

Thesis for the degree of Doctor of Philosophy

FLUORESCENT NUCLEOBASE ANALOGUES
AND THEIR USE FOR INVESTIGATING DNA
INTERACTIONS

ANKE DIERCKX

Department of Chemical and Biological Engineering
Chalmers University of Technology
Gothenburg, Sweden 2014

Fluorescent Nucleobase Analogues and their use for Investigating DNA
Interactions

ANKE DIERCKX

ISBN 978-91-7597-020-2

© ANKE DIERCKX, 2014

Doktorsavhandlingar vid Chalmers tekniska högskola

Ny serie nr 3701

ISSN 0346-718X

Department of Chemical and Biological Engineering

Chalmers University of Technology

SE-412 96 Gothenburg

Sweden

Telephone + 46 (0)31-772 1000

Cover: Left: Chemical structure of adenine analogues A^T, qA and a member of the A^{7T}-family. Also shown is a putative structure of qA inside a B-DNA decamer. Right: Tricyclic cytosine analogue FRET-pair tC^O (purple) and tC_{nitro} (orange) and their framework placed into a B-DNA decamer with a separation of 6 bp.

Chalmers Reproservice

Gothenburg, Sweden 2014

FLUORESCENT NUCLEOBASE ANALOGUES AND THEIR USE FOR INVESTIGATING DNA INTERACTIONS

Anke Dierckx

Department of Chemical and Biological Engineering
Chalmers University of Technology

Abstract

Ever since unravelling the structure of DNA, an expanding research field has emerged with ongoing efforts dedicated to increase our understanding of the molecule of life. Since the natural nucleobases are virtually non-emissive, it has been a challenge for decades to ‘light up’ DNA/RNA in order to investigate their properties utilizing fluorescence techniques. This thesis focuses on fluorescent nucleobase analogues (FBAs) as probes for fluorescently labeling DNA and investigating its interactions, for example, with proteins. These artificial nucleobases attempt to closely mimic the characteristics of natural bases, while introducing fluorescence properties to the system.

The first part of this work comprises the characterization of both new and established FBAs. Photophysical and base-mimicking properties of two fluorescent adenine analogues, triazole adenine (A^T) and quadracyclic adenine (qA) are presented. Both exhibit promising features compared to the widely used commercially available adenine mimic 2-aminopurine (2-AP). Even though A^T shows promising emission as a monomer and in certain DNA surroundings, it destabilizes the B-DNA duplex structure, most likely due to its C8-triazole extension. In order to overcome this effect, a new family of triazole adenine analogues extended on the 7-position was synthesized and photophysically characterized. The second thoroughly characterized adenine analogue, qA, is moderately fluorescent both as a monomer and inside DNA but in contrast to A^T , the two-ring extension on qA is suggested to be well accommodated in the major groove and renders the DNA-duplex unperturbed or even stabilized, depending on the surrounding sequence. Finally, the photostability of tC, an already established FBA of the tricyclic cytosine family, was investigated. The latter yields a single photoproduct with a decreased fluorescence, which destabilizes DNA duplexes.

In the second part of this work, the tricyclic cytosine FBA FRET-pair, tC^O - tC_{nitro} was applied in exploring the role of the mammalian transcription factor A in mitochondrial transcription. Furthermore, it was called upon to help resolve the order of events in which the different components of the transcription machinery initiate transcription.

Keywords: Fluorescence quantum yield, B-DNA, fluorescent nucleobase analogue, duplex stability, circular dichroism, adenine analogue, FRET

List of Publications

The thesis is based on the work described in the following articles, referred to by Roman numerals in the text:

- I Anke Dierckx, Peter Dinér, Afaf H. El-Sagheer, Joshi Dhruval Kumar, Tom Brown, Morten Grøtli and L. Marcus Wilhelmsson
Characterization of photophysical and base-mimicking properties of a novel fluorescent adenine analogue in DNA
Nucleic Acids Research, **2011**, 39, 4513-4524.
- II Christopher P. Lawson[§], Anke Dierckx[§], Francois-Alexandre Miannay, Eric Wellner, L. Marcus Wilhelmsson and Morten Grøtli
Synthesis and photophysical characterization of new fluorescent triazole adenine analogues
Submitted to Organic & Biomolecular Chemistry
[§]Both authors contributed equally to this work
- III Anke Dierckx[§], Francois-Alexandre Miannay[§], Nouha Ben Gaied, Søren Preus, Markus Björck, Tom Brown and L. Marcus Wilhelmsson
Quadracyclic adenine: A non-perturbing fluorescent adenine analogue
Chemistry A European Journal, **2012**, 18, 5987-5997.
[§]Both authors contributed equally to this work
- IV Søren Preus, Søren Jønck, Michael Pittelkow, Anke Dierckx, Thitinun Karpkird, Bo Albinsson and L. Marcus Wilhelmsson
The photoinduced transformations of fluorescent DNA base analogue tC triggers DNA melting
Photochemical & Photobiological Sciences, **2013**, 12, 1416-1422.
- V Yonghong Shi, Anke Dierckx, Paulina H. Wanrooij, Sjoerd Wanrooij, Nils-Göran Larsson, L. Marcus Wilhelmsson, Maria Falkenberg and Claes M. Gustafsson
Mammalian transcription factor A is a core component of the mitochondrial transcription machinery
Proceedings of the National Academy of Sciences, **2012**, 109, 16510-16515.
- VI Viktor Posse, Emily Hoberg, Anke Dierckx, Saba Shahzad, Camilla Koolmeister, Nils-Göran Larsson, L. Marcus Wilhelmsson, B. Martin Hällberg and Claes M. Gustafsson
The amino terminal extension of mammalian mitochondrial RNA polymerase ensures promoter specific transcription initiation
Nucleic Acids Research, **2014**, 42, 3638-3647.

Contribution Report

- Paper I: Main responsible for performing the photophysical characterization and writing the paper.
- Paper II: Main responsible for performing the photophysical characterization. Wrote the paper together with C.P.L..
- Paper III: Performed photophysical characterization and wrote paper together with F.A.M.. TDDFT calculations were done by S.P. and quantum yield measurements on the qA monomer by F.A.M..
- Paper IV: Performed fast photoconversion of tC in duplex DNA, recorded melting curves and CD spectra. Proofread paper.
- Paper V: Performed and analyzed spectroscopic measurements. Involved in writing corresponding paragraphs and proofread paper.
- Paper VI: Performed and analyzed spectroscopic measurements. Proofread paper.

Table of Contents

1	Introduction	1
2	Background	5
2.1	Deoxyribonucleic acid	5
2.2	Fluorescence and DNA	8
2.2.1	DNA ligands	9
2.2.2	Covalent modifications	9
3	Theory and Methodology	13
3.1	Interaction of light and matter	13
3.2	Absorption	15
3.3	Circular dichroism	16
3.4	Fluorescence	17
3.4.1	Fluorescence principles	17
3.4.2	Steady state fluorescence	20
3.4.3	Steady state fluorescence anisotropy	21
3.4.4	Time-resolved fluorescence	22
3.4.5	Förster resonance energy transfer (FRET)	24
3.4.6	The tricyclic cytosine FRET-pair	26
3.5	DNA melting	28
4	Results	31
4.1	Characterization of fluorescent nucleobase analogues	31
4.1.1	Characterization of fluorescent adenine analogues	31
4.1.2	Photostability of the tricyclic cytosine analogue tC	44
4.2	Application of the tricyclic cytosine FRET-pair to investigate protein-DNA interactions	46
5	Conclusion and outlook	51
6	Acknowledgements	55
7	References	57

1 Introduction

In 1869 a young Swiss physician, Friedrich Miescher, discovered a phosphorous-containing substance in the nuclei of leucocytes, which he named nuclein. The latter mainly consists of chromatin, a complex of chromosomal proteins and deoxyribonucleic acid, DNA, later identified as the carrier of our genetic information.^[1,2] A major breakthrough towards better understanding its structure and therefore its way of function came in 1953 when Watson and Crick proposed the famous double helical structure, in which two complementary strands consisting of the four natural nucleobases (adenine, guanine, cytosine and thymine), deoxyribose sugar and phosphate groups are paired in the shape of a right winding corkscrew.^[3] Another great development in nucleic acids research was the automation of DNA synthesis, allowing short (<200 bp) synthetic oligonucleotide sequences to be ordered from various commercial sources nowadays at low costs.^[4,5] Furthermore, the prestigious ‘human genome project’, which aimed at extracting the sequence of the entire human genome, was enabled due to the development of DNA sequencing methods.^[6] Even though much progress has been made, there are many aspects yet to be understood about this surprisingly simple, yet so important molecule. Ribonucleic acid (RNA) has also attracted much scientific interest as it has become apparent that it is not merely a messenger between our genetic information and proteins. In fact, important roles have been suggested for non-coding RNA sequences in gene regulation and other regulatory processes in the cell.^[7-10]

A quick search on the world wide web revealed that DNA is referred to nowadays as the instruction manual of every living cell or the blueprint of life to name but a few, thus it is no surprise that this molecule still attracts massive research attention today. Furthermore it has also generated substantial interest as a building block for nano-constructs due to its well-ordered structure and predictable assembly.^[11-13]

Since many complex processes are required for both DNA and RNA to execute their biological functions, it is a field of ongoing interest to provide a better understanding of the structure and dynamics of these molecules. Another major area of interest is the provision of clear insights into the complex network of interactions of DNA with other molecules such as proteins or RNA.

Besides a quest for better understanding the processes DNA and RNA are involved in, there is also a broad interest in improving the techniques which allow the retrieval of such information. There are different approaches to study the above mentioned aspects of DNA and RNA, such as the high resolution methods x-ray crystallography^[14] and Nuclear Magnetic Resonance (NMR)^[15]. Although

powerful at providing detailed structural information, a limitation of x-ray crystallography lies in the need for obtaining a crystallized sample, which hardly ever is a trivial task. Furthermore, results obtained for a crystalline molecule may differ significantly from the physiological conditions in which it is usually found. NMR, on the other hand, can be recorded in solution and is therefore a powerful tool to derive solution structures of nucleic acids. Additionally, it can yield dynamic information about the molecule. However, NMR suffers from the increasing complexity of the obtained signal with molecule size.

Another powerful tool for retrieving information about DNA and RNA is fluorescence, which is very sensitive, straightforward and requires only small amounts of sample in contrast to the above named methods. Moreover, it allows for real-time measurements in solution. Numerous proteins have been studied taking advantage of the fluorescence of the amino acid tryptophan.^[16] The nucleobases, the natural building blocks of DNA and RNA, are however virtually non-fluorescent ($\Phi_f \sim 10^{-4}$, $\tau_f < 1$ ps in aqueous solution at room temperature)^[17,18], which is probably a convenient way for nature to protect us from genetic damage caused by UV-light.^[19] As a result, DNA needs to be labelled with a fluorophore in order to study its properties using fluorescence techniques, thereby allowing a good signal to noise ratio due to negligible background contributions from the natural bases.

There are different strategies for fluorescently tagging nucleic acids, each with specific advantages depending on the aim of the study. Easy non-covalent labelling can be achieved by groove binders and intercalators.^[20-23] Covalent modifications are often achieved by tethering bright commercially available dyes (for example Cy-dyes, fluorescein and Alexa-dyes) to the DNA *via* a linker, allowing studies down to the single-molecule level.^[24-26] However, a number of intercalating dyes usually modify and elongate the DNA structure since they need to fit into the base-stack. Also bulky external dyes can interfere with for example DNA-protein interactions due to steric hindrance. Furthermore, the emission of external chromophores is usually insensitive to subtle structural changes inside the DNA. This is where fluorescent nucleobase analogues (FBAs) come in handy. These probes are significantly fluorescent (even though less bright than commercially available external dyes) and mimic the natural nucleobases, thereby minimally perturbing the duplex structure. FBAs can be specifically incorporated into DNA and for most members, their emission is sensitive to changes in the direct environment so they can report on local structural alterations in the DNA duplex.^[27-31]

One of the most widely utilized FBAs to date, 2-aminopurine (2-AP), was among the first to be discovered (Figure 1).^[32] 2-AP is an adenine isomorph that is

highly fluorescent as a monomer in solution, but is significantly quenched inside single- and double-stranded nucleic acid systems.^[32] It moderately destabilizes the DNA duplex^[33] and can form base-pairs with not just thymine but also cytosine, albeit less stable.^[34-36] Different classes of base analogues have been developed and explored over the last few decades, each with its own specific qualities. However, their design remains problematic due to the difficulties associated with prediction of emissive behaviour based on chemical structure. The development of FBAs with improved properties, *e.g.* higher fluorescence quantum yields, higher extinction coefficients, selective base-pairing and which cause minimal or no destabilization is an area of ongoing interest. Therefore the development and thorough characterization of new FBAs is highly significant to the field.

The focus of the first part of this thesis is concerned with the spectroscopic characterization of both new and already introduced FBAs (papers I-IV). Two fluorescent adenine analogues, triazole adenine (A^T) and quadracyclic adenine (qA) were thoroughly studied both as monomers and inside DNA (Figure 1). Due to its good emissive properties reported in a previous study,^[37] A^T was investigated as a nucleoside and incorporated into oligonucleotides with different base surroundings (paper I). A^T shows promising emissive properties evidenced by high quantum yields it displays in DNA depending on the base-surroundings. However, it causes a moderate destabilization of the duplex, which can probably be attributed to the C8 modification. The latter could force A^T to adopt a *syn* conformation around the glycosidic bond. As a follow-up study, trying to overcome this undesired destabilizing effect, a new series of triazole adenine analogues (A^{7T} -family) was designed, bearing the modification on the purine 7-position instead (paper II, Figure 1).

In an attempt to discover an adenine equivalent to the promising previously characterized tricyclic cytosines tC^O , tC and tC_{nitro} ,^[38-42] an extended quadracyclic adenine analogue, qA, was investigated (paper III, Figure 1). Even though qA does not show the stable emission inside DNA as was reported for both tC ^[38] and tC^O ^[39], it demonstrates a similar increase or preservation of duplex stability depending on the nature of the surrounding bases.

Finally, to approach the idea of FBAs for single-molecule studies, a good understanding of their photostability is vital. Therefore, the photodegradation of the tricyclic cytosine analogue tC (Figure 1) was characterized as a monomer and inside DNA. This led to the surprising discovery that intense irradiation can trigger its photoconversion to a product which destabilizes DNA duplexes (paper IV).

The second part of this work is concerned with the applications of FBAs, more specifically the FBA FRET-pair, tC^O - tC_{nitro} (Figure 1), in investigating

mammalian mitochondrial transcription (papers V-VI). A study by Börjesson *et al.* published in 2009 reported on the responsiveness of this FRET-pair to alterations in both distance and orientation due to its rigid positioning in the DNA duplex.^[40] The system is used here to report on local structural alterations in the DNA close to the transcription start site upon binding of the mammalian mitochondrial transcription factor A (TFAM) (paper V). In a follow-up study, the FRET-system was also used to help determine the sequence of events caused by TFAM, the TFB2M-factor and human mitochondrial RNA polymerase (POLRMT) at the promoter to initiate transcription (paper VI).

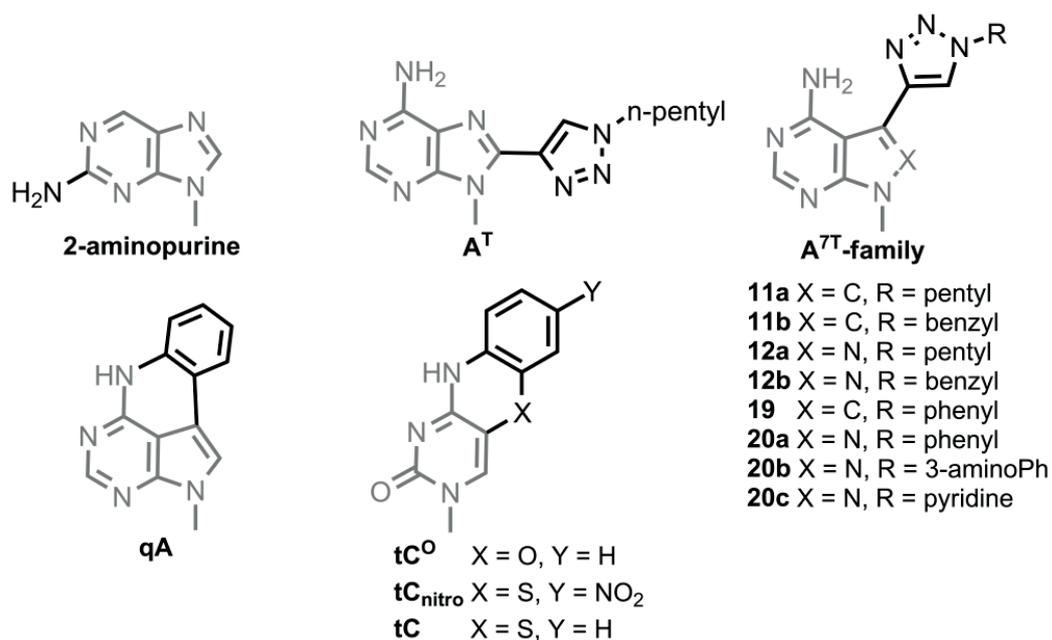


Figure 1 Structure of 2-aminopurine (2-AP) and the adenine analogues discussed in this work: A^T (paper I), the A^{7T}-family (paper II) and qA (paper III). Also the already established tricyclic cytosine family is shown (studied/applied in papers IV-VI). The backbone of natural adenine/cytosine is shown in grey.

2 Background

Since DNA is the central molecule in this thesis, its structure and function will be discussed. Fluorescent probes, more specifically fluorescent base analogues, will also be highlighted in this chapter as tools to study different aspects of DNA and RNA.

2.1 Deoxyribonucleic acid

In the 20th century, molecular genetics greatly advanced our understanding of the composition and structure of genes as well as how they exert their function. As mentioned in the introduction, it was the young Swiss physician Friedrich Miescher who discovered the phosphorus-containing substance nuclein in the nucleus of white blood cells in 1869.^[2] However, it was not until 1944 that Oswald Avery, Colin MacLeod and Maclyn McCarty reported a transformation experiment that provided the proof that DNA, one of the main components of nuclein, is the material that carries genetic information.^[43] By then it was also known that DNA was built of nitrogenous bases (adenine (A), cytosine (C), guanine (G) and thymine (T)), as well as phosphoric acid and the sugar deoxyribose (Figure 2). Adenine and guanine are referred to as purines whereas cytosine and thymine are referred to as pyrimidines according to the parent molecules they are related to. One main piece of the puzzle left to discover by the end of the 1940s was the structure of DNA and how it executes its function as gene carrier.^[1]

In 1950 Erwin Chargaff found the amount of purines and pyrimidines in DNA from various sources to be roughly equal.^[44] Furthermore, this also was true for the amounts of adenine and thymine as well as guanine and cytosine. Another crucial piece of information came from x-ray fibre diffraction patterns made by Rosalind Franklin and Maurice Wilkins.^[45,46] This led James Watson and Francis Crick to link the information together and publish their famous paper ‘Molecular structure of nucleic acids’ in 1953, in which they describe the double helical character of DNA.^[1,3]

The DNA double helix is built up of nucleotides, consisting of a nucleoside (a nucleobase linked to β -D-2'-deoxyribose *via* a glycosidic bond) and a phosphate group. These nucleotides are linked together by phosphodiester bonds to form a polynucleotide or single-stranded DNA chain. Usually, the nucleotides adopt the *anti*-conformation about the glycosidic bond, which means that the base is turned

away from the sugar ring to avoid steric hindrance between the sugar and either N3 in purines or the carbonyl oxygen in pyrimidines (Figure 2a). A single DNA sequence is able to hybridize to a complementary strand, having the sequence of bases running in the opposite direction (antiparallel), creating a double helix. This means that at both ends of the duplex, the 5'-phosphate group of the last nucleotide of one strand and the 3'-hydroxyl group of the other strand can be found. The complementary strand contains a sequence of nucleotides so that the nucleobases adenine and thymine as well as cytosine and guanine are always paired between both strands. This means that each sequence contains the information necessary to synthesize a complementary strand.^[1,47,48]

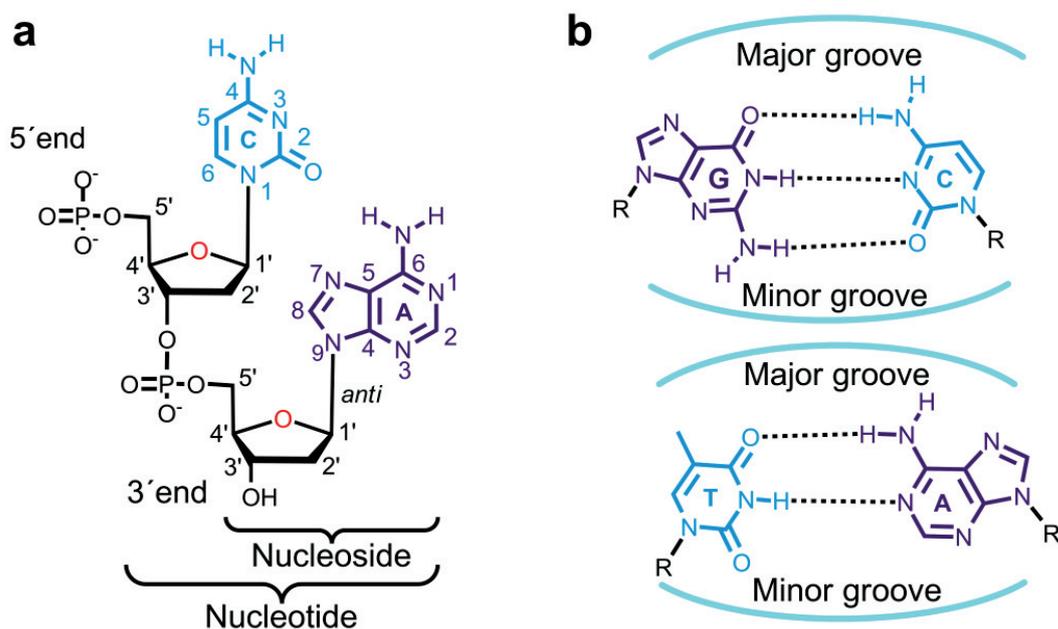


Figure 2 (a) Dinucleotide of 2'-deoxyadenosine linked to 2'-deoxycytidine by a phosphodiester bond. (b) Base pairing pattern between the natural nucleobases guanine (G) and cytosine (C) as well as between thymine (T) and adenine (A). Also the major-and minor groove sides of each base-pair are indicated. R indicates 2'-deoxyribose.

Hydrogen bonds are formed between the complementary bases contained in both strands of the DNA duplex: two hydrogen bonds between each adenine-thymine pair and three between cytosine and guanine (Figure 2b). Furthermore, π - π -stacking (a form of van der Waals interaction) and hydrophobic interactions between the aromatic nucleobases contribute significantly to the duplex stability. At physiological pH, the DNA molecule is a large polyanion because of the negatively charged phosphate groups in its backbone. This is why salt concentration has a large impact on duplex stability, since positively charged salt ions such as Na^+ or Mg^{2+} can shield the negative backbone charges, decreasing the

repulsion between both strands. Also duplex length, sequence, type of solvent and pH play a vital role in the stability of a DNA molecule. The DNA duplex can adopt different conformations which are greatly influenced by hydration and ions. In general DNA double helices are found in the B-form under physiological conditions in aqueous solution (Figure 3), characterized by a right winding helical twist about 36° and a 3.4 \AA rise between the base-pairs. About 10-10.6 base-pairs are contained per helical turn (depending on local sequence).^[1,47,48] Along the surface of the helix two distinct grooves can be detected, which are similar in depth but different in width, known as the major (11 \AA wide) and minor (6 \AA wide) groove.^[49] Other helical forms that exist are Z-DNA and A-DNA. The latter also is a right-winding helix, but more compressed than B-DNA, which can be obtained in dehydrated samples and DNA-RNA hybrids.^[47,48] Z-DNA on the other hand is a left winding helix and can be formed mainly in alternating purine-pyrimidine sequences such as poly[dG-dC]·poly[dG-dC] at high ionic strength.^[47,48,50] Several lines of evidence, such as the identification of Z-DNA binding proteins, suggest a biological role for this helical form *in vivo*.^[50] Also other non-B-DNA forms have been identified, such as the G-quadruplex, triplex and i-motif to name a few.^[50]

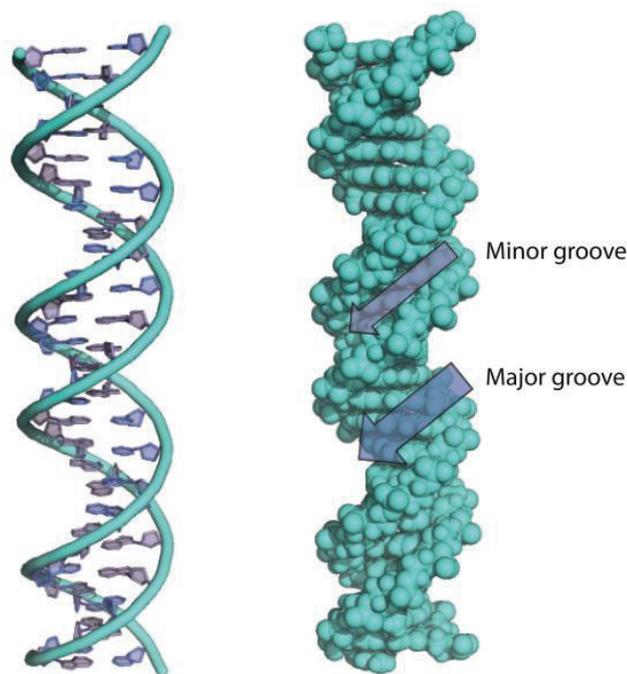


Figure 3 B-DNA double helix. Schematic representation (left) as well as space-filling model (right) with minor and major groove indicated.

After describing its basic structure, the function of DNA and connection to other processes in the cell will be briefly discussed here. First of all it is important to mention that the DNA molecule occurs in different forms depending on cell

type. Prokaryotes have a single, typically circular DNA molecule, which is associated with proteins. Eukaryotes, on the other hand, have their DNA organized in chromosomes, which are highly ordered complexes of the DNA double helix with proteins, such as histones. This means that complex dynamic processes must be involved for the information contained in the DNA sequence to become available. The central dogma of information transfer in cells can be summarized as follows: the sequence of a gene coded in DNA is transcribed into RNA, for example messenger RNA (mRNA) which is then processed and translated by ribosomes into proteins which can execute a specific function in the cell (Figure 4).^[1] The process of transcription in mammalian mitochondria is investigated in paper V and paper VI. Of course this is not the only process in which the genetic code needs to be read. Also upon cell division, all the genetic information needs to be replicated in order to provide both daughter cells with the same instruction manual. All these complex processes require interactions of DNA with a whole army of proteins such as transcription factors, polymerases and other molecules. It is apparent that dynamic DNA processes are constantly occurring in cells and that our genetic material largely exceeds the role of a passive information carrier.

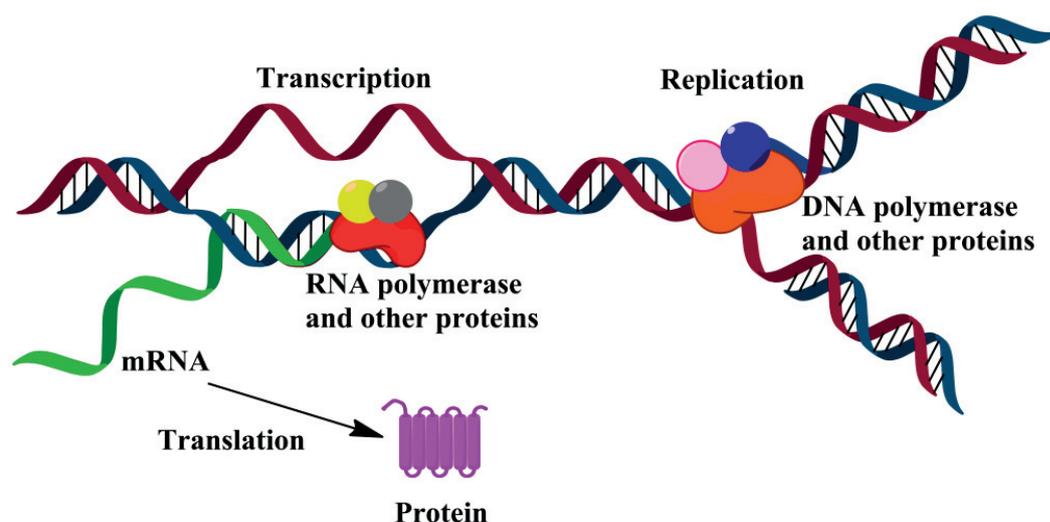


Figure 4 Schematic representation of replication and transcription of DNA as well as translation of messenger RNA.

2.2 Fluorescence and DNA

Due to the virtually non-emissive nature of the natural DNA/RNA building blocks,^[17,18] fluorophores need to be introduced into these systems in order to study their properties using fluorescence. In the introduction, main strategies to

render DNA fluorescent were briefly touched upon. This section divides these into non-covalent (DNA ligands) and covalent modifications of which fluorescent nucleobase analogues (FBAs) receive special attention as covalent modifications according to their relevance for this thesis.

2.2.1 DNA ligands

DNA can be easily labelled non-covalently with fluorophores that bind more or less non-specifically by hydrophobic and/or by electrostatic interactions with the negatively charged phosphate backbone. DNA ligands are usually planar, positively charged aromatic molecules (Figure 5). Furthermore, their quantum yield is typically increased several orders of magnitudes upon binding to DNA, yielding a good signal to noise ratio. Intercalators, such as oxazole yellow (YO), its homodimer YOYO^[23] and ethidium bromide^[20] insert themselves between two DNA bases, thereby elongating the duplex to some extent (Figure 5). Other examples of DNA ligands are groove binders such as 4',6-diamidino-2-phenylindole (DAPI)^[21] and Hoechst^[22]. Often, DNA ligands are used as an easy stain for visualizing DNA after electrophoresis or in microscopy. They are also used in artificial light harvesting in DNA nanoconstructs.^[51] However, the binding is rather non-specific and is a distribution, which makes it difficult to know where and exactly how many chromophores that are attached to each DNA duplex.

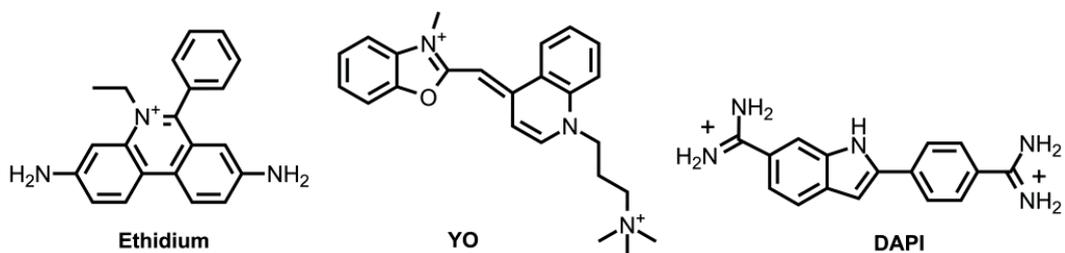


Figure 5 Examples of DNA ligands such as the intercalators ethidium and YO as well as the groove binder DAPI.

2.2.2 Covalent modifications

The most common way of covalently labelling nucleic acids is by tethering an external probe to it with a flexible hydrocarbon linker. In this way almost any commercially available bright dye can be attached as a step during solid-phase DNA synthesis or as a post-synthetic step. Examples are fluorescein, carboxytetramethylrhodamine (TAMRA) and the series of cyanine dyes or Alexa

dyes (Figure 6). Their brightness allows detection down to single-molecule level.^[24-26] However, as for a number of intercalating dyes, bulky external fluorophores can perturb the DNA and its interactions with other molecules, such as proteins or other ligands. Dyes have been shown to interact with DNA, for example through electrostatic interactions with the backbone, intercalation or end-stacking.^[52-56] Furthermore, the long linkers frequently used can contribute to dye diffusion and reorientation,^[53,56,57] increasing the difficulty of interpreting anisotropy or quantitative FRET experiments. Additionally, the positioning of the dye outside of the base-stack limits the retrieval of local site-specific information.

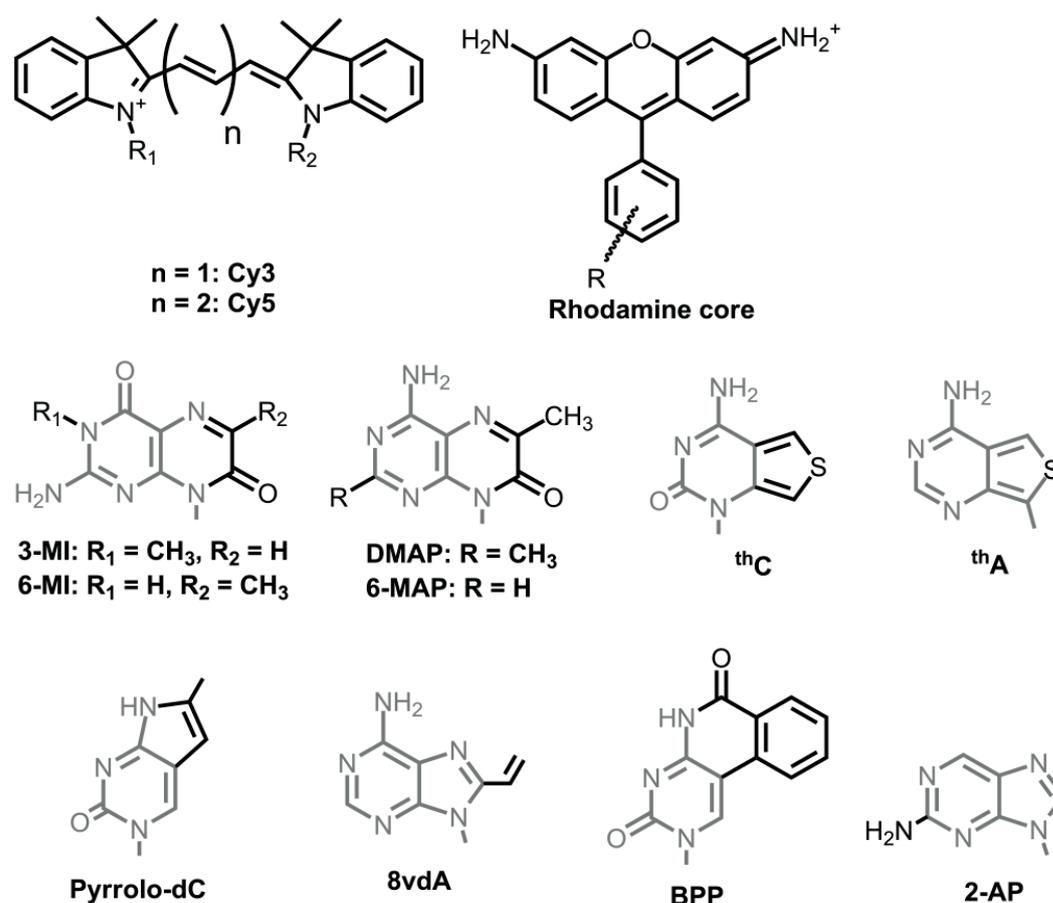


Figure 6 Examples of covalent external dyes for DNA (top row) as well as of previously characterized FBAs (bottom two rows), for which the basic skeleton of the natural nucleobase is shown in grey.

The main focus of this thesis is on the use of FBAs, covalently incorporated base-mimics, for rendering DNA fluorescent. In this work, the term FBAs is used for molecules that closely resemble one of the natural nucleobases in terms of hydrogen bonding capacity, while minimally perturbing the duplex integrity (conformation and stability) and at the same time introducing fluorescence into the nucleic acid system. Even though FBAs cannot compete with the external

commercial fluorophores in terms of brightness, they can be incorporated close to or at the very site of interest in the DNA helix, while preserving the native conformation of the DNA. The emission of most FBAs is sensitive to changes in their direct environment, making them responsive to local alterations. Furthermore, their more rigid incorporation into the base-stack allows them to report on the motion of the DNA helix, rather than the dye itself.

Over the last few decades, FBAs for each of the nucleobases have been reported and to date this research field continues to evolve rapidly (for a detailed overview see ^[27-31,58]). Depending on the type of experiment, different FBAs may be more suitable, based on the sensitivity of their emission to duplex structure, hydrogen bonding (with a complementary/mismatch base), environment (pH, base stacking), interactions of the DNA with other molecules such as polymerases, *etc.* However, the design of FBAs for each specific need has proven challenging. One faces structural constraints in order to leave the natural B-DNA as unperturbed as possible and preserving the hydrogen-bonding capacity of the base. Furthermore, there is a poor understanding of the correlation between structure and emissive properties of these molecules. Frequently, FBAs show a significant decrease in fluorescence quantum yield upon incorporation into nucleic acid systems. Exceptions are the adenine analogues 2-(3-phenylpropyl)adenosine (A-3CPh) and 2-(4-phenylbutyl)adenosine (A-4CPh), which show an increase in quantum yield upon incorporation into RNA^[59] as well as the tricyclic cytosines tC and tC^O, whose emission is relatively stable^[38,39].

The pteridines are a well-studied group of FBAs with the guanine analogues 6-MI and 3-MI^[60] as well as the adenine analogues 6MAP and DMAP^[61] as the most promising members (Figure 6).^[27,62] Although being moderately fluorescent inside DNA, they cause a destabilization of the duplex (with exception of 6-MI), similar to a single base pair mismatch for 3-MI.^[61,62] Other groups of FBAs comprise the pyrimidine analogues developed by Tor *et al.*,^[63-67] as well as their recently developed emissive RNA alphabet (thA, thG, thC and thU)^[68] (thA, thC shown in Figure 6). The base-discriminating fluorescent (BDF) nucleosides designed by Saito and co-workers could, as the name suggests, be applied for analysing single nucleotide polymorphisms (SNP).^[58] One example is the cytosine analogue benzopyridopyrimidine (BPP), which can form stable base-pairs with both adenine and guanine (Figure 6). Incorporation of a vinyl group on C8 of adenine produces an environmental sensitive fluorescent nucleobase, 8-vinyl-deoxyadenosine (8vdA), which is only minimally perturbing the duplex structure (Figure 6).^[69] The size-expanded DNA alphabet designed by Kool *et al.*, although perturbing the DNA duplex, has been practical for investigating steric effects in DNA.^[70-72] Pyrrolo-dC, a bicyclic cytosine analogue, hybridizes selectively with guanine and preserves duplex stability (Figure 6). It was formed by accident

during solid-phase synthesis upon attempting to incorporate the analogue furano-dT into DNA.^[73] A number of pyrrolo-dC-derivatives have also been developed.^[74-77]

One of the first successful and still most widely used FBAs is the adenine analogue 2-aminopurine (2-AP, Figure 6)^[32], which over the years has found its way into countless applications involving DNA and more recently also RNA.^[28-30] Its lowest energy absorption band is redshifted compared to natural adenine (~ 303 nm in water), allowing for selective excitation, whereas its emission is centred around 370 nm. Even though 2-AP shows a high fluorescence quantum yield in solution ($\Phi_f=68\%$ in water), it is quenched almost a hundred fold upon incorporation into nucleic acids, depending on the sequence environment.^[32,78] As was mentioned above, this is a common property for FBAs. Furthermore 2-aminopurine forms base-pairs with not only thymine, but also with cytosine, albeit less stably.^[34-36] As has been reported for other FBAs, 2-AP moderately destabilizes the DNA duplex^[33].

Over the past decades FBAs have been used in numerous applications, of which only a few will be highlighted here (for a more detailed overview see^[27-31]). Most of these exploit the sensitivity of the FBAs' emission to their microenvironment. Applications involve the investigation of certain aspects of DNA structure or dynamics. 6MAP has for example been used to analyse the premelting transitions of DNA A-tracts.^[79] Other examples involve 2-AP in studies concerning the effect of cations and sequence on base-stacking interactions at abasic sites^[80] or the dynamics of 2-AP mismatches in DNA^[81]. A second type of applications puts more focus on the interaction of DNA with other molecules. This is illustrated nicely by the incorporation of 2-AP in the promoter sequence to which bacteriophage T7 RNA polymerase binds, yielding information concerning kinetics of the promoter binding and open complex formation.^[82] Another application is the use of the guanine analogue 3-MI to study HIV-1 integrase.^[83,84] Recently, FBAs also started making their way into the field of nanotechnology. A first report was published in 2009 and involves the fluorescent cytosine analogue tC^O, reporting on the local stability of self-assembling DNA hexagons.^[85] FBAs are also becoming more and more important in the expanding field of RNA. An example is the use of 2-AP to study riboswitches, non-coding RNA elements that control gene expression, as a result of binding of small molecules.^[86,87] 6-phenylpyrrolocytosine (PhpC) was applied to follow cellular trafficking of siRNAs, whose gene silencing activity was virtually unaltered compared to natural siRNA.^[88] Also protein-RNA interactions have been studied with the help of FBAs. A nice illustration is a study concerning the change in emissive properties of 2-AP during the RNA editing process carried out by adenosine deaminases acting on RNA (ADARs).^[89,90]

3 Theory and Methodology

In this chapter the methods used to perform the experiments presented in this thesis will be introduced as well as their underlying theory. To begin with, the interaction of light and matter will be discussed briefly, leading to the description of absorption, circular dichroism and fluorescence spectroscopy. Furthermore, a section is dedicated to the principle of Förster Resonance Energy Transfer (FRET) and its application in DNA. Experimental details are described in the according sections of the published papers included in this thesis. For a thorough description of photochemistry^[91-93], circular dichroism^[94] and fluorescence techniques^[16,93,95], the reader is referred to the corresponding references.

3.1 Interaction of light and matter

The interaction of electromagnetic radiation with matter provides spectroscopists with desired information about the molecules or atoms they study. In order to describe this interaction, electromagnetic radiation must be treated as having a wave as well as particle character. In terms of classical physics, the electromagnetic field can be seen as a harmonic wave of an oscillating electric and magnetic field, travelling at the speed of light ($c=3.10^8$ m s⁻¹). The electric and magnetic fields oscillate in phase with a wavelength λ and frequency ν and describe waves which are perpendicular to each other and to the direction of propagation. The radiation can, however, also be described as a flow of photons, which can be seen as small energy packages with energy

$$E = h\nu \quad (1)$$

where h is Planck's constant ($h=6.626.10^{-34}$ J s) and ν is related to the speed of light, c , as

$$\nu = \frac{c}{\lambda} \quad (2)$$

Molecules usually exist in discrete energy levels (*i.e.* rotational, vibrational and electronic energy levels). This means that a particular energy gap has to be overcome to enable a promotion to a higher energy level. A molecule in state m can get excited to a state n if it absorbs a photon which holds exactly the amount of energy equal to the energy difference between both states. This is known as Bohr's frequency condition and is given as

$$\Delta E = E_n - E_m = h\nu \quad (3)$$

Quantum mechanics provides a theory to describe the energy levels that a molecule can occupy and to predict which transitions can occur between these levels. The state of a system (*e.g.* a molecule) is described by a wavefunction Ψ . The oscillating electric field of radiation (of the correct frequency) can cause an oscillation in the electron cloud of a molecule, initially in state Ψ_m and thereby exciting it to a higher electronic state described by Ψ_n . The transition will preferentially occur if the electric field of the radiation is polarized parallel to the polarization caused in the electron cloud during the process. This dipole induced by light determines the magnitude of the probability of this transition occurring and is called the transition dipole moment, $\vec{\mu}_{nm}$, which is given by:

$$\vec{\mu}_{nm} = \int \Psi_n^* \hat{\mu} \Psi_m d\tau \quad (4)$$

where the integral is taken over all space ($d\tau=dx dy dz$) of the product of the complex conjugate of the wave function of state n (Ψ_n^*) with the electric dipole moment operator, $\hat{\mu}$, and the wave function of the initial state Ψ_m . The electric dipole moment operator is described as

$$\hat{\mu} = \sum_i q_i \vec{r}_i \quad (5)$$

for which q_i and \vec{r}_i are the charge and position vector of the i^{th} electron respectively.

The probability for the transition occurring from state m to n is proportional to the square of the magnitude of the transition dipole moment, also known as the dipole strength, D_{nm}

$$P_{nm} \propto D_{nm} \quad (6)$$

where

$$D_{nm} = |\vec{\mu}_{nm}|^2 \quad (7)$$

The dipole strength is also related to an experimentally determinable parameter, the absorption coefficient $\varepsilon(\nu)$ ($\text{M}^{-1}\text{cm}^{-1}$) (also called extinction coefficient)

$$\int \varepsilon(\nu) d\nu \propto \nu_{nm} D_{nm} \quad (8)$$

where ν_{nm} is the transition frequency and $\epsilon(\nu)$ gives the frequency dependence of the absorption band.

Another measure of the intensity of the transition from states m to n at a frequency ν_{nm} is the oscillator strength f_{nm} , which represents the ratio of the intensity of the absorption to that expected from a three dimensional harmonic oscillator. The oscillator strength can be related to the dipole strength as well as to the integrated absorption coefficient:

$$f_{nm} \propto \nu_{nm} D_{nm} \propto \int \epsilon(\nu) d\nu \quad (9)$$

3.2 Absorption

The absorption of a sample is a measure for the amount of photons it absorbs upon exposing it to radiation of a suitable wavelength. Usually absorption is measured in a spectrophotometer, for which the basic components are a light source, a monochromator and a detector (Figure 7).

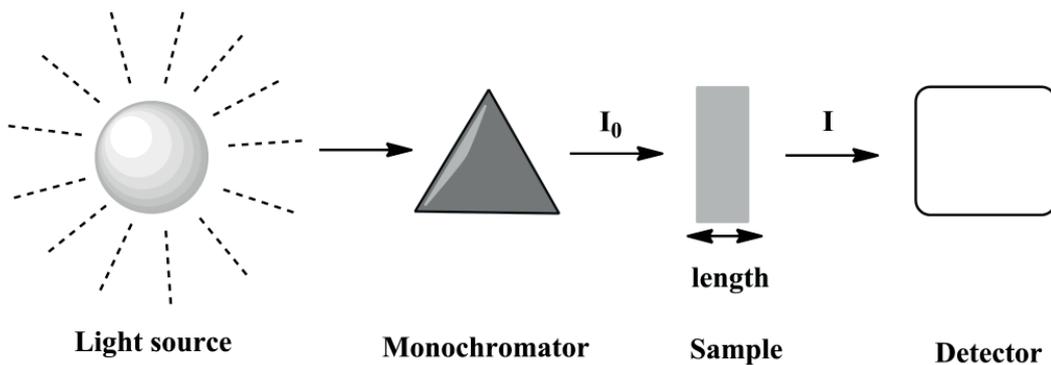


Figure 7 Basic setup of a spectrophotometer. Radiation of the correct wavelength is selected from the light source by a monochromator and led through air (I_0) or through a sample with length l (I) onto a detector.

The sample is placed in the light emitted by a source and its absorbance at a certain wavelength is related to the ratio of the intensities of the incoming (I_0) and transmitted radiation (I). It is also related to the absorption coefficient $\epsilon(\lambda)$, the concentration (c) and the path length (l) of the sample (often in a quartz cuvette). This relation can be described as

$$A(\lambda) = \log\left(\frac{I_0(\lambda)}{I(\lambda)}\right) = \epsilon(\lambda)cl \quad (10)$$

Absorption spectroscopy is often performed to determine the concentration of a sample with a known extinction coefficient $\varepsilon(\lambda)$ but also allows the retrieval of structural or conformational information. For new molecules, the absorption coefficient can be determined and suitable excitation energies for fluorescence measurements can be identified by making a simple absorption measurement.

3.3 Circular dichroism

The peculiar property of chiral molecules (*i.e.* molecules without a reflection plane) to display an unequal absorption of left and right circularly polarized light is known as circular dichroism (CD). A CD signal can also be induced for achiral chromophores upon incorporation into a chiral environment, such as for some DNA intercalators or groove binders. CD can be calculated as

$$CD(\lambda) = A_l(\lambda) - A_r(\lambda) = (\varepsilon_l(\lambda) - \varepsilon_r(\lambda))cl = \Delta\varepsilon(\lambda)cl \quad (11)$$

where $\Delta\varepsilon(\lambda)$ is the molar circular dichroism ($M^{-1}cm^{-1}$) and c and l are the concentration and path length of the sample respectively.

For biomolecules in solution, CD is often used to identify their secondary structure.^[94] The alignment of chromophores in a certain conformation of the molecule may cause shifted or split transitions due to exciton interactions, resulting in a specific CD signature. The CD spectrum of proteins, for example, varies depending on their composition of β -sheets, α -helices and random coils. Similarly, a DNA molecule in the A, B or Z conformation can be distinguished based on characteristic CD features. The UV-absorption band for DNA is mainly initiated by π - π^* transitions of the nucleobases. However, the bases themselves are achiral and thus do not possess a CD signal. The CD spectrum detected for DNA is due to the linkage of the bases with chiral sugars and to their helical stacking. Between 190 and 300 nm, the DNA CD arises mainly from the relative orientation of the bases to each other. For the most common helical form, B-DNA, this results in a CD band which is positive at ~ 275 nm, zero at ~ 258 nm and continuing to negative at ~ 240 nm. Towards the higher energy region, the spectrum continues with a band that increases towards less negative or positive at 220 nm, followed by a small negative peak and finally a large positive peak between 180-190 nm.

A-DNA and double-stranded RNA display a similar spectrum as B-DNA, however, the positive peak at ~ 275 nm is shifted towards 260 nm and increased in intensity. The negative peak between 230-250 nm is less intense and there is an

intense negative band at 210 nm followed by a very intense positive band at 190 nm. The left-winding Z-DNA duplex, which can, for example, be obtained for poly[dG-dC]·poly[dG-dC] at high ionic strength has a negative CD signal at 290 nm, a positive band at 260 nm and a large negative signal around 195-200 nm, while passing through zero between 180-185 nm. Representative spectra of B- and A-form DNA are shown in Figure 8.

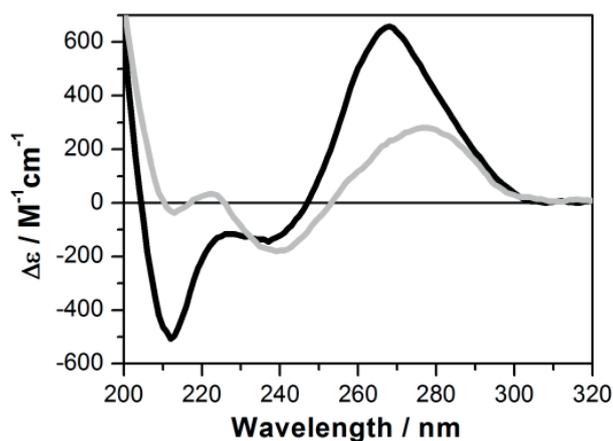


Figure 8 Representative CD spectra for a 1 μM solution of a 38-mer DNA duplex in the B-form (grey, 0.33 mM Na^+ , 0.18 mM phosphate buffer) or A-form (black, 0.33 mM Na^+ , 0.18 mM phosphate buffer, 78 % ethanol). The unit M^{-1} on the vertical axis refers to the duplex concentration. Spectra were smoothed by adjacent averaging over 5 points. The full sequence reads 5'-d(CCATCCCACCACGAGAGAGAGACGTCACCACCCTCC)-3'.

3.4 Fluorescence

In the previous paragraphs absorption was discussed as a process that can render a molecule in an electronically excited state. After that, the molecule has several options for losing its excess energy and returning to the ground state. One of these routes includes fluorescence, which will be discussed in the following paragraphs along with methods to study these properties and their applications.

3.4.1 Fluorescence principles

Fluorescence is the emission of radiation from an excited state to a lower electronic state of the same spin multiplicity (usually singlet states S_1 to S_0). This process along with the other pathways deactivating the excited state can be visualized in a Jablonski diagram (Figure 9).^[16,91,93,95]

An organic molecule in solution is usually found in the lowest vibrational level of the lowest electronic singlet state, S_0 , since the energetic bridge to higher vibrational levels is typically larger than the thermal energy ($E=k_B T$, where k_B is the Boltzmann constant and T is the temperature of the solution). This molecule is then able to absorb a photon (of the correct energy) (Abs) to get excited to a vibrationally excited level of a higher electronic singlet state (S_n with $n \geq 1$) (Figure 9). Subsequently, the vibrational excitation is quickly reduced due to collisions with solvent molecules (*i.e.* vibrational relaxation, VR), rendering the molecule at the lowest vibrational level of S_n . If this level is for example S_2 , the molecule can undergo internal conversion (IC) to transform to an isoenergetic excited vibronic state of S_1 . Hereafter, the molecule can once again rapidly relax by vibrational relaxation to the vibrationless ground level of S_1 (Figure 9).

So far, the processes discussed for the molecule to lose its excess energy are too fast to allow competition from radiative processes, such as fluorescence. However, internal conversion between S_1 and S_0 usually occurs on a longer timescale since the energy gap between these levels is larger; giving radiative processes a chance to take place. The competing processes from the vibrationless state of S_1 are internal conversion to S_0 (IC, k_{IC}), fluorescence (Fluo, k_f) or intersystem crossing to a triplet state T_1 (ISC, k_{ISC}), where k_n is the corresponding rate constant for process n (Figure 9). Apart from the transformation of S_1 to a vibrationally excited isoenergetic level of S_0 by internal conversion the molecule can also emit a photon as fluorescence, thereby arriving at a vibrational level of S_0 . The third option mentioned is intersystem crossing, in which the molecule in the singlet excited state changes the spin of an electron, thereby altering spin multiplicity, arriving at a vibrationally excited triplet state, T_1 in Figure 9. Again the molecule can undergo rapid vibrational relaxation to the ground vibrational level of T_1 , after which a photon can be emitted as phosphorescence. Alternatively the molecule can transform back to isoenergetic vibrational level of S_0 through intersystem crossing. In each of the above mentioned cases, the molecule will quickly lose its excess vibrational energy left in S_0 by vibrational relaxation.

A common measure for the fluorescence of a molecule is its fluorescence quantum yield, defined as the ratio of the number of photons emitted per amount of photons absorbed. This can also be understood as the probability of the excited molecule in S_1 to return to the ground state by emitting a photon amongst all possible processes. Using the rate constants that were introduced above, this can be written as

$$\Phi_f = \frac{k_f}{k_f + k_{IC} + k_{ISC}} \quad (12)$$

Frequently, fluorophores are also compared in terms of their brightness, which is the product of their fluorescence quantum yield and absorption coefficient (usually at the absorption maximum) ($\Phi_f \times \epsilon(\lambda)$).

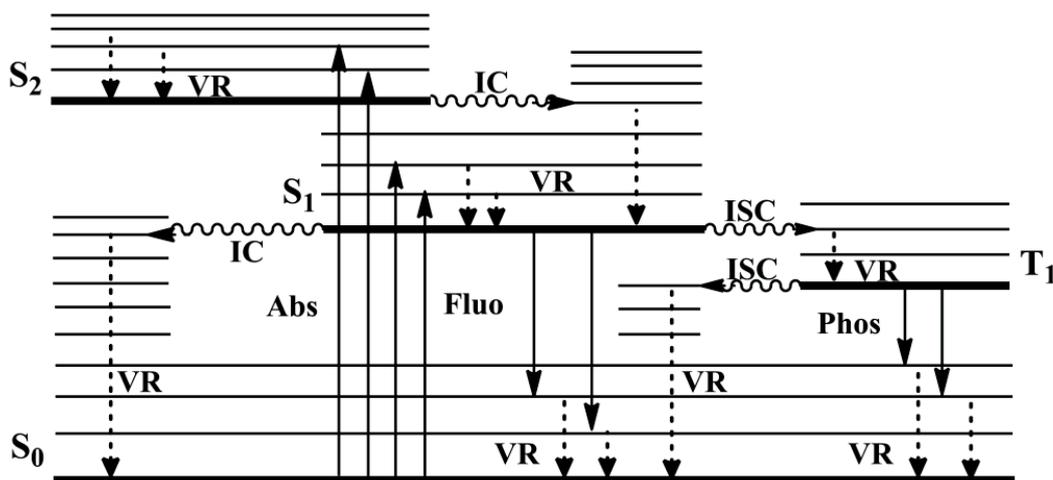


Figure 9 Jablonski diagram giving a schematic overview of common radiative processes a molecule can undergo (solid arrows) such as absorption (Abs), fluorescence (Fluo) and Phosphorescence (Phos). Also the non-radiative pathways are depicted by dashed arrows (vibrational relaxation, VR) or squiggly arrows (Internal conversion, IC and Intersystem crossing, ISC).

Another important parameter often used to describe the photophysical behavior of the molecule is its fluorescence lifetime, which is the average time the molecule spends in the excited state following excitation. This can be described as

$$\tau_f = \frac{1}{k_f + k_{IC} + k_{ISC}} \quad (13)$$

The fluorescence properties of the molecule can also be related to its absorbance by means of the radiative rate constant, k_f , and the absorption coefficient, $\epsilon(\tilde{\nu})$ and can be expressed with the Strickler-Berg equation^[96]:

$$k_f = 2.880 \times 10^{-9} n^2 \frac{g_m}{g_n} \langle \tilde{\nu}_f^{-3} \rangle^{-1} \int \epsilon(\tilde{\nu}) d \ln \tilde{\nu} \quad (14)$$

where g_m and g_n are the degeneracies of the ground (m) and upper (n) electronic state, n is the refractive index of the medium and $\tilde{\nu}_f$ is the wavenumber of emission. This equation is applicable for strong, broad-banded transitions in molecules.

3.4.2 Steady state fluorescence

A common practice to study the photophysical properties of a chromophore in bulk solution is to record an emission or excitation spectrum. This is done using a spectrofluorimeter, for which the basic setup is shown in Figure 10. In the case that an emission spectrum is to be obtained, a specific excitation wavelength will be selected from the light source by the excitation monochromator which is then focused onto the sample. Photons emitted by the sample are usually collected at right angle to the incident radiation in order to reduce the interference of unabsorbed excitation light. The emission monochromator is scanned through a selected wavelength region, thereby yielding a spectrum of the emission intensity per wavelength that reaches the detector. An excitation spectrum is measured in the same manner; however, this time the emission monochromator remains parked while scanning through the excitation energies.

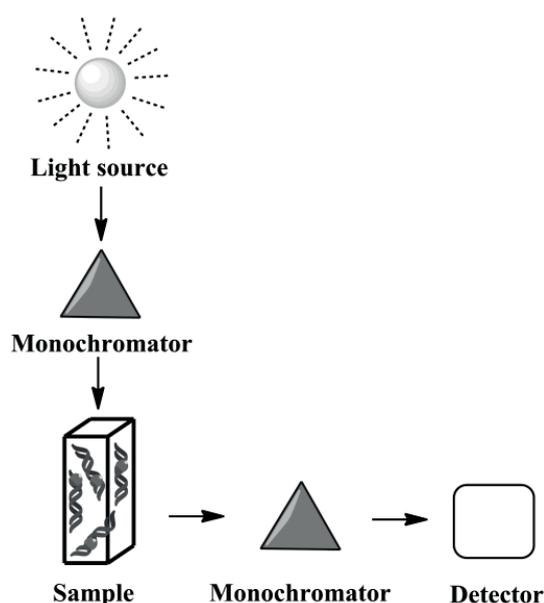


Figure 10 Setup of a basic spectrofluorimeter with a source of UV/visible radiation (for example a xenon arc lamp) that is led through a monochromator selecting the correct radiation energy to excite the sample. The emission monochromator then selects the correct wavelength of radiation to be observed by the detector (usually at right angle of the excitation light).

As mentioned in the previous paragraph, fluorescence quantum yields are frequently determined for molecules as a measure for their emissive capacity and carries information about the rate constants depleting the electronically excited S_1 state (Figure 9, equation 12). A common method to determine this value experimentally is by using a reference substance with a known quantum yield, $\Phi_{f,R}$. This can be done in practice using the following relationship:

$$\Phi_f = \Phi_{f,R} \frac{I}{I_R} \frac{A_R}{A} \frac{n^2}{n_R^2} \quad (15)$$

where I is the integrated emission intensity, A is the absorbance at the excitation wavelength and n is the refractive index of the solvent in which the sample/reference is contained. R denotes the parameters corresponding to the reference compound.

3.4.3 Steady state fluorescence anisotropy

As was described in section 3.1, the transition dipole moment for a transition in a molecule from a ground electronic state m to an upper state n , exhibits a magnitude and an orientation. While the main interest in the latter paragraph was the magnitude, related to the probability of the transition occurring, the focus here will be on the orientation of $\vec{\mu}_{nm}$. A molecule will preferentially absorb light that has its electric field oscillating in the same direction as the transition dipole moment. This implies that irradiating a sample with linear polarized light will favour exciting those molecules with their transition moments aligned with the polarization direction of the radiation. As fluorescence emission usually occurs from the S_1 to S_0 electronic energy level, this transition is polarized parallel to the S_0 - S_1 absorption transition. Consequently, the emission from a molecule being excited with polarized excitation light in the lowest energy absorption band (S_0 - S_1) will usually be co-linear, if it were not that a molecule in solution can reorient during the time of excitation. In this way, the emission will be depolarized depending on the time the molecule spends in the excited state as well as on its size, shape and environment. To extract information about the molecule in this way, the anisotropy, r , of a sample can be determined as follows:

$$r = \frac{I_{VV} - G I_{VH}}{I_{VV} + 2G I_{VH}} \quad (16)$$

for which $G = \frac{I_{HV}}{I_{HH}}$ is a correction factor and I_{XY} is the intensity of polarized emission upon excitation with polarized radiation. X and Y denote the polarization directions of the excitation and emission light, respectively, *i.e.* horizontal (H) or vertical (V).

If the molecule is immobilized on the time scale of fluorescence, the anisotropy will contain information about the angle (α_i) between the absorption and emission moment of the i^{th} transition. This can for example be achieved by

freezing the sample solution to a glass state. The anisotropy measured in this special case is called fundamental anisotropy ($r_{0,i}$) and is related to α_i by

$$r_{0,i} = \frac{1}{5}(3\cos^2\alpha_i - 1) \quad (17)$$

r_0 can adopt values between -0.2 and 0.4, corresponding to an angle of 90° and 0° , respectively, between the absorption and emission transition dipole moment. This means that for most molecules a fundamental anisotropy value close to 0.4 is expected upon excitation in the absorption region with a pure S_0 - S_1 transition ($\alpha_i = 0^\circ$).

3.4.4 Time-resolved fluorescence

The lifetime of a fluorophore can be described as a function of the rate constants depopulating the excited state, as was shown in equation 13. In practice we can retrieve this parameter by studying the kinetics of the photophysical process. If a population of molecules was excited with an infinitely sharp radiation pulse, the initially excited population will decay over time to the ground state. Since emission is a random process, it can be described by first order kinetics as follows:

$$I(t) = I_0 \exp(-t/\tau_f) \quad (18)$$

where the excited state population of molecules remaining at time t can be monitored by the fluorescence intensity $I(t)$ (which is proportional to it) and I_0 is the initial fluorescence intensity at time 0.

Often multi-exponential expressions are observed for chromophores in biological systems due to, for example, heterogeneity within the sample. In that case the fluorescence decay can be described as

$$I(t) = \sum_{i=1}^j \alpha_i \exp(-t/\tau_i) \quad (19)$$

for which α_i is the amplitude of the i^{th} lifetime (usually $j \leq 3$ to allow a reliable fit) and $\sum \alpha_i$ is normalized to unity. In that case an average lifetime is commonly used to retrieve desired information. This can be described as the intensity-averaged mean lifetime:

$$\langle \tau \rangle = \frac{\sum_{i=1}^j \alpha_i \tau_i^2}{\sum_{i=1}^j \alpha_i \tau_i} \quad (20)$$

or as the amplitude-weighted one:

$$\langle \tau \rangle = \frac{\sum_{i=1}^j \alpha_i \tau_i}{\sum_{i=1}^j \alpha_i} \quad (21)$$

The choice of which average lifetime to use depends on the phenomenon being studied. For Förster resonance energy transfer experiments the amplitude-weighted average (equation 21) is recommended.^[16,95,97]

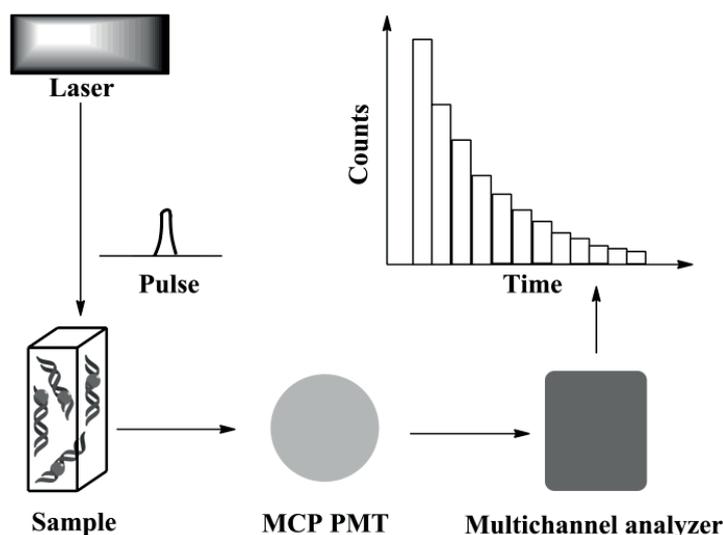


Figure 11 Schematic setup of a TCSPC measurement. A pulsed laser source excites the sample, after which the emitted photons are recorded by a microchannel plate photomultiplier tube (MCP PMT). The recorded data are fed into a multichannel analyzer, resulting in a histogram.

In this thesis fluorescence lifetimes were determined with the technique time-correlated single photon counting (TCSPC). The principle is based on the proportionality between the probability of detecting a photon at a certain time after excitation and the fluorescence intensity at this same time. A pulsed light source (laser) is used to repeatedly excite the sample and the time between excitation and the first photon reaching the detector is recorded (Figure 11). The time window is divided into smaller intervals, so-called channels, and every counted photon is recorded in the channel with the corresponding time interval. Measurements are repeated until enough photons (often 10000 for statistical reasons) are counted in the top channel. This finally yields a histogram with the amount of photons counted per time interval (Figure 11), representing the fluorescence decay as a function of time. To obtain a statistically correct measurement and to avoid two photons originating from the same excitation pulse reaching the detector, the setup is such that only every 100th pulse yields a photon arriving at the detector.

The lifetime can be retrieved by fitting the decay stored in the histogram ($R(t)$) to an exponential function. Therefore, also the instrument response function (IRF, $E(t)$) is recorded, which represents the way the laser pulse is seen by the system. The IRF can then be convoluted with a theoretical decay ($I(t-t')$) (in the form of equation 19) to best fit the measured decay as follows:

$$R(t) = E(t) \otimes I(t) = \int_0^t E(t')I(t - t')dt' \quad (22)$$

where emission, *i.e.* $I(t)$ starts at $t=t'$.

3.4.5 Förster resonance energy transfer (FRET)

Above (in section 3.4.1), the processes discussed for depopulating the excited S_1 state, concerned only ensembles of isolated molecules. However, molecules can influence each other in different ways when coming in proximity. In case two molecules are in each other's vicinity (without overlap of the molecular orbitals) and one is in an excited state, it can transfer its excitation energy to the second molecule provided that this one has an absorption transition of matching energy, known as the resonance condition. Not surprisingly, the molecules are usually referred to as the energy donor and acceptor. If they are not too close by, their interaction will mainly be coulombic in nature and can be described theoretically considering point dipoles at the centers of the molecule. As a rule of thumb the inter-chromophore distance is suggested to be at least four to five times the size of the molecule. The phenomenon is referred to as resonance energy transfer (RET) or Förster Resonance Energy Transfer (FRET) after Theodore Förster, who first described the relationship between the rate constant for energy transfer and spectroscopic observable properties^[98,99]. Often the term Fluorescence Resonance Energy Transfer (FRET) is also used, even though this name is somewhat misleading since no emission of photons takes place from the donor molecule in order for the energy to be relocated to the acceptor molecule.

According to Förster's theory,^[98,99] the resonance energy transfer rate, k_{ET} , is given as follows

$$k_{ET} = \frac{1}{\tau_{f,D}^0} \left[\frac{R_0}{r} \right]^6 \quad (23)$$

where r is the distance between the donor and acceptor molecules, $\tau_{f,D}^0$ is the lifetime of the donor in the absence of the acceptor and R_0 is the Förster radius, given as

$$R_0 = \left[\frac{9000(\ln 10)\phi_{f,D}^0 \kappa^2 J}{128N_A \pi^5 n^4} \right]^{1/6} \quad (24)$$

for which $\phi_{f,D}^0$ is the fluorescence quantum yield of the donor without the acceptor present, N_A is Avogadro's number, n is the refractive index of the medium, κ^2 is the orientation factor (see below) and J is the overlap integral between the donor emission $F_D(\lambda)$ (with the area under the curve normalized to unity) and the acceptor absorption coefficient spectrum $\varepsilon_A(\lambda)$, defined as

$$J = \int_0^\infty F_D(\lambda) \varepsilon_A(\lambda) \lambda^4 d\lambda \quad (25)$$

with

$$F_D(\lambda) = \frac{I_D(\lambda)}{\int_0^\infty I_D(\lambda) d\lambda} \quad (26)$$

where $I_D(\lambda)$ is the recorded donor emission spectrum and $\int_0^\infty F_D(\lambda) d\lambda = 1$.

The orientation factor (κ^2) mentioned above relates the relative orientation of the transition dipole moments of the donor and acceptor molecules to the rate of energy transfer. κ^2 is given as

$$\kappa^2 = (\cos\theta - 3\cos\alpha\cos\beta)^2 \quad (27)$$

where θ defines the angle between both transition moments, whereas α and β define the angle of the transition moment of the donor and acceptor, respectively, with the vector connecting their centers. κ^2 can adopt values between 0 for transition moments at right angle and 4 for a collinear arrangement. The orientation factor often introduces an uncertainty to R_0 and hence to the determination of r based on the recorded FRET efficiency (*vide infra*) since the exact orientation of chromophores in complex systems is usually unknown. Many dyes are externally attached to the system of interest by flexible linkers. In this situation often a κ^2 -value of 2/3 is assumed, which is valid in case of isotropic and dynamic averaging of the orientations of the transition dipoles.^[16,100]

The rate constant for energy transfer can now be used to obtain the efficiency of energy transfer (ϕ_{ET} or E_{FRET}) analogous to the fluorescence quantum efficiency of a molecule (equation 12), by including the rate constants of all the processes depleting the S_1 state:

$$\phi_{ET} = E_{FRET} = \frac{k_{ET}}{k_f + k_{IC} + k_{ISC} + k_{ET}} = \frac{R_0^6}{R_0^6 + r^6} \quad (28)$$

Now it also becomes apparent that the Förster radius, R_0 represents the distance between the donor and acceptor molecule corresponding to a FRET efficiency of 50%. Since the FRET efficiency is dependent on the sixth power of the distance between the donor and acceptor molecule it displays a sharp responsiveness to distance changes in the range around R_0 . In this way FRET is often used as a molecular ruler to obtain information about the relative separation between certain parts of a macromolecule.^[101,102]

In practice the FRET efficiency of a system can be determined by recording the fluorescence quantum yield or lifetime of the donor with $(\phi_{f,D}, <\tau_{f,D}>)$ and without the acceptor present $(\phi_{f,D}^0, <\tau_{f,D}^0>)$ (as can be derived from equation 28).

$$E_{FRET} = 1 - \frac{\phi_{f,D}}{\phi_{f,D}^0} = 1 - \frac{<\tau_{f,D}>}{<\tau_{f,D}^0>} \quad (29)$$

3.4.6 The tricyclic cytosine FRET-pair

The thought of retrieving structural information about DNA using FRET is a tempting one, considering the well-ordered geometry of this molecule.^[103] Often externally tethered dyes attached by a linker are applied in quantitative FRET measurements. This enables almost any commercially bright chromophore to be implemented, allowing studies down to single-molecule level.^[25,104,105] At the same time the precise determination of the donor-acceptor distance (r) and orientation factor κ^2 is often complicated due to dipole diffusion and reorientation.^[53,56,57] Elegant studies have been reported, trying to achieve a better control of dye position, resulting in FRET efficiency profiles responsive to the orientation and distance change between the donor-acceptor pairs upon increasing base-pair separation.^[25,55,106,107] A first study involved a perylene acceptor attached at the end of a stilbene-linked hairpin,^[106] whereas other reports involved end-stacked Cy3-Cy5 dyes attached by linkers^[25,55] or base-surrogates consisting of a more immobile pyrene-perylene FRET-pair^[107].

In recent years our group has characterized and implemented a nucleobase FRET-pair, consisting of the two tricyclic cytosine analogues, tC^O and tC_{nitro} (Figure 12a).^[40] The direct incorporation of these FBAs inside the DNA duplex allows an almost precise control of their relative orientation on the time-scale of

energy transfer. The emission of the fluorescent donor, tC^O , centered around 450 nm (in DNA),^[39] overlaps the lowest energy absorption band (approximately 440 nm) of the non-emissive acceptor tC_{nitro} ^[42] (Figure 12b). As for the earlier characterized compound tC ,^[38] the emission of tC^O is relatively insensitive to its microenvironment,^[39] in contrast to other FBAs. This results in an estimated Förster radius of 27.2 Å when assuming a κ^2 -value of 2/3, refractive index of 1.40 and an average value of the fluorescence quantum yield of tC^O in dsDNA of 22%.^[42] Consequently, this FRET-pair is more suited for studying short-range structural changes compared to commercial dyes such as Cy3-Cy5 ($R_0 \sim 60$ Å, $\kappa^2 = 2/3$)^[25]. As is to be expected for rigidly incorporated probes, increasing the base-pair separation between both FBAs results in a FRET profile that closely corresponds to that predicted for B-DNA (Figure 13).^[40]

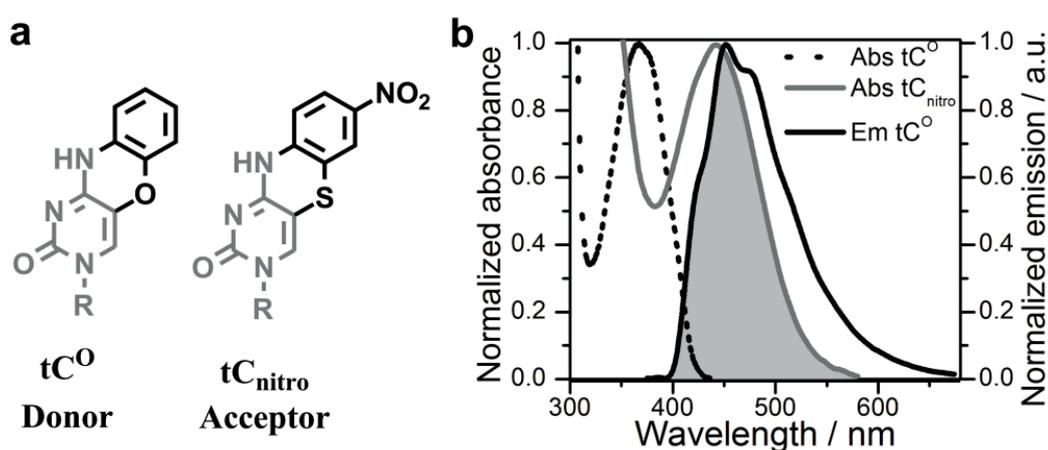


Figure 12 (a) tC^O - tC_{nitro} FRET-pair. The cytosine skeleton is grey. R is deoxyribose. (b) Absorption/emission spectrum of tC^O and tC_{nitro} inside duplex DNA with the overlap between tC^O emission and tC_{nitro} absorption highlighted in grey.

The main advantage of these stably incorporated probes is that they cause only minor perturbations to the DNA duplex,^[39,40] thereby allowing incorporation close to or inside the site of interest. The FRET-pair can therefore be applied to retrieve detailed structural information. In a recent study Preus *et al.* report on a new methodology for the simulation and analysis of FRET in nucleic acids, featuring a MATLAB based software, FRETmatrix.^[108] The project implemented the FBA FRET-pair to yield a 3D structure of kinked DNA. In this thesis, however, the FRET-pair is rather implemented in yes/no types of questions, *i.e.* to report on subtle local structural alterations in a specific DNA sequence upon interactions with proteins.

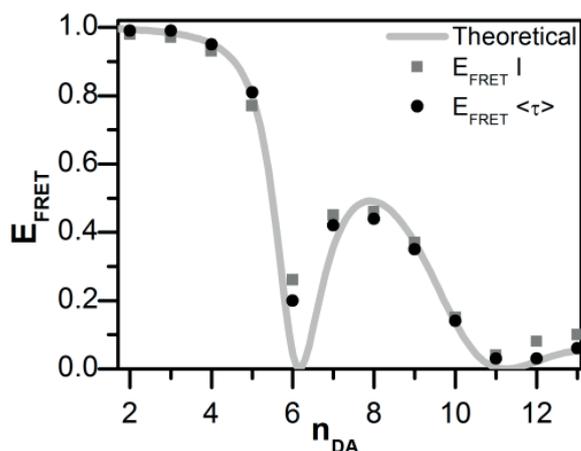


Figure 13 FRET efficiency (E_{FRET}) profiles for tC^O - tC_{nitro} FRET-pair as a function of the number of base-pairs separating them (n_{DA}). Efficiencies were determined with time-resolved ($E_{FRET} \langle \tau \rangle$) and steady state fluorescence ($E_{FRET} I$). Also a fitted profile based on B-DNA parameters is shown (Theoretical). (Data from our group which were published in ^[40])

3.5 DNA melting

As was mentioned in 2.1, the two polynucleotide strands creating a DNA double helix are held together by non-covalent forces. However, upon heating the duplex, enough thermal energy will be provided at some point for both strands to part, referred to as DNA melting. The temperature at which half of the duplexes are denatured is called the melting temperature, which is dependent on the length, sequence and concentration of the DNA as well as on its surroundings such as the salt concentration and pH.

Melting temperatures are often reported as a measure of the stability of DNA duplexes. They can be obtained by recording a UV/Vis melting curve, *i.e.* by monitoring the change in absorbance at 260 nm as a function of the slowly rising temperature. As the double helices melt, the absorbance will increase by approximately 10-20 %. The decreased absorbance of the nucleobases inside duplex DNA, called hypochromicity, can be explained by the close, more ordered stacking of the bases in the duplex compared to single-stranded DNA. In this sense the transition dipole moment of a stacked base will be shorter due to the influence of the induced dipoles in the neighbouring bases.^[109,110]

The melting temperature can be determined as the temperature at half maximum of the curve or the temperature corresponding to the maximum of the first derivative. In this thesis, the melting temperature is used mainly as a measure

of the (de)stabilization of the DNA caused by the incorporation of a modified nucleobase.

Another way in which the melting process can be monitored is by using a fluorophore whose emission is responsive to the duplex denaturation. As was mentioned in section 2.2.2, the fluorescence of most FBAs is quenched in double-compared to single-stranded DNA. In this way the rise in fluorescence of an FBA can be monitored as the DNA duplex separates in response to the increasing temperature. As for the absorption method, the melting temperature can be determined as the maximum of the first derivative. The half maximum of the signal can also be used, given that appropriate corrections are made for the inherent temperature responsive change in fluorescence of the FBA. An advantage of fluorescence melting, in contrast to UV melting is that it can also report on the local stability of the DNA.

4 Results

In this chapter the main results contained in this work will be discussed. For a complete overview of results and experimental procedures, the reader is referred to the appended papers (I-VI).

4.1 Characterization of fluorescent nucleobase analogues

The characterization of photophysical and base-mimicking properties of new FBAs is essential for identifying applications they are suited for. Furthermore, this can yield new insights into the correlation between their structure and spectroscopic properties for future design of novel FBAs. The evaluation of different aspects of a series of adenine analogues (papers I-III) and the cytosine analogue tC (paper IV) are presented here.

4.1.1 Characterization of fluorescent adenine analogues

To increase the diversity of FBAs with improved structural and spectroscopic properties, two fluorescent adenine analogues, triazole adenine (A^T) (paper I, Figure 14) and quadracyclic adenine (qA) (paper III, Figure 14) were synthesized. Their photophysical properties, both as nucleosides and inside DNA, are discussed here.

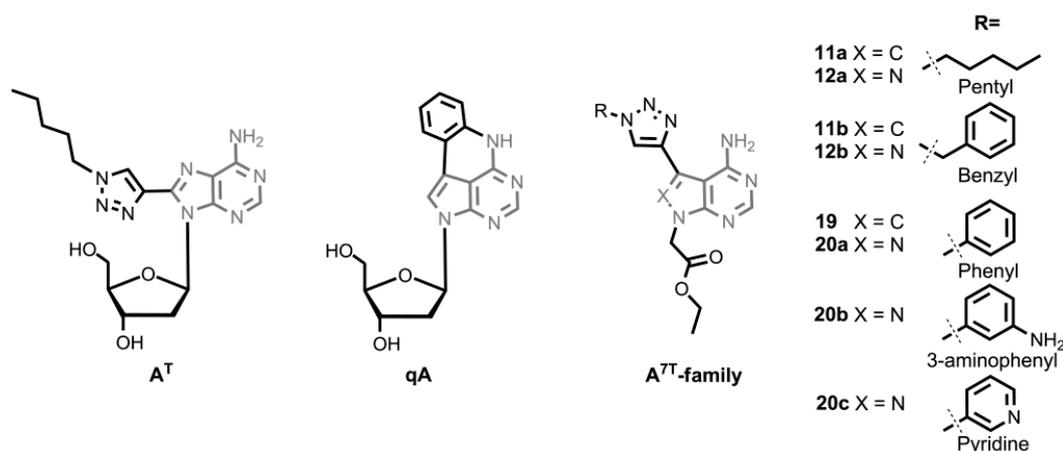


Figure 14 Nucleosides of A^T , qA and the ethyl esters of the new A^{TT} -family.

As a follow-up study, trying to overcome the DNA duplex destabilizing effect of A^T (*vide infra*), a new family of 7-modified triazole adenines (A^{7T} -family, Figure 14) was developed, of which the photophysical characterization of the monomeric form is also presented (paper II).

Spectroscopic characterization of the monomers

Before FBAs are incorporated into DNA, it is important to understand their photophysical properties as monomers. The spectral features of the adenine analogues qA, A^T and the A^{7T} -family (Figure 14) are presented in Figure 15 and Table 1. The absorption of all compounds is red-shifted compared to adenine (centered around 260 nm), allowing for selective excitation due to an absorption tail extending further than that of the natural nucleobases. Quadracyclic adenine (qA) has a well resolved lowest energy absorption band, located at 335 nm (in water), whereas triazole adenine (A^T) has its lowest energy absorption centered around 282 nm in water and therefore not fully separated from the natural bases. The A^{7T} -family (280-293 nm in methanol) absorbs in a similar region as A^T (286 nm in methanol) as well as a series of previously characterized C8-modified triazole adenines (289-296 nm in THF).^[37] A^T shows similar absorption envelopes in water and methanol, which is also observed for qA (Figure 15a). The A^{7T} -family was studied in methanol due to a poor solubility in water (**20b** and **20c** were also studied in ACN or DCM, *vide infra*). Within this group of analogues, a more structured absorption can be observed for the compounds containing a nitrogen (Figure 15e) instead of carbon (Figure 15c) on the 8-position.

The shape and position of the emission spectra of all triazole adenines (A^T and the A^{7T} -family) are very similar in methanol, with exception of compound **20b** (Figure 15d and f). The latter shows an extensive Stokes shift of $\sim 12800\text{ cm}^{-1}$ ($\Delta\lambda=175\text{ nm}$) compared to the other compounds in its series. This is due to a more red-shifted emission maximum (468 nm) in methanol, which may be caused by intramolecular charge transfer (ICT) in the excited state from the electron donating amino group. The shift of its emission maximum to shorter wavelengths in dichloromethane (DCM, apolar and aprotic) ($\lambda_{em,max}=393\text{ nm}$) and in acetonitrile (ACN, polar and aprotic) ($\lambda_{em,max}=428\text{ nm}$) are in accordance with a possible excited ICT state, which may be stabilized by polar protic solvents (Table 1, Figure 15f).^[111,112] As a comparison, the related conjugated analogue **20c** was also investigated in DCM and showed no changes in emission maximum. Furthermore, no significant changes could be detected in the absorption spectrum of both **20b** and **20c** in DCM, methanol or ACN (only for **20b**).

qA shows a lower emission energy than the triazole adenines (except for **20b** in methanol) both in water (456 nm) and methanol (440 nm) with a more structured spectral envelope detected in methanol (Figure 15b).

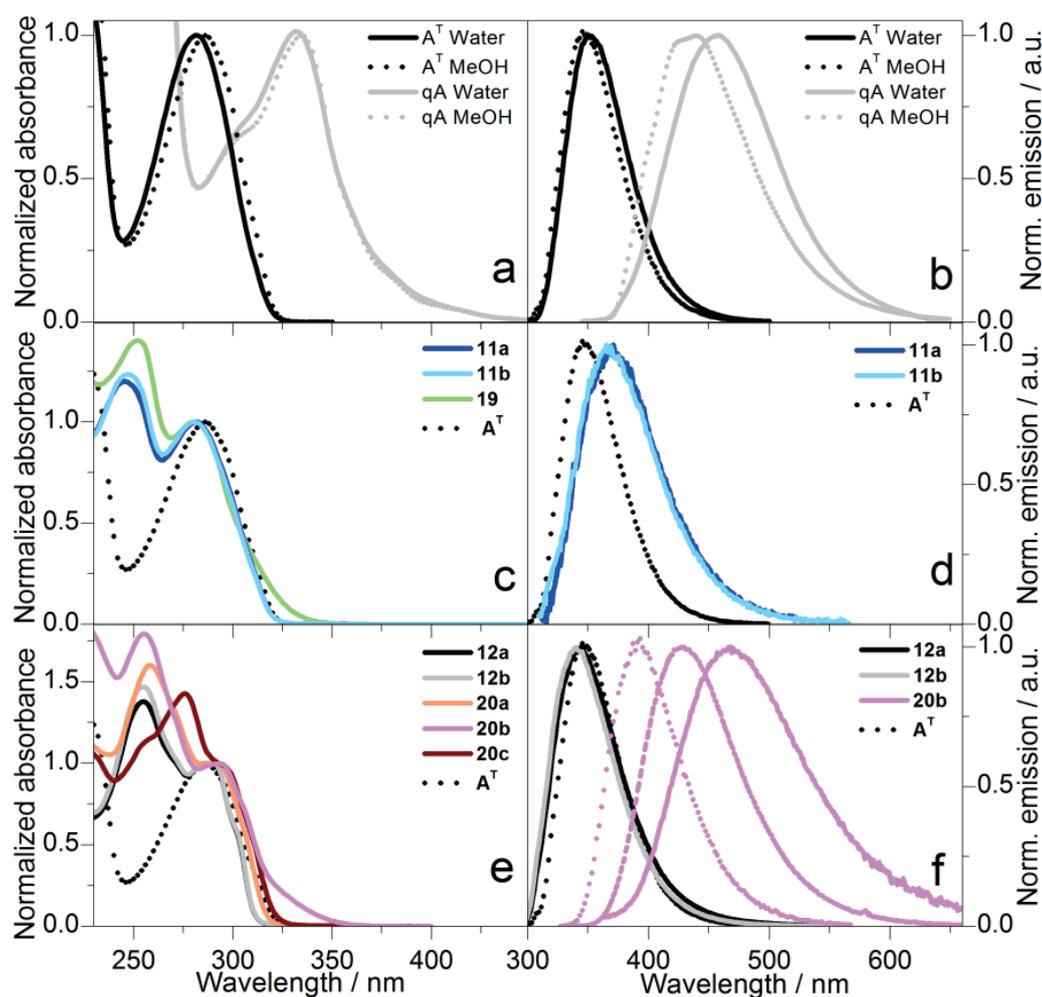


Figure 15 Normalized absorption (a) and emission spectra (b) of A^T and qA in water and methanol (MeOH). Also absorption (c, e) and emission (for $\Phi_f > 0.2\%$) (d, f) of members of the A^{7T} -family with (c, d) and without nitrogen (d, f) on position 8 of adenine. Also A^T in MeOH is included for comparison as well as emission of **20b** (f) in DCM (pink, dotted) and ACN (pink, dashed).

A^T is highly fluorescent both in water ($\Phi_f = 61\%$) and in methanol ($\Phi_f = 49\%$) (Table 1). Its quantum yield is similar to 2-aminopurine (2-AP) (68%)^[32] as well as the previously reported 8-modified adenine analogue, 8-vinyl-deoxyadenosine (8vdA) (66%)^[69] in aqueous solution and exceeds the quantum yields of the pteridine adenine analogues 6MAP (39%) and DMAP (48%)^[61]. The A^{7T} -family, on the other hand, is significantly less emissive than A^T in methanol with compound **12a**, a sister compound of A^T , exhibiting the highest quantum yield ($\Phi_f = 5.2\%$) of the series (Table 1). The second sister compound, **11a**, lacking the nitrogen on the 8-position, has a much lower quantum yield ($\Phi_f = 0.75\%$). Within each group of the A^{7T} -compounds, based on a nitrogen or carbon being incorporated on the 8-position, all other compounds show lower quantum yields

than the two sister compounds of A^T in methanol. As was also observed for a previous class of C8-triazole adenines, the compounds with extended aromatic systems (**19**, **20a**, **20b** and **20c**) are the least fluorescent within their groups.^[37] These compounds also show absorption spectra that extend to longer wavelengths and have the highest extinction coefficients, which could be explained by their more extended conjugated system.

Table 1 Fluorescence quantum yield (Φ_f), lowest energy absorption maximum ($\lambda_{A,max}$), emission maximum ($\lambda_{Em,max}$) and corresponding extinction coefficient (ϵ_{max}) at $\lambda_{A,max}$ of each adenine analogue (Cpd) dissolved in water, methanol (MeOH), acetonitrile (ACN) or dichloromethane (DCM). Also the brightness ($\propto \Phi_f \times \epsilon_{max}$) is shown.

Cpd	Φ_f (%)	$\lambda_{A,max}$ (nm)	$\lambda_{Em,max}$ (nm)	ϵ_{max} (M ⁻¹ cm ⁻¹)	Solvent	$\Phi_f \times \epsilon_{max}$ (M ⁻¹ cm ⁻¹)
A ^T	61	282	353	16500 ^c	Water	10065
	49	286	346	16500	MeOH	8085
11a	0.75	282	369	11200	MeOH	84
11b	0.39	281	366	11800	MeOH	46
12a	5.2	290	343	10700	MeOH	556
12b	4.4	290	340	10800	MeOH	475
19	<0.1	280	- ^b	14500	MeOH	<15
20a	<0.2	288	345	13600	MeOH	<27
20b	0.62	293	468	14000	MeOH	87
	24	295	428	14000 ^c	ACN	3360
	21	297	393	14000 ^c	DCM	2800
20c	<0.1	293 ^a	359	15300 ^c	MeOH	<15
	<0.1	295 ^a	359	15300 ^c	DCM	<15
qA	6.8	335	456	5000	Water	340
	10.6	335	440	6300	MeOH	668

^aAbsorption maximum ($\lambda_{A,max}$) estimated for lowest energy absorption shoulder

^bEmission maximum omitted due to the weak fluorescence. ^cExtinction coefficient assumed to be the same as in methanol.

Interestingly, compound **20b** shows a dramatic increase in quantum yield in ACN ($\Phi_f \sim 24\%$) and DCM ($\Phi_f \sim 21\%$) compared to in methanol ($\Phi_f = 0.62\%$), indicating efficient quenching by protic solvents (Table 1). This could be due to a more efficient internal conversion in case of hydrogen bonding with the solvent.^[111] As a comparison, the quantum yield of **20c** in DCM remained unaltered compared to in methanol (Table 1). A similar quenching behavior was also observed for the proposed twisted intramolecular charge transfer-state of N⁶,N^{6'}-dimethyladenosine.^[113]

Finally, qA shows a lower emission than A^T, with a moderate quantum yield of 6.8 % in water. The extinction coefficient corresponding to its lowest energy

absorption maximum ($5000 \text{ M}^{-1}\text{cm}^{-1}$) is reduced approximately three-fold in comparison with A^{T} ($16500 \text{ M}^{-1}\text{cm}^{-1}$).

In general, A^{T} is by far the brightest ($\propto \Phi_f \times \varepsilon$) of all ten adenine analogues discussed in this work, both in water and methanol (Table 1). Its brightness in water ($\sim 10000 \text{ M}^{-1}\text{cm}^{-1}$) is also more than double as high as for 2-AP ($\sim 4000 \text{ M}^{-1}\text{cm}^{-1}$)^[32,114].

Characterization of the lowest energy absorption band of A^{T} and qA

To get more insight into the different transitions contained in the lowest energy absorption band of A^{T} and qA, their fundamental anisotropy was recorded in a vitrified matrix (Figure 16). This information is important in case the analogues are to be applied in anisotropy- or FRET-measurements (*vide infra*, anisotropy of A^{T} inside DNA in high viscosity sucrose solutions). Furthermore, this allows for comparison of the energy of transitions with predictions from quantum chemical calculations.

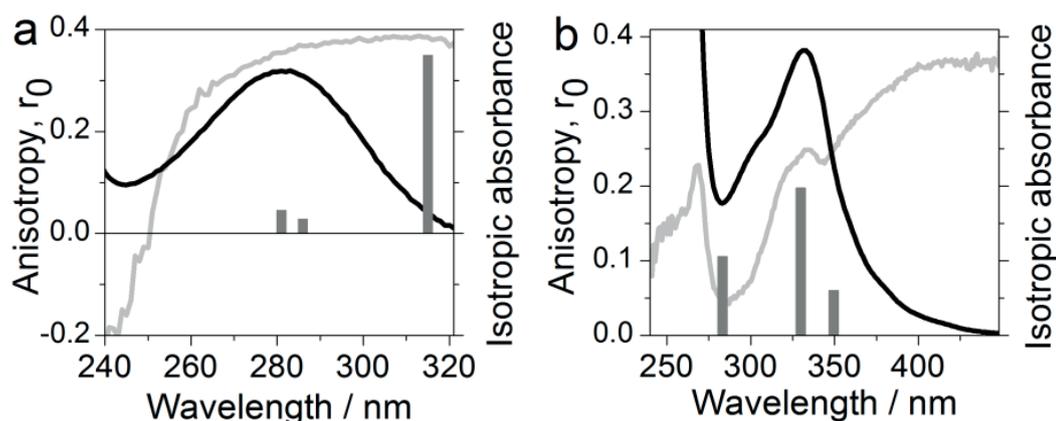


Figure 16 Excitation anisotropy spectrum of A^{T} (a) and qA (b) in a vitrified H_2O /ethylene glycol matrix (1:2 mixture) at -100°C (grey line). Also the isotropic absorbance is shown (black) as well as the relative intensities predicted for the 3 lowest energy transitions (grey bars) for A^{T} (methyl instead of pentyl substituent on triazole ring, left to right $f=0.094$; 0.059 and 0.72)^[37] and qA (left to right $f=0.1049$; 0.1968 and 0.0594 ; TDDFT calculations).

For A^{T} , a virtually constant value of $r_0=0.38$ was recorded between 280 and 320 nm, indicating a single electronic transition dominating this region (Figure 16a). A strong and isolated lowest energy transition ($f=0.72$) was predicted at 315 nm for a related methyl substituted C8-triazole adenine (calculations from a previous study)^[37]. This prediction corresponds reasonably well to our findings for A^{T} , even though the transition is slightly shifted. Furthermore, the value of r_0 is close to the theoretical maximum (0.4) for these excitation energies, suggesting virtually parallel absorption and emission transition moments.

The fundamental anisotropy spectrum for qA shows three main plateaus between 280 and 450 nm (Figure 16b). The first one, associated with the lowest energy absorption transition is found between 390 and 450 nm with $r_{\theta}=0.36$. As for A^T , this value is close to the theoretical maximum, suggesting a single electronic transition in this region, polarized parallel to the emission transition moment. A second plateau is located around the absorption maximum (~ 330 nm, $r_{\theta}=0.25$) and a third one around 300 nm ($r_{\theta}=0.05$), where also a shoulder can be observed in the absorption spectrum. These findings suggest three absorption transitions between 280 and 450 nm, with the lowest energy transition around 400 nm (S_0 - S_1). The positions as well as relative intensities of these three transitions seem to correspond well to theoretical predictions, albeit somewhat shifted (Figure 16b). Based on the second plateau (~ 330 nm, $r_{\theta}=0.25$), the angle between the transitions from S_0 to S_1 and from S_0 to S_2 can be estimated to be $\sim 30^\circ$. The third plateau (~ 300 nm, $r_{\theta}=0.05$) suggests the angle between S_0 - S_1 and S_0 - S_3 to be at least 50° . This is in line with a larger angle predicted between S_0 - S_1 and S_0 - S_3 (83°) than between S_0 - S_1 and S_0 - S_2 (15°) by calculations.

The predicted oscillator strength for the lowest energy transition of qA ($f=0.0594$) is approximately 12 times lower than that predicted for the A^T -derivative ($f=0.72$)^[37]. Knowledge of the lifetimes ($\tau_{f,A^T}=1.8$ ns and $\tau_{f,qA}=3.2$ ns; water) and quantum yields of the analogues ($\Phi_{f,A^T}=61\%$ and $\Phi_{f,qA}=6.8\%$; water) also allows the calculation of their radiative rate constants, k_f , which for qA is roughly 16 times decreased compared to A^T . This corresponds reasonably well to the predicted decrease in the oscillator strength as k_f and f are related (see equation 9 and 14).

Structure and stability of DNA duplexes containing A^T and qA

In order to investigate the influence of A^T (paper I) and qA (paper III) on the DNA duplex integrity, they were incorporated in a series of decamers for which the bases directly flanking the analogues were varied (Table 2). In this way the effect of as many combinations of purines and pyrimidines as possible was evaluated, as immediate 5' and 3' neighbors, depending on the amount of compound available. For A^T , ten different combinations of flanking bases were tested, whereas a larger amount of qA allowed us to investigate twelve sequences. This also included a decamer holding two qA molecules (CT,TA).

The melting temperature of all A^T and qA-modified duplexes was recorded and compared to the corresponding natural ones in order to evaluate the effect of the adenine modifications on duplex stability (Table 2). A^T causes an average drop in melting temperature of 8°C , indicating a moderate destabilization of the DNA

duplexes, which is most pronounced for sequences in which A^T is flanked by a 3' purine. This destabilizing effect is probably due to the triazole group and pentyl chain incorporated on C8 of adenine. Steric clashes of these substituents with the DNA backbone would account for a significant energy cost in case A^T was to be accommodated in the duplex in its natural *anti*-conformation. Even though smaller C8-substituents on purines are only insignificantly perturbing the duplex integrity,^[69,115] more bulky modifications have been observed to be destabilizing,^[116-118] similar to A^T. For some C8-modified purines, a preference for the *syn* conformation has been reported,^[119,120] whereas a *syn/anti* equilibrium has been suggested for others inside DNA.^[121] Also for A^T, the C8-triazole modification may cause a stabilization of the *syn*-conformation, relative to the *anti*-orientation.

Table 2 DNA melting temperatures of different 10-mer duplexes containing A^T ($T_m^{A^T}$, left) or qA (T_m^{qA} , right). Sequences are named after the bases directly neighboring A^T or qA. Also the difference in melting temperature (ΔT_m) is shown between the modified and unmodified (containing natural adenine) DNA duplex.

Neighbouring bases A ^T ^a	$T_m^{A^T}$ (°C) ^b	ΔT_m (°C)	Neighbouring bases qA ^a	T_m^{qA} (°C) ^b	ΔT_m (°C)
<i>CA</i>	43	-9	AA	46.3	-0.3
GG	45	-9	GA	48.3	0.0
AA	41	-9	GG	51.9	0.4
<i>TA</i>	39	-8	GC	52.3	1.5
TG	42	-8	<i>AC</i>	52.7	1.9
GA	43	-7	<i>TA</i>	48.5	3.2
<i>CT</i>	45	-7	CA	53.2	3.3
<i>AC</i>	46	-7	TG	48.7	3.4
<i>CC</i>	49	-7	<i>CC</i>	58.2	5.3
GC	49	-5	<i>CT</i>	56.1	6.5
			<i>TT</i>	50.6	6.9
			<i>CT,TA</i>	55.6	9.2

^aFull sequence is 5'-d(CGCA^TAYTCG)-3', with exception of sequence *CT,TA*; 5'-d(CGCA^TAATCG)-3' and *TT*, 5'-d(CGATTATGCG)-3'. **A** represents A^T or qA and the bases flanking them, X and Y, give the sequence name XY. Purines are shown in bold and pyrimidines in italic. ^bSamples were prepared in phosphate buffer (pH 7.5, 500 mM (A^T) or 200 mM (qA) Na⁺) at duplex concentrations of 2.5 μM.

In order to investigate how stable A^T is stacked into the DNA, its anisotropy was recorded inside duplexes **AA** and *TA* in viscous sucrose solutions. This high viscosity hinders the motion of the DNA helices on the time-scale of fluorescence, resulting in a limiting anisotropy value of 0.34 reached by A^T. The decrease compared to r_0 of 0.38 for the monomer in a vitrified matrix was predicted to be

due to an internal wobble of 16° inside the base-stack, compared to $\sim 5^\circ$ for canonical bases and 21° for ethidium bromide (personal communication from S.Preus).^[122] Consequently, despite a possible conformational change around the glycosidic bond, A^T seems to be stacked more stably in the duplex than the intercalating dye ethidium bromide, indicating that it still has base-pairing capacity to thymine (*vide infra*).

In contrast to A^T and other adenine analogues such as 2-AP,^[33] 6MAP and DMAP,^[61] qA generally stabilizes the DNA duplexes (Table 2). This can be seen by an average increase in melting temperature of 3°C compared to the natural duplexes. The largest stabilization is seen when qA is flanked by a pyrimidine on the 5' side, whereas this effect is minimal for a 5' neighboring purine. Sequences **AA**, **GA** and **GG**, having two purines neighboring qA, are virtually equally stable as the corresponding duplexes containing a natural adenine instead. The overall stabilization caused by qA is probably due to better π - π stacking with its neighboring bases due to the extended ring system. Similar observations have been made for the tricyclic cytosines tC ^[41] and tC^O ^[39], which also showed the largest stabilizing effect in case they had a 5' pyrimidine neighbor.

The two ring extension for qA, on position 6 and 7 of adenine, is probably well accommodated in the major groove. Also other purines modified on the 7-position have shown a stabilizing effect on DNA duplexes.^[123-125] Overall, qA seems to be stably incorporated into the DNA duplex. Moreover, fluorescence melting curves for duplex **AA** (Figure 17) yielded a virtually identical melting temperature (46.8°C) to that recorded with UV/Vis melting (46.3°C), indicating no pre-melting around qA.

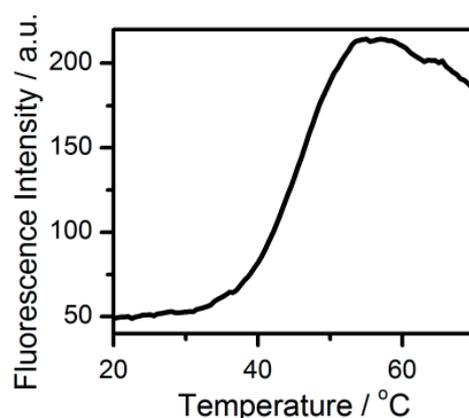


Figure 17 Fluorescence melting curve of qA inside sequence **AA** in phosphate buffer (pH 7.5, 200 mM Na^+ , duplex concentration of 2.5 μM).

In order to evaluate the effect of A^T and qA on the overall DNA conformation, the CD spectra of all modified sequences were recorded (Figure

18). Both A^T - and qA-modified duplexes show a general signature of regular B-form DNA, for which the typical parameters were described in Section 3.3. However, for A^T , some heterogeneity was observed in the shape of their CD spectra between sequences compared to the corresponding unmodified duplexes (Figure 18a). There are sequences with an almost identical CD-profile to the natural duplexes (*CT* and *AA*) and others which show slight differences corresponding to the A^T absorption bands (*CA*, *TG* and *CC*). Moreover, more distinct differences were observed for sequences *TA*, *AC*, *GA*, *GC* and *GG*, of which the latter three show the most perturbed CD spectra. As an example, the CD spectrum of *GG* with A^T or adenine is shown in Figure 18c, as well as the difference in absorption between both duplexes. As can be seen, the deviations in the CD-signature between both samples correspond well to the difference in absorption.

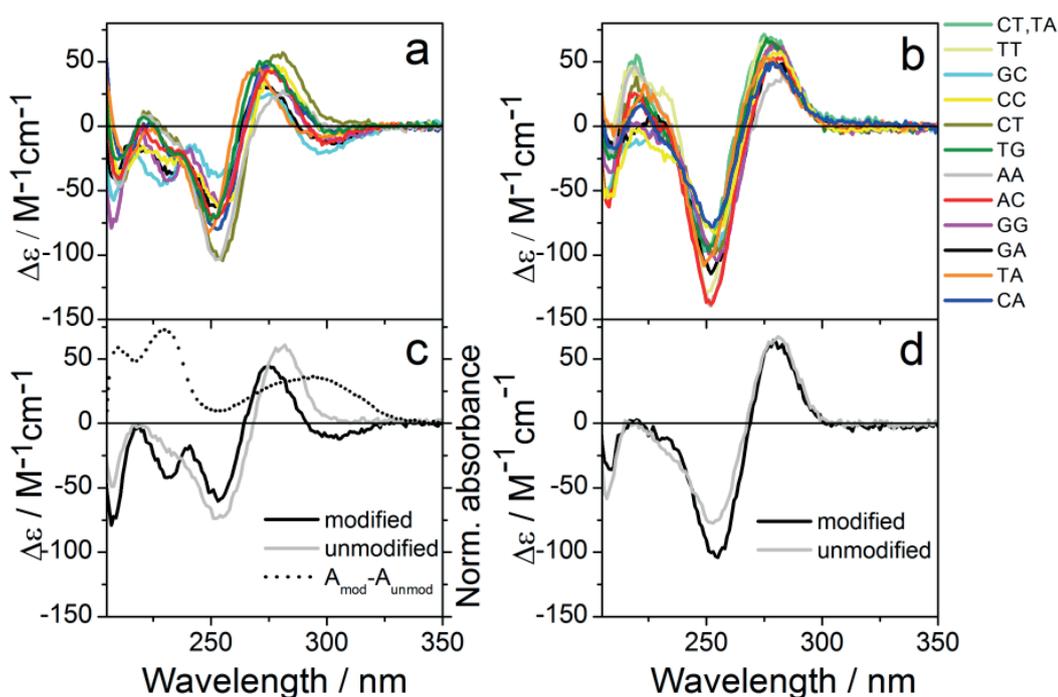


Figure 18 CD spectra of DNA duplexes containing A^T (a) and qA (b) (sequences in Table 2; *CT,TA* and *TT* are unique for qA). No CD signal is seen for qA above 350 nm. Also the comparison of CD signatures of sequence *GG* containing A^T (c) and qA (d) (black) to the spectrum of the unmodified *GG* duplex (grey) is shown. For A^T , the difference in the corresponding absorption spectra is shown as well ($A_{mod} - A_{unmod}$, dotted). M^{-1} on the x-axis refers to the duplex concentration ($2.5 \mu M$).

Contrary to A^T , the CD spectra of all samples containing qA are very similar in shape (Figure 18b). Furthermore, they all correspond well to those of the unmodified sequences. As for A^T , sequence *GG* is shown as an example (Figure 18d). This finding may be expected since qA seems to be better tolerated inside

DNA than A^T according to the melting studies. Additionally, also for sequence CT,TA containing two qA molecules, a similar CD-signature was observed, indicating that also the incorporation of several qA molecules preserves the B-DNA structure. As for some A^T -modified duplexes (CT and AA), no specific CD signal can be detected corresponding to the lowest energy absorption band of qA, centered around 337 nm. Similar observations have been made for the cytosine analogue tC^O ,^[39] for which we do not yet have a satisfactory explanation.

Base-pairing preference of A^T and qA

The base-pairing specificity of A^T and qA was investigated by recording the melting temperature for three different duplexes in which the analogues were placed opposite of a guanine, cytosine or adenine (Table 3). Sequences were chosen so that the effects of neighboring purines (GA), pyrimidines (CT) and a combination of both (CA for A^T and TA for qA) could be evaluated.

An additional decrease in melting temperature of $\sim 8^\circ\text{C}$ was observed for A^T opposite of cytosine or guanine in comparison to the A^T -thymine case, yielding a total destabilization of approximately 16°C . This value is comparable to an adenine-adenine mismatch in these sequences causing a decrease in melting temperature of 14°C . It is therefore plausible that A^T still exhibits hydrogen bonding capacity to thymine despite its destabilizing effect. Surprisingly, no additional decrease in melting temperature was observed when placing A^T opposite of adenine compared to the $A^T.T$ case indicating that A^T can form equally stable base-pairs with thymine as with adenine. A potential $A^T.A$ base pair is presented in Figure 19 and compared to a possible $A^T.T$ pair with A^T in the *syn* conformation around the glycosidic bond. Perturbations in the base-pairing patterns are not unexpected upon adding (often bulky) substituents to the natural bases. Also other FBAs such as 2-AP^[34-36] and BPP^[58] have been shown to have deviating base-pairing specificities compared to their natural counterparts.

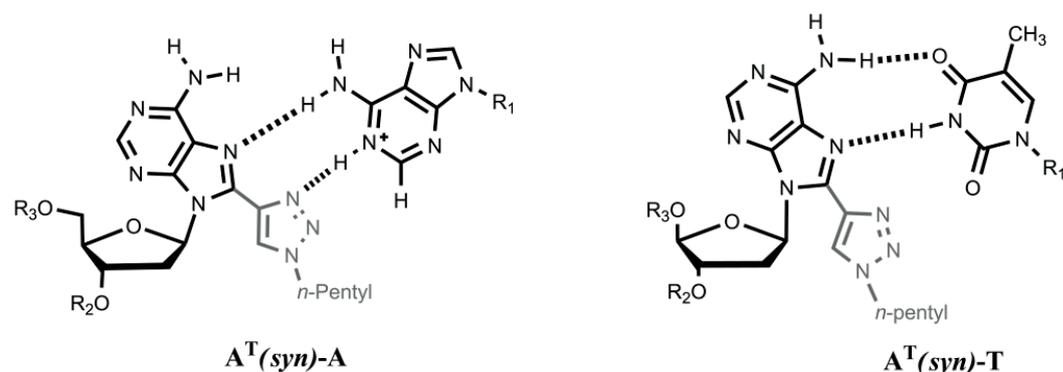


Figure 19 Putative $A^T.A$ (left) and $A^T.T$ (right) base-pair with A^T in the *syn* conformation around the glycosidic bond. R_1 , R_2 and R_3 represent the rest of the DNA structure. The triazole substituent on A^T is shown in grey.

In the case of qA, a significant destabilization was observed for a mismatch with cytosine, guanine and adenine compared to the qA-thymine case, indicating that qA shows specific base-pairing to thymine (Table 3). The decrease in melting temperature was the lowest for qA opposite of guanine and the same observation was made for the mismatched sequences with natural adenine. These findings indicate that the base-pairing pattern of qA is unperturbed compared to natural adenine. This is also in line with the assumption that the extended rings are well accommodated in the major groove, leaving qA positioned similar to natural adenine in the base-stack.

Table 3 Melting temperatures of sequences containing A^T (T_m^{AT-X}) or qA (T_m^{qA-X}) mismatched with G,C or A are presented as well as the difference with the melting temperature (ΔT_m^{AT} and ΔT_m^{qA}) of the corresponding matching duplex containing the analogues opposite of T.

Neighboring bases ^a	X ^c	T_m^{AT-X} (°C) ^d	ΔT_m^{AT} (°C)	T_m^{qA-X} (°C) ^d	ΔT_m^{qA} (°C)
GA	G	36	-7	38.8	-9.5
	C	36	-7	34.1	-14.2
	A	43	0	34.2	-14.1
CT	G	36	-9	47.6	-8.5
	C	36	-9	39.3	-16.8
	A	46	1	40.5	-15.6
CA (A ^T) ^b	G	36	-7	37.0	-11.5
	C	36	-7	32.8	-15.8
TA (qA)	A	42	-1	33.5	-15.0

^aFull sequences in Table 2. ^bSequence CA was used for A^T and TA in the case of qA. ^cBase opposite of A^T or qA. ^dConditions were similar as described for Table 2.

Photophysical properties of A^T and qA in DNA

The fluorescence properties (Φ_f and $\lambda_{Em,max}$) of A^T and qA inside DNA are summarized in Table 4 and Table 5, respectively. As is common upon incorporating FBAs into nucleic acid systems,^[27,29,30,32,69,73,115] the quantum yield of both A^T and qA generally is significantly decreased in single- and double-stranded DNA. Furthermore, variations can be seen between neighboring bases, which means that the emission of both probes is responsive to their immediate surroundings.

For A^T ($\Phi_f=61\%$ in water), the average quantum yield for all combinations of flanking bases is decreased approximately 12 times for the single strands and additionally 5 fold in double-stranded DNA. Furthermore, large variations can be seen depending on the neighboring bases. These observations may be due to different levels of stacking with the surrounding bases in single- and double-

stranded DNA as well as different conformations of A^T inside DNA. In general the highest quantum yields are detected in the case where A^T is flanked by an adenine on the 3'-side (exception duplex **GA**), whereas the lowest quantum yields are detected for a neighboring 3' guanine. The fact that sequence **GG** is among those with the lowest quantum yield may be due to photoinduced electron transfer (PET). Guanine has the lowest oxidation potential among the natural bases and could therefore quench the excited state of the neighboring A^T by PET. Decreases in quantum yield due to neighboring guanines have also been observed for other FBAs such as 2-AP^[126,127], tC^O^[39] and BPP^[58]. In contrast to sequences **GG** and **TG**, those with a guanine as the 5'-neighbor of A^T (**GA** and **GC**) have higher quantum yields in single- and double- stranded DNA. Similar observations had been made for tC^O previously,^[39] indicating that not only proximity to a guanine but also stacking and relative orientation are important parameters for PET.

Table 4 Fluorescence quantum yields (Φ_f) of the A^T-modified single- and double-stranded oligonucleotides, $\Phi_{f,SS}$ and $\Phi_{f,DS}$, respectively. Also their emission maxima are shown ($\lambda_{Em,max}$).

Neighboring bases ^a	$\Phi_{f,SS}$ (%)	$\lambda_{Em,max}$ (nm)	Neighboring bases ^a	$\Phi_{f,DS}$ (%)	$\lambda_{Em,max}$ (nm)
AA	21	353	AA	5.0	350
CA	9.0	351	TA	1.6	353
TA	7.6	352	CA	1.1	354
GC	4.2	358	AC	0.9	349
AC	3.1	357	GA	0.8	351
CC	2.1	355	CT	0.5	351
GA	1.8	353	GC	0.5	351
CT	1.4	358	CC	0.4	352
TG	0.6	355	TG	0.3	351
GG	0.5	355	GG	0.3	354

^aSequences are named as described under Table 2.

A decrease in average fluorescence quantum yield of ~4 and additionally ~5 times can be detected for qA incorporated into single-stranded and duplex DNA, respectively, compared to the monomer in water ($\Phi_{f,\lambda_{Ex337nm}}=6.8\%$) (Table 5). As for A^T, this may be explained by various levels of stacking with the surrounding bases. In the case of qA, the highest quantum yields are detected when it is positioned between two purines, including sequence **GG**, which is in contrast to findings for A^T and the FBAs mentioned above. However, we cannot draw any conclusions regarding PET since we have not determined the redox potentials of qA. As for A^T and other FBAs, the quenching of qA inside DNA is probably due to a combination of mechanisms, such as base stacking and collisions with the neighboring bases,^[78] hypochromism^[39,128] as well as PET,^[127] which are dependent on the surrounding sequence and structural parameters.

Maximum quantum yields for A^T (21%) and qA (5.8%) in single-stranded DNA exceed values reported for nucleic acids containing 2-AP.^[32] For A^T, this value is approximately 10 times higher compared to 2-AP^[32] and roughly five and two times higher than values reported for 6MAP and DMAP^[61], respectively. Inside duplexes, the maximum quantum yield of A^T (5%) and qA (0.6%) is roughly five times higher or of the same order, respectively, as 2-AP.^[32] The overall brightness ($\propto \langle \Phi_{f,DNA} \rangle \times \epsilon_{max}$) of A^T inside duplex DNA ($\sim 190 \text{ M}^{-1}\text{cm}^{-1}$) is, thus, roughly 10 times higher than qA ($\sim 20 \text{ M}^{-1}\text{cm}^{-1}$) and 3 times higher than 2-AP ($\sim 60 \text{ M}^{-1}\text{cm}^{-1}$) in a similar environment.^[32,114]

Table 5 Fluorescence quantum yields (Φ_f) of the qA-modified single- and double-stranded oligonucleotides, $\Phi_{f,SS}$ and $\Phi_{f,DS}$ respectively. Also their emission maxima are shown ($\lambda_{Em,max}$).

Neighboring bases ^[a]	$\Phi_{f,SS}$ (%) ^b	$\lambda_{Em,max}$ (nm) ^[b]	Neighboring bases ^[a]	$\Phi_{f,DS}$ (%) ^b	$\lambda_{Em,max}$ (nm) ^[b]
AA	5.8	432	AA	0.6	434
GG	5.5	433	GG	0.6	435
GA	4.4	432	GC	0.6	453
AC	2.0	434	GA	0.5	438
GC	1.6	438	CA	0.3	444
CA	1.0	433	CT,TA	0.3	457
TA	0.6	434	AC	0.2	441
CT,TA	0.4	439	TA	0.2	451
TG	0.4	437	TG	0.2	453
CC	0.4	440	CT	0.2	459
CT	0.3	443	CC	0.2	452
TT	0.2	434	TT	0.2	452

^aSequences are named as described under Table 2.

Conclusion about A^T, qA and A^{7T}-family

The emission of both qA and A^T is sensitive to their micro-environment. As for other FBAs, this sensitivity could be applied in studies concerning dynamics, structure^[79,80,129] and interactions of DNA with other molecules^[82,84,130]. Further, qA is a better adenine isomorph, whereas A^T is the better choice in case a higher brightness is crucial.

Even though the members of the new A^{7T}-family are dramatically less emissive than A^T in methanol, it will be interesting to investigate the most promising compounds, **12a** ($\Phi_f=5.2\%$) and **12b** ($\Phi_f=4.4\%$), in DNA since it is very likely that they will be well accommodated in the DNA duplex. Also **20b** ($\Phi_{f,DCM}=21\%$, $\Phi_{f,ACN}=24\%$) would be interesting to investigate in a DNA context where it is more shielded from protic solvent molecules that efficiently quench its fluorescence ($\Phi_{f,MeOH}=0.62\%$). However, one should consider that also inside

DNA there will be effects of hydrogen bonding from the opposite base and the triazole extension exposed into the major groove.

4.1.2 Photostability of the tricyclic cytosine analogue tC

In order to develop FBAs which are brighter and more photostable to allow single-molecule studies, a good understanding of their photodegradation is important. Therefore, the characterization of the photostability of tC, a member of the tricyclic cytosine family, is presented here (paper IV) (Figure 20a). The photoconversion of tC to its photoproduct tC[#] was studied both as a monomer and inside duplex DNA.

Upon intense irradiation of the tC monomer in its lowest energy absorption band (~375 nm) the latter is slowly replaced by a blue-shifted lowest energy absorption band centered around 310-320 nm (Figure 20b). The two isosbestic points (at 285 nm and 343 nm) indicate the conversion of tC to a single photoproduct tC[#], with a strongly decreased blue-shifted emission ($\Phi_f < 2\%$) compared to tC. Analysis of the photoproduct by NMR spectroscopy and mass spectrometry as well as the increase in the photoconversion rate upon saturation with O₂ suggest tC[#] to be the sulfoxide form of tC (Figure 20a). Furthermore, the lowest energy electronic transition of tC[#] predicted by DFT-calculations (317 nm) corresponds well to the observed absorption band of the photoproduct in water. The sulfoxide could be formed after tC undergoes intersystem crossing from its first excited singlet state to a triplet state and reacts with O₂.

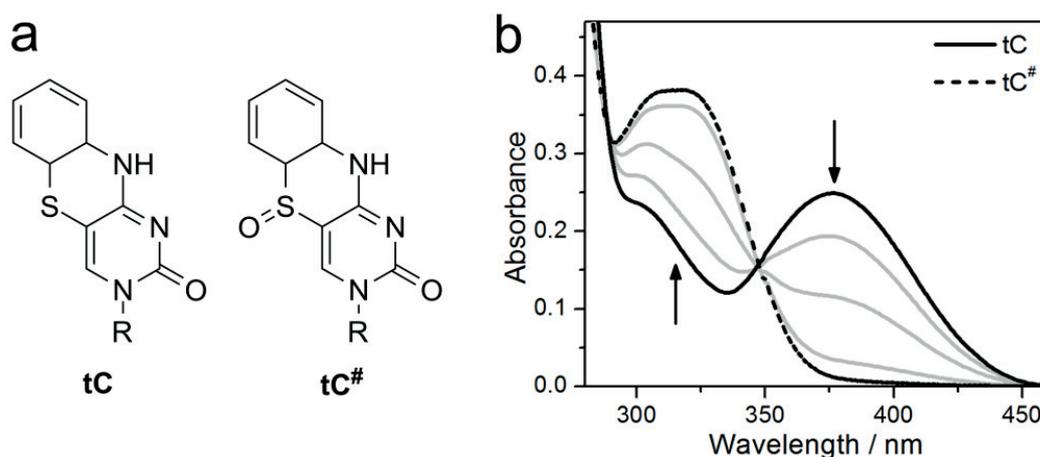


Figure 20 (a) Tricyclic cytosine analogue tC and its suggested photoproduct tC[#]. (b) Evolution of the UV-Vis absorption spectrum of the potassium salt of tC in water upon irradiation at 375 nm with a 150 W Xe lamp. The spectra were recorded over an irradiation period of 10 hours.

The photostability of tC was not only studied as a monomer, but also inside duplex DNA, where it was incorporated into a 10-mer oligonucleotide flanked by two adenines (Table 6, Figure 21). The photoconversion of tC inside double-stranded DNA was found to be significantly slower under similar conditions as for the monomeric form. This also supports the hypothesis that oxidation of the sulfide leads to the photoproduct, since tC would be less accessible to O₂ inside the base stack.

Table 6 Melting temperatures (T_m) of DNA duplexes containing a central cytosine, tC or its photoproduct tC[#].

Sequence ^a	X ^b	T_m (°C)
5'-d(CGCAACATCG)-3'	G	41.6
5'-d(CGCAAtCATCG)-3'	G	43.6
5'-d(CGCAAtC [#] ATCG)-3'	G	26.0
5'-d(CGCAAtCATCG)-3'	T	24.0

^aComplementary sequence is 5'-d(CGATXTTGGC)-3' ^bX is the base in the complementary sequence that is opposite of the central base C, tC or tC[#].

Interestingly, the stability of the DNA is significantly decreased upon photoconversion of tC ($T_m=43.6^\circ\text{C}$) to tC[#] ($T_m=26.0^\circ\text{C}$), resulting in partly denatured duplexes at room temperature (Table 6). This destabilization is comparable to a tC-thymine mismatch ($T_m=24^\circ\text{C}$) and could be explained by steric effects of the oxygen atom. The latter is predicted to be perpendicular to the surface of tC[#] (DFT calculations) and would therefore interfere with the base flanking tC[#] on the 5' side.

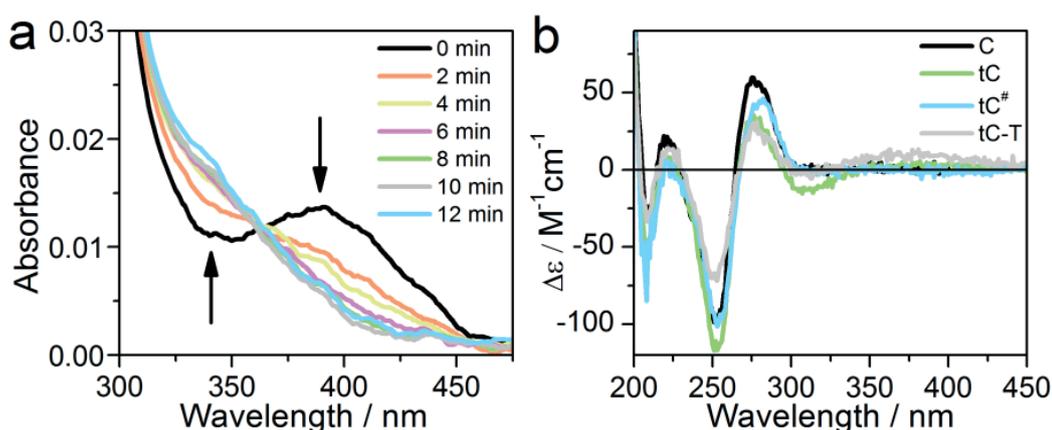


Figure 21 (a) Evolution of the UV-Vis absorption spectrum of tC inside duplex-DNA during fast photolysis upon irradiation with a Nd:YAG laser at 420 nm for 0-12 minutes (7 ns pulses, 10 Hz, 2.1 mJ/pulse). The solution was bubbled with O₂ every 2-4 minutes. Spectra are smoothed by adjacent averaging (10 points). (b) CD-spectra of DNA duplexes containing a central natural cytosine, tC or tC[#]. Sequences are shown in Table 6.

On the other hand, it should be mentioned that CD-spectra for the DNA duplexes containing tC[#] opposite of guanine or a tC-thymine mismatch still show the same spectral features of B-DNA (see 3.3) as the tC-G case and the natural duplex (Figure 21b). Therefore we cannot exclude the possibility that the observed destabilization is (in part) due to a reduced base-pairing capacity of tC[#] to G. Finally, when taking a closer look at the recorded CD spectra, the tC[#]-profile looks more similar to the unmodified duplex (C) than that of the tC-modified duplex. This could be explained by the difference in their isotropic absorption (Figure 20b).

As was mentioned in the paragraph above, the photodegradation of tC inside DNA was much slower (12-24 hours, 200 W Hg lamp) under similar conditions as compared to the monomer (~4 hours). However, fast photolysis (<10 min) can be achieved using more intense irradiation (Figure 21a). This indicates the potential of the photoconversion of tC as a trigger for DNA melting with potential applications in photoinduced therapeutics and in functional DNA-based nanodevices.^[131-133]

4.2 Application of the tricyclic cytosine FRET-pair to investigate protein-DNA interactions

The mitochondrial DNA (mtDNA) is a circular molecule that encodes 22 tRNAs (transfer RNA), 2 rRNAs (ribosomal RNA) and 13 proteins involved in the respiratory chain. Transcription is initiated from two sites, *i.e.* the light- and heavy-strand promoters (LSP and HSP1).^[134] A third transcription site has been located as well, HSP2, of which the function and sequence requirements still remain to be established.^[135,136]

In mammalian cells, transcription involves the mitochondrial RNA polymerase (POLRMT, human or Polrmt, mouse) and transcription factor B2 (TFB2M, human and Tfb2m, mouse). Also a third factor, the transcription factor A (TFAM, human or Tfam, mouse), has been reported to be involved,^[136-138] even though its role has been debated lately^[139]. In paper V the role of TFAM in mammalian transcription is established, whereas paper VI focuses on the assembly of the transcription machinery to the promoter. The nucleobase FRET-pair, consisting of tC^O and tC_{nitro} (see 3.4.6)^[40], was applied to monitor structural changes around the transcription start site in both studies.

Structural influence of TFAM on the heavy strand promoter

TFAM binds site-specifically to a site upstream of the transcription start site of LSP and HSP1 (Table 7).^[140,141] However, it is also known to bind non-sequence specifically as an mtDNA packaging factor, which wraps and bends the DNA.^[142-145] Even though a recent report disputed the role of TFAM in mammalian mitochondrial transcription,^[139] *in vitro* transcription experiments presented in paper V establish the necessity of TFAM. Only at lower salt concentrations (mainly for HSP1) or on a negative supercoiled template (only for HSP1), transcription can be detected in the absence of TFAM. This observation may indicate that the role of TFAM in transcription initiation is to induce negative supercoils in promoter DNA, thereby stimulating local melting by POLRMT and TFB2M. Furthermore, the TFAM-binding site is essential, since transcription from both LSP and HSP1 is abolished for a template with mutations in the specific TFAM-binding site.

The tC^O - tC_{nitro} FRET-pair was applied to investigate whether TFAM induces structural changes around the transcription start site. The tC^O donor and tC_{nitro} acceptor were therefore incorporated close to the HSP1 transcription start site with 6bp distances separating them in the opposite strands (Table 7). Since the probes are rigidly incorporated in the DNA duplex, their FRET efficiency is expected to respond to subtle structural changes around the site of incorporation.

Table 7 DNA sequence of the heavy strand promoter (HSP1), in which the FRET donor tC^O and acceptor tC_{nitro} were incorporated at sites X_1 - X_7 (see Table 8). The TFAM binding site is underlined and the transcription start site is highlighted grey. Mutations in the sequence are italicic.

DNA sequence ^a	
5'-d(-----CACACACCGCTGCTAACCCCATACCCCGAACCAACCAAAX ₁ X ₂ CCAAAGGCAC-----)-3'	
3'-d(-----GTGTGTGGCGACGATTGGGGTATGGGGCTTGTTGGTTTGGGGTTTXX ₃ X ₄ GTG-----)-5'	
5'-d(-----CAACACAATAGTAGCCAAAACGCCCGAACCAACCAAAX ₅ CCCAAAGGCAC-----)-3'	
3'-d(-----GTTGTTGTTATCATCGGTTTTCGGGGGGCTTGTTGGTTTGGGGTTTXX ₆ CGTG-----)-5'	
5'-d(-----CACACACCGCTGCTAACCCCATACCCCGAACCAACCAAAX ₇ TATCCCCGCAC-----)-3'	

^aFull sequence used is 70 nt long. The remaining sequence is depicted as -----.

To begin with, four donor-acceptor combinations were tested with tC^O and tC_{nitro} incorporated both in the coding (X_1, X_2) and template strand (X_3, X_4) (Table 7 and Table 8). The FRET efficiencies obtained for these samples (17-21%; Table 8) confirm the integrity of the duplex since values are similar to those recorded in a previous study for a 6 bp separation between tC^O and tC_{nitro} in B-DNA (using fluorescence lifetimes: $E_{\text{FRET}}=20\%$)^[40]. Next, TFAM was added to these sequences, resulting in increases in FRET efficiencies between 8.3-15%. This

change in FRET indicates that TFAM causes significant structural changes around the transcription start site.

We also determined the FRET efficiency of a sequence having a 7bp bubble spanning the transcription start site (D₄-A_{BUB}, Table 8) and found a value of 65%. This value is considerably larger than for the corresponding matching (D₄-A₄) duplex with TFAM bound ($E_{FRET,TFAM}=29\%$), indicating that TFAM may induce premelting around the transcription start site but that structural changes are not as invasive as a 7bp unpaired region.

Finally, the structural effect of TFAM was investigated in a duplex with mutations in the TFAM-binding sequence (D_{MUT}-A_{MUT}). No transcription could be detected *in vitro* for sequences with mutations in this site. However, the same change in FRET efficiency was observed for the mutated duplex as for the corresponding wild-type sequence (D₁-A₁) upon addition of TFAM (Table 8). Consequently, the structural changes induced by TFAM around the transcription start site seem rather due to its capacity to bind and bend DNA in a sequence-independent manner. Also a gel retardation experiment showed no difference in TFAM binding to a wild type or mutated promoter sequence. These findings indicate that TFAM not only induces structural alterations in the promoter sequence but is probably also engaged in complex interactions with the rest of the transcription machinery, for which the TFAM binding site is important.

Table 8 FRET efficiencies of tC^O-tC_{nitro} in four different DNA duplexes containing the wild-type TFAM binding sequence as well as in a duplex with a mutated TFAM binding site (D_{MUT}-A_{MUT}). Also a duplex with a 7 bp bubble spanning the transcription start site was tested (D₄-A_{BUB}) to which no TFAM was added.

Duplex	tC ^{O,a}	tC _{nitro} ^a	E_{FRET} (%) ^b	$E_{FRET,TFAM}$ (%) ^c	ΔE_{FRET} (%) ^d
D1-A1	X ₁	X ₃	21 ± 0.2	31 ± 0.1	10 ± 0.2
D2-A2	X ₂	X ₄	17 ± 0.3	29 ± 0.4	12 ± 0.5
D3-A3	X ₄	X ₂	17 ± 0.3	32 ± 0.8	15 ± 0.9
D4-A4	X ₃	X ₁	21 ± 0.4	29 ± 0.4	8.3 ± 0.6
D _{MUT} -A _{MUT}	X ₅	X ₆	21 ± 0.2	30 ± 0.8	10 ± 0.8
D4-A _{BUB}	X ₃	X ₇	65 ± 0.06		44 ± 0.4 ^e

^aPosition of tC^O or tC_{nitro} in the corresponding sequences shown in Table 7. ^bFRET efficiency upon addition of the corresponding buffer TFAM is dissolved in (except for D₄-A_{BUB}). ^cSamples contain a ratio of 3:1 TFAM:DNA. ^dDifference in FRET efficiency with and without TFAM present. ^e $\Delta E_{FRET}=E_{FRET,D4-ABUB}-E_{FRET,D4-A4}$

Investigating the step-wise assembly of the transcription machinery

In paper VI, the role of the N-terminal extension (NTE) of Polrmt (mouse) together with Tfam (mouse) was investigated in ensuring promoter-specific transcription. A model is put forward in which Tfam recruits Polrmt to the

mitochondrial promoters, thereby relieving the inhibitory effect of the NTE of Polrmt. In a next step, Tfb2m binds to the complex, allowing tight interactions with the transcription start site and initiating transcription. This model is supported by results from DNase I footprinting on the LSP promoter.

Also in this study we used the tC^O - tC_{nitro} FRET-pair to investigate whether FRET-data would agree with the proposed model for transcription initiation. For these experiments, the human HSP1, TFAM, POLRMT and TFB2M were applied, also allowing the verification of the model at the mitochondrial promoter of another organism. Duplex D_4 - A_4 was used in these studies, with tC^O - tC_{nitro} incorporated 6bp apart in opposite strands around the HSP1 transcription start site as described above (Table 7 and Table 8). The change in FRET efficiency was monitored upon addition of the three proteins in different orders (Table 9). POLRMT alone has virtually no effect on the FRET efficiency ($\Delta E_{FRET} = -0.2$ - 0.6%). Subsequent addition of TFAM to this system only yields a small change in FRET efficiency ($\Delta E_{FRET} = 2\%$) compared to the effect of TFAM without POLRMT present ($\Delta E_{FRET} = 8.3\%$) (Table 9). Similarly, the presence of both POLRMT and TFB2M leaves the FRET efficiency virtually unaltered ($\Delta E_{FRET} = 0.4\%$). However, upon subsequent addition of TFAM to the latter case, a change in FRET efficiency (12-13%) is observed which is larger than for TFAM alone (8.3%). A possible explanation for these results could be that TFAM recruits POLRMT to the promoter, with the latter suppressing non-specific interactions of TFAM with the transcription start site. Next, binding of TFB2M allows melting around the transcription start site and initiation of transcription. The structural changes around the transcription start site suggested by our nucleobase FRET-pair are in agreement with the model proposed above for the mouse LSP system.

Table 9 FRET efficiency of tC^O - tC_{nitro} in duplex D_4 - A_4 (for sequences see Table 7 and Table 8) upon addition of POLRMT, TFAM and TFB2M ($E_{FRET,protein}$) or corresponding buffers ($E_{FRET,buffer}$). Proteins were added in the order shown (1-3, 1-4 or 1 (only TFAM)).

DNA+protein	DNA:protein ^a	$E_{FRET,buffer}$ (%)	$E_{FRET,protein}$ (%) ^c	ΔE_{FRET} (%) ^b
1) POLRMT	1:1.1	21 ± 0.2	21 ± 0.1	0.6 ± 0.2
2) TFAM	1:1.5	21 ± 0.2	23 ± 0.5	1.6 ± 0.5
3) TFAM	1:3	22 ± 0.1	24 ± 0.1	2.0 ± 0.1
1) POLRMT	1:1.1	21 ± 0.3	21 ± 0.2	-0.2 ± 0.4
2) TFB2M	1:1.1	21 ± 0.5	21 ± 0.5	0.4 ± 0.7
3) TFAM	1:1.5	21 ± 0.2	33 ± 0.5	12 ± 0.5
4) TFAM	1:3	21 ± 0.4	34 ± 0.2	13 ± 0.4
1) TFAM	1:3	21 ± 0.4	29 ± 0.4	8.3 ± 0.6

^aRelative molar ratios of DNA and proteins added. ^bDifference in FRET efficiency with and without proteins added.

5 Conclusion and outlook

The main focus of this thesis has been on fluorescent nucleobase analogues, FBAs, which are probes for studying different aspects of DNA and RNA such as their structure, dynamics and interactions with other molecules.

In the first part of this thesis, the characterization of photophysical and base-mimicking properties of FBAs has been central. The spectroscopic characterization of ten adenine analogues; A^T, qA and the A^{7T}-family; was presented, of which the first two were also incorporated into DNA (paper I-III). These adenine analogues show interesting variations in their emissive and base-mimicking properties based on their structure. A^T (paper I) shows a high fluorescence quantum yield as a monomer ($\Phi_f=61\%$ in water). Inside DNA, values reach up to 21% for single strands and 5% for double strands. Its overall brightness inside duplex DNA is roughly three times higher than the most commonly used FBA, 2-AP, in a similar environment.^[32,114] However, as for other C8-modified purines,^[116-118] A^T moderately destabilizes the duplex structure. Surprisingly, this modification also enables A^T to base-pair with adenine besides thymine.

In order to overcome the destabilizing effect caused by A^T, a new series of 7-modified triazole adenines, A^{7T}-family (paper II), was developed, whose photophysical properties as monomers have been discussed here. The members of this new family were found to be significantly less emissive than A^T in methanol. Nevertheless, it will be interesting to investigate the fluorescent sister compound of A^T, **12a** ($\Phi_f=5.2\%$) and the related **12b** ($\Phi_f=4.4\%$) in DNA, since it is very likely that they will be well accommodated in the duplex. A stabilizing effect has namely been reported for other 7-substituted 8-aza-7-deazaadenines.^[125] Moreover, even though most FBAs are quenched inside nucleic acid systems, increasing quantum yields have been reported for some analogues.^[59] In the same line, compound **20b** may show promising fluorescence properties inside the base-stack where it is more shielded from protic solvent molecules ($\Phi_f\sim 21\text{-}24\%$ in ACN or DCM compared to 0.62% in methanol). However, hydrogen bonding to the opposite base and to the triazole moiety protruding in the major groove may have strong influences. Furthermore, it could be interesting to explore a four-ringed variant of the A^{7T}-family in which the carbon on position 5 of the triazole ring is connected to the exocyclic amino moiety of adenine, thereby creating a more rigid system, potentially with an increased emission.

The last adenine analogue studied in this work, qA (paper III) shows a moderate quantum yield as a monomer ($\Phi_f=6.8\%$ in water) and maximum values of 5.8 % and 0.6 % in single- and double-stranded DNA, respectively. However,

inside duplex DNA, the overall brightness of qA is approximately ten times lower than for A^T. On the other hand, qA is a better isomorphous adenine analogue since it shows specific base-pairing to thymine with the same preference for the type of opposite base as natural adenine. Furthermore, qA on average stabilizes the DNA duplexes, as has also been found for the ring-extended tricyclic cytosines tC^[41] and tC^{O[39]}. The increase in melting temperature is dependent on the directly neighboring bases and indicates that the two-ring extension is well accommodated in the major groove. At the moment, new derivatives of qA are being characterized which, according to calculations, are predicted to have increased oscillator strengths. First results have shown improved photophysical properties, in accordance with these predictions. This will hopefully lead to brighter quadracyclic adenine analogues that retain the advantageous isomorphous properties of qA.

In the interest of single-molecule types of studies, there is a large potential for FBAs with an improved brightness and photostability. For that reason, it is important to understand the photodegradation of these probes. In paper IV the photostability of the already established tricyclic cytosine, tC, was therefore characterized. It was found that its photoproduct, tC[#], very likely is the sulfoxide form of tC. Interestingly, tC[#] was found to significantly destabilize DNA duplexes, causing partially denatured duplexes at room-temperature. This property could have future applications in light-triggered therapeutics or in functional DNA based nanodevices.^[131-133]

In the second part of this work the tricyclic cytosine FRET-pair tC^O-tC_{nitro}^[40] was applied to study protein-DNA interactions in mitochondrial transcription. The probes were incorporated around the transcription start site of the heavy strand promoter (HSP1). Next, changes in FRET efficiency were monitored upon addition of different components of the mitochondrial transcription machinery. Since tC^O and tC_{nitro} are rigidly incorporated inside DNA, the observed signal changes are expected to correspond to structural alterations of the duplex. In paper V, the main focus was on the mammalian transcription factor A (TFAM), whereas paper VI monitored the assembly of the whole transcription machinery. In both studies, data obtained with the FRET-pair were in agreement with findings from for example *in vitro* transcription and DNaseI footprinting. The FRET-pair is being applied at the moment to study other protein-DNA interactions including more detailed structural studies using the Matlab-based software FRETmatrix developed by Preus *et al.*^[108]. Also for these types of studies brighter and more photostable FBAs will be a great asset. Furthermore, it will be highly advantageous for the FBA FRET system, including FRETmatrix, if it can be expanded with analogues of the other natural nucleobases so that incorporation

sites are not limited to cytosines. Potentially, the new qA analogues will provide suitable adenine candidates for the task as FRET donor and/or acceptor.

Up till today, FBAs, including those discussed in this work, cannot compete with external commercial dyes regarding optical properties such as brightness. However, their main advantage is that they can be incorporated near or at the very site of interest and therefore report on subtle structural changes in the nucleic acid system, while causing minimal perturbations. This was nicely illustrated by our DNA-protein interaction studies involving the FBA FRET-pair, tC^O - tC_{nitro} .

6 Acknowledgements

I would like to express my gratitude to the following people:

-First of all, to my supervisor, Marcus Wilhelmsson, for giving me the great privilege of becoming your first PhD student. Thank you for the fantastic guidance, support and discussions the past years. I consider myself very lucky.

-My co-supervisor Bo Albinsson, for interesting discussions, valuable advice and for welcoming me into the Balb group.

-All my co-authors for interesting and exciting collaborations. Morten and Chris, thank you for the nice discussions and advice.

-All current and former group members. Francois, Moa and Blaise: Thank you for all the support and the nice atmosphere. Søren: thank you for all the help throughout the years and the friendship. It has been a joy working with all of you.

-The Beuning group at Northeastern University for a fantastic spring in Boston, especially Lisa and Philip for all the help.

-Former and current office mates for a very nice atmosphere.

-Jakob, Mélina and Tamas: thank you for friendship, great conversations and support throughout the years!

-Current and former members of the climbing team for great Monday/Tuesday evenings in and outside of fabben: Johanna, Peter, Jong-Ah, Maria A., Johan, Louise and Fabian.

-Everyone at the division of Physical chemistry, past and present, for making Sweden my second home.

-Blaise, Moa, Chris, Jakob and Basti for reading this thesis.

-Friends, in Sweden, Germany and back home in Belgium. Julie, thank you for our fantastic friendship.

-My family for unconditional love and support.

-Basti, for loving, caring and (a lot of) listening.

7 References

- [1] Weaver, R.F. (2005) *Molecular biology* 3ed. McGraw-Hill, New York, NY.
- [2] Dahm, R. (2005) Friedrich Miescher and the discovery of DNA. *Dev. Biol.*, **278**, 274-288.
- [3] Watson, J.D. and Crick, F.H.C. (1953) Molecular structure of nucleic acids - A structure for deoxyribose nucleic acid. *Nature*, **171**, 737-738.
- [4] Reese, C.B. (2005) Oligo- and poly-nucleotides: 50 years of chemical synthesis. *Org. Biomol. Chem.*, **3**, 3851-3868.
- [5] atdbio: <http://www.atdbio.com/>.
- [6] Lander, E.S., Int Human Genome Sequencing, C., Linton, L.M., Birren, B., Nusbaum, C., Zody, M.C., Baldwin, J., Devon, K., Dewar, K., Doyle, M. *et al.* (2001) Initial sequencing and analysis of the human genome. *Nature*, **409**, 860-921.
- [7] Mattick, J.S. and Makunin, I.V. (2006) Non-coding RNA. *Human Molecular Genetics*, **15**, R17-R29.
- [8] Bartel, D.P. (2009) MicroRNAs: Target Recognition and Regulatory Functions. *Cell*, **136**, 215-233.
- [9] Montange, R.K. and Batey, R.T. (2008) Riboswitches: Emerging themes in RNA structure and function. *Annu. Rev. Biophys.*, **37**, 117-133.
- [10] Serganov, A. and Patel, D.J. (2007) Ribozymes, riboswitches and beyond: regulation of gene expression without proteins. *Nat Rev Genet*, **8**, 776-790.
- [11] Seeman, N.C. (2010) Structural DNA Nanotechnology: Growing Along with Nano Letters. *Nano Lett.*, **10**, 1971-1978.
- [12] Modi, S., Bhatia, D., Simmel, F.C. and Krishnan, Y. (2010) Structural DNA Nanotechnology: From Bases to Bricks, From Structure to Function. *J. Phys. Chem. Lett.*, **1**, 1994-2005.
- [13] Pinheiro, A.V., Han, D.R., Shih, W.M. and Yan, H. (2011) Challenges and opportunities for structural DNA nanotechnology. *Nat. Nanotechnol.*, **6**, 763-772.
- [14] Holbrook, S.R. (2008) Structural principles from large RNAs. *Annu. Rev. Biophys.*, **37**, 445-464.
- [15] Foster, M.P., McElroy, C.A. and Amero, C.D. (2007) Solution NMR of large molecules and assemblies. *Biochemistry*, **46**, 331-340.
- [16] Lakowicz, J.R. (2006) *Principles of Fluorescence Spectroscopy*. 3 ed. Springer, New York, NY.
- [17] Onidas, D., Markovitsi, D., Marguet, S., Sharonov, A. and Gustavsson, T. (2002) Fluorescence properties of DNA nucleosides and nucleotides: A refined steady-state and femtosecond investigation. *J. Phys. Chem. B*, **106**, 11367-11374.
- [18] Callis, P.R. (1983) Electronic states and luminescence of nucleic-acid systems. *Annu. Rev. Phys. Chem.*, **34**, 329-357.

- [19] Serrano-Andres, L. and Merchan, M. (2009) Are the five natural DNA/RNA base monomers a good choice from natural selection? A photochemical perspective. *J. Photochem. Photobiol., C*, **10**, 21-32.
- [20] Waring, M.J. (1965) Complex formation between ethidium bromide and nucleic acids. *J. Mol. Biol.*, **13**, 269-282.
- [21] Larsen, T.A., Goodsell, D.S., Cascio, D., Grzeskowiak, K. and Dickerson, R.E. (1989) The structure of DAPI bound to DNA. *J. Biomol. Struct. Dyn.*, **7**, 477-491.
- [22] Fornander, L.H., Wu, L.S., Billeter, M., Lincoln, P. and Norden, B. (2013) Minor-Groove Binding Drugs: Where Is the Second Hoechst 33258 Molecule? *J. Phys. Chem. B*, **117**, 5820-5830.
- [23] Larsson, A., Carlsson, C., Jonsson, M. and Albinsson, B. (1994) Characterization of the binding of the fluorescent dyes YO and YOYO to DNA by polarized-light spectroscopy. *J. Am. Chem. Soc.*, **116**, 8459-8465.
- [24] Ha, T. (2001) Single-molecule fluorescence methods for the study of nucleic acids. *Curr. Opin. Struct. Biol.*, **11**, 287-292.
- [25] Iqbal, A., Arslan, S., Okumus, B., Wilson, T.J., Giraud, G., Norman, D.G., Ha, T. and Lilley, D.M.J. (2008) Orientation dependence in fluorescent energy transfer between Cy3 and Cy5 terminally attached to double-stranded nucleic acids. *Proc. Natl. Acad. Sci. U. S. A.*, **105**, 11176-11181.
- [26] Mollova, E.T. (2002) Single-molecule fluorescence of nucleic acids. *Curr. Opin. Chem. Biol.*, **6**, 823-828.
- [27] Hawkins, M.E. (2001) Fluorescent pteridine nucleoside analogs - A window on DNA interactions. *Cell Biochem. Biophys.*, **34**, 257-281.
- [28] Rist, M.J. and Marino, J.P. (2002) Fluorescent nucleotide base analogs as probes of nucleic acid structure, dynamics and interactions. *Curr. Org. Chem.*, **6**, 775-793.
- [29] Sinkeldam, R.W., Greco, N.J. and Tor, Y. (2010) Fluorescent Analogs of Biomolecular Building Blocks: Design, Properties, and Applications. *Chem. Rev.*, **110**, 2579-2619.
- [30] Wilhelmsson, L.M. (2010) Fluorescent nucleic acid base analogues. *Q. Rev. Biophys.*, **43**, 159-183.
- [31] Wilson, J.N. and Kool, E.T. (2006) Fluorescent DNA base replacements: reporters and sensors for biological systems. *Org. Biomol. Chem.*, **4**, 4265-4274.
- [32] Ward, D.C., Reich, E, Stryer, L. (1969) Fluorescence Studies of Nucleotides and Polynucleotides. *J. Biol. Chem.*, **244**, 1228-1237.
- [33] Xu, D.G., Evans, K.O. and Nordlund, T.M. (1994) Melting and premelting transitions of an oligomer measured by DNA-base fluorescence and absorption. *Biochemistry*, **33**, 9592-9599.
- [34] Sowers, L.C., Boulard, Y. and Fazakerley, G.V. (2000) Multiple structures for the 2-aminopurine-cytosine mispair. *Biochemistry*, **39**, 7613-7620.
- [35] Sowers, L.C., Fazakerley, G.V., Eritja, R., Kaplan, B.E. and Goodman, M.F. (1986) Base-pairing and mutagenesis - observation of a protonated

- base pair between 2-aminopurine and cytosine in an oligonucleotide by proton NMR. *Proc. Natl. Acad. Sci. U. S. A.*, **83**, 5434-5438.
- [36] Law, S.M., Eritja, R., Goodman, M.F. and Breslauer, K.J. (1996) Spectroscopic and calorimetric characterizations of DNA duplexes containing 2-aminopurine. *Biochemistry*, **35**, 12329-12337.
- [37] Dyrager, C., Börjesson, K., Diner, P., Elf, A., Albinsson, B., Wilhelmsson, L.M. and Grøtli, M. (2009) Synthesis and Photophysical Characterisation of Fluorescent 8-(1H-1,2,3-Triazol-4-yl)adenosine Derivatives. *Eur. J. Org. Chem.*, 1515-1521.
- [38] Sandin, P., Wilhelmsson, L.M., Lincoln, P., Powers, V.E.C., Brown, T. and Albinsson, B. (2005) Fluorescent properties of DNA base analogue tC upon incorporation into DNA - negligible influence of neighbouring bases on fluorescence quantum yield. *Nucleic Acids Res.*, **33**, 5019-5025.
- [39] Sandin, P., Börjesson, K., Li, H., Mårtensson, J., Brown, T., Wilhelmsson, L.M. and Albinsson, B. (2008) Characterization and use of an unprecedentedly bright and structurally non-perturbing fluorescent DNA base analogue. *Nucleic Acids Res.*, **36**, 157-167.
- [40] Börjesson, K., Preus, S., El-Sagheer, A.H., Brown, T., Albinsson, B. and Wilhelmsson, L.M. (2009) Nucleic Acid Base Analog FRET-Pair Facilitating Detailed Structural Measurements in Nucleic Acid Containing Systems. *J. Am. Chem. Soc.*, **131**, 4288-4293.
- [41] Engman, K.C., Sandin, P., Osborne, S., Brown, T., Billeter, M., Lincoln, P., Nordén, B., Albinsson, B. and Wilhelmsson, L.M. (2004) DNA adopts normal B-form upon incorporation of highly fluorescent DNA base analogue tC: NMR structure and UV-Vis spectroscopy characterization. *Nucleic Acids Res.*, **32**, 5087-5095.
- [42] Preus, S., Börjesson, K., Kilså, K., Albinsson, B. and Wilhelmsson, L.M. (2010) Characterization of Nucleobase Analogue FRET Acceptor tC_{nitro}. *J. Phys. Chem. B*, **114**, 1050-1056.
- [43] Avery, O.T., MacLeod, C.M. and McCarty, M. (1944) Studies on the chemical nature of the substance inducing transformation of pneumococcal types induction of transformation by a desoxyribonucleic acid fraction isolated from *Pneumococcus* type III. *J. Exp. Med.*, **79**, 137-158.
- [44] Chargaff, E. (1950) Chemical specificity of nucleic acids and mechanism of their enzymatic degradation. *Experientia*, **6**, 201-209.
- [45] Franklin, R.E. and Gosling, R.G. (1953) Molecular configuration in sodium thymonucleate. *Nature*, **171**, 740-741.
- [46] Wilkins, M.H.F., Stokes, A.R. and Wilson, H.R. (1953) Molecular structure of deoxyribose nucleic acids. *Nature*, **171**, 738-740.
- [47] Garrett, R.H. and Grisham, C.M. (2005) *Biochemistry*. 3 ed. Thomson Brooks/Cole, Belmont, CA.
- [48] Bloomfield, V.A., Crothers, D.M., Tinoco, I.Jr. (2000) *Nucleic acids: Structures, properties and functions*. University Science books, Sausalito, CA.
- [49] Stofer, E. and Lavery, R. (1994) Measuring the Geometry of DNA Grooves. *Biopolymers*, **34**, 337-346.

- [50] Choi, J. and Majima, T. (2011) Conformational changes of non-B DNA. *Chem. Soc. Rev.*, **40**, 5893-5909.
- [51] Woller, J.G., Hannestad, J.K. and Albinsson, B. (2013) Self-Assembled Nanoscale DNA-Porphyrin Complex for Artificial Light Harvesting. *J. Am. Chem. Soc.*, **135**, 2759-2768.
- [52] Unruh, J.R., Gokulrangan, G., Lushington, G.H., Johnson, C.K. and Wilson, G.S. (2005) Orientational dynamics and dye-DNA interactions in a dye-labeled DNA aptamer. *Biophys. J.*, **88**, 3455-3465.
- [53] Norman, D.G., Grainger, R.J., Uhrin, D. and Lilley, D.M.J. (2000) Location of cyanine-3 on double-stranded DNA: Importance for fluorescence resonance energy transfer studies. *Biochemistry*, **39**, 6317-6324.
- [54] Singh, M.P., Jennings, T.L. and Strouse, G.F. (2009) Tracking Spatial Disorder in an Optical Ruler by Time-Resolved NSET. *J. Phys. Chem. B*, **113**, 552-558.
- [55] Ouellet, J., Schorr, S., Iqbal, A., Wilson, T.J. and Lilley, D.M.J. (2011) Orientation of Cyanine Fluorophores Terminally Attached to DNA via Long, Flexible Tethers. *Biophys. J.*, **101**, 1148-1154.
- [56] Iqbal, A., Wang, L., Thompson, K.C., Lilley, D.M.J. and Norman, D.G. (2008) The structure of cyanine 5 terminally attached to double-stranded DNA: Implications for FRET studies. *Biochemistry*, **47**, 7857-7862.
- [57] Neubauer, H., Gaiko, N., Berger, S., Schaffer, J., Eggeling, C., Tuma, J., Verdier, L., Seidel, C.A.M., Griesinger, C. and Volkmer, A. (2007) Orientational and dynamical heterogeneity of rhodamine 6G terminally attached to a DNA helix revealed by NMR and single-molecule fluorescence spectroscopy. *J. Am. Chem. Soc.*, **129**, 12746-12755.
- [58] Okamoto, A., Saito, Y. and Saito, I. (2005) Design of base-discriminating fluorescent nucleosides. *J. Photochem. Photobiol., C*, **6**, 108-122.
- [59] Zhao, Y., Knee, J.L. and Baranger, A.M. (2008) Characterization of two adenosine analogs as fluorescence probes in RNA. *Bioorg. Chem.*, **36**, 271-277.
- [60] Driscoll, S.L., Hawkins, M.E., Balis, F.M., Pfeleiderer, W. and Laws, W.R. (1997) Fluorescence properties of a new guanosine analog incorporated into small oligonucleotides. *Biophys. J.*, **73**, 3277-3286.
- [61] Hawkins, M.E., Pfeleiderer, W., Jungmann, O. and Balis, F.M. (2001) Synthesis and fluorescence characterization of pteridine adenosine nucleoside analogs for DNA incorporation. *Anal. Biochem.*, **298**, 231-240.
- [62] Hawkins, M.E., Pfeleiderer, W., Balis, F.M., Porter, D. and Knutson, J.R. (1997) Fluorescence properties of pteridine nucleoside analogs as monomers and incorporated into oligonucleotides. *Anal. Biochem.*, **244**, 86-95.
- [63] Greco, N.J. and Tor, Y. (2005) Simple fluorescent pyrimidine analogues detect the presence of DNA abasic sites. *J. Am. Chem. Soc.*, **127**, 10784-10785.

- [64] Srivatsan, S.G., Tor Y. (2007) Fluorescent Pyrimidine Ribonucleotide: Synthesis, Enzymatic Incorporation, and Utilization. *J. Am. Chem. Soc.*, **129**, 2044-2053.
- [65] Srivatsan, S.G., Greco, N.J. and Tor, Y. (2008) A highly emissive fluorescent nucleoside that signals the activity of toxic ribosome-inactivating proteins. *Angew. Chem., Int. Ed. Engl.*, **47**, 6661-6665.
- [66] Srivatsan, S.G., Weizman, H. and Tor, Y. (2008) A highly fluorescent nucleoside analog based on thieno 3,4-d pyrimidine senses mismatched pairing. *Org. Biomol. Chem.*, **6**, 1334-1338.
- [67] Sinkeldam, R.W., Hopkins, P.A. and Tor, Y. (2012) Modified 6-Aza Uridines: Highly Emissive pH-Sensitive Fluorescent Nucleosides. *Chemphyschem*, **13**, 3350-3356.
- [68] Shin, D., Sinkeldam, R.W. and Tor, Y. (2011) Emissive RNA alphabet. *J. Am. Chem. Soc.*, **133**, 14912-14915.
- [69] Ben Gaied, N., Glasser, N., Ramalanjaona, N., Beltz, H., Wolff, P., Marquet, R., Burger, A. and Mely, Y. (2005) 8-vinyl-deoxyadenosine, an alternative fluorescent nucleoside analog to 2'-deoxyribosyl-2-aminopurine with improved properties. *Nucleic Acids Res.*, **33**, 1031-1039.
- [70] Krueger, A.T. and Kool, E.T. (2008) Fluorescence of size-expanded DNA bases: Reporting on DNA sequence and structure with an unnatural genetic set. *J. Am. Chem. Soc.*, **130**, 3989-3999.
- [71] Liu, H.B., Gao, J.M. and Kool, E.T. (2005) Helix-forming properties of size-expanded DNA, an alternative four-base genetic form. *J. Am. Chem. Soc.*, **127**, 1396-1402.
- [72] Gao, J.M., Liu, H.B. and Kool, E.T. (2004) Expanded-size bases in naturally sized DNA: Evaluation of steric effects in Watson-Crick pairing. *J. Am. Chem. Soc.*, **126**, 11826-11831.
- [73] Berry, D.A., Jung, K.Y., Wise, D.S., Sercel, A.D., Pearson, W.H., Mackie, H., Randolph, J.B. and Somers, R.L. (2004) Pyrrolo-dC and pyrrolo-C: fluorescent analogs of cytidine and 2'-deoxycytidine for the study of oligonucleotides. *Tetrahedron Lett.*, **45**, 2457-2461.
- [74] Noé, M.S., Ríos, A.C. and Tor, Y. (2012) Design, Synthesis, and Spectroscopic Properties of Extended and Fused Pyrrolo-dC and Pyrrolo-C Analogs. *Org. Lett.*, **14**, 3150-3153.
- [75] Seela, F. and Sirivolu, V.R. (2008) Pyrrolo-dC oligonucleotides bearing alkynyl side chains with terminal triple bonds: synthesis, base pairing and fluorescent dye conjugates prepared by the azide-alkyne "click" reaction. *Org. Biomol. Chem.*, **6**, 1674-1687.
- [76] Suchy, M. and Hudson, R.H.E. (2014) Pyrimidine-Fused Heterocyclic Frameworks Based on an N4-Arylcytosine Scaffold: Synthesis, Characterization, and PNA Oligomerization of the Fluorescent Cytosine Analogue 5,6-BenzopC. *J. Org. Chem.*, **79**, 3336-3347.
- [77] Wojciechowski, F. and Hudson, R.H.E. (2008) Fluorescence and hybridization properties of peptide nucleic acid containing a substituted phenylpyrrolocytosine designed to engage guanine with an additional H-bond. *J. Am. Chem. Soc.*, **130**, 12574-12575.

- [78] Rachofsky, E.L., Osman, R. and Ross, J.B.A. (2001) Probing structure and dynamics of DNA with 2-aminopurine: Effects of local environment on fluorescence. *Biochemistry*, **40**, 946-956.
- [79] Augustyn, K.E., Wojtuszewski, K., Hawkins, M.E., Knutson, J.R. and Mukerji, I. (2006) Examination of the premelting transition of DNA A-tracts using a fluorescent adenosine analogue. *Biochemistry*, **45**, 5039-5047.
- [80] Stivers, J.T. (1998) 2-aminopurine fluorescence studies of base stacking interactions at abasic sites in DNA: metal-ion and base sequence effects. *Nucleic Acids Res.*, **26**, 3837-3844.
- [81] Guest, C.R., Hochstrasser, R.A., Sowers, L.C. and Millar, D.P. (1991) Dynamics of mismatched base-pairs in DNA. *Biochemistry*, **30**, 3271-3279.
- [82] Bandwar, R.P. and Patel, S.S. (2001) Peculiar 2-aminopurine fluorescence monitors the dynamics of open complex formation by bacteriophage T7 RNA polymerase. *J. Biol. Chem.*, **276**, 14075-14082.
- [83] Deprez, E., Tauc, P., Leh, H., Mouscadet, J.F., Auclair, C., Hawkins, M.E. and Brochon, J.C. (2001) DNA binding induces dissociation of the multimeric form of HIV-1 integrase: A time-resolved fluorescence anisotropy study. *Proc. Natl. Acad. Sci. U. S. A.*, **98**, 10090-10095.
- [84] Hawkins, M.E., Pfeleiderer, W., Mazumder, A., Pommier, Y.G. and Falls, F.M. (1995) Incorporation of a fluorescent guanosine analog into oligonucleotides and its application to a real-time assay for the HIV-1 Integrase 3'-processing reaction. *Nucleic Acids Res.*, **23**, 2872-2880.
- [85] Sandin, P., Tumpane, J., Börjesson, K., Wilhelmsson, L.M., Brown, T., Nordén, B., Albinsson, B. and Lincoln, P. (2009) Thermodynamic Aspects of DNA Nanoconstruct Stability and Design. *J. Phys. Chem. C*, **113**, 5941-5946.
- [86] Lemay, J.F., Penedo, J.C., Tremblay, R., Lilley, D.M.J. and Lafontaine, D.A. (2006) Folding of the adenine riboswitch. *Chem. Biol.*, **13**, 857-868.
- [87] Hennesly, S.P. and Sanbonmatsu, K.Y. (2011) Tertiary contacts control switching of the SAM-I riboswitch. *Nucleic Acids Res.*, **39**, 2416-2431.
- [88] Wahba, A.S., Azizi, F., Deleavey, G.F., Brown, C., Robert, F., Carrier, M., Kalota, A., Gewirtz, A.M., Pelletier, J., Hudson, R.H.E. *et al.* (2011) Phenylpyrrolocytosine as an Unobtrusive Base Modification for Monitoring Activity and Cellular Trafficking of siRNA. *ACS Chemical Biology*, **6**, 912-919.
- [89] Yi-Brunozzi, H.Y., Stephens, O.M. and Beal, P.A. (2001) Conformational changes that occur during an RNA-editing adenosine deamination reaction. *J. Biol. Chem.*, **276**, 37827-37833.
- [90] Stephens, O.M., Yi-Brunozzi, H.Y. and Beal, P.A. (2000) Analysis of the RNA-editing reaction of ADAR2 with structural and fluorescent analogues of the GluR-B R/G editing site. *Biochemistry*, **39**, 12243-12251.
- [91] Gilbert, A. and Baggott, J.E. (1991) *Essentials of molecular photochemistry*. Blackwell Scientific Publications, Oxford.

- [92] Atkins, P.P.W. and De Paula, J. (2002) *Atkins' Physical Chemistry*. Oxford University Press, Incorporated, New York.
- [93] Parson, W.W. (2007) *Modern Optical Spectroscopy*. Springer-Verlag, Heidelberg.
- [94] Nordén, B., Rodger, A., Dafforn, T. (2010) *Linear Dichroism and Circular Dichroism: A textbook on polarized-light spectroscopy*. The Royal Society of Chemistry, Cambridge, UK.
- [95] Valeur, B. (2001) *Molecular Fluorescence: Principles and Applications*. Wiley-VCH, Weinheim.
- [96] Strickler, S.J. and Berg, R.A. (1962) Relationship between absorption intensity and fluorescence lifetime of molecules. *J. Chem. Phys.*, **37**, 814-822.
- [97] Fiserova, E. and Kubala, M. (2012) Mean fluorescence lifetime and its error. *J. Lumin.*, **132**, 2059-2064.
- [98] Förster, T. (1948) Zwischenmolekulare energiewanderung und fluoreszenz. *Ann. Phys.-Berlin*, **2**, 55-75.
- [99] Förster, T. (1946) Energiewanderung und Fluoreszenz. *Naturwissenschaften*, **33**, 166-175.
- [100] Dale, R.E., Eisinger, J. and Blumberg, W.E. (1979) Orientational freedom of molecular probes - orientation factor in intra-molecular energy-transfer. *Biophys. J.*, **26**, 161-193.
- [101] Stryer, L. (1978) Fluorescence energy-transfer as a spectroscopic ruler. *Annu. Rev. Biochem.*, **47**, 819-846.
- [102] Stryer, L. and Haugland, R.P. (1967) Energy transfer - a spectroscopic ruler. *Proc. Natl. Acad. Sci. U. S. A.*, **58**, 719-726.
- [103] Preus, S. and Wilhelmsson, L.M. (2012) Advances in Quantitative FRET-Based Methods for Studying Nucleic Acids. *ChemBioChem*, **13**, 1990-2001.
- [104] Wozniak, A.K., Schroder, G.F., Grubmuller, H., Seidel, C.A.M. and Oesterhelt, F. (2008) Single-molecule FRET measures bends and kinks in DNA. *Proc. Natl. Acad. Sci. U. S. A.*, **105**, 18337-18342.
- [105] Roy, R., Hohng, S. and Ha, T. (2008) A practical guide to single-molecule FRET. *Nat. Methods*, **5**, 507-516.
- [106] Lewis, F.D., Zhang, L.G. and Zuo, X.B. (2005) Orientation control of fluorescence resonance energy transfer using DNA as a helical scaffold. *J. Am. Chem. Soc.*, **127**, 10002-10003.
- [107] Kato, T., Kashida, H., Kishida, H., Yada, H., Okamoto, H. and Asanuma, H. (2013) Development of a Robust Model System of FRET using Base Surrogates Tethering Fluorophores for Strict Control of Their Position and Orientation within DNA Duplex. *J. Am. Chem. Soc.*, **135**, 741-750.
- [108] Preus, S., Kilsa, K., Miannay, F.A., Albinsson, B. and Wilhelmsson, L.M. (2013) FRETmatrix: a general methodology for the simulation and analysis of FRET in nucleic acids. *Nucleic Acids Res.*, **41**.
- [109] Tinoco, I. (1960) Hypochromism in polynucleotides. *J. Am. Chem. Soc.*, **82**, 4785-4790.
- [110] Rhodes, W. (1961) Hypochromism and other spectral properties of helical polynucleotides. *J. Am. Chem. Soc.*, **83**, 3609-3617.

- [111] Chipem, F.A.S., Mishra, A. and Krishnamoorthy, G. (2012) The role of hydrogen bonding in excited state intramolecular charge transfer. *Physical Chemistry Chemical Physics*, **14**, 8775-8790.
- [112] Lakowicz, J.R. (2006), *Principles of Fluorescence Spectroscopy*. 3 ed. Springer, New York, pp. 205-235.
- [113] Albinsson, B. (1997) Dual fluorescence from N⁶,N⁶-dimethyladenosine. *J. Am. Chem. Soc.*, **119**, 6369-6375.
- [114] Smagowicz, J. and Wierzchowski, K.L. (1974) Lowest excited states of 2-aminopurine. *J. Lumin.*, **8**, 210-232.
- [115] Nadler, A., Strohmeier, J. and Diederichsen, U. (2011) 8-Vinyl-2'-deoxyguanosine as a Fluorescent 2'-Deoxyguanosine Mimic for Investigating DNA Hybridization and Topology. *Angew. Chem.-Int. Edit.*, **50**, 5392-5396.
- [116] Catalanotti, B., Galeone, A., Gomez-Paloma, L., Mayol, L. and Pepe, A. (2000) 2'-deoxy-8-(propyn-1-yl)adenosine-containing oligonucleotides: Effects on stability of duplex and quadruples structures. *Bioorg. Med. Chem. Lett.*, **10**, 2005-2009.
- [117] Rankin, K.M., Sproviero, M., Rankin, K., Sharma, P., Wetmore, S.D. and Manderville, R.A. (2012) C 8-Heteroaryl-2'-deoxyguanosine Adducts as Conformational Fluorescent Probes in the Nar I Recognition Sequence. *The Journal of organic chemistry*, **77**, 10498-10508.
- [118] Eason, R.G., Burkhardt, D.M., Phillips, S.J., Smith, D.P. and David, S.S. (1996) Synthesis and characterization of 8-methoxy-2'-deoxyadenosine-containing oligonucleotides to probe the syn glycosidic conformation of 2'-deoxyadenosine within DNA. *Nucleic Acids Res.*, **24**, 890-897.
- [119] Millen, A.L., McLaughlin, C.K., Sun, K.M., Manderville, R.A. and Wetmore, S.D. (2008) Computational and experimental evidence for the structural preference of phenolic C-8 purine adducts. *J. Phys. Chem. A*, **112**, 3742-3753.
- [120] Tavale, S.S. and Sobell, H.M. (1970) Crystal and molecular structure of 8-bromoguanosine and 8-bromoadenosine, 2 purine nucleosides in *syn* conformation. *J. Mol. Biol.*, **48**, 109-&.
- [121] Millen, A.L., Manderville, R.A. and Wetmore, S.D. (2010) Conformational Flexibility of C8-Phenoxy-2'-deoxyguanosine Nucleotide Adducts. *J. Phys. Chem. B*, **114**, 4373-4382.
- [122] Barkley, M.D. and Zimm, B.H. (1979) Theory of twisting and bending of chain macromolecules - analysis of the fluorescence depolarization of DNA. *J. Chem. Phys.*, **70**, 2991-3007.
- [123] Seela, F. and Thomas, H. (1995) Duplex stabilization of DNA - oligonucleotides containing 7-substituted 7-deazaadenines. *Helv. Chim. Acta*, **78**, 94-108.
- [124] Seela, F. and Zulauf, M. (1998) 7-deazaadenine-DNA: Bulky 7-iodo substituents or hydrophobic 7-hexynyl chains are well accommodated in the major groove of oligonucleotide duplexes. *Chem.-Eur. J.*, **4**, 1781-1790.
- [125] Seela, F. and Zulauf, M. (1999) Synthesis of oligonucleotides containing pyrazolo 3,4-d pyrimidines: The influence of 7-substituted 8-aza-7-

- deazaadenines on the duplex structure and stability. *J. Chem. Soc., Perkin Trans. 1*, 479-488.
- [126] Somsen, O.J.G., Hoek, v.A. and Amerongen, v.H. (2005) Fluorescence quenching of 2-aminopurine in dinucleotides. *Chem. Phys. Lett.*, **402**, 61-65.
- [127] Shafirovich, V., Dourandin, A., Huang, W.D., Luneva, N.P. and Geacintov, N.E. (2000) Electron transfer at a distance induced by site-selective photoionization of 2-aminopurine in oligonucleotides and investigated by transient absorption techniques. *Phys. Chem. Chem. Phys.*, **2**, 4399-4408.
- [128] Kenfack, C.A., Burger, A. and Mely, Y. (2006) Excited-state properties and transitions of fluorescent 8-vinyl adenosine in DNA. *J. Phys. Chem. B*, **110**, 26327-26336.
- [129] Zhang, X. and Wadkins, R.M. (2009) DNA Hairpins Containing the Cytidine Analog Pyrrolo-dC: Structural, Thermodynamic, and Spectroscopic Studies. *Biophys. J.*, **96**, 1884-1891.
- [130] Singleton, S.F., Roca, A.I., Lee, A.M. and Xiao, J. (2007) Probing the structure of RecA-DNA filaments. Advantages of a fluorescent guanine analog. *Tetrahedron*, **63**, 3553-3566.
- [131] Brieke, C., Rohrbach, F., Gottschalk, A., Mayer, G. and Heckel, A. (2012) Light-Controlled Tools. *Angew. Chem.-Int. Edit.*, **51**, 8446-8476.
- [132] Wang, G.J. and Zhang, J. (2012) Photoresponsive molecular switches for biotechnology. *J. Photochem. Photobiol., C*, **13**, 299-309.
- [133] Ceo, L.M. and Koh, J.T. (2012) Photocaged DNA Provides New Levels of Transcription Control. *ChemBioChem*, **13**, 511-513.
- [134] Falkenberg, M., Larsson, N.G. and Gustafsson, C.M. (2007), *Annu. Rev. Biochem.*, Vol. 76, pp. 679-699.
- [135] Montoya, J., Christianson, T., Levens, D., Rabinowitz, M. and Attardi, G. (1982) Identification of initiation sites for heavy-strand and light-strand transcription in human mitochondrial-DNA. *Proc. Natl. Acad. Sci. U. S. A.*, **79**, 7195-7199.
- [136] Litonin, D., Sologub, M., Shi, Y.H., Savkina, M., Anikin, M., Falkenberg, M., Gustafsson, C.M. and Temiakov, D. (2010) Human mitochondrial transcription revisited: only TFAM and TFB2M are required for transcription of the mitochondrial genes in vitro. *J. Biol. Chem.*, **285**, 18129-18133.
- [137] Falkenberg, M., Gaspari, M., Rantanen, A., Trifunovic, A., Larsson, N.G. and Gustafsson, C.M. (2002) Mitochondrial transcription factors B1 and B2 activate transcription of human mtDNA. *Nat. Genet.*, **31**, 289-294.
- [138] Kanki, T., Ohgaki, K., Gaspari, M., Gustafsson, C.M., Fukuoh, A., Sasaki, N., Hamasaki, N. and Kang, D.C. (2004) Architectural role of mitochondrial transcription factor A in maintenance of human mitochondrial DNA. *Mol. Cell. Biol.*, **24**, 9823-9834.
- [139] Shutt, T.E., Lodeiro, M.F., Cotney, J., Cameron, C.E. and Shadel, G.S. (2010) Core human mitochondrial transcription apparatus is a regulated two-component system in vitro. *Proc. Natl. Acad. Sci. U. S. A.*, **107**, 12133-12138.

- [140] Fisher, R.P., Topper, J.N. and Clayton, D.A. (1987) Promoter selection in human mitochondria involves binding of a transcription factor to orientation-independent upstream regulatory elements. *Cell*, **50**, 247-258.
- [141] Gaspari, M., Falkenberg, M., Larsson, N.G. and Gustafsson, C.M. (2004) The mitochondrial RNA polymerase contributes critically to promoter specificity in mammalian cells. *EMBO J.*, **23**, 4606-4614.
- [142] Fisher, R.P., Lisowsky, T., Parisi, M.A. and Clayton, D.A. (1992) DNA wrapping and bending by a mitochondrial high mobility group-like transcriptional activator protein. *J. Biol. Chem.*, **267**, 3358-3367.
- [143] Diffley, J.F.X. and Stillman, B. (1991) A close relative of the nuclear, chromosomal high-mobility group protein HMG1 in yeast mitochondria. *Proc. Natl. Acad. Sci. U. S. A.*, **88**, 7864-7868.
- [144] Parisi, M.A., Xu, B.J. and Clayton, D.A. (1993) A human mitochondrial transcriptional activator can functionally replace a yeast mitochondrial HMG-box protein both in vivo and in vitro. *Mol. Cell. Biol.*, **13**, 1951-1961.
- [145] Kaufman, B.A., Durisic, N., Mativetsky, J.M., Costantino, S., Hancock, M.A., Grutter, P. and Shoubridge, E.A. (2007) The mitochondrial transcription factor TFAM coordinates the assembly of multiple DNA molecules into nucleoid-like structures. *Mol. Biol. Cell*, **18**, 3225-3236.