesis of ligand $\mathbf{L^2}$ phosphoramidite (*R*/*S*)-**20**.

for this reaction are available in the literature, the most common acid catalysts being hydrochloric acid,[216,219–222] acetic acid,[223–225] or *para*-toluenesulfonic acid[226] (see also ref.[227]). Carrying out the reaction with hydrochloric acid at room temperature is straightforward, with a fast release of the desired product. The workup requires, however, washing steps with water to remove salts formed after the neutralization. This is disadvantageous, as the product is also soluble in water and is difficult to extract with common organic solvents, like ethyl acetate or dichloromethane. Using acetic acid or *para*-toluenesulfonic acid led to low yields, even when the reaction was carried out at elevated temperatures. Additionally, using acidic ion exchange resins was also unsuccessful, as the pyridine product became too strongly adsorbed in the resin and could only be partially recovered. Instead, the more acidic trifluoroacetic acid was used at 70 °C in a water/THF mixture. After neutralization with aqueous ammonia the solvents could be evaporated under reduced pressure, thereby also removing all used reagents. Purification by column chromatography then yielded the desired product **18** in good yields.

The last two steps of the synthesis are those routinely employed to prepare the ligand for the standard solid-phase DNA synthesis. First, the primary alcohol functionality of **18** has to be capped with the dimethoxytrityl (DMT) protecting group, whereas secondly the secondary alcohol is coupled with the phosphitylating agent (CEP-Cl) to yield the corresponding phosphoramidite **20**. This approach and the general usage of both functionalities will be discussed shortly at the beginning of the next chapter 6. Often, the first reaction is carried out with dimethoxytrityl chloride and in the presence of pyridine, the latter acting both as a base and solvent.[227–229] Using this protocol, the conversion to the product **19** was found to be rather low and the isolated phosphoramidite was contaminated with significant amounts of pyridine, which was very hard to remove. Thus, the overall yield remained low, even despite variations in stoichiometry and solvent. As a consequence two major alterations were made: first, the base was substituted for Hünig's base DIPEA, which forms a weakly soluble ammonium salt in THF with the byproduct HCl (as already stated above), and second, 4-dimethylaminopyridine (4-DMAP) was used as a catalyst. Care was taken, both that the reaction conditions were anhydrous to prevent hydrolysis of the DMT-Cl, and that during the column chromatography purification step the solvents contained a small amount of triethylamine or pyridine (0.1 % v/v) to prevent acid catalysed hydrolysis of the phosphoramidite. With these modifications to the procedure, the product **19** was obtained in good yields.

The second and last reaction, namely the coupling with the phosphitylating agent to yield the phosphoramidite (R/S)-**20**, was carried out according to the procedure as described above for the synthesis of **14**. Both pairs of diastereomers of **20** could be isolated in good yields and were successfully characterized by ^{1}H-, ^{13}C-, and ^{31}P-NMR. The trivalent phosphorus has a non-planar trigonal pyramidal bonding environment and thus is chiral, while the energy barrier for inversion is quite high, so both enantiomers are stable at room temperature and do not racemise.[230] The two stereocentres present in **20** results in the formation of stereoisomers, namely two pairs of diastereomers. As a

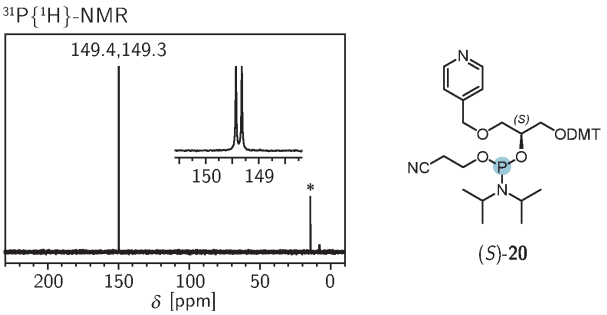


Figure 5.14.: $^{31}P\{^1H\}$-NMR of phosphoramidite (*S*)-**20**. The inset shows the two signals observed for the two diastereomeric pairs. *hydrolysed phosphitylating agent or product.

consequence, the NMR signals split into two identical sets, which renders signal assignment and complete characterization difficult. The two sets of signals are best seen in the $^{31}P\{^1H\}$-NMR spectrum of **20** (see figure 5.14). Here this effect has the advantage of an additional proof, apart from the chemical shift value, for the successful synthesis.

Similar to the synthesis of a reference ligand $\mathbf{L^{1*}}$ for ligand $\mathbf{L^1}$, the same was done for the second generation ligand $\mathbf{L^2}$, resulting in the synthesis of reference ligand $\mathbf{L^{2*}}$. Using nearly the same procedure as described for ligand $\mathbf{L^2}$ phosphoramidite **20**, the synthesis was achieved starting from benzyl bromide **21** (see figure 5.15).[225] Due to time considerations, this reference ligand was however not incorporated into DNA oligonucleotide strands. The synthetic procedures are nevertheless depicted in the experimental section, for the sake of completeness.

Now that the phosphoramidite syntheses are successfully accomplished, the ligands can be incorporated into DNA oligonucleotide strands, which will be discussed in the next chapter.

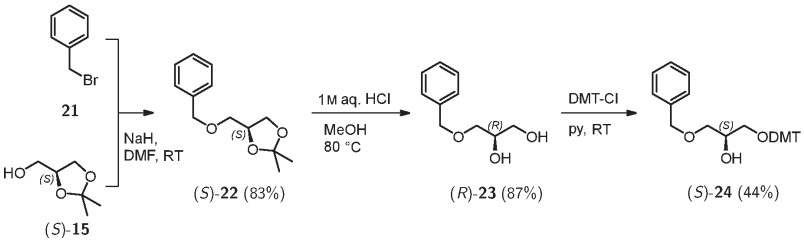


Figure 5.15.: Synthesis of DMT-protected ligand $\mathbf{L^{2*}}$ (*S*)-**24**.

6. Tetramolecular G-quadruplexes

Once the ligand syntheses are established, as described in the previous chapter, the different ligands can be incorporated into DNA oligonucleotide strands. In this chapter, the results for ligand modified tetramolecular G-quadruplexes will be presented, whereas the following chapter 7 will deal with unimolecular G-quadruplexes based on telomeric DNA oligonucleotides.

The actual DNA oligonucleotide synthesis follows the phosphoramidite method, which was developed several decades ago and is now routinely used by solid-phase synthesis and employing a DNA synthesizer. Each consecutive step in the assembly of the desired oligonucleotide follows a coupling cycle, which will be described shortly in the next paragraph (for an overview see figure 6.1).

Within the coupling cycle, several key features are implemented which in sum lead to the highly efficient and fast oligonucleotide synthesis. The first nucleobase is attached covalently to a solid, porous resin, namely controlled pore glass (CPG), a technique which was taken from the established solid-support synthesis of peptides on a polymeric resin as introduced by Robert B. Merrifield.[231] This significantly decreases synthesis time scale, as the laborious work-up and purification necessary for each step when employing a solution state synthesis is avoided, that is, the reagents are just pumped trough the solid resin, while byproducts are washed away.

Reactive functional groups in the purine and pyrimidine nucleobases are capped with base labile protecting groups (iso-butyryl for the C2-NH_2 group in case of guanine), while the 5′-OH is protected with the

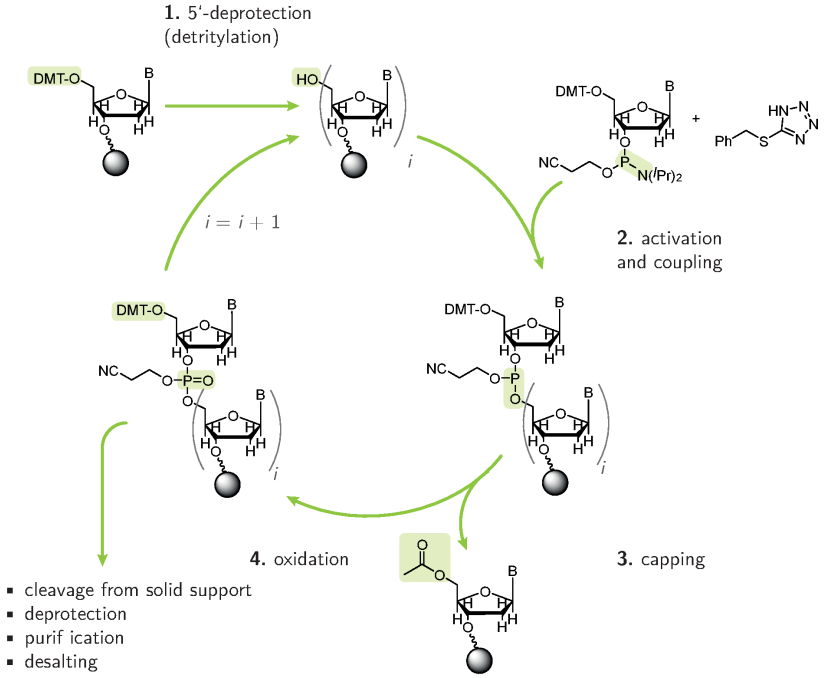


Figure 6.1.: Schematic illustration of the phosphoramidite based oligonucleotide synthesis cycle.

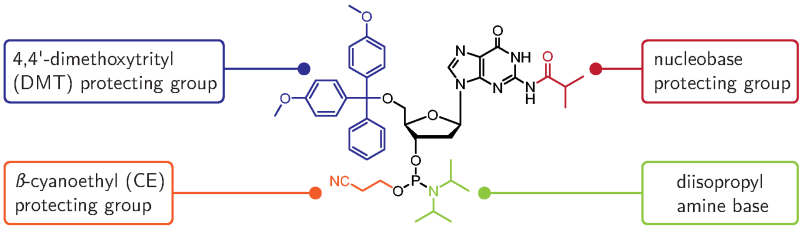


Figure 6.2.: Example for the deoxyribonucleotide phosphoramidite used in standard solid-phase DNA synthesis (here DMT-dG(*i*Bu)-CEP).

acid labile 4,4′-dimethoxytrityl (DMT) group.[228,229] In the first step (1: 5′-deprotection, also called detritylation), this trityl functionality is removed under acidic conditions, the released orange coloured trityl cation can be quantified by spectrophotometric means, thus giving an indication of the coupling efficiency. The DMT group is removed already under mild acidic conditions, minimizing the risk of nucleotide depurination.

The next step (2: activation and coupling) is the actual coupling step with the next nucleoside phosphoramidite. Additional to the before described protecting scheme, the phosphorus is attached to the nucleoside via said phosphoramidite, while the trivalent phosphorus itself is protected with a β-cyanoethyl group and additionally carries a diisopropylamine functionality (see figure 6.2).[207,208,232] The phosphoramidite is mixed with a tetrazole , e. g. 1H-tetrazole (in this work benzylthio-1H-tetrazole, BTT, is used), which acts both as a protonating agent and a nucleophile, substituting the diisopropyl group and thereby activating the phosphoramidite for the nucleophilic attack of the free hydroxyl functionality of the resin bound nucleotide.[233] The phosphoramidite is coupled to the nucleotide via a nucleophilic substitution reaction, with typical coupling efficiencies of ≥ 99 %.

To avoid accumulation of $(n-1)$ failure strands, unreacted nucleotides are deactivated by acetylation of the remaining free hydroxyl groups (3: capping).

As the trivalent phosphorus is still reactive towards nucleophilic substitution, it is oxidized to the unreactive pentavalent state with an oxidant (iodine) in the last step of the cycle (4: oxidation). At this point the cycle enters its starting state, and the whole cycle is repeated for the required amount of steps.

Once the full length product is obtained, it is removed from the solid support by cleavage of the covalent linker (basic medium), releasing the oligonucleotide with a free 3′-OH group. The base treatment also removes the protecting groups on the nucleobases, as well as on the phosphorus. At this point, the byproducts and also failure strands are removed in a purification step, typically comprising reversed-phase high

pressure liquid chromatography (RP-HPLC) and removal of non-nucleic-acid salts by reversed phase solid phase extraction (RP-SPE).

6.1. Tetramolecular G-quadruplexes of the type $[L^1d(G_n)]_4$ (n=3–5)

In this section the results for tetramolecular G-quadruplexes assembling from strands $\mathbf{L^1}d(G_n)$ (n=3–5) will be presented.

As already stated in the previous chapter 5, the first goal of this thesis is a proof-of-concept experiment, which demonstrates that a metal-base tetrad can indeed be incorporated into a G-quadruplex structure, analogous to the metal base-pairs in duplex DNA (chapter 4.2). To this means, ligand $\mathbf{L^1}$ was covalently attached at the 5′-end to a DNA strand consisting of four guanine nucleotides ($\mathbf{L^1}d(G_4)$).

DNA synthesis followed the protocol as described above (see also the experimental section 9.1). As in the synthesis of the phosphoramidites only a very simple work-up and purification procedure was employed, coupling efficiencies varied. Nevertheless, excellent coupling values similar to those of the standard nucleoside phosphoramidites (∼99 %) were regularly achieved. After synthesis, the oligonucleotide bound to the solid support was treated with aqueous ammonia over night at 55 °C for cleavage and deprotection. Solid phase extraction then yielded the clean oligonucleotide. The semi-preparative RP-HPLC purification step was omitted in this case, due to difficulties arising from the short oligonucleotide length and the small difference in polarity between the full length strand and the (n–1) failure strand. Only those samples were used, where ESI(–) MS revealed the product to be of sufficient purity.

Figure 6.3a shows the ESI-MS spectrum of a sample with $[\mathbf{L^1}d(G_4)]$. Not only one signal is detected but several, corresponding to the different charge states of the oligonucleotide, when the phosphate groups are partially deprotonated, i. e. $[M–2H]^{2–}$, $[M–3H]^{3–}$, $[M–4H]^{4–}$ (where M stands for the oligonucleotide with fully protonated phosphate groups). Deconvolution of this spectrum gives the neutral spectrum, which only shows the fully protonated species, plus those strands which transiently

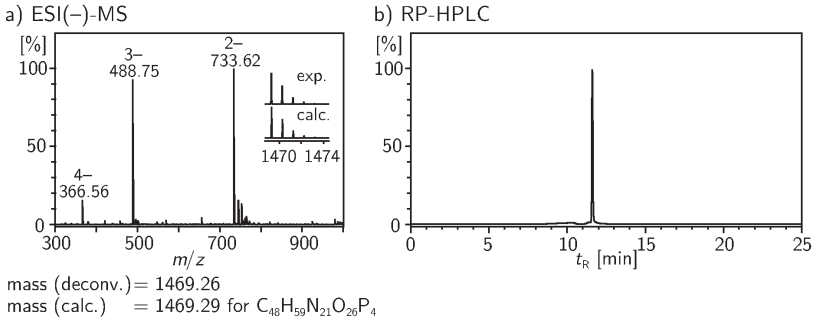


Figure 6.3.: Characterization of the oligonucleotide strand [$\mathbf{L^1}$d(G_4)]. a) the ESI(–) MS spectrum measured in H_2O/MeCN 60:40 (v/v) 0.02 % NH_3 and b) analytical RP-HPLC trace measured at 260 nm.

carry Na^+ or K^+ as the counterion. To complement the correct mass elucidation, analytical RP-HPLC was conducted, the corresponding trace depicted in figure 6.3b. The HPLC analysis clearly shows that the sample contains one major species, while integration of the signal indicates the oligonucleotide to be of >90 % purity.

Following the synthesis and characterization, the concentration of the DNA sample has to be determined, which requires knowledge about the UV extinction coefficient ϵ_{260}. Several methods are routinely used throughout the literature, the most common ones being the additive approach, which uses the sum of monomer extinction coefficients, and the nearest neighbour method, making use of mononucleotide extinction coefficients subtracted from dinucleotide values. Both methods, in conjunction with the commonly attributed values they make use of, give extinction coefficient values which are on average 14–16 % too large. However, revised values for the additive method based on high-resolution ^{1}H-NMR measurements are available.[234] In order to take into account the presence of the covalently attached modification (ligand $\mathbf{L^1}$), the extinction coefficient of the ligand was determined. It is assumed that the phosphate does not contribute to the UV-VIS spectrum, so that the free ligand (**3**) can be used for the determination. Measuring the

absorbance values at 260 nm for a series of different concentrations of **3** and applying a linear regression fit then enables the extinction coefficient determination according to the Beer-Lambert law (with absorbance A, extinction coefficient ϵ, concentration c, and optical path length d).

$$A = \epsilon \times c \times d \tag{6.1}$$

Once the concentration of the oligonucleotide stock solution is determined by measuring the absorbance of a DNA sample at 260 nm and calculation according to equation 6.1, the actual G-quadruplex samples can be prepared and analysed. G-quadruplex formation requires low temperatures and high monovalent cation concentration (as described in chapter 4.1), however even at 4 °C formation is very slow, with ∼4 d at 4 °C necessary to reach equilibrium at a 7 µM DNA concentration for the investigated strand. To accelerate the annealing process, DNA samples were first heated to 85 °C, slowly cooled to room temperature and then subjected to a freeze-thaw cycle as described in the literature.[235] Sodium or lithium cacodylate buffer (NaCaco or LiCaco, pH 7.2–7.3, 10 mmol L^{-1}) and sodium chloride (or sodium perchlorate) at 100 mmol L^{-1} concentration were used. The cacodylate buffer, albeit toxic, has the advantage of a low pK_a temperature-dependency and low transition metal binding affinity, which is not the case for other common buffers like Tris buffer (tris(hydroxymethyl)aminomethane adjusted with e. g. HCl) or phosphate buffer[236] (see also the rules for good biological buffers as proposed by Good et al.[237]).

The so prepared sample was tested for the presence of G-quadruplex formation by thermal denaturation, thermal difference spectra (see chapter 4.3.1), and CD (chapter 4.3.2). As depicted in figure 6.4 (only black curves), all three methods prove the formation of a G-quadruplex. The observed thermal difference profile is in accordance with that typical of G-quadruplexes, showing a distinct minimum at around 295 nm.[170] Following the temperature dependant absorption at this wavelength, then gives rise to the sigmoidal denaturation profile depicted in figure 6.4a. Additional proof comes from the corresponding CD spectrum, which is indicative of stacked G-quartet formation, i. e. the G-quadruplex

structure. Furthermore, the CD spectrum additionally shows that the topology is most likely parallel, all strands with the same 5′→3′ directionality, as anticipated.[173]. As stated before, the intense positive CD band at around 260 nm with a negative one at around 240 nm can be attributed to a purely homopolar coupling between the four G-tetrads, which is typical for parallel, tetramolecular G-quadruplexes. The parallel arrangement of the four oligonucleotide strands with respect to each other is a prerequisite for the successful coordination of the four pyridine donor functionalities to a transition metal ion.

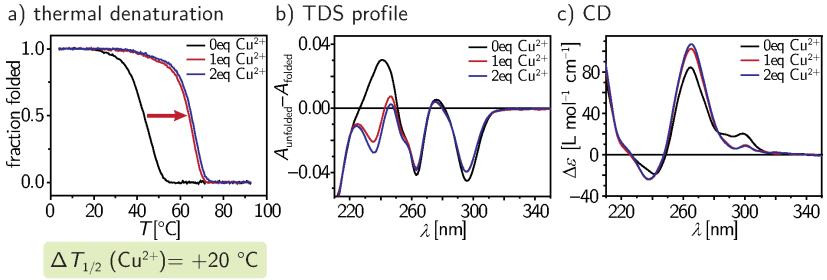


Figure 6.4.: Spectroscopic results for the G-quadruplex $[\mathbf{L^1}d(G_4)]_4$. Spectroscopic results for the G-quadruplex $[\mathbf{L^1}d(G_4)]_4$ upon addition of $CuSO_4$. a) Thermal denaturation curve (monitored at 295 nm) converted to the fraction folded values; b) corresponding thermal difference spectra ($\Delta A = A_{\text{unfolded}} - A_{\text{folded}}$); c) CD spectra at 4 °C. 10 mM LiCaco pH 7.3, 100 mM NaCl, 7.5 µM DNA, 0, 1.875, or 3.75 µM $CuSO_4$.

Next, the effect of different transition metal ions on the G-quadruplex thermal stability was tested, whereby only metal cations were chosen which are known to form square-planar, square-pyramidal, or octahedral (square-bipyramidal) coordination environments (like Fe^{2+}, Fe^{3+}, Co^{2+}, Ni^{2+}, Cu^{2+}, Zn^{2+}, Pd^{2+}, Au^{3+}).[12] Among this series, only Cu^{2+} and Ni^{2+} showed an effect, both on the melting temperature of the G-quadruplex and the CD spectrum (see figures 6.4 and 6.5). Both metal cations lead to a remarkable increase in denaturation temperature $\Delta T_{1/2} = +9$ °C (Ni^{2+}), $\Delta T_{1/2} = +20$ °C (Cu^{2+}), when one equivalent per G-quadruplex (0.25 equiv. per single strand) is added to the DNA

solution, independent of the addition being before or after the annealing process. Even in the presence of 4 equiv. of the metal cation, the same stabilisation effect is detected. This observation for Cu^{2+} is indeed remarkable, as a destabilizing effect for excess Cu^{2+} ($\sim$0.7 equiv./nt) ions on an unmodified G-quadruplex was reported in the literature.[238–241] Interestingly, in the case of Ni^{2+}, further metal addition leads to an additional increase in denaturation temperature (as compared to the first equiv.), whereas for Cu^{2+} the effect is negligible (even up to four equiv., data not shown). When comparing the two metal ions, Cu^{2+} seems to bind much stronger to the pyridine donors than compared to Ni^{2+}, in accordance with the predictions made in the Irving-Williams series.[12,242]

Concerning the CD spectra, metal addition leads to a rise in intensity of the 260 nm band, whereas the band at 240 nm decreases, in agreement with the increase in stability observed in the thermal denaturation studies. The CD analysis reveals an additional feature, namely a small positive band at around 295 nm, which is attributed to a heteropolar G-quartet coupling. When Cu^{2+} or Ni^{2+} is added, this band immediately looses intensity, and disappears entirely at excess of the metal salt. This phenomenon can be explained fourfold: first, a small portion of the G-quadruplexes could adopt not a purely parallel topology, but one with one or two strands antiparallel to each other, second, one G-quartet could adopt an all-*syn* conformation, third, the parallel G-quadruplexes partially stack on top of each other in an antiparallel fashion (directionality $3'\leftarrow5'|5'\rightarrow3'$), fourth, the pyridine units take part in π–π stacking to the top G-quartet, thereby altering the CD electronic transitions. The first and the second explanation are rather unlikely, as the antiparallel topology is not observed for tetramolecular G-quadruplexes and the change in strand orientation or *syn*/*anti* flip (strand deattachment, flipping directionality or conformation from *anti* to *syn*, reassociation) can be expected to be slow, in contrast to the fast process observed here. The other two explanations are both possible, the 5′–5′ stacking is, however, more likely, as this is commonly observed for tetramolecular G-quadruplexes.[33]

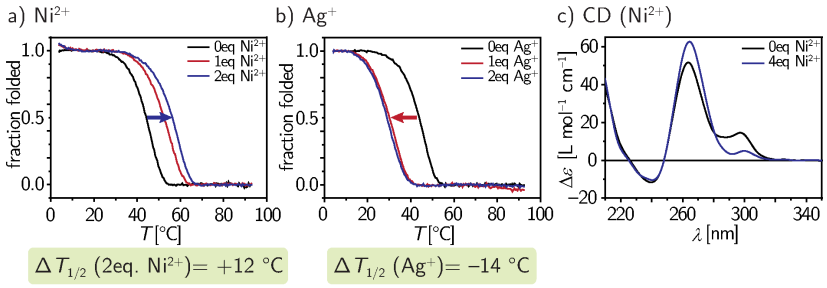


Figure 6.5.: Spectroscopic results for the G-quadruplex $[\mathbf{L^1}d(G_4)]_4$ upon addition of $NiSO_4$ and $AgClO_4$. Thermal denaturation curve (monitored at 295 nm) converted to the fraction folded values for a) $NiSO_4$ and b) $AgClO_4$; c) CD spectra at 4 °C for $NiSO_4$. 10 mM LiCaco pH 7.3, 100 mM NaCl ($NiSO_4$) or $NaClO_4$ ($AgClO_4$), 7.5 µM DNA, 0, 1.875, or 3.75 µM metal salt.

For comparison, the effect of Ag^+ ions on the G-quadruplex stability was also tested, showing that one equiv. $AgClO_4$ addition per G-quadruplex leads to a decrease in denaturing temperature with $\Delta T_{1/2}$ = –14 °C (see figure 6.5b), in agreement with the identical observation made for unmodified G-quadruplexes.[243] Most likely, the interaction results from the Ag^+ ions coordinating to guanine N7, thereby disrupting the G-quartet hydrogen bonding pattern.[201] Ag^+ ions usually prefer a linear coordination environment, thus rendering the $(pyridine)_4$ motif an unsuitable chelating structure. In contrast, both Cu^{2+} and Ni^{2+} should be easily coordinated by the four pyridines in a square-planar, or axially distorted octahedral (with two additional water ligands) fashion.

To estimate the binding stoichiometry for the Cu^{2+} ions, a titration experiment was conducted, in which increasing amounts of $CuSO_4$ were added to the preformed G-quadruplex, while measuring the absorbance changes at 236 nm. The wavelength was chosen as the one which showed the most pronounced absorption change, probably corresponding to a π–π^* transition of the pyridine units. From this experiment a 1:1 $Cu^{2+}/[\mathbf{L^1}d(G_4)]_4$ ratio can be estimated, in accordance with the anticipated binding mode (see figure 6.7).

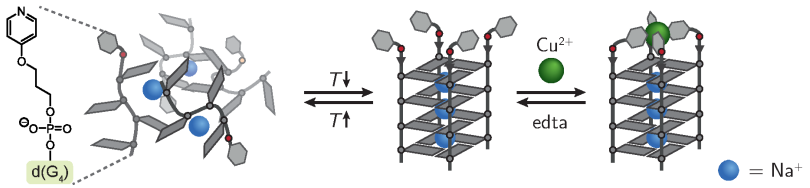


Figure 6.6.: Schematic assembly of the G-quadruplex $[\mathbf{L^1}\mathrm{d(G_4)}]_4$ from single strands at low temperatures, incorporation of Cu^{2+} and removal with edta.

The pronounced increase in stability can be fully reversed by removing the Cu^{2+} with a chelating agent like edta (Na_2H_2EDTA), which regenerates the denaturation profile of the Cu^{2+} unbound G-quadruplex (see the appendix). By recurrently adding both reagents, with an intermediate annealing process respectively, this stabilization and destabilization cycle can be repeated several times, demonstrating the robustness of the metal G-quadruplex assembly presented here (see figure 6.7b).

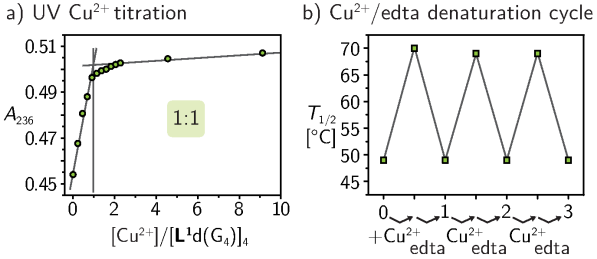


Figure 6.7.: UV-VIS absorbance titration measured at 236 nm upon addition of Cu^{2+} ions to $[\mathbf{L^1}\mathrm{d(G_4)}]_4$ at 4 °C. The vertical line denotes the approximate binding stoichiometry of 1:1 $Cu^{2+}/[\mathbf{L^1}\mathrm{d(G_4)}]_4$; b) Denaturation temperature cycle for the recurring stabilization and destabilization of the G-quadruplex $[\mathbf{L^1}\mathrm{d(G_4)}]_4$ by the alternate addition of Cu^{2+} or edta (7 µM $\mathbf{L^1}\mathrm{d(G_4)}$, 10 mM sodium cacodylate, 100 mM NaCl, $CuSO_4$ and edta additions in 1 equiv. steps). Reprinted with permission from D. M. Engelhard et al. [244] Copyright ©2013 John Wiley & Sons, Inc.

Up until here it was assumed that the increase in denaturation temperature $T_{1/2}$ is also reflected by an increase in thermodynamic stability. As discussed in the introductory chapter 4.3.1, this does not necessarily have to be the case. In order to address this predicament, the G-quadruplex association process was followed with UV-VIS spectroscopy, both in the presence and absence of Cu^{2+}. As shown in the appendix, association occurs on a very long time-scale compared to the denaturation, i. e. a time-scale of several days at 4 °C, but is significantly accelerated in the presence of Cu^{2+}. The thermodynamic stability constant K is defined as the quotient between k_{on} and k_{off} (see also equation 4.3), so both findings, higher denaturation temperature (so increase in k_{off}) and higher k_{on} work in the same direction and lead to a increase in K. Strictly speaking, this observation is only qualitative, as the two rates are determined at completely different temperatures (at $T_{1/2}$ = 44 °C and at 4 °C, respectively).

To shed light on the kind of interaction between the metal ions and the G-quadruplex, a reference strand $\mathbf{L^{1*}}d(G_4)$ was synthesized, in which the pyridine unit is exchanged for a phenyl motif, which should prevent metal ion coordination. Indeed, as evidenced by the thermal denaturation profiles, no increase in stability of the control G-quadruplex $[\mathbf{L^{1*}}d(G_4)]_4$ is observed for the addition of the three metal ions, respectively (see figure 6.8). In fact, addition of Ag^+ ions to the G-quadruplex led to the same decrease in denaturing temperature ($\Delta T_{1/2}$=–14 °C) as observed for the pyridine containing G-quadruplex, supporting the binding of the Ag^+ ions to the G-quadruplex part, like the guanine moieties. It should be noted that the thermal stability of the phenyl substituted G-quadruplex is higher than for its pyridine counterpart in absence of a transition metal ion. Assuming a planar stacking of the aromatic unit of the ligands onto the topmost (5′) guanine, this phenomenon can be explained by a more favourable π–π stacking between the more electron rich and more hydrophobic phenyl and guanine as compared to the stacking between pyridine and guanine.

Based on these promising results, the pyridine-modified G-quadruplexes with other G-tract lengths were investigated. Extending the

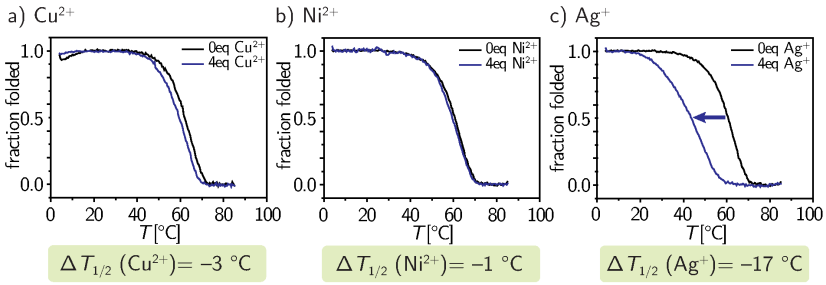


Figure 6.8.: Thermal denaturation curves (monitored at 295 nm) converted to the fraction folded values for the G-quadruplex $[\mathbf{L^{1*}}d(G_4)]_4$ upon addition of a) $CuSO_4$, b) $NiBF_4$, and c) $AgBF_4$. 10 mM NaCaco pH 7.2, 100 mM NaCl (Cu^{2+}) or $NaClO_4$ (Ni^{2+} and Ag^+), 7.4 µM DNA, 0 or 7.1 µM metal salt.

G-quadruplex core by one G-quartet obtaining $[\mathbf{L^1}d(G_5)]_4$ leads to an increase in stability by $\Delta T_{1/2} = +40$ °C in comparison to $[\mathbf{L^1}d(G_4)]_4$ as expected.[173] Analogues to the shorter G-quadruplex, addition of Cu^{2+} results in an, however small, increase in stability $\Delta T_{1/2} = +3$ °C. One explanation for this minor effect of the Cu^{2+} ion could be that the G-quadruplex core is already very rigid and stable and therefore more sensitive to conformational changes like the coordination of the ligand pyridines to the Cu^{2+} ion. In other words, the increase in stability due to the metal coordination may be partially compensated by conformational strain imposed on the G-quadruplex stem. Looking at the denaturation profile, the beginning of G-quadruplex renaturation is observed at low temperatures (<20 °C), which is not observed for the shorter G-quadruplex. This highlights again the enormous hysteresis effect occurring in tetramolecular G-quadruplexes. It also nicely supports the thesis that the copper metal ion does indeed cause an increase in thermodynamic stability, because beginning renaturation is observed at higher temperatures for the Cu^{2+} bound G-quadruplex, whereas it occurs at lower temperatures in the transition metal free G-quadruplex (see the appendix). Exchanging the Na^+ electrolyte for K^+ would even increase the thermodynamic stability of the whole assembly,

so instead, a metal ion with lower tendency to promote G-quadruplex formation, namely Cs^+, was tested. Although mostly thought to inhibit G-quadruplex formation, here Cs^+ ions do indeed lead to the characteristic UV hypochromic shift at 295 nm, typical of a tetramolecular G-quadruplex. Furthermore, formation is backed up by the corresponding CD spectrum, which depicts the characteristic features for a parallel G-quadruplex as described above. Cu^{2+} addition results in an increase in G-quadruplex denaturation temperature, but with a significantly higher magnitude as compared to the Na^+ case ($\Delta T_{1/2}$ = +11 °C). The lower overall stability and increased flexibility of the G-quadruplex as caused by the Cs^+ ions (which is also reflected in a much broader melting transition profile as compared to the very sharp one in case of Na^+) is influenced by the metal coordination to a greater extent.

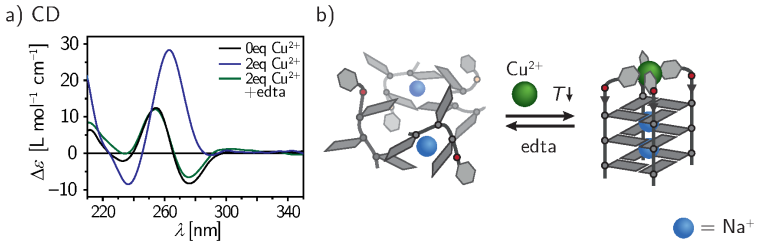


Figure 6.9.: a) CD spectra and b) corresponding scheme for Cu^{2+} induced formation of $[\mathbf{L^1}d(G_3)]_4$ at low temperatures and denaturation upon Cu^{2+} ion removal with edta.

Going to the sequence with only three consecutive guanines per strand, $[\mathbf{L^1}d(G_3)]_4$, no G-quadruplex formation is observed, even at lower temperatures. Neither do the UV (TDS or melting) spectral data indicate G-quadruplex formation, nor does the CD spectrum, which only shows the typical bands for single strands. This means, that the denaturation temperature is at least 44 °C lower than for $[\mathbf{L^1}d(G_4)]_4$. When the DNA sample is annealed in the presence of Cu^{2+} ions, however, G-quadruplex formation can be triggered, giving rise to the typical spectroscopic features as described for the $[\mathbf{L^1}d(G_4)]_4$ G-quadruplex. When the Cu^{2+} ions are removed by a chelating agent

like edta, immediate disruption of the G-quadruplex into single strands is observed (see figure 6.9).

All of the results presented above are based on optical spectroscopic techniques and need to be complemented by independent methods. A common biochemical technique to study the size and shape of biomolecules is (gel) electrophoresis. In this experiment the biomolecules migrate through a porous matrix under the influence of an electric field. Electrophoretic mobility is dependant on several factors, including the net charge, size, and shape of the respective molecule; pore size of the matrix; pH, temperature, and ionic strength of the buffer system; and the electric field.[245] The porous matrix is used to minimize convection and diffusion of the biomolecules and is typically an agarose or polyacrylamide gel for DNA separations. The polyacrylamide gels are made from acrylamide (2–30 % w/v) and a linking agent like N,N'-methylene bisacrylamide. In gels with a high acrylamide concentration molecule migration is dominated by molecular size, with only minor contributions from mass and charge. As the DNA is negatively charged, it will migrate into the direction of the anode (positive pole).

G-quadruplex samples $[\mathbf{L}\mathrm{d}(\mathrm{G_n})]_4$ ($\mathbf{L} = \mathbf{L^1}$ or $\mathbf{L^{1*}}$) were investigated with polyacrylamide gel electrophoresis (PAGE), to elucidate G-quadruplex formation and the presence of higher-order structures like G-wires (see the experimental section 9.1.8 for details). The gels were visualized with three independent methods, namely staining with protoporphyrin IX, (PPIX)[246,247] staining with silver-impregnation,[248,249] and UV-shadowing.[250] UV-shadowing is a non-invasive and fast technique, in which the gel is laid on top a fluorescent-dye impregnated silica plate (common plates used for thin layer chromatography) and then irradiated at 254 nm with a standard UV-lamp. The DNA absorbs at this wavelength, thus the fluorescent dye of the TLC plate is not excited. As a consequence, the DNA will appear as dark spots on a bright background. This method is unfortunately rather insensitive, as it requires quite large amounts of DNA, so an additional silver staining was performed. In contrast, silver staining is very sensitive, but requires a more elaborate preparation procedure and is more prone to experi-

Table 6.1.: Overview of denaturation temperatures $T_{1/2}$ for the G-quadruplexes $[\mathbf{L}d(G_n)]_4$ with $\mathbf{L} = \mathbf{L^1}$ or $\mathbf{L^{1*}}$, and n = 3–5.

G-quadruplex	**salt**	**metal**	**equiv.**	$T_{1/2}$ [°C]	$\Delta T_{1/2}$ [°C]
$[\mathbf{L^1}d(G_3)]_4$	NaCl†	$CuSO_4$	0	<4	
			4	15.5	>+11.5
$[\mathbf{L^{1*}}d(G_3)]_4$	NaCl†	$CuSO_4$	0	<4	
			4	<4	–
$[\mathbf{L^1}d(G_4)]_4$	NaCl	$CuSO_4$	0	43.4	
			1	64.0	+20.6
			2	65.2	+21.8
		$NiSO_4$	0	44.2	
			1	52.4	8.2
			2	56.3	12.1
	$NaClO_4$	$AgClO_4$	0	43.9	
			1	30.1	–13.8
			2	29.0	–14.9
$[\mathbf{L^{1*}}d(G_4)]_4$	NaCl†	$CuSO_4$	0	61.9	
			4	59.4	–2.5
		$NiBF_4$	0	60.6	
			4	59.3	–1.3
	$NaClO_4$			60.6	
		$AgBF_4$	4	43.6	–17.0
$[\mathbf{L^1}d(G_5)]_4$	NaCl	$CuSO_4$	0	83.9	
			1	87.4	+3.5
			2	88.1	+4.2
	CsCl	$CuSO_4$	0	53.5	
			1	64.7	+11.2
			2	66.1	+12.6

†Measured in sodium cacodylate buffer pH 7.2.

mental errors and impurities. Here, the basic principle lies in treating the gel with a Ag^+ containing solution and after washing, a reducing agent is added. Ag^0 will start to form, predominantly from locations of high concentration, i. e. at the DNA molecules due to the electrostatic interactions of the positively charged Ag^+ with the negatively charged DNA. The reaction is stopped when an optimal signal-to-noise ratio is obtained, that is, when the DNA bands appear clearly as black spots on a faintly coloured background. The techniques greatly relies on the purity of all reagents and diligence during the experiment to avoid contamination or artefacts. The PPIX staining relies on the high affinity of (planar) porphyrins like protoporphyrin IX towards parallel G-quadruplexes, by stacking to the 5′ or 3′ G-quartet.[246,247] The PPIX serves as a dual probe, both for the G-quadruplex formation and also for the parallel topology. Upon irradiation of the PPIX stained gels with UV light (254 nm) PPIX bound to the parallel G-quadruplex exhibits strong fluorescence.

As depicted in figure 6.10, the obtained gel nicely shows the formation of discrete G-quadruplexes, based on the slow migrating bands (lane 4) compared to the fast moving band of the single strand (lane 3). The band of the G-quadruplex sample moves approximately equally fast compared to the 15 nt ssDNA control (lane 1), which is in agreement with the respective nucleobase count (16x dG plus 4x ligand $\mathbf{L^1}$). Of course the electrophoretic mobility is not directly comparable between the (presumably) random coil ssDNA and the very compact G-quadruplex structures. Furthermore, no difference is observed in comparison with the reference G-quadruplex $[\mathbf{L^{1*}}\mathrm{d(G_n)}]_4$, demonstrating that the ligand does not have a great influence on the electrophoretic mobility. The silver staining reveals the partial formation of higher order structures, although these very slow migrating bands are not visible in the case of UV-shadowing, suggesting that only a small fraction of the oligonucleotides tend to aggregate. The PPIX staining clearly shows fluorescence in lanes 4–7 and 8, in accordance with the formation of parallel G-quadruplexes, whereas no fluorescence is detected in the single- and double-stranded control samples (lane 1 and 2). The fluorescence in

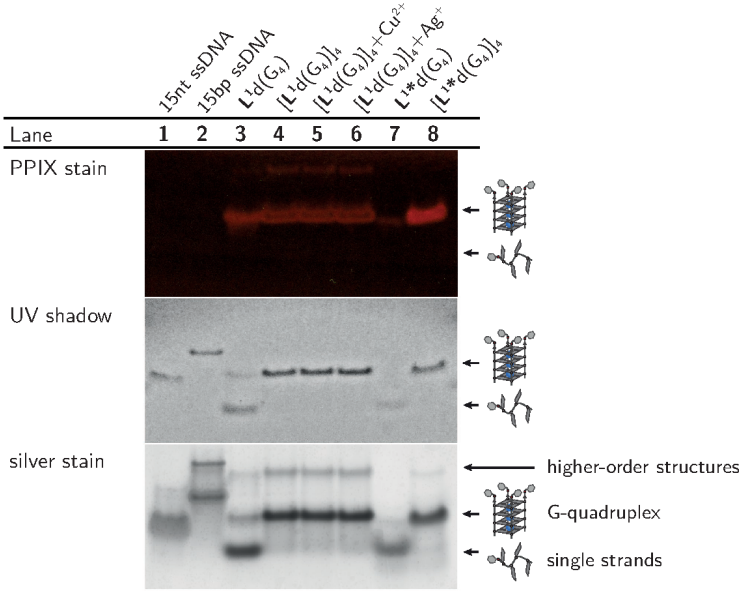


Figure 6.10.: Non-denaturing gel electrophoresis for $\mathbf{L^1}d(G_4)$ and control strand $\mathbf{L^{1*}}d(G_4)$. G-quadruplex bands were first visualized by staining the gel with PPIX (upper picture), by UV-shadowing (middle picture) and then with silver staining. The arrows denote the same migration length in the gels. Reprinted with permission from ref.[244] Copyright ©2013 John Wiley & Sons, Inc.

lane 3 (and also 4) can be attributed to partial G-quadruplex formation under the native gel electrophoresis conditions. No effect of the added metal ion is observed on the migration length of the G-quadruplex (lane 5 and 6), either suggesting that the bound metal ion does not change the overall size of the G-quadruplex (implying that no effect is observed due to the change in charge), or, far more likely, that the metal ion is not present any longer in the G-quadruplex bands. The latter could be explained by the applied electric field, which impedes movement of the cationic metal ions towards the equally positive pole.

Concerning the full characterization of this new class of G-quadruplexes, the last piece of the puzzle is the elucidation of the copper coordination environment. With the control experiments on $[\mathbf{L^{1*}}\mathrm{d(G}_n\mathrm{)}]_4$, the copper–pyridine interaction for $[\mathbf{L^{1}}\mathrm{d(G}_n\mathrm{)}]_4$ could be established, however, up to this point the geometry around the Cu^{2+} centre remained unclear. To address this question, electron paramagnetic resonance spectroscopy (EPR) may be utilized.

EPR (or ESR for electron spin resonance) is a spectroscopic technique, similar to NMR, in which electron magnetic dipole moment transitions are detected in the presence of an external magnetic field. It is limited to paramagnetic substances, that is compounds with unpaired electrons, e. g. radicals, open-shell transition metal ions, or compounds in triplet states. In the strong, homogeneous magnetic field of the EPR spectrometer, the degenerate energy levels of the electron magnetic-dipole moment, with spin $S = \frac{1}{2}$, split into two states (Zeeman interaction, see figure 6.11).[251,252] The energy gap ΔE between these two states is then defined as:

$$\Delta E = g_\mathrm{e} \times \mu_\mathrm{B} \times B_0 \tag{6.2}$$

with g_e the Landé or Zeeman g-factor, which is $g_\mathrm{e} = 2.0023$ for a free electron, μ_B the Bohr magneton, and B_0 the strength of the external magnetic field. When the sample is irradiated at microwave frequency and the radiation energy is in resonance with the magnetic dipole transition of the free electron, excitation from the $M_S = -\frac{1}{2}$ state to the $M_S = +\frac{1}{2}$ state occurs and as a result an absorptive signal is

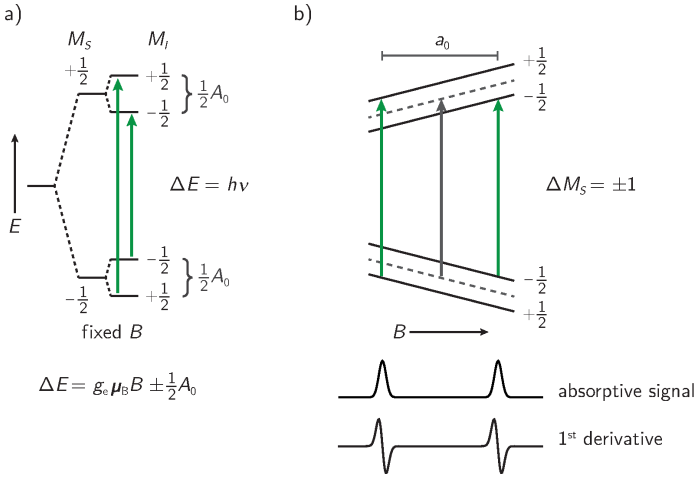


Figure 6.11.: Schematic energy levels and EPR transitions for a system consisting of one unpaired electron and one nucleus ($I = \frac{1}{2}$) under the influence of an external magnetic field. a) Zeeman splitting ($\Delta M_S = \pm 1$) due to the external magnetic field and hyperfine interaction ($\Delta M_I = \pm 1$) by coupling to the nuclear spin magnetic moment. b) EPR transitions and corresponding absorptive and first derivative signals observed due to Zeeman splitting and hyperfine interaction when B of the external field is varied. Reprinted (redrawn) with permission from ref.[251] Copyright ©2006 John Wiley & Sons, Inc.

detected. Often, instead of the absorptive signal, the first derivative is depicted. Due to contributions of local magnetic fields, the observed g-factor of a paramagnetic compound deviates from the value of the free electron g_e. As magnetic nuclei also have an intrinsic magnetic dipole moment, dipole-dipole interactions between the nuclear spins and the electron spin is observed, giving rise to a further energy state splitting, called the hyperfine splitting (A_0 at constant magnetic field; a_0 as a magnetic field difference, see figure 6.11). The absorptive signal consequently splits into two signals (for $I = \frac{1}{2}$) with half intensity, as now two energetically unequal transitions, detected at different magnetic field values, are allowed. The number of lines which are observed is given by $(2NI + 1)$, with N the number of interacting nuclei and I the nuclear spin quantum number. Additional complexity arises from an orientation dependency relative to the external field, i. e. anisotropy, of the g-factor and the hyperfine splitting, which provides information about the local geometrical environment of the unpaired electron.[251]

These considerations imply that for a paramagnetic transition metal ion, EPR spectroscopy can provide information about the spin state of the electron and the orbital in which it is located, and the local coordination environment and geometry of the transition metal ion; rendering it a perfect tool for the task set above. Consequently, EPR measurements were conducted on the $Cu^{2+}[\mathbf{L}d(G_4)]_4$ G-quadruplex in collaboration with Roberta Pievo and Marina Bennati (Max-Plank-Institute for Biophysical Chemistry, Göttingen, Germany), within the framework of the DFG IRTG 1422.

The X-band CW-EPR spectrum of a frozen solution of Cu^{2+}-$[\mathbf{L}d(G_4)]_4$ shows a signal pattern (at around 300 mT) characteristic of a mononuclear Cu^{2+} species. No half-field signal (160 mT) is detected over the measured temperature range of 10–70 K, suggesting that the sample does not contain species of higher nuclearity, which would give rise to transitions with $\Delta M = \pm 2$. The anisotropic spectrum can be dissected into several contributions, namely hyperfine coupling of the electron spin to the Cu^{2+} nucleus ($I = \frac{3}{2}$) and additional (super-) hyperfine coupling to nitrogen atoms ($I = 1$). All of these parameters could be

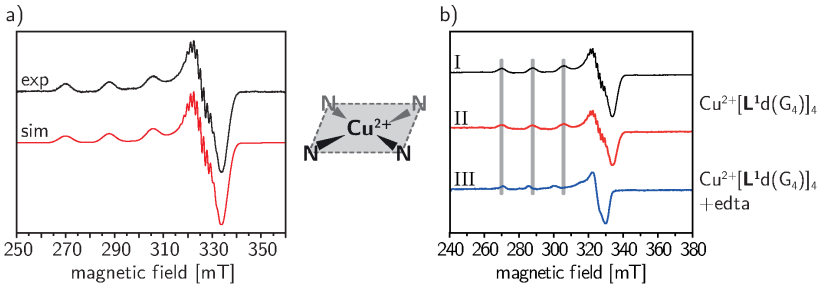


Figure 6.12.: a) X-band CW-EPR spectrum of a frozen solution of $Cu^{2+}[\mathbf{L^1}d(G_4)]_4$ and simulation (shown in red). b) Comparison between spectra with Cu^{2+} addition (I) after or (II) prior annealing, and (III) after edta (Na_2H_2EDTA) addition. Experimental parameters: 9.4 GHz, T=50 K, microwave power: 0.6325 mW, conversion time: 123 ms, modulation amplitude: 10 G, modulation frequency: 100 kHz, 1 h signal averaging. The spectrum was simulated using Matlab-based toolbox Easyspin (version 4.51).[34] Simulation parameters (for (a)) are: $g_{x,y,z}$ = 2.0470, 2.0687, 2.2635; A(Cu) = [47 45 552] MHz; A(N) = [40 39 40] MHz; line width = 5 G. b) Sodium cacodylate pH 7.1 10 mM, NaCl 100 mM, $\mathbf{L}d(G_4)$ 605–630 µM, $CuSO_4$ 145–165 µM, 159 µM 10 % (v/v) glycerol). Adapted with permission from ref. [244] Copyright ©2013 John Wiley & Sons, Inc.

nicely fitted to the experimental spectrum using slightly rhombic **g**- and $\mathbf{A}^{Cu}$-tensors. Figure 6.12a (red curve) shows the simulated spectrum for a CuN_4 species, in perfect agreement with the experimental spectrum (black line). It should be noted that in principle the spectrum could also be simulated with only three nitrogen atoms, as the hyperfine splitting to the ^{14}N atoms in the region 330–340 mT is only weakly-resolved. The simulation with four gave, however, still the better fit.[253] Evaluation of the three g-values suggests that the unpaired electron is located in the $d_{x^2-y^2}$ orbital in the ground state, in accordance with a square-planar, square-pyramidal, or distorted octahedral coordination environment for the Cu^{2+} ion.[244,254] The resolution of the obtained EPR spectrum does not allow to elucidate, if e. g. water molecules are weakly coordinated in the axial positions.

As shown in Figure 6.12b (I) and (II), there is no difference in the Cu^{2+} EPR spectra of the G-quadruplex samples where the copper salt was added after (directly before freezing the sample for the EPR experiment) or prior to G-quadruplex formation. No deviation in signal intensity (apart from the slightly different Cu^{2+} concentrations), number of hyperfine signals and splitting or g-values can be observed. This indicates first that the coordination of Cu^{2+} to the G-quadruplex is relatively fast ($<$ 1 min at room temperature), so major structural rearrangements during coordination seem unlikely. Secondly, the presence of copper ions does not seem to lead to alternative coordination behaviours when added during G-quadruplex annealing, which includes heating to 85 °C for 10 min. This is in accordance with a preorganized coordination environment for the copper ions, resulting from the parallel arrangement of the oligonucleotide strands within the G-quadruplex $[\mathbf{L^1}d(G_4)]_4$ and the pyridine tetrad motif. When the Cu^{2+} bound G-quadruplex sample is treated with an edta solution, a spectrum typical for the Cu^{2+}-edta complex[255] is obtained (figure 6.12b III), thereby confirming the reversibility of the metal incorporation as proposed by the UV-VIS and CD measurements.

All in all, this chapter demonstrated that pyridine-donor functionalized guanine-rich oligonucleotide strands $\mathbf{L}d(G_n)$ assemble into

tetramolecular G-quadruplexes and are capable of coordinating a Cu^{2+} ion in a square-planar fashion via the pyridine units, as evidenced by UV-VIS, CD, gel electrophoresis and EPR measurements. Based on the full characterization of this initial system, the concept implemented here can be investigated in more detail, which will be the focus of the next chapters.

6.2. Linker length variation in tetramolecular G-quadruplexes [Ld(G4)]4 (L = L1, L1b–d)

The ligand design for **L1** was based on an initial molecular modelling by attaching the pyridine via a propylene linker to the 5′-end of a known tetramolecular G-quadruplex structure. To test the hypothesis that the linker length chosen there represents the conformational and sterical most suited approach, these linker lengths were systematically changed. Both the attachment to the oligonucleotide part (phosphate plus $(dG)_4$) and the pyridine moiety are kept the same, only the propylene linker is shortened (ligand **L1b**) or elongated by one methylene unit (**L1c**). Additionally, a very short linker comprising only a methylene group directly attached to the pyridine was tested (**L1d**), on the hypothesis that this short linker should prevent transition metal ion coordination or results in a significantly lower stability of the G-quadruplex. A run of four guanine nucleotides was chosen, as for the corresponding G-quadruplex [**L1**$d(G_4)]_4$ the complete UV-VIS based denaturation profile including low and high temperature baselines is observed, both for the Cu^{2+} free and the Cu^{2+} bound case. The results for the UV-VIS detected thermal denaturation temperatures are summarized in table 6.2, whereas the individual melting curves can be found in the appendix.

For all G-quadruplexes investigated, the denaturation occurs in the chosen temperature window (T = 4–95 °C), while the shape of each melting curve is the same for all investigated systems, meaning that the slope is comparable. Thus, differences in denaturation temperature can be compared directly and should be in correlation with the thermodynamic stabilities, as discussed for ligand **L1**. The first striking difference be-

Table 6.2.: Overview of denaturation temperatures $T_{1/2}$ for the G-quadruplexes $[\mathbf{L}\mathrm{d}(\mathrm{G}_4)]_4$ with $\mathbf{L} = \mathbf{L^1}$, $\mathbf{L^{1b}}$, $\mathbf{L^{1c}}$, $\mathbf{L^{1d}}$. The linker length d_{linker} is given as the distance between $C4$ of the pyridine ring and O_{P}, the oxygen attached to the $3'$ phosphorus.

G-quadruplex	$d_{\mathbf{linker}}$ [Å]	**metal**	**equiv.**	$T_{1/2}$ [°C]	$\Delta T_{1/2}$ [°C]
$[\mathbf{L^{1d}}\mathrm{d}(\mathrm{G}_4)]_4$	2.4	$CuSO_4$	0	36.5	
			1	61.6	+25.1
			2	63.3	+26.8
$[\mathbf{L^{1b}}\mathrm{d}(\mathrm{G}_4)]_4$	4.8	$CuSO_4$	0	37.2	
			1	62.1	+24.9
			2	63.5	+26.3
$[\mathbf{L^{1}}\ \mathrm{d}(\mathrm{G}_4)]_4$	6.1	$CuSO_4$	0	43.4	
			1	64.0	+20.6
			2	65.2	+21.8
$[\mathbf{L^{1c}}\mathrm{d}(\mathrm{G}_4)]_4$	7.3	$CuSO_4$	0	53.5	
			1	57.4	+3.9
			2	57.4	+3.9

tween the melting curves is the denaturation temperature in the absence of a transition metal ion. Although in a similar range, they differ quite considerably: The G-quadruplex with the shortest linker ($\mathbf{L^{1d}}$) shows the lowest stability ($T_{1/2}$ = 36.5 °C), while that with the longest ($\mathbf{L^{1c}}$) exhibits a much higher denaturation temperature ($T_{1/2}$ = 53.5 °C), with the other linker lengths in between those two. This trend is not linear in relation to the absolute linker length, as seen by the results for $\mathbf{L^{1d}}$ and $\mathbf{L^{1b}}$, which show no significant difference in denaturation temperature. As already mentioned above, the effect of the ligand can be most likely attributed to attractive pyridine–guanine interactions, like π–π stacking. Assuming this effect, the longer linker length presumably provides the necessary conformational flexibility to correctly position the pyridine, thus enhancing the interaction. The pyridine moieties could then also stimulate the 5′–5′ stacking of two G-quadruplexes, as observed both in the solution and solid state[33] and also in conjunction with planar drugs bridging between two G-quadruplexes.[205]

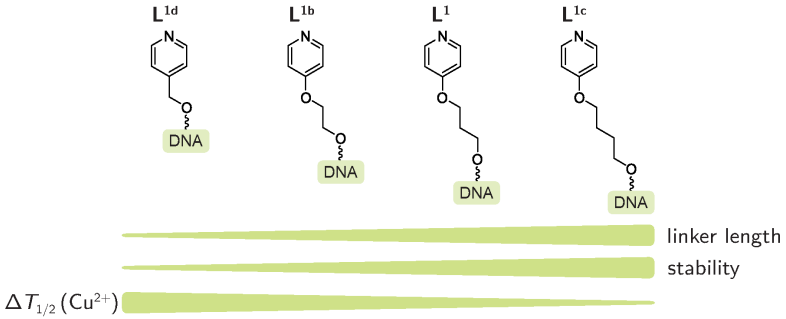


Figure 6.13.: Comparison of the linker lengths in $\mathbf{L^1}$, $\mathbf{L^{1b}}$, $\mathbf{L^{1c}}$, $\mathbf{L^{1d}}$.

This picture changes completely when Cu^{2+} is added to the respective G-quadruplexes $[\mathbf{L}d(G_4)]_4$ ($\mathbf{L}$ = $\mathbf{L^1}$, $\mathbf{L^{1b}}$, $\mathbf{L^{1c}}$, $\mathbf{L^{1d}}$). The remarkable increase in denaturation temperature for $Cu^{2+}[\mathbf{L^1}d(G_4)]_4$ is even surpassed in case of the shorter ligands $\mathbf{L^{1b}}$ and $\mathbf{L^{1d}}$ ($\Delta T_{1/2}$ = +25 °C in both cases). Contrarily, the longest ligand $\mathbf{L^{1c}}$ exhibits only a very moderate increase in stability ($\Delta T_{1/2}$ = +4 °C). When looking at the absolute values of the Cu^{2+} bound G-quadruplexes it is

quite obvious that the stability is almost the same for the three shorter ligands. This observation fits perfectly into the picture of the attractive pyridine–guanine interaction, which is interrupted by the coordination to the Cu^{2+} centre. Due to the steric hindrance of the pyridine-hydrogen atoms, the pyridine moieties adopt a propeller arrangement around the metal ion, giving up the co-planarity to the G-quartet. As a consequence, the stability is now governed by the G-quadruplex stem and the pyridine–metal coordination and only with minor effects from the linker, resulting in the observed similar denaturation temperatures. In case of the longest ligand $\mathbf{L^{1c}}$, the melting temperature of the copper-bound G-quadruplex is distinctly lower than in case of the other ligands. Either the coordination of the pyridine to the Cu^{2+} ion leads to steric hindrance and conformational unfavourable torsion angles of the alkyl chain, or to a loss of conformational freedom and therefore to a decrease in entropy, giving rise to the drop in thermodynamic stability.

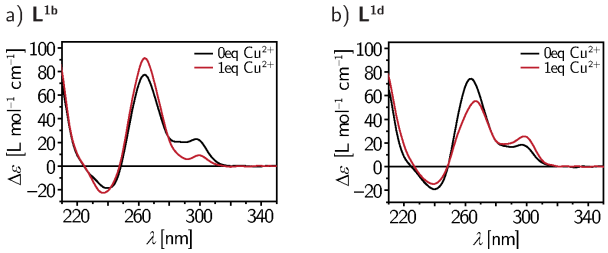


Figure 6.14.: Comparison of the CD spectra before and after Cu^{2+} addition for $[\mathbf{L}d(G_4)]_4$. a) $\mathbf{L} = \mathbf{L^{1b}}$, b) $\mathbf{L} = \mathbf{L^{1d}}$.

To complement these results, CD spectra were measured for all G-quadruplexes, both for the Cu^{2+} bound and unbound cases (see figure 6.14). For all ligands, the CD profile is similar to that of ligand $\mathbf{L^1}$, including the signal intensities, typical of a tetramolecular G-quadruplex with parallel strand polarities. This suggests that the variation in the linker length does not influence the G-quadruplex formation tendency of the investigated oligonucleotide strands. Furthermore, the pyridine units are equally preorganized at one end to accommodate the Cu^{2+} ion. Like in the case for $[\mathbf{L^1}d(G_4)]_4$, a minor band at around 295 nm is observed

for all linker lengths, which was earlier attributed to the 5′-5′ stacking of the G-quadruplexes in combination with pyridine stacking onto the top G-quartet. When Cu^{2+} is added, signal intensity of the positive 260 nm band increases, whereas the negative one at 240 nm also gets more pronounced, in conjunction with the observed increase in stability. Surprisingly, for the shortest ligand $\mathbf{L^{1d}}$ the opposite is true. Both the signal at 240 nm and 260 nm weaken upon Cu^{2+} addition, whereas the band at 295 nm increases in intensity, meaning that the heteropolar coupling between guanines of opposite glycosidic torsion angles becomes more pronounced. This is remarkable, as still the great increase in stability is observed in the UV-VIS based thermal denaturation studies. At this point it is necessary to investigate, if the short linker length actually allows for the coordination of the copper ion by the pyridines of the same G-quadruplex, or if the linker is in fact too short.

To get insights into this problem, the molecular modelling as done for ligand $\mathbf{L^1}$ was repeated for the other ligands. This time, only the top 5′ G-quartet was used as the anchoring platform for the ligand Cu^{2+} complex and was fixed in position during the geometry optimization, while the level of theory was slightly enhanced, using the semiempirical PM6 approach (as opposed to the previous molecular mechanics level). No further constraints were fixed neither to maintain the square-planar coordination geometry, nor the position of the Cu^{2+} ion on the central axis. Nevertheless, both were retained during the optimization steps, only with minor deviations from the perfect geometries. When comparing the two different levels of theory for the $[\mathbf{L^1}dG]_4$ case, it becomes immediately obvious that there are quite significant structural differences (see figure 6.15). Whereas in the MMFF optimized structure the ligand's propylene linker adopts a straight conformation, it is much more bend in the PM6 case. Interestingly, the Cu^{2+}–Na^+ separation remains quite the same, with 5.8 Å (MMFF) for the initial model and 6.1 Å (PM6) for the newly optimized structure. It should be noted that depending on the initial starting geometry for the optimization, in some cases the Cu^{2+}-pyridine plane moves away from the central horizontal axis (axis perpendicular to the G-quartets) and rotates in relation to the G-quartet

plane, resulting in a structure with both elongated and bend linkers. This suggests that even in the coordinated state, the ligands exhibit a conformational flexibility. These observations are retained with the other ligands, while the bending of the linker is the most pronounced in case of the longest ligand $\mathbf{L^{1c}}$ (see figure 6.15c).

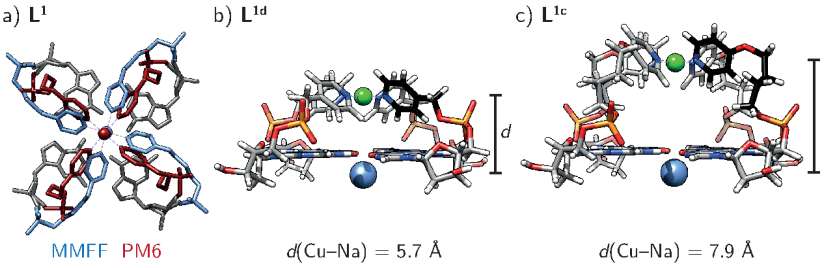


Figure 6.15.: Comparison of the molecular models consisting of a fixed G-quartet (PDB 3TVB) and 5′ attached ligand with Cu^{2+} coordination optimized at the PM6 level. a) comparison of structures for $Cu^{2+}[\mathbf{L^1}d(G)]_4$ optimized at MMFF (blue) or PM6 (red). Hydrogen atoms have been omitted for clarity; b) side view of $Cu^{2+}[\mathbf{L^{1d}}d(G)]_4$; c) side view of $Cu^{2+}[\mathbf{L^{1c}}d(G)]_4$.

Astonishingly, for the shortest ligand $\mathbf{L^{1d}}$ an analogous Cu^{2+}-pyridine coordination site can be modelled, with apparently no unusual bond lengths or angles (see figure 6.15). On the one hand, this observation fits to the observed increase in stability upon Cu^{2+} addition, as obtained from the UV-VIS thermal denaturation experiments, but on the other hand cannot explain the unusual signal trend in the CD spectra as discussed above. Although the Cu^{2+}–Na^+ distance is much shorter than for the other ligands, in accordance with the decrease in linker length, 7.9 Å for $\mathbf{L^{1c}}$ to 5.7 Å for $\mathbf{L^{1d}}$, it is still much higher than the average G-quartet–G-quartet distance (∼3 Å), so any interaction should be unlikely. One explanation could be the spontaneous flipping of the top G-quartet (from *anti* to *syn* or vice versa) upon Cu^{2+} complexation. All four guanines have to flip simultaneously, otherwise at least one of the strand would have to be antiparallel to the others, based on the rules established earlier. This seems rather unlikely, as this would

require hydrogen bond rupture, flipping, and then rebonding, in total a supposedly very slow process, in contrast to the fast kinetics observed in the experiment (immediate change of CD bands). Another possibility would be the already described 5′–5′ stacking of two G-quadruplexes. As the linker is very short, it could favour intermolecular complexation of two pyridines from each G-quadruplex to one Cu^{2+} ion (see figure 6.16), however, this hypothesis cannot be proven or falsified by the available experimental data.

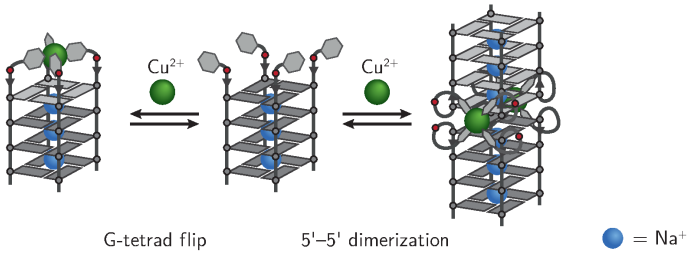


Figure 6.16.: Schematic models for the G-quadruplex $Cu^{2+}[\mathbf{L^{1d}}d(G)]_4$ leading to hypothetical (left) G-tetrad flipping, or (right) 5′–5′ dimerization.

Nevertheless, all of these results indicate that variations in the ligand structure are tolerated by the G-quadruplex assembly, giving freedom to design new and more unusual coordination environments and to go to topological more complicated G-quadruplexes.

6.3. Second-generation ligand L^2 in tetramolecular G-quadruplexes

A significant step towards G-quadruplexes with more advanced topologies, including unimolecular assemblies with strand connecting loops, is the incorporation of ligand $\mathbf{L^2}$ into G-quadruplex forming oligonucleotides. The advantage of ligand $\mathbf{L^2}$ over the first-generation ligands of type $\mathbf{L^1}$ is the possibility to attach it in any position (5′, 3′, and internal) of the oligonucleotide. Although the structural differences

between $\mathbf{L^1}$ and $\mathbf{L^2}$ are only subtle (see also figure 5.10 on page 75), the slight increase in steric bulk could have a negative impact on the formation of the Cu^{2+}–pyridine complex. For a complete list of the denaturation temperatures for the G-quadruplexes discussed in this chapter see the tables 6.3 and 6.4 at the end (page 123).

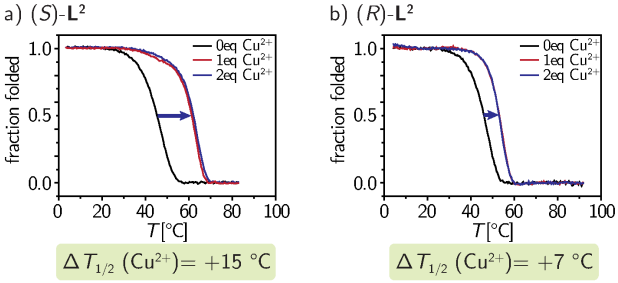


Figure 6.17.: Thermal denaturation curves (monitored at 295 nm) converted to the fraction folded values for the G-quadruplex $[\mathbf{L^2}d(G_4)]_4$ upon addition of $CuSO_4$, with a) (*S*)-$\mathbf{L^2}$ and b) (*R*)-$\mathbf{L^2}$. 10 mM LiCaco pH 7.3, 100 mM NaCl, 7.5 µM DNA, 0, 1.875 or 3.75 µM metal salt.

So, G-quadruplexes of the type $[\mathbf{L^2}^{R/S}d(G_4)]_4$ were synthesized and compared to the results obtained for ligand $\mathbf{L^1}$ in the ability to bind Cu^{2+}. Fortunately, the thermal denaturation temperatures are nearly equal to that obtained for $\mathbf{L^1}$, both for (*R*)-$\mathbf{L^2}$ and (*S*)-$\mathbf{L^2}$. The newly introduced chiral centre within the ligand apparently does not have a great influence on the thermal stability of the corresponding G-quadruplex. When $[\mathbf{L^2}^S d(G_4)]_4$† is treated with Cu^{2+} an increase in the denaturation temperature slightly lower to that found for the $\mathbf{L^1}$ case is observed ($\Delta T_{1/2} = +15$ °C), possibly reflecting the slightly increased steric hindrance of ligand $\mathbf{L^2}$ relative to $\mathbf{L^1}$. This observation changes significantly when the enantiomer (*R*)-$\mathbf{L^2}$ is used, as here,

†The stereodiscriptor (*R*) or (*S*) refers to that of the chiral centre in solketal **15** used in the synthesis of $\mathbf{L^2}$, because it can change within $\mathbf{L^2}$ depending on the position of the ligand inside the DNA: For 3′ and internal positions the stereodiscriptor is retained, whereas in the 5′ position it flips from (*S*) to (*R*) and vice versa.

only a moderate increase in stability is detected ($\Delta T_{1/2} = +7$ °C). These differences in stability have to arise from different conformational arrangements of the respective chiral ligand with respect to the chiral 5′ G-quartet. Moreover, the Cu^{2+}–pyridine complex can be present as two enantiomers, due to the pyridine rings propeller arrangement. In combination with the chiral ligand, a pair of diastereomers is formed, one of which can be more stable than the other. When these diastereomers interact with the chiral G-quadruplex, one combination of propeller arrangement and ligand chirality may fit better than the others.† This phenomenon was also described for chiral phenanthroline-based silver complexes interacting with duplex DNA.[256]

The effect observed here is indeed promising in terms of future applications, as it allows for the fine-tuning of the metal coordination environment, e.g. for catalysis.

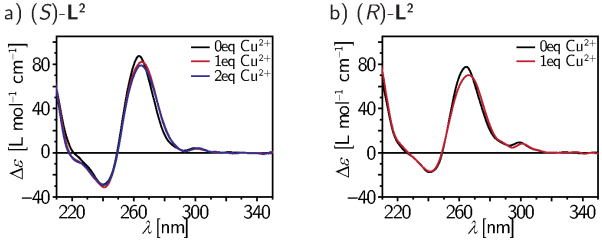


Figure 6.18.: CD spectra for the G-quadruplex $[\mathbf{L^2}d(G_4)]_4$ upon addition of $CuSO_4$, with a) (*S*)-$\mathbf{L^2}$ and b) (*R*)-$\mathbf{L^2}$. 10 mM LiCaco pH 7.3, 100 mM NaCl, 7.5 µM DNA, 0, 1.875 or 3.75 µM metal salt.

The CD spectra (see figure 6.18) confirm the formation of a parallel G-quadruplex topology $[\mathbf{L^2}d(G_4)]_4$, with only slight differences regarding signal intensity and position when comparing the (*S*) and the (*R*) form. The pronounced contribution of heteropolar coupling, observed as a band at 295 nm in case of ligand $\mathbf{L^1}$, is neglectable for ligand $\mathbf{L^2}$, both in the (*S*) or (*R*) variant. Upon Cu^{2+} addition the CD band at around 260 nm decreases slightly in intensity and shifts to higher wavelengths, which is in pronounced contrast to the effect observed with $\mathbf{L^1}$, where a signal

†Thanks to Prof. J. Müller (Münster university) for drawing attention to this effect.

increase was detected. These findings indicate that for $[\mathbf{L^1}\mathrm{d(G_4)}]_4$ the decrease of signal at 295 nm is closely related to the increase observed at 260 nm. This also means that the previous assumption about the relationship between G-quadruplex stability and CD signal intensity becomes rather unlikely. The increase of signal at 260 nm and decrease at 240 nm is apparently only an effect of the structural rearrangements in the G-quadruplex due to the metal ion coordination and thus alterations of the electronic transitions.

Despite the observed subtle differences, these results nevertheless encourage to now make use of the possibility to incorporate the ligand $\mathbf{L^2}$ at other positions than at the 5′ end.

First, the ligand was attached at the opposite end, namely at the 3′ end. As the used CPG support already contained one of the four standard nucleobases, G-quadruplexes of the type $[\mathrm{d(G_4)}\mathbf{L^2}\mathrm{dT})]_4$ were synthesized. The 3′ thymine can of course have an impact on both the G-quadruplex formation and the Cu^{2+} complexation, so the results for this G-quadruplex are not directly comparable to those obtained from the one described above. Although the CD spectra indicate formation of parallel tetramolecular G-quadruplexes, the thermal denaturation curves do not exhibit a single transition but several, all quite insensitive to the addition of Cu^{2+} ions (see the appendix). This suggests the formation of higher-order structures [173,257,258], probably with slipped strands like in interlocked G-quadruplexes [259], and not the presence of discrete G-quadruplexes. Taking this into account, the pyridine donor functionalities would not be preorganized to accommodate the Cu^{2+} ion, thus preventing coordination. No effect was also observed in the case that the Cu^{2+} ion is present during the annealing step, so the metal is not able to act as a template to steer the annealing process into the direction of discrete G-quadruplexes.

To circumvent this problem, in a second attempt the G-quadruplex was capped at the 5′ end with one thymine, each, and the thermal denaturation and CD studies were repeated. With this modification the resulting G-quadruplex $[\mathrm{d(TG_4)}\mathbf{L^2}\mathrm{dT})]_4$ behaved in analogy to the 5′ ligand appended one described above. Both the CD spectra and thermal

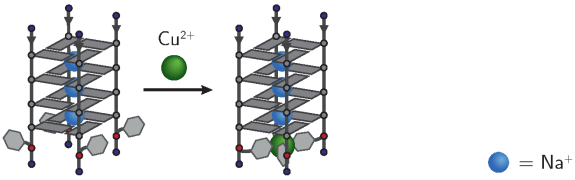


Figure 6.19.: Schematic models for the G-quadruplex $[d(TG_4)\mathbf{L^2}dT]_4$ and incorporation of Cu^{2+}.

denaturation profiles indicate clean G-quadruplex formation, while the stabilizing effect of the Cu^{2+} ion is restored (see the appendix), although the increase in stabilization is not as pronounced as for the $[\mathbf{L^2}d(G_4)]_4$ G-quadruplex ($\Delta T_{1/2}$ (*S*-$\mathbf{L^2}$) = +7 °C, $\Delta T_{1/2}$ (*R*-$\mathbf{L^2}$) = +8 °C).

The next obvious possibility is to combine both versions and to construct a G-quadruplex with two Cu^{2+} coordination sites, one at each end. The G-quadruplex now resembles a supramolecular cage comparable to the artificial M_2L_4 cages (with e. g. M = Pd^2+ or Cu^2+) recently reviewed in the literature (see figure 6.20).[200] The only difference to the previous G-quadruplex $[d(TG_4)\mathbf{L^2}dT)]_4$ is the substitution of the 5′ deoxythymidine against the ligand $\mathbf{L^2}$ resulting in $[\mathbf{L^2}d(G_4)\mathbf{L^2}dT)]_4$.

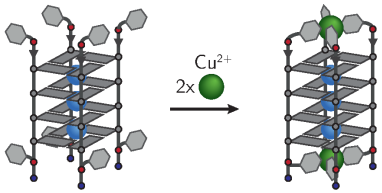


Figure 6.20.: Schematic models for the G-quadruplex $[\mathbf{L^2}d(G_4)\mathbf{L^2}dT]_4$ and incorporation of Cu^{2+}.

When no Cu^{2+} is added, the thermal stabilities are quite similar, with $[\mathbf{L^2}d(G_4)\mathbf{L^2}dT)]_4$ being slightly less stable than in the case with T at the 5′ end. Compared to $[\mathbf{L^2}d(G_4)]_4$ the G-quadruplex is however significantly more stable, in accordance with the reported beneficial

role of extra nucleobases at the 3′ end.[173] For all of the tetramolecular G-quadruplexes discussed above, metal complexation seems to be strong for binding of Cu^{2+}, as after addition of the first equivalent metal per G-quadruplex no further stabilization is observed. In case of $[\mathbf{L^2}d(G_4)\mathbf{L^2}dT)]_4$, there are two binding sites, so the thermal denaturation profiles show an additional increase in thermal stability upon the second equivalent of Cu^{2+}. Considering the increase in denaturation temperature for $[d(TG_4)\mathbf{L^{2\,}}^{S}dT)]_4$ ($\Delta T_{1/2} = +7$ °C) and $[\mathbf{L^{2\,}}^{S}d(G_4)dT]_4$ ($\Delta T_{1/2} = +15$ °C), one could expect an increase of $\Delta T_{1/2} = \sim+22$ °C for $[\mathbf{L^{2\,}}^{S}d(G_4)\mathbf{L^{2\,}}^{S}dT)]_4$ with 2 equiv. of Cu^{2+}. The effect of the Cu^{2+} ions, however, does not seem to be just additive. Instead, a rise of $\Delta T_{1/2} = +14$ °C is detected. Still, looking at the absolute melting temperatures, the G-quadruplex $(Cu^{2+})_2[\mathbf{L^{2\,}}^{S}d(G_4)\mathbf{L^{2\,}}^{S}dT]_4$ is more stable than both $Cu^{2+}[d(TG_4)\mathbf{L^{2\,}}^{S}dT]_4$ and $Cu^{2+}[\mathbf{L^{2\,}}^{S}d(G_4T)]_4$.

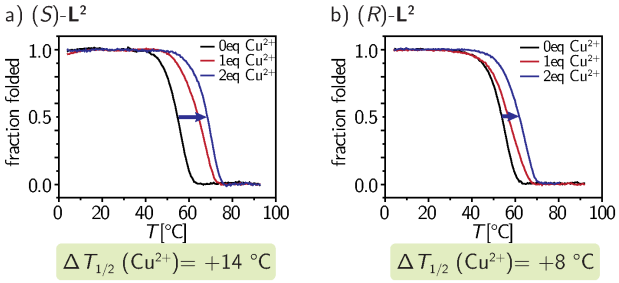


Figure 6.21.: Thermal denaturation curves (monitored at 295 nm) converted to the fraction folded values for the G-quadruplex $[\mathbf{L^2}d(G_4)\mathbf{L^2}dT]_4$ upon addition of $CuSO_4$, with a) (*S*)-$\mathbf{L^2}$ and b) (*R*)-$\mathbf{L^2}$. 10 mM LiCaco pH 7.3, 100 mM NaCl, 7.5 µM DNA, 0, 1.875 or 3.75 µM metal salt.

An interesting effect is observed for the G-quadruplex with two ligands (*R*)-$\mathbf{L^2}$. Upon addition of one equivalent Cu^{2+} the denaturation profile becomes much broader, and sharpening again when a further equivalent is added, while the average between the 0 and 2 equiv. curve is identical to the one observed for 1 equiv. This phenomenon cannot be explained if the Cu^{2+} ions were bound preferentially at one of the two

binding sites. Although it could indicate cooperativity between the two $(\text{pyridine})_4$ pockets, which would result in the 50:50 population of the metal free and the two metal bound state, this is rather unlikely. Instead one can expect the binding sites to be similar, as they are composed of identical donor functionalities, leading to a statistical mixture (no Cu^{2+}, 1 Cu^{2+} either at top or bottom, 2 Cu^{2+} ions) in turn resulting in an average denaturation profile. For the ligand enantiomer (*S*)-**L²** this effect is not that pronounced, but the 1 equiv. melting curve is still slightly broader than the 0 or 2 equiv. curve. As $\Delta T_{1/2}$ is greater than for (*R*)-**L²**, the average would be a denaturation curve clearly exhibiting a double transition. This is not the case, so a cooperative effect of the two binding sites can be ruled out.

The CD spectra for both G-quadruplexes are similar to those obtained for the tetramolecular G-quadruplexes discussed above. Only the overall signal intensity seems to be greater than in the previous cases. An increase in CD signal does not have to correlate with the thermodynamic stability, as discussed above. This can also be seen for the effect of Cu^{2+} ion addition, for which a slight decrease in CD signals is detected, although both the UV-VIS thermal denaturation and the thermal difference spectra indicate a higher thermodynamic stability and thus a higher concentration of the G-quadruplex at the same temperature.

Similar to the results discussed in chapter 6.1, it is interesting to examine the effect of shorter or longer G-tracts within the assembly. When an additional G-quartet is incorporated into the just discussed G-quadruplex, yielding $[\mathbf{L^2}\text{d}(\text{G}_5)\mathbf{L^2}\text{dT}]_4$, the assembly should become even more stable. Indeed, the denaturation temperature shifts more than 40 °C to higher values. Even at 95 °C (the upper limit for the investigated aqueous samples) only a small portion of the G-quadruplex denatures, whereas full denaturation can only be triggered by disruption of the hydrogen bonds in the G-quartets by addition of base (aq. LiOH). This high stability makes it difficult to investigate the effect of Cu^{2+} ion addition to the G-quadruplex. It seems that the Cu^{2+} ions are inducing a lowering of the thermal stability, which could be explained by the

high rigidity of the G-quadruplex with five G-quartets. To identify if this is indeed the case, the same G-quadruplexes were prepared with Cs^+ instead of Na^+ as the electrolyte cation. The much bigger Cs^+ ion lowers the thermodynamic stability significantly, so that the whole denaturation profile becomes visible again. Going from Na^+ to Cs^+ as the G-quadruplex stabilizing cation not only lowers $T_{1/2}$, but also effects the shape of the profile itself, which becomes much broader and with additional transitions in the low and medium temperature range. These findings indicate the reduced thermal stability and the higher conformational flexibility, resulting in the formation of higher order structures. As anticipated, addition of Cu^{2+} ions again leads to a distinct increase in thermal stability ($\Delta T_{1/2}$ = +6 °C (1 equiv.), +12 °C (2 equiv. Cu^{2+})) with a simultaneous sharpening of the denaturation curves.

The opposite approach towards the thermal stability is to reduce the number of G-quartets, in this case to only three, resulting in $[\mathbf{L^2}d(G_3)\mathbf{L^2}dT]_4$. In analogy to $[\mathbf{L^1}d(G_4)]_4$ no G-quadruplex formation is visible in the absence of Cu^{2+} and also for both enantiomers of ligand $\mathbf{L^2}$. This is best seen again in the CD spectra, in which a positive band at around 255 nm and a negative one at 275 nm indicates the presence of only the oligonucleotide single strands (see figure 6.22a black curve).

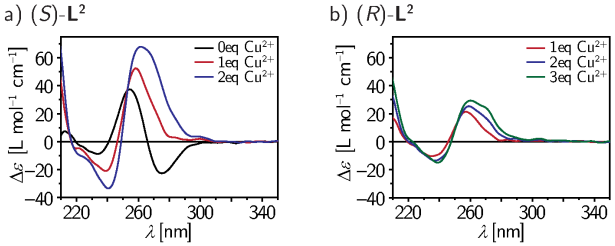


Figure 6.22.: CD spectra for the G-quadruplex $[\mathbf{L^2}d(G_3)\mathbf{L^2}dT]_4$ upon addition of $CuSO_4$ and after a denaturation and annealing cycle, with a) (*S*)-$\mathbf{L^2}$ and b) (*R*)-$\mathbf{L^2}$. 10 mM LiCaco pH 7.3, 100 mM NaCl, 7.5 µM DNA, 0, 1.875, 3.75, 5.625 µM metal salt.

When the annealing step is repeated in the presence of Cu^{2+} ions, the CD spectrum nicely shows the formation of tetramolecular G-quadruplexes. Increasing amounts of Cu^{2+} further trigger the assembly, and whereas for the G-quadruplex with (*R*)-**L²** this effect is only weakly distinct, it is the more pronounced for (*S*)-**L²**. The relatively weak overall CD signals for the (*R*)-**L²** variant can most likely be attributed to a low thermal stability.

The UV-VIS based thermal denaturation profiles support this assumption. The G-quadruplex [**L²**d(G$_3$)**L²**dT]$_4$ with (*R*)-**L²** behaves quite similar to [**L¹**d(G$_3$)]$_4$ with denaturation temperatures around $T_{1/2}$ = 20 °C, even in the presence of higher amounts of Cu^{2+}. In contrast, the variant with (*S*)-**L²** exhibits a much higher thermal stability ($T_{1/2}$ = 34 °C). Also here, higher amounts of coordinating metal cation do not lead to dramatic changes in the melting temperature. Nevertheless, the overall jump in stability is quite remarkable, going from single strands directly to a system which does not denature at room temperature.

In an attempt to stabilize the Cu^{2+}-free G-quadruplex, the experiments were repeated with KCl as the electrolyte, something which was not done for ligand **L¹**. Astonishingly, even with the highly stabilizing K^+ ion, no G-quadruplex formation could be detected. In this sense, the system behaved exactly like with Na^+ as the electrolyte. Cu^{2+} ions were indeed able to trigger G-quadruplex assembly, when present during renaturation. The resulting G-quadruplex now has an even higher resistance to thermal denaturation, with a pronounced influence of additional equivalents of Cu^{2+}. Taking 0 °C as the lowest obtainable temperature (due to the high amount of salt, no freezing is observed at this temperature), this means an outstanding stabilization of $\Delta T_{1/2}$ = +63 °C (*R*)-**L²**, 3 equiv. Cu^{2+}) and $\Delta T_{1/2}$ = +70 °C (*S*)-**L²**, 3 equiv. Cu^{2+}), respectively. These results nicely demonstrate that the Cu^{2+} ions act as a template for the G-quadruplex formation and do not only bind to the already preformed and preorganized "chelate G-quadruplex ligand".

Going back to the discussion about the preferential binding site of the Cu^{2+} ions, the melting profiles exhibit an interesting feature.

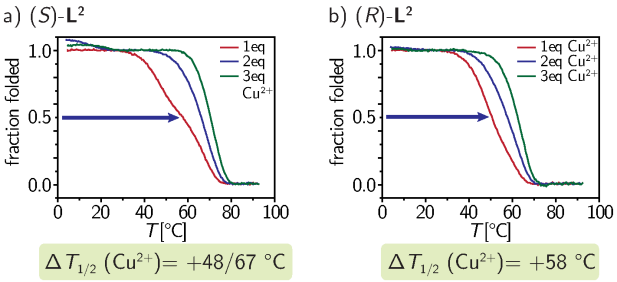


Figure 6.23.: Thermal denaturation curves (monitored at 295 nm) converted to the fraction folded values for the G-quadruplex $[\mathbf{L^2}\mathrm{d(G_3)}\mathbf{L^2}\mathrm{dT}]_4$ upon addition of $CuSO_4$, with a) (*S*)-$\mathbf{L^2}$ and b) (*R*)-$\mathbf{L^2}$. 10 mM LiCaco pH 7.3, 100 mM NaCl, 7.5 µM DNA, 1.875, 3.75, or 5.125 µM metal salt (addition prior annealing).

When only one equivalent of Cu^{2+} is present, the corresponding G-quadruplex melting profile shows a double-transition, best visible for $[\mathbf{L^2}\mathrm{d(G_3)}\mathbf{L^2}\mathrm{dT}]_4$ with (*S*)-$\mathbf{L^2}$ (see figure 6.23a red curve). Denaturation temperature elucidation by fitting the fraction folded curve with a double sigmoidal function (biphasic dose response, implemented in OriginPro 8.5[260]) results in two temperatures $T_{1/2}$ = 48 and 67 °C, respectively. The second temperature is identical to the one found for the denaturation curve in the presence of 2 eq. of Cu^{2+}. These findings strongly support the assumption made above that the curve with 1 equiv. Cu^{2+} is in fact the additive between all four binding states. For the G-quadruplex $[\mathrm{d(TG_4)}\mathbf{L^2}\mathrm{dT}]_4$ with the ligand $\mathbf{L^2}$ at the 3′ end it was evident that the binding affinity of Cu^{2+} is lower than if the ligand is attached to the 5′ end. This was based on the gradual increase of the denaturation temperature with higher amounts of Cu^{2+}, in contrast to the "all or nothing" case for $[\mathbf{L^2}\mathrm{d(G_4)}]_4$. Keeping this in mind, the further increase in stability as observed for $[\mathbf{L^2}\mathrm{d(G_3)}\mathbf{L^2}\mathrm{dT}]_4$ with higher amounts of Cu^{2+} (from two to three equivalents) can be rationalized in the same way. Taken together, the two metal binding sites seem to act independently from another, so no cooperativity of the Cu^{2+} binding occurs.

A peculiar phenomenon is observed in the corresponding CD spectra. While different amounts of Cu^{2+} do neither change intensity nor position of the CD bands for (*S*)-**L²**, both are affected in case of (*R*)-**L²**. Here, the more Cu^{2+} is present, the more CD signal is lost, while the band starts to broaden (or even to split into two signals). Either the already quite old DNA samples (at the time of the CD measurements) were already partially degraded, or an interaction is only present for the (*R*) enantiomer of the ligand which is yet not accounted for. So these experiments have to be repeated with new DNA samples.

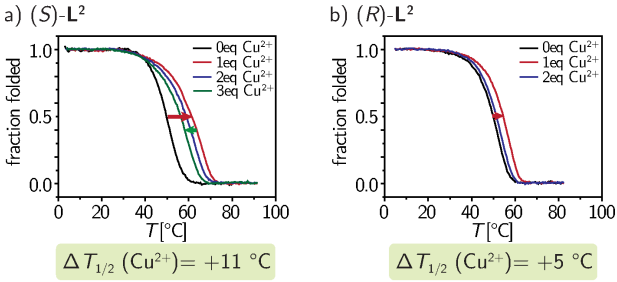


Figure 6.24.: Thermal denaturation curves (monitored at 295 nm) converted to the fraction folded values for the G-quadruplex $[(\mathbf{L^2})_2\mathrm{d}(\mathrm{G}_4)]_4$ upon addition of $CuSO_4$, with a) (*S*)-**L²** and b) (*R*)-**L²**. 10 mM LiCaco pH 7.3, 100 mM NaCl, 7.5 µM DNA, 0, 1.875, 3.75, or 5.125 µM metal salt.

The question which now arises is what happens if the two Cu^{2+} binding sites are located not on opposite faces but on the same side of the G-quadruplex stem. For this purpose G-quadruplexes of the type $[(\mathbf{L^2})_2\mathrm{d}(\mathrm{G}_4)]_4$ were synthesized and tested for their ability to bind Cu^{2+} ions. In the absence of Cu^{2+}, as seen in figure 6.24, the G-quadruplex is slightly more thermally stable than its counterpart $[\mathbf{L^2}\mathrm{d}(\mathrm{G}_4)]_4$ with only one ligand **L²** at each 5′-end. Likewise, the chirality of the ligand **L²** does not have an impact on the thermal stability, when no copper is bound. Addition of one equivalent of Cu^{2+} per G-quadruplex leads to the already established increase in thermal stability, the extent is however slightly lower when compared to that for $[\mathbf{L^2}\mathrm{d}(\mathrm{G}_4)]_4$. Higher

amounts of the transition metal ion now have a detrimental effect: although the denaturation temperature is still higher than in the metal free state, it is lower than for one equivalent Cu^{2+} (see figure 6.24). In other words, the first Cu^{2+} ions stabilizes $[\mathbf{L^2}d(G_4)]_4$, while further metal ions destabilize the assembly.

Looking at the structure of the G-quadruplex, different Cu^{2+} binding modes are possible, both for one or two bound metal ions per G-quadruplex (see figure 6.25). Similar to the previous discussion where the two copper ions were located on opposite sites of the G-quadruplex stem it is also possible that a mixture of several binding modes will be present, although the likelihood for a preferred arrangement is much greater here due to the direct linkage of the ligands. In lack of further structural information no definitive conclusion about the coordination site of the Cu^{2+} can be drawn.

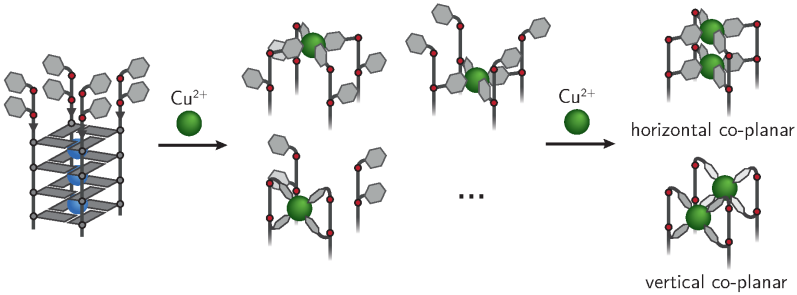


Figure 6.25.: Schematic models for the G-quadruplex $[(\mathbf{L^2})_2d(G_4)]_4$ and possible binding modes for the incorporation of Cu^{2+}. The dots denote further possible arrangements.

Concerning the thermal stabilities it is likely that the binding modes between the first Cu^{2+} ion and the second are different, with a structural reorganisation upon binding of the second Cu^{2+} ion. Electrostatic repulsion between the two highly charged copper cations can be neglected, because of the large distance between the two ions, even for the "horizontal co-planar" coordination (with respect to the G-quartet plane, figure 6.25, distance is ~6 Å). Furthermore, the "horizontal co-planar"

coordination should lead to a further, but perhaps smaller, increase in thermodynamic stability. Thus, the "vertical co-planar" coordination of the two Cu^{2+} ions seems more likely. The CD spectra of the corresponding G-quadruplexes show that the bound Cu^{2+} ions do not greatly influence the stacking of the G-quartets, as the CD bands are more or less insensitive to the metal ion addition (see the appendix). Based on these considerations, the observed effects can be rationalised as follows: The first Cu^{2+} ion binds, similar to all the previous examples, in a "horizontal co-planar" coordination to the one pyridine of the four oligonucleotide strands, each. This binding mode stabilises the G-quadruplex and gives rise to the observed increase in the melting temperature. The second Cu^{2+} shifts the equilibrium towards the "vertical co-planar" coordination, which should have a much less beneficial effect on the G-quadruplex stability, thus the observed decrease in the melting temperature.

It is interesting to note that for this G-quadruplex $[(\mathbf{L^2})_2\mathrm{d}(\mathrm{G}_4)]_4$, the effect of the different ligand enantiomers is quite pronounced. Whereas for (*S*)-$\mathbf{L^2}$ the rise in stability upon addition of 1 eq. Cu^{2+} is quite distinct ($\Delta T_{1/2} = +11$ °C), this effect is much less pronounced in case of the system with (*R*)-$\mathbf{L^2}$ ($\Delta T_{1/2} = +5$ °C). Apparently, the ligand (*S*)-$\mathbf{L^2}$ can accommodate the necessary conformational arrangement for the Cu^{2+} binding much better than the enantiomer (*R*)-$\mathbf{L^2}$.

In summary, this chapter dealt with novel tetramolecular G-quadruplexes, which carry pyridine functionalities covalently attached at the 5′ or 3′ position of each oligonucleotide strand. The four pyridine moieties are capable of coordination to a transition metal ion, like Cu^{2+} or Ni^{2+}, leading to a metal-base tetrad in analogy to the native G-quartet of G-quadruplexes. The systems were fully characterized not only with UV-VIS thermal denaturation experiments and CD spectroscopy but also with complementary techniques, including gel electrophoresis and EPR. In combination, the tetramolecular G-quadruplex formation, parallel strand orientation, thermal stabilities and the nature and impact of the square-planar coordination of the bound transition metal ion could be established. The tolerance of the G-quadruplex core towards ligand

alterations was demonstrated for different linker lengths, which did not interfere with G-quadruplex formation. Furthermore, also two of the metal binding sites could be incorporated into the G-quadruplex, either at opposite or at same sides. Transition metal ion coordination is fully reversible and can even trigger G-quadruplex formation by acting as a template. Moreover, the ability of the two ligand enantiomers to form the metal base-tetrad was evaluated, revealing that the (S) form has superior properties when compared to the (R) enantiomer.

These promising results for the symmetric tetramolecular G-quadruplexes presented here encouraged to establish more complexity in the G-quadruplex assembly and to go to biological more relevant topologies, that is unimolecular G-quadruplexes, which will be dealt with in the next chapter.

Table 6.3.: Overview of denaturation temperatures $T_{1/2}$ for the tetramolecular G-quadruplexes with ligand (R/S)-$\mathbf{L^2}$.

G-quadruplex	$\mathbf{L^2}$	**salt**	**metal**	**equiv.**	$T_{\mathbf{1/2}}$ [°C]	$\Delta T_{\mathbf{1/2}}$ [°C]
$[\mathbf{L^2}\mathrm{d}(\mathrm{G_4})]_4$	*S*	NaCl	$CuSO_4$	0	45.2	
				1	61.2	+16.0
				2	62.0	+16.8
	R	NaCl	$CuSO_4$	0	46.0	
				1	53.0	+7.0
				2	52.9	+6.9
$[\mathrm{d}(\mathrm{TG_4})\mathbf{L^2}\mathrm{dT}]_4$	*S*	NaCl	$CuSO_4$	0	56.5	
				1	63.9	+7.4
				2	66.4	+9.9
	R	NaCl	$CuSO_4$	0	56.6	
				1	64.3	+6.7
				2	66.7	+10.1
$[\mathbf{L^2}\mathrm{d}(\mathrm{G_4})\mathbf{L^2}\mathrm{dT}]_4$	*S*	NaCl	$CuSO_4$	0	54.6	
				1	64.4	+9.8
				2	68.6	+14.0
	R	NaCl	$CuSO_4$	0	54.0	
				1	57.1†	+3.1
				2	61.8	+7.8
$[\mathbf{L^2}\mathrm{d}(\mathrm{G_5})\mathbf{L^2}\mathrm{dT}]_4$	*S*	NaCl	$CuSO_4$	0–2	≥95	
		CsCl	$CuSO_4$	0	52.1‡	
				1	66.4‡	+13.9
				2	74.5‡	22.0
	R	NaCl	$CuSO_4$	0	≥90	
				1–2	≥88	–2
		CsCl	$CuSO_4$	0	53.1‡	
				1	59.1‡	+6.0
				2	65.2‡	+12.1

†Apparent denaturation temperature. See the text for details.

‡Very broad and multiple denaturation transitions. Only rough $T_{1/2}$ determination.

Table 6.4.: Continuation of table 6.3 Overview of denaturation temperatures $T_{1/2}$ for the tetramolecular G-quadruplexes with ligand (*R*/*S*)-**L²**.

G-quadruplex	L²	salt	metal	equiv.	$T_{1/2}$ [°C]	$\Delta T_{1/2}$ [°C]
[**L²**d(G$_3$)**L²**dT]$_4$	*S*	NaCl	CuSO$_4$	0	≤0	
				1	33.8	≥+33.8
				2	37.5	≥+37.5
				3	38.3	≥+38.3
		KCl	CuSO$_4$	0	≤0	
				1	47.8/66.8†	≥+47.8/66.8
				2	66.8	≥+66.8
				3	70.4	≥+70.4
	R	NaCl	CuSO$_4$	0	≤0	
				1	∼20	≥+20
				2	∼23	≥+23
		KCl	CuSO$_4$	0	≤0	
				1	50.6‡	≥+50.6
				2	58.2	≥+58.2
				3	62.5	≥+62.5
[(**L²**)$_2$d(G$_4$)]$_4$	*S*	NaCl	CuSO$_4$	0	50.1	
				1	61.5	+11.4
				2	59.2	+9.1
				3	56.2	+6.1
	R	NaCl	CuSO$_4$	0	49.9	
				1	54.6	+4.7
				2	51.3	+1.4

†Double transition. See the text for details.

‡Apparent denaturation temperature. See the text for details.

7. Unimolecular telomeric G-quadruplexes

The previous chapter focused on the results for ligand modified tetramolecular G-quadruplexes. Based on these results, the ligand system is brought to a next level by incorporating up to four ligands into one oligonucleotide strand, which is able to assemble into unimolecular G-quadruplexes.

It could be shown that donor-functionalized G-rich oligonucleotide strands can be assembled into tetramolecular G-quadruplexes and are capable of binding a transition metal ion like Cu^{2+} in a square-planar fashion. Improvement of the ligand system led to the design and synthesis of ligand $\mathbf{L^2}$, which can be incorporated into oligonucleotide strands at any position. Moreover, up to two ligands were inserted into one single Guanine-rich strand enabling the uptake of more than one transition metal ion in the corresponding tetramolecular G-quadruplex. As this approach proved to be successful, the question arose if even more instances of the ligand $\mathbf{L^2}$ are tolerated by the G-quadruplex core. This would enable the synthesis of a single oligonucleotide strand, containing four of the pyridine donor-functionalities, thereby bringing the concept of the metal base-quartet to the even more variable unimolecular G-quadruplexes. Here, most structural data are available for the human telomere repeating unit d(TTAGGG), so the first part focuses on this sequence.

7.1. The metal base-quartet motif in the human telomere sequence

Incorporation of the transition metal–pyridine coordination complex into a unimolecular G-quadruplex requires a little more consideration than in the case of the tetramolecular assemblies. The tetramolecular G-quadruplex assemblies do not need connecting loops, which are required for the folding of unimolecular G-quadruplexes. So, not only is the most stable unimolecular topology very sensitive to alterations in the sequence, but also do the strand connecting loops need some spatial freedom. That means that first, exchange of a canonical nucleobase for the ligand **L²** may alter the strand arrangement in the resulting G-quadruplex topology, second, complex formation between the transition metal ion and the pyridine units could be prevented by the close proximity of loop nucleotides, and third, metal ion coordination could lead to major structural rearrangements. Looking into the literature reported structures based on NMR spectroscopy and X-ray analyses, more than ten different topologies have been found for the human telomere sequence, five of them showing a unimolecular assembly (see figure 7.1).[261] Among these, the sequence d[A(GGGTTA)$_3$GGG] (htel22) was chosen, as the solution state structure of the corresponding G-quadruplex, in the presence of Na^+ ions, comprises two parallel lateral loops at the face of one G-quartet (figure 7.1f). The lateral loops should leave enough space above the G-quartet plane to accommodate the metal complex, something which was not expected for diagonal or propeller type loops. Furthermore, as described in the introduction, the group of J.-L. Mergny have also used successfully two adjacent lateral loops for the formation of a T–Hg^{2+}–T metal base-pair positioned right above the top G-quartet in a unimolecular G-quadruplex.[165]

The effect of the modifications on the topology of the anticipated G-quadruplex cannot be predicted, so the first step was to test how the ligand incorporation effects the respective folding behaviour and thermodynamic stability. Two modified versions of the htel22 sequence were investigated, which differ in the position of the ligand incorporation.

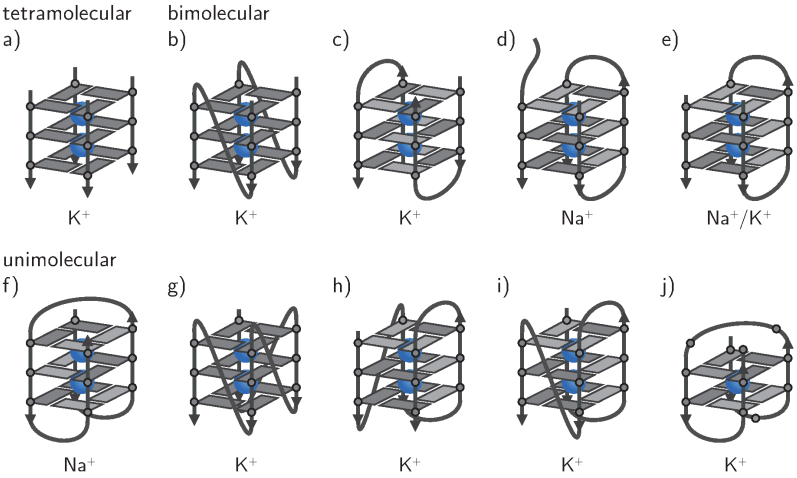


Figure 7.1.: Schematic structures of human telomeric G-quadruplexes. a) Tetramolecular G-quadruplex for d[TTAGGG] or d[TTAGGGT];[262] b) parallel bimolecular G-quadruplex for d[$(TAGGGT)_2$];[263,264] c) antiparallel bimolecular G-quadruplex for d[$(TAGGGT)_2$];[264] d) asymmetric bimolecular G-quadruplex for d[$GGG(TTAGGG)_2T$];[265] e) asymmetric bimolecular G-quadruplex for d[$GGG(TTAGGG)_2T$] and d[TAGGGT];[265] f) basket-form unimolecular G-quadruplex for d[$A(GGGTTA)_3GGG$];[266] g) propeller-type form unimolecular G-quadruplex for d[$A(GGGTTA)_3GGG$];[263] h) form 1 unimolecular G-quadruplex for d[$TA(GGGTTA)_3GGG$];[267–273] i) form 2 unimolecular G-quadruplex for d[$TA(GGGTTA)_3GGGTT$];[269,273,274] j) basket-type form unimolecular G-quadruplex for d[$(GGGTTA)_3GGGT$].[275] Guanines are depicted as grey spheres and grey rectangles; *anti* guanines are coloured in light gray, *syn* guanines in dark grey. Figure adapted (redrawn) from ref.[261]

In htel22-$\mathbf{L^2}_4$a, four loop nucleotides are exchanged for the ligand (*S*)-$\mathbf{L^2}$, whereas the G-quartet core remains in its wild-type. The second variant is htel22-$\mathbf{L^2}_4$b, in which the four guanines of one G-tetrad are replaced by four ligands (*S*)-$\mathbf{L^2}$. The unmodified G-quadruplex sequence htel22 was used as a reference.

Table 7.1.: Investigated G-quadruplex forming sequences based on the human telomeric repeat. Guanines involved in G-quartet formation are underlined.

DNA	**sequence 5′→3′**
htel22	d[A GGG TTA GGG TTA GGG TTA GGG]
htel22-$\mathbf{L^2}_4$a	d[A GGG $\mathbf{L^2}$T$\mathbf{L^2}$ GGG TTA GGG $\mathbf{L^2}$T$\mathbf{L^2}$ GGG]
htel22-$\mathbf{L^2}_4$b	d[A GG$\mathbf{L^2}$ TTA $\mathbf{L^2}$GG TTA GG$\mathbf{L^2}$ TTA $\mathbf{L^2}$GG]

7.1.1. Exchange of loop nucleotides

Like with the tetramolecular G-quadruplexes discussed in the previous chapter 6, UV-VIS thermal denaturation experiments and CD spectroscopy were used for all strands to test their ability to fold into G-quadruplexes. First the unmodified htel22 and the modified htel22-$\mathbf{L^2}_4$a sequence were compared with each other. In both cases, UV-VIS thermal difference spectra (see the appendix) and melting curves indicate G-quadruplex formation. In contrast to the previously studied tetramolecular assemblies, the G-quadruplex structures investigated here exhibit no hysteresis phenomenon in the de- and renaturation curves, indicative of the formation of unimolecular G-quadruplexes.[276]

The observed melting temperatures T_m for the unmodified sequence htel22 are in accordance with the literature values, both for the samples containing NaCl or KCl as electrolyte.[277] For the ligand $\mathbf{L^2}$ modified sequence htel22-$\mathbf{L^2}_4$a in the NaCl containing sample, the melting temperature is distinctly lower, when compared to the wild-type strand htel22 (ΔT_m = –7.3 °C). Interestingly, this difference in stability be-

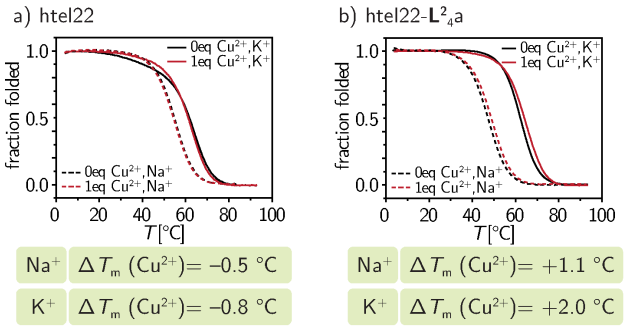


Figure 7.2.: Thermal denaturation curves (monitored at 295 nm) converted to the fraction folded values for the G-quadruplexes a) htel22, b) htel22-$\mathbf{L^2}_4$a upon addition of $CuSO_4$ in the presence of either NaCl (dotted lines) or KCl (solid lines). 10 mM LiCaco pH 7.3, 100 mM NaCl/KCl, 4.0 µM DNA, 0 or 4.0 µM $CuSO_4$.

tween wild-type and modified G-quadruplex is nearly negligible in the case of KCl as the stabilizing salt ($\Delta T_m = -1.8$ °C). This observation can be attributed to the different topologies of htel22 in Na^+ and K^+ solutions. G-quadruplex stability is not only governed by the G-quartet core but also by stacking of loop nucleotides onto the G-quartets and the related loop flexibility.[278] Indeed, explicit loop nucleobase stacking onto the G-quartets is observed for the solution structure of htel22.[266] Evidently, in the Na^+ case, these contributions are removed when substituting canonical nucleotides with the ligand $\mathbf{L^2}$, thus giving rise to the lower thermal stability. For the K^+ containing sample, these stabilising effects seem to be already absent in the wild-type sequence htel22, so introduction of the ligand $\mathbf{L^2}$ does not change the overall stability to a great extent. It is also noteworthy that for htel22 with the salt KCl, the melting transition is much broader than for htel22-$\mathbf{L^2}_4$a. This is either due to differences in stability as discussed in the introductory method chapter 4.3.1[169] or indicates the presence of lower melting, alternative topologies for htel22.

Next, the effect of Cu^{2+} addition was tested for both G-quadruplex assemblies. Similar to the results obtained for the unmodified tetramolec-

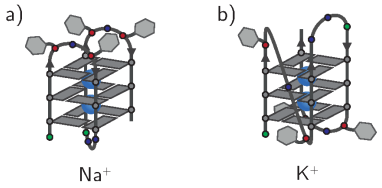


Figure 7.3.: Proposed topologies for htel22-$\mathbf{L^2}_4$a in the presence of a) Na^+ or b) K^+.

ular G-quadruplexes, the native sequence htel22 showed a very minor decrease in thermal stability when exposed to Cu^{2+} cations, both for Na^+ and K^+ containing buffer. Surprisingly, the effect was also negligible in case of htel22-$\mathbf{L^2}_4$a. The observed slight increases in thermal stability, $\Delta T_m = +1.1$ °C (Na^+) and $\Delta T_m = +2.0$ °C (K^+) are supposedly inside the margin of error. Furthermore, the spatial arrangement of the ligands inside the K^+ form of htel22-$\mathbf{L^2}_4$a should prevent direct intramolecular binding of a Cu^{2+} ion, unless the cation induces a topological rearrangement from the mixed (3+1) strand orientation to the antiparallel topology found in the Na^+ form (see figure 7.3).

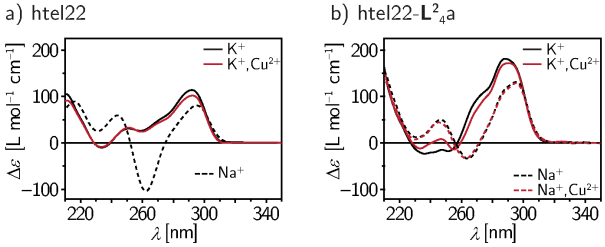


Figure 7.4.: CD spectra for the G-quadruplexes a) htel22, b) htel22-$\mathbf{L^2}_4$a upon addition of $CuSO_4$ in the presence of either NaCl (dotted lines) or KCl (solid lines). 10 mM LiCaco pH 7.3, 100 mM NaCl/KCl, 4.0 µM DNA, 0 or 4.0 µM $CuSO_4$.

The measured CD spectra for htel22 show the typical profiles for the antiparallel, basket-type topology in the presence of Na^+, as indicated by a positive CD signal at 295 nm and 244 nm, respectively, and a

negative CD band at 262 nm.[277] Likewise, the CD profile of the K^+ form with two maxima at 292 nm and 251 nm is indicative of the hybrid (3+1) assembly or a mixture of several of these, as reported in the literature.[277,279] Addition of $CuSO_4$ does not alter the CD spectrum to a great extent, in accordance with the assumption that Cu^{2+} ions do not significantly interfere with the G-quadruplex core.

In case of the G-quadruplex htel22-$\mathbf{L^2}_4$a, similar results were obtained. For the Na^+ form, the CD profile is, however, not that distinct, which can be explained by the absence of loop nucleotide stacking onto the G-quartets, or the presence of small amounts of other topologies apart from the antiparallel basket-type assembly. The CD profiles in the K^+ case are more comparable to the ones of the native htel22 sequence, which fits well to the observed similarities between the thermal stabilities of htel22 and htel22-$\mathbf{L^2}_4$a in the UV-VIS based melting experiments. The observed differences in the extinction coefficient values can be attributed to an inadvertent variation in the oligonucleotide concentration. No significant effect of Cu^{2+} ions on the CD profiles were observed, both for the Na^+ and the K^+ containing samples. Again, this fits well to the assumption made above that the spatial orientation of the ligands within htel22-$\mathbf{L^2}_4$a does not allow for the formation of the Cu^{2+}[pyridine]$_4$ coordination complex.

7.1.2. Topology change induced by *N*-methyl mesoporphyrin IX

As described above, the anticipated topology of htel22-$\mathbf{L^2}_4$a in K^+ solution should prevent intramolecular binding of the Cu^{2+} ions. However, in the UV-VIS melting experiment a slight increase in thermal stability was observed upon addition of Cu^{2+} ions. To test if the sample with htel22-$\mathbf{L^2}_4$a, KCl, and $CuSO_4$ partially contains an antiparallel topology, which allows Cu^{2+} coordination, binding of the planar porphyrin *N*-methylmesoporphyrin IX (NMM) was investigated.

NMM is known to bind to parallel G-quadruplexes and is able to convert hybrid structures like the (3+1) type G-quadruplex into the parallel stranded orientation, as this involves the flipping of only one of

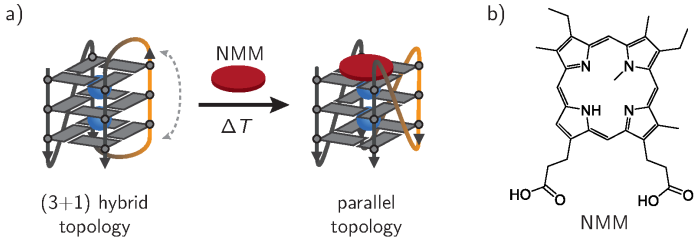


Figure 7.5.: Schematic illustration of the *N*-methyl mesoporphyrin IX (NMM) induced topology change of a (3+1) type G-quadruplex to a propeller-type parallel G-quadruplex. The flipping strand is highlighted in orange.

the four strand sections (see figure 7.5).[280,281] Although the binding process of NMM and thus the induced topology change is normally achieved at a temperature below the melting temperature of the G-quadruplex, the strand flipping could in principle also occur via the completely unfolded G-quadruplex. Complete unfolding is more likely, as not only the guanines of one strand section have to flip positions, but also those of least one of the G-quartets have to change their *syn–anti* conformations.

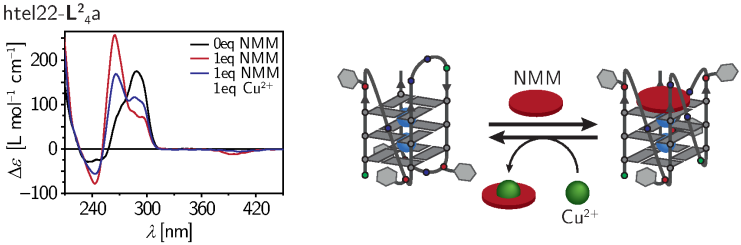


Figure 7.6.: *N*-methyl mesoporphyrin IX (NMM) induced topology change for the G-quadruplex htel22-$\mathbf{L^2}_4$a. Left: CD spectrum for htel22-$\mathbf{L^2}_4$a in the presence of no additive, 1 eq. NMM, and 1 eq. NMM together with 1 eq. $CuSO_4$; right: schematic illustration of the topology change upon NMM binding and the reverse reaction induced by addition of Cu^{2+} ions.

G-quadruplex samples for htel22-$\mathbf{L^2}_4$a were incubated with NMM both in the presence and absence of Cu^{2+} ions for 12 h at 30 °C and subsequently the CD spectra measured. The CD band at 289 nm decreases in intensity, whereas from the shoulder at 265 nm a new positive band emerges. Likewise, the negative signal at 243 nm becomes more distinct. Moreover, a new band at 395 nm appears, which can be attributed to a π–π^* transition of the bound NMM induced by chirality transfer. These spectral changes are indicative of the formation of a G-quadruplex assembly with all parallel strand polarities. So, like with the native htel22 strand, NMM is able to induce the same topology change from the supposedly (3+1) assembly of htel22-$\mathbf{L^2}_4$a to the all parallel G-quadruplex structure. Addition of $CuSO_4$ does indeed lead to a reversal of this trend: the CD signal at 265 nm decreases in intensity, whereas the signal at 289 nm intensifies. Also, the induced CD signal of the NMM at higher wavelengths starts to disappear. Apparently, the bound NMM is partially removed, which can be explained by Cu^{2+} binding to NMM. Indeed, complexation of Cu^{2+} to NMM was already reported in the literature, including a crystal structure showing that the Cu^{2+} cation is bound by the four nitrogen donor atoms.[282,283] All these results indicate that although the modified sequence htel22-$\mathbf{L^2}_4$a behaves both structurally and regarding the reactivity towards binding of NMM like the native htel22 G-quadruplex, but does unfortunately not allow the binding of Cu^{2+} ions to the pyridine moieties of the incorporated ligands $\mathbf{L^2}$.

7.1.3. Exchange of one G-quartet

Next, the sequence htel22-$\mathbf{L^2}_4$b was investigated, in which one G-quartet of the unmodified G-quadruplex htel22 is exchanged for four of the ligand $\mathbf{L^2}$, leaving the loop nucleotides untouched. This modification should not show the just discussed limitations, as the number of loop nucleotides is the same as in the native htel22 sequence. The overall thermal stability, however, can be expected to be much lower, since the G-quadruplex htel22-$\mathbf{L^2}_4$b is able to form only two stacked G-quartets, instead of the three in the htel22 G-quadruplex. This assumption is

justified by the thermal denaturation experiments, from which melting temperatures just above the lower temperature limit can be deduced. Again, no significant hysteresis is observed, suggesting the formation of unimolecular G-quadruplexes. Quite remarkably, the G-quadruplex htel22-$\mathbf{L^2}_4$b modified with (*R*)-$\mathbf{L^2}$ (T_m = 18.0 °C) is thermally more stable than that with the enantiomer (*S*)-$\mathbf{L^2}$ (T_m = 12.5 °C). Apparently, the two enantiomers exhibit different interactions, either attractive ones, like π–π stacking, or repulsive interactions with the G-quartet core or the loop nucleotides.

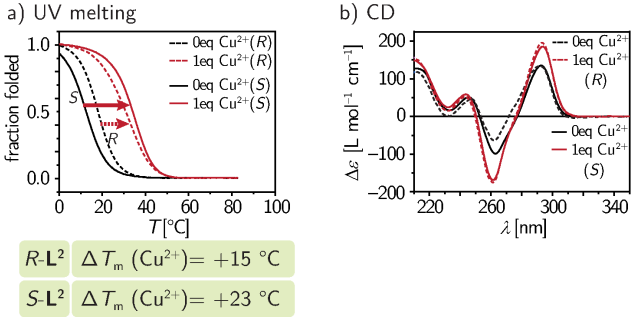


Figure 7.7.: a) Thermal denaturation curves (monitored at 295 nm) converted to the fraction folded values and b) CD spectra for the G-quadruplexes htel22-$\mathbf{L^2}_4$b with either ligand (*R*)-$\mathbf{L^2}$ or (*S*)-$\mathbf{L^2}$ upon addition of $CuSO_4$. 10 mM LiCaco pH 7.3, 100 mM KCl, 4.0 µM DNA, 0 or 4.0 µM $CuSO_4$.

This trend changes, however, when $CuSO_4$ is added to the G-quadruplex solutions. First of all, a remarkable increase in the thermodynamic stability is observed for the respective G-quadruplexes of both ligand $\mathbf{L^2}$ enantiomers. The melting temperatures increase by ΔT_m = +15 °C and ΔT_m = +23 °C for the G-quadruplex with ligand (*R*)-$\mathbf{L^2}$ and (*S*)-$\mathbf{L^2}$, respectively. Now the G-quadruplex with ligand (*S*)-$\mathbf{L^2}$, which showed a lower thermal stability without metal, now exhibits the higher absolute melting temperatures. This effect is an indicator of the structural rearrangement of the ligands upon binding to the Cu^{2+} ion. In the copper-free state, ligand (*R*)-$\mathbf{L^2}$ seems to form more favourable

stabilising interactions with other parts of the G-quadruplex structure, compared to the enantiomer (*S*)-**L²**. This beneficial spatial arrangement of (*R*)-**L²**, however, seems to be counteracting the coordination to the Cu^{2+} ion, which in contrast appears to be easier accommodated by the (*S*)-**L²** enantiomer.

To get more insights into the structure of the G-quadruplex assemblies, CD measurements were performed. The metal-free samples of htel22-$\mathbf{L^2}_4$b show two positive bands at 246 nm and 293 nm, respectively, and a negative one at 263 nm, which is typical for a purely antiparallel orientation of the G-rich strands.[279] This observation is in accordance with those made for the native sequence htel22, suggesting an analogues folding. The basket-like (type 3) topology with only two G-quartets closely resembles that formed by the human telomere sequence d[GGG(TTAGGG)$_3$T].[180,275] The CD spectra and thermal stabilities for this sequence, however, differ considerably from that of the modified G-quadruplex htel22-$\mathbf{L^2}_4$b. This difference can be explained by the formation of base triplets (GGG and GGA) stacking onto the G-quartets in case of the unmodified G-quadruplex, which cannot form in case of the modified sequence.[180,275] Likewise, the higher melting temperatures of the two-G-quartet type 3 G-quadruplex can also be rationalised by the presence of these interactions. The observed decrease in thermal stability when canonical nucleotides are exchanged for a modified compound were also reported for duplex DNA with GNA modifications.[99]

The measured CD spectra give also more insights into the binding of the Cu^{2+} ion to the modified G-quadruplex. Upon addition of $CuSO_4$, the overall shape of the CD profiles do not change, but the signals gain in intensity, both the positive one at 293 nm and the negative at 263 nm. This suggests an increase in G-quadruplex concentration and is in agreement with the increase in thermodynamic stability observed in the UV-VIS based thermal denaturation experiments. Furthermore, these results also demonstrate that the strand orientation, i. e. the G-quadruplex topology, is not altered by the metal ion complexation. The observed changes in the CD spectrum are quite fast, the process

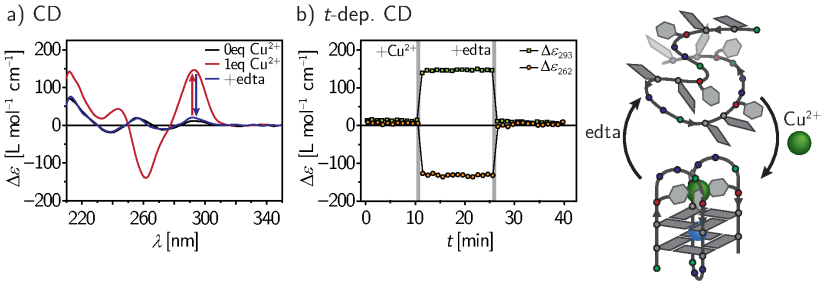


Figure 7.8.: a) CD spectra and b) corresponding time-dependent CD signals measured at 25 °C for the G-quadruplex htel22-$\mathbf{L^2}_4$b with ligand (*S*)-$\mathbf{L^2}$ upon addition of $CuSO_4$ and Na_2H_2EDTA (edta). 10 mM LiCaco pH 7.3, 100 mM KCl, 4.0 µM DNA, 0 or 4.0 µM $CuSO_4$, 0 or 4.8 µM edta.

is completed around ten minutes after $CuSO_4$ addition, supporting the assumption that the Cu^{2+} ion binds to the four pyridines of one unimolecular G-quadruplex and does not incur formation of higher order aggregates. To further prove this, the Cu^{2+} complexation was repeated at 25 °C, at which the metal-free G-quadruplex is not stable. Immediately after addition, the Cu^{2+} ions induce formation of the G-quadruplex structure from the unfolded strands. As seen in the time dependence of the CD signals, the system rapidly jumps into the fully annealed state. In contrast, when the Cu^{2+} ion is removed by addition of Na_2H_2EDTA (edta) and thus formation of the Cu(edta) complex, consequently the G-quadruplex immediately disassembles and the single-strands are released (see figure 7.8).

7.1.4. Molecular dynamics for htel22-$\mathbf{L^2}_4$b

At this point it is beneficial to compare the experimental data and derived structural conclusions with models for the copper-free and copper bound G-quadruplex htel22-$\mathbf{L^2}_4$b. In principle, the same approach used in chapter 6.2 (page 107) for the ligand modified tetramolecular G-quadruplexes could be used. However, the effect of the loop nucleotides

Table 7.2.: Overview of denaturation temperatures T_m for the G-quadruplexes based on the human telomeric repeat.

G-quadruplex	$\mathbf{L^2}$	salt	$\mathbf{CuSO_4}$ eq.	T_m [°C]	ΔT_m [°C]
htel22	*S*	NaCl	0	54.7	
			1	54.2	–0.5
		KCl	0	63.9	
			1	63.1	–0.8
htel22-$\mathbf{L^2}_4$a	*S*	NaCl	0	47.4	
			1	48.5	+1.1
		KCl	0	62.1	
			1	64.1	+2.0
htel22-$\mathbf{L^2}_4$b	*R*	KCl	0	18.1	
			1	33.0	+14.9
	S	KCl	0	12.5	
			1	35.5	+23.0

on the folding would be quite difficult to take into account. Consequently a different approach, namely molecular dynamics simulations, was implemented for the ligand modified G-quadruplex htel22-$\mathbf{L^2}_4$b. As shortly described in the introductory chapter, molecular dynamics simulations enable the construction of molecular models of biomolecules and the investigation of their time-dependant conformational freedom under the imposed temperature, pressure, and salt concentration.

Classical molecular dynamics simulations require a force field, which contains all the bonded and non-bonded interactions between atoms, and an initial starting structure. A typical force field used for nucleic acids assemblies is the AMBER *parmbsc0* force field, which is known to adequately describe the DNA secondary structure of G-quadruplexes. [284,285]

A problem arising at this point is the covalent linkage of ligand $\mathbf{L^2}$ to the rest of the G-quadruplex. All the parameters concerning this modification, including atom charges, force constants, bond lengths, angles, and torsions, are of course absent in the available AMBER

force field, so they have to be derived by other means. To maintain the integrity of the new parameters with the existing ones, the same calculation schemes were used. For an overview of the used parameters as well as a thorough step-to-step description of the necessary calculations, see the molecular dynamics section 9.1.10 in the experimental section and the section A.1 in the appendix (page 195), respectively.

In short, geometry optimised structures of the ligand $\mathbf{L^2}$ and the fragment $Cu^{2+}(\mathbf{L^2})_4$ were used to obtain structural parameters like bond lengths, angles, and torsions. Restricted electrostatic potential (RESP) fitting yielded atom point charges.[286–288] For the respective force constants, either analogues values already implemented in the AMBER *parmbsc0* force field were used, or approximated using literature values in case of parameters involving the Cu^{2+} ion.[289,290] This extended force field could then be used in all following MD simulations. It should be noted that the labile nature of the Cu^{2+}–*N* dative bonds of the $Cu^{2+}(\text{pyridine})_4$ moiety could not be simulated adequately and instead the bonds had to be modelled by inert covalent linkages between these atoms. This allowed to constrain the metal coordination environment to be square-planar, but prevented dynamics regarding dative bond breaking and formation.[195,196]

The initial models of htel22-$\mathbf{L^2}_4$b for the MD simulations were constructed from the NMR-derived structure of the native htel22 sequence (PDB entry 143D, first model).[266] One of the G-quartets was manually deleted using the software Chimera[291,292] and replaced either by four instances of the ligand $\mathbf{L^2}$, or by the copper–ligand complex $Cu^{2+}(\mathbf{L^2})_4$. Within the MD program Gromacs 5.04,[293–298] these rough models were placed in a solvent box with explicit water molecules and potassium cations for charge neutralisation, and then subjected to several rounds of energy minimisation. This ensured that the whole solvent–solute system is in equilibrium, with correct temperature and pressure, and no unusual structural features, like too close contacts between solvent molecules and the G-quadruplex, or unfavourable bond lengths or angles within the G-quadruplex structure. Based on these models, 20 ns MD production run simulations were performed.

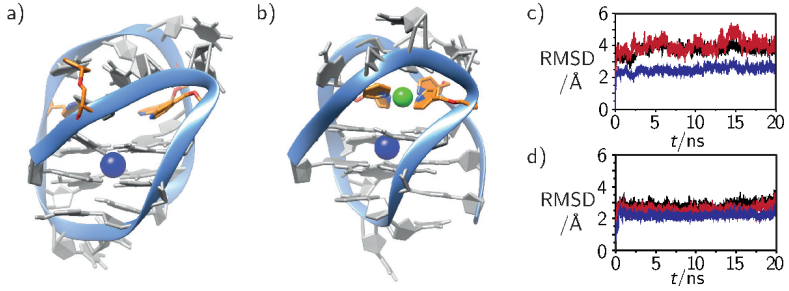


Figure 7.9.: Molecular models for the htel22-$\mathbf{L^2}_4$b G-quadruplex with ligand (*S*)-$\mathbf{L^2}$ derived from molecular dynamics simulations. htel22-$\mathbf{L^2}_4$b a) without and b) with bound Cu^{2+} ion; the phosphate backbone is shown as a blue ribbon, K^+ and Cu^{2+} are represented as blue and green spheres, respectively; the ligand $\mathbf{L^2}$ is highlighted in orange; hydrogen atoms have been omitted for clarity. c) and d) rmsd time traces for the same models, respectively; black curve: non-hydrogen nucleic acid atoms, blue: guanosine non-hydrogen, red: ligand non-hydrogen atoms.

Normally, MD simulations are performed for unmodified nucleic acid structures, so the modifications implemented here represent a significant deviation from the standard procedures. However, the general trends derived from the spectroscopy experimental data, as seen above, are supported. The integrity of the G-quadruplex structure and the conformational freedom of the different parts, e. g. the ligands, during the time-course of the MD simulation can best be evaluated by the root mean square deviation (rmsd) values, which give the average deviation of atom positions from those in a reference structure. Here, the energy minimised starting topology was used as the reference. Figure 7.9 shows the G-quadruplex topologies of htel22-$\mathbf{L^2}_4$b ((*S*)-$\mathbf{L^2}$) with and without bound Cu^{2+}, extracted from the last frames of the respective MD trajectories. For htel22-$\mathbf{L^2}_4$b without copper, the rmsd values show a high flexibility, both for atom positions of the whole G-quadruplex (mean rmsd 3.8 Å) and those of only the ligand atoms (rmsd 4.0 Å, see figure 7.9c, black and red curve, respectively). Interestingly, the rmsd

values of only the guanosine atoms (figure 7.9c blue curve) show a much lower deviation, with average values around 2.5 Å. This observation demonstrates the rigidity of the G-quadruplex core, held together by the hydrogen bonding pattern and the π–π stacking of the G-quartets. In contrast, no preferential conformation is observed through the whole MD trajectory for the ligand $\mathbf{L^2}$ parts, which can freely move around, partially flipping between pointing in- and outward of the G-quadruplex core (see figure 7.10).

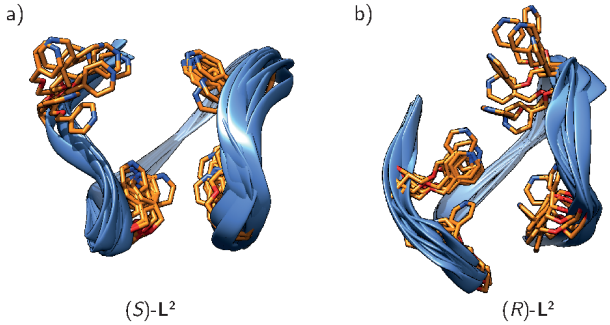


Figure 7.10.: Overlay of molecular models for the htel22-$\mathbf{L^2}_4$b G-quadruplex with ligand a) (*S*)-$\mathbf{L^2}$ and b) (*R*)-$\mathbf{L^2}$ derived from the respective molecular dynamics simulation trajectories. The phosphate backbone is shown as a blue ribbon; the ligand $\mathbf{L^2}$ is highlighted in orange; all other atoms as well as ligand hydrogen atoms have been omitted for clarity.

The whole situation changes, when the same MD simulation is performed with the copper bound G-quadruplex. The overall rmsd values drop distinctly to an average of 2.8 Å, an effect which can be attributed to lower spatial fluctuations of the ligands (2.6 Å) and also the guanosines (2.2 Å). The ligands are now held in place by the Cu^{2+} ion, greatly lowering their conformational freedom and simultaneously, the G-quartets of the G-quadruplex stem become more rigid, too. Of course, the covalent bonds between the copper ion and the pyridine nitrogen atoms ensure the binding of the ligand, so any temporary dative bond breaking or reformation cannot be accounted for. Still,

the MD simulations are in accordance with the UV-VIS based melting experiments and the CD spectra for htel22-$\mathbf{L^2}_4$b, showing that the overall G-quadruplex shape is unchanged when the Cu^{2+} binds to the pyridine ligands, with no topological rearrangement of the G-quadruplex core.

It should be noted that the G-quartets and the guanines in both MD models do not maintain a perfect co-planarity with respect to each other during the time course of the MD simulations, possibly reflecting the low thermodynamic stability of the assemblies as demonstrated by their UV-VIS derived melting temperatures.

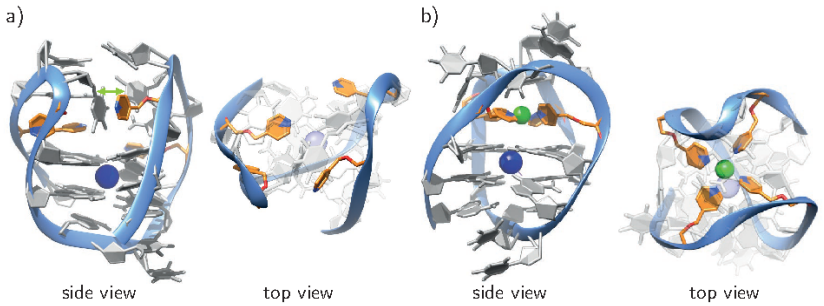


Figure 7.11.: Molecular models for the htel22-$\mathbf{L^2}_4$b G-quadruplex with ligand and (*R*)-$\mathbf{L^2}$ derived from molecular dynamics simulations. htel22-$\mathbf{L^2}_4$b a) without and b) with bound Cu^{2+} ion; the phosphate backbone is shown as a blue ribbon, K^+ and Cu^{2+} are represented as blue and green spheres, respectively; the ligand $\mathbf{L^2}$ is highlighted in orange; hydrogen atoms have been omitted for clarity.

The same MD simulations were performed on the G-quadruplex with the (*R*) enantiomer of ligand $\mathbf{L^2}$ and showed nearly identical results compared to htel22-$\mathbf{L^2}_4$b with (*S*)-$\mathbf{L^2}$. The observed rmsd values are similar, both for the Cu^{2+} bound and unbound G-quadruplex, suggesting that the variation of the ligand's configuration does not lead to significant changes in the topologies. Interestingly, a pyridine unit of one ligand can be seen in coplanarity and a close proximity of 3.3 Å to one of the loop thymine aromatic rings, suggesting the formation of an attractive

π–π interaction. This structural feature is stable through the whole MD trajectory and was not observed for the G-quadruplex with ligand (*S*)-$\mathbf{L^2}$. Although it can be just a coincidence that this interaction is observed in the one case and not in the other, it explains the higher thermodynamic stability of the G-quadruplex with (*R*)-$\mathbf{L^2}$ as compared to that with (*S*)-$\mathbf{L^2}$.

As mentioned earlier in chapter 6.3 on page 111, the chirality of the copper complex may influence the stability of the whole G-quadruplex. So probably, for htel22-$\mathbf{L^2}_4$b with (*S*)-$\mathbf{L^2}$ the copper complex fits better with the pyridines into the G-quadruplex, than the one with (*R*)-$\mathbf{L^2}$.

In this section it was shown that the four-fold incorporation of the ligand $\mathbf{L^2}$ into a DNA sequence based on the human telomeric repeat enables the formation of unimolecular G-quadruplexes, capable of binding a Cu^{2+} ion. The copper complexation is not only fast, and reversible but can also be used to trigger the G-quadruplex annealing process. The substantial experimental data based on UV-VIS thermal denaturation experiments and CD spectroscopy could be verified by implementation of the ligand and copper–ligand complex modification into MD simulations on the G-quadruplex models.

7.2. The metal base-quartet in the *tetrahymena* telomere sequence

The previous chapter has shown that the design of ligand $\mathbf{L^2}$ is very suitable for the metal base-quartet motif and versatile enough to fit into even unimolecular G-quadruplex assemblies, while care has to be taken concerning the loop lengths. All G-quadruplex sequences investigated so far were designed to easily accommodate the copper–pyridine complex, that is, existing structural data were used to fit the metal–ligand environment inside. The next obvious step is to not only use the Cu^{2+} ion merely to stabilise the already most stable topology, but to use the unique binding motif to steer the assembly into a different topology. By this, the copper ion exhibits not only a stabilising role but also a functional one.

Table 7.3.: Sequences for the G-quadruplexes based on the *tetrahymena* telomeric repeat.

DNA	**sequence 5′→3′**
ttel24	d[TT GGG GTTG GGG TTG GGG TT GGG G]
ttel24-T_4	d[TT GGG TTTT GGG TTG GGT TT TGG G]
ttel24-$\mathbf{L^2}_4$	d[TT GGG $\mathbf{L^2}$TT$\mathbf{L^2}$ GGG TTG GG$\mathbf{L^2}$ TT $\mathbf{L^2}$GG G]

As the basis for this experimental setup, the *tetrahymena* telomeric repeat d[TTGGGG] was used, which differs from the human telomeric repeat only in the exchange of a loop adenine against a guanine nucleotide. The sequence d[$(TTGGGG)_4$] (ttel24) was already investigated and is known to adopt a (3+1) hybrid, unimolecular G-quadruplex structure with three stacked G-quartets in Na^+ buffer solution.[299] No structural data derived from NMR experiments or X-ray analyses are available, however, for K^+ containing samples. Substitution of four nucleotides for the ligand (S)-L^2 should result in a G-quadruplex assembly with a compromise between loop flexibility, that is loop length, and G-quadruplex stability, as even with the ligands, formation of three G-quartets should be possible. A difficulty initially overlooked is the positions of the ligand modifications in the ttel24 sequence. In the sodium form of the ttel24 G-quadruplex, some of the guanines do not take part in the G-quartet formation, resulting in three loops with different composition, namely d[GTTG], d[TTG], and d[TT]. The chosen modification positions are 6, 9, 18, and 21, which are two guanines from the loops and two from the G-quartets. As a consequence, the G-quadruplex cannot fold into the wild-type topology any more. For comparability, another reference strand ttel24-T_4 was therefore synthesised, which has thymines at the positions of the ligand $\mathbf{L^2}$ (see table 7.3).

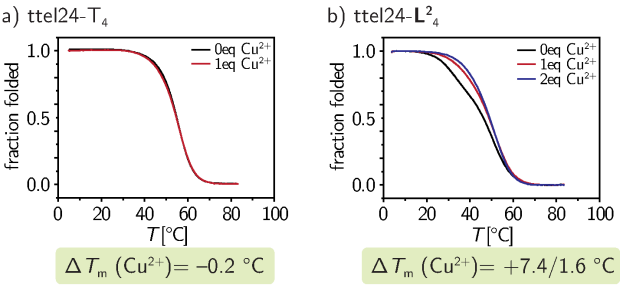


Figure 7.12.: Thermal denaturation curves (monitored at 295 nm) converted to the fraction folded values for the G-quadruplexes a) ttel24-T_4, b) ttel24-$\mathbf{L^2}_4$ upon addition of $CuSO_4$. 10 mM LiCaco pH 7.3, 100 mM KCl, 4.0 µM DNA, 0, 4.0, 8.0 µM $CuSO_4$.

7.2.1. Folding monitored by UV-VIS and CD

First, the thermal stabilities and CD profiles of the ttel24-T_4 G-quadruplex were compared to those of the modified sequence ttel24-$\mathbf{L^2}_4$. The UV-VIS based melting curves of the ttel24- T_4 sequence exhibit one single transition in KCl buffer at $T_m = 54.8$ °C, with no observable effect on the melting temperature when $CuSO_4$ is added. For the denaturation in Na^+ containing solution, the melting temperature was observed to be quite low, with $T_m = 28.2$ °C, so all further experiments were conducted using KCl as the salt additive. It should be noted that for the strand ttel24-T_4, one DNA sample batch did not show a single transition but a broad denaturation profile with two transitions. The analytical RP-HPLC traces and the ESI(–) mass spectra were identical for the different stock solutions, so it has to be assumed that the observed differences arose from the sample preparation, like variations in buffer pH, or an unidentified contamination.

A denaturation profile with a double transition is observed for the modified sequence ttel24-$\mathbf{L^2}_4$, the overall melting transition is, however, in the same range as the unmodified sequence ttel24-T_4. This observation suggests a more complicated melting behaviour with a deviation from the simple two-state equilibrium between unfolded and

folded G-quadruplex.[300] Moreover it also indicates potential formation of higher-order structures or the presence of multiple similar G-quadruplex conformers.[277] When the melting profile is fitted with a double-sigmoidal function, denaturation temperatures of T_m = 32.7 °C and 50.8 °C for ttel24-$\mathbf{L^2}_4$ can be calculated. When the denaturation experiments are repeated in the presence of $CuSO_4$, the melting profiles change partially. While the high temperature transition is mainly unaffected, the low temperature transition shifts to higher values. Altogether, the profile now resembles more that of one single transition, rather than two separate ones. The available experimental data do not allow to determine if the high temperature transition corresponds to a separate G-quadruplex topology to which the Cu^{2+} ion cannot bind, or if it belongs to the same topology which is also responsible for the lower temperature transition.

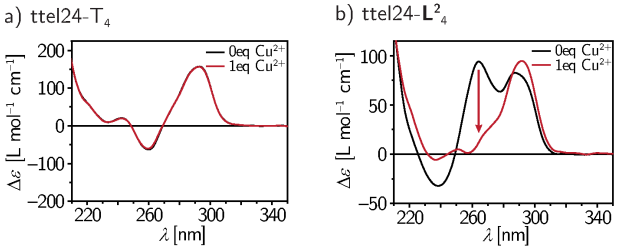


Figure 7.13.: CD spectra for the G-quadruplexes a) ttel24-T_4 and b) ttel24-$\mathbf{L^2}_4$ upon addition of $CuSO_4$. 10 mM LiCaco pH 7.3, 100 mM KCl, 4.0 µM DNA, 0 or 4.0 µM $CuSO_4$.

Further insights into the nature of the Cu^{2+} binding could be derived from the CD spectra. In case of the unmodified ttel24-T_4 sequence, the CD spectrum exhibits an intense positive band at 293 nm and a less intense negative one at 260 nm. The overall shape of the CD signal resembles that reported for the sequence ttel22 d[GGGG(TTGGGG)$_3$], measured in Na^+ containing solution, and that for the (3+1) assembly of the human telomeric sequence discussed above. In contrast, the CD profile of ttel24-$\mathbf{L^2}_4$ differs considerably, with two intense positive bands at 264 nm and 288 nm, respectively, and a smaller negative signal at

238 nm. This suggests a different folding topology for the modified strand and also potentially the presence of a mixture of different G-quadruplex conformers. Despite the differences between the CD spectra, both can be classified as 'type 2', according to Karsisiotis et al.[279] The profile is typically found for G-quadruplexes with three G-quartets, which stack in both a homopolar and a heteropolar fashion.

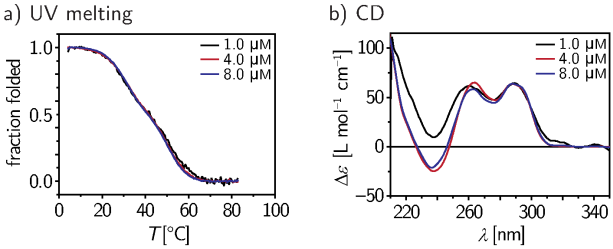


Figure 7.14.: a) Thermal denaturation curves (monitored at 295 nm) converted to the fraction folded values and b) CD spectra for the G-quadruplex ttel24-$\mathbf{L^2}_4$ at three different strand concentrations. 10 mM LiCaco pH 7.3, 100 mM KCl, 1.0, 4.0, or 8.0 µM DNA.

To elucidate if higher order structures are present in the G-quadruplex sample of the modified sequence ttel24-$\mathbf{L^2}_4$, concentration dependant UV melting and CD experiments were performed. For all three concentrations, namely 1.0, 4.0, and 8.0 µM, the denaturation profiles were superimposable and the CD spectra did also not change significantly, so the presence of higher-order structures like G-wires or stacked G-quadruplexes is unlikely. The co-existence of a mixture of different monomeric, unimolecular G-quadruplex topologies cannot be ruled out by these experiments, however.

Apart from the ttel24 sequence, two other NMR-derived solution state structures of G-quadruplexes with similar sequences compared to ttel24-T_4 have been deposited in the PDB data base, but are not yet published elsewhere. The two sequences are d[$TG_3T_3G_3TTG_3T_3G_3$] (PDB entry 2MFU)[301] and d[$G_3T_4G_3TG_3T_4G_3$] (PDB 2MFT)[302], respectively, and assemble into (2+2) antiparallel G-quadruplexes in

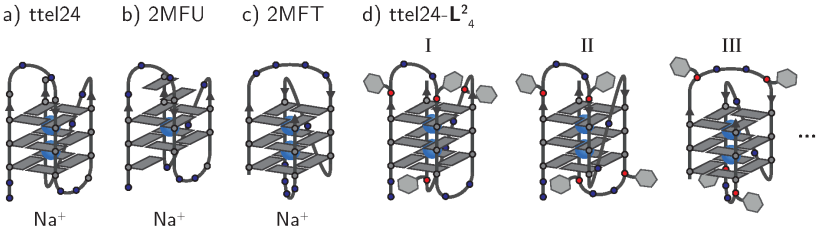


Figure 7.15.: Schematic G-quadruplex models for the sequences a) ttel24 in Na^+ solution;[299] b) d[$TG_3T_3G_3TTG_3T_3G_3$] in Na^+ solution;[301] c) d[$G_3T_4G_3TG_3T_4G_3$] in Na^+ solution;[302] d) I–III: examples for potential folding topologies of ttel24-$\mathbf{L^2}_4$ in K^+ solution.

Na^+ solution (see figure 7.15b and c). Although this can by no means be generalised, it seems that for these sequences, the long dT_4 sequence favours a diagonal loop, dT_3 lateral loops, and dT_2 propeller-type loops. That the dT_2 connection mostly forms propeller-type loops was also reported by Cheatham III and coworkers.[30] Unfortunately, no spectroscopic data like CD spectra are available for the 2MFU or 2MFT G-quadruplexes, so it cannot be verified if the antiparallel (2+2) strand arrangement still gives rise to a 'type 2' CD spectrum as observed for ttel24-$\mathbf{L^2}_4$. The glycosidic bond orientations in the reported structures reveal, however, that the G-quartets do indeed show both homo- and heteropolarity stacking, so a CD spectrum similar to that found for the modified sequence seems very likely. Based on these structures, models for the modified sequence ttel24-$\mathbf{L^2}_4$ can be derived (see figure 7.15d I–III). All three potential topologies have one common structural feature: The ligand $\mathbf{L^2}$ modifications are not located on the same side of the G-quadruplex, but on opposite ones, meaning that the square-planar $Cu^{2+}(\mathbf{L^2})_4$ complex cannot be formed without changes in the strand orientation.

7.2.2. Folding monitored by ^{1}H-NMR spectroscopy

To get insights into the polymorphic nature of the ttel24-$\mathbf{L^2}_4$ sequence, ^{1}H-NMR measurements were conducted, with the ttel24-T_4 sample as a reference. In principle, NMR experiments can be used to determine the exact folding topology of a G-quadruplex forming sequence. Unambiguous assignments of all the exchangeable and non-exchangeable proton signals have to be achieved by low enrichment with ^{15}N, separately for every guanine.[303] This requires, however, the presence of clearly distinguishable signals, so signal assignment is often impaired by the presence of a conformer mixture. Moreover, since the oligonucleotide strand ttel24-$\mathbf{L^2}_4$ contains the ligand modifications, this procedure of enrichment would be way too laborious to be feasible. Nevertheless, NMR spectroscopy can be used to identify the coexistence of G-quadruplex conformers and the number of stacked G-quartets. In the ^{1}H-NMR spectrum, the region in the chemical shift range δ = 10–12 ppm is characteristic of G-quartet formation, as here the imino resonances of the guanines' $H1$ protons appear.[303] In contrast, those involved in Watson-Crick base-pairing give rise to resonances at 13–14 ppm.[303]

For the G-quadruplex sample of ttel24-T_4, the ^{1}H-NMR spectrum shows eight sharp and one more broader imino proton signal between 11.2–12.0 ppm (see figure 7.16a). The sharp and well defined signals strongly suggest the existence of only one stable G-quadruplex topology with no contribution of other assemblies. For a unimolecular G-quadruplex with three stacked G-quartets, 12 independent guanine imino proton signals would be expected, compared to 8 for the stacking of two G-quartets. Integration of the observed signals indicates that a total of approximately 12 protons give rise to the detected 9 signals, suggesting the presence of overlapping proton signals. This fits to the reported folding with three G-quartets of the very similar 2MFT sequence (figure 7.15c), which only has one thymine less in the middle loop and lacks the 5′-TT overhang, although the reported structure was measured in Na^+ instead of K^+ containing solution. Comparison of the proton signal assignments for the two-tetrad 2MFU and the three-tetrad 2MFT G-quadruplexes deposited with the respective structures in the PDB

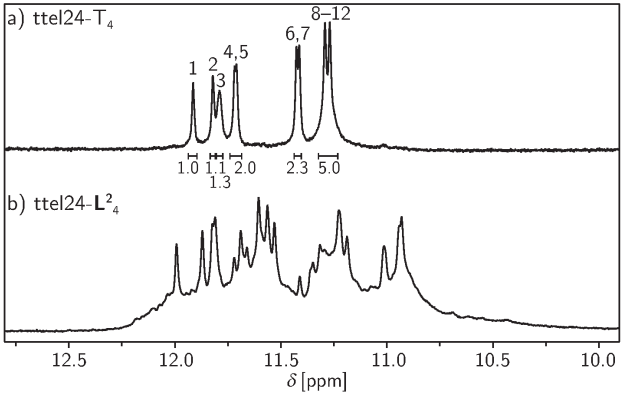


Figure 7.16.: 500.25 MHz ^{1}H-NMR spectra of G-quadruplexes a) ttel24-T_4 and b) ttel24-$\mathbf{L^2}_4$. H_2O/D_2O 10:1, 10 mM LiCaco pH 7.3, 100 mM KCl, 0.5 µM (ttel24-T_4) or 0.5 µM (ttel24-$\mathbf{L^2}_4$) DNA, 298 K. For a) the consecutive proton numbering is just based on the observed integral values of the respective signal.

database further support the presence of three stacked G-quartets in ttel24-T_4. Guanine imino protons not involved in G-quartet formation give rise to ^{1}H-NMR signals shifted to higher field $\delta \leq 11$ ppm, compared to those inside the G-quartet hydrogen bonding pattern with values $\delta =$ 11–12.5 ppm. This holds true even for the observed G-triplet in the structure of the human telomeric G-quadruplex d[$(GGGTTA)_3GGGT$], which shows chemical shift values of the involved guanine imino protons at $\delta \sim 10.5$ ppm.[275] No signals below $\delta = 11.2$ ppm are observed for ttel24-$\mathbf{T}_4$, strongly suggesting that all guanine imino protons are involved in G-quartet formation.

In stark contrast, the ^{1}H-NMR spectrum of ttel24-$\mathbf{L^2}_4$ exhibits quite broad signals, and numerous imino proton signals between 10.5 and 12.2 ppm. Although some of the signals are more sharp and clearly distinguishable from the rest of the less intense and strongly overlapping signals, the occurrence of a mixture of topologies has to be assumed for ttel24-$\mathbf{L^2}_4$. No ^{1}H-NMR spectra were recorded for the copper-bound form of ttel24-$\mathbf{L^2}_4$, as the paramagnetism of the Cu^{2+} ion will lead to very strong signal broadening and thus impairing structural assignments.

7.2.3. Cu^{2+} induced topology switching

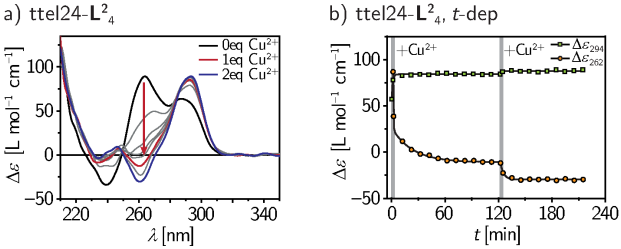


Figure 7.17.: Time-dependant a) CD spectra and b) CD extinction coefficient values at 262 and 294 nm for the G-quadruplexes ttel24-$\mathbf{L^2}_4$ upon addition of $CuSO_4$ at 30 °C. 10 mM LiCaco pH 7.3, 100 mM KCl, 4.0 µM DNA, 0, 4 or 8.0 µM $CuSO_4$.

Coming back to the CD spectroscopic measurements, the effect of the Cu^{2+} ion was studied in more detail. When Cu^{2+} is added to the G-quadruplex solution with ttel24-$\mathbf{L^2}_4$, a remarkable effect on the CD profile is observed. Very slowly, the band at 263 nm and also the one at 240 nm starts to disappear, whereas the band at 293 nm gains intensity. This process is very slow and equilibrium is not reached within 5 h at 4 °C, suggesting a major structural rearrangement of the oligonucleotide strand. This assumption is also in accordance with the fast kinetics of Cu^{2+} binding observed for the htel22-$\mathbf{L^2}_4$b G-quadruplex, which required no topological change of the G-quadruplex core. The same CD based experiment was repeating at a higher temperature to increase the reaction rate of the observed process. At 30 °C, equilibrium is reached within 120 min, at which point the CD profile resembles that typical of a G-quadruplex with antiparallel strands, as seen for the strand htel22-$\mathbf{L^2}_4$b. Addition of another equivalent of $CuSO_4$ shifts the equilibrium further to the 'type 3' CD profile.[279] The same equilibrium profile is obtained when directly 2 eq. of Cu^{2+} are added to the G-quadruplex solution.

From all these observations it can be inferred that the modified sequence ttel24-$\mathbf{L^2}_4$ forms in its initial state a G-quadruplex topol-

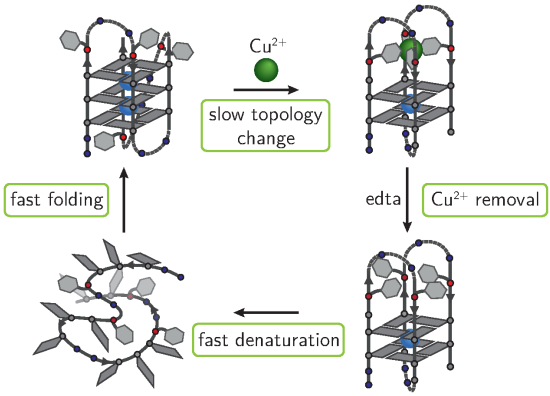


Figure 7.18.: Proposed G-quadruplex topology change induced by formation of the $Cu^{2+}(\mathbf{L^2})_4$ coordination complex within ttel24-$\mathbf{L^2}_4$ and the fast reverse reaction via the unfolded strand triggered by removal of the Cu^{2+} ion with edta.

ogy, unable to accommodate the copper–pyridine coordination complex. Upon Cu^{2+} addition, a change in topology from a mixture of conformers, likely consisting mainly of (3+1) hybrid structures, to a stable assembly with antiparallel strand orientations is invoked. The depicted antiparallel model comprises only two stacked G-quartets, although a topology with three would also be possible. The G-quadruplex topology with two G-quartets is more likely, as first, only a small increase in thermodynamic stability was observed for the ttel24-$\mathbf{L^2}_4$ sequence upon Cu^{2+} addition, suggesting that the stabilising effect of the copper ion is compensated by the loss of one G-quartet, and second, the assembly with three G-tetrads would comprise only shorter and more unstable lateral loops with only two nucleotides and no longer diagonal loops with four nucleotides.[30]

As seen above, the UV absorption based thermal denaturation experiments did not reveal much about the stabilising effect of the Cu^{2+} ion, so they were repeated by measuring the temperature dependant CD signals, both for the copper-bound and copper-free G-quadruplex ttel24-$\mathbf{L^2}_4$. In the absence of Cu^{2+}, the CD spectra reveal the presence

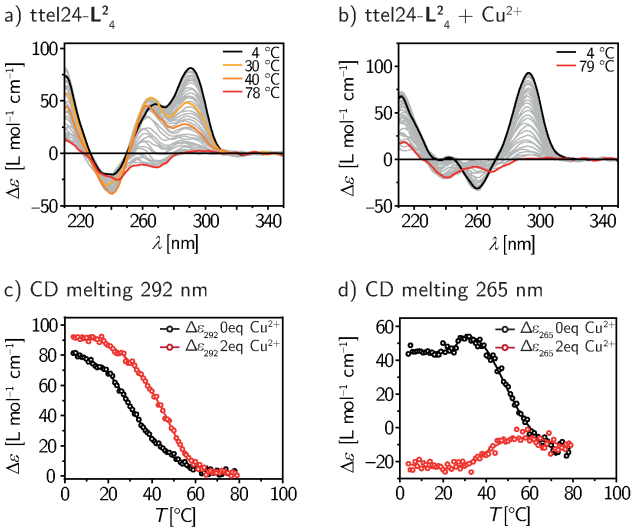


Figure 7.19.: Temperature dependant CD spectra, a) without and b) with Cu^{2+}, and denaturation profiles (c and d) for the G-quadruplex ttel24-$\mathbf{L^2}_4$. 10 mM LiCaco pH 7.3, 100 mM KCl, 4.0 µM DNA, 0 or 8.0 µM $CuSO_4$. The solid lines in the denaturation profiles are only a guide to the eye.

of at least two denaturation intermediates, best seen at around 30 °C and 40 °C (see figure 7.19a). The corresponding melting profile followed at 265 nm shows an increase up to approximately 35 °C, before the signal follows a sigmoidal-shaped decrease in intensity (figure 7.19d black curve). For the copper-bound G-quadruplex, the situation is completely different. The denaturation profile measured at 292 nm clearly shows a distinct increase in melting temperatures from T_m (0 eq. Cu^{2+}) = 28/33 °C to T_m (2 eq. Cu^{2+}) = 35/50 °C (figure 7.19c). These values are similar to those obtained in the UV-VIS based thermal denaturation experiments. Concomitantly, the melting profile followed at 265 nm reveals only one transition. These results further support the occurrence of a simple two-state equilibrium between folded and unfolded G-quadruplex in the presence of Cu^{2+} ions. Furthermore, the absence of intermediate structures is in accordance with a G-quadruplex topology comprising only two G-quartets.

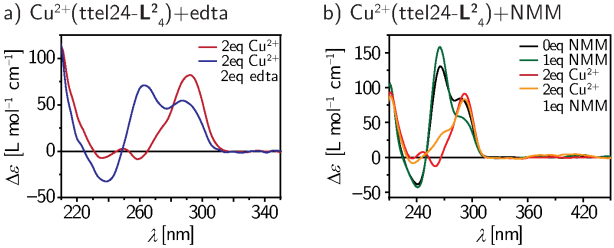


Figure 7.20.: CD spectra measured at 30 °C for the copper-bound or copper-free G-quadruplex ttel24-$\mathbf{L^2}_4$ upon addition of a) 2eq. Na_2H_2EDTA (edta), or b) *N*-methyl mesoporphyrin IX (NMM). 10 mM LiCaco pH 7.3, 100 mM KCl, 4.0 µM DNA, 0 or 8.0 µM $CuSO_4$, a) 0 or 8 µM Na_2H_3EDTA, b) 0 or 4.3 µM NMM;

Like for the G-quadruplexes htel22-$\mathbf{L^2}_4$b, the effect of the Cu^{2+} ion can be reversed by addition of a chelate agent like edta. Interestingly, the reverse topological change from the antiparallel assembly back to the mixed hybrid structures is significantly faster than for the copper induced changes. When edta is added to the equilibrated Cu^{2+} bound G-

quadruplex ttel24-$\mathbf{L^2}_4$, the initial CD profile corresponding to the hybrid topology mixture is restored within 10 min minutes (see figure 7.20a). This observation indicates that upon removal of the stabilising Cu^{2+} ion from the ligand $\mathbf{L^2}$ complex, the oligonucleotide strand immediately unfolds completely and then refolds. The fast folding kinetics of a unimolecular G-quadruplex from the single strands was also observed for the unmodified G-quadruplex htel22 (data not shown). Thus, the integral part of the Cu^{2+} ion in stabilising the otherwise unfavourable antiparallel G-quadruplex topology is highlighted.

To get insights into the topology mixture present for the G-quadruplex ttel24-$\mathbf{L^2}_4$ in the absence of Cu^{2+} ions, the binding of the porphyrin NMM was investigated here, too. As already shown above for the htel22 sequences, NMM discriminates between parallel/hybrid, and antiparallel G-quadruplex assemblies. When ttel24-$\mathbf{L^2}_4$ is incubated with NMM shortly at 30 °C, the CD spectra exhibit the same changes as observed for htel22-$\mathbf{L^2}_4$a, with an increase of the CD signals at 265 nm and a concomitant loss of signal at 293 nm (see figure 7.20b). Apparently, the sample does indeed contain parallel or hybrid structures as proposed above. Likewise, the presence of already antiparallel topologies becomes unlikely, as these should be unaffected by the presence of NMM and should still give rise to the characteristic CD profile with a positive peak at around 295 nm. Although NMM was already shown to bind Cu^{2+} ions, the same experiment was repeated for the copper-bound G-quadruplex. Here, addition of NMM immediately leads to an increase of the signal at 265 nm, indicating the partial removal of the Cu^{2+} ion from the G-quadruplex and the subsequent unfolding and refolding analogues to the effect of edta. Due to this ability of NMM to bind Cu^{2+}, no information regarding an induced topology change can be derived.

In summary, this chapter highlighted the power of the metal base-tetrad, based on the complexation of Cu^{2+} to four units of ligand $\mathbf{L^2}$, to not only stabilise a preorganised G-quadruplex structure, but also to induce the topology switching from a mixture of conformers to a stable G-quadruplex with a different topology. Indications for the nature of

the involved topologies came from UV-VIS based thermal denaturation experiments and thorough CD spectroscopy analysis. Further information came from basic ^{1}H-NMR experiments were performed on the copper-free G-quadruplexes. Furthermore, the switching process can be reversed by removing the Cu^{2+} with the chelating agent edta, and can be steered into another topology by binding to the porphyrin NMM.

8. Outlook

In this thesis it could be shown that the metal base-tetrad is a versatile and easy to handle modification of G-quadruplexes, enabling to stabilise, trigger the formation, or switch the topology of tetramolecular and unimolecular G-quadruplexes. Based on these results, future progress can be divided into two parts: First, some of the phenomena observed for the G-quadruplex systems and the related experimental data discussed in this thesis need further investigation. Secondly, the metal base-tetrad motif opens the field of G-quadruplex nanotechnology to other novel applications with regard to structure and functionality.

Regarding open questions which still need to be addressed in the near future, the first concerns the linker length variation in the ligand-appended tetramolecular G-quadruplexes $[\mathbf{L^{1,1b-d}}\mathrm{d(G_4)}]_4$ (chapter 6.2). Here, the thermal denaturation experiments would best be complemented by additional molecular dynamics simulations on both the copper-free and -bound G-quadruplexes. This would enable to get further insights into the observed different thermal stabilities and the arrangement of the respective ligands above the top G-quartet. The ligand $\mathbf{L^{1d}}$ with the shortest linker has a slightly different attachment to the pyridine (-CH_2O-) compared to the other investigated ligands $\mathbf{L^{1,1b-d}}$ (-OCH_2-). So, any effect arising from this difference should be evaluated, e. g. by comparison to a ligand with identical attachment as $\mathbf{L^{1d}}$ but with a longer linker length. Moreover, the peculiar effect of Cu^{2+} ions on the CD profile of the $[\mathbf{L^{1d}}\mathrm{d(G_4)}]_4$ G-quadruplex should be analysed in more detail by complementary methods, including gel electrophoresis and atomic force microscopy, in order to evaluate the formation of higher-order structures.

Concerning the G-quadruplexes $[\mathbf{L^2}d(G_{3-5})\mathbf{L^2}dT]_4$ with the metal base-tetrads positioned on opposite sides, it would be interesting to determine the Cu^{2+}–Cu^{2+} distance. If feasible, the metal base-tetrad can indeed be used as a spin label, as proposed in chapter 2. The distance between the copper ions should be highly dependant on the number of G-quartets, as the copper ions are positioned just above the top and below the bottom G-quartet, and because each additional G-quartet in the G-quadruplex stem results in an elongation of the G-quadruplex along the central axis. A suitable method for the determination of distances between paramagnetic centres is EPR spectroscopy, namely the techniques PELDOR and ENDOR. Distance measurements between metal base-pairs in duplex DNA,[304,305] copper–porphyrins attached to DNA,[306] and nitroxide spin labels in a unimolecular G-quadruplex[307] were already reported in the literature. EPR measurements for the $(Cu^{2+})_2[\mathbf{L^2}d(G_{3-5})\mathbf{L^2}dT]_4$ G-quadruplexes are already in progress in collaboration with the group of O. Schiemann at the university of Bonn (Germany). Both the G-quadruplex stem and the copper–pyridine complex are rather rigid, which should give quite narrow distance distributions when compared to freely rotating spin labels attached for example to proteins. In conjunction with molecular dynamics simulations to obtain theoretical Cu^{2+}–Cu^{2+} distances, these experiments could demonstrate the feasibility of the metal base-tetrad modified G-quadruplexes as a molecular ruler and benchmark system for distance determinations.

The synthesis of the reference ligand $\mathbf{L^{2*}}$ was established in this thesis, however the ligand was not incorporated into oligonucleotide strands. Substituting one of the two instances of $\mathbf{L^2}$ with $\mathbf{L^{2*}}$ in the G-quadruplex $[(\mathbf{L^2})_2d(G_4)]_4$ should help to answer the question of the preferred binding sites of the two Cu^{2+} ions. It was unclear if the respective metal ion was coordinated by the pyridine moieties of four oligonucleotide strands, or by two of the pyridines of just two strands. The substitution would prevent binding of the Cu^{2+} ion by the pyridine donors of only two of the four strands.

In general, the molecular dynamics simulations performed in this thesis are still quite basic, with e. g. only approximated force constants on the Cu^{2+} ion, so these still need some improvements. Apart from more better parametrisation of the ligands and the metal ion, the general feasibility of the covalently bound metal ion and ligands could be evaluated by simulating already known structures of DNA assemblies carrying covalently bound ligand modifications and coordinated metal ions. These systems could be for example crystal structures or NMR derived topologies of DNA duplexes modified with metal base-pairs. When this is achieved, free energy calculations can be used to assess the relative thermodynamic stabilities of different artificial DNA constructs, thus complementing the experimental results obtained with UV-VIS based thermal denaturation studies.

The second part of possible future work comprises alterations of the metal base-tetrad or the G-quadruplex stem for use in DNA nanotechnology. The first possibility is to change the nature of the ligand donor functionality, e. g. by substituting the pyridine nitrogen donors for thiols or carboxylates. Keeping in mind that with ligands based on the $\mathbf{L^2}$ glycol backbone, any position in an oligonucleotide sequence is addressable, the next obvious step is to introduce different ligand functionalities into the same sequence. Like with the sequence htel22-$\mathbf{L^2}_4$b, the folding into unimolecular G-quadruplexes would spatially preorganise the different ligands for metal ion coordination. With this, completely new, designed coordination environments for transition metal ions would be obtainable. These constructs could then be used to mimic the active sites of natural metalloenzymes and would also enable the switching of the enzyme activity by control of the re- and denaturation of the G-quadruplex framework. Of course, coordination environments could easily be constructed from more than four donor functionalities, thus offering the possibility to coordinate metal ions that prefer higher coordination numbers.

The second possibility is to use a G-quadruplex modified with a metal coordination site to either control binding of a biomolecule like a protein or to use the metal ion as a sensor for the binding event.

An example is the thrombin binding aptamer, introduced in chapter 4.2.4, which forms a unimolecular G-quadruplex. In the folded form, the aptamer binds to thrombin and thus inhibits its protease activity. As seen for htel22-$\mathbf{L^2}_4$b, Cu^{2+} can efficiently trigger the G-quadruplex formation from the single-strands. So when the metal base-tetrad is introduced into the thrombin binding aptamer, the activity of thrombin can be modulated via the uptake or removal of the Cu^{2+} ion. For this system, work is already in progress.

A third possible application of the metal base-tetrad inside G-quadruplexes would be a combination of the just discussed two, namely to attach the modified G-quadruplex sequence onto a stable platform, like a porous resin or a DNA origami framework. The metal mediated G-quadruplex formation could then be used to switch an anticipated interaction with another biomolecule on or off or to invoke (macroscopic) structural changes due to the different spatial dimensions of the folded G-quadruplex and the unfolded single-strands. Thus, the metal base-tetrad bearing G-quadruplex serves as a spatially controllable platform.

All in all, the metal base-tetrad introduced in this thesis offers a versatile platform for future applications of G-quadruplexes with controlled stability, topology, and thus functionality.

9. Methods and experimental procedures

Part of the methods section and experimental procedures described here have been already published. Reprinted (adapted) with permission from D. M. Engelhard et al.[244] Copyright ©2013 John Wiley & Sons, Inc.

9.1. Methods

9.1.1. NMR spectroscopy

NMR spectra were measured on a *Bruker* Avance III 300 (300.13 MHz, 5 mm BBFO BB-19 F/1 H ATM z-Grad probe, autosampler BACS 120) or *Bruker* Avance III HD 500 (500.25 MHz, 5 mm CryoProbeProdigy BB/1 H-19 F ATM z-Grad). ^{1}H- and ^{13}C-NMR chemical shifts are given in ppm with the residual solvent signal as an internal standard: $CDCl_3$ (7.26, resp. 77.160 ppm), [D_6]dimethylsulfoxide (2.500 ppm, resp. 39.520 ppm), D_2O (4.790 ppm), CD_3CN (1.940 ppm, resp. 118.260 ppm). ^{31}P-NMR chemical shifts are given in ppm relative to 85 % H_3PO_4 (0.000 ppm, external reference, data stored in the NMR spectrometer). Signal multiplicities are composed of the following abbreviations: s (singlet), d (doublet), t (triplet), q (quartet), sept (septett), m (multiplett). $CDCl_3$ was filtered through basic alumina prior to measurement of the phosphoramidites to prevent acid catalysed decomposition.

9.1.2. Mass spectrometry

EI mass spectra (positive mode, ionisation energy 70 eV) were measured on a *Finnigan* MAT 8200, ESI (positive or negative mode) mass spectra

Table 9.1.: DNA oligonucleotide synthesis reagents.

name	reagent	composition
DCA	detritylation	3 % (v/v) dichloroacetic acid in anhydrous toluene
ACT	activator	0.25–0.3 mol L^{-1} BTT in anhydrous MeCN
Cap A	capping A	16 % (v/v) *N*-methyl imidazole in anhydrous THF or 6.5 % (w/v) 4-dimethylaminopyridine in anhydrous THF
Cap B	capping B	2,6-lutidine/acetic anhydride/anhydrous THF 1:1:8 ($v/v/v$)
Oxi	oxidizer	0.05 mol L^{-1} iodine in pyridine/water 9:1 (v/v) or 0.02 mol L^{-1} iodine in THF/pyridine/water 7:2:1 ($v/v/v$)
MeCN		anhydrous acetonitrile

on a *Bruker* HCT Ultra, maXis, or MicrOTOF mass spectrometer. All DNA samples were measured in H_2O/MeOH 10:1 or 9:1, or in H_2O/acetonitrile 6:4.

9.1.3. DNA synthesis

Oligonucleotide strand **htel22** was purchased HPLC-purified from *IBA GmbH* (Goettingen, Germany). All other oligonucleotides were synthesized on a *K&A Laborgeraete GbR* Synthesizer H-8 on a 1 µmol scale using the standard phosphoramidite methods on CPG. All solvents for the DNA synthesis were purchased in anhydrous quality, except for the oxidizer solution, from *Carl Roth*, *Sigma Aldrich* or *Fisher Scientific* and additionally dried over molecular sieves prior to the DNA synthesis for at least 12 h. Reagents were purchased from *Carl Roth* and *Sigma Aldrich*.

The activator BTT was synthesized according to a modified literature procedure (see subsection 9.2.17).[308] Standard 4,4′-Dimethoxytrityl

(DMT) and β-cyanoethyl (CE) protected DNA phosphoramidites (DMT-dT-CEP, DMT-dG(*i*Bu)-CEP, DMT-dA(Bz)-CEP) were purchased from *Sigma Aldrich*. Controlled pore glass (CPG) solid supports (1000 Å, 30 µmol g^{-1}) were either purchased already packed in luer type cartridges from *K&A Laborgeraete GbR* or manually packed (CPG from *Sigma Aldrich*, cartridges from *Biosearch Technologies* or *Link Technologies*). The oligonucleotide synthesis followed the built-in methods of the DNA synthesizer and was only slightly modified. Each individual step of one cycle contained an acetonitrile wash and drying with argon at the end. First, the cartridge was treated three times with DCA. Second, coupling was achieved by mixing the activator with the respective phosphoramidite 1:1 (v/v). The coupling time was ∼0.5 min for standard phosphoramidites and ∼3.5 min for the ligand phosphoramidites. Third, the cartridge was treated with a 1:1 (v/v) mixture of Cap A and B, which was followed, fourth, by the oxidiser solution. Here, an additional washing step with acetonitrile was introduced compared to the standard routine. This cycle was repeated for the required number of steps. After DNA synthesis, the solid supports were removed from the cartridges and then treated with concentrated aqueous ammonia solution at 55 °C over night for cleavage and deprotection. The supernatant solution was filtered and the solid support was washed two times with 200 µL H_2O. The ammonia was removed under reduced pressure using a vacuum concentrator (S-Concentrator BA-VC-300H, *H. Saur*, Reutlingen, Germany). The DNA samples were heated shortly to 55–60 °C and filtered again, prior to purification by RP-HPLC (see below). Subsequently, according to the following protocol, the 5′-OH DMT protecting group was removed and each sample desalted using *Glen Research* Glen-Pak or *Waters* Sep-Pak C18 cartridges: First, the cartridge was conditioned by slowly passing 2x2 mL of acetonitrile, 2 mL MeCN/0.1 mol L^{-1} triethylammonium acetate (TEAA) pH 7, 2x2 mL 0.1 mol L^{-1} TEAA through the cartridge with a syringe. The oligonucleotide was taken up in 2 mL of 0.1 mol L^{-1} TEAA and slowly passed three times through the cartridge. This was followed by a washing step with 2x2 mL 0.1 mol L^{-1} TEAA and a treatment with 2 % (v/v) trifluoroacetic acid. In case of

oligonucleotides with no 5′-DMT protecting group, the acid treatment was omitted. The bound oligonucleotide was washed with a total of 4 mL ultrapure H_2O and then eluted from the cartridge by addition of 2 mL MeCN/H_2O 1:1 (v/v). Evaporation of the solvents and redissolvation in ultrapure water (300–1000 µL) then yielded the pure oligonucleotide.

The concentration of each DNA sample was determined via the absorbance at 260 nm at 25 °C and using revised extinction coefficients for the nucleosides.[234] For all ligands it was assumed that UV absorption at 260 nm derives solely from the aromatic unit. Therefore compounds **3**, **8**, **18**, **13**, **11a**, **11b** were used to determine the respective extinction coefficient. A standard dilution experiment was used, in which the absorbance is measured for several concentrations of the respective ligand (dissolved in water) and the extinction coefficient is calculated from the slope of the linear regression (see figure 9.1).

9.1.4. RP-HPLC

DNA samples were purified DMT-ON by RP-HPLC (*Agilent* 1260 Infinity; columns: semipreparative (10 mm diameter) or analytical (4.6 mm) *Machery Nagel* 250 Nucleodur 100-5 C18ec or *Agilent* Zorbax 300SB-C18). The column oven temperature was set to 60 °C. HPLC grade solvents were purchased from *Sigma Aldrich* and *VWR International*. Solvent A: 0.05 or 0.1 mol L^{-1} TEAA pH 7, B: 70:30 MeCN/0.05 or 0.1 mol L^{-1} TEAA pH 7. A gradient of 0–80 % B in 30 min was used; with a flow rate of 3 mL min^{-1} for semi-preparative and 1 mL min^{-1} for analytical separations. In appendix A.2 the analytical RP-HPLC traces for the synthesized oligonucleotides after RP-HPLC purification, detritylation and desalting are shown. Oligonucleotides [$\mathbf{L^1}d(G_3)$], [$\mathbf{L^{1*}}d(G_3)$], [$\mathbf{L^1}d(G_4)$], and [$\mathbf{L^{1*}}d(G_4)$] were only desalted.

9.1.5. DNA sample preparation

Each DNA sample contained 10 mmol L^{-1} lithium cacodylate buffer (pH 7.3), 100 mmol L^{-1} of a salt (NaCl, $NaClO_4$, KCl, or CsCl). In some cases 10 mmol L^{-1} sodium cacodylate buffer (pH 7.2) was used.

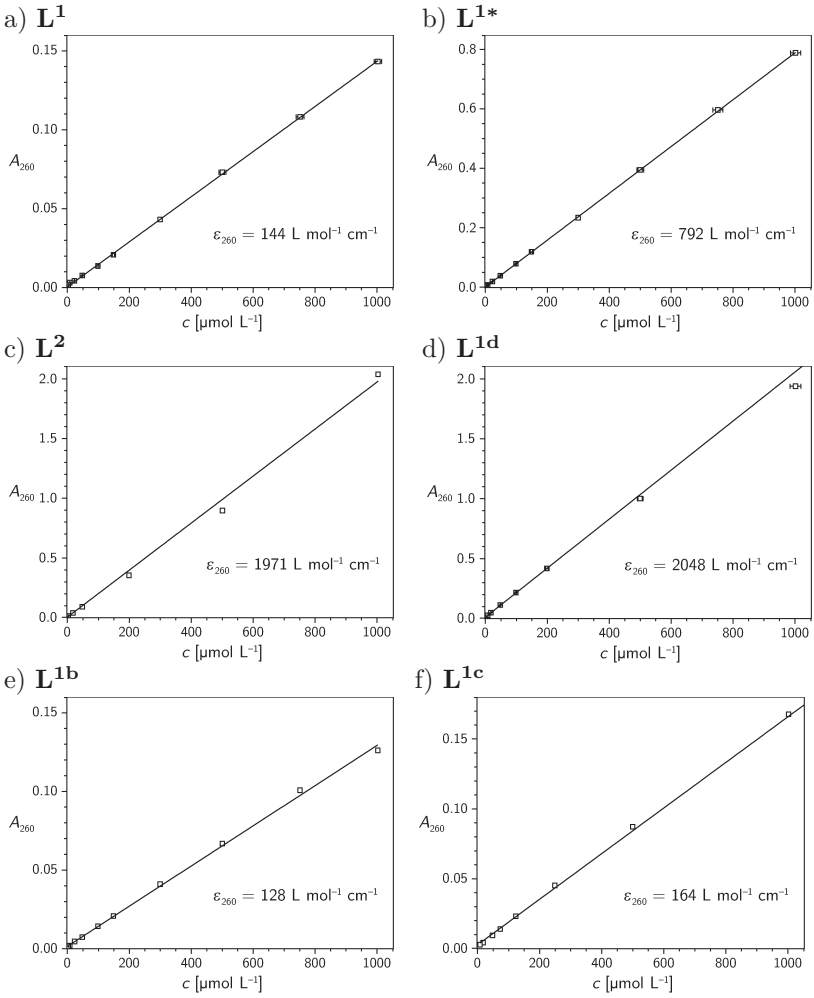


Figure 9.1.: Dilution experiments for the extinction coefficient (ϵ_{260}) determination of compounds **3** (a), **8** (b), **18** (c), **13** (d), **11a** (e) and **11b** (f). The absorbance at 260 nm was measured first for the highest concentration. The sample was then diluted successively and the absorbance measured for each step.

Oligonucleotide single-strand concentrations were 7.5 µmol L^{-1} for the tetramolecular G-quadruplexes and 4 µmol L^{-1} for the unimolecular forming sequences. For all experiments, samples were prepared with ultrapure water (type I, 18.2 MΩ cm), obtained with a Puranity TU 3 UV *VWR International GmbH*, Darmstadt.

Samples were heated to 85 °C (or 95 °C) for 10 min, then slowly cooled to 4 °C at 0.5 °C min^{-1} and then left at this temperature for several hours (typically over night). In case of the tetramolecular G-quadruplexes, formation was induced by a subsequent freeze-thaw procedure (>1 h at –25 °C or –30 °C).

9.1.6. UV-VIS spectroscopy

Both UV-VIS spectra and thermal denaturation curves (melting curves) were recorded on a *Jasco* V-650 spectrophotometer equipped with a PAC-743 6-cell thermostat for temperature control. The temperature was measured in a water-filled cuvette. Data interval was set to 1 nm, UV-Vis bandwidth to 2.0 nm, response time to medium. In order to avoid condensation of water on the cuvette surface (*Hellma Analytics* 114-QS, quartz glass, 1 cm path length) or cell window, a constant flow of dried air was pumped through the measurement cell. Evaporation of water at high temperatures and resulting changes in the absorption behaviour was minimized by a thin layer of silicon oil applied onto the sample and tightly stoppering the cuvette. The volume of the additive solutions (transition metal salts, Na_2H_2EDTA) were kept as small as possible compared to the total volume of the sample (1.0 mL) to avoid major changes in absorption due to dilution.

For normal UV-VIS spectra the scan rate was 200 nm min^{-1}. Spectra were recorded both before and after thermal denaturation. All samples were temperature equilibrated (0 °C/4 °C or 85 °C/95 °C) for at least two minutes prior to measurement of full spectra. For the unimolecular G-quadruplexes thermal difference spectra were calculated by subtracting the low-temperature (0 °C/4 °C) UV-VIS absorbance spectrum from the one measured at high-temperature (85 °C/95 °C). In case of the tetramolecular G-quadruplexes, the spectrum after de-

naturation (at 4 °C) was subtracted from the one before denaturation (also 4 °C). In both cases, a negative band (hypochromic shift) at approximately 295±2 nm and positive bands at 243±2 nm and 273±2 nm (hyperchromic shift) indicated G-quadruplex formation.[170]

In case of the thermal denaturation curves, absorption of the sample was recorded in a 0.5 °C interval with a temperature gradient set to 0.5 °C min^{-1}, which corresponds to ∼0.4 °C min^{-1} including the measurement time. To prevent major deviations from this temperature gradient, the absorption of always three cuvettes was measured. Data points were recorded from 0 °C or 4 °C to 85 °C or 95 °C and recursively back. Two to three full cycles were recorded in case of the unimolecular samples.

UV-VIS spectra were background corrected (buffer and electrolyte) and zeroed using the absorption at 350 nm. Melting curves were background corrected using the absorption at 350 nm, averaged in case of the unimolecular G-quadruplexes and converted to the fraction folded values by linear fitting of the low- and high temperature baselines.[276] Melting temperatures were determined by reading the respective value at the fraction folded value $\alpha = 0.5$. In cases where the baselines could not be appropriately fitted, or the melting curves showed more than one transition, melting temperatures were estimated using the first derivative method or by fitting of the curves with a double sigmoidal function.

9.1.7. CD spectroscopy

CD spectra were measured on a *Jasco* J-810 Spectropolarimeter (350–200 nm, 50 nm min^{-1} (in a few cases 200 nm min^{-1}), 1 nm (few cases 0.2 nm) data interval, 1.0 nm bandwidth, standard sensitivity, D.I.T. 1 s, 1–5x data accumulation) equipped with a PTC-423s thermostat for temperature control. All spectra were background corrected (buffer and electrolyte), zeroed to the average signal between 330–350 nm, and smoothed (FFT). Ellipticity in mdeg was converted to molar circular dichroism (per strand, not per nucleotide) using equations 4.4 and 4.5 (section 4.3.2, page 55). The background was measured in the same cuvette and cuvette orientation as the sample. To avoid condensation

of water onto the cuvette surface or cell window, a constant nitrogen gas flow was maintained. Samples were prepared as described above in section 9.1.5. The volume of the metal salt or Na_2H_2EDTA solution was kept as small as possible compared to the total volume of the sample (1.0 mL) to avoid major changes in absorption due to dilution. For the time-dependant CD spectra, a spectrum was recorded every minute (with a scan rate of 200 nm min^{-1}); solutions of the metal salts or Na_2H_2EDTA were added right between two measurements, with a subsequent short mixing of the whole sample. For the thermal denaturation experiments, whole spectra of the sample were recorded (200 nm min^{-1}) in a 1.0 °C interval with a temperature gradient of 0.5 °C min^{-1} from 4 °C to 80 °C. The spectra were treated in the same way as described above. An additional smoothing (15 pt Savitzky-Golay) was applied to remove the high noise level due to the fast data acquisition rate. No smoothing was applied for the single-wavelength melting curves.

9.1.8. Gel-electrophoresis

Non-denaturing gel electrophoresis experiments were carried out using a *CBS Scientific* Dual-Vertical Mini-Gel system MGV-402 at 115 V (80 mA, runtime ∼1 h) in constant voltage mode using a Tris-borate 0.09 mol L^{-1} pH 8.3 running buffer with 100 mmol L^{-1} NaCl. The gels (0.75 mm thickness, 10 cm width, 8 cm height, 26 % (w/v) acrylamide, 0.7 % (w/v) bisacrylamide) were cast in the same buffer. Ammonium peroxodisulfate (APS) was used as the radical initiator, whereas *N*,*N*,*N*′,*N*′-tetramethylethylenediamine (TEMED) was used as a stabiliser. A short stacking gel was prepared on top of the separating gel (5 % (w/v) acrylamide, 0.14 % (w/v) bisacrylamide). G-quadruplex samples and the 15bp dsDNA control were annealed as described in section 9.1.5 in 10 mmol L^{-1} sodium cacodylate buffer with 100 mmol L^{-1} NaCl. The single stranded samples were prepared without buffer in deionized water (type I) and heated to 85 °C just prior to the electrophoresis experiment. DNA quantities: 15 nt ssDNA 300 pmol, 15 bp dsDNA 40 pmol, [$\mathbf{L^1}$d(G_4)] (as ssDNA) 1.97 nmol, G-quadruplex of [$\mathbf{L^1}$d(G_4)] 450 pmol with 0 or 370 pmol $CuSO_4$ or $AgBF_4$, [$\mathbf{L^{1*}}$d(G_4)] (as ssDNA) 890 pmol, G-

Table 9.2.: Reference sequences used in the gel-electrophoresis experiments.

DNA	sequence
15nt ssDNA	5′-TGCAAGCTTGGCACT-3′
15bp dsDNA	5′-TGCAAGCTTGGCACT-3′ 3′-ACGTTCGAACCGCGA-5′

quadruplex with [$\mathbf{L^{1*}}$d(G_4)] 220 pmol. Staining was achieved either by treatment of the gel with an aqueous solution of PPIX (protoporphyrin IX, *Frontier Scientific*, 25 µL of a solution containing 1.7 mg mL^{-1} PPIX in dimethylsulfoxide dissolved in 50 mL deionized water) for 10 min on a rocking platform or by reductive silver impregnation.[249] The PPIX stained gels were irradiated using a UV transilluminator (254 nm) and photographed with a digital camera equipped with a *Hoya* HMC filter (G). Silver stained gels were scanned. Alternatively, UV-shadowing (254 nm) was used to visualize the bands. All pictures were converted to black and white and adjusted for brightness and contrast.

9.1.9. EPR spectroscopy

EPR spectroscopic measurements were carried out using a *Bruker* Elexsys E500 CW-EPR and a Super-X microwave bridge operating at 9.3–9.5 GHz. The spectrometer was equipped with a standard *Bruker* X-band ER4119-SHQE cavity and a liquid helium cryostat (*Oxford Instruments*). For the temperature-variable EPR spectra the following conditions were used. The G-quadruplex was annealed by slow cooling from 85 °C to room temperature and a subsequent freeze-thaw process (–32 °C). After addition of the copper salt and 10 % (v/v) glycerol to the quadruplex solution, the sample was immediately frozen in liquid nitrogen. Final concentrations: 9 µmol L^{-1} sodium cacodylate pH 7.1, 89 mmol L^{-1} NaCl, 632 µmol L^{-1} [$\mathbf{L^1}$d(G_4)], 163 µmol L^{-1} $CuSO_4$. Experimental parameters: 9.4 GHz, microwave power: 0.6325 mW at

20–70 K and 0.1002 mW at 10 K, conversion time: 160 ms, modulation amplitude: 10 G, modulation frequency: 100 kHz. For the spectra acquisitions involving addition of edta the same conditions were used, except for the conversion time (123 ms). The EPR signal of the copper quadruplex samples were averaged for one hour.

9.1.10. Molecular dynamics simulation

For the molecular dynamics simulations, the following protocol was used. A separate molecular model of the ligand $\mathbf{L^2}$ was capped at the 5′-*O* and 3′-*O* ends with methyl phosphate groups, which function as a substitute for the covalent bridge to the next nucleotides. Geometry optimisation at the MP2/6-31G(d,p) level was performed, using Gaussian '09 [309] according to the reported parametrization protocols. [284] This model then provides all necessary bond lengths, angles, and torsions. Atom point charges were calculated by electrostatic potential (ESP) fitting at the HF/6-31G(d) level of theory, followed by restricted electrostatic potential (RESP) fitting [286–288] implemented in the *antechamber* program within the *AmberTools14* package. [310] Charges of the *P* and *O* atoms of the –PO_2OMe caps were constrained to their values in the original AMBER force field. [284] For the copper ion bound to the ligands, a slightly different approach had to be used. Here, the complex of one Cu^{2+} bound by four of the capped ligands L2 was geometry optimised at the HF level with the 6-31G(d) basis set on the ligth atoms (*C*,*H*,*N*,*O*,*P*) and the LANL2DZ basis set on the Cu^{2+} atom. The LANL2 effective core potential (ECP) was used for the inner electrons of the Cu^{2+} ion. Unconstrained RESP fitting led a negative point charge on the Cu^{2+} ion, so a smaller model with only four pyridines as ligands was used instead to calculate a more reliable charge (+0.345) on the copper cation. This value was then used as a constraint for the RESP fitting of the larger $Cu^{2+}(\mathbf{L^2})_4$ model. For future MD simulations, the use of the RED tools (RESP ESP charge Derive)† or the RED server is highly encouraged, as the charge calculation becomes more standardized and is even slightly simpler. [311]

†See q4md-forcefieldtools.org

For the ligand alone, force constants were generated with the *parmchk2* program within *AmberTools14*, which searches for similar atom combinations and uses the corresponding values. Force constants for parameters that involve the Cu^{2+} ion were approximated using literature values.[289,290] A precise parametrisation of the copper coordination environment was not necessary here, so no further efforts were undertaken to improve these force constants. Instead, torsional constraints were imposed to maintain the symmetrical propeller-like arrangement of the pyridine rings around the central Cu^{2+} ion.[289] For future MD simulations the VFFDT program to determine metal related force constants based on frequency calculations could be used.[312] The problem of ambiguous bond angle definitions was avoided by altering the atom names of adjacent pyridine-*N* atoms. For all simulations it was assumed that the ligand's enantiomers have the same force field parameters, so that the whole parametrisation process was only done for ligand (*S*)-**L²** and then just copied to the (*R*)-form.

MD simulations were performed using the GROMACS 5.0.4[294–298] software package and followed mainly literature recommendations for the simulation of G-quadruplexes.[188] The pre-installed *parmbsc0* force field was used including the modifications as described above. Recently a revised force field called *parmbsc1* was developed, which could be of use in future MD simulations.[313] The initial topologies used as the starting point for the subsequent MD calculations were derived from the reported structure of the antiparallel human telomeric G-quadruplex htel22 (PDB entry 143D, first model).[266] All manipulations were carried out using UCSF Chimera.[291,292] The guanine nucleotides of one G-quartet were manually removed and instead the ligand models optimised before inserted. The same was done for the case of the copper–ligand complex. As the NMR structure does not contain K^+ ions, those were manually positioned in the central G-quadruplex channel, between the two G-quartets. In GROMACS, the model was positioned in the middle of a rhombic dodecahedron box (cutoff 1 nm) and solvated with explicit TIP3P water molecules. The negative net charge of the solute was neutralised by addition of the corresponding amount of randomly

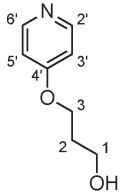
positioned K^+ ions. Then, the system was relaxed by a two-stage energy minimisation, with the first round comprising 500 steps of steepest descent and the second 2500 steps of conjugate gradient minimisation. The non-bonded Lennard-Jones cutoff was set to 1 nm and the non-bonded pair list updated every 50 steps. For the coulombic interactions, Particle-mesh Ewald summation (PME) was used.[314] Next, the system was equilibrated with positional constraints (1000 kJ mol^{-1} A^{-2}) on the model's heavy-atoms and the following parameters: 10000 steps, time step 2 fs; temperature coupling v-rescale, 298 K; pressure coupling isotropic, Berendsen, 1 bar, time constant for coupling 2 ps, compressibility 4.5 10^{-5}; velocity generation. In a second round 100000 steps of equilibration were performed with identical parameters, with the exception of the constraints, which were reduced to 100 kJ mol^{-1} A^{-2}, and without the generation of initial velocities. The so obtained *NpT* ensemble was then subjected to a 20 ns MD production run with no constraints, except for torsional constraints on the copper complex to maintain the correct arrangement of the pyridines around the copper ion. The calculated trajectories were first centred, aligned and then analysed with VMD[315,316] for root mean square deviation (rmsd) calculations. Models were visualized with UCSF Chimera.[291,292]

9.2. Synthetic procedures

If necessary, reactions were carried out under an inert nitrogen atmosphere using standard Schlenk techniques. Chemicals and standard solvents were purchased from *Sigma Aldrich*, *Acros Organics*, *Carl Roth*, *TCI Europe*, *VWR*, or *ABCR* and used as received. All chemicals and reagents used in the DNA experiments were of highest purity to avoid interference from trace contaminations. 4-Bromomethyl-pyridine hydrobromide **17** was either purchased from *Acros Organics* or synthesized from 4-(hydroxymethyl)-pyridine **13** according to a literature procedure.[218] Dry solvents over molecular sieves were purchased from *Acros Organics*. Reactions were monitored with thin layer chromatography (TLC), using silica coated aluminium plates (*Merck*, silica 60,

fluorescence indicator F254, thickness 0.25 mm). For the column chromatography silica (*Merck*, silica 60, 0.02-0.063 mesh ASTM) was used as the stationary phase.

9.2.1. Synthesis of 3-(pyridin-4-yl-oxy)-1-propanol (3)



3

Compound **3** was synthesized according to a slightly modified literature procedure.[206]

4-Chloropyridine hydrochloride **1** (1.00 g, 6.67 mmol, 1.00 eq.) was suspended in dimethylsulfoxide (5 mL) and then treated with a suspension of freshly distilled propane-1,3-diol **2** (2.40 mL, 33.2 mmol, 4.98 eq.) in dimethylsulfoxide (5 mL). This mixture was sonificated shortly and heated to 40 °C until the solids had dissolved. To the pale yellow solution sodium hydroxide (700 mg, 17.5 mmol, 2.62 eq.) was added and the reaction mixture was then heated to 100 °C and stirred at this temperature for 19.5 h. The colour of the suspension changed quickly from yellow to brown and then slowly to dark red. The reaction was poured into 50 mL of a water/ice mixture and the resulting solution was extracted three times with ethyl acetate. The combined organic fractions were dried over sodium sulfate, filtered and subsequently the solvents evaporated. The crude product was purified twice by column chromatography on silica (ethyl acetate/ethanol 100:0 → 0:100, ethyl acetate/ethanol 10:1) to give the title compound (0.542 g, 53 %) as a white solid.

chemical formula $C_8H_{11}NO_2$

molecular weight 153.18 g mol^{-1}

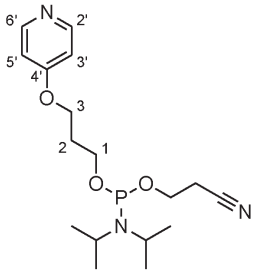
^{1}H-NMR (300 MHz, $CDCl_3$) δ = 1.84 (s_{broad}, 1H, -OH), 2.06 (tt, J = 6.0, 6.0 Hz, 2H, 2-H_2), 3.86 (t, J = 6.0 Hz, 2H, 1-H_2), 4.16 (t, J = 6.0 Hz, 2H, 3-H_2), 6.75 – 6.85 (m, 2H, 3′-H, 5′-H), 8.37 – 8.46 (m, 2H, 2′-H, 6′-H) ppm.

^{13}C-NMR (75 MHz, $CDCl_3$) δ = 31.9, 59.7, 65.1, 110.4, 151.2, 165.0 ppm.

MS (ESI, pos., MeOH) m/z (%) = 96.0 (40) $[C_5H_6NO]^+$, 154.1 (100) $[M+H]^+$; calc. for $C_8H_{12}NO_2$ 154.0865 $[M+H]^+$, found 154.0863.

MS (EI, pos.) m/z = 95.0 $[M–(CH_2)_3O]^+$, 153.1 $[M]^+$.

9.2.2. Synthesis of 3-(pyridin-4-yl-oxy)-propan-1-yl (2-cyanoethyl) *N,N*-diisopropylphosphoramidite (5)



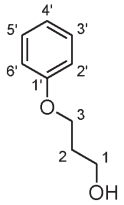
5

3-(4-Pyridyloxy)-1-propanol **3** (49.3 mg, 0.322 mmol, 1.00 eq.) was coevaporated with 2 mL of anhydrous tetrahydrofurane in an inert N_2 atmosphere. 2 mL of anhydrous tetrahydrofurane and diisopropylethylamine (220 µL, 1.29 mmol, 4.02 eq.) was added, together with a few grains of dried molecular sieves (3 Å) and the solution stirred at room temperature for 1 h. 2-cyanoethyl-*N,N*-diisopropylchlorophosphoramidite **4** (93.0 µL, 0.417 mmol, 1.30 eq.) was added dropwise via a

syringe and the resulting solution was stirred for one hour at room temperature. The reaction mixture was evaporated under reduced pressure, the residue was dissolved in 1 ml of chloroform/methanol (5:1) with 0.1 % (*v*/*v*) pyridine and then purified by column chromatography under an atmosphere of protecting gas (desactivated silica, chloroform/methanol 10 : 1, all solvents degassed). The product was obtained as a pale yellow oil after removal of the solvents by evaporation in high vacuum. The phosphoramidite **5** was stored at −32 °C and used within several days for the automated DNA synthesis.

chemical formula	$C_{17}H_{28}N_3O_3P$
molecular weight	353.40 g mol^{-1}
^{1}H-NMR (300 MHz, $CDCl_3$)	δ = 1.26 (dd, J = 6.8, 4.2 Hz, 12H, $N(CH(C\underline{H}_3)_2)_2$), 2.10 (tt, J = 6.2, 6.1 Hz, 2H, 2-H_2), 2.61 (t, J = 6.4 Hz, 2H, $OCH_2C\underline{H}_2CN$), 3.59 (dsept, J = 10.2, 6.8 Hz, 2H, $N(C\underline{H}(CH_3)_2)_2$), 3.70 – 3.92 (m, 4H, 1-H_2, $OC\underline{H}_2CH_2CN$), 4.14 (t, J = 6.2 Hz, 2H, 3-H_2), 6.76 – 6.86 (m, 3H, 3′-H, 5′-H), 8.36 – 8.47 (m, 2H, 2′-H, 6′-H) ppm.
^{31}P-NMR (121 MHz, $CDCl_3$)	δ = 147.89 ppm.
MS (ESI, pos., MeCN)	m/z = 354.17 $[M+H]^+$.

9.2.3. Synthesis of 3-phenoxy-propan-1-ol (8)



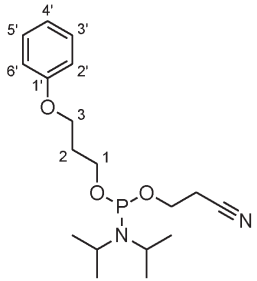
8

Compound **8** was synthesized according to a literature procedure.[210] Phenol (1.01 g, 10.7 mmol, 1.00 eq.) was dissolved in dimethylformamide (3 mL) and added to a suspension of potassium carbonate (1.97 g, 11.7 mmol, 1.10 eq.). The suspension was heated to 60 °C and stirred at this temperature for 40 min. 3-chloro-1-propanol (0.990 mL, 14.3 mmol, 1.34 eq.) was added slowly and the resulting mixure stirred at 60 °C for 5.5 h and then slowly cooled to room temperature. Water (20 mL) was added, the reaction mixture neutralized with 1 mol L^{-1} aqueous hydrochloric acid and then extracted four times with dichloromethane (total 100 mL). The combined organic fractions were dried over magnesium sulfate, filtered and the solvents evaporated. The crude product was purified by column chromatography on silica (hexane/ethyl acetate 8:2) to give the title compound **8** (1.12 g, 7.36 mmol, 69 %) as a pale yellow oil.

chemical formula	$C_9H_{12}O_2$
molecular weight	152.19 g mol^{-1}
^{1}H-NMR (300 MHz, $CDCl_3$)	δ = 1.90 (s_{broad}, 1H, 1-OH), 1.99 (tt, J = 6.0, 6.0 Hz, 2H, 2-H_2), 3.81 (t, J = 6.0 Hz, 2H, 1-H_2), 4.07 (t, J = 6.0 Hz, 2H, 3-H_2), 6.81 – 6.95 (m, 3H, 2′-H, 4′-H, 6′-H), 7.16 – 7.30 (m, 2H, 3′-H, 5′-H) ppm.

^{13}C-NMR (75 MHz, $CDCl_3$)	δ = 32.1, 60.6, 65.8, 114.6, 121.0, 129.6, 158.9 ppm.
MS (ESI, pos., MeOH)	m/z = 175.1 $[M+Na]^+$; calc. for $C_9H_{12}O_2Na$ 175.0731 $[M+Na]^+$, found 175.0730.

9.2.4. Synthesis of 3-phenoxy-propan-1-yl (2-cyanoethyl) *N,N*-diisopropylphosphoramidite (9)

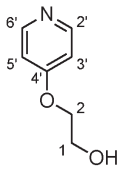


9

3-Phenoxy-1-propanol **8** (92.7 mg, 0.609 mmol, 1.00 eq.) was dissolved in 4 mL of anhydrous tetrahydrofurane under an inert N_2 atmosphere. Dried molecular sieves (3 Å) were added and the solution was concentrated to 2 mL under reduced pressure. *N,N*-diisopropylethylamine (400 µL, 2.35 mmol, 3.86 eq.) was added, and the solution stirred at room temperature for 50 min. 2-cyanoethyl-*N,N*-diisopropylchlorophosphoramidite **4** (180 µL, 0.807 mmol, 1.32 eq.) was added dropwise via a syringe and the resulting solution was stirred for one hour at room temperature. The reaction mixture was purified directly by column chromatography under an atmosphere of protecting gas (deactivated silica, chloroform, all solvents degassed). The product was obtained as a pale yellow oil after removal of the solvents by evaporation in high vacuum. The phosphoramidite **9** was stored at −32 °C and used within several days for the automated DNA synthesis.

chemical formula	$C_{18}H_{29}N_2O_3P$
molecular weight	352.42 g mol^{-1}
^{1}H-NMR (300 MHz, $CDCl_3$)	δ = 1.18 (dd, J = 6.7, 4.0 Hz, 12H, $N(CH(C\underline{H}_3)_2)_2$), 2.09 (tt, J = 6.2, 6.2 Hz, 2H, 2-H_2), 2.60 (t, J = 6.7 Hz, 2H, $OCH_2C\underline{H}_2CN$), 3.60 (dsept, J = 10.2, 6.8 Hz, 2H, $N(C\underline{H}(CH_3)_2)_2$), 3.70 – 3.94 (m, 4H, 1-H_2, $OC\underline{H}_2CH_2CN$), 4.08 (t, J = 6.2 Hz, 2H, 3-H_2), 6.85 – 6.99 (m, 3H, 2′-H, 4′-H, 6′-H), 7.23 – 7.33 (m, 2H, 3′-H, 5′-H) ppm.
^{13}C-NMR (75 MHz, $CDCl_3$)	δ = 20.5 (d, J = 7.0 Hz), 24.8 (t, J = 6.9 Hz), 31.2 (d, J = 7.4 Hz), 43.2 (d, J = 12.4 Hz), 58.5 (d, J = 19.4 Hz), 60.4 (d, J = 17.6 Hz), 64.5, 114.7, 117.8, 120.8, 129.6, 159.1 ppm.
^{31}P-NMR (121 MHz, $CDCl_3$)	δ = 147.65 ppm.

9.2.5. Synthesis of 2-(pyridin-4-yl-oxy)-ethan-1-ol (11a)



11a

Compound **11a** was synthesized based on a literature procedure.[211] See also chapter 5 for details.

To ethylene glycol **10a** (3.00 mL, 53.8 mmol, 3.90 eq.) dissolved in dimethylsulfoxide (15 mL) was added sodium hydroxide (1.22 g, 30.6 mmol, 2.22 eq.) and the suspension heated to 100 °C. 4-chloropyridine hydrochloride **1** (2.07 g, 13.8 mmol, 1.00 eq.) in dimethylsulfoxide (25 mL) was added slowly and the resulting mixture was heated at 100 °C over night. The mixture was allowed to cool to room temperature and extracted several times with chloroform (total 250 mL). The solvents of the organic layer were removed under reduced pressure. The crude product was purified by column chromatography on silica (ethyl acetate/methanol 9:1) to give the title compound (0.390 g, 2.80 mmol, 20 %) as a white solid.

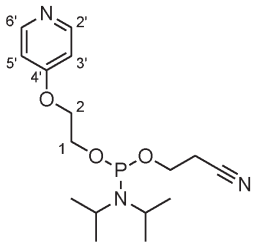
chemical formula	$C_7H_9NO_2$
molecular weight	139.15 g mol^{-1}
^{1}H-NMR (300 MHz, $CDCl_3$)	δ = 2.20 (s_{broad}, 1H, 1-OH), 3.96 – 4.05 (m, 2H, 1-H_2), 4.11 – 4.18 (m, 2H, 2-H_2), 6.80 – 6.89 (m, 2H, 3′-H, 5′-H), 8.42–8.48 (m, 2H, 2′-H, 6′-H) ppm.
^{1}H-NMR (300 MHz, DMSO-d_6)	δ = 3.65 – 3.77 (m, 2H, 1-H_2), 4.00 – 4.10 (m, 2H, 2-H_2), 4.92 (t, J = 5.3 Hz, 1H, 1-OH), 6.90 – 6.98 (m, 2H, 3′-H, 5′-H), 8.30 – 8.41 (m, 2H, 2′-H, 6′-H) ppm.

^{13}C-NMR (75 MHz, DMSO-d_6) δ = 59.3, 69.6, 110.4, 150.9, 164.5 ppm.

MS (ESI, pos., MeOH) m/z = 96.0 (70) $[C_5H_6NO]^+$, 140.1 (100) $[M+H]^+$; calc. for $C_7H_{10}NO_2$ 140.0708 $[M+H]^+$, found 140.0706.

MS (EI, pos.) m/z (%) = 51.0 (20), 78.0 (50) $[C_5H_4N]^+$, 96.0 (100) $[C_5H_6NO]^+$, 108.1 (22) $[C_6H_6NO]^+$, 139.1 (41) $[M]^+$

9.2.6. Synthesis of 2-(pyridin-4-yl-oxy)-ethan-1-yl (2-cyanoethyl) *N,N*-diisopropylphosphoramidite (12a)



12a

Phosphoramidite **12a** was synthesized from 2-(4-pyridyloxy)-1-ethanol (49.3 mg, 0.452 mmol) **11a** according to the procedure as described for **20** (subsection 9.2.13).

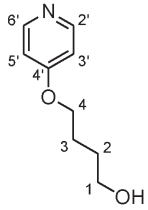
chemical formula $C_{16}H_{26}N_3O_3P$

molecular weight 339.38 g mol^{-1}

^{31}P-NMR (121 MHz, $CDCl_3$) δ = 149.28 ppm (+ signals from hydrolyzed/oxidized CEP-Cl).

MS (ESI, pos., MeCN) m/z = 241.1 (50), 340.3 (100) $[M+H]^+$.

9.2.7. Synthesis of 4-(pyridin-4-yl-oxy)-butan-1-ol (11b)



11b

Compound **3** was synthesized according to a slightly modified literature procedure.[206]

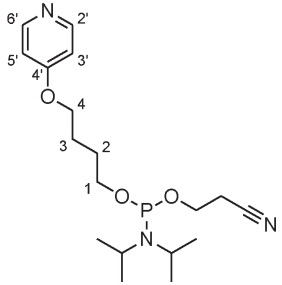
4-Chloropyridine hydrochloride **1** (0.50 g, 3.33 mmol, 1.00 eq.) was suspended in dimethylsulfoxide (30 mL) and treated with butane-1,4-diol **10b** (1.47 mL, 16.67 mmol, 5.00 eq.). Sodium hydroxide (333 mg, 8.33 mmol, 2.50 eq.) was added and the reaction mixture was then heated to 100 °C and stirred at this temperature for one day. The reaction mixture was concentrated in vacuo and then poured into 50 mL of a water/ice mixture. The resulting solution was extracted several times with ethyl acetate (total 350 mL). The combined organic fractions were dried over magnesium sulfate, filtered and the solvent evaporated. The crude product was purified by column chromatography on silica (ethyl acetate/ethanol 5:1) to give the title compound (0.272 g, 1.63 mmol, 49 %) as a white solid.

chemical formula	$C_9H_{13}NO_2$
molecular weight	167.21 g mol^{-1}
^{1}H-NMR (300 MHz, $CDCl_3$)	δ = 1.64 – 1.81 (m, 2H, 2-H_2), 1.84 – 1.97 (m, 2H, 3-H_2), 2.45 (s, 1H, 1-OH), 3.72 (t, J = 6.3 Hz, 2H, 1-H_2), 4.04 (t, J = 6.3 Hz, 2H, 4-H_2), 6.73 – 6.83 (m, 2H, 3′-H, 5′-H), 8.34 – 8.43 (m, 2H, 2′-H, 6′-H) ppm.

^{13}C-NMR (75 MHz, $CDCl_3$) δ = 25.6, 29.3, 62.3, 67.8, 110.4, 151.0, 165.2 ppm.

MS (ESI, pos., MeOH) m/z (%) = 96.0 (100) $[C_5H_6NO]^+$, 168.1 (80) $[M+H]^+$, 190.1 (45) $[M+Na]^+$; calc. for $C_9H_{13}NO_2$ 168.1020 $[M+H]^+$, found 168.1019.

9.2.8. Synthesis of 4-(pyridin-4-yl-oxy)-butan-1-yl (2-cyanoethyl) *N,N*-diisopropylphosphoramidite (12b)



12b

Phosphoramidite **12b** was synthesized from 4-(4-pyridyloxy)-1-butanol (60.1 mg, 0.359 mmol) **11b** according to the procedure as described for **20** (subsection 9.2.13).

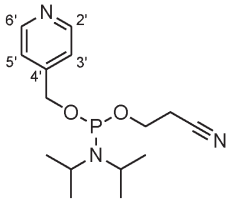
chemical formula $C_{18}H_{30}N_3O_3P$

molecular weight 367.43 g mol^{-1}

^{31}P-NMR (121 MHz, $CDCl_3$) δ = 147.50 ppm (+ minor signals from hydrolyzed/oxidized CEP-Cl).

MS (ESI, pos., MeCN) m/z = 241.1 (60), 368.3 (100) $[M+H]^+$, 430.2 (70).

9.2.9. Synthesis of (pyridin-4-yl-methyl) (2-cyanoethyl) *N,N*-diisopropylphosphoramidite (14)

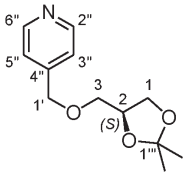


14

Phosphoramidite **14** was synthesized from pyridine-4-yl methanol (51.0 mg, 0.367 mmol) **13** according to the procedure as described for **20** (subsection 9.2.13).

chemical formula	$C_{15}H_{24}N_3O_2P$
molecular weight	309.35 g mol^{-1}
^{31}P-NMR (121 MHz, $CDCl_3$)	δ = 149.19 ppm (+ signals from hydrolyzed/oxidized CEP-Cl).

9.2.10. Synthesis of (*S*)-3-(pyridin-4-yl-methoxy)-1,2-*O*-isopropyliden-propan-1,2-diol ((*S*)-16)



(*S*)-**16**

Sodium hydride (60 % suspension in mineral oil, 0.914 g, 22.9 mmol, 2.90 eq.) was washed with anhydrous tetrahydrofurane (2 x 20 mL) in an inert atmosphere, and shortly dried under reduced pressure. Dry dimethylformamide (20 mL) was added, the suspension cooled to 0 °C

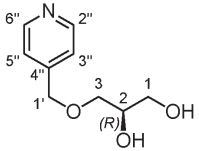
with an ice/water bath and then (*S*)-(2,2)-dimethyl-1,3-dioxolan-4-yl)-methanol (*S*)-**15** (1.12 mL, 9.07 mmol, 1.15 eq.) was added dropwise. The resulting suspension was stirred at 0 °C for 20 min, and 4-bromomethyl-pyridine hydrobromide **17** (2.00 g 7.91 mmol, 1.0 eq.) was added in portions. The reaction mixture was stirred at 0 °C for additional 30 min, and then stirred at room temperature over night. The reaction was stopped by slow addition of 100 mL water, whereas 4 mL of brine was added for better phase separation in the next step. The reaction mixture was extracted three times with ethyl acetate, the combined fractions washed three times with saturated aqueous sodium chloride solution and dried over magnesium sulfate. The organic phase was filtered and solvents removed under reduced pressure. The product (*S*)-**16** was obtained as a yellow oil (1.68 g, 7.52 mmol, 95 %), which was used in the next step without further purification.

chemical formula	$C_{12}H_{17}NO_3$
molecular weight	223.27 g mol^{-1}
^{1}H-NMR (300 MHz, $CDCl_3$)	δ = 1.37 (q, J = 0.7 Hz, 3H, 1‴-Me), 1.43 (q, J = 0.7 Hz, 3H, 1‴-Me), 3.54 (dd, J = 10.0, 5.2 Hz, 1H, 3-H$_a$), 3.59 (dd, J = 10.0, 5.6 Hz, 1H, 3-H$_b$), 3.77 (dd, J = 8.3, 6.3 Hz, 1H, 1-H$_a$), 4.08 (dd, J = 8.3, 6.5 Hz, 1H, 1-H$_b$), 4.29–4.37 (m, 1H, 2-H), 4.58 (d, J = 14.3 Hz, 1H, 1′-H$_a$), 4.62 (d, J = 14.3 Hz, 1H, 1′-H$_b$), 7.25–7.28 (m, 2H, 3″-H, 5″-H), 8.55–8.60 (m, 2H, 2″-H, 6″-H) ppm.
^{13}C-NMR (75 MHz, $CDCl_3$)	δ = 25.5, 26.9, 66.7, 71.9, 71.9, 74.8, 109.7, 121.9, 147.3, 150.0 ppm.
MS (ESI, pos., MeCN)	m/z = 224.1 (100) $[M+H]^+$; calc. for $C_{12}H_{18}NO_3$ 224.1281 $[M+H]^+$, found 224.1281.

Compound (*R*)-**16** (1.67 g, 7.48 mmol, 95 % yield) was synthesized from (*R*)-(2,2)-dimethyl-1,3-dioxolan-4-yl)methanol (*R*)-**15** (2.00 g,

7.91 mmol) using the same procedure as for the enantiomer (*S*)-**16** and showed identical NMR and mass spectrometric data.

9.2.11. Synthesis of (*R*)-3-(pyridin-4-yl-methoxy)propan-1,2-diol ((*R*)-18)



(*R*)-**18**

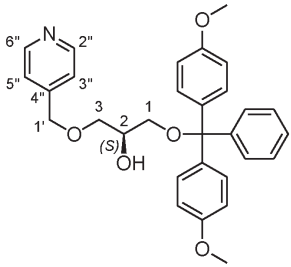
Compound (*S*)-**16** (1.45 g, 6.50 mmol, 1.00 eq.) was dissolved in tetrahydrofurane/water (30 mL, 1:1 *v*/*v*) and trifluoroacetic acid (540 µL, 7.00 mmol, 1.08 eq.) added dropwise. The reaction mixture was stirred at 70 °C for 2 h and then cooled again to room temperature. The acidic solution was neutralized by addition of aqueous ammonia (25 %) and the solvents were removed *in vacuo*. The crude product was purified by column chromatography on silica (chloroform/methanol 10:1) to give the title compound (0.940 g, 5.10 mmol, 79 %) as a yellow oil.

chemical formula	$C_9H_{13}NO_3$
molecular weight	183.21 g mol^{-1}
^{1}H-NMR (300 MHz, $CDCl_3$)	δ = 2.24 (s_{broad}, 2H, 1-OH, 2-OH), 3.59 (dd, J = 9.7, 6.6 Hz, 1H, 3-H_a), 3.63 (dd, J = 9.7, 4.5 Hz, 1H, 3-H_b), 3.67 (dd, J = 11.4, 5.6 Hz, 1H, 1-H_a), 3.76 (dd, J = 11.4, 3.9 Hz, 1H, 1-H_b), 3.92–4.00 (m, 1H, 2-H), 4.56–4.62 (m, 2H, 1′-H_a, 1′-H_b), 7.22–7.29 (m, 2H, 3″-H, 5″-H), 8.44–8.71 (m, 2H, 2″-H, 6″-H) ppm.
^{13}C-NMR (75 MHz, $CDCl_3$)	δ = 64.0, 70.8, 71.9, 72.5, 121.9, 147.3, 149.9 ppm.

MS (ESI, pos., MeCN) $m/z = 184.1$ (100) $[M+H]^+$; calc. for $C_9H_{14}NO_3$ 184.0968 $[M+H]^+$, found 184.0969.

Compound (*S*)-**18** (0.942 mg, 5.10 mmol, 65 % yield) was synthesized from (*R*)-**16** (1.45 g, 6.50 mmol) using the same procedure as for the enantiomer (*R*)-**18** and showed identical NMR and mass spectrometric data.

9.2.12. Synthesis of (*S*)-1-(4,4′-dimethoxytrityl-oxy)-3-(pyridin-4-yl-methoxy)propan-2-ol ((*S*)-19)



(*S*)-**19**

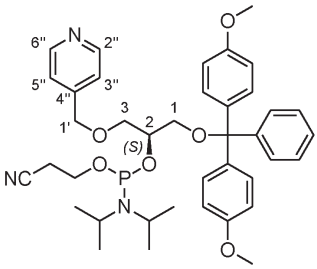
Diol (*R*)-**18** (0.766 g, 4.18 mmol, 1.00 eq.) was co-evaporated with anhydrous tetrahydrofurane (3 x 5 mL) in an inert atmosphere and then dissolved in anhydrous tetrahydrofurane (3 mL). 4-dimethylaminopyridine (30 mg, 0.24 mmol, 0.06 eq.) was added and the solution cooled to 0 °C with an ice/water bath. A cooled solution of 4,4′-dimethoxytrityl chloride (1.57 g, 4.63 mmol, 1.10 eq.) in anhydrous tetrahydrofurane (5 mL) was added, followed by diisopropylethylamine (1.40 mL, 8.23 mmol, 2.00 eq.). The reaction mixture was slowly warmed to room temperature and then stirred overnight. Methanol (1 mL) and ethyl acetate (50 mL) were added and the organic layer was washed with sat. aqueous sodium hydrogen carbonate solution (3 x 25 mL), water (3 x 25 mL), and saturated aqueous sodium chloride solution (3 x 25 mL). The organic layer was dried over magnesium sulfate, filtered, and the solvents removed under reduced pressure. The crude product

was purified by column chromatography on silica (ethyl acetate, 1 % NEt_3) to give the title compound (1.62 g, 3.33 mmol, 80 %) as a white solid.

chemical formula	$C_{30}H_{31}NO_5$
molecular weight	485.58 g mol^{-1}
^{1}H-NMR (300 MHz, $CDCl_3$)	δ = 2.49 (s_{broad}, 1H, 2-OH), 3.22 – 3.32 (m, 2H, 1-H_a, 1-H_b), 3.61 (dd, J = 9.7, 5.8 Hz, 1H, 3-H_a), 3.66 (dd, J = 9.7, 4.5 Hz, 1H, 3-H_a), 3.75 – 3.83 (m, 6H, 2x OMe), 3.97 – 4.07 (m, 1H, 2-H), 4.51 – 4.60 (m, 2H, 1′-H_a, 1′-H_b), 6.80 – 6.87 (m, 4H, H_{ar}), 7.16–7.19 (m, 2H, 3″-H, 5″-H), 7.19 – 7.36 (m, 7H, H_{ar}), 7.41 – 7.47 (m, 2H, H_{ar}), 8.53 – 8.59 (m, 2H, 2″-H, 6″-H) ppm.
^{13}C-NMR (75 MHz, $CDCl_3$)	δ = 55.4, 64.4, 70.1, 71.7, 72.3, 86.4, 113.3, 121.8, 127.0, 128.0, 128.3, 130.2, 136.0, 144.9, 147.5, 149.9, 158.7 ppm.
MS (ESI, pos., MeCN)	m/z = 303.1 (20) $[DMT]^+$, 486.2 (100) $[M+H]^+$, 524.2 (23) $[M+K]^+$; calc. for $C_{30}H_{32}NO_5$ 486.2275 $[M+H]^+$, found 486.2279.

Compound (*R*)-**19** (1.82 g, 3.76 mmol, 73 % yield) was synthesized from (*S*)-**18** (0.940 g, 5.13 mmol) using the same procedure as for the enantiomer (*S*)-**19** and showed identical NMR and mass spectrometric data.

9.2.13. Synthesis of (*S*)-1-(4,4′-dimethoxytrityl-oxy)-3-(pyridin-4-yl-methoxy)propan-2-yl (2-cyanoethyl) *N*,*N*-diisopropylphosphoramidite ((*S*)-20)



(*S*)-**20**

Compound (*S*)-**19** (182 mg, 0.375 mmol, 1.00 eq.) was co-evaporated with anhydrous tetrahydrofurane (2 x 1 mL) in an inert atmosphere, dissolved in anhydrous tetrahydrofurane (6 mL) and diisopropylethylamine (255 µL, 1.50 mmol, 4.00 eq.) was added dropwise. 2-cyanoethyl-*N*,*N*-diisopropylchlorophosphoramidite **4** (CEP-Cl, 92.0 µL, 0.413 mmol, 1.10 eq.) was added dropwise and the resulting solution was stirred for one hour at room temperature, during which time a white precipitate formed. The reaction mixture was filtered through a syringe filter to remove most of the precipitated diisopropylethylammonium chloride. The solvent was removed under reduced pressure to give the phosphoramidite (*S*)-**20** as a yellow oil. The crude product was used without further purification either directly, or stored at −25 °C and used within several days for the automated DNA synthesis.

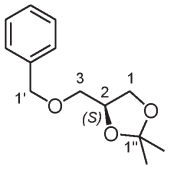
chemical formula $C_{39}H_{48}N_3O_6P$

molecular weight 685.80 g mol^{-1}

^{1}H-NMR (300 MHz, $CDCl_3$)	δ = 1.13–1.20 (m, 2x12H, $N(CH(C\underline{H}_3)_2)_2$), 2.40–2.49 (m, 2H, $OCH_2C\underline{H}_2CN$), 2.52–2.61 (m, 2H, $OCH_2C\underline{H}_2CN$), 3.15–3.35 (m, 2x2H, 2x 1-H_2), 3.39–3.89 (m, 2x12H, 2x OMe, 2x 3-H_2, $OC\underline{H}_2CH_2CN$, $N(C\underline{H}(CH_3)_2)_2$), 4.07–4.27 (m, 2x1H, 2-H), 4.45–4.62 (m, 2x2H, 2x 1′-H), 6.74–6.84 (m, 2x4H, H_{ar}), 7.13–7.36 (m, 2x9H, 3″-H, 5″-H, H_{ar}), 7.40–7.48 (m, 2x2H, H_{ar}), 8.48–8.56 (m, 2x2H, 2″-H, 6″-H) ppm (due to the two diastereomers present, all NMR signals are doubled).
^{31}P-NMR (121 MHz, $CDCl_3$)	δ = 149.33, 149.44 ppm (+ minor signals from hydrolyzed/oxidized CEP-Cl).

Phosphoramidite (*R*)-**20** was synthesized from (*R*)-**19** using the same procedure as for the enantiomer (*S*)-**20** and showed an identical ^{31}P-NMR spectrum.

9.2.14. Synthesis of (*S*)-3-(phenyl-methoxy)-1,2-*O*-isopropyliden-propan-1,2-diol ((*S*)-22)



(*S*)-**22**

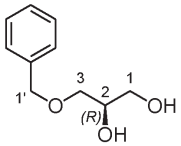
Sodium hydride (60 % suspension in mineral oil, 0.440 g, 11.0 mmol, 2.81 eq.) was treated with anhydrous dimethylformamide (22 mL) in an inert atmosphere, and the suspension cooled to 0 °C with an ice/water bath. (*S*)-(2,2)-dimethyl-1,3-dioxolan-4-yl)methanol (*S*)-**15** (0.517 g, 3.91 mmol, 1.00 eq.) was added dropwise and the resulting suspension was stirred at 0 °C for 30 min, after which benzyl bromide **21** (0.834 g, 4.88 mmol, 1.25 eq.) was added in portions. The reaction mixture was stirred at room temperature over night. The reaction was stopped by slow addition of 20 mL water. The reaction mixture was extracted with dichloromethane (3 x 15 mL), the combined fractions washed with saturated aqueous sodium chloride solution (5 x 15 mL) and then dried over magnesium sulfate. The organic layer was filtered and the solvent removed under reduced pressure. The product was purified by column chromatography on silica (petroleum ether/ethyl acetate 95:5) to give the title compound as a yellow oil (0.719 g, 3.23 mmol, 83 %).

chemical formula $C_{13}H_{18}O_3$

molecular weight 222.28 g mol^{-1}

^{1}H-NMR (300 MHz, $CDCl_3$)	δ = 1.29 (q, J = 0.7 Hz, 3H, 1″-Me), 1.35 (q, J = 0.7 Hz, 3H, 1″-Me), 3.40 (dd, J = 9.8, 5.7 Hz, 1H, 3-H_a), 3.49 (dd, J = 9.8, 5.7 Hz, 1H, 3-H_b), 3.67 (dd, J = 8.3, 6.4 Hz, 1H, 1-H_a), 3.98 (dd, J = 8.3, 6.4 Hz, 1H, 1-H_b), 4.16–4.30 (m, 1H, 2-H), 4.48 (d, J = 12.1 Hz, 1H, 1′-H_a), 4.53 (d, J = 12.1 Hz, 1H, 1′-H_b), 7.14–7.34 (m, 5H, H_{ar}) ppm.
^{13}C-NMR (75 MHz, $CDCl_3$)	δ = 25.5, 26.8, 67.0, 71.2, 73.6, 74.9, 109.5, 127.8, 127.8, 138.1 ppm.

9.2.15. Synthesis of (*R*)-3-(phenyl-methoxy)propan-1,2-diol ((*R*)-23)



(*R*)-**23**

Compound (*S*)-**22** (0.591 g, 2.66 mmol, 1.00 eq.) was dissolved in methanol (6 mL) and 1 mol L^{-1} aqueous hydrochloric acid (1.4 mL, 1.4 mmol, 0.53 eq.) added dropwise. The reaction mixture was stirred at 80 °C for 1 h, cooled to room temperature and the solvent removed *in vacuo*. The crude product was purified by column chromatography on silica (chloroform/methanol 95:5) to give the title compound (0.423 g, 2.32 mmol, 87 %) as a yellow oil.

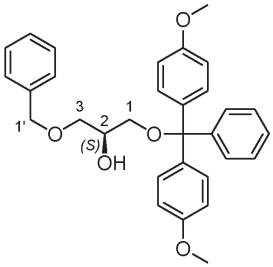
chemical formula	$C_{10}H_{14}O_3$
molecular weight	182.22 g mol^{-1}

^{1}H-NMR (300 MHz, $CDCl_3$) δ = 2.72 (s_{br}, 2H, 1-OH, 2-OH), 3.44 (dd, J = 9.7, 6.1 Hz, 1H, 3-H_a), 3.48 (dd, J = 9.7, 4.4 Hz, 1H, 3-H_b), 3.52 (dd, J = 11.5, 5.8 Hz, 1H, 1-H_a), 3.61 (dd, J = 11.5, 3.8 Hz, 1H, 1-H_b), 3.74–3.87 (m, 1H, 2-H), 4.41–4.54 (m, 2H, 1′-H_a, 1′-H_b), 7.15–7.34 (m, 5H, H_{ar}) ppm.

^{13}C-NMR (75 MHz, $CDCl_3$) δ = 64.1, 70.8, 71.9, 73.7, 127.9, 128.0, 128.6, 137.8 ppm.

MS (ESI, pos., MeCN) m/z (%) = 205.1 (100) $[M+Na]^+$, 309.1 (95); calc. for $C_{10}H_{14}O_3Na$ 205.0836 $[M+Na]^+$, found 205.0835.

9.2.16. Synthesis of (*S*)-1-(4,4′-dimethoxytrityl-oxy)-3-(phenyl-methoxy)propan-2-ol ((*S*)-24)



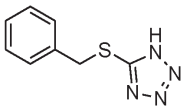
(*S*)-**24**

Diol (*R*)-**23** (0.272 g, 1.49 mmol, 1.00 eq.) was co-evaporated with anhydrous pyridine (3 mL) in an inert atmosphere and dissolved in anhydrous pyridine (20 mL). The solution was cooled to 0 °C with an ice/water bath and a solution of 4,4′-dimethoxytrityl chloride (0.98 g, 2.89 mmol, 1.94 eq.) was added. The reaction mixture was slowly warmed to room temperature and stirred for two days. Small amounts of saturated aqueous sodium hydrogen carbonate solution and water were added and the aqueous layer was extracted with ethyl acetate (4 x 20 mL). The combined organic fractions were washed with saturated aqueous

sodium chloride solution (3 x 30 mL). The organic layer was dried over sodium sulfate, filtered, and the solvent removed under reduced pressure. The crude product was purified by column chromatography on silica (petroleum ether/ethyl acetate 4:1→2:1, 1 % *v*/*v* triethylamine) to give the title compound (0.321 g, 0.662 mmol, 44 %) as a yellow oil.

chemical formula	$C_{31}H_{32}O_5$
molecular weight	484.59 g mol^{-1}
^{1}H-NMR (300 MHz, $CDCl_3$)	δ = 2.68 (s_{broad}, 1H, OH), 3.16 (dd, J = 9.3, 5.3 Hz, 2H, 1-H_a), 3.20 (dd, J = 9.3, 5.6 Hz, 2H, 1-H_b), 3.52 (dd, J = 9.7, 6.1 Hz, 1H, 3-H_a), 3.58 (dd, J = 9.7, 4.4 Hz, 1H, 3-H_b), 3.65 – 3.81 (m, 6H, 2x OMe), 3.88 – 4.02 (m, 1H, 2-H), 4.46 – 4.53 (m, 2H, 1′-H_a, 1′-H_b), 6.71 – 6.83 (m, 4H, H_{ar}), 7.09 – 7.35 (m, 12H, H_{ar}), 7.09 – 7.35 (m, 2H, H_{ar}) ppm.
^{13}C-NMR (75 MHz, $CDCl_3$)	δ = 55.3, 64.5, 70.0, 71.8, 73.4, 86.2, 113.2, 126.8, 127.7, 127.8, 127.9, 128.2, 128.5, 130.2, 136.1, 138.2, 145.0, 158.6 ppm.
MS (ESI, pos., MeCN)	m/z (%) = 205.1 (25) $[M\text{-}DMT+H+Na]^+$, 303.1 (100) $[DMT]^+$, 507.2 (66) $[M+Na]^+$, 991.4 (32) $[2M+Na]^+$; calc. for $C_{31}H_{32}O_5Na$ 507.2139 $[M+Na]^+$, found 507.2142.

9.2.17. Synthesis of 5-(benzylthio)-1*H*-tetrazole (BTT)



BTT

BTT was synthesized according to a modified literature procedure,[308] provided by professor G. Clever (TU Dortmund).

Benzyl thiocyanate (35.78 g, 0.240 mol, 1.00 eq.) was dissolved in dioxane (190 mL) and treated with a solution of ammonium chloride (12.94 g, 0.242 mol, 1.01 eq.) in water (48 mL). The solution was degassed three times under reduced pressure with an ultrasonic bath. Sodium azide (15.60 g, 0.240 mol, 1.00 eq.) was added slowly and the resulting suspension stirred at 85 °C to 90 °C for a total of 6 h. The reaction mixture was slowly cooled to room temperature and concentrated under reduced pressure to approximately 50 mL. An ice-water mixture (200mL) and subsequently concentrated aqueous hydrochloric acid was added dropwise until pH 2 was reached. The precipitate was filtered off and then resuspended in toluene (150 mL). The suspension was heated to 110 °C and stirred at this temperature for 1 h. The aqueous layer was removed with a pipette, and acetonitrile (10 mL) was added to the remaining organic layer. The resulting solution was stirred at 110 °C for additional 30 min, after which it was filtered. The filtrate was slowly cooled to room temperature upon which the title compound started to crystallize, which was filtered and dried in vacuo to give BTT (26.04 g, 0.135 mol, 56 %) as a white crystalline solid.

chemical formula	$C_8H_8N_4S$
molecular weight	192.24 g mol^{-1}
^{1}H-NMR (300 MHz, CD_3CN)	δ = 4.48 (s, 2H, -CH_2-), 7.14-7.49 (m, 5H, H_{ar}) ppm.
^{13}C-NMR (75 MHz, CD_3CN)	δ = 37.5, 128.8, 129.6, 129.9, 137.6, 155.1 ppm.
MS (EI, pos.)	m/z = 91 (100) $[C_7H_7]^+$, 192 (16) $[M]^+$.

A. Appendix

The appendix comprises a step-to-step guideline for the molecular dynamics simulations performed in this thesis, and an overview of the analytical and spectroscopic data obtained for the different oligonucleotides.

A.1. Molecular dynamics

As outlined in chapter 7.1.4, the molecular dynamics simulations carried out on the modified G-quadruplexes require a parametrisation of any features not present in the published/native force field. In this appendix the work flow and step to step tasks will be explained.

1. **force field** This is an important point, as this choice affects all of the next tasks. In this work, the AMBER *ff99parmbsc0* force field was chosen, as a whole range of G-quadruplex structures have been simulated with it.

2. **parametrization** If new ligand systems, bound transition metals, etc. have to be simulated, their parameters have to be determined, as those will not be present in the native force field. For the AMBER force field, this includes force constants (bonds, angles, dihedrals) and charges (atoms).

3. **MD simulation** The actual MD production run (ns time range) is the last step in a series of simulations for solute and solvent relaxation and equilibration. The MD simulation takes into account the solute (DNA), solvent (water), electrolyte for charge neutrality (KCl), pressure (1 bar) and temperature (298 K).

All programs used here have to be correctly compiled/installed on a Linux based system. See the respective homepages/manuals for detailed instructions: AmberTools,[310] Gromacs,[293] VMD (for trajectory visualization),[315] and Chimera (for molecule modification and graphics production).[291]

A.1.1. Parametrisation

As stated above, the parametrisation of an artificial compound for the use within the AMBER force field consists of two parts: charges and force constants.

Charge calculation

For the charge calculation, first the ligand's geometry has to be optimized, for which the HF 6-31G(d) level of theory is used. This is done using Gaussian '09. The 5′- and 3′-O ends have to be capped, e. g. with a $-PO_2OMe$ group. The gas phase HF (or MP2) level of theory is used to be in compliance with the original AMBER force field generation. If several stable conformations exist, these can all be used for the charge calculation in the next step. See for example an AMBER online tutorial for a dye covalently attached to the 5′-end of a DNA strand.†

ligand_opt.com

```
# opt hf/6-31g(d) geom=connectivity
        ⋮
//coordinates//
        ⋮
```

Once a stable conformation is obtained (see Figure A.1 and A.2), ESP (ElectroStatic Potential) charges are calculated by fitting the electrostatic potential of the system onto points positioned on several layers around each atom. For future MD simulations see also the RED tools.[311]

†http://ambermd.org/tutorials/advanced/tutorial1/

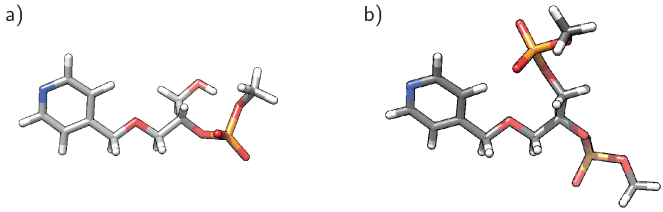


Figure A.1.: Optimized structure (MP2/6-31G(d)) for the ligand (S)-$\mathbf{L^2}$ with a) a $-PO_2OMe$ cap at the 3′-end, for 5′ DNA modification; b) a $-PO_2OMe$ cap each at the 3′- and the 5′-end, for internal DNA modification.)

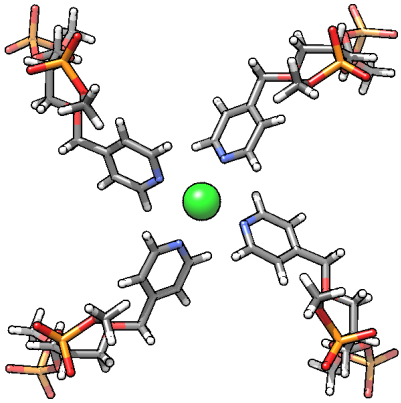


Figure A.2.: Optimized structure (UHF/6-31G(d)/LANL2DZ) for the internal [Cu($\mathbf{L^{2}}^{S}$)$_4$] fragment.

ligand_opt_esp.com

```
# hf/6-31g(d) guess=(read,only) geom=(allcheck) Integral(
    Grid=UltraFineGrid) pop=(minimal,mk) iop
    (6/41=10,6/42=17,6/50=1) scf=xqc
        ⋮
//coordinates//
        ⋮
output.gesp
```

The gas phase HF level of theory is used to be in compliance with the original AMBER force field generation. The `iop` values define the number of concentric layers (10) and points on each layer (the value of 17 gives around 2500 points per atom) for the ESP charge fitting. The `6/50=1` enables the generation of the output file. In case a transition metal ion is present, the input file has to be slightly changed.

complex_opt_esp.com

```
# uhf/genecp guess=(read,only) geom=(allcheck) Integral(Grid
    =UltraFineGrid) pop=(minimal,mk,readradii) iop
    (6/41=10,6/42=17,6/50=1) scf=xqc
        ⋮
//coordinates//
        ⋮
Cu 0
LANL2DZ
****
C N O P H 0
6-31g(d)
****

Cu 0
LANL2

Cu 2.0

output.gesp
```

Here the `readradii` option is needed, as for copper no radius is stored. The value itself has to be slightly bigger than the van-der-Waals radius. The LAN2DZ basis set together with the LANL2 effective core potential was used for the copper ion, but of course any other suitable

basis set may be used. The so obtained **output.gesp** file is now used for a RESP charge fitting, as the ESP charges depend on the basis set used, whereas the RESP charges do not. First, a **constraint.respain** file has to be created in which charge constraints will be defined.

constraint.respain

```
//Phosphate (-PO2-OMe) is constrained to a charge of -0
    .6921.
//Atoms -PO2-O- are constrained to their original value in
    the amber force field.
//predefined charges in a format of (CHARGE partial_charge
    atom_ID atom_name)
//resp charge from Cupy4 for Cu
CHARGE 0.344973 1 Cu1
//charges for 3' end
CHARGE 1.1659 98 P1
CHARGE -0.7761 102 O13
CHARGE -0.7761 106 O17
CHARGE -0.4954 110 O21
        ⋮
//charges for 5' end
CHARGE 1.1659 130 P5
CHARGE -0.7761 134 O25
CHARGE -0.7761 138 O29
CHARGE -0.5232 142 O33
        ⋮
//charge groups in a format of (GROUP num_atom net_charge),
//more than one group may be defined.
//atoms in the group in a format of (ATOM atom_ID atom_name)
//
// 3'end -PO2OMe constrained to -0.6921
//group 1 3'end
GROUP 8 -0.6921
ATOM 98 P1
ATOM 102 O13
ATOM 106 O17
ATOM 110 O21
ATOM 114 C37
ATOM 115 H45
ATOM 116 H46
ATOM 117 H47
        ⋮
```

In the AMBER force field, all internal nucleoside building blocks have a charge of −1 each. The 5′- and 3′-nucleosides together have a charge of −1, which is distributed over the two ends, namely −0.6921 for the 3′-end, and −0.3079 for the 5′-end. The same approach is used for the uncoordinated, free ligand. In case of the coordinated ligand, an additional positive charge of +0.5 for each ligand arises from the bound Cu^{2+}-ion (see Figure A.3).

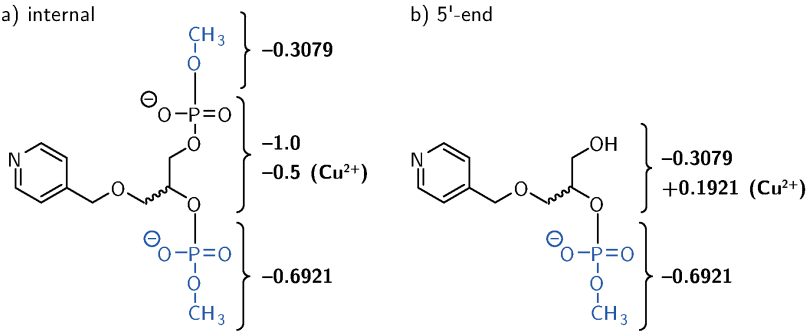


Figure A.3.: Methyl capped ligands for AMBER charge constraints scheme.

In the first round of RESP charge fitting, the Gaussian output file is first converted to a structure file (**.ac**) with `antechamber` and a charge file (**.dat**) with `espgen`. Together, they are converted to a RESP job file using `respgen`. Subsequently, the fitting is invoked with `resp` which also takes into account the constraints as defined above in the **constraint.respain** file.

```
antechamber -i output.gesp -fi gesp -o CuL24_i.ac -fo ac -at
     amber -pf y
espgen -i output.gesp -o CuL24_i.dat
respgen -i CuL24_i.ac -o CuL24_i.respin1 -f resp1 -a
    constraint.respain
mv QIN CuL24_i.qin
//the output file QIN is renamed to avoid ambiguities.//
resp -O -i CuL24_i.respin1  -o CuL24_i.respout1 -e CuL24_i.
    dat -t CuL24_iresp1.qout -p CuL24_iresp1.chg -q CuL24_i.
    qin
```

Afterwards, a second round of fitting is invoked, in which all equivalent hydrogen atoms like in methylene or methyl groups are constrained to identical values.

```
respgen -i CuL24_i.ac -o CuL24_i.respin2 -f resp2 -a
    constraint.respain
resp -O -i CuL24_i.respin2  -o CuL24_i.respout2 -e CuL24_i.
    dat -t CuL24_iresp2.qout -p CuL24_iresp2.chg -q
    CuL24_iresp1.qout
```

An antechamber file is then created with the new RESP charges. As this is a plain text file, it can be easily converted into a **.pdb** file by changing e. g. the atom names and types. Additionally, an **.frcmod** file can be created, in which missing force constant parameters are listed (see the next subsection).

```
antechamber -i CuL24_i.ac -fi ac -o CuL24_i_resp.ac -fo ac
    -c rc -cf CuL24_iresp2.qout -at amber -pf y
parmchk2 -i CuL24_i_resp.ac -o CuL24_i.frcmod -f ac -p /home
    /clever-lab/amber14/dat/leap/parm/parm99.dat
```

For the atom names any can be used, as long as they exactly match those as defined in the **dna.rtp** entries. However, for the ligand backbone, the same nomenclature as for the native nucleosides was used. In case of the atom types, the best matching ones from the AMBER force field are used (Figure A.4).

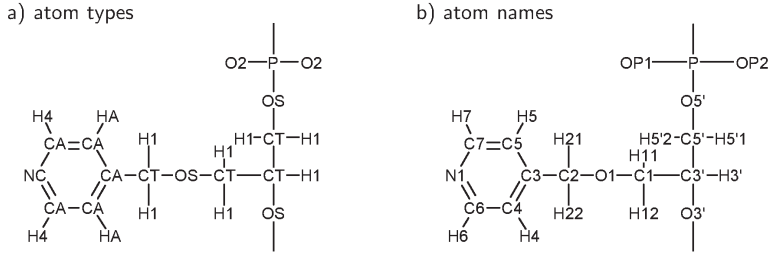


Figure A.4.: Ligand $\mathbf{L^2}$ AMBER atom types and names.

Force constants

Once the charges are derived for the ligand, both for the free and Cu^{2+}–bound state, the next step is to identify and parametrize missing force constants. These include angles in the unbound ligand, but especially those angles and bonds involved in the Cu^{2+} coordination.

Values for the bond lengths and angles can be taken from the optimized structures. Finding the corresponding force constant values is not so trivial. Here, force constants for the bonds involved in metal ion coordination were estimated based on literature values (see section 9.1.10 on page 170). For future MD simulations, the use of the VFFDT program or a more thorough force constant determination is encouraged.[312]

A.1.2. Molecular Dynamics Simulation

Once the charges and force constants have been obtained, they have to be integrated into the AMBER force field within Gromacs. The folder (.ff) with the *parmbsc0* force field is copied to a separate location and renamed (here **amber99sb_parmbsc0_L2.ff**).† In the **atomtypes.atp** file, new atoms/atom names have to be inserted. In the case of the square-planar copper coordination, this is necessary to avoid ambiguity between *trans* and *cis* *N*–Cu–*N* angles.

atomtypes.atp

```
                                                    ⋮
NC    14.01000 ;sp2 N in 6 memb.ring w/LP (ADE,GUA)
NX    14.01000 ;same as NC, new for copper coordination
NY    14.01000 ;same as NC, new for copper coordination
                                                    ⋮
```

In the file **dna.rtp** there is an entry for each DNA building block with atoms and their partial charges. New entries for the 5′ (`L25`, `L5X`, `L5Y`), internal (`L2I`, `LX`, `LY`), 3′ ligands, with and without copper(II) coordination, are inserted. One for the copper(II) ion is also included (`CUM`).

†For future MD simulations the *parmbsc1* force field may be used.[313]

dna.rtp

```
                                                    ⋮
[ LX ]
 [ atoms ]
     P      P             1.16590      1
   O1P     O2            -0.77610      2
     ⋮      ⋮                   ⋮      ⋮
   O3'     OS           -0.151263     27
 [ bonds ]
     P    OP1
     ⋮      ⋮
    C7     H7
  -O3'     P      ; connection to the next base.

[ CUM ]
 [ atoms ]
   CU      CU            0.344973      1
```

In the **ffbonded.itp** all new force constant parameters are added. Care must be taken to the units used (e.g. kcal vs kJ, see the Gromacs manual). Also, sometimes different definitions for the force constants are used.

ffbonded.itp

```
[ bondtypes ]
                                      ⋮
; New Cu-N bonds for L2
 CU   NX         1   0.22359     66944.0
 CU   NY         1   0.22359     66944.0
 CA   NX         1   0.13390    404174.4 ; same as CA-NC
 CA   NY         1   0.13390    404174.4 ; same as CA-NC
                                      ⋮
[ angletypes ]
                                      ⋮
; New angles necessary for ligand L2.
; Generated with parmchk2 and converted
; (force constant factor 8.368 = 2·cal/J)
 CA   CT   OS     1   109.500   418.400 ; same as CM-CT-OS
 CA   CT   H1     1   109.500   418.400 ; same as CA-CT-HC
 CA   CA   NC     1   121.500   585.760 ; same as CM-CA-NC
 CA   NC   CA     1   118.600   585.760 ; same as CA-NC-CQ
 H4   CA   NC     1   119.100   418.400 ; same as H4-CM-N*
```

```
; New angles necessary for the square planar copper
    coordination.
 NX   CU   NX     1   175.369   209.000 ; square planar, trans
 NX   CU   NY     1    89.906   669.000 ; square planar, cis
 NY   CU   NY     1   175.369   209.000 ; square planar, trans
 CA   NX   CU     1   121.266   585.760 ; square planar
 CA   NY   CU     1   121.266   585.760 ; square planar
 CA   CA   NX     1   121.500   585.760 ; same as CM-CA-NC
 CA   NX   CA     1   118.600   585.760 ; same as CA-NC-CQ
 H4   CA   NX     1   119.100   418.400 ; same as H4-CM-N*
 CA   CA   NY     1   121.500   585.760 ; same as CM-CA-NC
 CA   NY   CA     1   118.600   585.760 ; same as CA-NC-CQ
 H4   CA   NY     1   119.100   418.400 ; same as H4-CM-N*
                                    ⋮
[ dihedraltypes ]
                                    ⋮
; Improper dihedral to keep Cu square planar
NY   NY   CU   NX   4   180.00   50.0000   2 ; initial guess for kd
NX   NX   CU   NY   4   180.00   50.0000   2 ; initial guess for kd
                                    ⋮
 X   CA   NX    X   9   180.00   20.0832   2 ; same as X-CA-NC-X
 X   CA   NY    X   9   180.00   20.0832   2 ; same as X-CA-NC-X
```

The file **residuetypes.dat** is also copied to the current path and the names of the new ligands are inserted. The names must match those as defined in **dna.rtp**.

The next step is to generate a suitable PDB structure file. This can be accomplished by downloading a PDB file for a native G-quadruplex structure and manually inserting the ligand (e.g. with Chimera). The PDB file then has to be adjusted to be consistent with the PDB file format and the force field used within Gromacs: Atom names must match those in dna.rtp, atom types must be ATOM (except for the non-bonded ions like K^+ and Na^+), residue names must match those in **residuetypes.dat**. The residues have to be in the same chain identifier, otherwise the copper(II) coordination cannot be recognized. Once the PDB file meets all the requirements, it can be converted into a topology file (**.top**) readable by GROMACS.

```
gmx pdb2gmx -f structure.pdb -water tip3p
```

In the options prompt the desired force field is selected. If one has a modified force field, the corresponding folder should be in the path of the structure file. The generated topology file lacks all information about the transition metal coordination, therefore, these have to be included manually.[†] This includes Cu–*N* bonds, and all angles in which the Cu^{2+} ion is involved. The same is true for dihedrals, but requires force constants for the dihedral torsions as well.

Once the topology file is adjusted as necessary, the next steps comprise the generation of a solvent box around the molecule, addition of solvent molecules and neutralization of the net charge of the system. For the bounding box a dodecahedron is used, which is more efficient than a simple cube.

```
gmx editconf -f conf.gro -bt dodecahedron -d 1.0 -o box.gro
//cutoff is 1.0 nm around the solute molecule//
gmx solvate -cp box.gro -cs spc216.gro -p topol.top -o
    solvated.gro
//the box just created is filled with water molecules (type
    spc216)//
gmx grompp -f em.mdp -p topol.top -c solvated.gro -o em.tpr
//A job file is created by using the parameters from the .
    mdp file//
gmx genion -s em.tpr -o em_K.gro -p topol.top -pname K
    -neutral
//the system is neutralized with potassium cations//
```

The **em.mdp** file used in the second last step is already the parameter file for the first round of energy minimization. Only values differing from the default ones need to be listed (see the GROMACS manual for details).

em.mdp

```
integrator = steep ;using steepest descent algorithm
nsteps = 500 ;number of steps
nstlist = 50
cutoff-scheme = verlet
```

[†]In analogy to disulfide bridges in proteins, the parameters could also be automatically included by altering the specbond.dat file. This proved to be too difficult, as e. g. the bond lengths have to be within a certain margin.

```
vdw-type = cut-off
rvdw = 1.0      ;short-range van der Waals cutoff (in nm)
coulombtype = pme ;particle mesh ewald summation
rcoulomb = 1.0 ;short-range electrostatic cutoff (in nm)
```

In the next step a short energy minimization cycle is performed, so that overlong bonds etc., which may be present due to the manual insertion of the ligand, can relax to more useful values.

```
gmx grompp -f em.mdp -p topol.top -c em_K.gro -o em_K.tpr
gmx trjconv -s solvated.gro -f solvated.gro -o solvated.pdb
    -pbc atom -ur compact
gmx trjconv -s em_K.gro -f em_K.gro -o em_K.pdb -pbc atom
    -ur compact
//pdb file generation with conversion from the dodecahedron
    to cubic coordinates//
gmx mdrun -v -deffnm em_K
//actual energy minimization; takes all files starting with
    em_K as input//
```

A second round of energy minimization is invoked with a more reliable but slower algorithm, using an **em_cg.mdp** file.

em_cg.mdp

```
integrator = cg ;conjugate gradient alogrithm
nstcgsteep = 1000 ;steepest descent every number of steps
nsteps = 3000
nstlist = 50
cutoff-scheme = verlet
vdw-type = cut-off
rvdw = 1.0
coulombtype = pme
rcoulomb = 1.0
```

Again, the job file is created with `gmx grompp` and then executed with `gmx mdrun`.

```
gmx grompp -f em_cg.mdp -p topol.top -c em_K.gro -o em_cg_K.
    tpr
gmx mdrun -v -deffnm em_cg_K
gmx trjconv -s em_cg_K.gro -f em_cg_K.gro -o em_cg_K.pdb
    -pbc atom -ur compact
```

Following the energy minimization, the solvent molecules have to relax their positions around the solute molecule. During the two cycles, the atoms in the G-quadruplex structure are fixed, but with different constraint force constants.

pr.mdp

```
define = -DPOSRES ;switch for DNA position restraints, must
    be activated in the topology file
integrator = md
nsteps = 10000  ;nsteps * dt = 20 ps
dt = 0.002  ;2 fs
nstlist = 50
nstxout = 500 ;save coordinates every n * dt =  1.0 ps
nstlog = 500  ;update log file every n * dt =  1.0 ps
rlist = 1.0
coulombtype = pme
rcoulomb = 1.0
cutoff-scheme = verlet
vdw-type = cut-off
rvdw = 1.0
tcoupl = v-rescale  ;modified Berendsen thermostat
tc-grps = DNA Water_and_Ions  ;two coupling groups
tau-t = 0.1 0.1  ;time constant, in ps
ref-t = 298 298  ;reference temperatures in K
Pcoupl = Berendsen
pcoupltype = isotropic
tau-p = 2.0  ;time constant for coupling in ps
compressibility = 4.5e-5 4.5e-5 4.5e-5 0 0 0
ref-p = 1.0  ;reference pressure in bar
refcoord-scaling = all
nstenergy = 100
constraints = all-bonds  ;constrain all bonds
gen-vel = no  ;do not assign velocities from Maxwell
    distribution
```

Additionally, the ensemble is heated to 298 K and the pressure is increased to 1 atm (generation of the *NpT* ensemble). The file **pr2.mdp** differs from pr.mdp only in the number of steps (nsteps = 100000), that velocities are generated (option gen-vel=yes or remove the whole entry), and in the constraints (values in **posres.itp** changed from 1000 to 100).

```
gmx grompp -f pr.mdp -p topol.top -c em_cg_K.gro -o pr.tpr
gmx mdrun -v -deffnm pr
gmx grompp -f pr2.mdp -p topol2.top -c pr.gro -o pr2.tpr
gmx mdrun -v -deffnm pr2
gmx trjconv -s pr2.gro -f pr2.gro -o pr2.pdb -pbc atom -ur
    compact
```

The system is now ready for the molecular dynamics production run, for which the **run.mdp** file is used.

run.mdp

```
integrator = md
nsteps = 10000000
dt = 0.002
nstlist = 50
nstxout = 500
nstlog = 500
rlist = 1.0
coulombtype = pme
rcoulomb = 1.0
cutoff-scheme = verlet
vdw-type = cut-off
rvdw = 1.0
tcoupl = v-rescale
tc-grps = DNA Water_and_Ions
tau-t = 0.1 0.1
ref-t = 298 298
nstxtcout = 1000
nstenergy = 1000
Pcoupl = Berendsen
pcoupltype = isotropic
tau-p = 2.0
compressibility = 4.5e-5 4.5e-5 4.5e-5 0 0 0
ref-p = 1.0
refcoord-scaling = all
constraints = all-bonds
```

The run is invoked again by:

```
gmx grompp -f run.mdp -p topol.top -c pr2.gro -o run.tpr
gmx mdrun -v -deffnm run
```

During the simulation the G-quadruplex tends to diffuse out of the solvent box, giving rise to strange artefacts, due to the periodic boundary conditions. This can be removed by centering the molecule within the box.

```
gmx trjconv -s run.tpr -f run.xtc -o run_center.xtc -pbc mol
    -center
gmx trjconv -f run_center.xtc -o run_center_cell.xtc -pbc
    atom -ur compact
```

The generated trajectory (**.xtc**) can now be visualized with VMD.

A.1.3. Visualization

Trajectories in VMD

A PDB file can be loaded into VMD with the NewMolecule option, after which the trajectory is loaded (Load data into molecule). For details on the representation styles and making movies, see the VMD manual.

Graphics with Chimera

All graphics were visualized using the Chimera software. The respective PDB file generated with `trjconv` in Gromacs was loaded into Chimera and the following commands applied (if applicable/desired for the respective model and point of view).

```
background solid white
nucleotides sidechain atoms :DG:DA:DT:DG3:DA5
  //hide nucleotide objects, show atoms
fillring thick :DG:DA:DT:DG3:DA5
  //fill aromatic rings
fillring thin :LX:LY
  //name of the artificial compound, here ligand phosphate
fillring thin :L2I
color orange :LX:LY@C=
  // color carbon atoms of ligand orange
```

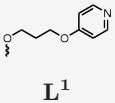
```
display :L2I@C=,N=,O=
  //hydrogens are omitted
ribcolor cornflower blue
  //ribbon color
color dark gray,a :DG:DA:DT:DG3:DA5
  // color all atoms
color blue :23
  //color for potassium or sodium ion
color green :24
  //color for copper ion
~display :K
  //hide all potassium ions
vdwdefine +0.5 :24
  //increase van-der-Waals radius of residue 24
represent sphere :23
  //change representation of residue 23 to sphere
transparency 80,a :DT:DA:DG:DG3:DA5:23
  //make atoms transparent (80%)
set dcstart 0.3
  //starting distance for fog
set dcend 1.0
  //end distance for fog
```

A.2. Oligonucleotide analytical data

In this section analytical data (ESI mass spectra and RP-HPLC traces) and metal complexation studies (UV thermal difference spectra, thermal denaturation curves, CD spectra) are shown for each of the synthesized oligonucleotides. Unless otherwise noted, samples were annealed in 10 mM lithium cacodylate buffer pH 7.2–7.3 (sodium cacodylate in case of oligonucleotides de1, de2 ($\mathbf{L^{1*}}$), de19), with 100 mM sodium chloride. Equivalents of metal salt added ($CuSO_4$, $NiSO_4$, $AgClO_4$) refer to the respective G-quadruplex, not the single strands. In case of the RP-HPLC traces, two different analytical C18 columns had to be used, therefore the given retention times cannot be directly compared. For detailed experimental descriptions see the experimental part (chapter 9).

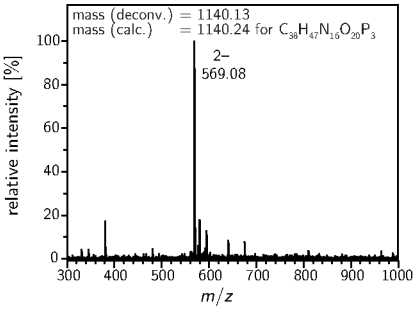
A.2.1. Oligonucleotide de1 (L^1)

Sequence (5′ → 3′):
LGGG

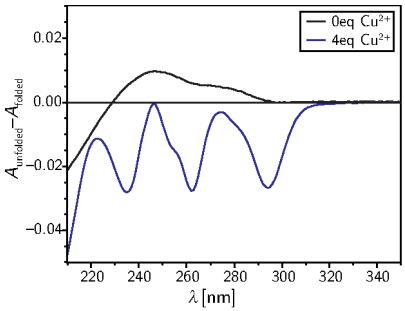


L^1

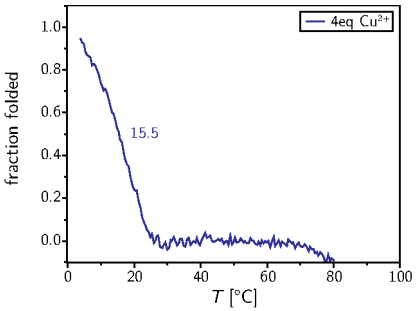
ESI(−)-MS



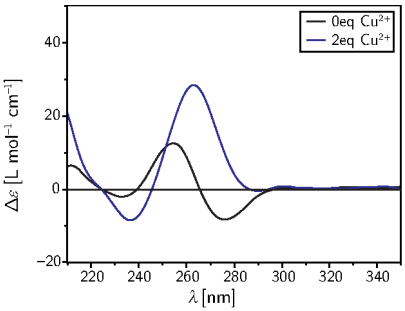
TDS



UV melting curves

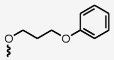


CD



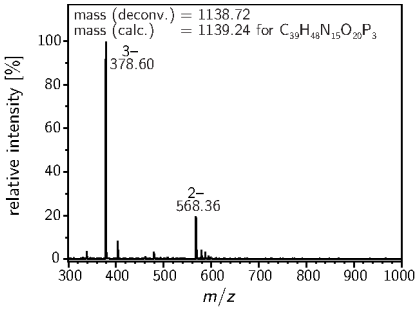
A.2.2. Oligonucleotide de1 (L^{1*})

Sequence (5′ → 3′):
LGGG

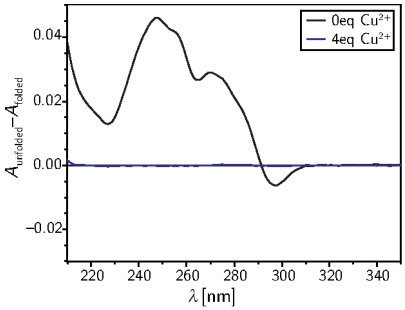


L^{1*}

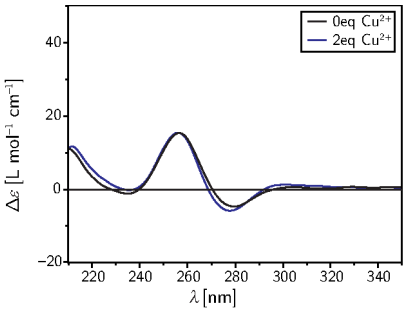
ESI(−)-MS



TDS

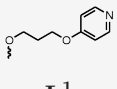


CD

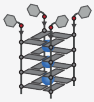


A.2.3. Oligonucleotide de2 (L^1)

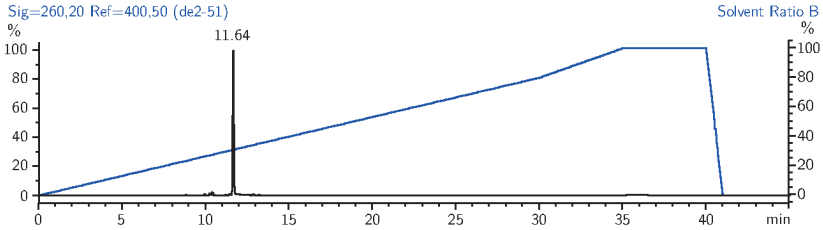
Sequence (5′ → 3′):
LGGGG



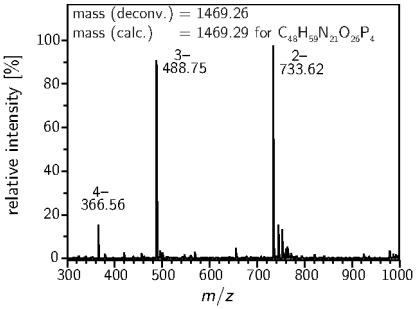
L^1



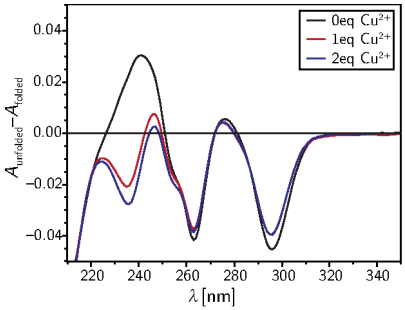
Analytical RP-HPLC



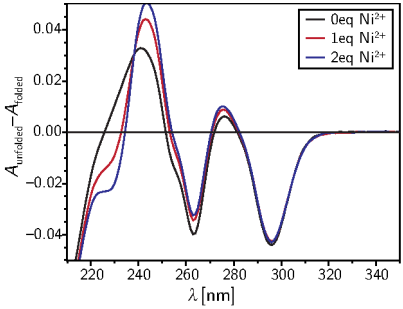
ESI(−)-MS



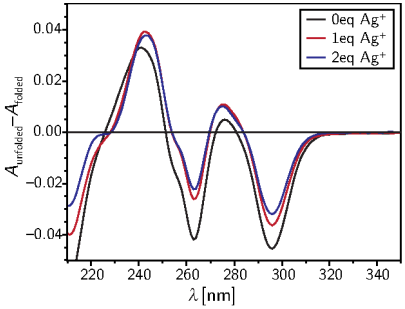
TDS (Cu^{2+})



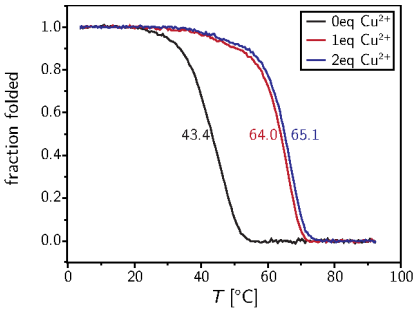
TDS (Ni^{2+})



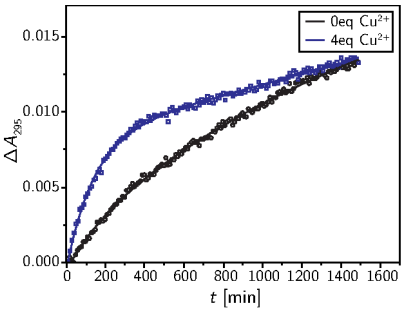
TDS (Ag^{+})



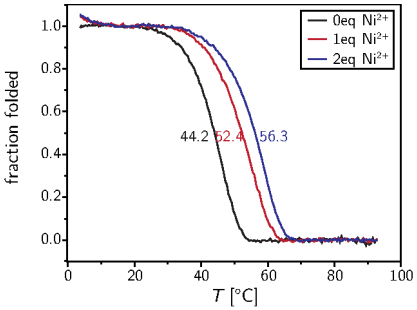
UV melting curves (Cu^{2+})



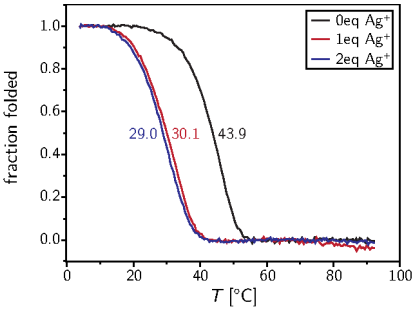
UV annealing curves (Cu^{2+})



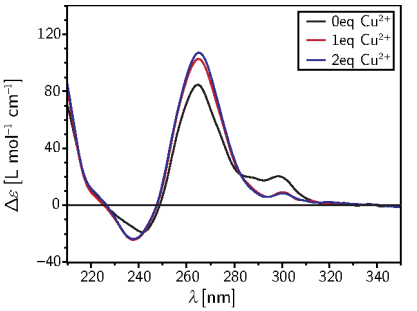
UV melting curves (Ni^{2+})



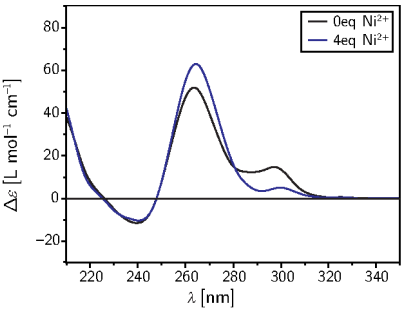
UV melting curves (Ag^{+})



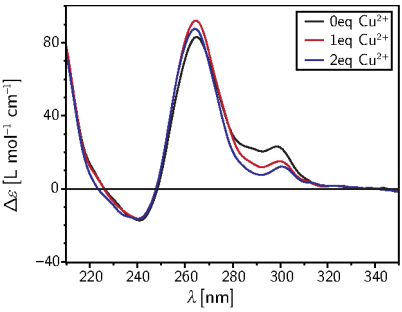
CD (Cu^{2+})



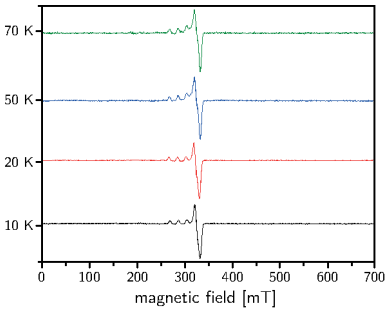
CD (Ni^{2+})



CD (Ag^+)

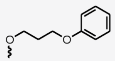


***T*-dep. EPR (Cu^{2+})**

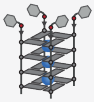


A.2.4. Oligonucleotide de2 (L^{1*})

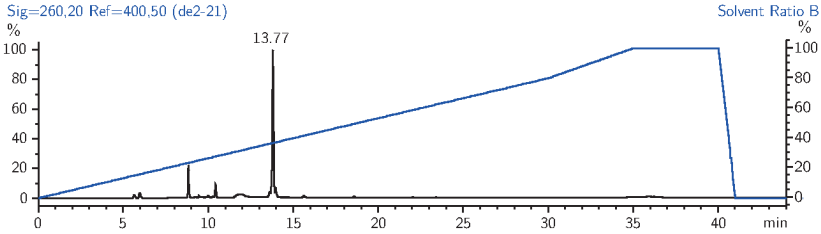
Sequence ($5' \rightarrow 3'$):
LGGGG



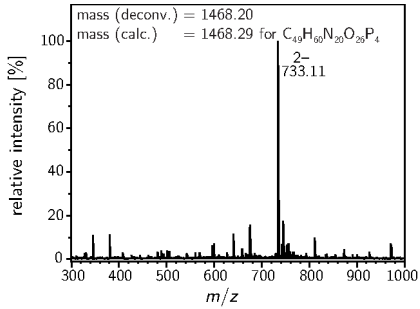
L^{1*}



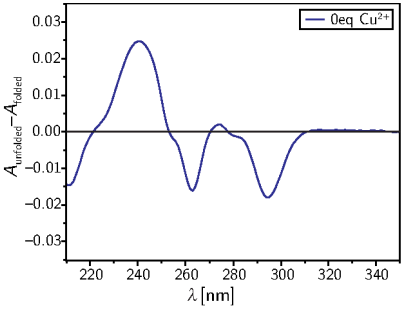
Analytical RP-HPLC



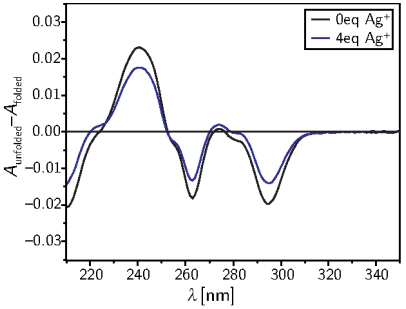
ESI(−)-MS



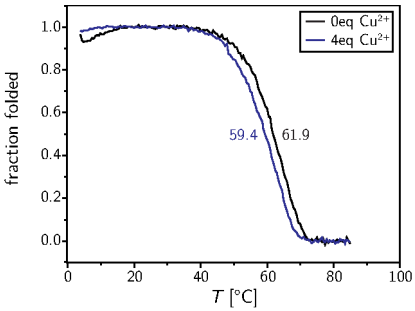
TDS (Cu^{2+})



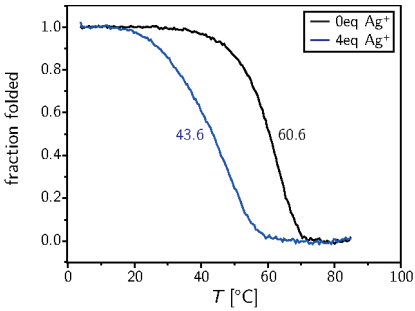
TDS (Ag^{+})



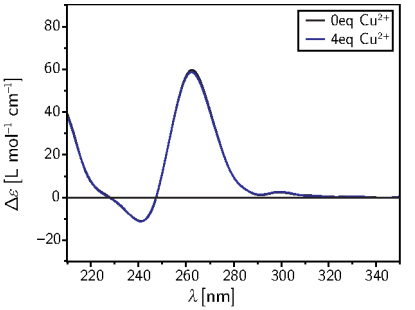
UV melting curves (Cu^{2+})



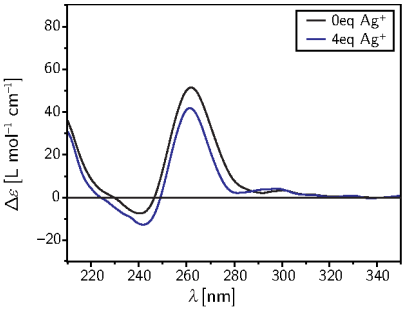
UV melting curves (Ag^{+})



CD (Cu^{2+})

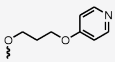


CD (Ag^{+})

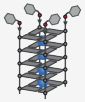


A.2.5. Oligonucleotide de19 (L^1)

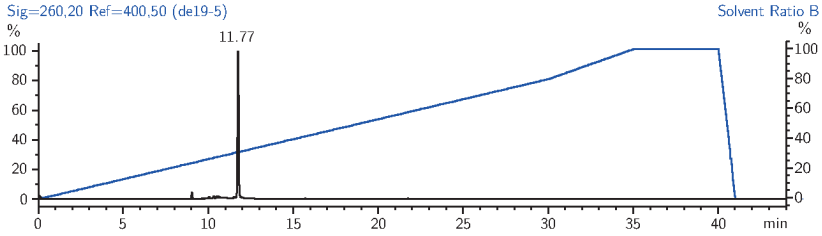
Sequence ($5' \rightarrow 3'$):
LGGGGG



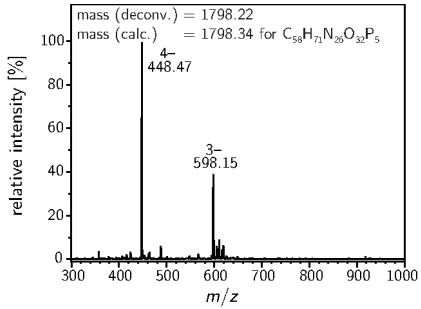
L^1



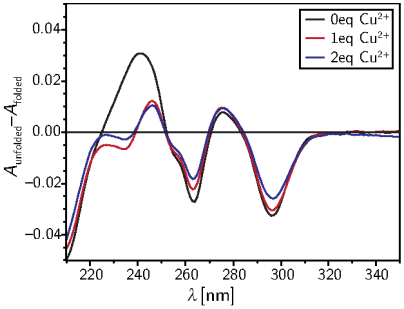
Analytical RP-HPLC



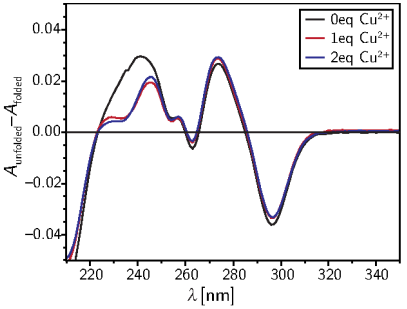
ESI(−)-MS



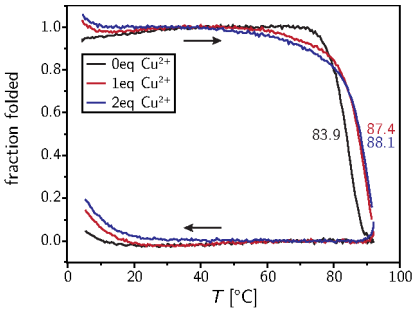
TDS (Na^+)



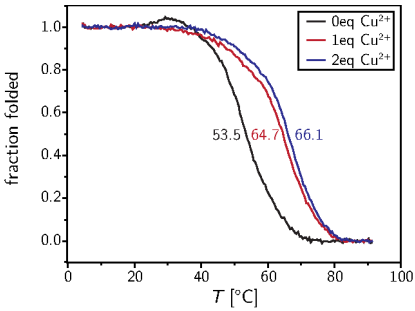
TDS (Cs^+)



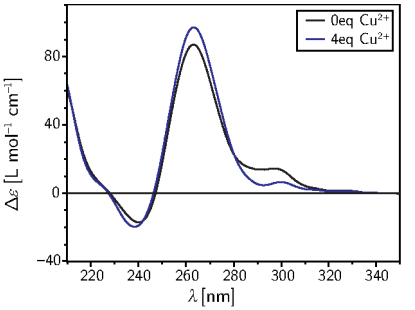
UV melting curves (Na^+)



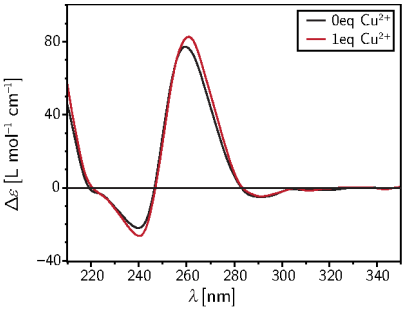
UV melting curves (Cs^+)



CD (Na^+)



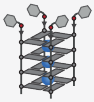
CD (Cs^+)



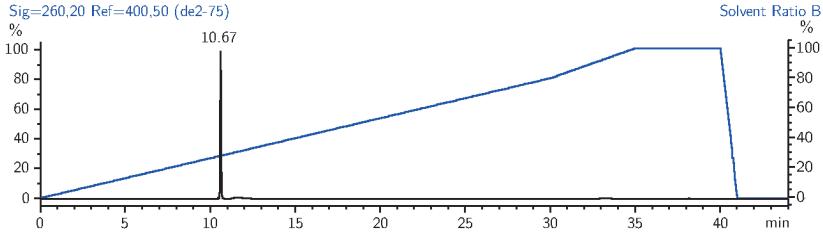
A.2.6. Oligonucleotide de2 (L^{1d})

Sequence (5′ → 3′):
LGGGG

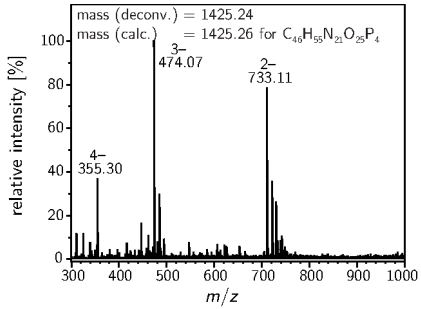
L^{1d}



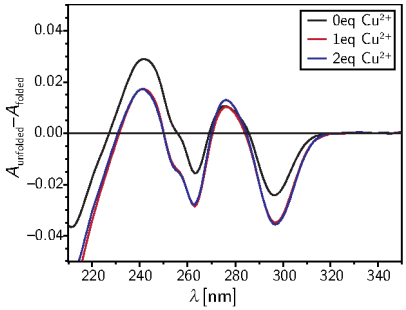
Analytical RP-HPLC



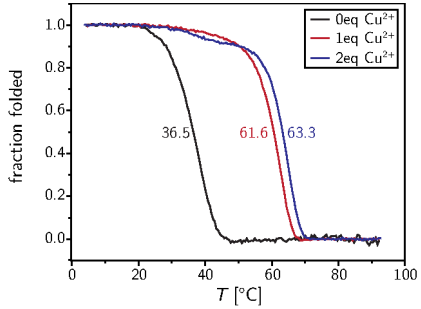
ESI(−)-MS



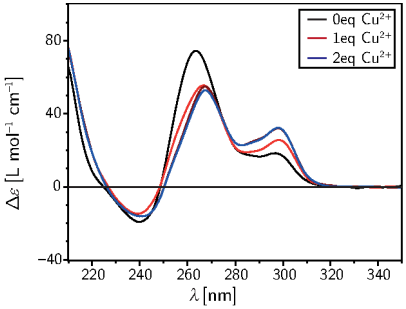
TDS



UV melting curves

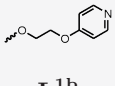


CD

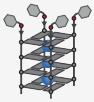


A.2.7. Oligonucleotide de2 (L^{1b})

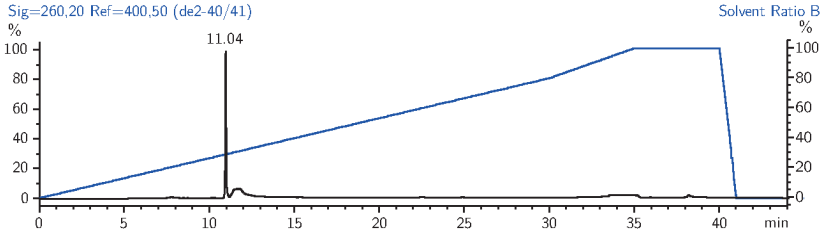
Sequence (5′ → 3′):
LGGGG



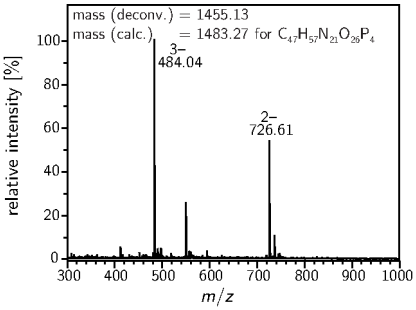
L^{1b}



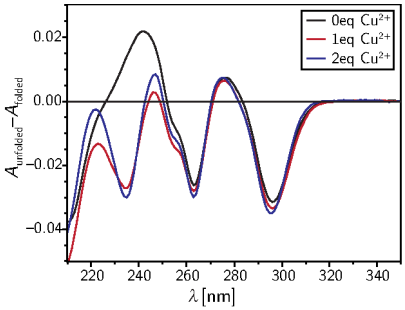
Analytical RP-HPLC



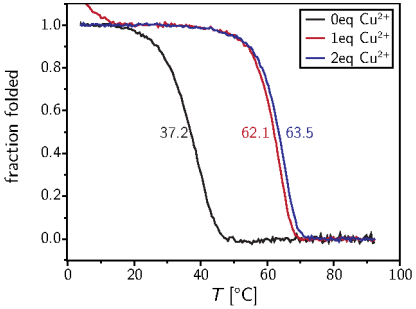
ESI(−)-MS



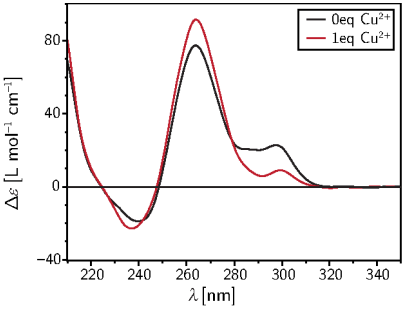
TDS



UV melting curves

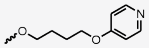


CD

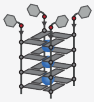


A.2.8. Oligonucleotide de2 (L^{1c})

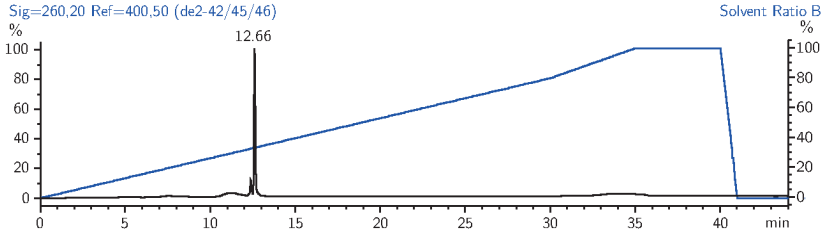
Sequence (5′ → 3′):
LGGGG



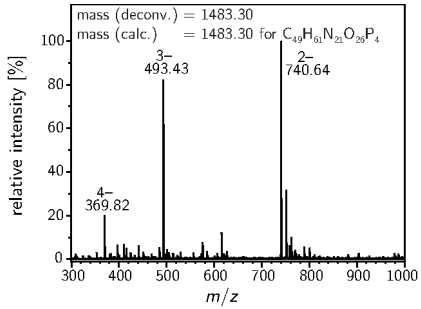
L^{1c}



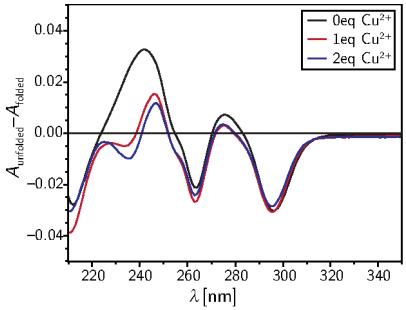
Analytical RP-HPLC



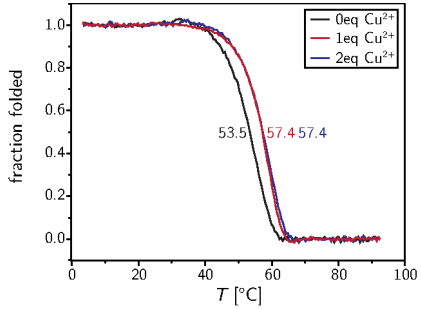
ESI(−)-MS



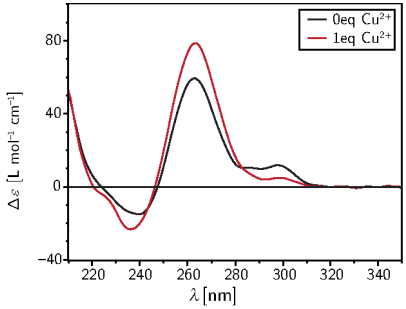
TDS



UV melting curves

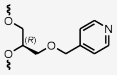


CD

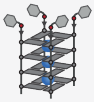


A.2.9. Oligonucleotide deg5R ((*R*)-L²)

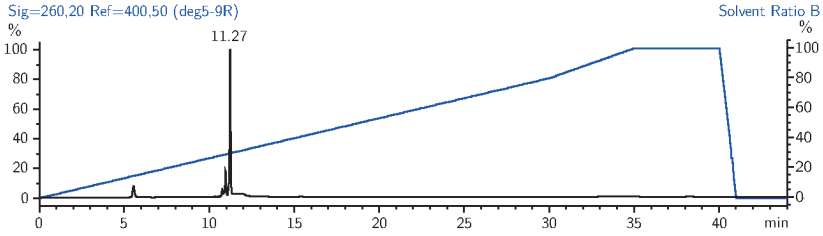
Sequence (5′ → 3′):
LGGGG



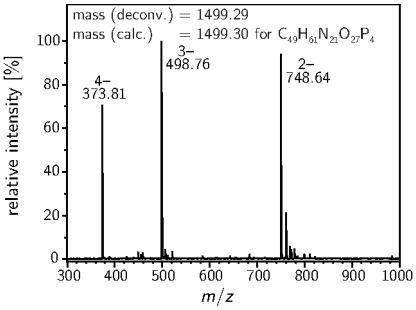
(*R*)-L²



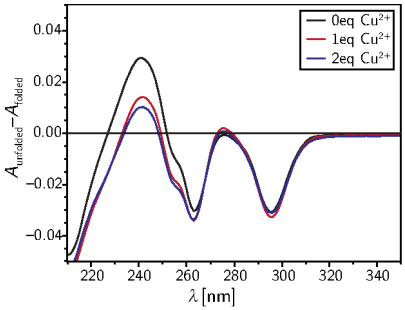
Analytical RP-HPLC



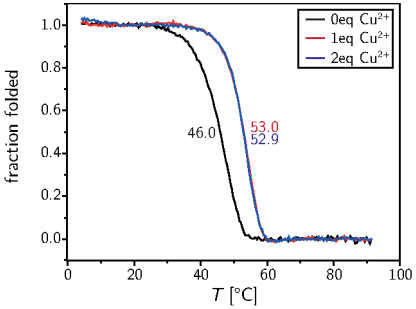
ESI(−)-MS



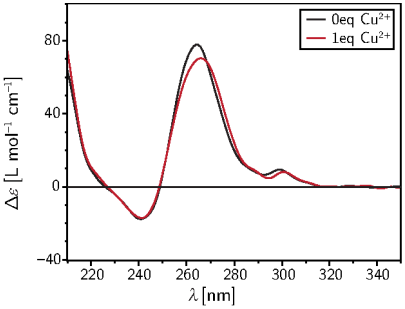
TDS



UV melting curves

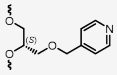


CD

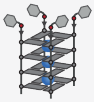


A.2.10. Oligonucleotide deg5S ((*S*)-L^2)

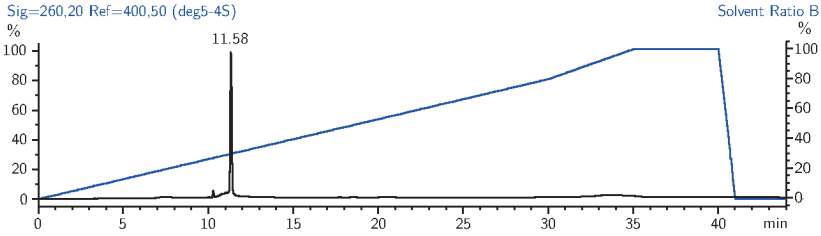
Sequence (5′ → 3′):
LGGGG



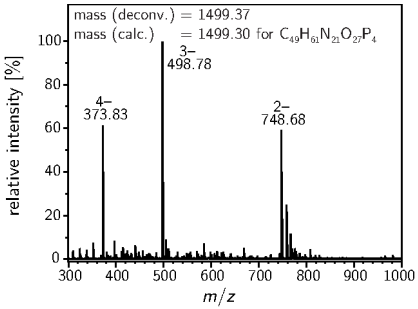
(*S*)-L^2



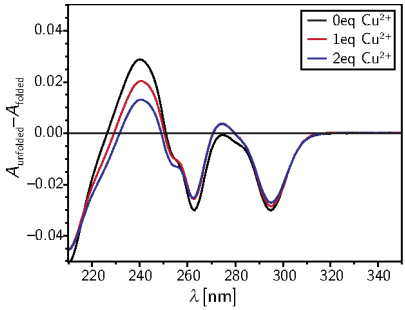
Analytical RP-HPLC



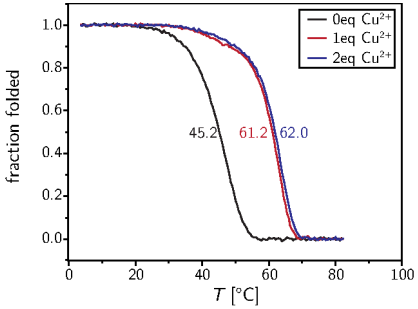
ESI(−)-MS



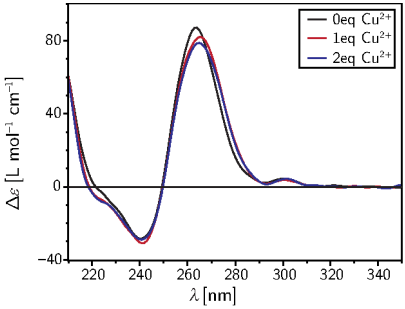
TDS



UV melting curves

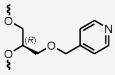


CD

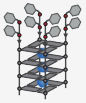


A.2.11. Oligonucleotide deg6RR ((*R*)-L^2)

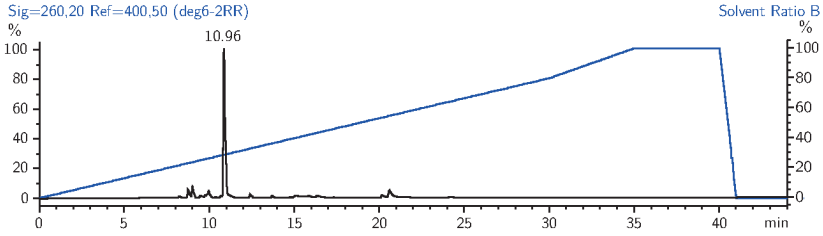
Sequence ($5' \rightarrow 3'$):
LLGGGG



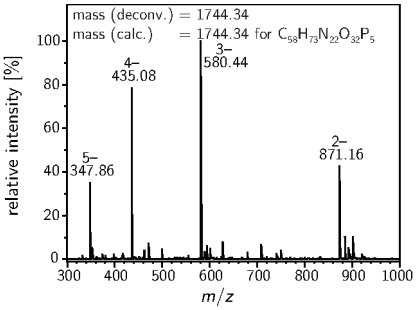
(*R*)-L^2



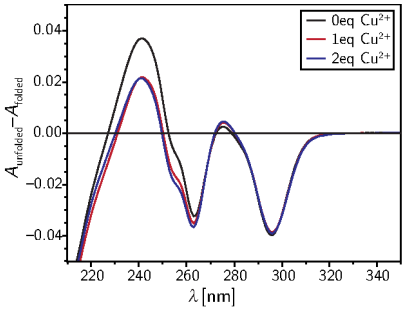
Analytical RP-HPLC



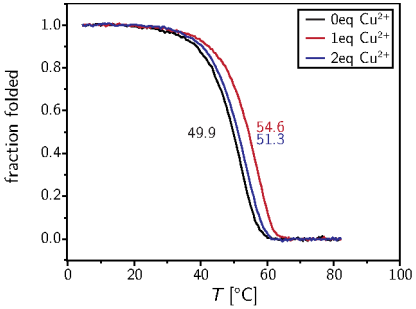
ESI(−)-MS



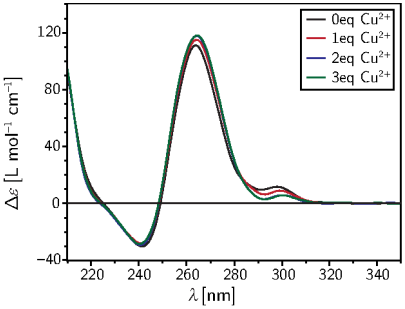
TDS



UV melting curves

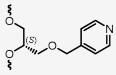


CD

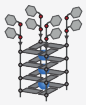


A.2.12. Oligonucleotide deg6SS ((*S*)-L^2)

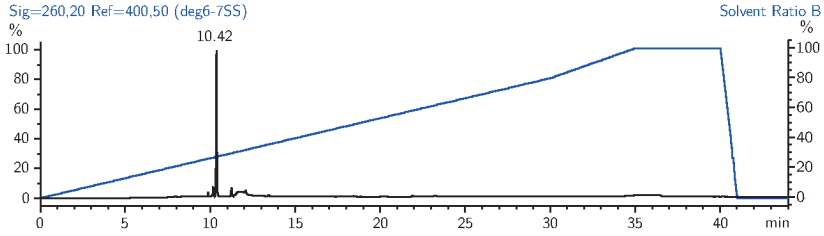
Sequence (5′ → 3′):
LLGGGG



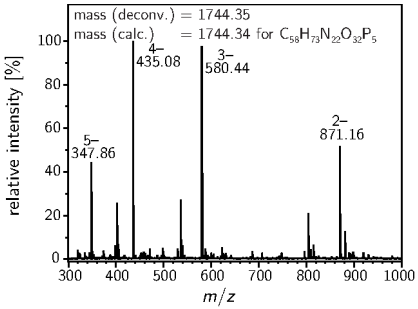
(*S*)-L^2



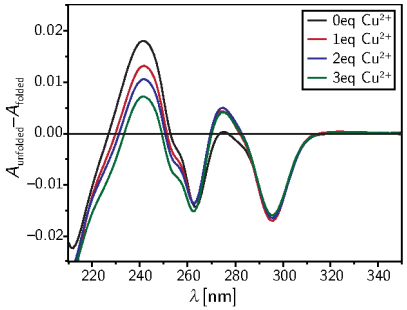
Analytical RP-HPLC



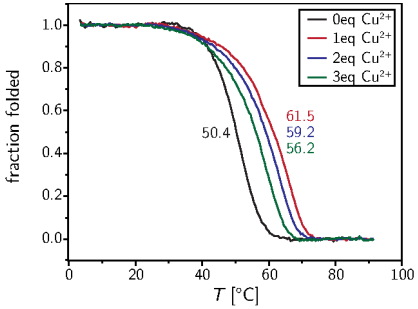
ESI(−)-MS



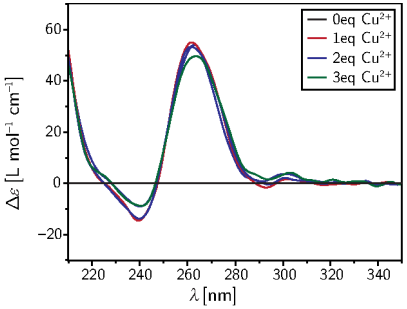
TDS



UV melting curves

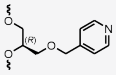


CD



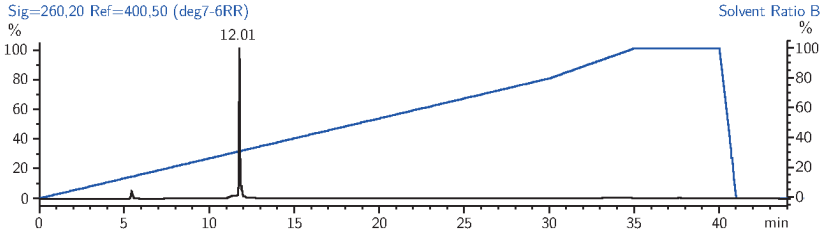
A.2.13. Oligonucleotide deg7RR ((*R*)-L^2)

Sequence (5′ → 3′):
LGGGLT

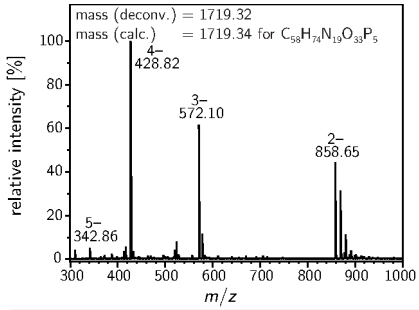


(*R*)-L^2

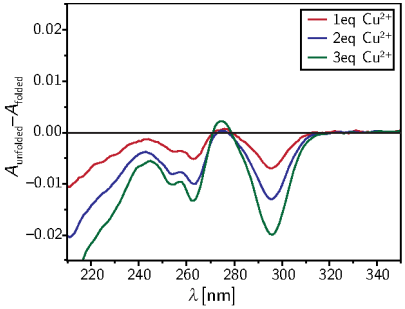
Analytical RP-HPLC



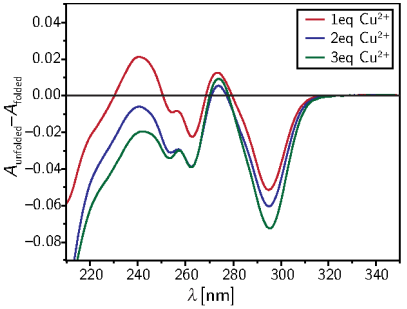
ESI(−)-MS



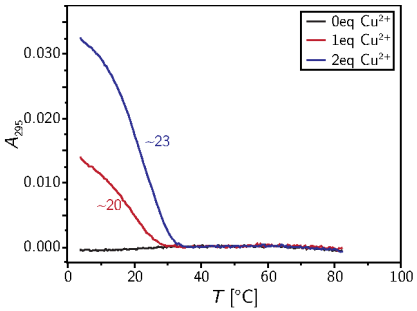
TDS (Na^+)



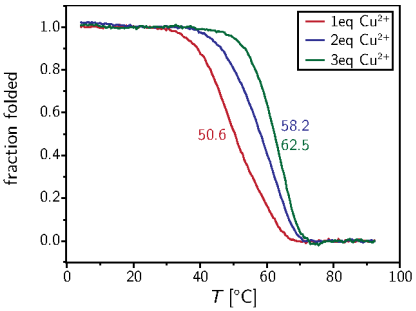
TDS (K^+)



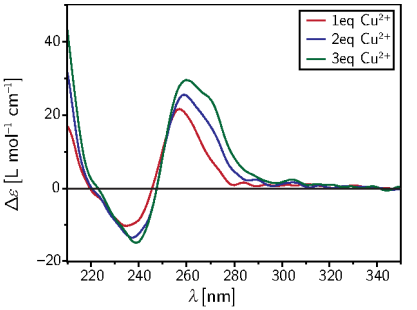
UV melting curves (Na^+)



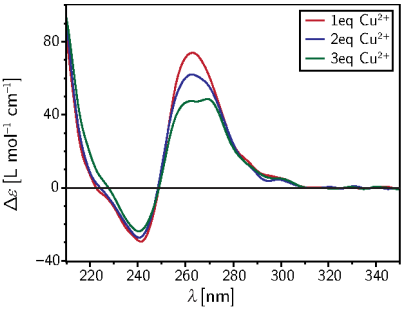
UV melting curves (K^+)



CD (Na^+)

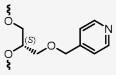


CD (K^+)



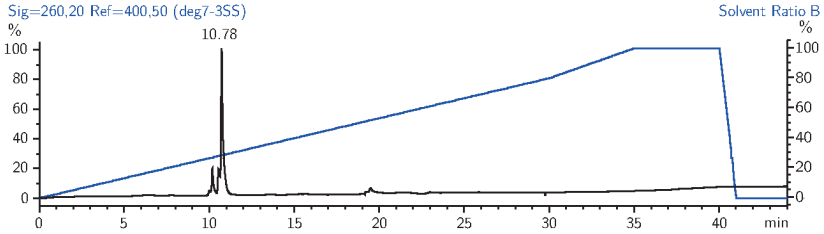
A.2.14. Oligonucleotide deg7SS ((*S*)-L^2)

Sequence (5′ → 3′):
LGGGLT

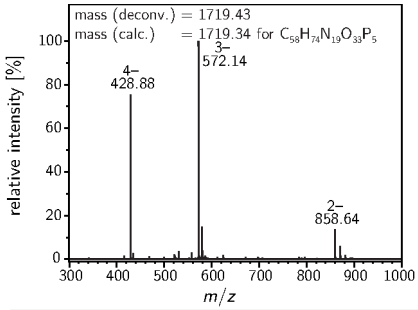


(*S*)-L^2

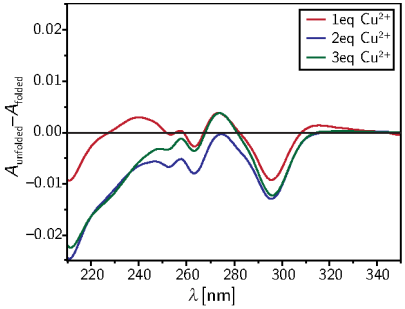
Analytical RP-HPLC



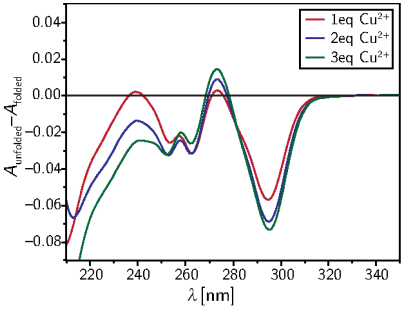
ESI(−)-MS



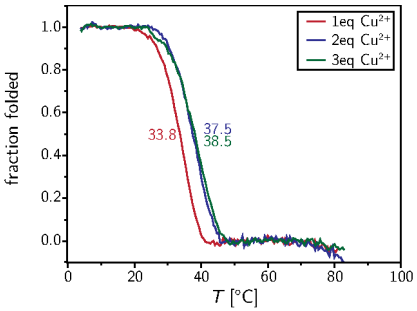
TDS (Na^+)



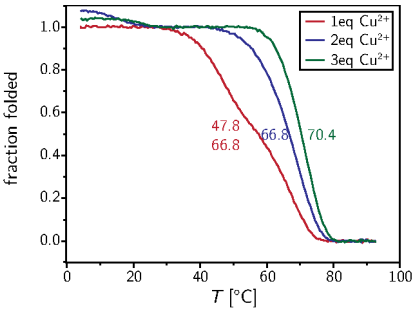
TDS (K^+)



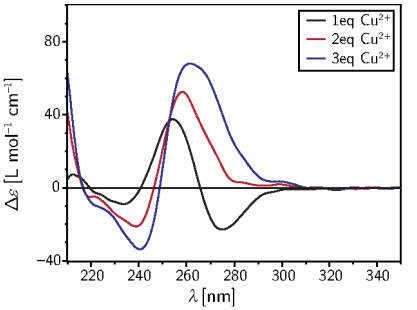
UV melting curves (Na^+)



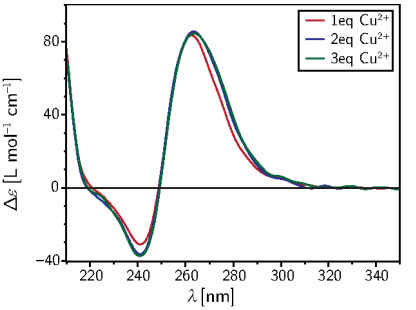
UV melting curves (K^+)



CD (Na^+)

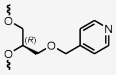


CD (K^+)



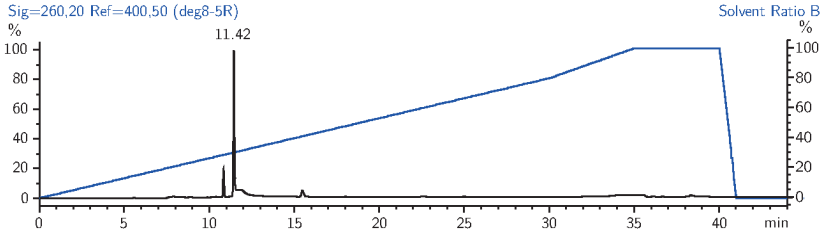
A.2.15. Oligonucleotide deg8R ((*R*)-L^2)

Sequence (5′ → 3′):
GGGGLT

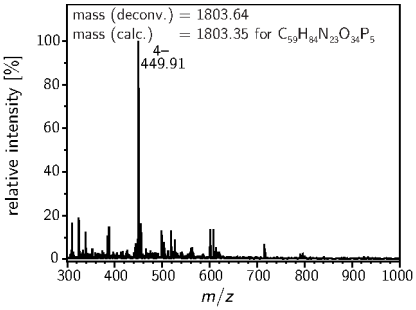


(*R*)-L^2

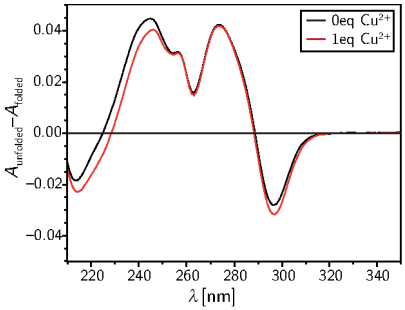
Analytical RP-HPLC



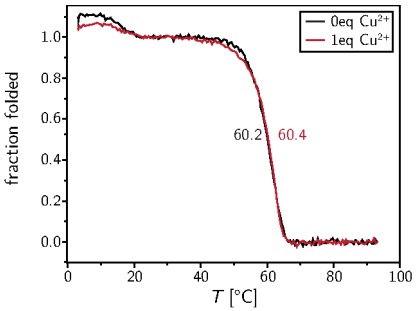
ESI(−)-MS



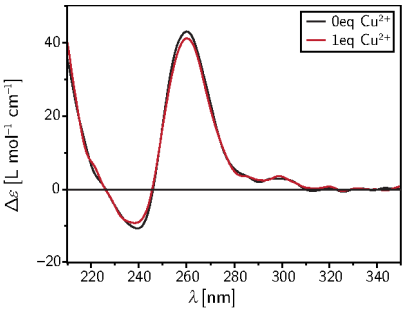
TDS



UV melting curves

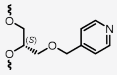


CD



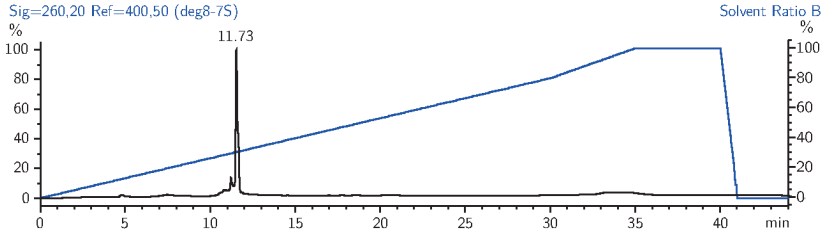
A.2.16. Oligonucleotide deg8S ((*S*)-L^2)

Sequence (5′ → 3′):
GGGGLT

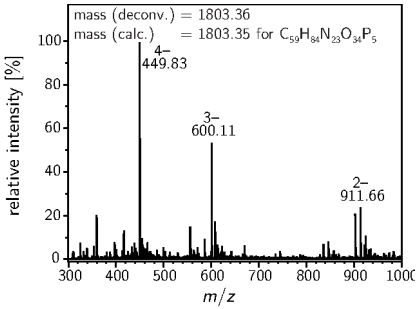


(*S*)-L^2

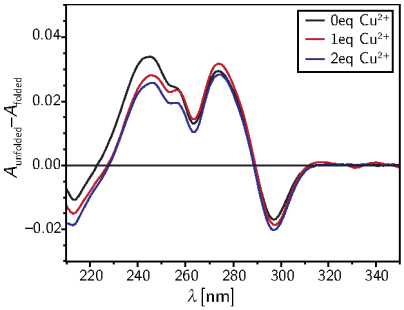
Analytical RP-HPLC



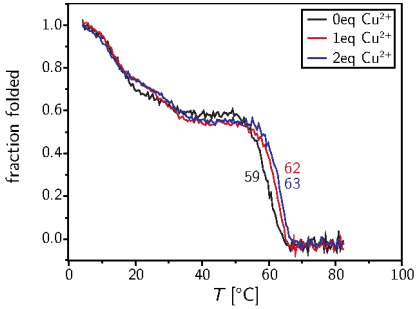
ESI(−)-MS



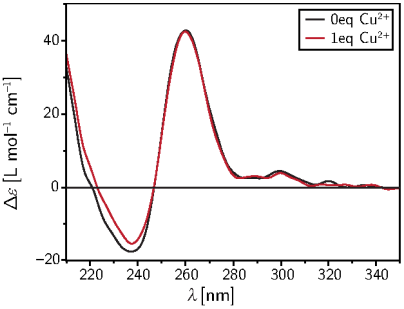
TDS



UV melting curves

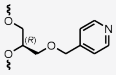


CD



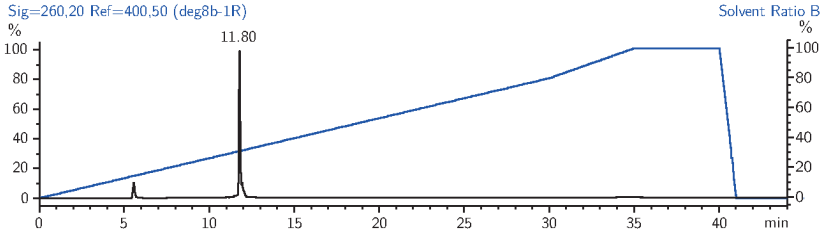
A.2.17. Oligonucleotide deg8bR ((*R*)-L^2)

Sequence (5′ → 3′):
TGGGGLT

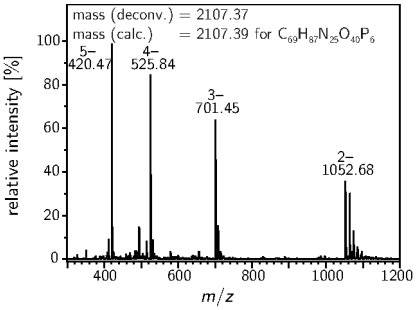


(*R*)-L^2

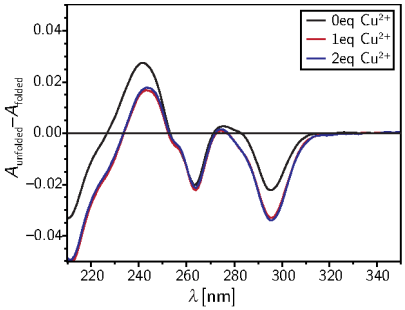
Analytical RP-HPLC



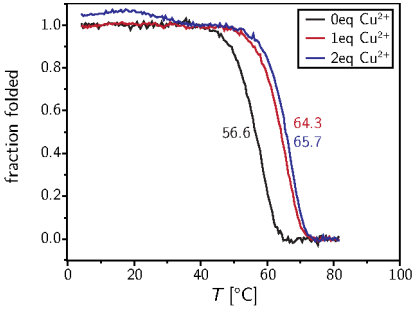
ESI(−)-MS



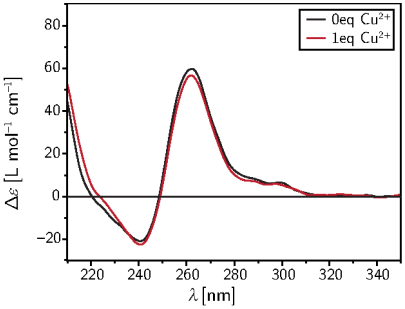
TDS



UV melting curves

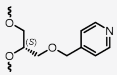


CD



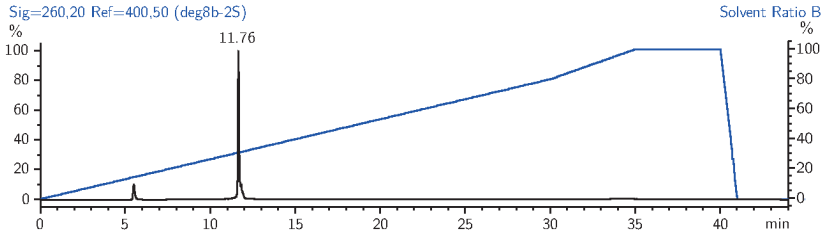
A.2.18. Oligonucleotide deg8bS ((*S*)-L^2)

Sequence (5′ → 3′):
TGGGGLT

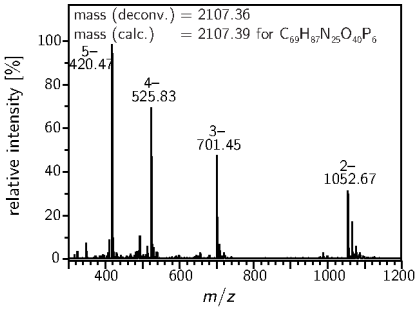


(*S*)-L^2

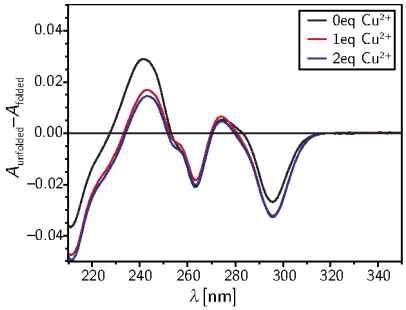
Analytical RP-HPLC



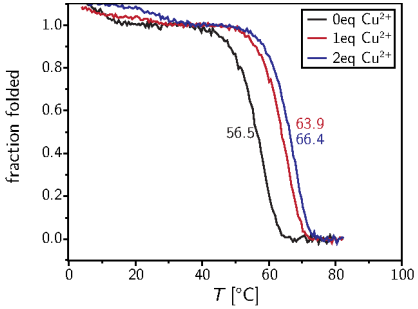
ESI(−)-MS



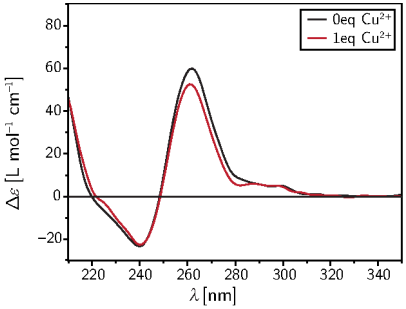
TDS



UV melting curves

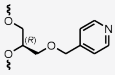


CD

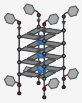


A.2.19. Oligonucleotide deg10RR ((*R*)-L^2)

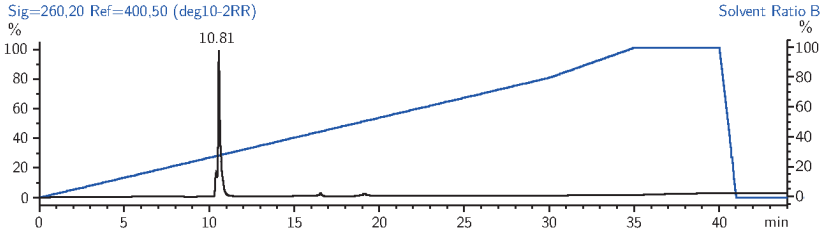
Sequence (5′ → 3′):
LGGGGLT



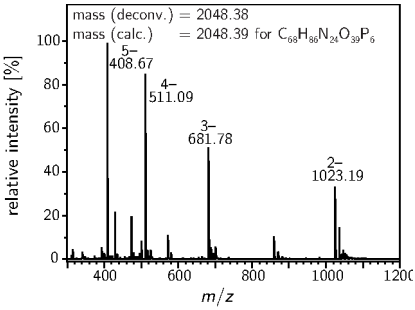
(*R*)-L^2



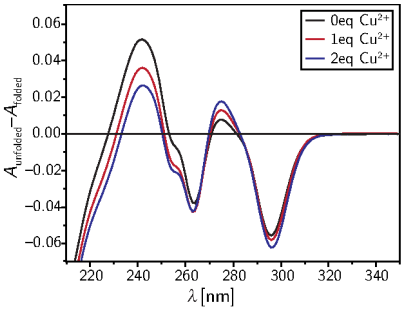
Analytical RP-HPLC



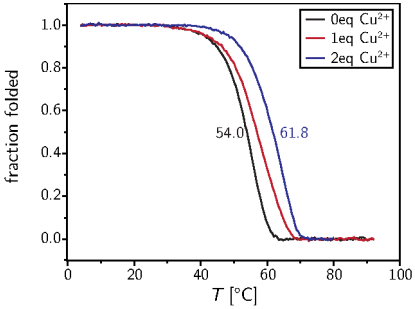
ESI(−)-MS



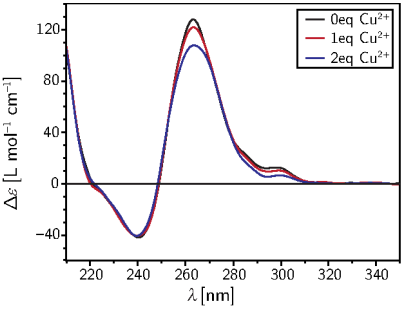
TDS



UV melting curves

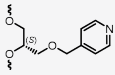


CD

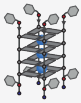


A.2.20. Oligonucleotide deg10SS ((*S*)-L^2)

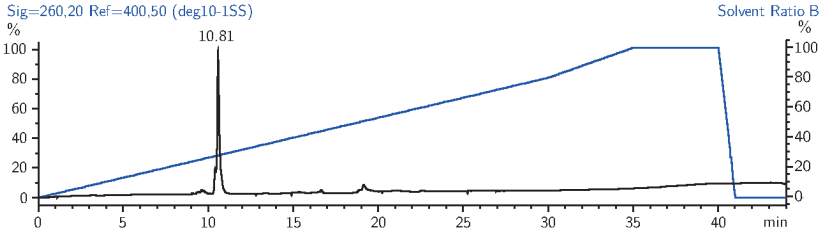
Sequence (5′ → 3′):
LGGGGLT



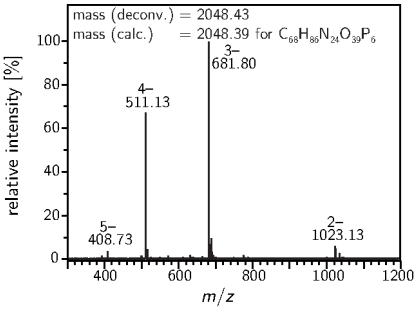
(*S*)-L^2



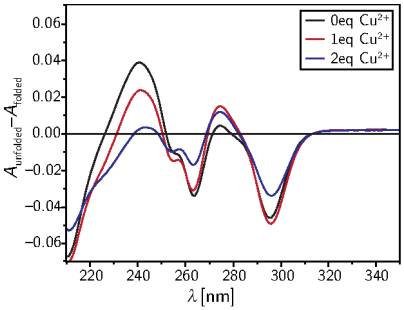
Analytical RP-HPLC



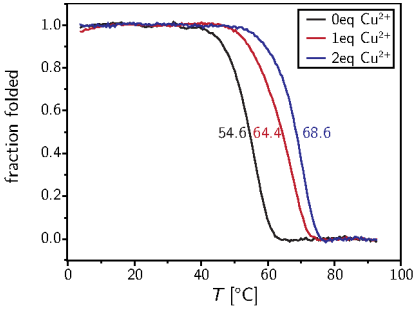
ESI(−)-MS



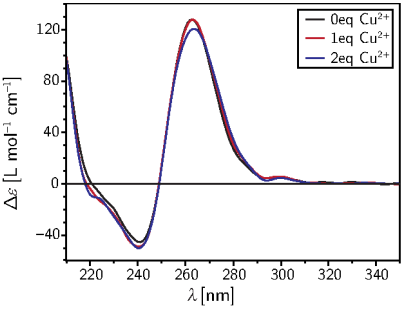
TDS



UV melting curves

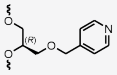


CD

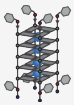


A.2.21. Oligonucleotide deg12RR ((*R*)-L^2)

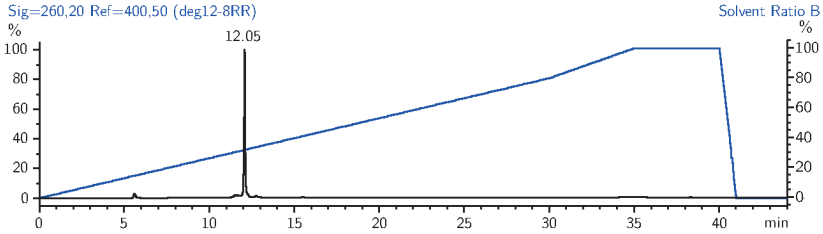
Sequence (5′ → 3′):
LGGGGGLT



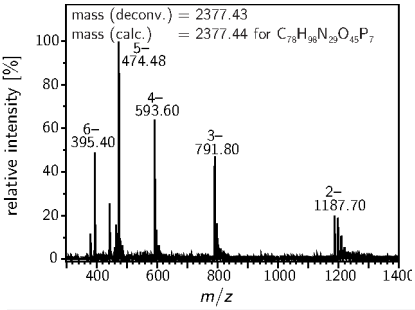
(*R*)-L^2



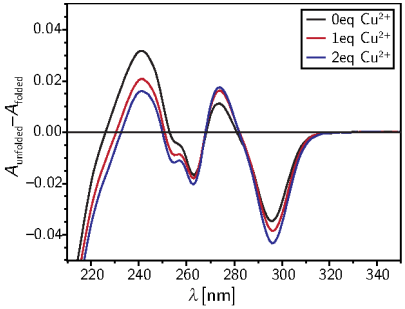
Analytical RP-HPLC



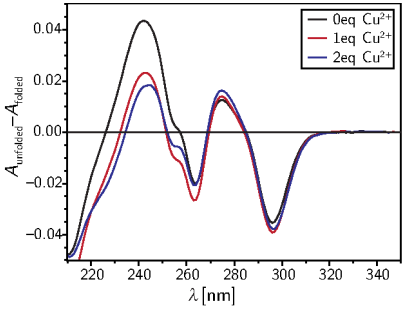
ESI(−)-MS



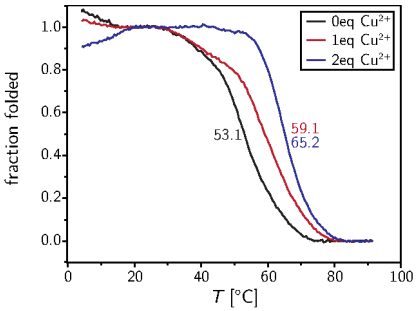
TDS (Cs$^+$)



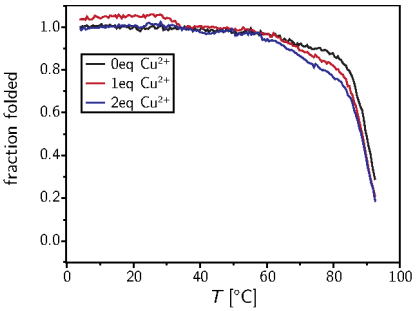
TDS (Na$^+$)



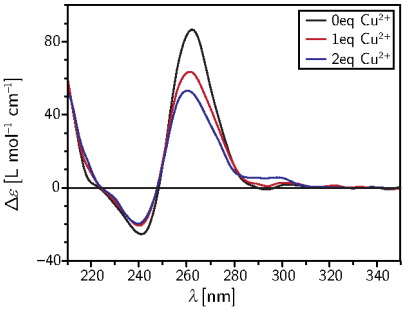
UV melting curves (Cs^+)



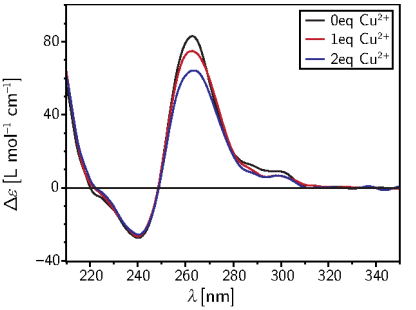
UV melting curves (Na^+)



CD (Cs^+)

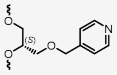


CD (Na^+)

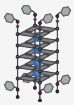


A.2.22. Oligonucleotide deg12SS ((*S*)-L^2)

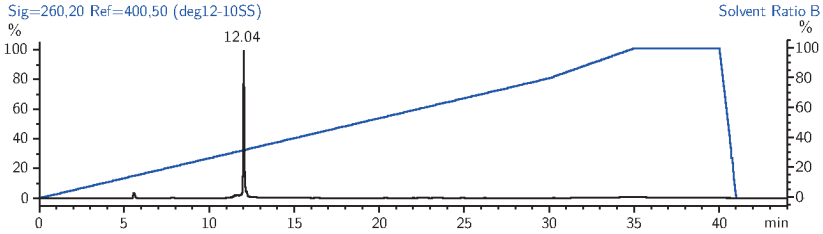
Sequence (5′ → 3′):
LGGGGGLT



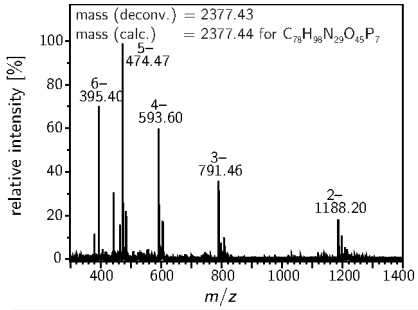
(*S*)-L^2



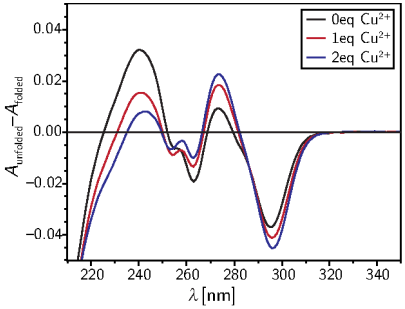
Analytical RP-HPLC



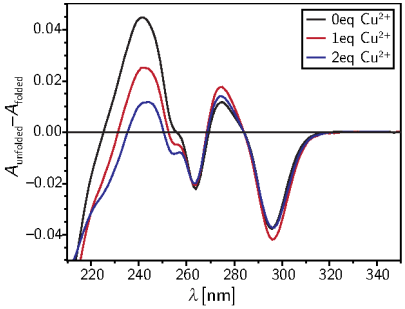
ESI(−)-MS



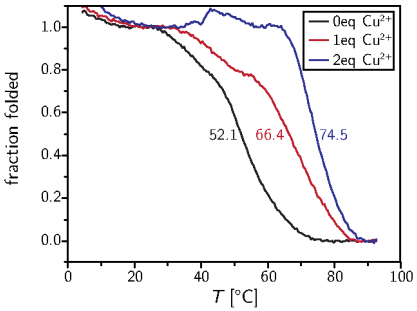
TDS (Cs^+)



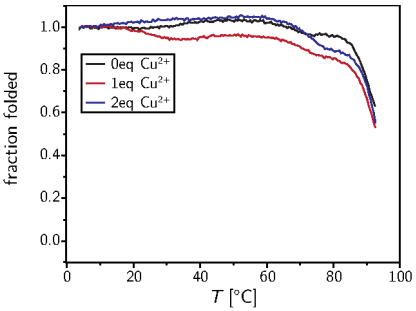
TDS (Na^+)



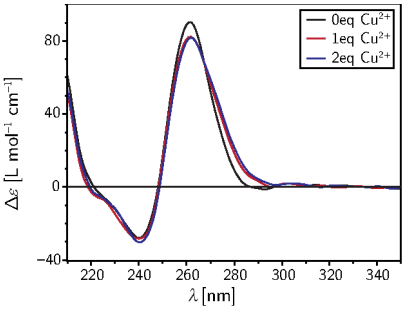
UV melting curves (Cs^+)



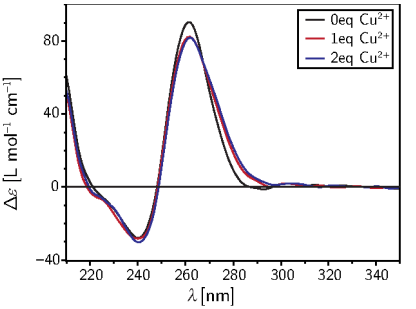
UV melting curves (Na^+)



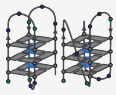
CD (Cs^+)



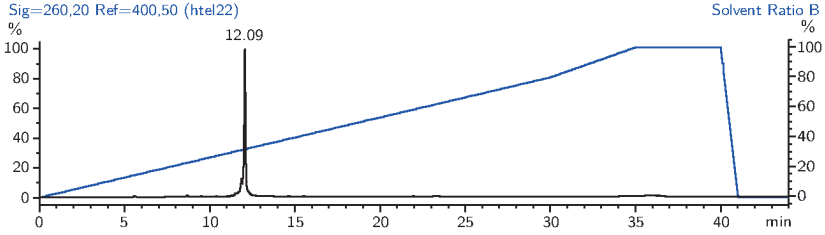
CD (Na^+)



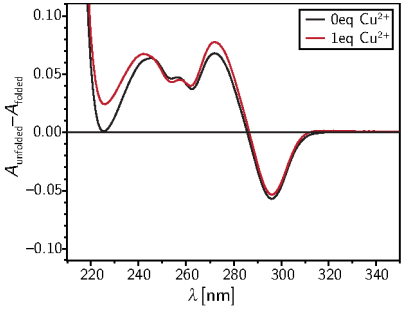
A.2.23. Oligonucleotide htel22

Sequence ($5' \rightarrow 3'$):		
AGGGTTAGGGTTAGGGTTAGGG	**no ligand**	

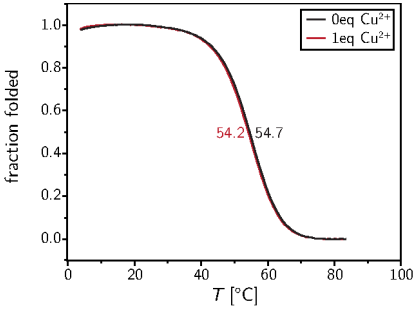
Analytical RP-HPLC



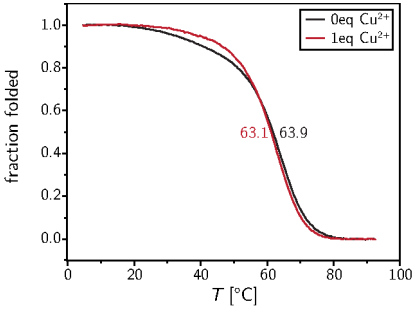
TDS (K^+)



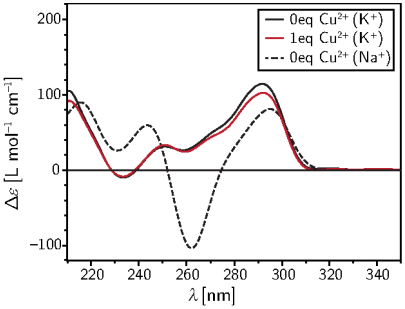
UV melting curves (Na^+)



UV melting curves (K^+)

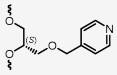


CD

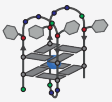


A.2.24. Oligonucleotide htel22-L^2$_4$b deg11-4S ((*S*)-L^2)

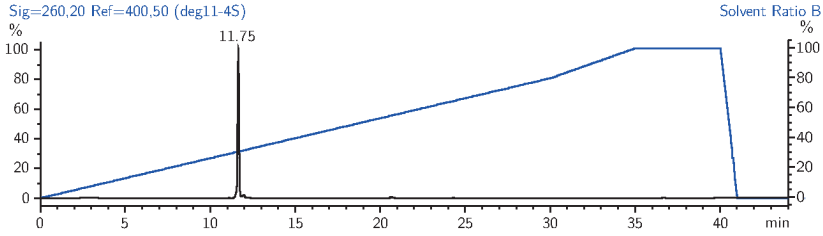
Sequence (5′ → 3′):
AGGLTTALGGTTAGGLTTALGG



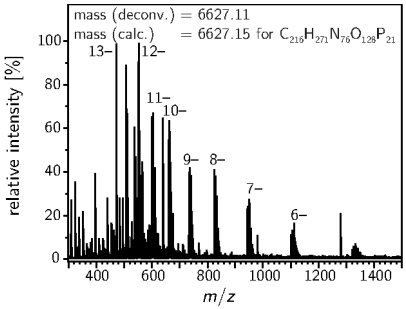
(*S*)-L^2



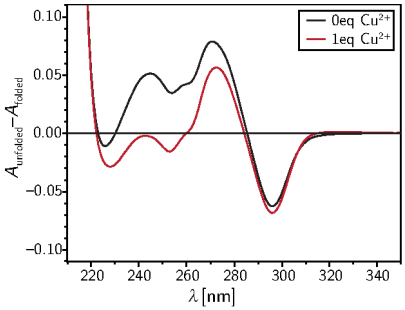
Analytical RP-HPLC



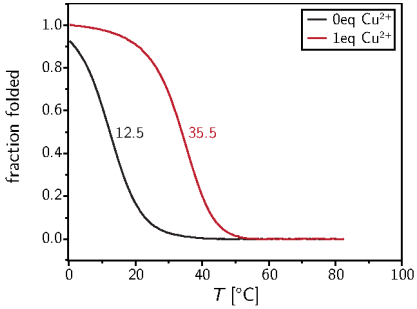
ESI(−)-MS



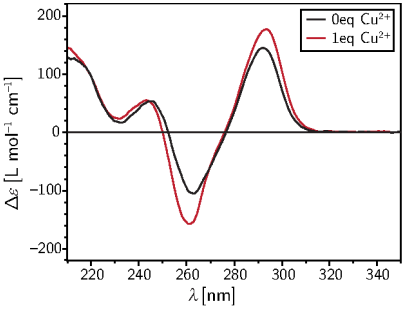
TDS (K$^+$)



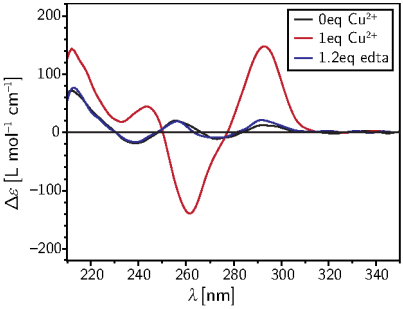
UV melting curves (K$^+$)



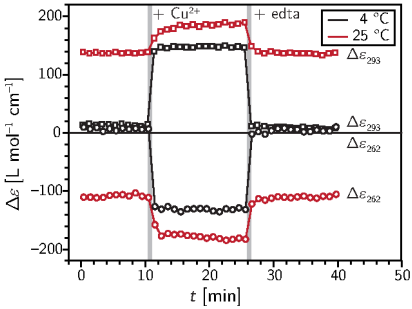
CD (K$^+$)



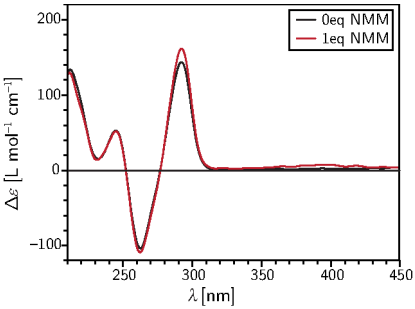
CD (K^+), edta, 25 °C



CD (K^+, *t*-dep. Cu^{2+}/edta

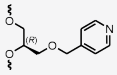


CD (K^+), +NMM

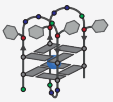


A.2.25. Oligonucleotide htel22-L^2$_4$b deg11-3R ((*R*)-L^2)

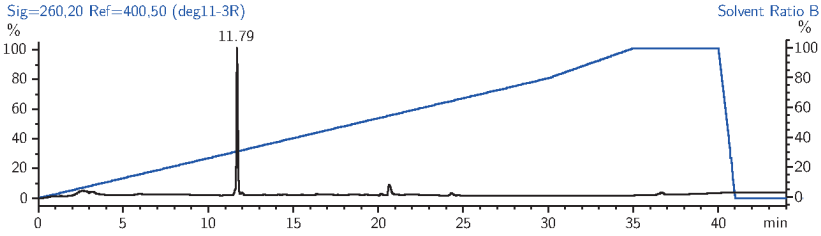
Sequence (5′ → 3′):
AGGLTTALGGTTAGGLTTALGG



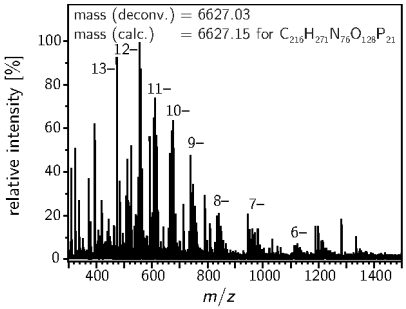
(*R*)-L^2



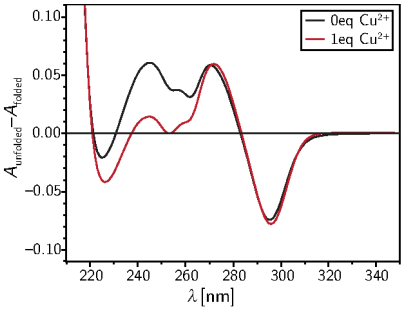
Analytical RP-HPLC



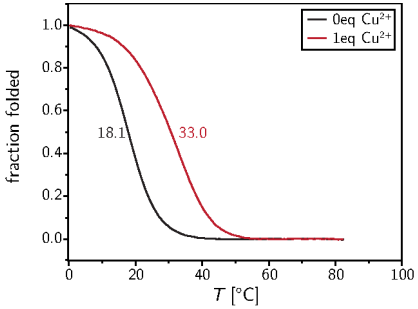
ESI(−)-MS



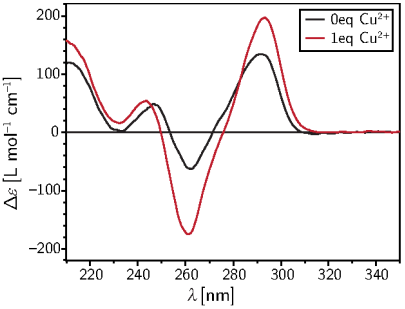
TDS (K$^+$)



UV melting curves (K$^+$)

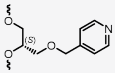


CD (K$^+$)



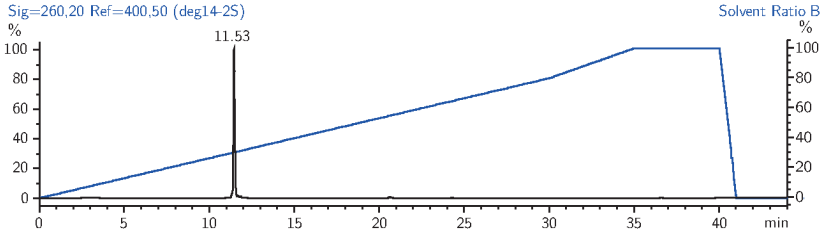
A.2.26. Oligonucleotide htel22-L^2$_4$a deg14-2S ((*S*)-L^2)

Sequence (5′ → 3′):
AGGGLTLGGGTTAGGGLTLGGG

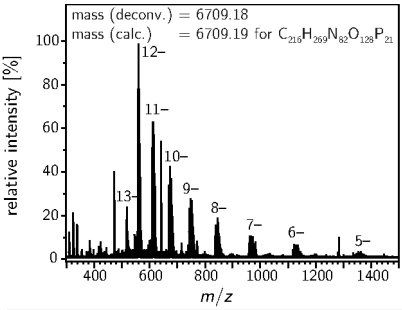


(*S*)-L^2

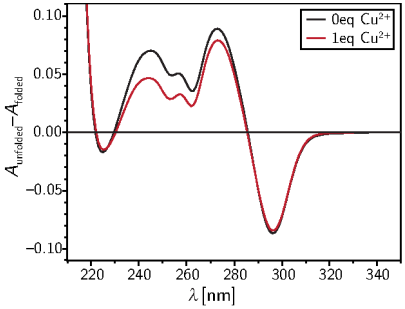
Analytical RP-HPLC



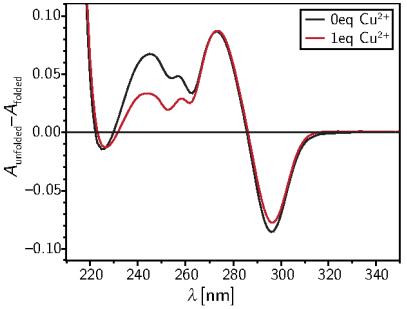
ESI(−)-MS



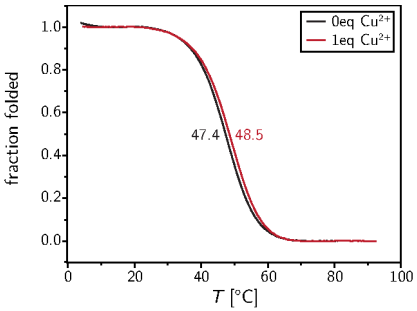
TDS (Na^+)



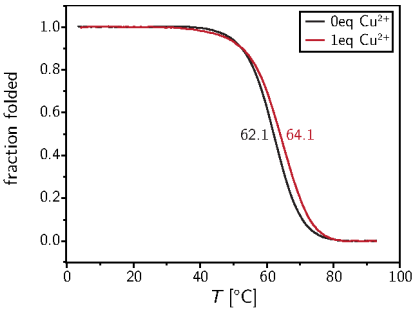
TDS (K^+)



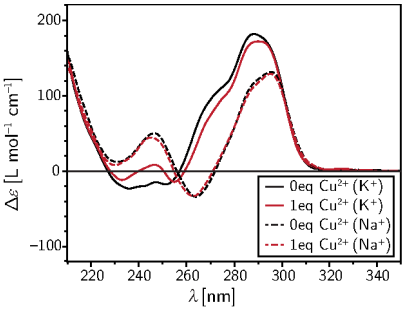
UV melting curves (Na^+)



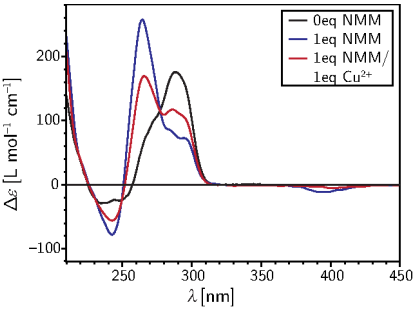
UV melting curves (K^+)



CD, (K^+)



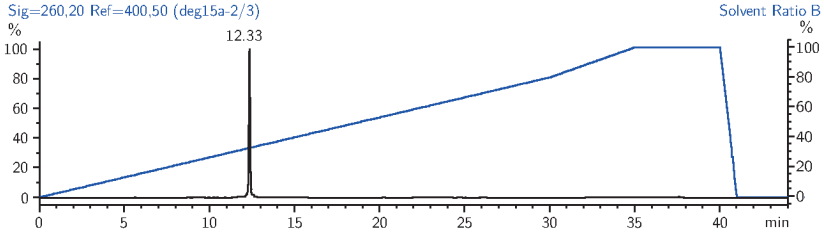
CD (K^+ +NMM)



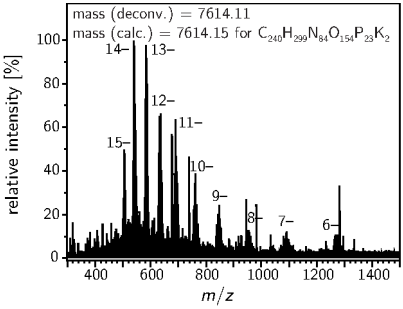
A.2.27. Oligonucleotide ttel24-T_4 deg15a

Sequence (5′ → 3′):		
TTGGGTTTTGGGTTGGGTTTTGGG	**no ligand**	

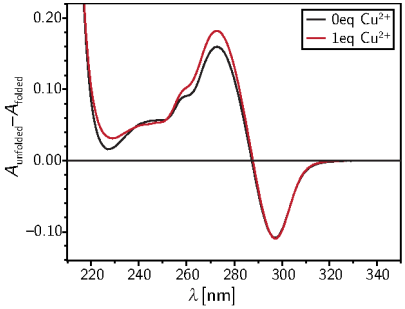
Analytical RP-HPLC



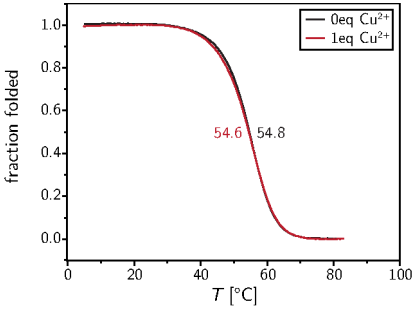
ESI(−)-MS



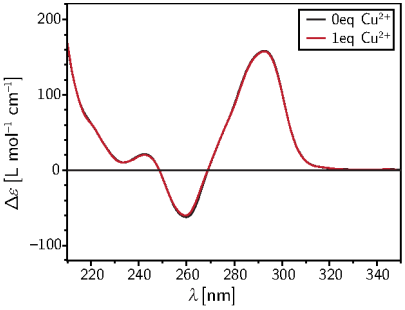
TDS (K^+)



UV melting curves (K^+)

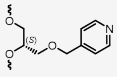


CD (K^+)

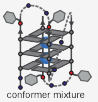


A.2.28. Oligonucleotide ttel24-L$^2_4$ deg15S ((*S*)-L^2)

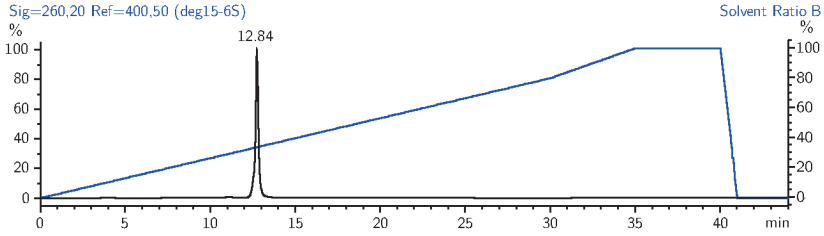
Sequence (5′ → 3′):
TTGGGLTTLGGGTTGGGLTTLGGG



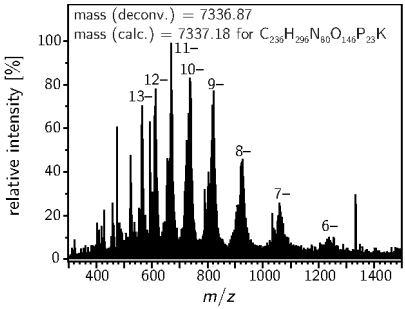
(*S*)-L^2



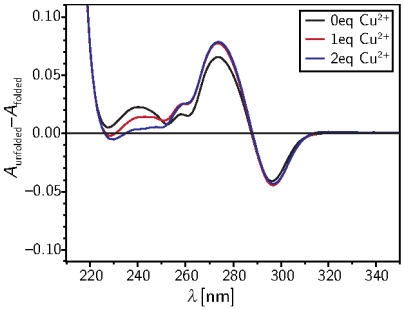
Analytical RP-HPLC



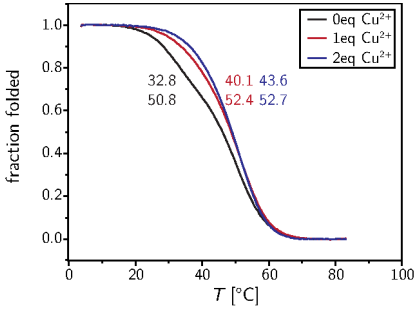
ESI(−)-MS



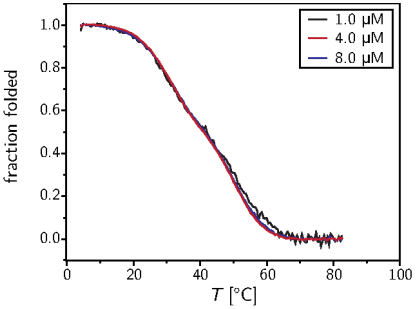
TDS (K$^+$)



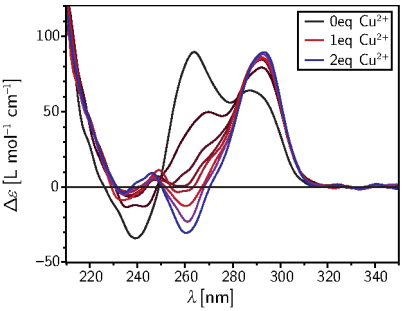
UV melting curves (K$^+$)



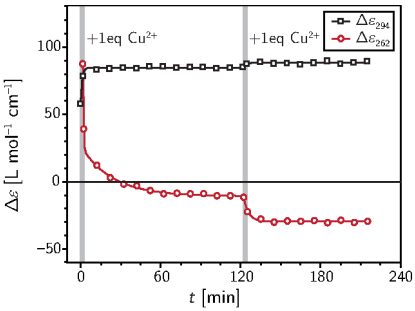
UV melting (K$^+$), conc. dep.



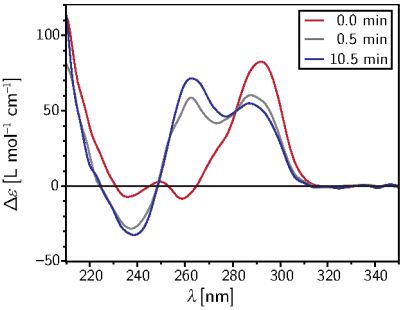
CD (K^+)



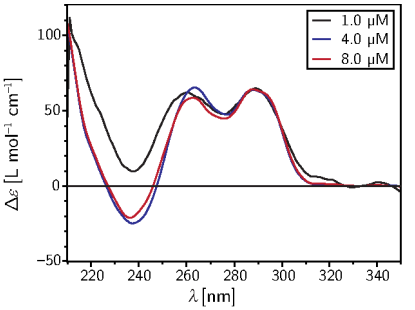
CD (K^+), *t*-dep.



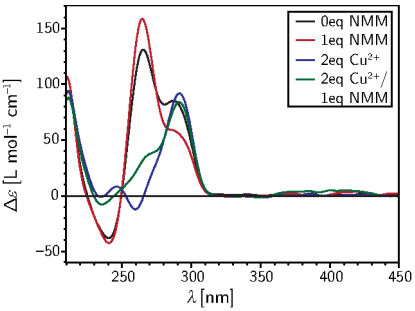
CD (K^+), Cu^{2+}, +edta, 30 °C



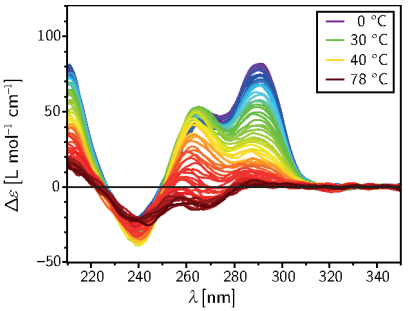
CD (K^+), conc. dep.



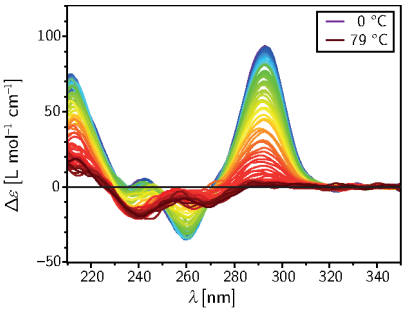
CD (K^+) +NMM



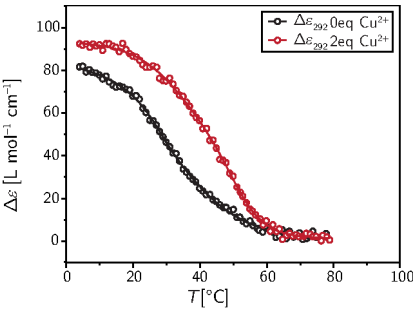
CD (K^+), *T*-dep., no Cu^{2+}



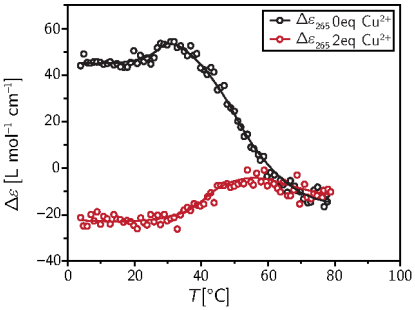
CD (K^+), *T*-dep., with Cu^{2+}



CD (K^+), melting curve



CD (K^+), melting curve



Bibliography

[1] D. L. Nelson, M. M. Cox, *Lehninger Biochemie*, Springer Berlin Heidelberg, Berlin, Heidelberg, **2009**.

[2] *The Nobel Prize in Chemistry 2015 - Press Release*, **2015**. `http://www.nobelprize.org/nobel_prizes/chemistry/laureates/2015/press.html`.

[3] J. Choi, T. Majima, *Chem. Soc. Rev.* **2011**, *40*, 5893–5909.

[4] S. Neidle, *FEBS J.* **2010**, *277*, 1118–1125.

[5] M. Mitchell, A. Gillis, M. Futahashi, H. Fujiwara, E. Skordalakes, *Nat. Struct. Mol. Biol.* **2010**, *17*, 513–518.

[6] J. W. Steed, J. L. Atwood, *Supramolecular chemistry 2nd ed.*, Wiley, Chichester, UK, **2009**.

[7] G. H. Clever, C. Kaul, T. Carell, *Angew. Chem. Int. Ed.* **2007**, *46*, 6226–6236; *Angew. Chem.* **2007**, *119*, 6340–6350.

[8] F. Coste, J. M. Malinge, L. Serre, W. Shepard, M. Roth, M. Leng, C. Zelwer, *Nucleic Acids Res.* **1999**, *27*, 1837–1846.

[9] W. Kaim, B. Schwederski, *Bioanorganische Chemie: Zur Funktion chemischer Elemente in Lebensprozessen 4th ed.*, of *STUDIUM*, Vieweg+Teubner, Wiesbaden, **2010**.

[10] N. M. Bell, J. Micklefield, *ChemBioChem* **2009**, *10*, 2691–2703.

[11] Y. Takezawa, M. Shionoya, *Acc. Chem. Res.* **2012**, *45*, 2066–2076.

[12] A. F. Holleman, E. Wiberg, N. Wiberg, *Lehrbuch der anorganischen Chemie 102nd ed.*, de Gruyter, Berlin and New York, **2007**.

[13] J. M. Schurr in *Nucleic Acid-Metal Ion Interactions*, N. V. Hud (Ed.), Royal Society of Chemistry, Cambridge, **2008**, pp. 307–349.

[14] N. V. Hud, A. E. Engelhart in *Nucleic Acid-Metal Ion Interactions*, N. V. Hud (Ed.), Royal Society of Chemistry, Cambridge, **2008**, pp. 75–117.

[15] A. Sigel, H. Sigel, R. K. O. Sigel (Eds.), *Interplay between Metal Ions and Nucleic Acids*, of *Metal Ions in Life Sciences*, Springer Netherlands, Dordrecht, **2012**.

[16] D. M. J. Lilley, *Quart. Rev. Biophys.* **2000**, *33*, 109–159.

[17] Y. Liu, S. C. West, *Nat. Rev. Mol. Cell. Biol.* **2004**, *5*, 937–944.

[18] M. D. Frank-Kamenetskii, S. M. Mirkin, *Annu. Rev. Biochem.* **1995**, *64*, 65–95.

[19] H. A. Day, P. Pavlou, Z. A. E. Waller, *Bioorg. Med. Chem.* **2014**, *22*, 4407–4418.

[20] N. B. Leontis, *Nucleic Acids Res.* **2002**, *30*, 3497–3531.

[21] D. Rhodes, H. J. Lipps, *Nucleic Acids Res.* **2015**, *43*, 8627–8637.

[22] M. A. Keniry, *Biopolymers* **2000**, *56*, 123–146.

[23] T. Simonsson, *Biol. Chem.* **2001**, *382*, 621–628.

[24] J. T. Davis, *Angew. Chem. Int. Ed.* **2004**, *43*, 668–698; *Angew. Chem.* **2004**, *116*, 684–716.

[25] S. Neidle, S. Balasubramanian (Eds.), *Quadruplex Nucleic Acids // Quadruplex nucleic acids*, of *RSC biomolecular sciences*, Royal Society of Chemistry and RSC Pub, Cambridge, **2006**.

[26] S. Burge, G. N. Parkinson, P. Hazel, A. K. Todd, S. Neidle, *Nucleic Acids Res.* **2006**, *34*, 5402–5415.

[27] A. T. Phan, V. Kuryavyi, D. J. Patel, *Curr. Opin. Struct. Biol.* **2006**, *16*, 288–298.

[28] M. Egli, P. S. Pallan, *Curr. Opin. Struct. Biol.* **2010**, *20*, 262–275.

[29] T. M. Bryan, P. Baumann, *Mol. Biotechnol.* **2011**, *49*, 198–208.

[30] X. Cang, J. Šponer, T. E. Cheatham III, *J. Am. Chem. Soc.* **2011**, *133*, 14270–14279.

[31] M. L. Bochman, K. Paeschke, V. A. Zakian, *Nat. Rev. Genet.* **2012**, *13*, 770–780.

[32] J. B. Chaires, D. Graves (Eds.), *Quadruplex Nucleic Acids*, of *Topics in Current Chemistry*, Springer Berlin Heidelberg, Berlin and Heidelberg, **2013**.

[33] C. J. Lech, B. Heddi, A. T. Phan, *Nucleic Acids Res.* **2013**, *41*, 2034–2046.

[34] J. Zhou, A. Bourdoncle, F. Rosu, V. Gabelica, J.-L. Mergny, *Angew. Chem. Int. Ed.* **2012**, *51*, 11002–11005; *Angew. Chem.* **2012**, *124*, 11164–11167.

[35] J. L. Huppert, *Nucleic Acids Res.* **2005**, *33*, 2908–2916.

[36] J. L. Huppert, *FEBS J.* **2010**, *277*, 3452–3458.

[37] A. E. Engelhart, J. Plavec, Ö. Persil, N. V. Hud in *Nucleic Acid-Metal Ion Interactions*, N. V. Hud (Ed.), Royal Society of Chemistry, Cambridge, **2008**, pp. 118–153.

[38] G. N. Parkinson, M. P. H. Lee, S. Neidle, *Nature* **2002**, *417*, 876–880.

[39] G. Laughlan, A. I. Murchie, D. G. Norman, M. H. Moore, P. C. Moody, D. M. Lilley, B. Luisi, *Science* **1994**, *265*, 520–524.

[40] W. Guschlbauer, J. F. Chantot, D. Thiele, *J. Biomol. Struct. Dyn.* **1990**, *8*, 491–511.

[41] A. Wong, G. Wu, *J. Am. Chem. Soc.* **2003**, *125*, 13895–13905.

[42] D. J. Patel, A. T. Phan, V. Kuryavyi, *Nucleic Acids Res.* **2007**, *35*, 7429–7455.

[43] S. Haider, G. N. Parkinson, S. Neidle, *J. Mol. Biol.* **2002**, *320*, 189–200.

[44] X. Cang, J. Šponer, T. E. Cheatham III, *Nucleic Acids Res.* **2011**, *39*, 4499–4512.

[45] J. L. Huppert, S. Balasubramanian, *Nucleic Acids Res.* **2006**, *35*, 406–413.

[46] P. Murat, S. Balasubramanian, *Curr. Opin. Genet. Dev.* **2014**, *25*, 22–29.

[47] M. Fry, *Front. Biosci.* **2007**, *12*, 4336.

[48] C. Sissi, B. Gatto, M. Palumbo, *Biochimie* **2011**, *93*, 1219–1230.

[49] V. Brázda, L. Hároníková, J. C. C. Liao, M. Fojta, *Int. J. Mol. Sci.* **2014**, *15*, 17493–17517.

[50] G. Biffi, D. Tannahill, J. McCafferty, S. Balasubramanian, *Nat. Chem.* **2013**, *5*, 182–186.

[51] A. Laguerre, K. Hukezalie, P. Winckler, F. Katranji, G. Chanteloup, M. Pirrotta, J.-M. Perrier-Cornet, Wong, Judy M Y, D. Monchaud, *J. Am. Chem. Soc.* **2015**.

[52] G. W. Collie, G. N. Parkinson, *Chem. Soc. Rev.* **2011**, *40*, 5867–5892.

[53] S. Balasubramanian, L. H. Hurley, S. Neidle, *Nat. Rev. Drug. Discov.* **2011**, *10*, 261–275.

[54] J. Bidzinska, G. Cimino-Reale, N. Zaffaroni, M. Folini, *Molecules* **2013**, *18*, 12368–12395.

[55] R. Simone, P. Fratta, S. Neidle, G. N. Parkinson, A. M. Isaacs, *FEBS Lett.* **2015**, *589*, 1653–1668.

[56] M. Métifiot, S. Amrane, S. Litvak, M.-L. Andreola, *Nucleic Acids Res.* **2014**, *42*, 12352–12366.

[57] A. D. Cian, J.-L. Mergny, *Nucleic Acids Res.* **2007**, *35*, 2483–2493.

[58] M. Franceschin, *Eur. J. Org. Chem.* **2009**, *2009*, 2225–2238.

[59] N. W. Luedtke, *Chimia* **2009**, *63*, 134–139.

[60] S. N. Georgiades, N. H. Abd Karim, K. Suntharalingam, R. Vilar, *Angew. Chem. Int. Ed.* **2010**, *49*, 4020–4034; *Angew. Chem.* **2010**, *122*, 4114–4128.

[61] T. Biver, *Coord. Chem. Rev.* **2013**, *257*, 2765–2783.

[62] S. A. Ohnmacht, S. Neidle, *Bioorg. Med. Chem. Lett.* **2014**, *24*, 2602–2612.

[63] G. F. Salgado, C. Cazenave, A. Kerkour, J.-L. Mergny, *Chem. Sci.* **2015**, *6*, 3314–3320.

[64] E. Arunan, G. R. Desiraju, R. A. Klein, J. Sadlej, S. Scheiner, I. Alkorta, D. C. Clary, R. H. Crabtree, J. J. Dannenberg, P. Hobza, H. G. Kjaergaard, A. C. Legon, B. Mennucci, D. J. Nesbitt, *Pure Appl. Chem.* **2011**, *83*.

[65] P. J. Cragg, *Supramolecular chemistry: From biological inspiration to biomedical applications*, Springer, Dordrecht and London, **2010**.

[66] K. Ariga, T. Kunitake, *Supramolecular chemistry: Fundamentals and applications : advanced textbook*, Springer, Heidelberg, **2006**.

[67] F. Diederich, P. J. Stang, R. R. Tykwinski, *Modern supramolecular chemistry: Strategies for macrocycle synthesis*, Wiley-VCH, Weinheim, **2008**.

[68] N. C. Seeman, *J. Theor. Biol.* **1982**, *99*, 237–247.

[69] N. C. Seeman, *Mol. Biotechnol.* **2007**, *37*, 246–257.

[70] N. C. Seeman, *Annu. Rev. Biochem.* **2010**, *79*, 65–87.

[71] N. C. Seeman, *Nano Lett.* **2010**, *10*, 1971–1978.

[72] I. Willner, B. Willner, *Nano Lett.* **2010**, *10*, 3805–3815.

[73] C. Teller, I. Willner, *Curr. Opin. Biotechnol.* **2010**, *21*, 376–391.

[74] Y. Krishnan, F. C. Simmel, *Angew. Chem. Int. Ed.* **2011**, *50*, 3124–3156; *Angew. Chem.* **2011**, *123*, 3180–3215.

[75] E. Stulz, *Chem. Eur. J.* **2012**, *18*, 4456–4469.

[76] C.-H. Lu, B. Willner, I. Willner, *ACS Nano* **2013**, *7*, 8320–8332.

[77] Y. Tang, B. Ge, D. Sen, H.-Z. Yu, *Chem. Soc. Rev.* **2014**, *43*, 518–529.

[78] L. Zhang, M. Li, Y. Qin, Z. Chu, S. Zhao, *Analyst* **2014**, *139*, 6298–6303.

[79] F. Wang, X. Liu, I. Willner, *Angew. Chem. Int. Ed.* **2015**, *54*, 1098–1129; *Angew. Chem.* **2015**, *127*, 1112–1144.

[80] J. H. Chen, N. C. Seeman, *Nature* **1991**, *350*, 631–633.

[81] P. W. K. Rothemund, *Nature* **2006**, *440*, 297–302.

[82] S. M. Douglas, H. Dietz, T. Liedl, B. Högberg, F. Graf, W. M. Shih, *Nature* **2009**, *459*, 414–418.

[83] S. D. Perrault, W. M. Shih, *ACS Nano* **2014**, *8*, 5132–5140.

[84] H. M. T. Choi, V. A. Beck, N. A. Pierce, *ACS Nano* **2014**, *8*, 4284–4294.

[85] A. Shaw, V. Lundin, E. Petrova, F. Fördős, E. Benson, A. Al-Amin, A. Herland, A. Blokzijl, B. Högberg, A. I. Teixeira, *Nat. Methods* **2014**, *11*, 841–846.

[86] F. Zhang, J. Nangreave, Y. Liu, H. Yan, *J. Am. Chem. Soc.* **2014**, *136*, 11198–11211.

[87] S. M. Douglas, I. Bachelet, G. M. Church, *Science* **2012**, *335*, 831–834.

[88] Y.-J. Chen, B. Groves, R. A. Muscat, G. Seelig, *Nat. Nanotech.* **2015**, *10*, 748–760.

[89] R. M. Zadegan, M. L. Norton, *Int. J. Mol. Sci.* **2012**, *13*, 7149–7162.

[90] N. L. Rosi, C. A. Mirkin, *Chem. Rev.* **2005**, *105*, 1547–1562.

[91] F. Wang, C.-H. Lu, I. Willner, *Chem. Rev.* **2014**, *114*, 2881–2941.

[92] Y. Singh, P. Murat, E. Defrancq, *Chem. Soc. Rev.* **2010**, *39*, 2054–2070.

[93] M. A. Campbell, J. Wengel, *Chem. Soc. Rev.* **2011**, *40*, 5680–5689.

[94] N. Ueda, T. Kawabata, K. Takemoto, *J. Heterocycl. Chem.* **1971**, *8*, 827–829.

[95] T. Seita, K. Yamauchi, M. Kinoshita, M. Imoto, *Makromol. Chem.* **1972**, *154*, 255–261.

[96] L. Zhang, A. Peritz, E. Meggers, *J. Am. Chem. Soc.* **2005**, *127*, 4174–4175.

[97] M. K. Schlegel, L.-O. Essen, E. Meggers, *J. Am. Chem. Soc.* **2008**, *130*, 8158–8159.

[98] M. K. Schlegel, L.-O. Essen, E. Meggers, *Chem. Commun.* **2010**, *46*, 1094.

[99] E. Meggers, L. Zhang, *Acc. Chem. Res.* **2010**, *43*, 1092–1102.

[100] P. Nielsen, M. Egholm, R. Berg, O. Buchardt, *Science* **1991**, *254*, 1497–1500.

[101] P. E. Nielsen, G. Haaima, *Chem. Soc. Rev.* **1997**, *26*, 73.

[102] P. E. Nielsen in *Drug-Nucleic Acid Interactions*, Vol. 340 of *Methods in Enzymology*, Elsevier, **2001**, pp. 329–340.

[103] S. Barluenga, N. Winssinger, *Acc. Chem. Res.* **2015**, *48*, 1319–1331.

[104] A. H. El-Sagheer, T. Brown, *Chem. Soc. Rev.* **2010**, *39*, 1388–1405.

[105] C. I. Schilling, N. Jung, M. Biskup, U. Schepers, S. Bräse, *Chem. Soc. Rev.* **2011**, *40*, 4840–4871.

[106] B. Saccà, C. M. Niemeyer, *Chem. Soc. Rev.* **2011**, *40*, 5910.

[107] F. Diezmann, O. Seitz, *Chem. Soc. Rev.* **2011**, *40*, 5789.

[108] V. L. Malinovskii, D. Wenger, R. Häner, *Chem. Soc. Rev.* **2010**, *39*, 410–422.

[109] S. H. Weisbrod, A. Marx, *Chem. Commun.* **2007**, 1828–1830.

[110] S. H. Weisbrod, A. Baccaro, A. Marx, *Nucleic Acids Symp. Ser.* **2008**, 383–384.

[111] F. Wojciechowski, C. J. Leumann, *Chem. Soc. Rev.* **2011**, *40*, 5669–5679.

[112] J. Müller, *Eur. J. Inorg. Chem.* **2008**, 3749–3763.

[113] G. H. Clever, M. Shionoya, *Coord. Chem. Rev.* **2010**, *254*, 2391–2402.

[114] A. Ono, H. Torigoe, Y. Tanaka, I. Okamoto, *Chem. Soc. Rev.* **2011**, *40*, 5855.

[115] P. Scharf, J. Müller, *ChemPlusChem* **2013**, *78*, 20–34.

[116] K. Tanaka, M. Shionoya, *J. Org. Chem.* **1999**, *64*, 5002–5003.

[117] E. Meggers, P. L. Holland, W. B. Tolman, F. E. Romesberg, P. G. Schultz, *J. Am. Chem. Soc.* **2000**, *122*, 10714–10715.

[118] S. Johannsen, N. Megger, D. Böhme, Sigel, Roland K O, J. Müller, *Nat. Chem.* **2010**, *2*, 229–234.

[119] H. Weizman, Y. Tor, *Chem. Commun.* **2001**, 453–454.

[120] N. Düpre, L. Welte, J. Gómez-Herrero, F. Zamora, J. Müller, *Inorg. Chim. Acta* **2009**, *362*, 985–992.

[121] K. Tanaka, Y. Yamada, M. Shionoya, *J. Am. Chem. Soc.* **2002**, *124*, 8802–8803.

[122] Y. Takezawa, W. Maeda, K. Tanaka, M. Shionoya, *Angew. Chem. Int. Ed.* **2009**, *48*, 1081–1084; *Angew. Chem.* **2009**, *121*, 1101–1104.

[123] J.-L. H. A. Duprey, Y. Takezawa, M. Shionoya, *Angew. Chem. Int. Ed.* **2013**, *52*, 1212–1216; *Angew. Chem.* **2013**, *125*, 1250–1254.

[124] G. Bianké, R. Häner, *ChemBioChem* **2004**, *5*, 1063–1068.

[125] H. Yang, H. F. Sleiman, *Angew. Chem. Int. Ed.* **2008**, *47*, 2443–2446; *Angew. Chem.* **2008**, *120*, 2477–2480.

[126] H. Yang, K. L. Metera, H. F. Sleiman, *Coord. Chem. Rev.* **2010**, *254*, 2403–2415.

[127] C. K. McLaughlin, G. D. Hamblin, H. F. Sleiman, *Chem. Soc. Rev.* **2011**, *40*, 5647.

[128] M. Gellert, M. N. Lipsett, D. R. Davies, *Proc. Natl. Acad. Sci. U. S. A.* **1962**, *48*, 2013–2018.

[129] J. T. Davis, G. P. Spada, *Chem. Soc. Rev.* **2007**, *36*, 296–313.

[130] S. Neidle, S. Balasubramanian in *Quadruplex Nucleic Acids // Quadruplex nucleic acids*, of *RSC biomolecular sciences*, S. Neidle, S. Balasubramanian (Eds.), Royal Society of Chemistry and RSC Pub, Cambridge, **2006**, pp. 253–296.

[131] A. Wong, R. Ida, L. Spindler, G. Wu, *J. Am. Chem. Soc.* **2005**, *127*, 6990–6998.

[132] A. Ghoussoub, J.-M. Lehn, *Chem. Commun.* **2005**, 5763–5765.

[133] M. S. Kaucher, W. A. Harrell, J. T. Davis, *J. Am. Chem. Soc.* **2006**, *128*, 38–39.

[134] H. Fenniri, B.-L. Deng, A. E. Ribbe, *J. Am. Chem. Soc.* **2002**, *124*, 11064–11072.

[135] P. Murat, D. Cressend, N. Spinelli, Van der Heyden, Angéline, P. Labbé, P. Dumy, E. Defrancq, *ChemBioChem* **2008**, *9*, 2588–2591.

[136] R. Bonnet, T. Lavergne, B. Gennaro, N. Spinelli, E. Defrancq, *Chem. Commun.* **2015**, *51*, 4850–4853.

[137] G. Oliviero, J. Amato, N. Borbone, A. Galeone, L. Petraccone, M. Varra, G. Piccialli, L. Mayol, *Bioconjugate Chem.* **2006**, *17*, 889–898.

[138] G. Oliviero, N. Borbone, J. Amato, S. D'Errico, A. Galeone, G. Piccialli, M. Varra, L. Mayol, *Biopolymers* **2009**, *91*, 466–477.

[139] R. Ferreira, M. Alvira, A. Aviñó, I. Gómez-Pinto, C. González, V. Gabelica, R. Eritja, *ChemistryOpen* **2012**, *1*, 106–114.

[140] Y. Xu, Y. Suzuki, M. Komiyama, *Angew. Chem. Int. Ed.* **2009**, *48*, 3281–3284; *Angew. Chem.* **2009**, *121*, 3331–3334.

[141] P. Murat, R. Bonnet, Van der Heyden, Angéline, N. Spinelli, P. Labbé, D. Monchaud, M.-P. Teulade-Fichou, P. Dumy, E. Defrancq, *Chem. Eur. J.* **2010**, *16*, 6106–6114.

[142] R. Bonnet, P. Murat, N. Spinelli, E. Defrancq, *Chem. Commun.* **2012**, *48*, 5992.

[143] G. Oliviero, J. Amato, N. Borbone, S. D'Errico, A. Galeone, L. Mayol, S. Haider, O. Olubiyi, B. Hoorelbeke, J. Balzarini, G. Piccialli, *Chem. Commun.* **2010**, *46*, 8971–8973.

[144] L. A. Yatsunyk, O. Piétrement, D. Albrecht, P. L. T. Tran, D. Renčiuk, H. Sugiyama, J.-M. Arbona, J.-P. Aimé, J.-L. Mergny, *ACS Nano* **2013**, *7*, 5701–5710.

[145] O. Mendoza, M. Porrini, G. F. Salgado, V. Gabelica, J.-L. Mergny, *Chem. Eur. J.* **2015**, *21*, 6732–6739.

[146] A. Rajendran, M. Endo, K. Hidaka, P. Lan Thao Tran, J.-L. Mergny, H. Sugiyama, *Nucleic Acids Res.* **2013**.

[147] A. Rajendran, M. Endo, K. Hidaka, H. Sugiyama, *Angew. Chem. Int. Ed.* **2014**, *53*, 4107–4112; *Angew. Chem.* **2014**, *126*, 4191–4196.

[148] R. P. Fahlman, D. SEN, *J. Mol. Biol.* **1998**, *280*, 237–244.

[149] S. Nagatoishi, T. Nojima, B. Juskowiak, S. Takenaka, *Angew. Chem. Int. Ed.* **2005**, *44*, 5067–5070; *Angew. Chem.* **2005**, *117*, 5195–5198.

[150] P. Alberti, J.-L. Mergny, *Proc. Natl. Acad. Sci. U. S. A.* **2003**, *100*, 1569–1573.

[151] N. Hamaguchi, A. Ellington, M. Stanton, *Anal. Biochem.* **2001**, *294*, 126–131.

[152] M. A. Aleman-Garcia, R. Orbach, I. Willner, *Chem. Eur. J.* **2014**, *20*, 5619–5624.

[153] B. Deng, Y. Lin, C. Wang, F. Li, Z. Wang, H. Zhang, X.-F. Li, X. C. Le, *Anal. Chim. Acta* **2014**, *837*, 1–15.

[154] Y. Li, D. SEN, *Biochemistry* **1997**, *36*, 5589–5599.

[155] P. Travascio, Y. Li, D. Sen, *Chem. Biol.* **1998**, *5*, 505–517.

[156] X. Cheng, X. Liu, T. Bing, Z. Cao, D. Shangguan, *Biochemistry* **2009**, *48*, 7817–7823.

[157] S. Shimron, J. Elbaz, A. Henning, I. Willner, *Chem. Commun.* **2010**, *46*, 3250–3252.

[158] J. Chen, D. Liu, A. H. F. Lee, J. Qi, A. S. C. Chan, T. Li, *Chem. Commun.* **2002**, 2686–2687.

[159] C. Wang, Y. Li, G. Jia, Y. Liu, S. Lu, C. Li, *Chem. Commun.* **2012**, *48*, 6232.

[160] Y. C. Huang, D. Sen, *J. Am. Chem. Soc.* **2010**, *132*, 2663–2671.

[161] Y. Li, D. Sen, *Chem. Biol.* **1998**, *5*, 1–12.

[162] S.-P. Liu, S. H. Weisbrod, Z. Tang, A. Marx, E. Scheer, A. Erbe, *Angew. Chem. Int. Ed.* **2010**, *49*, 3313–3316; *Angew. Chem.* **2010**, *122*, 3385–3388.

[163] B. Y.-W. Man, D. S.-H. Chan, H. Yang, S.-W. Ang, F. Yang, S.-C. Yan, C.-M. Ho, P. Wu, C.-M. Che, C.-H. Leung, D.-L. Ma, *Chem. Commun.* **2010**, *46*, 8534.

[164] T. Li, B. Li, E. Wang, S. Dong, *Chem. Commun.* **2009**, 3551–3553.

[165] N. M. Smith, S. Amrane, F. Rosu, V. Gabelica, J.-L. Mergny, *Chem. Commun.* **2012**, *48*, 11464–11466.

[166] Z.-S. Wu, C.-R. Chen, G.-L. Shen, R.-Q. Yu, *Biomaterials* **2008**, *29*, 2689–2696.

[167] Y. Xu, Y. Suzuki, T. Lönnberg, M. Komiyama, *J. Am. Chem. Soc.* **2009**, *131*, 2871–2874.

[168] D. Miyoshi, H. Karimata, Z.-M. Wang, K. Koumoto, N. Sugimoto, *J. Am. Chem. Soc.* **2007**, *129*, 5919–5925.

[169] J.-L. Mergny, L. Lacroix, *Oligonucleotides* **2003**, *13*, 515–537.

[170] J.-L. Mergny, J. Li, L. Lacroix, S. Amrane, J. B. Chaires, *Nucleic Acids Res.* **2005**, *33*, e138.

[171] L. A. Marky, K. J. Breslauer, *Biopolymers* **1987**, *26*, 1601–1620.

[172] S. J. Schroeder, D. H. Turner in *Biophysical, Chemical, and Functional Probes of RNA Structure, Interactions and Folding: Part A*, Vol. 468 of *Methods in Enzymology*, D. Herschlag (Ed.), Elsevier, **2009**, pp. 371–387.

[173] J.-L. Mergny, A. de Cian, A. Ghelab, B. Saccà, L. Lacroix, *Nucleic Acids Res.* **2005**, *33*, 81–94.

[174] W. W. Parson, *Modern optical spectroscopy: With examples from biophysics and biochemistry*, Springer, Berlin and New York, **2007**.

[175] E. Zimmermann, Dissertation, Gerhard-Mercator-Universität Duisburg, Duisburg, **2002**.

[176] J. Kypr, I. Kejnovská, D. Renciuk, M. Vorlícková, *Nucleic Acids Res.* **2009**, *37*, 1713–1725.

[177] A. Randazzo, G. P. Spada, M. Webba da Silva, *Top. Curr. Chem.* **2013**, *330*, 67–86.

[178] S. Masiero, R. Trotta, S. Pieraccini, S. de Tito, R. Perone, A. Randazzo, G. P. Spada, *Org. Biomol. Chem.* **2010**, *8*, 2683–2692.

[179] D. M. Gray, J.-D. Wen, C. W. Gray, R. Repges, C. Repges, G. Raabe, J. Fleischhauer, *Chirality* **2008**, *20*, 431–440.

[180] A. Marchand, A. Granzhan, K. Iida, Y. Tsushima, Y. Ma, K. Nagasawa, M.-P. Teulade-Fichou, V. Gabelica, *J. Am. Chem. Soc.* **2015**, *137*, 750–756.

[181] M. J. Abraham, D. van der Spoel, E. Lindahl, B. Hess, and the GROMACS development team, *GROMACS User Manual version 5.1*, **2015**. `www.gromacs.org`.

[182] S. A. Adcock, J. A. McCammon, *Chem. Rev.* **2006**, *106*, 1589–1615.

[183] T. E. Cheatham III, D. A. Case, *Biopolymers* **2013**, *99*, 969–977.

[184] J. Šponer, P. Banáš, P. Jurečka, M. Zgarbová, P. Kührová, M. Havrila, M. Krepl, P. Stadlbauer, M. Otyepka, *J. Phys. Chem. Lett.* **2014**, *5*, 1771–1782.

[185] M. A. Young, T. E. Cheatham III, M. A. Young, *Biopolymers* **2000**, *56*, 232–256.

[186] T. E. Cheatham III, *Curr. Opin. Struct. Biol.* **2004**, *14*, 360–367.

[187] J. Šponer, X. Cang, T. E. Cheatham III, *Methods (San Diego, CA, U. S.)* **2012**, *57*, 25–39.

[188] S. Haider, S. Neidle, *Methods Mol. Biol.* **2010**, *608*, 17–37.

[189] J. Šponer, N. Špačková, *Methods (San Diego, CA, U. S.)* **2007**, *43*, 278–290.

[190] M. Krepl, M. Zgarbová, P. Stadlbauer, M. Otyepka, P. Banáš, J. Koča, T. E. Cheatham III, P. Jurečka, J. Šponer, *J. Chem. Theory Comput.* **2012**, *8*, 2506–2520.

[191] E. Fadrná, N. Špačková, J. Sarzyñska, J. Koča, M. Orozco, T. E. Cheatham III, T. Kulinski, J. Šponer, *J. Chem. Theory Comput.* **2009**, *5*, 2514–2530.

[192] P. Stadlbauer, M. Krepl, T. E. Cheatham III, J. Koca, J. Šponer, *Nucleic Acids Res.* **2013**, *41*, 7128–7143.

[193] R. D. Gray, J. O. Trent, J. B. Chaires, *J. Mol. Biol.* **2014**, *426*, 1629–1650.

[194] L. Petraccone, N. C. Garbett, J. B. Chaires, J. O. Trent, *Biopolymers* **2010**, *93*, 533–548.

[195] R. J. Deeth, A. Anastasi, C. Diedrich, K. Randell, *Coord. Chem. Rev.* **2009**, *253*, 795–816.

[196] P. Comba, *Coord. Chem. Rev.* **2003**, *238-239*, 9–20.

[197] M. Yoneya, T. Yamaguchi, S. Sato, M. Fujita, *J. Am. Chem. Soc.* **2012**, *134*, 14401–14407.

[198] M. Yoneya, S. Tsuzuki, T. Yamaguchi, S. Sato, M. Fujita, *ACS Nano* **2014**, *8*, 1290–1296.

[199] B. R. Gibney, D. P. Kessissoglou, J. W. Kampf, V. L. Pecoraro, *Inorg. Chem.* **1994**, *33*, 4840–4849.

[200] M. Han, D. M. Engelhard, G. H. Clever, *Chem. Soc. Rev.* **2014**, *43*, 1848–1860.

[201] R. M. Izatt, J. J. Christensen, J. H. Rytting, *Chem. Rev.* **1971**, *71*, 439–481.

[202] B. Lippert in *Nucleic Acid-Metal Ion Interactions*, N. V. Hud (Ed.), Royal Society of Chemistry, Cambridge, **2008**, pp. 39–74.

[203] B. Iglesias, R. Alvarez, A. R. de Lera, *Tetrahedron* **2001**, *57*, 3125–3130.

[204] *Spartan '08*, **2008**, Wavefunction, Inc., Irvine, CA. `https://www.wavefun.com/products/spartan.html`.

[205] G. R. Clark, P. D. Pytel, C. J. Squire, *Nucleic Acids Res.* **2012**, *40*, 5731–5738.

[206] L. F. A. Hennequin, P. Ple, WO 02/092577, **2002**.

[207] S. L. Beaucage, M. H. Caruthers, *Tetrahedron Lett.* **1981**, *22*, 1859–1862.

[208] L. J. McBride, M. H. Caruthers, *Tetrahedron Lett.* **1983**, *24*, 245–248.

[209] J. E. Marugg, A. Burik, M. Tromp, G. A. van der Marel, J. H. van Boom, *Tetrahedron Lett.* **1986**, *27*, 2271–2274.

[210] C. Sirichaiwat, C. Intaraudom, S. Kamchonwongpaisan, J. Vanichtanankul, Y. Thebtaranonth, Y. Yuthavong, *J. Med. Chem.* **2004**, *47*, 345–354.

[211] A. R. Katritzky, R. Murugan, *J. Chem. Soc., Perkin Trans. 2* **1987**, 1867.

[212] B. Nilsson, J. Tejbrant, B. Pelcman, E. Ringberg, M. Thor, J. Nilsson, M. Jönsson, US 6465 B1, **2002**.

[213] T. Teraji, K. Sakane, J. Goto, 0074653 A2, **1983**.

[214] K. Ikegashira, T. Ikenogami, N. Ogawa, T. Matsumoto, T. Oka, T. Matsuo, T. Yamasaki, WO2010131738 (A1), **2010**.

[215] F. S. Gibson, M. S. Park, H. Rapoport, *J. Org. Chem.* **1994**, *59*, 7503–7507.

[216] M. Lemaire, F. Posada, J.-G. Gourcy, G. Jeminet, *Synthesis* **1995**, 627–629.

[217] D. Albanese, M. Salsa, D. Landini, V. Lupi, M. Penso, *Eur. J. Org. Chem.* **2007**, *2007*, 2107–2113.

[218] R. L. Bixler, C. Niemann, *J. Org. Chem.* **1958**, *23*, 575–584.

[219] B. K. Albrecht, J. E. Audia, A. Côté, V. S. Gehling, J.-C. Harmange, M. C. Hewitt, C. G. Naveschuk, A. M. Taylor, R. G. Vaswani, 2012/075456 A1, **2012**.

[220] K. Clinch, G. B. Evans, R. F. G. Fröhlich, R. H. Furneaux, P. M. Kelly, J. M. Mason, V. L. Schramm, P. C. Tyler, WO 2008/030118 A1, **2008**.

[221] M. Fuji, F. Watanabe, Y. Fujii, H. Hashizume, T. Okuno, K. Shirahase, I. Teshirogi, M. Ohtani, *J. Org. Chem.* **1997**, *62*, 6804–6809.

[222] B. Kristinsson, G. Haraldsson, *Synlett* **2008**, *2008*, 2178–2182.

[223] C. Fernandez, Z. Gándara, G. Gómez, B. Covelo, Y. Fall, *Tetrahedron Lett.* **2007**, *48*, 2939–2942.

[224] W. Du, S. S. Kulkarni, J. Gervay-Hague, *Chem. Commun.* **2007**, 2336.

[225] M. Abe, R. Niibayashi, S. Koubori, I. Moriyama, H. Miyoshi, *Biochemistry* **2011**, *50*, 8383–8391.

[226] G. G. Haraldsson, P. Thordarson, A. Halldorsson, B. Kristinsson, *Tetrahedron: Asymmetry* **1999**, *10*, 3671–3674.

[227] P. G. M. Wuts, T. W. Greene, *Greene's protective groups in organic synthesis 4th ed.*, Wiley-Interscience, Hoboken, N.J., **2007**.

[228] M. Smith, D. H. Rammler, I. H. Goldberg, H. G. Khorana, *J. Am. Chem. Soc.* **1962**, *84*, 430–440.

[229] H. Schaller, G. Weimann, B. Lerch, H. G. Khorana, *J. Am. Chem. Soc.* **1963**, *85*, 3821–3827.

[230] C. Kölmel, C. Ochsenfeld, R. Ahlrichs, *Theoret. Chim. Acta* **1992**, *82*, 271–284.

[231] R. B. Merrifield, *J. Am. Chem. Soc.* **1963**, *85*, 2149–2154.

[232] R. L. Letsinger, J. L. Finnan, G. A. Heavner, W. B. Lunsford, *J. Am. Chem. Soc.* **1975**, *97*, 3278–3279.

[233] S. Berner, K. Mühlegger, H. Seliger, *Nucleic Acids Res.* **1989**, *17*, 853–864.

[234] M. J. Cavaluzzi, P. N. Borer, M. J. Cavaluzzi, *Nucleic Acids Res.* **2004**, *32*, 13e.

[235] Q. Zhai, M. Deng, L. Xu, X. Zhang, X. Zhou, *Bioorg. Med. Chem. Lett.* **2012**, *22*, 1142–1145.

[236] *Biological Buffers*, **2008**. https://www.applichem.com/fileadmin/Broschueren/BioBuffer.pdf.

[237] N. E. Good, G. D. Winget, W. Winter, T. N. Connolly, S. Izawa, Singh, Raizada M. M., *Biochemistry* **1966**, *5*, 467–477.

[238] J. Duguid, V. A. Bloomfield, J. Benevides, G. J. Thomas, *Biophys. J.* **1993**, *65*, 1916–1928.

[239] V. Andrushchenko, van de Sande, J H, H. Wieser, *Biopolymers* **2003**, *72*, 374–390.

[240] C. Paris, F. Geinguenaud, C. Gouyette, J. Liquier, J. Lacoste, *Biophys. J.* **2007**, *92*, 2498–2506.

[241] D. Monchaud, P. Yang, L. Lacroix, M.-P. Teulade-Fichou, J.-L. Mergny, *Angew. Chem. Int. Ed.* **2008**, *47*, 4858–4861; *Angew. Chem.* **2008**, *120*, 4936–4949.

[242] H. Irving, R. J. P. Williams, *J. Chem. Soc.* **1953**, 3192.

[243] X.-H. Zhou, D.-M. Kong, H.-X. Shen, *Anal. Chem.* **2010**, *82*, 789–793.

[244] D. M. Engelhard, R. Pievo, G. H. Clever, *Angew. Chem. Int. Ed.* **2013**, *52*, 12843–12847; *Angew. Chem.* **2013**, *125*, 13078–13082.

[245] U. Dechert in *Gentechnische Methoden*, M. Jansohn, S. Rothhämel (Eds.), Spektrum Akademischer Verlag, Heidelberg, **2012**, pp. 37–93.

[246] T. Li, E. Wang, S. Dong, *Anal. Chem.* **2010**, *82*, 7576–7580.

[247] D.-L. Ma, H.-Z. He, K.-H. Leung, H.-J. Zhong, D. S.-H. Chan, C.-H. Leung, *Chem. Soc. Rev.* **2013**, *42*, 3427–3440.

[248] B. J. Bassam, G. Caetano-Anollés, P. M. Gresshoff, *Anal. Biochem.* **1991**, *196*, 80–83.

[249] B. J. Bassam, P. M. Gresshoff, *Nat. Protoc.* **2007**, *2*, 2649–2654.

[250] S. M. Hassur, H. W. Whitlock, *Anal. Biochem.* **1974**, *59*, 162–164.

[251] J. A. Weil, J. R. Bolton, *Electron Paramagnetic Resonance*, John Wiley & Sons, Inc, Hoboken, NJ, USA, **2006**.

[252] P. W. Atkins, J. de Paula, *Physikalische Chemie 4th ed.*, Wiley-VCH, Weinheim, **2006**.

[253] R. Šípoš, T. Szabó-Plánka, A. Rockenbauer, N. V. Nagy, J. Šima, M. Melník, I. Nagypál, *J. Phys. Chem. A* **2008**, *112*, 10280–10286.

[254] B. Hathaway, D. Billing, *Coord. Chem. Rev.* **1970**, *5*, 143–207.

[255] F. Bou-Abdallah, N. D. Chasteen, *J. Biol. Inorg. Chem.* **2007**, *13*, 15–24.

[256] B. Jash, J. Neugebauer, J. Müller, *Inorg. Chim. Acta* **2016**, *452*, 181–187.

[257] D. Sen, W. Gilbert, *Biochemistry* **1992**, *31*, 65–70.

[258] M. K. Uddin, Y. Kato, Y. Takagi, T. Mikuma, K. Taira, *Nucleic Acids Res.* **2004**, *32*, 4618–4629.

[259] Y. Krishnan-Ghosh, D. Liu, S. Balasubramanian, *J. Am. Chem. Soc.* **2004**, *126*, 11009–11016.

[260] *OriginPro 8.5.0G SR1*, **2010**, OriginLab Corporation, Northhampton, MA 01060 USA. `http://www.originlab.com`.

[261] A. T. Phan, *FEBS J.* **2010**, *277*, 1107–1117.

[262] Y. Wang, D. J. Patel, *Biochemistry* **1992**, *31*, 8112–8119.

[263] G. N. Parkinson, M. P. H. Lee, S. Neidle, *Nature* **2002**, *417*, 876–880.

[264] A. T. Phan, D. J. Patel, *J. Am. Chem. Soc.* **2003**, *125*, 15021–15027.

[265] N. Zhang, A. T. Phan, D. J. Patel, *J. Am. Chem. Soc.* **2005**, *127*, 17277–17285.

[266] Y. Wang, D. J. Patel, *Structure (Oxford, U. K.)* **1993**, *1*, 263–282.

[267] Y. Xu, Y. Noguchi, H. Sugiyama, *Bioorg. Med. Chem.* **2006**, *14*, 5584–5591.

[268] A. Ambrus, D. Chen, J. Dai, T. Bialis, R. A. Jones, D. Yang, *Nucleic Acids Res.* **2006**, *34*, 2723–2735.

[269] A. T. Phan, K. N. Luu, D. J. Patel, *Nucleic Acids Res.* **2006**, *34*, 5715–5719.

[270] K. N. Luu, A. T. Phan, V. Kuryavyi, L. Lacroix, D. J. Patel, *J. Am. Chem. Soc.* **2006**, *128*, 9963–9970.

[271] J. Dai, C. Punchihewa, A. Ambrus, D. Chen, R. A. Jones, D. Yang, *Nucleic Acids Res.* **2007**, *35*, 2440–2450.

[272] A. Matsugami, Y. Xu, Y. Noguchi, H. Sugiyama, M. Katahira, *FEBS J.* **2007**, *274*, 3545–3556.

[273] A. T. Phan, V. Kuryavyi, K. N. Luu, D. J. Patel, *Nucleic Acids Res.* **2007**, *35*, 6517–6525.

[274] J. Dai, M. Carver, C. Punchihewa, R. A. Jones, D. Yang, *Nucleic Acids Res.* **2007**, *35*, 4927–4940.

[275] K. W. Lim, S. Amrane, S. Bouaziz, W. Xu, Y. Mu, D. J. Patel, K. N. Luu, A. T. Phan, *J. Am. Chem. Soc.* **2009**, *131*, 4301–4309.

[276] J.-L. Mergny, L. Lacroix, *Curr. Protoc. Nucleic Acid Chem.* **2009**, *37*, 17.1.1–17.1.15.

[277] P. L. T. Tran, J.-L. Mergny, P. Alberti, *Nucleic Acids Res.* **2011**, *39*, 3282–3294.

[278] A. Guedin, P. Alberti, J.-L. Mergny, *Nucleic Acids Res.* **2009**, *37*, 5559–5567.

[279] A. I. Karsisiotis, N. M. Hessari, E. Novellino, G. P. Spada, A. Randazzo, M. Webba da Silva, *Angew. Chem. Int. Ed.* **2011**, *50*, 10645–10648; *Angew. Chem.* **2011**, *123*, 10833–10836.

[280] J. M. Nicoludis, S. P. Barrett, J.-L. Mergny, L. A. Yatsunyk, *Nucleic Acids Res.* **2012**, *40*, 5432–5447.

[281] J. M. Nicoludis, S. T. Miller, P. D. Jeffrey, S. P. Barrett, P. R. Rablen, T. J. Lawton, L. A. Yatsunyk, *J. Am. Chem. Soc.* **2012**, *134*, 20446–20456.

[282] H. Qin, J. Ren, J. Wang, E. Wang, *Chem. Commun.* **2010**, *46*, 7385–7387.

[283] S. Shipovskov, T. Karlberg, M. Fodje, M. D. Hansson, G. C. Ferreira, M. Hansson, C. T. Reimann, S. Al-Karadaghi, *J. Mol. Biol.* **2005**, *352*, 1081–1090.

[284] W. D. Cornell, P. Cieplak, C. I. Bayly, I. R. Gould, K. M. Merz, D. M. Ferguson, D. C. Spellmeyer, T. Fox, J. W. Caldwell, P. A. Kollman, *J. Am. Chem. Soc.* **1995**, *117*, 5179–5197.

[285] A. Pérez, I. Marchán, D. Svozil, J. Šponer, T. E. Cheatham III, C. A. Laughton, M. Orozco, *Biophys. J.* **2007**, *92*, 3817–3829.

[286] C. I. Bayly, P. Cieplak, W. Cornell, P. A. Kollman, *J. Phys. Chem.* **1993**, *97*, 10269–10280.

[287] P. Cieplak, W. D. Cornell, C. Bayly, P. A. Kollman, *J. Comput. Chem.* **1995**, *16*, 1357–1377.

[288] J. Wang, P. Cieplak, P. A. Kollman, *J. Comput. Chem.* **2000**, *21*, 1049–1074.

[289] R. Mera-Adasme, K. Sadeghian, D. Sundholm, C. Ochsenfeld, *J. Phys. Chem. B* **2014**, *118*, 13106–13111.

[290] T. Xue, S. Cui, Y. Wang, B. Yang, G. Chen, *Theor. Chem. Acc.* **2015**, *134*.

[291] *UCSF Chimera 1.10.2*, **2015**, Resource for Biocomputing, Visualization, and Informatics, University of California, San Francisco. `www.cgl.ucsf.edu/chimera`.

[292] E. F. Pettersen, T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt, E. C. Meng, T. E. Ferrin, *J. Comput. Chem.* **2004**, *25*, 1605–1612.

[293] *Gromacs 5.0.4*, **2015**. `www.gromacs.org`.

[294] S. Pronk, S. Páll, R. Schulz, P. Larsson, P. Bjelkmar, R. Apostolov, M. R. Shirts, J. C. Smith, P. M. Kasson, D. van der Spoel, B. Hess, E. Lindahl, *Bioinformatics* **2013**, *29*, 845–854.

[295] B. Hess, C. Kutzner, D. van der Spoel, E. Lindahl, *J. Chem. Theory Comput.* **2008**, *4*, 435–447.

[296] D. van der Spoel, E. Lindahl, B. Hess, G. Groenhof, A. E. Mark, H. J. C. Berendsen, *J. Comput. Chem.* **2005**, *26*, 1701–1718.

[297] B. Hess, H. Bekker, H. J. C. Berendsen, J. G. E. M. Fraaije, *J. Comput. Chem.* **1997**, *18*, 1463–1472.

[298] G. Bussi, D. Donadio, M. Parrinello, *J. Chem. Phys.* **2007**, *126*, 14101.

[299] Y. Wang, D. J. Patel, *Structure (Oxford, U. K.)* **1994**, *2*, 1141–1156.

[300] R. D. Gray, R. Buscaglia, J. B. Chaires, *J. Am. Chem. Soc.* **2012**, *134*, 16834–16844.

[301] A. Karsisiotis, P. Dillon, M. Webba da Silva, *Solution NMR structure of quadruplex d(TGGGTTTGGGTTGGGTTTGGG) in sodium conditions*, **2014**. `DOI10.2210/pdb2mfu/pdb`.

[302] A. I. Karsisiotis, M. Webba da Silva, *Solution NMR structure of the d(GGGTTTTGGGTGGGTTTTGGG) quadruplex in sodium conditions*, **2014**. `DOI10.2210/pdb2mft/pdb`.

[303] M. Adrian, B. Heddi, A. T. Phan, *Methods (San Diego, CA, U. S.)* **2012**, *57*, 11–24.

[304] K. Tanaka, A. Tengiji, T. Kato, N. Toyama, M. Shionoya, *Science* **2003**, *299*, 1212–1213.

[305] G. H. Clever, S. J. Reitmeier, T. Carell, O. Schiemann, *Angew. Chem. Int. Ed.* **2010**, *49*, 4927–4929; *Angew. Chem.* **2010**, *122*, 5047–5049.

[306] T. Nguyen, P. Håkansson, R. Edge, D. Collison, B. A. Goodman, J. R. Burns, E. Stulz, *New J. Chem.* **2014**, *38*, 5254–5259.

[307] V. Singh, M. Azarkh, T. E. Exner, J. S. Hartig, M. Drescher, *Angew. Chem. Int. Ed.* **2009**, *48*, 9728–9730; *Angew. Chem.* **2009**, *121*, 9908–9910.

[308] R. Welz, S. Müller, *Tetrahedron Lett.* **2002**, *43*, 795–797.

[309] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery Jr., J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, Ö. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski, D. J. Fox, *Gaussian 09, Revision D.01*, **2013**, Gaussian, Inc., Wallingford CT. `www.gaussian.com`.

[310] D. A. Case, J. T. Berryman, R. M. Betz, D. S. Cerutti, T. E. Cheatham III, T. A. Darden, R. E. Duke, T. J. Giese, H. Gohlke, A. W. Goetz, N. Homeyer, S. Izadi, P. Janowski, J. Kaus, A. Kovalenko, T. S. Lee, S. LeGrand, T. Luchko, R. Luo, B. Madej, K. M. Merz, G. Monard, P. Needham, H. Nguyen, H. T. Nguyen, I. Omelyan, A. Onufriev, D. R. Roe, A. Roitberg, R. Salomon-Ferrer, C. L. Simmerling, W. Smith, J. Swails, R. C. Walker, J. Wang, R. M. Wolf, X. Wu, D. M. York, P. A. Kollman, *AMBER 2014*, **2014**, University of California, San Francisco. `http://www.ambermd.org`.

[311] F.-Y. Dupradeau, A. Pigache, T. Zaffran, C. Savineau, R. Lelong, N. Grivel, D. Lelong, W. Rosanski, P. Cieplak, *Phys. Chem. Chem. Phys.* **2010**, *12*, 7821–7839.

[312] S. Zheng, Q. Tang, J. He, S. Du, S. Xu, C. Wang, Y. Xu, F. Lin, *J. Chem. Inf. Model.* **2016**, *56*, 811–818.

[313] I. Ivani, P. D. Dans, A. Noy, A. Pérez, I. Faustino, A. Hospital, J. Walther, P. Andrio, R. Goñi, A. Balaceanu, G. Portella, F. Battistini, J. L. Gelpi, C. Gonzalez, M. Vendruscolo, C. A. Laughton, S. A. Harris, D. A. Case, M. Orozco, *Nat. Methods* **2016**, *13*, 55–58.

[314] U. Essmann, L. Perera, M. L. Berkowitz, T. Darden, H. Lee, L. G. Pedersen, *J. Chem. Phys.* **1995**, *103*, 8577.

[315] *VMD 1.9.2*, **2014**, Theoretical and Computational Biophysics group, NIH Center for Macromolecular Modeling and Bioinformatics, at the Beckman Institute, University of Illinois at Urbana-Champaign. `www.ks.uiuc.edu/Research/vmd/`.

[316] W. Humphrey, A. Dalke, K. Schulten, *J. Mol. Graphics* **1996**, *14*, 33–38.

List of Figures

List of Tables

Acronyms

α	fraction folded, optical rotation
A	adenine
AAAS	American Association for the Advancement of Science
ABTS	2,2′-azino-bis(3-ethylbenzothiazoline-6-sulphonic acid)
ACS	American Chemical Society
ACT	activator
AFM	atomic force microscopy
APS	ammonium peroxodisulfate
bp	base-pair
BTT	5-(benzylthio)-1*H*-tetrazole
c	concentration
C	cytosine
Cap A/B	capping A/B
CD	circular dichroism
CE	2-cyanoethyl
CEP	2-cyanoethyl *N*,*N*-diisopropylphosphoramidite
CPG	controlled pore glass
CW	continous wave
CuAAC	Cu^+ catalysed alkyne azide 1,3-dipolar cycloaddition
δ	chemical shift
d	distance
d	2′ deoxy (-nucleoside)
D.I.T.	data integration time
DCA	dichloroacetic acid
deconv.	deconvoluted
DIPEA	diisopropyl ethyl amine
DMAP	4-dimethylamino pyridine
DMF	dimethylformamide
DMSO	dimethyl sulfoxide
DMSO	dimethyl sulfoxide

DMT	4,4′-dimethoxytrityl
DNA	deoxyribonucleic acid
ds	double-strand
ϵ	extinction coefficient
ECP	effective core potential
edta	ethylenediaminetetraacetic acid disodium salt
EDTA	ethylenediaminetetraacetate
EDTP	ethylenediaminetetraphosphoric acid
EI	electron ionization
EPR	electron paramagnetic resonance
eq.	equivalent(s)
ESI	electrospray ionisation
ESP	electrostatic potential
FFT	fast Fourier transform
G	guanine
GMP	guanosine monophosphate
GNA	glycol nucleic acid
HF	Hartree-Fock
Hz	Hertz
J	coupling constant
LiCaco	lithium cacodylate
LNA	locked nucleic acid
M	mol L^{-1}
MD	molecular dynamics
MMFF	molecular mechanics force field
MeCN	acetonitrile
MeOH	methanol
mRNA	messenger ribonucleic acid
NEt_3	triethyl amine
$NEt\mathit{i}Pr_2$	diisopropyl ethyl amine
NMM	*N*-methyl mesoporphyrin IX
NMR	nuclear magnetic resonance
NaCaco	sodium cacodylate
PAGE	polyacrylamide gel-electrophoresis
PBC	periodic boundary conditions
PDB	protein data bank
PNA	peptide nucleic acid
PPIX	protoporphyrin IX

ppm	parts per million
R	rotational strength
RESP	restricted electrostatic potential
RNA	ribonucleic acid
RP-HPLC	reversed phase high pressure liquid chromatography
RP-SPE	reversed-phase solid phase extraction
RSC	Royal Society of Chemistry
rt	room temperature
ss	single-strand
θ	ellipticity
$T_{\mathbf{1/2}}$	non-equilibrium denaturation temperature
$T_{\mathbf{m}}$	equilibrium melting temperature
T	temperature
T	thymine
TBA	thrombin binding aptamer
TDS	thermal difference spectrum/spectra
TEAA	triethyl ammonium acetate
TEMED	N,N,N',N'-tetramethylethylenediamine
THF	tetrahydrofurane
TLC	thin layer chromatography
UNA	unlocked nucleic acid
UV-VIS	ultraviolet-visible
UV	ultraviolet