Fluorescence spectroscopy of DNA G-quadruplexes : ensemble and single-molecule measurements

Dao, Nguyen Thuan

2013

Dao, N. T. (2013). Fluorescence spectroscopy of DNA G-quadruplexes : ensemble and single-molecule measurements. Doctoral thesis, Nanyang Technological University, Singapore.

<https://hdl.handle.net/10356/51168>

<https://doi.org/10.32657/10356/51168>

NANYANG
TECHNOLOGICAL
UNIVERSITY

FLUORESCENCE SPECTROSCOPY OF
DNA G-QUADRUPLEXES:
ENSEMBLE AND SINGLE-MOLECULE
MEASUREMENTS

DAO NGUYEN THUAN

SCHOOL OF PHYSICAL AND MATHEMATICAL
SCIENCES

2013

**Fluorescence Spectroscopy of
DNA G-quadruplexes:
Ensemble and Single-molecule
measurements**

by

DAO Nguyen Thuan

Supervisor: Assoc. Prof. PHAN Anh Tuan

School of Physical and Mathematical Sciences

A thesis submitted to the Nanyang Technological University in
partial fulfillment of the requirement for the degree of Doctor of
Philosophy

2013

ACKNOWLEDGEMENTS

First of all, I would like to thank my advisor, Assoc. Prof. Phan Anh Tuan for his guidance, generous support and a lot of suggested ideas that I have received from him during my study. He is not only an excellent researcher, but also an outstanding mentor. The scientific training that I have learned, combined with the opportunity to develop my own experimental skills are what I am really grateful for. I truly appreciate the experience that I have acquired while training how to do science in Dr. Phan's Biophysics lab.

I would like to thank Prof. Maria-Elisabeth Michel-Beyerle for her continued mentorship and advice that has been invaluable, along with Prof. Alfred Huan Cheng Hon and Prof. Shen Zexiang for their support in the beginning of my study. I also would like to thank Prof. Gagik G. Gurzadyan, an expert in optical spectroscopic technique, who is always helpful and pleasant to discuss with whenever I need.

The post-docs in the Biophysics lab, including Dr. Amrane Samir and Dr. Heddi Brahim are great to work alongside with. I want to thank Brahim for a lot of meaningful and cheerful discussions that we had. I also would like to thank Dr. Reinhard Haselsberger, my collaborator from BioFemto lab for his assistance in single photon counting experiment.

I would like to thank Prof. Taekjip Ha for his support during my trip to his lab in UIUC as a visiting scientist. It is a great opportunity for me to learn smFRET technique from his lab. I'm indebted to Dr. Prakrit Vaibhav Jena, Lee Kyung Suk, Reza Vafabakhsh for their assistance and help in practicing this technique.

My fellow graduate students, including Hu Lanying, Do Ngoc Quang, Herry Martadinata, Lim Kah Wai, Vineeth Thachappilly Mukundan, Michael Adrian, Christopher Lech, Cheong Vee Vee and Chung Wan Jun are wonderful people that I'm grateful to know. I'm happy to thank them for their collaborations, advices and discussions. I also want to thank Kah Wai, who has been a very nice friend. We often had considerable amount of time to discuss together while having dinner and walking back

to the lab. Thanks Herry, Vee Vee and our former project officer Tan Soo Pei Juliet for being so helpful and efficient in purchasing DNA samples and chemicals.

I wish all the success for the post members of the Biophysics lab, including undergraduates as Yue Jia En Doris, Lim Ci Ji, Rickson Surya Winardhi, Baeg Myeong Jun, Teo Ming Hoon, Joefina Lim Kim Cheow, Jocelyn Lim Mei Wen, Li Zhe and graduate as Dr. Miao Yuanhua. My special thank to Rickson for his contribution in constructing the first version of the smFRET system in our lab, from there I could continue and complete this system.

My undergraduate professors, Prof. Nguyen Van Hieu, Prof. Nguyen Huu Duc, Assist. Prof. Pham Thu Nga and Assist. Prof. Nguyen Nang Dinh have supported and encouraged me throughout my study, which I appreciate very much.

I would like to thank Nanyang Technological University for offering me the scholarship during the four years of my study.

I'm happy to thank my parents, Dao Nguyen Thanh and Chu Thi Hoa for their love, encouragement and unconditional support in every step of my life. I thank my brother Dao Nguyen Huy for being a lifeline, just a phone call away. I also want to thank my parents-in-law, Vu Nguyen Pha and Nguyen Hong Phuong for their encouragement and their help in taking care of my wife and my son when I'm not around.

Last but not least, I love to thank my wife, Vu Thu Thuy for her love, trust, understanding, sharing and for the incredible role that she has played in my life.

Finally, thank you, Dao Nguyen Nam Duong, for making our life beautiful! We always love you, son!

ABSTRACT

Guanine-rich DNA can form structures called G-quadruplexes, which can be biologically important and also have potential applications in different fields including chemistry, nanotechnology and material science. G-quadruplexes are highly polymorphic and the activities of these molecules strongly depend on their structures. This study will focus on conformation and stability as well as excited state dynamics of G-quadruplexes using both ensemble and single-molecule fluorescence measurements, coupled with other techniques such as UV absorption, CD, and NMR spectroscopy.

It has been shown that G-quadruplex DNA has stronger intrinsic fluorescence than their less-structured counterparts. We characterize and compare fluorescence properties of various well-defined G-quadruplex structures. We show that the increase of intrinsic fluorescence of G-rich DNA sequences when they form G-quadruplexes can be used to monitor the folding and unfolding of G-quadruplexes as a function of cations and temperature. The temperature-dependent fluorescence spectra of different G-quadruplexes also exhibit characteristic patterns. We demonstrate that the stability and possibly also the structure of G-quadruplexes can be characterized and distinguished by their intrinsic fluorescence spectra.

Excited states created in DNA under UV light have recently received much attention of the scientific community in efforts to understand the photoreactions behind the genetic damage induced by UV light. We report the observation of excimer formation by stacking of two G-quadruplex blocks based on the spectroscopic signatures of this excimer: no change in absorption, red shifted fluorescence emission and lifetime decay in the order of nanoseconds. This is also the first time that fluorescence and NMR experiments are performed concurrently on identical samples, in order to maintain the high-resolution NMR structural information, when studying excimer emission. We show that for the formation of the geometry-optimized excimers, the specific base-stacking overlap at the G-quadruplex interface is critical. This specific overlap pattern has not been observed in B-DNA or G-

quadruplex monomers. We also demonstrate the ability to manipulate the excimer formation by changing DNA flanking sequence, DNA concentration, salt concentration, and temperature. This excimer emission not only could be used as a signature for detection of stacking of G-quadruplexes, but also offer a unique means for studying the excited state dynamics between G-tetrad stacking layers.

Currently, single-molecule FRET (smFRET) has become a very powerful biophysical technique by offering the ability to monitor the structural changes of individual molecules in real time. To study conformation and stability of G-quadruplex DNA at the single molecule regime, we have successfully constructed a prism-type smFRET system based on an Olympus IX71 microscope. We have combined extensively the smFRET technique and various ensemble measurements including NMR spectroscopy to study the structure, dynamics and stability of two distinct well-defined G-quadruplex folds: a parallel G-quadruplex formed by $d[(G_3T)_3G_3]$ and an antiparallel G-quadruplex formed by $d[(G_3T_4)_3G_3]$. While the $d[(G_3T)_3G_3]$ sequence forms an extremely stable parallel G-quadruplex conformation even at very low salt concentration, smFRET result of the $d[(G_3T_4)_3G_3]$ sequence reveals its structural polymorphism even at high salt concentration. In addition, smFRET data obtained from these sequences have been compared with those of the human telomeric $d[(G_3TTA)_3G_3]$ sequence in order to understand the effect of the loop length and its composition to the stability of G-quadruplexes. The result obtained from these studies could be applied in the design and prediction of G-quadruplex topologies.

Our results on the conformation and stability of G-quadruplex DNAs as well as the photophysics of their excited states will help to understand more detail of this important form of DNA. The ensemble and single molecule fluorescence techniques which have been developed and employed in our work can be applied to study different forms of nucleic acid structures and their complexes, together with their interactions with other biomolecules such as proteins or small molecules.

TABLE OF CONTENTS

Chapter 1: Introduction	- 1 -
1.1 Deoxyribonucleic acid (DNA)	- 1 -
1.2 G-quadruplexes.....	- 2 -
1.3 Ensemble and single-molecule fluorescence measurement	- 6 -
1.4 In-solution and on-surface measurement	- 8 -
1.5 Objective and scope of the work	- 9 -
 Chapter 2: Methodology	 - 11 -
2.1. DNA sample preparation	- 11 -
2.1.1. DNA synthesis.....	- 11 -
2.1.2. Deprotection, purification and dialysis.....	- 12 -
2.2. NMR spectroscopy.....	- 14 -
2.2.1. Basics of NMR spectroscopy	- 14 -
2.2.2. 1D NMR of G-quadruplexes	- 15 -
2.3. UV absorption spectroscopy	- 19 -
2.4. Circular dichorism spectroscopy	- 21 -
2.5. Steady state fluorescence spectroscopy	- 24 -
2.5.1. Principles of a steady state fluorometer.....	- 24 -
2.5.2. Right-angle vs. front-face geometry	- 25 -
2.5.3. Quantum yield measurement	- 26 -
2.6. Time correlated single photon counting (TCSPC).....	- 27 -
2.7. Forster resonance energy transfer (FRET).....	- 28 -

2.8. Single-molecule FRET (smFRET)	30 -
2.8.1. Construction of single-molecule FRET system	31 -
2.8.2. Mapping calibration	35 -

Chapter 3: Following G-quaruplex formation by its intrinsic fluorescence- 37 -

3.1. Following cation-dependent G-quadruplex formation of human telomeric (HT) sequence	39 -
3.2. Fluorescence quantum yield of HT sequence	45 -
3.3. Time-resolved fluorescence decay of human telomeric sequence. -	46 -
3.4. Following temperature-dependent G-quadruplex formation of HT sequence.....	48 -
3.5. Formation of G-quadruplex by other sequences: effect of cation and folding topology	51 -
3.5.1. Oxytricha telomeric sequence: effect of cation.....	51 -
3.5.2. Giardia telomeric sequence: effect of folding topology.....	53 -
3.6. Distinction between Z-DNA and G-quadruplex formation	55 -
3.7. Conclusion.....	56 -

Chapter 4: Excimer formation by stacking of G-quadruplex blocks..- 57 -

4.1. Structure and spectroscopy of a stacked G-quadruplex dimer	59 -
4.1.1. NMR structure	61 -
4.1.2. Steady-state absorption and fluorescence spectra.....	61 -
4.1.3. Excited-state dynamics	66 -
4.2. Correlation between the 385-nm fluorescence band and the formation of a 5'-5' stacked G-quadruplex dimer	68 -
4.2.1. T ₀ , T ₁ , and T ₂ in the presence of high and low K ⁺ concentration	70 -

4.2.2. T1 and T2: Following stepwise G-quadruplex dimer formation by parallel monitoring NMR and fluorescence spectra.....	- 72 -
4.2.3. T0, T1, and T2 at low DNA concentration	- 73 -
4.3. The role of base stacking overlap in the 385-nm fluorescence band.....	- 75 -
4.4. Assignment of the 385-nm band to excimer emission	- 81 -
4.5. Is there a structural detail responsible for the formation of geometry-optimized excimers?.....	- 83 -
4.6. Conclusion	- 85 -

Chapter 5: Conformation and dynamics of G-quadruplexes studied by single-molecule FRET..... - 87 -

5.1. Experimental procedures	- 89 -
5.2. smFRET analysis of human telomeric G-quadruplexes: comparison with previously reported works	- 91 -
5.2.1. Observation of 3 FRET values: one unfolded and two folded states.....	- 91 -
5.2.2. Single molecule FRET time traces reveal state transitions always through unfolded (U) state	- 92 -
5.3. Conformation and stability of parallel and antiparallel G-quadruplexes studied by smFRET.....	- 94 -
5.3.1. Ensemble measurements	- 95 -
5.3.2. Single molecule FRET measurements.....	- 99 -
5.4. Conclusion	- 103 -

Chapter 6: Summary and future work - 105 -

6.1. Summary	- 105 -
6.2. Perspective and Future directions.....	- 106 -

BIBLIOGRAPHY.....	- 109 -
--------------------------	----------------

APPENDICES.....	- 117 -
------------------------	----------------

A. Quartz slides cleaning protocol for smFRET measurement.....	- 117 -
---	----------------

B. PEG coating protocol for immobilization of DNA on quartz surface for smFRET measurement.....	- 119 -
--	----------------

C. Chamber fabrication protocol for smFRET measurement.....	- 120 -
--	----------------

D. Procedure of correcting FRET efficiency in single molecule time trace.....	- 121 -
--	----------------

ABBREVIATIONS

CD – Circular dichroism

Cy3, Cy5 – Cyanine dyes

DMT – Dimethyltryptamine

DNA – Deoxyribonucleic acid

DSS – 4,4-dimethyl-4-silapentane-1-sulfonic acid

EMCCD – Electron multiplying charge coupled device

FID – Free induction decay

FRET – Förster (Fluorescence) resonance energy transfer

IFE – Inner filter effect

KPi – mixture of $\text{KH}_2(\text{PO}_4)_3$ and $\text{K}_2\text{H}(\text{PO}_4)_3$

ND filter – Neutral density filter

NMR – Nuclear magnetic resonance

OD – Optical density

PBP – Pellin-Broca prism

PMT – Photomultiplier tube

ppm – parts per million

smFRET – single molecule Förster (Fluorescence) resonance energy transfer

SNR – Signal to noise

T50 – Common buffer for smFRET study, consisting of 50 mM Na^+ , 20 mM Tris (pH 8.0)

TCA – Trichloroacetic acid

TCSPC – Time-correlated single photon counting

TEAA – mixture of triethylamine and acetic acid

TIRF – Total internal reflection fluorescence

TMR – Tetramethylrhodamine

TRIS – tris(hydroxymethyl)aminomethane

UV – Ultraviolet

Chapter 1: Introduction

1.1 . Deoxyribonucleic acid (DNA)

Deoxyribonucleic acid (DNA) carries the genetic codes associated with functioning and inheritance of living organisms and viruses. DNA has the main role as a molecular genetic storage.

The revelation of the structure of the DNA double helix (Figure 1.1) by James Watson and Francis Crick in 1953 was one of the most important scientific discoveries of the twentieth century (Watson et al. 1953). The impact of their model has been extremely important, opening the foundation for modern molecular biology, and leading to an explosion of many discoveries and new techniques in biology in the last half-century. Not only did this structure satisfy the known physical and chemical properties of DNA, it also helps to understand DNA's biological functions. For example, the complementary property of the two strands of the DNA double helix provided a potential mechanism that allows it to be copied. And the bases defined along a DNA sequence suggested how genetic information could be encoded.

Figure 1.1: (A) A schematic representation of the B-DNA double helix. (B) A polynucleotide chain, showing nucleotide units that are connected by the phosphodiester bonds. (C) Base pairing in duplex DNA: Purine bases (adenine and guanine) pair with pyrimidine bases (thymine and cytosine) through hydrogen bonds (dotted line) (adapted from (Bates et al. 1993)).

1.2. G-quadruplexes

Less than a decade after Watson and Crick proposed their structure for DNA, Gellert and colleagues studied guanylic acid by X-ray diffraction (Gellert et al. 1962), revealing that the “gel-like” properties of their aqueous solution were caused by the assembly of tetrameric units into large helical structures. They figured out that four molecules of guanylic acid could form a square layer arrangement by hydrogen bonds, now referred to as a G-tetrad (Figure 1.3-A). After that, G-rich sequences found at telomeres and in the immunoglobulin chain switch region were shown to

have the potential to adopt stable structures with four-stranded topology, which now are known as G-quadruplexes (Henderson et al. 1987; Sen et al. 1988; Sundquist et al. 1989; Williamson et al. 1989).

As there are many sequences capable of forming G-quadruplexes, a question arises regarding structures of these sequences within cells. Telomeres are protective caps at the end of the eukaryotic chromosomes (Figure 1.2-A), preventing them from degradation and playing a crucial role in cell survival and proliferation. After each round of cell division, the telomeres shorten, and the cells eventually reach the limit of their replicative capacity, resulting in cell apoptosis. Since single-stranded overhang G-rich sequences are found at the end of the chromosome, which potentially can form G-quadruplex structures, there has been much interest in telomeric DNA in this regard. Human telomeric DNA consist of thousands of tandem repeats of $(TTAGGG)_n$ (Moyzis et al. 1988), with an overhang of about 200 nucleotides long at the 3' end. The telomerase is not active in human somatic cells, but is activated in 80 – 85 % cancer cells, undesirably prolonging the life of these cancer cells (Kim et al. 1994). The formation of G-quadruplex structures at the overhang of the telomere's end was shown to inhibit the telomerase activity (Figure 1.2-B) (Mergny et al. 1998a; Oganessian et al. 2006; Zahler et al. 1991).

Figure 1.2: (A) Chromosomes with telomere protective caps in red. (B) Proposed mechanism of the inhibition of telomerase activity: the pink and blue strands indicate G-rich sequences $(TTAGGG)_n$ and complementary sequences $(AATCCC)_n$ respectively (adapted from (Mergny et al. 1998a)).

More than 370,000 G-rich sequences in the human genome have been identified to have the potential to form G-quadruplex structures (Todd et al. 2005). It is noted that more than 40% of human gene promoter has putative G-quadruplex forming sequences (Huppert et al. 2007). Many G-rich sequences capable of forming G-quadruplex structures have been found within DNA and RNA, which are suspected of regulating telomere replication and mRNA translation (Bugaut et al. 2012; Zahler et al. 1991). G-quadruplex DNA has become a promising drug target for anticancer and anti-HIV therapy (Chou et al. 2005; Han et al. 2000; Jing et al. 2005; Patel et al. 2007; Pedersen et al. 2011).

Among the five nucleotides (A, T, G, C, U) in DNA and RNA, guanine (G) has the property to form the most stable complex because of its unique hydrogen bonding with neighbor guanines. G-quadruplexes mostly are stabilized by cations which stay in between G-quartet layers and occupy the central cavity. The positive charge of these cations help to neutralize the electrostatic repulsion between the phosphate backbone anions (Simonsson 2001). Generally, in the alkali series, G-quadruplex stability are favored mostly by K^+ , followed by Rb^+ , Na^+ , Cs^+ , and Li^+ (Liu et al. 2012; Shim et al. 2009). There could be two parameters that might govern the effect of stability. First, hydration energy of monovalent cations could affect G-quadruplex stability (Akhshi et al. 2012). The second parameter is the size and charge of the cations, which determine the location of the cations inside the G-quadruplex structure (Williamson 1994). For example, Na^+ ions were found to be located between 2 positions: in the cavity between two quartets and in the plane of (within) a quartet (Ida et al. 2008) while the bigger ions, K^+ , were found to be situated symmetrically between two consecutive quartets (Hud et al. 1999; Ida et al. 2008; Sket et al. 2010). The smaller ions, Li^+ , can also easily fit in 2 positions like the Na^+ ions, but its smaller size might be the reason that it is less effective in stabilizing G-quadruplex than Na^+ or K^+ ions (Bardin et al. 2008).

Figure 1.3: (A) A schematic of a G-quartet: Four guanine residues form a planar structure G-tetrad through Hoogsteen hydrogen bonding. This tetrad will be shown as a square in the following diagrams. (B-C) Guanine in (B) *anti* (magenta) and (C) *syn* (cyan) glycosidic conformations. (D-G) Four types of G-tetrad cores: (D) (4 + 0) parallel, (E) (3 + 1) antiparallel, (2+2) antiparallel G-tetrad core (F) (up–up–down–down) and (G) (up–down–up–down). (H-K) Three types of loops (colored red): (H) diagonal, (I) edgewise and (K) propeller (double-chain-reversal). Arrows indicate the strand orientations, from 5' to 3' direction. (adapted from (Phan 2010)).

In the literature, G-quadruplexes were reported to exhibit extensive structural polymorphism. Inside the G-tetrad core, the guanine itself can adopt two conformations: *anti* and *syn* (Figure 1.3-(B, C)) while the core strands can adopt four possible orientations (Figure 1.3-(D-G)). G-quadruplexes may form from one (intramolecular) or two or more (intermolecular) DNA strands. Moreover, outside the G-tetrad core, the nucleotide linkers between G-tetrad layers can adopt three different types of loop structures (Figure 1.3-(H-K)). G-quadruplex conformation depends on both the DNA sequence and the nature of the stabilizing cations. Thus it is difficult to predict which structure the sequence might fold into, and each sequence needs to be characterized extensively under different buffer conditions. For example, as mentioned before, potassium usually is the most stabilizing cations and tends to favor parallel conformations (Miura et al. 1995a). But, antiparallel G-quadruplex

stabilized by potassium and parallel G-quadruplexes stabilized by sodium are also found to be quite stable (Balagurumoorthy et al. 1994; Hu et al. 2009; Phan et al. 2003).

In recent decades of G-quadruplex studies, several ensemble techniques have been applied to monitor its conformation. These include nuclear magnetic resonance (NMR) (Smith et al. 1992), X-ray crystallography (Kang et al. 1992), circular dichroism (CD) (Hardin et al. 1991), Raman spectroscopy (Miura et al. 1995b), calorimetry (Jin et al. 1992) and ultraviolet (UV) absorption (Mergny et al. 1998b).

1.3. Ensemble and single-molecule fluorescence measurement

As we measure the properties of DNA molecules by ensemble spectroscopic techniques, such as fluorescence, UV, CD or NMR, we are likely to be measuring the average characteristics of a very large ensemble of molecules. These ensemble techniques have many advantages as: ease of operating, widely use, commercial availability and help us obtain lots of information in molecules' properties. Typically, fluorescence, UV or CD measurement often require at least 0.6 – 6 nmol of DNA molecules (1 – 10 μ M in 600 μ l cuvette), while NMR measurement requires around 0.05 – 0.5 μ mol of DNA molecules (0.1 – 1 mM in 500 μ l NMR tube). That's indeed a huge number of molecules ($\sim 10^{17}$) if we multiply by the Avogadro number.

So in most ensemble techniques, any measurements we make are the average across many millions or billions of molecules. However in most of the cases, especially with complex biological molecules, some of these molecules (although might be the minor population), can exhibit different properties over time, and moreover in a random (or stochastic) manner. In ensemble measurement, it is often very difficult to synchronize the change in conformation of biological molecules and mostly impossible to detect the short-lived conformation or the extremely minor species (Figure 1.4-a). Single molecule techniques, for example, single-molecule FRET (smFRET) offers an ability of monitoring the structural changes of individual molecules in real time (Figure 1.4-(b, c)). This technique allows not only the detection of the existence of several conformations over time (even the extremely minor one), but also the identification of rarely visited and short-lived state (Lee et al. 2005).

Figure 1.4: Illustration of single-molecule FRET study. (a) Ensemble FRET measurement represents the average value from many molecules with different states (colors and shapes). (b) FRET signal of individual molecule reveals its different states in real time. (c) Postsynchronizing of studied events during data analysis. (imported from (Selvin et al. 2007)).

smFRET also allows the observation of the relative population of different conformations as a function of time to monitor the progresses of a reaction such as reaction rates and pathways. Following the first demonstration of using smFRET in studying human telomeric G-quadruplex (Ying et al. 2003), there have been numerous studies of G-quadruplex structures (Lee et al. 2005; Lee et al. 2008; Lee et al. 2009; Okamoto et al. 2008) as well as interaction of G-quadruplexes with other DNA sequences (Fegan et al. 2010) or small molecules (Jena et al. 2009), that employed smFRET technique. All studies used the same designed molecular system established by Ying et al. (Ying et al. 2003), which was constructed from the hybridization of 2 strands (Figure 1.5) : the G-quadruplex strand includes a Cy5 at the 5'-end, a G-quadruplex forming sequence of interest, a stem part at the 3'-end; and the complementary strand includes a complementary stem part with a TMR (or Cy3) in the middle. It is possible that the Cy5 labeling at the 5'-end or the duplex stem may interfere with the formation of G-quadruplexes. However, Ying et al. (Ying et al. 2003) showed that dye modification and the duplex stem do not affect the formation of G-quadruplexes by comparing data of smFRET with UV melting and CD spectroscopy of modified and non-modified sequences.

Figure 1.5: Schematic of the G-quadruplex-duplex forming complex used in smFRET study by employing confocal microscope. Figure imported from (Ying et al. 2003).

1.4. In-solution and on-surface measurement

The detection and characterization of biomolecules are often performed when these molecules are in solution or attached on a solid surface. Most of the spectroscopic ensemble techniques, e.g. fluorescence, UV, CD, and NMR are often performed in solution (Deniz et al. 1999; Schuler et al. 2002). Single-molecule FRET technique by TJ Ha (Selvin et al. 2007) originally proposed to study properties of molecules attached on the slide surface (Figure 1.6-(a, b)). DNA hybridization on chips is also often detected for DNA on solid surface. As compared to measurements on surface, measurements in solution have the advantage that it can more closely mimic the condition of molecules in vivo therefore the obtained properties can be expected to be similar to in vivo condition.

One of the promising approaches for observing single molecules over time without immobilizing them on the surface is to trap these molecules within small vesicles (diameter 50–200 nm) that are anchored on the surface (as an example in smFRET measurement – Figure 1.6-(c, d)). If we assume the interactions of these molecules with the membrane are minimal, then they can be considered to move “freely” inside the membrane. Therefore, this scheme provides a closer condition which could mimic the biological environment. Moreover, the solution inside the membrane can easily be exchanged over time, which allows one to study the dynamics of these molecules in different environments. By using this approach, the equilibrium of folding vs. unfolding of single-protein molecules can be measured in extended period, which has been demonstrated by Rhoades et al. (Rhoades et al. 2003).

Figure 1.6: An example of surface and solution measurement for smFRET: (a) BSA proteins (or DNAs) tandem bound nonspecifically to the surface by neutravidin-biotin interaction. (b) Mixture of biotin-PEG and m-PEG is attached to aminosilanized slide surface. c) Ni²⁺ or Cu²⁺ is designed to be carried on PEG-slide's surface, which can bind efficiently to 6His-tagged proteins while not interfere with their activities. (d) Vesicles permeable to small molecules can be selected trap biomolecules by using dimyristoyl phosphatidylcholine (DMPC). Specific binding of biotinylated lipid to PEG allowed specific tethering of the vesicles on the quartz's surface (imported from (Roy et al. 2008)).

1.5. Objective and scope of the work

This research aims to develop and apply both ensemble and single-molecule fluorescence measurements to study the structure, stability and photophysical properties of G-quadruplex DNA in order to provide a better understanding about this important form of DNA. Besides fluorescence methods, other biophysical techniques such as UV-VIS (ultra violet-visible) absorption, CD (circular dichroism), NMR (nuclear magnetic resonance) were also employed to provide a comprehensive study on G-quadruplexes DNA. The developed methodology could also be applied to

different forms of nucleic acid structures and their complexes, together with their interaction with other biomolecules such as protein or small molecules.

Chapter 2: Methodology

2.1. DNA sample preparation

2.1.1. DNA synthesis

DNA oligonucleotides were synthesized on an ABI 394 DNA/RNA synthesizer using chemicals from Glen Research and Cambridge Isotope Laboratories. Generally, the synthesis processes are included of four main steps: detritylation, coupling, capping, and stabilization.

Detritylation: Before the synthesis process start, the respective columns containing the last nucleotide at the 3' end of the sequences were selected and placed into the synthesizer (the synthesizer will synthesize DNA sequences from 3' end to 5' end). In each cycle of base addition, the columns are washed with acetonitrile to remove traces of the preceding agent, following by a process known as detritylation. TCA (Trichloroacetic acid) is delivered to the columns to remove the DMT (Dimethoxytrityl) group that is attached to the 5' end (Figure 2.1), leaving a hydroxyl to react with an incoming phosphoramidite in the next step. Because the cleaved DMT is relatively stable in acid solution and has bright orange color, the presence of this color indicates that the synthesis is running smoothly.

Coupling: Following the removal of DMT, the support is washed extensively by flowing excess acetonitrile to remove residual TCA, which will compete with the 5' hydroxyls of the activated phosphoramidite and might decrease coupling efficiency. The column is then reversely flushed with argon to remove residual acetonitrile. In the following step, the next phosphoramidite and tetrazole (activator - a weak acid) are delivered to the columns. Upon mixing, they react to form a tetrazolyl phosphoramidite intermediate. Then through the reaction with the hydroxyl group of the nucleoside, the 5' to 3' linkage is formed. It should be noted that this coupling is not 100% efficient.

Figure 2.1: Protection groups of a nucleoside (picture adapted from IDT's protocol for chemical synthesis of oligonucleotides).

Capping: If there is a reactive hydroxyl group retained in the 5' end, it would be able to couple in the next round and could cause a coupling failure. To remove the coupling failures, an acetylating reagent is added to the columns. This reagent reacts specifically with free hydroxyl groups to conceal the oligonucleotide which has failed coupling reaction. Again, the excess reagents are removed by reversely flushed argon.

Stabilization: After the failure sequences are concealed, the final step is stabilizing the phosphate linkage between the recently added base with the whole oligonucleotide chain. As iodine and water are added to the column, the oxidation reaction of phosphite into phosphate occurred, which results in stabilization of the phospho-triester bond. After the column are reversely flushed with argon and washed with acetonitrile, a cycle of nucleotide addition is completed.

These 4 steps above are repeated until the 1st nucleotide at the 5' end of the desire sequence is attached.

2.1.2. Deprotection, purification and dialysis

Deprotection: If the DMT option in the synthesizer is active, a DMT group is attached only to the final product. The final products contained both our desired and failure sequences of base-protected oligonucleotides are collected in the synthesized column. They are deprotected by base hydrolysis using ammonium hydroxide at high temperature (55 °C) for 12 to 16 hours or room temperature for

longer time (36 to 48 hours). The ammonium hydroxide will cleave the oligonucleotides from the resin and remove the protection groups (Figure 2.1).

Purification: After evaporation of ammonium hydroxide, DNA samples were purified by Poly-Pak™ II cartridges following Glen Research protocol. Basically, the cartridge is washed with 4 ml of 100% Acetonitrile, followed by 4 ml of 2M TEAA (mixture of triethylamine and acetic acid) to enhance the binding of DMT with the resin inside the cartridge. The DNA samples are loaded inside the cartridge (3 times loading to maximize the chance of binding). 6 ml of ammonium hydroxide (1:20) is flushed through the cartridge to remove the remained failure sequences (due to their weaker binding with the resin as compare to the DMT group of the desired sequence). 4 ml of de-ionized water is used to clean the excess ammonium hydroxide. 4 ml of TFA (Trifluoroacetic acid) (2%) is flushed to the cartridge to cut the DMT group, followed by 4 ml of de-ionized water to wash out the excess TFA. The purified DNA is collected by slowly flushing the cartridge with 4 ml of 20% Acetonitrile. At the end of this step, the cartridge can be washed with 4 ml of 100% Acetonitrile and can be reused.

Dialysis: After purification, the DNA samples in 4 ml of 20% acetonitrile are loaded into a membrane bag (pore size is 500 or 1000 daltons). The samples are dialyzed against water for 2 hours (to remove acetonitrile), then against solution containing respective salt (Li^+ , K^+ , Na^+) for 8 hours or overnight (to remove any contaminations and slowly let the cations come inside the membrane), and finally against water for 2 hours (to remove excess cations). In some cases, the DNA samples are dialyzed against solution containing 10 mM Tris-HCl, pH 7.0 to desalt and minimize any presence of alkaline cations inside the membrane.

After dialysis, the solution is collected, freeze-dried and lyophilized. The final products are the purified DNA sequences. Appropriate buffer will be added to the DNA samples and a small percentage (1%) will be used to quantify the amount of DNAs. The common NMR buffer used in this case is 90 mM K^+ buffer, pH 7.0 which composed of 90 mM KCl and 5 mM Tris-HCl, pH 7.0.

2.2. NMR spectroscopy

2.2.1. Basics of NMR spectroscopy

Nuclear magnetic resonance (NMR) is a phenomenon in which magnetic nuclei under a magnetic field absorb and emit electromagnetic signal with a resonant frequency. Not all the nuclei experience this phenomenon, depending upon whether they have spin or not. NMR active nuclei, which have a spin number of ($s = \frac{1}{2}$) such as ^1H and ^{13}C are the most commonly studied isotopes in NMR. In the presence of an external magnetic field, these nuclei would be split into two spin states ($s = 2n + 1$) into ($m_s = +\frac{1}{2}$) and ($m_s = -\frac{1}{2}$). These two states have different energies, with ($m_s = +\frac{1}{2}$) being aligned with the magnetic field and ($m_s = -\frac{1}{2}$) being in the opposite direction. After excitation by a radio frequency pulse at 90° , the net magnetization of the nucleus rotates into a direction which is perpendicular to the external magnetic field (Figure 2.2-(a, b)), and the electron spin will precess around the magnetic field with a frequency called Larmor frequency.

During the rotation of the net magnetization, there is a precessing transverse magnetization oscillates at a well-defined frequency. The rotating magnetic moment generates a rotating magnetic field, which then induces an electric current in a probed wire coil placed near the sample. The oscillating electric current induced by the precessing nuclear transverse magnetization is the NMR signal called free-induction decay (FID) (Figure 2.2-c). Fourier transform from time domain to a frequency domain of this FID signal results in an NMR spectrum (Figure 2.2-d). The frequencies at which NMR absorptions occur are calculated by

$$\nu = \gamma B \quad (\text{Eq.2.1})$$

with B is the strength of the magnetic field and γ is the gyromagnetic ratio of the nucleus. For convenience, the unit in x-axis of typical NMR spectrum will be changed to chemical shift (Figure 2.2-e), which is related to the frequency of a referenced compound (DSS, in our case). Typically, the chemical shift is rather small so it is multiplied by 10^6 and has a unit displayed in parts per million, or *ppm*.

Figure 2.2: Schematic diagram highlighting fundamentals of NMR spectroscopy. (a) Precession motion of a nucleus under magnetic field. (b) Induction of an NMR signal. (c) FID signal of oscillating electric current as a function of time. After Fourier transform, it is plotted as a function of frequency (d), but it is more convenient to use the chemical shift scale (e) in which frequencies are expressed relative to that of a reference compound. (adapted figures from (Keeler 2005; Malcolm 2008)).

2.2.2. 1D NMR of G-quadruplexes

In ^1H NMR experiment, different protons will exhibit different chemical shifts due to several factors, including inductive effect, solvent effect and anisotropy effect (Keeler 2005). Figure 2.3 shows the trend of chemicals shift for different protons of nucleic acids. The intensity of the NMR peaks usually reflects the number of protons in the spectrum.

Figure 2.3: (a) Trend of the chemical shifts for protons of nucleic acids. (b) Schematic view of guanine nucleotide monophosphate and other heterogenous bases (c, d, e). Numbering of the bases is shown in pink.

NMR experiments were recorded on either 600 or 700 Mhz Bruker spectrometer at 25 °C. The DNA concentrations of NMR samples were 0.1 – 2 mM and usually were suspended in 500 µl solution of a common buffer, containing 70 mM KCl and 20 mM KPi (pH 7.0). The NMR peak of D₂O was used as a lock signal, therefore a small amount of D₂O often added to the solution (10% in volume). Water has a concentration of approximately 55 M ([1000 g/liter] * [1 mol/ 18g] = [1000/18] mol/liter ~ 55 M), which is around 100.000 times more as compared with the DNA concentration. Therefore, the signal of water will be extremely large, overwhelm the DNA's signals, causing a dynamic range problem. Thus, one needs to suppress the water signal in order to optimize the signal-to-noise (SNR) ratio. Here, a pulse sequence named Jump and Return (jr) (Plateau et al. 1982) was employed to suppress the water peak at 4.70 ppm.

Figure 2.4: (A) Imino proton NMR spectra of the human telomeric $d[TT(G_3TTA)_3G_3A]$ sequence in different ion conditions. (B) Schematic structure of the G-quadruplex formation by the sequence. Loops are colored red; anti and syn guanines are colored cyan and magenta, respectively (imported from (Dao et al. 2011)).

Guanine imino protons in a G-quartet exhibit their characteristic chemical shifts within the range of 10–12 ppm (Feigon et al. 1992), as compared to the range of 13–14 ppm for those in Watson–Crick duplex base pairing (Patel et al. 1974) (Figure 2.3-a). For instance, the twelve sharp guanine imino proton peaks at 10–12 ppm (Figure 2.4-A, blue line) indicate the formation of a three-layered G-quadruplex with twelve guanines taking part in the G-tetrad core (Figure 2.4-B) as the DNA sequence $d[TT(G_3TTA)_3G_3A]$ (human telomeric) is in a buffer solution containing K^+ ions (Dao et al. 2011; Luu et al. 2006). The exactly same DNA sequence in Li^+ buffer does not show any peaks in 10–12 ppm range (Figure 2.4-A, green line), indicating there is no (or very little) G-quadruplex formation in this condition. Moreover, it is also possible to predict the population of G-quadruplex conformation from 1D NMR spectrum. For example, the same ratio of the peak's intensities (or explicitly, the integrated areas of the peaks) in Figure 2.4 suggests there might be only one intramolecular G-quadruplex conformation formed, while some minor peaks in 10–12 ppm range indicate the presence of a minor conformation.

Figure 2.5: (A) Fully assigned imino proton NMR spectrum of the *Tetrahymena* telomeric d[TG₄T₂G₄T] sequence in Na⁺ solution. (B) and (C) Folding schematic of the corresponding asymmetric dimeric G-quadruplex. Four different strand conformations are colored magenta, cyan, green and yellow, respectively (imported from (Adrian et al. 2012)).

Figure 2.5 gives an example of the power of NMR spectroscopy in investigating multiple G-quadruplex conformations formed by the *Tetrahymena* telomeric d[TG₄T₂G₄T] sequence in Na⁺ solution, where four different conformations coexist (Phan et al. 2004). Well-resolved proton spectra (Figure 2.5-A) allowed simultaneous folding determination and kinetics characterization of two inter-converting asymmetric dimeric G-quadruplexes in solution (Figure 2.5-(B, C)). However, in most of the cases where there are more than one conformation presented in the sample, the spectrum is usually heavily overlapped. This problem makes the structural analysis at high resolution very difficult and usually more complicated experiments (such as 2D and 3D NMR) are needed in order to solve the structure. By the addition of extra dimensions in 2D and 3D NMR, the overlapped spectrum can be resolved and the analysis of these complicated spectra is made possible.

2.3. UV absorption spectroscopy

Nucleic acids absorb in the ultraviolet region of the spectrum due to the aromatic ring systems of the constituent purines and pyrimidines (Fleck et al. 1962). UV absorption spectra were measured at 20 °C using either a quartz cuvette with 1 mm path length for high DNA concentration (~0.2 mM) or a quartz cuvette with 1 cm path length for low DNA concentration (~4 µM). Two UV dual-beam spectrophotometers were used: a Varian CARY-300 and a Jasco V-660. The CARY-300 spectrophotometer can provide a linear working range up 4.0 optical density (OD). Basically, a beam from a light source (deuterium: 190 – 350 nm or halogen: 330 – 900 nm) is selected by adjustable slits and monochromators and directed to a beam splitter. The beams after that will pass through a sample and a reference simultaneously and directed to a photo multiply tube (PMT) detector. A chopper placed before the detector allow the sample and reference beams reach the detector at the same point, result in eliminating any errors due to non-uniformity of the detector (Figure 2.6).

The absorbance (also called “optical density” – OD) of the sample was measured by Equation 2.2, where $A_{(\lambda)}$ is the absorbance at particular wavelength of measurement, I_0 and I is the intensity of the beam before and after passing through the sample, respectively. In Beer–Lambert law, the absorbance is linear with extinction coefficient of the sample (ϵ), sample concentration (c) and pathlength of the cuvette (l). DNA concentration (c) was calculated using nearest-neighbour approximation for the absorption coefficients of the unfolded sequences (Cantor et al. 1970).

$$A_{\lambda} = \log_{10} \left(\frac{I_0}{I} \right) = \epsilon cl \quad (\text{Eq.2.2})$$

Figure 2.6: Schematic and operation of a double-beam spectrometer UV-VIS machine. (Figure adapted from Varian's CARY-300 brochure).

Melting experiments monitored the absorption as a function of temperature at 295 nm for G-quadruplex forming sequences (Mergny et al. 1998b) or at 260nm for Watson-Crick duplex forming sequences. Samples were covered with approximately 100 μ l of paraffin oil to prevent water evaporation. They were equilibrated at 90 $^{\circ}$ C for 10 minutes, then cooled to 20 $^{\circ}$ C and heated to 90 $^{\circ}$ C twice consecutively at a rate of 0.2 $^{\circ}$ C per minute. Data were collected every 1.2 $^{\circ}$ C during both cooling and heating processes.

Figure 2.7: (A) Absorbance spectra of $d[AG_3(TTAG_3)_3]$ oligonucleotide at 90 °C (full line) and 1 °C (dotted line). The square black dot curve is the different absorbance of the high vs. low temperature spectra (right axis). (B) Melting curves obtained at pH 7.0 for this sequence at two different wavelengths: 260 nm (open circles) and 295 nm (black triangles). (adapted from (Mergny et al. 1998b)).

Figure 2.7 gives an example of absorbance measured at 295 nm, which allows to precisely monitor G-quadruplex formation and dissociation by varying temperature. The different absorption curve of the high (90 °C) vs. low (1 °C) temperature spectra, observed upon G-quadruplex formation by $d[AG_3(TTAG_3)_3]$ sequence, represents a maximal hyperchromism at 295 nm, as shown in Figure 2.7-A. A clean sigmoidal melting curve was obtained at 295 nm, after normalization resulted in a melting temperature ~ 55 °C (Figure 2.7-B). The heating and cooling curves are super imposable, indicating equilibrium processes during the melting experiment. The simplicity and sensitivity (only 1 – 5 μ M DNA is required) of this technique make it very useful and widely used in all cases where any G-quadruplex formations are suspected.

2.4. Circular dichorism spectroscopy

Circular dichroism (CD) spectroscopy measures the different absorbance of left-handed polarized versus right-handed polarized light from chiral samples.

Figure 2.8: Principle of operation of a Circular Dichroism spectrophotometer (figure adapted from <http://www.photophysics.com>).

Most circular dichroism machines operate on the same principles, which are demonstrated in Figure 2.8. A beam from a source of white unpolarised light passes through a set of monochromator and linear polariser to generate a monochromatic linearly polarised beam. Instead of a quarter-wave plate, a photo-elastic modulator (PEM) converts this linear polarized beam into either left- or right-circularly polarised (LCP or RCP) light alternating at 50 KHz frequency. When there is no CD active sample, the LCP and RCP have equal light intensities, thus there is only a steady output from the detector. When an active CD sample is placed in, it will absorb LCP or RCP light preferentially. The result of unequal amount of LCP and RCP light reaching the detector generate a signal that varies with the frequency of the PEM (50 KHz) which is measured by a lock-in amplifier as the vAC signal. The average of light intensity overtime is measured, and is termed vDC. The vAC signal is then divided by vDC signal and multiplies by a calibration factor (G) to give units of milidegrees or Δ absorbance. As the monochromator scans the wavelength range, a CD spectrum is obtained.

In this work, CD spectra at 20 °C were recorded on a Jasco J-815 spectrometer using a standard quartz cuvette with 1-cm path-length. Scans from 220 to 320 nm were acquired with scanning speed of 200 or 500 nm/min, step of 1 nm, and bandwidth of 1 nm. For each spectrum, three continuous scans were acquired and averaged, then were subtracted with spectral contribution from the buffer. The data were also zero-corrected by subtracting the CD signal at 320 nm. The DNA concentration was 5 – 20 μ M in 500 μ l reaction volume.

Figure 2.9: Example of CD signature spectra from different DNA conformations in a common buffer contained 90mM K^+ , pH 7.0.

In protein; alpha-helix, beta-sheet, and random coil structures have their defined CD spectra. While in DNA, not only between duplex and G-quadruplex but also within different G-quadruplex conformations, there are signature bands that characterize for these structures. The G-quartets are piled on top each other and rotated with respect to the adjacent one differently in different G-quadruplex conformations. This rotation causes different couplings between transition dipole moments located in near neighbour guanines, which originates CD signal of the G-quadruplex (Clark 1994). The CD spectrum from 220 nm to 320 nm region is usually measured to characterize for the conformation of DNA (Randazzo et al. 2012). Figure 2.9 gives an example of CD spectra from different DNA conformations formed by different sequences: B-DNA duplex (black curve) formed by d[(GC)₆]; parallel (4+0) G-quadruplex (G4) formed by d[TG_4T]; anti-parallel (2+2) G4 formed by d[$G_4T_4G_4$] and anti-parallel (3+1) G4 formed by d[$TT(G_3TTA)_3G_3A$]. The formation and unfolding of different G-quadruplex conformations also can be monitored with the CD signal of the characteristic peaks, for example, at 290 nm for anti-parallel G4 (3+1) G-quadruplex (Kypr et al. 2009). CD spectroscopy does not provide specific structure information at atomic level such as X-ray and NMR spectroscopy, but it is a quick method that does not require large amounts of sample (5 – 20 μ M) or extensive data processing. Therefore, CD spectroscopy is widely used in all basic conformational findings on DNA.

2.5. Steady state fluorescence spectroscopy

2.5.1. Principles of a steady state fluorometer

Fluorescence is the radiative emission from molecules as their electrons move from excited states to ground states. The fluorescence phenomenon usually is governed by three important processes, which happen in different timescales. These processes associated with transitions between electronic states of molecules and usually being illustrated by a Jablonski diagram (Figure 2.10-A). **Excitation** of most of the fluorophore by incoming photons happens in femtoseconds (10^{-15} s), while **relaxation** of excited electrons to the lowest energy level in the excited state by vibrational relaxation is much slower, at picoseconds (10^{-12} s) regime. The final process, **emission** of longer wavelength photons (red-shifted effect between absorption and emission peaks) as the excited electrons return to ground state, occurs in the period of nanoseconds (10^{-9} s).

Figure 2.10: (A) A simple Jablonski diagram showing transitions between molecular electronic energy levels. (B) A schematic working principles of a fluorometer.

A typical fluorometer contains an excitation source, monochromators and a fluorescence detector (Figure 2.10-B). The excitation source is usually a xenon or deuterium lamp which has a broadband and strong UV emission. Single wavelength (or explicitly, a very narrow bandwidth) beam is selected from this broad band excitation light by an excitation monochromator. Fluorescence from the sample is collected in different geometry and dispersed by emission monochromator before coming to the detector, which usually is a photomultiplier tube (PMT). Scanning the excitation monochromator (while fixing the emission wavelength) gives the excitation

spectrum and scanning the fluorescence monochromator (while fixing the excitation wavelength) gives the fluorescence spectrum.

2.5.2. Right-angle vs. front-face geometry

The fluorescence characteristic of a sample is usually collected by employing either one of the two geometries. In the common right-angle geometry (Figure 2.11-A), which is used widely for solution samples, a beam is directed to one face of a square cuvette to excite the sample, and its emission is collected through an adjacent face. However, with a high absorbing solution, the excitation beam is absorbed mostly near the front surface, with little light can enter the center of the cuvette, and thus, the detector receives very little or no light. This is an “inner filter effect” (IFE). In the second geometry, called front-face geometry (Figure 2.11-B), which is used mainly for solid samples, the excitation and emission light is directed and collected at the same cuvette surface.

Figure 2.11: Two different geometry settings for fluorescence measurements: (A) right-angle; (B) front-face geometry.

In right-angle geometry, when the optical density is more than 0.05 at both the excitation and the fluorescence wavelength, OD_{exc} and OD_{em} , both the excitation and the emission intensities are attenuated by $10^{-0.5OD_{exc}}$ and $10^{-0.5OD_{em}}$. Therefore, the corrected fluorescence intensity is given approximately (Lakowicz 2006) by

$$F_{corr} = F_{obs} \times 10^{(OD_{exc} + OD_{em})/2} \quad (\text{Eq.2.3})$$

accounting for both, the inner filter effect and the fluorescence re-absorption.

Inner filter effects can be eliminated by using front-face geometry. Optimally, in this geometry, the incident light enters with an angle of 30° or 60° respect to the cell's surface, depending on the orientation of the cuvette holder (Eisinger et al. 1979). Front-face geometry can provide several advantages in measuring the fluorescence of any DNA / protein solution with a strong absorbing property (mM as compared to μ M in right-angle geometry). Unlike right-angle optics, there are only certain molecules of DNAs/proteins near the surface are excited in front-face geometry. Therefore, no correction is needed for inner filter effect (IFE) and fluorescence re-absorption in this case.

In this work, fluorescence emission spectra were recorded on a Jobin-Yvon-Spex Fluorolog 3-11 fluorometer. All spectra are scanned with an integration of 0.2 s, an excitation and emission slit width of 4 nm, and step size of 1 nm. Three continuous scans are acquired and averaged to give an emission spectrum in all cases. The instrument-specific photomultiplier tube correction file supplied by the manufacture was applied to all spectra. The DNA concentration was the same as in UV absorption spectroscopy experiments, unless otherwise stated. Since the fluorescence measurements were performed on samples with a large optical density (OD~0.5 for right-angle geometry), the resulting spectra have to be corrected for emission losses caused by the inner filter effect (IFE). Temperature-dependent fluorescence spectra were acquired either on the Fluorolog 3-11 or on PTI Quanta Master 4 using the same experimental conditions. Temperature control was achieved either by an external circulating water bath NESLAB RTE attached to the sample compartment (PTI Quanta Master 4) or by a Peltier device (Fluorolog 3-11). Samples were covered with approximately 100 μ l of paraffin oil to prevent water evaporation. The heating and cooling rates were ~2 °C per minute.

2.5.3. Quantum yield measurement

The quantum yield of DNA sequences in Tris-HCl buffer (pH 6.8) and in the presence of K⁺ was measured by using 2'-deoxyguanosine 5'-monophosphate (dGMP) as the reference (Changenet-Barret et al. 2010; Lakowicz 2006). DNA samples of various concentrations (OD = 0.01 – 0.1) were used for measurements. The solvent was either deionized water containing 5 mM Tris-HCl buffer only or 5 mM Tris-HCl buffer supplemented with 5 mM KCl. Absorption and fluorescence spectra were recorded using a 1-cm-pathlength cuvette. For fluorescence measurements,

samples were excited at 265 nm to match the condition reported in (Changenet-Barret et al. 2010). The background spectra of the solvent were subtracted from the respective spectra of the samples. The quantum yield was determined using the equation $Q = Q_R (\alpha / \alpha_R)$, where Q and Q_R are the quantum yield of the sample and reference, respectively; α and α_R are the slopes of the integrated fluorescence intensity-vs-absorbance for the sample and reference, respectively. For simplicity, we considered the solvent refractive index for the sample (5 mM Tris-HCl buffer) and for the reference (deionized water) to be similar.

2.6. Time correlated single photon counting (TCSPC)

Figure 2.12: Principle of TCSPC. The middle graph represent the pulse output from a discriminator (imported from (Lakowicz 2006)).

The basic principle of TCSPC is shown in Figure 2.12. The sample is excited with laser pulses (green curve), resulting in a final decay profile (red curve), which is expected after many excitation pulses and numerous photons emission, as shown at the top of the figure. At a certain delay time, the detector collects number of photons emitted from the sample. The time difference between the excitation pulse and the observed photon is measured by a discriminator and stored in a histogram. The middle graph in Figure 2.12 shows the number of detected photons as a function of

the time difference. The signal is then accumulated over time to construct the waveform decay from the sample (the bottom graph in Figure 2.12). It is important to note that, for TCSPC, the excitation beam is adjusted so that there is typically 1 photon per more than hundred excitation pulses.

Time-resolved fluorescence was measured with a time-correlated single photon counting (TCSPC) set-up from PicoQuant. The output of a Titan:Sapphire Laser (780 nm - 1000 nm, 80 MHz, 100 fs) was frequency tripled (THG) to obtain a 260-nm excitation and focused on the sample (peak intensity, 2 kWcm⁻²). A portion of the excitation light was used as the start signal for the measurement cycle controlled by histogram accumulating real-time processor (PicoHarp 300). The induced fluorescence was collected by a time-resolved spectrometer FluoTime 200 with wavelength resolution of 1 nm or better. The signal was recorded by a Multi-Channel-Plate Photomultiplier Tube (MCP-PMT) with an overall IRF (Instrument Response Function) FWHM of 40 ps. Both the start signal and the fluorescence signal (stop signal) are triggered by a programmable CFD (Constant Fraction Discriminator) and a TDC (Time-to-Digit Converter). The PicoHarp 300 system records the arrival times of individual photons in a histogram which measures the different times between laser excitation pulse and the corresponding fluorescence photon arrival (start-stop-times). The fluorescence decays were fitted using the convolution of the IRF and the multiexponential decay function with lifetimes τ_i with $i \leq 3$:

$$I(t) = \int_{-\infty}^t IRF(t') \sum_{i=1}^n A_i e^{-\frac{t-t'}{\tau_i}} dt' \quad (\text{Eq.2.4})$$

By deconvoluting the recorded signal $I(\tau)$ (software FluoFit), a time resolution of 15 ps can be achieved.

2.7. Forster resonance energy transfer (FRET)

FRET is a non-radiative energy transfer process that occurs between two fluorophores (dyes) based on long range dipole-dipole (Forster) coupling. Therefore, FRET efficiency is dependent on the distance between these two fluorophores. The FRET process can be illustrated as following: after the first fluorophore (donor, **D**; Figure 2.13-(A,C)) is excited, the energy can be transferred to another fluorophore with lower energy bandgap (acceptor, **A**) as long as two conditions are fulfilled: first, both fluorophores are close enough to each other (< 10 nm); and second, the

absorption spectrum of the acceptor must overlaps with the emission spectrum of the donor. As the energy transfer process happens, the emission intensity of the acceptor increases in expense with the decreasing of emission intensity of the donor.

Figure 2.13: Principles of fluorescence resonance energy transfer (FRET): (A) The donor dye molecule (green) is excited directly from a light source. If the two dyes are closed, the energy can be transferred non-radiatively from the donor to the acceptor (red). (B) If the distance between two dyes is increased, lesser energy is transferred, resulting in stronger donor emission. (C) Simplified Jablonski diagram showing energy bandgaps and the pathways of the energy transfer process. (D) Plot of FRET efficiency as a function of the donor-to-acceptor distance at different Forster radius R_0 .

The FRET efficiency E between the two excited states $S_{1,D} \rightarrow S_{1,A}$ is given by Equation 2.5, where k_D is the emission efficient constant of the donor (without the acceptor) and k_T is the energy transfer efficient constant. According to Forster, E is inversely relates to the power of six of the distance R between the two dyes,

$$E = \frac{k_T}{k_T + k_D} = \frac{1}{1 + (R/R_0)^6} \quad (\text{Eq. 2.5})$$

The Forster radius R_0 is the distance between 2 dyes when the FRET efficiency equals to 50%, which ranged from 2 and 7 nm for typical pairs of fluorophores and can be calculated by Equation 2.6 as:

$$R_0^6 = \frac{9000(\ln 10)\Phi_D\kappa^2}{128\pi^5 N n^4} \int_0^\infty f_{D(\lambda)} \varepsilon_A(\lambda) \lambda^4 d\lambda \quad (\text{Eq. 2.6})$$

where Φ_D is the donor quantum yield (without the acceptor), κ^2 is the orientation parameter, N is Avogadro's number, and n is the solution's refractive index. The integral overlap of the absorption spectrum of the acceptor with the emission spectrum of the donor is calculated as a function of wavelength λ with $\varepsilon_A(\lambda)$ is the molar extinction coefficient of the acceptor and $f_{D(\lambda)}$ is the normalized donor emission spectrum ($\int f_{D(\lambda)} d\lambda = 1$). The orientation parameter κ^2 represents the orientation between two dipoles: the emission dipole of the donor and the absorption dipole of the acceptor. The value of κ^2 factor can be from 0 (for perpendicular dipoles) to 4 (for parallel dipoles). In most of the case, for simplicity, we assume that the dyes can rotate unrestrictedly; therefore this factor will take an average value of all possible relative orientation, which results in $\kappa^2 = 2/3$.

Previously, fluorescence resonance energy transfer (FRET) was applied as a spectroscopic ruler to study G-quadruplex formation in bulk solution (Green et al. 2006; Mergny et al. 2001; Ying et al. 2003) by using fluorescence dyes attached at the termini of the DNA sequences.

2.8. Single-molecule FRET (smFRET)

Some of the deficiencies of ensemble FRET can be overcome by single-molecule FRET, which is essentially the monitoring of FRET at single molecule level. To achieve this high sensitivity, first, the samples have to be diluted a lot to get into a condition in which individual molecules can be resolved. Second, only a small volume of the sample should be excited and detected to reduce the contribution of “unwanted molecules”, therefore reducing the background signal. Finally, the fluorophores used to study must be bright and photostable to reduce signal-to-noise ratio and allow observing of FRET signal over a significant period of time.

All these conditions were satisfied in the first successful detection of single-molecule FRET, reported by T. J. Ha (Ha et al. 1996), a pioneer researcher in smFRET. To study various conformations and interactions of G-quadruplexes DNA and biomolecules, we construct a smFRET system based on an Olympus IX71 microscope followed a well published protocol by TJ Ha (Selvin et al. 2007). In this section, we describe the construction and test a smFRET system in our lab in NTU.

2.8.1. Construction of single-molecule FRET system

Figure 2.14: Schematic for smFRET spectroscopy. (a) Two different types of excitation: PTIR (prism type) or OTIR (objective type). Emission from tethered molecules is collected by the objective lens, and the slit limits the imaging area into half size of the CCD detector. The image is collimated by 2 lenses and separated into the donor and acceptor channels and imaged side by side on the CCD camera. (b) Scheme of total internal reflection (TIR) excitation setup (zoom-in of square box in (a)). In PTIR type, a focused laser beam enters at a wide incident angle ($\theta_c > 68^\circ$) to the prism positioned on top of the sample, which then creates an evanescent field at the quartz-water interface. Notes of the optic as followed: $\lambda/2$: half waveplate; PBS: polarizing beam splitter; L1–4: lenses; EF: evanescent field; DM1-2: dichroic mirrors; BE: beam expander; LP: long pass filter; θ_c : incident angle (imported from (Roy et al. 2008)).

Prism-type total internal reflection (PTIR) (Figure 2.14), which employed evanescent field excitation, was used to excite a very small and confine volume of DNA molecules that were immobilized on the surface of a quartz slide. The FRET efficiency was measured between a single donor and a single acceptor attached to two ends of an oligonucleotide fragment. The fluorescence was recorded using a charge-coupled device (CCD) camera. To be able to confirm the detected signal is from a single molecule, one had to observe both the coincident inversely change in donor and acceptor channels, as well as single-bleaching steps of the fluorophores.

For example, if the acceptor was photo-bleached and therefore couldn't take the energy from the donor, one must observe the donor emission intensity increased and the acceptor emission vanished. But if the donor is photo-bleached before the acceptor, both emission intensities should reduce to the background level.

Figure 2.15: Picture of our PTIR setup: (A) A total schematic view of the system. (B) The excitation path where the laser is focused by an achromatic lens and totally reflected by a prism. (C) The detection paths: green is the Cy3 emission, and red is the Cy5 emission. The selected emissions are focused on an ultra sensitive CCD camera.

An excitation beam at 532 nm from a solid-state diode Nd:Yag laser was directed to the quartz slide at an angle around 17° with respect to the horizontal plane. The laser power was adjusted proportionally using the attenuator (Neutral Density filter). A Pellin-Broca prism (PBP) made of fused silica with refractive index of 1.46 was used to create TIR (Total Internal Reflection). Meanwhile, the quartz slide was made of fused quartz with approximately the same refractive index, i.e. 1.46. To match these prism-quartz interface (filling the optical gap), immersion oil with refractive index of 1.51 was used. The minimum angle for TIR (Total Internal Reflection) for the Pellin-Broca prism used here are 24° . It is calculated from the quartz slide to water interface (from $n = 1.46$ to $n = 1.33$), which has the critical angle:

$$\theta_{critical} = \sin^{-1} \left(n_{solution} / n_{quartz} \right) \quad (\text{Eq. 2.7})$$

resulting in a value of 65.6° , or 24.4° from the horizontal plane. A plano-convex lens was placed at 50 mm (same distance as its focal length) distanced from the bottom of the prism to achieve the smallest excitation area (Figure 2.15-B).

Figure 2.16: Spectrum profiles of TMR, Cy5 dyes' emission (normalized) and filter Z532/632m and dichroic mirror 632 DCXR.

The fluorescence signal was collected by 60x magnification water immerse objective lens which has NA = 1.2. The working distance of this objective lense allowed ~150 μm penetration depth into the solution. A special customized filter (code Z532/632m-PH special) from Chroma was place in the filter cube of the microscope to block the laser scattering at 523 nm and reduce the overlap of the donor and acceptor emissions (Figure 2.16). A home-made adjustable slit (using 2 razor blades) was used to make the image become rectangular shape. The slit width was about 1.5 mm, which was adjusted to align these 2 rectangular images into the EMCCD chip to make dual emission channels. A 100 mm achromatic plano-convex lens was placed 100 mm away from the image plane, to collimate the beam. A 630 nm dichroic mirror (code 630DCXR from Chroma) was used to separate the donor (TMR or Cy3) and acceptor (Cy5) emission. Finally, these 2 collimated beams were focused using a 250 mm achromatic plano-convex lens, with an EMCCD camera positioned 250 mm away from the lens (Figure 2.15-C). To be able to detect single molecule signal, a state-of-art Andor Ixon DV-897BV EMCCD (Electron Multiplying Charge-Coupled Device) was used. With electrical cooling up to $\sim 85^\circ\text{C}$ and up to 90 % quantum yield (in visible region), it could achieve a very low dark current (due to thermionic emission of electrons) which is crucial in single-molecule experiments.

For TIRF, it is very important that the excitation volume must be extremely small for minimizing unwanted fluorescence from neighbor molecules and hence enhance the contrast of the image. Let's estimate the excitation volume in our system. The diameter of the focused spot is approximately given by:

$$2a \approx \frac{4}{\pi} \lambda \frac{f}{D} \quad (\text{Eq. 2.8})$$

where f is the focal length of the lens, λ is the wavelength used & D is the diameter of the incident beam (Hinterdorfer et al. 2009). In our setup, the focused spot is calculated to be around 35 μm . However, in practice it is larger due to the presence of optical aberration, and typically it is around 50 μm . The intensity of the evanescent wave will decay exponentially with distance, according to the formula:

$$I(z) = I(0) \exp\left(-\frac{z}{d}\right) \quad (\text{Eq. 2.9})$$

$$d = \frac{\lambda}{4\pi \sqrt{(n_1^2 \sin^2 \theta - n_2^2)}} \quad (\text{Eq. 2.10})$$

where n_1 , n_2 are the higher and lower of the two refractive indices, θ is the incident angle of the incoming beam (Hinterdorfer et al. 2009). Assuming the intensity decays to 1/e of the maximum gives $d \sim 100$ nm. Therefore, the excitation volume is about 0.1 fL, which is extremely small and advantageous to reduce both Raman & Rayleigh scattering.

In our setup, achromatic lenses were used to reduce any image deformation due to optical aberration. The magnification after 2 achromatic lenses was 2.5 times. Combined together with the 60x magnification objective lens, the total magnification will be 150 times. In our camera, each pixel corresponds to 16 μm in size. The chip area which captures the light is 8.2 mm x 8.2 mm, therefore each channel is 4.1 mm x 8.2 mm. As the total magnification is 150 times, the sample's imaging area will be 27.33 μm x 54.67 μm . The left channel is the green/donor channel, where the TMR (Cy3) emission will lie. On the other hand, right channel is the red/acceptor channel, where Cy5 emission will lie. They are made correspond to each other. By extracting the anti-correlated data from the intensities of the spots in both channels, we can deduce the FRET efficiency, and hence structural conformation corresponding to the certain FRET value.

2.8.2. Mapping calibration

To guarantee an accurate correlation between the donor and acceptor images, an overlapping map is created from an image of fluorescent micro-beads (Fbeads, code F8806, Invitrogen) with emission in both detection channels. From the emission profile (Figure 2.17-A), most emission of the Fbeads is in donor channel, resulting in very bright spots on donor channel and very dim spots on acceptor channel (Figure 2.17-B). An exposure to laser beam for a few minutes quenches the Fbeads' emission, resulting in a more balance of the spot's brightness in two channels (Figure 2.17-C), which makes it is clearer and easier to pick up the spots. 3-4 fluorescence spots in the donor channel and their corresponding spots in the acceptor channel are manually selected. An automated algorithm (IDL software) then generates a mapping table between the two channels that corrects for offset displacement, rescaling and distortion.

Figure 2.17: (A) Spectrum profiles of the Fbeads' excitation (blue) and emission (red) and the filter Z532/632m (yellow). (B+C) Image of the Fbeads (0.04% solids in solution) stick on quartz surface before (B) and after a few minutes exposed to laser beam (C).

Chapter 3: Following G-quaruplex formation by its intrinsic fluorescence*

Guanine-rich DNA and RNA sequences can form four-stranded helical structures called G-quadruplexes based on stacking of G•G•G•G tetrads (Davis 2004; Neidle 2009; Patel et al. 2007). G-quadruplexes formed by natural DNA sequences in the telomeres and oncogenic promoters and by RNA sequences in the 5' untranslated region (5'UTR) of oncogenic transcripts have been established as attractive anticancer targets (Balasubramanian et al. 2009; Patel et al. 2007). On the other hand, engineered G-quadruplexes can have potential applications ranging from medicine to supramolecular chemistry and nanotechnology (Davis 2004).

G-quadruplex structures are highly polymorphic (Davis 2004; Neidle 2009; Patel et al. 2007). They give rise to specific spectroscopic signatures in infrared (IR) (Gabelica et al. 2008), UV-absorption thermal difference spectra (TDS) (Mergny et al. 1998b; Mergny et al. 2005), CD (Kypr et al. 2009), and NMR (Feigon et al. 1992) spectroscopy. The intrinsic fluorescence yield of DNA is very low (Gepshtein et al. 2008; Markovitsi et al. 2004; Mendez et al. 2009; Miannay et al. 2009) as it corresponds to quantum yields of the order of $\phi \sim 10^{-4}$, and the detection of DNA usually relies on targeting DNA selectively with fluorescent dyes (Lakowicz et al. 2001). This approach has been used also in search of selective quadruplex binders (Alzeer et al. 2009; Monchaud et al. 2008; Yang et al. 2009). In analogy to molecular beacons as fluorescent probes in DNA conformational analysis (Venkatesan et al. 2008), G-quadruplex formation has been followed by pyrene excimer emission

* The work in this chapter has been published in:

Dao, N. T., R. Haselsberger, et al. (2011). "Following G-quadruplex formation by its intrinsic fluorescence." *Febs Letters* **585**(24): 3969-3977.

(Dembska et al. 2010) and fluorescence resonance energy transfer (FRET) in the ensemble (Green et al. 2003; Mergny et al. 2001) and at the single-molecule level (Lee et al. 2005; Okamoto et al. 2008; Ying et al. 2003). In both cases the fluorescent reporters were attached at the termini of DNA sequences. Recently, the detection of G-quadruplex formation using an internal guanine derivative as a fluorescent probe has been reported (Dumas et al. 2010; Dumas et al. 2011; Nadler et al. 2011). However, all these strategies involving exogenous fluorophores might affect the G-quadruplex structure (Ford et al. 1995).

Name	Sequence	Length	Molar extinction coefficient
B-DNA duplex	CGCGAATTCGCG	12	110700
Z-DNA duplex	CGCGCGCGCGCG	12	101500
Human telomere	TTGGGTTAGGGTTAGGGTTAGGGA	24	244300
<i>Oxytricha</i> telomere	GGGGTTTTGGGG	12	115200
<i>Giardia</i> telomere	TAGGGTAGGGTAGGGTAGGG	20	212200
<i>Giardia</i> telomere (Form 1)	TAGGGTAGGGTAGGGTAIGG	20	212200
<i>Giardia</i> telomere (Form 2)	TAGGGTAGGGTGGGTAGGG	19	198200

Table 3.1: DNA oligonucleotides used in this study. The molar extinction coefficient was calculated by nearest-neighbor approximation method.

In recent reports (Mendez et al. 2009; Miannay et al. 2009), it has been pointed out that DNA G-quadruplexes have a higher intrinsic fluorescence quantum yield than their less-structured counterparts. This notion has been further substantiated by large amplitude longer lifetime components of excited states (Miannay et al. 2009). Here we confirm these fluorescence properties and focus on the structural aspects of various well-defined G-quadruplexes structures (Table 3.1). For all the sequences used in this study, we verified the formation and the structure of G-quadruplexes using NMR, UV absorption and CD spectroscopy (Feigon et al. 1992; Kypr et al. 2009; Mergny et al. 1998b; Mergny et al. 2005). Using similar DNA concentrations as for UV absorption or CD spectroscopy, we show that the formation of G-quadruplexes as a function of cations and temperature can be followed by their fluorescence and fluorescence excitation spectra. These features are shown to discriminate, for example, between G-quadruplex and Z-DNA structures.

3.1. Following cation-dependent G-quadruplex formation of human telomeric (HT) sequence

Figure 3.1-(A-E) shows a compilation of structural and spectroscopic data which are representative of unfolded and folded structures of the 24-nt human telomeric sequence d[TT(G₃TTA)₃G₃A] (henceforth termed **HT**). At room temperature, NMR imino protons at 10-12 ppm were observed for **HT** in K⁺ solution, but not in Li⁺ solution (Figure 3.1-B), indicating G-quadruplex formation of this sequence in K⁺ solution (Figure 3.1-C), but not in Li⁺ solution. The NMR spectrum observed here in K⁺ solution was similar to the one reported previously (Luu et al. 2006), in which this human telomeric sequence has been determined to form a (3+1) G-quadruplex.

Figure 3.1-A displays the absorption spectra of HT recorded in different solutions: 5 mM Tris-HCl buffer, pH 6.8 (red), after addition of 5 mM LiCl (green), and after further addition of 5 mM KCl (blue). The difference between these absorption spectra (Figure 3.2) indicative of G-quadruplex formation in the presence of K⁺ is closely related to the absorption thermal difference spectra (TDS) reported in (Mergny et al. 2005).

Figure 3.1-D shows the corresponding fluorescence spectra of **HT** upon excitation at 260 nm. The fluorescence spectra of this sequence in bare Tris-HCl buffer and upon adding 5 mM LiCl are strictly superimposed. Further addition of 5 mM KCl leads to the development of a broad unstructured fluorescence band peaking between 330 and 340 nm as reported before (Mendez et al. 2009; Miannay et al. 2009; Vayá et al. 2010). As a G-quadruplex structure of **HT** was detected by NMR only in K⁺ solution, but not in Li⁺ solution (Luu et al. 2006), we assign this fluorescence band to G-quadruplex formation. It should be noted that the added volumes of LiCl and KCl from high-concentration stocks are small, resulting in a minute change (< 0.5%) in sample volume and thus in concentration. The fluorescence spectra of samples obtained by adding K⁺ to the Li⁺ solution are similar to those containing only K⁺ (Figure 3.3), consistent with the notion that specific cations (but not simply the ionic strength) are responsible for the G-quadruplex formation. Similar results were also obtained upon excitation in the blue and the red wings of the absorption spectrum (Figure 3.4).

Figure 3.1: NMR and optical spectra of the human telomeric $d[TT(G_3TTA)_3G_3A]$ sequence (**HT**) in different solutions. Spectra of **HT** in different solutions are color-coded as follows: 5 mM Tris-HCl buffer, pH 6.8 (red), after addition of 5 mM LiCl (green), and after further addition of 5 mM KCl (blue). The black dashed curve shows the fluorescence of the bare Tris-HCl buffer. (A): Absorption spectra (in OD). (B): Imino proton NMR spectra (normalized to sugar and aromatic protons). (C): The schematic structure of **HT**. (D): Fluorescence spectra of **HT** ($\lambda_{exc} = 260$ nm). (D'): Fluorescence spectra (D) upon subtraction of the Tris-buffer emission. (D''): Fluorescence spectra (D') upon Inner Filter Correction (IFE). (E), (E'), and (E'') are the respective fluorescence excitation spectra monitored at 330 nm, the 295-nm peak in (R) being the Raman signal of water. Fluorescence and absorption spectra were recorded at 20°C. The DNA concentration was $\sim 5 \times 10^{-5}$ M per base in fluorescence and UV absorption measurements, and $\sim 1 \times 10^{-2}$ M per base in NMR experiments.

Figure 3.2: Difference absorption spectra of the **HT** sequence, obtained by subtracting the absorption spectra of **HT** in Li^+ (green) and K^+ (blue) containing buffer from the absorption spectrum of **HT** in bare Tris-HCl buffer.

Figure 3.3: Fluorescence spectra at 20°C of the **HT** sequence upon 260-nm excitation in different ion conditions. The DNA sample was dialyzed in Li^+ and put in 5 mM Tris-HCl buffer solution. Fluorescence spectra of sample after (i) adding 5 mM Li^+ and 5 mM K^+ (black curve) and (ii) after adding only 5 mM K^+ (triangle-blue curve). The spectra are not IFE corrected.

Figure 3.4: Fluorescence of the HT sequence excited in the blue and red wings of the absorption spectrum in different solutions. Colors coded as in Figure 3.1. The spectrum in black is the reference fluorescence of the Tris-HCl buffer only. Excitation at 254 nm (A), 270 nm (B), and 290 nm (C). The spectra are not IFE corrected.

Figure 3.1-E shows an increasing amplitude of an extremely broad and unstructured excitation spectrum, probing the fluorescence at its maximum (340 nm) in K^+ solution. This characteristic increase was only observed upon addition of K^+ ions, but not Li^+ ions. Similar results were obtained, when the fluorescence excitation was probed in the blue and red wings of the fluorescence band (Figure 3.5).

Figure 3.5: Fluorescence excitation spectra of the **HT** sequence in different ion conditions monitored at 320-nm (A) and 360-nm (B). Colors coded as in Figure 3.1. The spectra are not IFE corrected.

Since the fluorescence measurements depicted in the spectra (D) and (E) were performed on samples with a large optical density ($OD = 0.5$), the resulting spectra have to be corrected for emission losses caused by the Inner Filter Effects (IFE). These losses take place in two paths: (i) excitation: where the molecules absorb at

excitation wavelength as the excitation beam penetrates through the sample; (ii) emission: where the neighbor molecules reabsorb at the emission wavelength of the emitting light.

Due to the negligible overlap of absorption and fluorescence spectra, corrections for reabsorption of fluorescence are minimal and can be neglected. Fluorescence and excitation spectra were corrected in two consecutive steps: (i) Subtraction of the background emission of the pure Tris-HCl buffer from the measured fluorescence and fluorescence excitation spectra, (D) and (E), yielding now (D') and (E'). (ii) Subsequently, the spectra (D') and (E') were corrected for IFE using the approximate Equation 2.3. The corrected fluorescence and excitation spectra are shown in Figure 3.1 as (D'') and (E'').

The consequences of correcting the observed spectra (D) and (E) are two folds: (i) The corrected fluorescence spectra display an increase of amplitude by approximately a factor of 2 maintaining otherwise the spectral envelope. We note in passing that the IFE corrected fluorescence spectrum (D'') still shows a small-amplitude offset-emission in the 450-500 nm range. This unknown emission also reported in the literature, e.g. (Miannay et al. 2009), responds to a broad excitation band peaking around 325 nm (Figure 3.6) that is well beyond the measurable absorption of DNA. At shorter wavelengths, the excitation spectrum follows the absorption contours. (ii) Apart from a significant increase of amplitude, the IFE-corrected fluorescence excitation spectra follow obediently the absorption spectra as shown in Figure 3.7. The similarity between absorption and fluorescence excitation spectra underlines the notion that the origin of the fluorescence band peaking around 330-340 nm is indeed absorption of DNA rather than of a photodegradation product.

Figure 3.6: (A) Absorption spectrum (thick black curve), excitation spectra probed at 405 nm (circle dot, black curve) and 450 nm (circle dot, red curve) and fluorescence spectrum under 330-nm excitation (thin black curve) of the **HT** sequence in 5 mM Tris-HCl buffer, pH 6.8. (B) Fluorescence decays of **HT** probed at 330 nm (black curve) and 450 nm (red curve) by time correlated single-photon counting (TCSPC) technique after excitation at 260 nm. The instrument response function (IRF) (gray curve) was measured using the Raman emission of water at 285 nm.

Figure 3.7: Comparison between the normalized absorption (continuous line) and the normalized fluorescence excitation (dashed line) spectra of **HT** in Tris-HCl buffer and in the presence of different cations, Li^+ and K^+ . (A): in 5 mM Tris-HCl buffer, pH 6.8; (B): after addition of 5 mM LiCl; (C): after further addition of 5 mM KCl.

3.2. Fluorescence quantum yield of HT sequence

Using the reported quantum yield reference $(1.3 \pm 0.1) \times 10^{-4}$ of dGMP (Changenet-Barret et al. 2010), the plots of the integrated fluorescence intensity vs. absorbance (Figure 3.8) allowed us to estimate the quantum yield of **HT** in bare Tris-HCl buffer and in the presence 5 mM K^+ , being $(1.4 \pm 0.1) \times 10^{-4}$ and $(3.5 \pm 0.1) \times 10^{-4}$ respectively. Thus, the quantum yield of **HT** in bare Tris-HCl buffer was similar to that of dGMP, whereas a 2.7-fold increase in quantum yield was observed when **HT** formed a G-quadruplex structure in the presence of K^+ . The quantum yield obtained here for the **HT** G-quadruplex is within the range of the previously reported values, from 2.4×10^{-4} to 9.5×10^{-4} (Table 3.2). The spread of these values may be due to differences in G-quadruplex structures, as well as other experimental conditions.

Figure 3.8: Plots of integrated fluorescence intensity (from 300 to 500 nm) vs. absorbance of 2'-deoxyguanosine 5'-monophosphate (dGMP) – black square; **HT** sequence in 5 mM Tris-HCl buffer (pH 6.8) – red circle; **HT** in 5 mM Tris-HCl buffer (pH 6.8) supplemented with 5 mM KCl – triangle green. The absorbance of the samples was collected right before the fluorescence experiments. In each case (deionized water, 5 mM Tris-HCl, and 5 mM KCl in 5 mM Tris-HCl) the respective background spectrum has been subtracted prior to integration. The data were fit by linear least-square regression $y = Ax$, giving $y_1 = 2.86 \times 10^7 x$ ($R^2 = 0.996$) for dGMP in deionized H_2O , $y_2 = 3.02 \times 10^7 x$ ($R^2 = 0.999$) for **HT** in Tris-HCl buffer; and $y_3 = 7.70 \times 10^7 x$ ($R^2 = 0.983$) for **HT** in Tris-HCl buffer supplemented with 5 mM KCl.

Sequence	Solvent	Excitation Wavelength (nm)	Quantum yield ($\times 10^{-4}$)	Reference
HT	5 mM Tris-HCl (pH = 6.8)	265	1.4 ± 0.1	This work
HT	5 mM Tris-HCl (pH = 6.8), 5 mM KCl	265	3.5 ± 0.1	This work
GpG	phosphate (pH = 7)	248	1.3	(Vigny et al. 1977)
polyG	phosphate (pH = 7)	248	4.7	(Vigny et al. 1977)
G ₉	10 mM potassium phosphate (pH 7.2), 50 mM KCl	275.5	2.39	(Mendez et al. 2009)
G-wires	30 mM sodium phosphate (pH 7.4), 100 mM NaCl	265	9.5 ± 0.1	(Changenet-Barret et al. 2010)

Table 3.2: Quantum yields of G-rich sequences in different buffer solutions.

3.3. Time-resolved fluorescence decay of human telomeric sequence

The excited-state dynamics of the **HT** sequence in the presence of Li^+ and K^+ ions has been probed using time-correlated single-photon counting (TC-SPC), a technique which favors minimal photodegradation of DNA due to the low overall excitation doses. As most of the **HT** fluorescence, excited at 260 nm and probed in the 330-340 nm range, decays on the time scale of the IRF of the system, the ultrafast time components cannot be resolved and are therefore subsumed within the limiting resolution of 15 ps. Nevertheless, the temporal fluorescence profile shown in Figure 3.9 points to a slightly slower fluorescence decay in the presence of K^+ , than in the presence of Li^+ . These qualitative results are consistent with previous reports of the excited-state dynamics of single- and double-stranded DNA (Gepshtein et al. 2008; Middleton et al. 2009; Schwalb et al. 2008; Vayá et al. 2010), G-rich sequence (Gepshtein et al. 2008; Schwalb et al. 2008) as well as G-quadruplexes (Miannay et al. 2009).

Figure 3.9: Fluorescence decay of **HT** excited at 260 nm in different solutions as encoded in Figure 3.1. The instrument response function (IRF) of the (TCSPC) set-up was measured using the Raman emission of water at 285 nm.

The essential difference between the three-exponential fit parameters (table 3.3) for the fluorescence decay of **HT** in Li^+ and K^+ solutions is the observation of a 50 ps component with a small fractional amplitude of ~2% in the case of K^+ as compared to Li^+ which shows only an ultrafast decay. Concomitantly, the fractional amplitudes of the ultrafast components (<15 ps) are 98% in K^+ as compared to 100% in Li^+ solution. We time-resolved also the fluorescence offset in the 450-500 nm region of the steady-state fluorescence spectra D' in Figure 3.1. In contrast to the decay pattern probed at 330 nm, the decay of this background emission in the 450-500 nm is dominated by long 2 ns and 500 ps components (Figure 3.6). In principle, the origin of this emission might be one or more impurities at low concentration showing up only due to a favorable combination of extinction coefficient and excited state lifetimes. However, since the excitation spectrum of this emission at shorter wavelengths follows the absorption spectrum of **HT** (Figure 3.6), low-lying excited states of DNA might well be involved. For example, already in 1960, $n\text{-}\pi^*$ transitions have been proposed to exist at the low-energy side of the strong 260-nm $\pi\text{-}\pi^*$ absorption band of DNA (Rich et al. 1960). Based on electroabsorption spectra the low-energy absorption in the 300-nm region of single-stranded polynucleotides in water/glycerol mixtures has been explained by interaction of $n\text{-}\pi^*$ and $\pi\text{-}\pi^*$ transitions of neighboring bases (Krawczyk et al. 2007). Most recently, the weak absorption tail beyond 300 nm observed for DNA single- and double-strands has

been attributed to charge-transfer states that are strongly stabilized in aqueous solution (Banyasz et al. 2011).

Buffer	λ_{em} [nm]	A ₁	τ_1 [ns]	Fract. Int. [%]	Fract. Amp. [%]	A ₂	τ_2 [ns]	Fract. Int. [%]	Fract. Amp. [%]	A ₃	τ_3 [ns]	Fract. Int. [%]	Fract. Amp. [%]	χ^2
Li ⁺	330	27	3.5	1.4	< 0.1	47	0.45	0.3	< 0.1	6594000	$\leq$ 0.01	98.3	100.0	1.081
K ⁺	330	37	1.1	4.3	< 0.1	2636	0.05	13.4	1.7	149980	$\leq$ 0.01	83.3	98.3	1.119

Table 3.3: Fitting parameters of the TCSPC decays of **HT** sequence in Tris-HCl buffer before and after adding different salts.

3.4. Following temperature-dependent G-quadruplex formation of HT sequence

As the formation of G-quadruplexes results in increasing amplitude of the fluorescence, we followed the folding and unfolding of the **HT** quadruplex by measuring the temperature dependent spectra (Figure 3.10-A) and plotting the amplitudes of the 330-nm fluorescence vs. temperature as displayed in Figure 3.10-B. As expected, a similar temperature-dependence was observed for the 260-nm fluorescence excitation signals (Figure 3.10-C). Contributions from temperature-dependent pH-sensitivity of the Tris-buffer can be ignored since similar fluorescence melting results were obtained in Tris-HCl and cacodylate buffer (Figure 3.11). These measurements, performed at low (5 mM) and high (90 mM) K⁺ concentration, illustrate the role of K⁺ concentration on the stability of the G-quadruplex. We emphasize that both, the data of the cooling and heating processes are fully reversible and well-superimposed (Figure 3.11). The thermal reversibility of the drastic changes in the fluorescence amplitudes rules out significant contributions from photoproducts accumulating under the experimental conditions of our measurements. The melting characteristics shown in Figure 3.10-(B,C) are consistent with the ones recorded in UV absorption at 295 nm (Figure 3.11-D), with the melting temperature (T_m) being 50±5 and 60±5 °C for **HT** at low and high K⁺ concentration. In contrast to UV melting curves resting on the loss of the 295-nm hyperchromicity of DNA G-quadruplexes upon unfolding, the melting curves displayed in Figure 3.10-(B,C), are attributed to the decrease of excited state lifetimes upon G-quadruplex unfolding at high temperature. Detailed studies on

temperature-dependent excited state dynamics of DNA structures in different environments are in progress.

Figure 3.10: (A): Temperature dependence of fluorescence spectra of **HT** in 5 mM Tris-HCl buffer (pH 6.8) supplemented with 5 mM KCl upon excitation at 260 nm. Only the cooling process is shown. (B): 330-nm peak intensity of fluorescence in the presence of 5 and 90 mM K^+ as a function of temperature. (C): Temperature dependence of the excitation intensity at 260 nm, probed at the 330-nm peak fluorescence of **HT** in 5 and 90 mM K^+ . Full (empty) circles represent the cooling (heating) processes in 5 mM K^+ solution. Full (empty) triangles represent the cooling (heating) processes in 90 mM K^+ solution. The DNA concentration was $\sim 5 \times 10^{-5}$ M per base.

Figure 3.11: (A-C) Temperature-dependent fluorescence of **HT** in different buffer solutions upon 260-nm excitation. Cooling and heating cycles representing **HT** in 5 mM Tris-HCl, pH = 6.8 (A and A') and in 5 mM cacodylate, pH=6.8 (B and B'). The DNA sample was dialyzed in solution containing LiCl, and then put into the respective buffer solution. Subsequently, 5 mM KCl was added. (C) 330-nm peak intensity of **HT** in both buffer solutions (A, A' and B, B') as a function of temperature. The spectra are not IFE corrected. (D) UV-melting characteristics of the **HT** sequence at low and high K^+ concentration, monitored in absorbance at 295 nm. Full circles (triangles) represent the cooling process and empty ones the heating process in 5 mM K^+ (90 mM K^+) solution.

3.5. Formation of G-quadruplex by other sequences: effect of cation and folding topology

3.5.1. *Oxytricha* telomeric sequence: effect of cation

In order to find out whether the fluorescence spectra depend on the nature of the cations K^+ and Na^+ in a case where the quadruplex folding is not dependent on the nature of these cations, we recorded temperature-dependent fluorescence spectra of the *Oxytricha* telomeric $d[G_4T_4G_4]$ sequence in K^+ , Na^+ and Li^+ solutions (Figure 3.12-(A-C)). Here, the formation of the same dimeric basket-type G-quadruplex fold in K^+ and in Na^+ solution (Figure 3.12-(A',B')) was verified by NMR imino proton spectra (Figure 3.12-(A'',B'')), in accordance with the respective NMR spectra reported previously (Schultze et al. 1999). Although 1D imino-proton NMR spectra of this sequence in K^+ and Na^+ solutions exhibit different patterns, probably due to difference in the conformation of the diagonal T4 loops and the nature of the stabilizing cations, this sequence has previously been shown by NMR to have the same folding topology in both K^+ and Na^+ solutions (Schultze et al. 1999).

Fluorescence spectra of $d[G_4T_4G_4]$ in K^+ solution (Figure 3.12-A) and in Na^+ solution (Figure 3.12-B) are similar, with the latter being only slightly broader and red-shifted as compared to the former. For the sample in K^+ solution, the 330-nm fluorescence peak was still observed at 90°C, suggesting the formation of residual higher-order G-quadruplex structures stable at high temperature. Fluorescence spectra of $d[G_4T_4G_4]$ in Li^+ solution indicated the lack of G-quadruplex formation (Figure 3.12-C), consistent with the absence of significant imino proton peaks in the NMR spectrum (Figure 3.12-C'').

Figure 3.12 (A-C): Temperature-dependent fluorescence spectra (shown for the cooling process) upon 260-nm excitation of the *Oxytricha* telomeric d(GGGGTTTTGGGG) sequence in different solutions containing 90 mM of (A) K⁺, (B) Na⁺, and (C) Li⁺ at a DNA concentration of $\sim 5 \times 10^{-5}$ M per base. Loops are colored red; anti and syn guanines are colored cyan and magenta, respectively. A', B': Schematic structure of G-quadruplexes and (A''-C'') imino proton NMR spectra of *Oxytricha* telomeric d(GGGGTTTTGGGG) sequence in (A'') K⁺, (B'') Na⁺, and (C'') Li⁺ solution.

3.5.2. *Giardia* telomeric sequence: effect of folding topology

Alternatively, to probe the effect of different G-quadruplex topologies on the fluorescence spectrum in cases where the same cation K^+ controls the folding, we studied the *Giardia* telomeric $d[(TAG_3)_4]$ sequence (Figure 3.13). The *Giardia* telomeric $d[(TAG_3)_4]$ sequence differs from the human telomeric sequence only by one Thymine deletion within the (TTA) linker in each repeat. It has been studied by our group in order to investigate the effect of loop length and sequence on the folding topology of G-quadruplexes (Hu et al. 2009). We showed by NMR that this sequence coexists in two G-quadruplex conformations (Figure 3.13-A') in K^+ solution (named Form 1 and Form 2), and each form represents the major structure upon single-residue modification (Hu et al. 2009). Substitution of G18 by an inosine favors Form 1, which is a basket-type G-quadruplex containing two G-tetrads, a $G \cdot (A-G)$ triad and two $A \cdot T$ base pairs. Deletion of A12 favors Form 2, which is a propeller-type parallel-stranded G-quadruplex. The formation of these two G-quadruplex conformations was again verified by NMR imino proton spectra (Figure 3.13-(A''-C'')). Temperature-dependent fluorescence spectra for the three sequences (natural, Form 1 and Form 2) exhibited a similar extent of fluorescence changes in cooling and heating cycles (Figure 3.13-(A-C)). The spectra of Form 1 and Form 2 showed different fine structures, with two peaks at 325 and 340 nm for Form 1 (Figure 3.13-B), and only one peak at 330 nm for Form 2 (Figure 3.13-C). This could reflect the base stacking in the structure of the loops of Form 1. Spectral pattern of the natural sequence resembles to a weighted average between the spectra of Form 1 and Form 2, consistent with this sequence forming a mixture of the two conformations. Residual fluorescence at high temperatures was only observed for the natural and Form 1 sequences, but not for Form 2 sequence.

Figure 3.13 (A-C): Temperature-dependent fluorescence spectra (shown for the cooling process) upon 260-nm excitation of *Giardia* telomeric sequences in three forms: the natural form, $d[(\text{TAG}_3)_4]$, and the modified Forms 1, $d[(\text{TAG}_3)_3\text{TAIGG}]$, and 2, $d[(\text{TAGGG})_2\text{TG}_3\text{TAG}_3]$, all in 90 mM K^+ at a DNA concentration of $\sim 5 \times 10^{-5}$ M per base. (A'-C'): Schematic structures and (A''-C'') imino proton NMR spectra of G-quadruplexes. (A', A'') natural $d(\text{TAGGG})_4$ sequence; (B', B'') Form 1, and (C', C'') Form 2. Loops are colored as in Figure 3.12.

3.6. Distinction between Z-DNA and G-quadruplex formation

Thus, the intrinsic fluorescence provides a method to identify and follow the G-quadruplex formation complementing other spectroscopic techniques (Feigon et al. 1992; Kypr et al. 2009; Mergny et al. 1998b; Mergny et al. 2005). For instance, since Z-DNA and G-quadruplex DNA give similar UV-absorption TDS signals (Mergny et al. 2005) at the signature wavelength of 295 nm (Figure 3.14-A), the two structures can be easily distinguished by their fluorescence thermal difference spectra (Figure 3.14-B), as the increase in the 330-nm fluorescence signal due to G-quadruplex formation is significantly higher than that due to the formation of a Z-DNA duplex. It should be noted that the DNA strand concentration used in the fluorescence studies here was only about $\sim 2 \mu\text{M}$, similar to those typically used for UV absorption or CD spectroscopy.

Figure 3.14: (A) Absorption and (B) fluorescence thermal difference spectra (TDS) upon 260-nm excitation of Z-DNA (red) and G-quadruplex DNA (black). Z-DNA is formed by the $d[(\text{CG})_6]$ sequence in 4 M NaCl and G-quadruplex is formed by **HT** in 5 mM KCl. Absorption TDS (Mergny et al. 2005) were obtained by subtracting the UV absorption spectrum at 20°C from that at 90°C while fluorescence TDS were obtained by subtracting the fluorescence spectrum at 90°C from that at 20 °C. The DNA concentration was $\sim 5 \times 10^{-5} \text{ M}$ per base in fluorescence and $\sim 8 \times 10^{-5} \text{ M}$ per base in absorption measurements.

3.7. Conclusion

In conclusion, large increase in fluorescence amplitude accompanied by parallel absorption and fluorescence excitation spectra and high reversibility of temperature-dependent fluorescence signals support the notion that these are characteristic features of G-quadruplex structures. Thus, fluorescence spectroscopy constitutes a method for the identification of G-quadruplex formation. Fluorescence spectroscopy is shown to complement other spectroscopic techniques and to provide independent information on specific NMR structures, as well as on the stability and polymorphism of G-quadruplexes.

Chapter 4: Excimer formation by stacking of G-quadruplex blocks*

DNA strongly absorbs ultraviolet light. As excited states in DNA trigger genetic damage, their dynamics have been extensively studied and still are until the present date (Gustavsson et al. 2010; Middleton et al. 2009). Since the millennium such studies are particularly fueled by both the reach of the UV spectral region in femtosecond spectroscopy and the advances in computational methods. The outcome of the first femtosecond studies on the four isolated DNA bases in aqueous solution was that their lowest excited singlet state deactivates on the sub-picosecond time scale (Gustavsson et al. 2002; Pecourt et al. 2000; Peon et al. 2001). Such ultrafast decay rates certainly help to minimize the extent of deleterious photoreactions. However, this optimization principle for photochemical stability is to some extent violated in the case of oligonucleotides, where multiple decay routes exist with much longer time constants. There, in addition to the ultra-short excited state lifetime of the majority population, long fluorescence decay components spanning hundreds of picoseconds have been identified (Crespo-Hernandez et al. 2008; Kwok et al. 2006; Miannay et al. 2009; Middleton et al. 2009; Plessow et al. 2000; Schwalb et al. 2008; Takaya et al. 2008; Vayá et al. 2010).

In spite of the well-taken notion (Gustavsson et al. 2010) that the early evolution of $\pi\pi^*$ singlet states in DNA may be more complex, a large body of experimental and computational studies (Gustavsson et al. 2010; Middleton et al. 2009; Santoro et al. 2009) converges in attributing the nature of sub-picosecond fluorescence to internal

* The work in this chapter is being written for publication:

“Excimer formation by stacking of G-quadruplex blocks”, Dao, N. T., Haselsberger, R., Michel-Beyerle, M. E. and Phan, A. T.

conversion brought about by a barrierless transition from the excited state to the electronic ground state via highly efficient conical intersections. In contrast, the nature of the slow fluorescence decay components is still under lively debate, as illustrated by two recent papers (Su et al. 2012; Vayá et al. 2012b). Two key issues in the different approaches to long-lived fluorescence in polynucleotides are the time scale on which excitons on a “string” of fluctuating nucleobases localize and, if they do so, the nature of the emissive states formed. Fluorescence anisotropy determined for calf thymus DNA (Vayá et al. 2012b) by fluorescence upconversion led to a model explaining long fluorescence components by the involvement of charge transfer states giving rise to delayed excitations. An alternative explanation of long fluorescence components in DNA oligomers rests on femtosecond transient absorption data (Middleton et al. 2009). The recent study (Su et al. 2012) on single-stranded adenine sequences that differ in length and the extent of base stacking led to the assignment of short-lived signals to unstacked nucleobases, while the long-lived emission on the 100-ps time scale was attributed to stacked nucleobases. Stacked nucleobases showing monomer characteristics in their ground states may well favor strong electronic interactions upon excitation and form excimers or exciplexes, depending on the extent of charge transfer between the interacting nucleobases. In this picture, long lifetimes of excited pairs of stacked nucleobases may reflect their slow decay to a dissociative ground state.

Whatever the nature of long-lived states in DNA may be, it is obvious that the excited state dynamics in nucleobase oligomers must be very sensitive to their individual structure, π -base stacking and hydrogen-bond base pairing. And exactly this sensitivity poses the crucial problem in understanding DNA fluorescence dynamics. Thus, it is not surprising that structurally diverse oligonucleotides studied under different experimental conditions prompted different views on DNA photophysics. In this situation it is highly desirable to investigate the fluorescence features of a DNA structure where the overlap of nucleobases can be controlled. In this intention, we focus in the present paper on the fluorescence studies from specific guanine quadruplex structures that allow manipulating the stacking interaction between nucleobases under the control of parallel-recorded NMR spectra.

In recent reports, it has been pointed out that G-quadruplexes have higher intrinsic fluorescence quantum yields than their single-stranded counterparts (Dao et al. 2011; Gepshtein et al. 2008; Hua et al. 2012; Mendez et al. 2009). As in single- and double stranded DNA systems, the increase of quantum yield in G-quadruplexes

is caused by increased amplitudes of long-lived fluorescence components. G-quadruplex structures may form from guanine-rich DNA and RNA sequences. They comprise stacks of G•G•G•G tetrads and stabilized through coordination of mono- or divalent cations. G-quadruplexes are highly polymorphic, may be intramolecular or intermolecular, and their properties depend strongly on their structure (Burge et al. 2006; Davis 2004; Patel et al. 2007; Phan 2010). Beyond their putative existence in biology and role in telomere function (Biffi et al. 2013; Lipps et al. 2009; Maizels 2006), G-quadruplex structures show potential applications in different fields including cancer (Balasubramanian et al. 2011; Patel et al. 2007) and HIV therapy (Chou et al. 2005; Patel et al. 2007) as well as in functional supramolecular assemblies (Davis 2004), where they may serve as building elements.

A recent NMR structural study of the DNA sequence d[G₃T]₄ - named T30695 in reference (Do et al. 2011) and **T0** in this work - showed that this sequence adopts a quadruplex dimer structure formed by the stacking of two propeller-type parallel-stranded G-quadruplex units at their 5'-ends. In the following we focus on the fluorescence properties of dimer-forming G-quadruplex structures and examine interfacial G-stacking in parallel NMR and fluorescence experiments.

For the exemplary sequence **T0** we report on an intense, red-shifted fluorescence band with a distinct maximum at 385 nm that is dominated by a 1-ns lifetime component. This phenomenon is attributed to excimer emission resulted from close coplanar G-G stacking of two **T0** G-quadruplex monomer blocks at their 5' G-tetrads. Since we put special emphasis on the correlation of structural and spectroscopic parameters, unless otherwise specified, the majority of optical measurements are performed at the condition of the concurrent NMR experiments implying high DNA concentration. Apart from testing out the influence of parameters as DNA concentration, ionic strength, and temperature on the 385-nm fluorescence, we also demonstrate the ability to manipulate excimer interaction by manipulating the interfacial G-G overlap using different DNA sequences.

4.1. Structure and spectroscopy of a stacked G-quadruplex dimer

We performed a structural NMR study in conjunction with optical spectroscopy on the DNA sequence d[(G₃T)₄], denoted as **T0** in this paper (Table 4.1). There are only few studies on the photophysics of DNA that are based on solution NMR structures and to the best of our knowledge the existing studies refer to chromophore-modified

DNA single- and duplex systems (Siegmund et al. 2009; von Feilitzsch et al. 2008). In contrast, here we investigated the photophysics of DNA itself and started out to follow absorption/fluorescence features of G-quadruplexes using samples that were identical with the ones analyzed in NMR experiments. In Figure 4.1, a triad of characteristic “tandem” experiments on the sequence **T0** in the absence and presence of K⁺ ions is shown. This triad collects the NMR structural information (A, B) together with steady-state spectra of absorption and fluorescence (C) as well as ps-time resolved fluorescence decay traces (D). These parallel-recorded structure/fluorescence features of the DNA sequence **T0** are exemplary and will serve as a reference for other sequences studied in Sections 4.2 and 4.3.

Name	Sequence	Length
T0 (T30695^a)	GGGTGGGTGGGTGGGT	16
T1	TGGGTGGGTGGGTGGGT	17
T2	TTGGGTGGGTGGGTGGGT	18
T0_S^c	GGGSGGGSGGGSGGG	16
T0_3'	GGGTGGGTGGGTGGG	15
T0_I2 (J19^b)	GIGTGGGTGGGTGGGT	16
93del	GGGGTGGGAGGAGGGT	16
HT	TTGGGGTTAGGGGTTAGGGGTTAGGGA	24
Telo	TTAGGG	
TERRA	GGGUUAGGGU	
GGA4	GGAGGAGGAGGA	

Table 4.1: Oligonucleotides used in this study. Except for TERRA, which is an RNA sequence, the remaining oligonucleotides used are DNA sequences. ^aAs named in (Do et al. 2011). ^bAs named in (Do et al. 2012). ^cS represents a spacer, a deoxyribose without the base.

4.1.1. NMR structure

At the condition of NMR experiments, i.e. high DNA concentration (0.25 mM), imino protons at 10-12 ppm, indicative of the G-quadruplex formation, were observed for **T0** in the presence of 50 mM K^+ , but not in the absence of K^+ (Figure 4.1-B). Here, the NMR spectrum of **T0** in K^+ solution has the same characteristics as the one reported recently (Do et al. 2011), where this sequence was shown to form a dimeric structure involving two propeller-type parallel G-quadruplex blocks being stacked on each other at their 5' end (Figure 4.1-A). The excited state properties of this G-quadruplex dimer are the central topic of this paper.

4.1.2. Steady-state absorption and fluorescence spectra

The absorption spectra of **T0** at high DNA concentration (0.25 mM) in the absence and presence of K^+ ions (Figure 4.1-C) are invariant with the exception of a small ($\approx 5\%$) net hyperchromic effect in the region of the maximum (~ 260 nm). The well-established hyperchromicity of G-quadruplex structures in the 295-nm region (Mergny et al. 1998b) is less pronounced. Upon excitation of **T0** at 260 nm (close to the absorption maximum) in the absence of K^+ , a fluorescence spectrum peaking at 330 nm, typical for single-stranded DNA, was recorded (Figure 4.1-C). Surprisingly, in the presence of 50 mM K^+ a novel strong fluorescence band developed with a maximum at 385 nm (Figure 4.1-C).

Figure 4.1: NMR, absorption and fluorescence spectra as well as fluorescence decay traces of the **T0** sequence (0.25 mM) in Tris-HCl buffer (5 mM, pH 6.8) in the absence (black) and presence of 50 mM KCl (blue) at 25°C. (A) Sequence and schematic structure of the stacked G-quadruplex dimer formed by **T0** in K⁺ solution. Guanines in the stacking layer are highlighted in blue, all other guanines are colored cyan. (B) Imino proton NMR spectra. (C) Absorption (dashed curves) and fluorescence (continuous curves) spectra of **T0**. (D) Fluorescence decay of **T0**, excited at 260 nm and probed at 430 nm. The instrument response function (IRF) of the TCSPC set-up was measured using the Raman emission of water at 285 nm. The added volume of KCl from a high-concentration stock was very small, resulting in a minor change (<0.5%) of the sample volume and thus of the DNA concentration.

Figure 4.2: Comparison between the normalized fluorescence excitation spectra (IFE corrected) probed at 385 nm, 430 nm, and 350 nm and the normalized absorption spectrum of **T0** at 0.5 OD in 50 mM K^+ , 5 mM Tris-HCl buffer.

The maximum of the fluorescence spectrum of **T0** in K^+ solution displayed in Figure 4.1-C is red-shifted by $\approx 4000\text{ cm}^{-1}$ from the 330-nm peak of **T0** in the absence of K^+ . However, in order to circumvent complex corrections at high concentrations (measured in the reflection mode), fluorescence excitation spectra (Figure 4.2) have been recorded at the smaller OD of 0.5 in the standard 90° geometry. Upon IFE correction, the fluorescence excitation spectra probed at 350 nm, 385 nm, and 430 nm follows the envelope of the absorption band. It is interesting to note that the (i) the maximum of the fluorescence excitation spectrum is slightly blue-shifted, (ii) its half-width at half maximum is by 1000 cm^{-1} smaller than the respective absorption spectrum, and (iii) the shoulder in the 280 nm range is more pronounced. The fluorescence quantum yield $Q_{\text{T0}_{\text{dimer}}}$ based on tryptophan (Lakowicz 2006) as external standard has been determined to be $Q_{\text{T0}_{\text{dimer}}} = 1.52 \pm 0.1 \times 10^{-3}$ (Figure 4.3). This value is in excellent agreement with a less direct estimate $Q_{\text{T0}_{\text{dimer}}} = 1.6 \times 10^{-3}$ (Figure 4.3) that rests on the quantum yield of the monomeric G-quadruplex from the human telomere (**HT**) sequence (Dao et al. 2011).

Figure 4.3: (A) Plot of integrated fluorescence intensity (from 300 nm to 500 nm, excited at 260 nm) vs. absorbance of Tryptophan – red square dot in miliQ water; and **T0** – circle blue dot in 50mM K⁺ and 5 mM Tris-HCl buffer (pH 7.0). The absorbances of the samples were collected right before the fluorescence experiment. A background spectrum (50 mM KCl in 5 mM Tris-HCl buffer) was subtracted prior to integration. The data were fit by linear least-square regression, forcing an intercept at 0 to give $y_1 = 2.3015 \times 10^{12}x$ ($R^2 = 0.9837$) for Tryptophan and $y_2 = 2.698 \times 10^{10}$ ($R^2 = 0.988$) for **T0** sequence. The ratio between 2 slopes is $\alpha_3 / \alpha_4 = 85.3$. If we take the quantum yield of Tryptophan in miliQ water as 0.13 ± 0.01 (ref. 29 in the paper), then the quantum yield of **T0** in 50mM K⁺, 5mM Tris-HCl is $1.52 \pm 0.1 \times 10^{-3}$. (B) Plot of integrated fluorescence intensity (from 300 nm to 500 nm, excited at 265 nm) vs. absorbance of human telomeric (**HT**) d[TT(GGGTTA)₃GGGA] sequence – black triangle dot; and **T0** in 5 mM KCl – circle blue dot in 50mM K⁺ and 5 mM Tris-HCl buffer (pH 7.0). After background subtraction and fitting gives $y_1 = 0.77 \times 10^8x$ ($R^2 = 0.983$) for **HT** sequence and $y_2 = 3.42 \times 10^8x$ ($R^2 = 0.997$) for **T0** sequence. The ratio between 2 slopes is $\alpha_1 / \alpha_2 = 0.225$. If we take the quantum yield of **HT** in K⁺ buffer as $3.5 \pm 0.1 \times 10^{-4}$ (ref. 17 in the paper), then the quantum yield of **T0** in 50mM K⁺, 5mM Tris-HCl is $1.6 \pm 0.1 \times 10^{-3}$.

Note: For simplicity, we considered the solvent refractive index for the sample (5 mM Tris-HCl buffer, 50 mM KCl) and for the reference (miliQ water) to be similar.

The 385-nm fluorescence band was also observed for various **T0**-modified sequences whose NMR spectra indicated the formation of a similarly stacked G-quadruplex in K⁺ solution (Figure 4.4). These sequences include (i) **T0-3'** derived from **T0** by removing the thymine at the 3' end, (ii) **T0-I2** modified by one inosine substitution (Do et al. 2011) in **T0** at position 2, and (iii) **T0-S** derived from **T0-3'** by substituting all thymine linkers by a spacer (i.e. a deoxyribose without the base). We note in passing that the observation of the 385-nm fluorescence band for **T0-S**, a sequence devoid of thymine, rules out the possibility that the 385-nm emission is

related to thymine dimer formation (Gurzadyan et al. 1993; Marguet et al. 2005). Reversible fluorescence spectra of **T0** (Figure 4.5) upon melting and cooling allow to exclude photodegradation products to be the origin of the 385-nm fluorescence band.

Figure 4.4: (A) Imino-proton NMR spectra at 25 °C of **T0** and modified sequences **T0-3'**, **T0-I2**, and **T2-S** in solution containing 5 mM Tris-HCl (pH 7.0) and 50 mM K^+ . (B-D) Fluorescence spectra of (B) **T0** and **T0-3'**; (C) **T0** and **T0-I2**; and (D) **T0** and **T0-S** under 260-nm excitation. The DNA concentration was ~ 0.2 mM (~ 3.0 OD in a 10 x 1 mm quartz cuvette with 1-mm path length).

Figure 4.5: Temperature-dependent fluorescence spectra upon 260 nm excitation of **T0** sequence ($\sim 4 \mu\text{M}$) in the solution containing 50 mM K^+ : (A) cooling and (B) heating curves.

4.1.3. Excited-state dynamics

The excited-state dynamics of the **T0** sequence in the absence and presence of K^+ has been monitored by TCSPC. In bare Tris-buffer devoid of K^+ ions, the fluorescence decay of **T0** occurs within the instrument function of the spectrometer (15 ps) and most probably faster. Therefore, without higher time-resolution in fluorescence up-conversion experiments, the 3-exponential fit of the decay traces (Figure 4.1-D) and their amplitudes has to be taken with caution since such a fit without resolving the fast component may reflect only a trend, in particular, if the ultrashort lifetimes are much shorter than 10 ps and as long as the contributions of the slow components are small. Thus, 4% of excited DNA decay within 100 ps and less than 1% within 1 ns (Table 4.2) may constitute only a lower limit. Nevertheless, these small populations of “long components” together contribute almost 70% to the steady-state fluorescence spectrum.

T0

DNA concentration (K^+ concentration)	τ_1 [ns]	FI [%]	FA [%]	τ_2 [ns]	FI [%]	FA [%]	τ_3 [ns]	FI [%]	FA [%]	χ^2
32 (0 mM K^+)	2.1	49	0.7	0.15	20	3.9	0.010	31	95.4	1.028

T0

DNA concentration (K ⁺ concentration)	τ_1 [ns]	FI [%]	FA [%]	τ_2 [ns]	FI [%]	FA [%]	τ_3 [ns]	FI [%]	FA [%]	χ^2
32 (50 mM K ⁺)	1.0	62	14.4	0.44	35	18.5	0.010	3	67.1	1.058
32 (2 mM K ⁺)	1.1	86	12.8	0.21	9	7.1	0.010	5	80.1	1.019
0.5 (50 mM K ⁺)	1.3	91	12.4	0.11	4	6.3	0.010	5	81.3	1.053

T1

DNA concentration (K ⁺ concentration)	τ_1 [ns]	FI [%]	FA [%]	τ_2 [ns]	FI [%]	FA [%]	τ_3 [ns]	FI [%]	FA [%]	χ^2
32 (50 mM K ⁺)	1.3	50	9.8	0.57	47	20.8	0.010	3	69.4	1.090
32 (2 mM K ⁺)	1.3	78	10.8	0.48	17	6.3	0.010	5	82.9	1.074
0.5 (50 mM K ⁺)	1.3	59	4.4	0.49	22	4.4	0.020	19	91.2	0.997

T2

DNA concentration (K ⁺ concentration)	τ_1 [ns]	FI [%]	FA [%]	τ_2 [ns]	FI [%]	FA [%]	τ_3 [ns]	FI [%]	FA [%]	χ^2
32 (50 mM K ⁺)	1.1	70	8.2	0.31	23	9.0	0.010	7	82.8	1.176
32 (2 mM K ⁺)	1.2	48	1.1	0.07	15	5.3	0.010	37	93.6	1.057
0.5 (50 mM K ⁺)	1.2	14	2.1	0.39	70	3.5	0.010	16	94.4	1.011

Table 4.2: Fitting parameter (lifetime components (τ), fractional intensities (**FI**) and amplitudes (**FA**)) for the TCSPC decays of original **T0** and modified **T1**, **T2** sequences (excited at 260 nm, monitored at 430 nm) in 50 mM KCl and 5 mM Tris-HCl (pH=7.0) at different DNA and K⁺ concentrations.

As expected from the 385-nm fluorescence band for **T0** in 50 mM K⁺ solution, the amplitudes of long components are drastically increasing at the expense of the short 15-ps one (Figure 4.1-D). The 3-exponential fit (Table 4.2) resulted in an increase of the 1-ns component by a factor of 14 contributing now with 62% to the steady-state fluorescence. The middle component is similarly affected with the lifetime increasing by about a factor of 3 and the amplitude by almost a factor of 5. These long components together determine now 97% of the steady-state spectrum. Concomitantly, the contribution of the fast component decreases from 31% to 3%.

4.2. Correlation between the 385-nm fluorescence band and the formation of a 5'-5' stacked G-quadruplex dimer

Previously, we have shown that the monomer-dimer equilibrium of the 5'-5' stacking of propeller-type parallel G-quadruplexes (Figure 4.6-B) can be controlled by a number of parameters including the concentration of cations, DNA concentration, temperature, as well as the DNA flanking sequence (Do et al. 2011; Do et al. 2012). In this study, we use these parameters to manipulate the relative populations of the monomer and stacked dimer in order to investigate the relation between the new 385-nm emission and the stacking of G-quadruplex blocks. There are two equilibria to be considered: the first one pertains to the formation and stabilization of the monomer quadruplexes and the second one to the dimerisation of these monomeric units. K⁺ concentration is an important parameter that can be used to manipulate both equilibria with high K⁺ concentration favoring the G-quadruplex dimer (Do et al. 2012). The equilibrium controlling the quadruplex monomer rests on the coordination of K⁺ to the G-tetrads, while the factors influencing the second equilibrium are less understood. Previous studies (Do et al. 2011; Do et al. 2012; Kato et al. 2005; Martadinata et al. 2009; Mukundan et al. 2011; Wang et al. 1992) have shown that the addition of non-guanine residues to the flanking ends reduces the stacking propensity between the two G-quadruplex blocks, hence, shifts the equilibrium towards the monomer. In the following, we examine the development of the 385-nm fluorescence band as a function of the relative populations of the monomer and stacked dimer in a comparative study of **T0** along with the sequences **T1** and **T2** containing one and two thymines added at the 5' end of **T0**, respectively (Figure 4.6-A and Table 4.1).

Figure 4.6: Comparison of NMR, absorption and fluorescence spectra of the **T0** (blue) and two modified sequences, **T1** (green) and **T2** (magenta), in the presence of high (50 mM) and low (2 mM) K⁺ concentrations. (A) Sequences of **T0**, **T1**, and **T2**. (B) Equilibrium between monomeric G-quadruplex blocks and stacked G-quadruplex dimer. (C, C') Imino-proton NMR spectra (normalized to DSS protons). The imino-proton peaks of dimer and monomer are marked with a circle and cross, respectively. (D, D') Absorption and fluorescence spectra.

4.2.1. T0, T1, and T2 in the presence of high and low K⁺ concentration

NMR imino proton spectra of **T0**, **T1** and **T2** are presented together with their absorption and fluorescence spectra in Figure 4.6. The imino proton at ~10.9 ppm provides a marker for the G-quadruplex dimer, while the imino proton at ~11.7 ppm provides a marker for the G-quadruplex monomer. At high (50 mM K⁺) solution (Figure 4.6-C), the imino proton spectrum of **T2** shows both peaks at 10.9 and 11.7 ppm indicating the presence of both dimer and monomer, while the imino proton spectra of **T0** and **T1** show only the peak at 10.9 ppm characteristic for the stacked G-quadruplex dimer (Do et al. 2011; Do et al. 2012). The ratio between the areas below the two peaks represents the relative populations of the two species. For example, at 50 mM K⁺ the strand concentration of the dimer is comparable to that of the monomer in the case of **T2** (Figure 4.6-C). After confirming these structural features by NMR, the samples were loaded into a cuvette and their absorption and fluorescence were measured. Although there are slight differences in the absolute absorbance (Figure 4.6-D), upon normalization the absorbance of these samples is well overlapping (Figure 4.7-A). All samples exhibit a broad fluorescence emission band peaking around 385 nm, however, the emission intensity of **T0** and **T1** is twice of that of **T2** (Figure 4.6-D). These data reveal an excellent correlation between the intensity of the 385-nm emission band and the population of the stacked G-quadruplex dimer. If we take the maximum emission intensity of **T0** and **T1** as a reference for the fluorescence from 100% stacked dimer, then the emission of **T2** will be from ~ 50% stacked dimer, which is consistent with the NMR data.

Figure 4.7: Normalized absorption and fluorescence emission spectra of the **T0**, **T1**, and **T2** sequence in 5 mM Tris-HCl buffer in the presence of 2 different salt concentrations: (A) 50 mM K⁺; (A') 2 mM K⁺. Color coded as in Figure 4.6. The DNA concentration was ~0.2 mM (~3 OD in a 10 x 1 mm quartz cuvette with 1-mm path length).

At a low (2 mM) K^+ concentration, the NMR imino proton spectra in Figure 4.6-C' show for **T2** only the monomer peaks, for **T0** only the dimer peaks, while for **T1** the two sets of peaks of both monomer and dimer are observed. As expected, the 385-nm emission also varies consistently with the ratio between the dimer and monomer (Figure 4.6-D'): **T2** shows one emission band around 330 nm characterized for a monomeric G-quadruplex (Dao et al. 2011; Mendez et al. 2009; Miannay et al. 2009) **T0** shows an emission band at 385 nm with the same intensity as observed at high salt condition (50 mM K^+), consistent with 100% stacked G-quadruplex dimer formation. **T1** shows an emission band at 385 nm with approximately half of the emission intensity of **T0**, consistent with 50% dimer formation. Normalized fluorescence spectra from these samples clearly show a transition from monomer to dimer, with the peak shift from 330 nm for **T2** towards 385 nm for **T1** and **T0** (Figure 4.7-A'). Notice that, besides the main emission band at 385 nm, a shoulder appears in the fluorescence spectrum of **T1** at 2 mM K^+ (Figure 4.7-A') and **T2** at the high 50 mM K^+ concentration (Figure 4.7-A). As in the concomitant NMR spectra, this fluorescence feature points to a mixture of both monomer and stacked G-quadruplex dimer.

The dependence of excited-states dynamics on K^+ concentration has been probed by TCSPC (Figure 4.8). In case of the easily dimerizing sequences **T0** and **T1**, neither the spectral shape of the 385-nm fluorescence band, nor the lifetime nor amplitude of the long 1-ns component are significantly affected. In contrast, the contribution of the middle component decreases as the lifetime shortens at a smaller fractional amplitude when the K^+ concentration decreases (Table 4.2). Concomitantly, the amplitude of the shortest component is increasing. We also note that only in the case of **T2**, where a decrease of the K^+ concentration inhibits dimerization (Figure 4.6-D'), the amplitude of the long component decreases by a factor of 8 and its contribution to the steady-state fluorescence drops from 70% (in 50 mM K^+) to 48% (in 2 mM K^+).

Figure 4.8: Fluorescence decays of **T0**, **T1**, and **T2** in 5 mM Tris-HCl buffer in the presence of (A) 50 mM K^+ and (B) 2 mM K^+ , excited at 260 nm and probed at 430 nm as encoded in Figure 4.6. The instrument response function (IRF) of the TCSPC set-up was measured using the Raman emission of water at 285 nm. Colors coded as in Figure 4.6. The DNA concentration was ~ 0.2 mM (~ 3 OD in a 10 x 1 mm quartz cuvette with 1-mm path length).

4.2.2. T1 and T2: Following stepwise G-quadruplex dimer formation by parallel monitoring NMR and fluorescence spectra

The experiments described above show an obvious relation between the dimer formation measured by NMR and the 385-nm emission of **T0**, **T1**, and **T2** (Figure 4.6). In order to quantitatively confirm the correlation between the stacked dimer formation and the 385-nm emission, we performed a titration experiment by adding K^+ aliquots to the solutions containing either the **T1** or **T2** sequences.

As increasing amounts of K^+ were added, NMR imino proton spectra show the formation of the G-quadruplex monomer, followed by the gradual transformation into the stacked dimer (Figure 4.9-A,A'). Fluorescence spectra parallel recorded show initially the 330 nm emission band of the monomer that is followed by a rising band at 385 nm (Figure 4.9-B,B'). The ratio between the areas of the two imino proton peaks, one at 10.9 ppm characteristic for the monomer and one at 11.7 ppm for the dimer, allowed us to estimate the relative amount of the dimer. Such an estimate is less feasible on the basis of the concomitant fluorescence spectra: the spectra of the monomer and dimer overlap and contribute to the steady-state fluorescence according to their decay-associated spectra. Nevertheless, the titration experiments show that the increase in steady-state fluorescence occurs qualitatively in parallel with the development of peaks for the dimer in the NMR spectra. Due to the complexity of the fluorescence spectra involving superimposed features of monomer and dimer emissions and the interplay of the different decay components depending

on details of the experiment, any quantitative evaluation has to wait for the decay-associated fluorescence spectra at different K^+ concentrations and temperature and time-resolved measurements of fluorescence anisotropy.

Figure 4.9: (A, A') Imino proton NMR and (B, B') fluorescence spectra from salt titration experiment of **T1** and **T2**. The imino proton at 10.9 ppm (cross) is a marker for the dimer, while the imino proton at 11.7 ppm (circle) provides a marker for the monomer.

4.2.3. T0, T1, and T2 at low DNA concentration

In order to study the fluorescence property of the 385-nm emission formed by stacking of two G-quadruplex blocks while maintaining all the structural information at atomic level obtained by NMR, most of DNA samples were used so far are at high-concentration ($\sim 0.25 \text{ mM}$). Although a very short penetration depth of excited light into the DNA sample due to its high concentration (in front-face configuration) results in no requirement for IEF correction and the re-absorption of fluorescence are minimal due to the negligible overlap of absorption and fluorescence spectra, it is obvious that one might questions whether there is any difference if fluorescence

spectra were measured in a typical condition: lower DNA concentration ($<5\ \mu\text{M}$) at right-angle geometry. Moreover, since we suggested the new 385-nm emission caused by stacking of G-quadruplex blocks, it is also expected that this emission should depend on the DNA concentration. To address these questions, we have performed a systematic fluorescence measurement of the **T0** sequences in low DNA condition, right-angle geometry.

In 50 mM K^+ solution, the formation of parallel G-quadruplexes at low ($4\ \mu\text{M}$) DNA concentrations from all three sequences **T0**, **T1**, and **T2** is supported by CD spectra (Figure 4.10-B). Their absorption spectra are invariant (Figure 4.10-A) and do not depend on DNA concentrations in the range studied (Figure 4.7-A). In contrast, their fluorescence spectra are significantly different (Figure 4.11-A). **T0** shows the strongest emission at 385 nm with the same profile as the one at high DNA concentration, **T2** shows only the emission at 330 nm characterized for the G-quadruplex monomer, and **T1** shows two emission bands at 330 and 385 nm, characterized for a mixture of the monomer and dimer. Normalized fluorescence spectra clearly point to the transition from monomer to dimer, with a peak shift from 330 nm for **T2** to 385 nm for **T0** (Figure 4.10-A).

The complexity of steady-state fluorescence spectra becomes evident when we compare the fluorescence decay pattern at high and low DNA concentrations (Table 4.2). For **T0**, the lifetime and amplitude of the 1-ns component stay constant, while the contribution of the 100-ps component decreases favoring conversely the short, unresolved 10-ps component. Comparing the fluorescence decay pattern from the three sequences (Figure 4.11-B) shows the slowest decay for **T0**, followed by that of **T1** and **T2**. This trend is also reflected in the fitting results (Table 4.2). These observations are consistent with the data at high DNA concentration. However, we are aware that this comparison has to be taken with caution since – for the sake of comparability – the fluorescence decay pattern has been measured always at 430 nm in the red wing of the emission spectrum.

Figure 4.10: (A) Normalized absorption, fluorescence emission (upon IFE correction) and (B) CD spectra of the **T0**, **T1**, and **T2** sequences at low DNA concentration ($\sim 4 \mu\text{M}$) in a buffer containing 50 mM K^+ . Color coded as Figure 4.6.

Figure 4.11: (A) Absorption and fluorescence spectra (IFE corrected) of **T0**, **T1**, and **T2** at low DNA concentration. (B) Fluorescence decay excited at 260 nm and probed at 430 nm. DNA concentration was $4 \mu\text{M}$. Solution contained 5 mM Tris-HCl (pH 7.0) and 50 mM KCl.

4.3. The role of base stacking overlap in the 385-nm fluorescence band

It is known that various stacking modes are possible at the stacking interface between two G-quadruplex blocks (Lech et al. 2012). In this section, we examine the effect of different stacking modes on the 385-nm fluorescence emission band. NMR was first used to verify the formation of G-quadruplex structures that have been

shown previously to exhibit different stacking modes at the interface between G-quadruplex blocks. Subsequently, their fluorescence spectra were recorded.

The sequence d[G₄TG₃AG₂AG₃T] (named **93del**) has been shown to form a robust dimeric structure in K⁺ solution (Figure 4.12-B), where two parallel G-quadruplex subunits are interlocked with each other (Phan et al. 2005). At the interface between the two G-quadruplex blocks, the G-tetrads stack with an overlap of the 5-member rings of guanine bases (Figure 4.13-B). Here the formation of this interlocked G-quadruplex for **93del** was confirmed by the observation of the same NMR spectrum (Figure 4.14-A) as previously reported. (Phan et al. 2005) In contrast to the 385-nm emission band for **T0**, we only observed an emission band around 330 nm for **93del** in 50 mM K⁺ solution (Figure 4.15-A).

Figure 4.12: Ribbon view of a representative structure of stacked or interlocked G-quadruplexes formed by different sequences: (A) **T0**, (B) **93del**, (C) **TERRA** and (D) **GGA4**. Guanine bases at the interfaces between two G-quadruplex blocks are colored as follows: **T0**-green, **93del**-orange, **TERRA**-red, **GGA4**-green, while the remaining guanine bases are colored cyan. Backbone and sugar atoms are colored gray.

Figure 4.13: Different overlaps of guanine bases at the stacking or interlocking interface of the G-quadruplexes formed by different sequences: (A) **T0**, (B) **93del**, (C) **TERRA** and (D) **GGA4**. Colors are coded as in Figure 4.12.

Figure 4.14: NMR imino proton spectra of (A) **93del** in 50 mM K^+ , (B) **telo** in 50 mM K^+ , (C) **TERRA** in 90 mM K^+ and (D) **GGA4** in 200 mM K^+ . The same characteristics of these NMR spectra with ones reported before confirm the formation of the same G-quadruplexes as previously determined for these sequences. The DNA concentration was ~ 0.2 mM.

Figure 4.15: Normalized fluorescence spectra of various structures formed by stacking or interlocking of two parallel G-quadruplex blocks: (A) **T0** and interlocked G-quadruplex **93del**; (B) **T0** and 3'-3' stacked G-quadruplex **telo**; (C) **T0** and 5'-5' stacked G-quadruplexes **TERRA** and **GGA4**. Note that here we only focus on the emission peaks rather than their different intensities, which might be investigated further through decay-associated fluorescence spectra.

Two tetrameric parallel G-quadruplexes formed by the sequence d(TTAGGG) (named **telo**) have been shown to stack on their 3' end (Kato et al. 2005; Wang et al. 1992). The observed NMR spectrum of **telo** (Figure 4.14-B) confirmed the formation of this 3'-3' stacked G-quadruplex. Our recent modeling (Lech et al. 2012) suggests that at the 3'-3' interface between the two parallel G-quadruplex blocks, the G-tetrads stack with an overlap of the 6-member rings of guanine bases (Figure 4.16). Again, for this sequence **telo** we observed only an emission band around 330 nm in K⁺ solution (Figure 4.15-B).

Figure 4.16: Illustrative examples of base stacking modes in the core (A-C) and the interface (D-E) of stacked G-quadruplex with their PDB ID codes. The stacking modes (D) and (E) were observed only in crystal. This figure is adapted from reference (Lech et al. 2012).

Subsequently, we examined two different 5'-5' stacked G-quadruplexes, one formed by the human telomeric RNA sequence r(GGGUUAGGGU) (named **TERRA**) (Martadinata et al. 2012) (Figure 4.12-C) and the other by the DNA sequence d(GGA)₄ (named **GGA4**) (Matsugami et al. 2001) (Figure 4.12-D). The formation of these 5'-5' stacked G-quadruplexes are confirmed by NMR spectra (Figure 4.14-C and 4.14-D). In the case of **TERRA**, at the interface between the two G-quadruplex blocks, the G-tetrads stack with an overlap of the 5-member rings of guanine bases (Figure 4.13-C), in the case of **GGA4**, the G-tetrads stack with an overlap between

the 5- and 6-member rings of guanine bases (Figure 4.13-D). Even though both structures are 5'-5' stacked parallel G-quadruplexes as that of **T0**, our results show that only **GGA4** exhibits the 385-nm emission, while **TERRA** only shows an emission band around 330 nm (Figure 4.15-C).

4.4. Assignment of the 385-nm band to excimer emission

The highlight of the present work is a firm establishment of the correlation between the formation of a stacked G-quadruplex dimer from the sequence d[(G₃T)₄] named **T0** and the development of a strong fluorescence band peaking at 385 nm. Fluorescence features of oligonucleotides in the 400-nm region are not new, as they have been reported for oligomers of all four nucleobases (Plessow et al. 2000; Stuhldreier et al. 2010; Vayá et al. 2012a). What is new is that the intensity of this 385-nm fluorescence band from a stacked G-quadruplex dimer can be controlled in NMR spectra and manipulated to the extent that it is possible to safely assign the fluorescence to excimer emission. This NMR structure supports the notion of excimer emission. It reveals a G-quadruplex dimer, in which the two G-tetrads at the 5'-end interface of the monomeric quadruplex subunits are stacked in a specific way at van der Waals distance. Upon dimer formation the fluorescence quantum yield $Q_{\text{T0dimer}} = 1.5 \times 10^{-3}$ increases by a factor of ≈ 15 as compared to Q_{T0} in bare Tris-buffer, i.e. in the absence of the quadruplex-stabilizing K⁺ ions. The spectroscopic criteria pointing to excimer fluorescence are (i) the broad, red-shifted fluorescence band, excited within the envelope of the absorption band that is independent of dimer formation and (ii) a long lifetime component of 1 ns that dominates the integrated steady-state emission. We speculate that G-G dimers held together across the stacking interface between two G-quadruplex blocks may reach optimal conformation for excimer formation by meticulously small tilting of angles or lateral displacements within the “pre-formatted” guanine pairs.

While the NMR spectra identify the concentrations of monomer and dimer quadruplexes by separated imino proton resonances, a quantitative comparison of NMR and integrated fluorescence spectra is complicated by the fact that the fluorescence spectra of the quadruplex dimer and monomer are partially overlapping. In contrast, loosely bound excimers and other disordered structures may explicitly show up in the fluorescence decay pattern and contribute to the steady-state fluorescence spectrum. In the 3-exponential fit of the fluorescence decay, there is a

long component with a lifetime of 1 ns and a short one with 10 ps. The latter is subsumed within the instrument function of the TCSPC experiments and giving therefore most probably only an upper limit. In addition, there is a middle component in the 100-500-ps range that contributes significantly to the steady-state fluorescence spectrum. An interesting phenomenon is that the amplitude of this middle component is most sensitive to external conditions as its amplitude decreases drastically upon lowering the DNA or the K^+ concentration. These trends described for the sequence **T0** are maintained also for the related sequences **T1** and **T2**, which carry one or two thymines flanking the 5'-end G-tetrads, known from NMR experiments for shifting the monomer-dimer equilibrium to the monomer side.

We speculate that the sub-nanosecond components might be caused by all kinds of geometries and interactions including those within the tetrads not involved in the stacking at their 5' interface and loosely bound dimer configurations that give rise to excimer-type fluorescence. In fact, there are situations of molecular pairs in polymer and crystalline environments where experimental findings prompted the extension of the excimer concept from geometry-optimized to loosely bound excimers. Recently (Yoo et al. 2010), excimer-like states in intramolecularly coupled perylenediimides in polymer films have been identified. Resolving their dynamics and spectral signature in single molecule fluorescence microscopy revealed that the spread of fluorescence lifetimes is corresponding to the conformational inhomogeneity of the system. Similarly, in the nineteen seventies when molecular crystals research was culminating, the existence and properties of different excimer-like structures have been explored in a perylene crystal. In the case of the α -modification of the perylene that bears two pairs of aromatic molecules in coplanar arrangement at 3.4 Å separation in its elementary cell, an unusual type of emission at higher energies occurs at low temperature in addition to the usual broad excimer fluorescence band. This new emission has been identified as a variant of the geometry-optimized excimer state. In this variant owing to a lattice distortion the two perylene molecules forming the pair cannot achieve the optimum geometry in the excited state. Conversely, the binding energy of the loose complex corresponds to $\approx 800\text{ cm}^{-1}$, whereas the binding energy of the geometry-optimized excimer state corresponds to $\approx 2700\text{ cm}^{-1}$ (Cohen et al. 1978; Vonnfreidorf et al. 1978).

4.5. Is there a structural detail responsible for the formation of geometry-optimized excimers?

Putting different G-quadruplex dimers to NMR and fluorescence test, we showed that for geometry-optimized excimers the pattern of base-stacking overlap at the G-quadruplex interface is critical. Stacking and interlocking between G-quadruplex blocks has been observed (Do et al. 2011; Do et al. 2012; Kato et al. 2005; Krishnan-Ghosh et al. 2004; Martadinata et al. 2009; Matsugami et al. 2001; Mukundan et al. 2011; Sket et al. 2010; Smargiasso et al. 2008; Wang et al. 1992) and might contribute to higher-order structure formation of telomeric DNA and RNA (Marsh et al. 1994; Martadinata et al. 2009; Phan 2010). In order to understand how G-quadruplex stacking can lead to an excimer emission, it is important to consider the detailed structures of base stacking in G-quadruplexes (Lech et al. 2012). A G-quadruplex contains a core formed by stacking of several G-tetrads and supported by four backbone strands. There are three stacking modes (Figure 4-16-(A-C)) found in the G-tetrad core involving (i) an overlap of 5-member rings (termed “5-ring” mode), (ii) a partial overlap of 6-member rings (termed “partial 6-ring” mode), and (iii) a partial overlap between the 5- and 6-member rings of guanine bases (termed “partial 5/6-ring” mode) (Lech et al. 2012). The geometries of the stacked guanines within the G-tetrad core are constrained due to the covalent backbone linkage between adjacent nucleotides. Alternatively, the stacking interface between G-quadruplex blocks is more flexible and can lead to stacked guanines that adopt a wider range of geometries. Four major stacking modes depicted in Figure 4.16-(D-G) were observed at the interface involving (i) a partial overlap of 6-member rings (termed “partial 6-ring”), (ii) an overlap of 6-member rings (termed “6-ring”), (iii) an overlap of 5-member rings (termed “5-ring”), and (iv) an overlap between the 5- and 6-member rings of guanine bases (termed “5/6-ring”) (Lech et al. 2012). Among all these stacking modes, the latter (“5/6-ring”) involves the largest overlap between guanine bases.

Figure 4.17 shows the structure of the **T0_I2** sequence (Do et al. 2011), an inosine-modified equivalent of **T0**, as an example of a G-quadruplex with the “5/6-ring” interface stacking mode. In this structure, there is little overlap of between aromatic rings of guanines in the G-tetrad core of a monomer G-quadruplex block (blue to cyan, Figure 4.17-B), while there is a large overlap of the 5- and 6-member rings of guanines at the stacking interface between two blocks (blue to blue, Figure 4.17-C). This might explain why the excimer emission is only observed when the stacked G-quadruplex dimer is formed.

Figure 4.17: (A) Ribbon view of the representative dimeric parallel G-quadruplex structure of **T0-I2**, an inosine-modified equivalent of **T0**, in K^+ solution (PDB code, 2LE6). Guanines in the stacking layer are highlighted in blue, all other guanines are colored cyan. Backbone and sugar are colored gray. Top view of two stacked G-tetrad layers within one monomer (B) and between two monomers (C). O6 atoms are colored brown.

Examining the base overlap patterns at the interface of different structures formed by stacking or interlocking of two parallel G-quadruplex blocks studied in Fig. 5 shows that the 385-nm emission band is observed only when there is the interface stacking of the “5/6-ring” mode is present. The 5'-5' stacked G-quadruplex formed by the TERRA sequence is an example showing that 5'-5' G-quadruplex stacking is not a sufficient condition for observing the 385-nm excimer emission.

Nevertheless, the stacking modes observed experimentally by X-ray and NMR techniques represent only the most abundant and/or average structure at the stacking interface. The two G-quadruplex blocks might be flexible and adopt a range of conformations around these modes. Geometry-optimized excimers together with loosely bound excimer-like state could occur within this range of conformations at the stacking interface.

4.6. Conclusion

In parallel NMR and fluorescence spectroscopy using identical DNA samples of a stacked G-quadruplex dimer, the red-shifted 385-nm fluorescence band is shown to originate from G-G excimers formed at the interface of two stacked 5'-end G-tetrads. These geometry-optimized excimers decay on the time scale of 1 ns. The fluorescence quantum yield increases by a factor of 15 upon addition of K^+ ions that stabilize the G-quadruplex monomer and dimer structures. Although we are attributing the 385-nm emission band to an excimer formed by two stacked guanines, we are well aware that charge transfer states may be energetically close to the $\pi\pi$ -excited states (Lange et al. 2009), so that the term *exciplex* for the emitting state may be appropriate. Nevertheless, we keep the term *excimer* since this paper emphasizes the structural aspects rather than the mechanistic ones. For the formation of geometry-optimized excimers the base-stacking overlap at the G-quadruplex interface has been shown to be critical. The most favorable overlap condition is the one where the 5- and 6-membered rings of the guanine bases overlap.

Chapter 5: Conformation and dynamics of G-quadruplexes studied by single-molecule FRET*

As discussed in the Introduction, G-quadruplexes are polymorphic. In ensemble measurements, various G-quadruplex structures have already been reported for human telomeric DNA sequences (Phan 2010). At the single-molecule level, human telomeric DNA sequence $d[(G_3TTA)_3G_3]$ studied by smFRET (Lee et al. 2005) has revealed its extreme conformational diversity and dynamics (3 FRET values reported, which are assigned to one unfolded state and two folded states, and each state has both long-lived and short-lived species) at low salt concentration (2 mM K^+). At high salt concentration (100 mM K^+), there are still two stable folded states as seen from distinct FRET values (Lee et al. 2005; Ying et al. 2003) and they can interconvert on a time scale of a minute.

There are only a few smFRET studies on G-quadruplexes reported so far. Single G-to-T base mutations, with the sequence $d[(G_3TTA)_2GTGTTAG_3]$ (Lee et al. 2005; Lee et al. 2009); and $d[(G_3TTA)_2TGTTAG_3]$ and $d[(G_3TTA)_2GGTTTAG_3]$ (Lee et al. 2009) have been shown to induce various folded structures (as observed by various FRET values). The position of single base mutation also affects the stability of G-quadruplexes: a mutation of a guanine from the central G-tetrad leads to G-quadruplexes that unfold faster at all tested K^+ concentration and temperature, while a mutation of a guanine at an outer G-tetrad induces heterogeneous dynamic at intermediate K^+ concentration (10 mM K^+) and leads to a long-lived folded state at

* The work in this chapter is being written for publication:

“Effect of the loop length on conformation and stability of parallel vs. antiparallel G-quadruplexes study by single molecule FRET”, Dao, N. T. and Phan, A. T.

high K^+ concentration (300 mM K^+). There is also an attempt of using a modified guanine (^{Br}G) substitution (Okamoto et al. 2008) at different positions in order to control the G-quadruplex folding topology in the human telomeric DNA, taking advantage of the property of the ^{Br}G to favor syn conformation. However, these substitutions still lead to a mixture of different conformations as seen from smFRET measurements; and the validity of the proposed folding pathway and the novel “triple-strand-core model” are still uncertain. A smFRET study has been also reported for the *Oxytricha* telomeric sequence $d[(G_4T_4)_3G_4]$ (Lee et al. 2008). G-quadruplexes formed by this sequence show an extreme stability (even at 2 mM K^+) as compared to those formed by the human telomeric DNA sequence (Lee et al. 2005). However, the results from the *Oxytricha* G-quadruplex $d[(G_4T_4)_3G_4]$ should not be directly compared with those of the human telomeric DNA G-quadruplexes $d[(G_3TTA)_3G_3]$, because these two sequences differ from each other by the number of guanines in the G-tracts as well as the length of their linker sequences.

So far, smFRET studies of G-quadruplexes have not been supported by high-resolution structural data. A single FRET peak can correspond to a mixture of multiple structures. In this chapter, we combine extensively the smFRET technique and various ensemble measurements including NMR spectroscopy to study the structure, dynamics and stability of two distinct well-defined G-quadruplex folds. This combination provides both atomic-resolution and single-molecule views on the molecules studied.

5.1. Experimental procedures

We first immobilized DNA molecules on a quartz surface in order to monitor their FRET signal overtime. Here, we used a “sandwich layer” schematic (PEG-biotin – Neutravidin – Biotin) to anchor DNA molecules on the aminosilanized quartz surface, following a well-established protocol by TJ Ha’s group (Selvin et al. 2007). The molecules of interest were constructed from hybridization of two DNA strands (Figure 5.2-A): (1) the first strand included Cy5 at the 5'-end (red color), the G-quadruplex forming sequence of interest ($d[(G_3TTA)_3G_3]$, $d[(G_3T)_3G_3]$, or $d[(G_3T_4)_3G_3]$) (black color), a 29-nt stem sequence (orange color) and Biotin at the 3'-end (blue color); (2) the second strand included the complementary stem sequence (orange color) with TMR (green color) in the middle. The first and the second strands were annealed at the 1:2 ratio by heating up to 85 °C and cooling slowly to room temperature. A sample chamber (up to 5 channels) was made by stacking a quartz slide (with drill holes) to a coverslip separated by one or two layers double-sided tape. The sample chamber was then sealed with epoxy. Detailed protocols for cleaning of the quartz slides and chamber fabrication are described in Appendices A and B. Protocols for preparation of coated PEG-biotin on the quartz slide’s surface are presented in Appendix C.

A mixture of oxygen scavenger system (0.8% w/v D-glucose, 1 mg/ml glucose oxidase, 0.04 mg/ml catalase) and triplex state quencher (Trolox, saturated dispend in de-ionized water) was used to reduce photo-bleaching and blinking associated with triplex states energy transfer which have been reported for Cy dyes (especially Cy5) (Zhuang et al. 2000), and to enhance the photo-stability of the dyes (Rasnik et al. 2006).

The following steps are usually followed in order to achieve optimal condition of single-molecule immobilization on the quartz surface for smFRET study:

1. Flush 100 µl of the T50 buffer to each channel to clean the channel.
2. Flush in 50 µl of (1/2500) of Neutravidin stock solution and incubate for 5 minutes.
3. Flush in 100 µl T50 to clean the residual Neutravidin molecules which are still floating in the solution, leaving only those attached to the Biotins on the surface.
4. Flush in 50 µl of 20 – 50 pM DNA sample (diluted in the T50 buffer) and incubate for 5 minutes

5. Flush in 100 μ l T50 in order to clean the rest of the DNA molecules that didn't attach to the surface.
6. Check smFRET to see whether there are enough spots ($\sim 300 - 400$ spots) on the surface. If there are not enough spots, repeat step 4 and 5.
7. Flush in 100 μ l T0-Tris (20 mM Tris, pH 8.0, no salt) that can clear the ions on the channels (for example Na^+ in the T50 buffer).
8. Flush 50 μ l imaging buffer consisting of the oxygen scavenger system and Trolox in appropriate buffer (5 mM Li^+ , 2 mM K^+ , 10 mM K^+ ...) into the channel.
9. Acquire at least ~ 15 sets of movie data with 0.1 second integration time and $\sim 100 - 200$ seconds duration for histogram and time trace analysis.
10. Repeat step 7 – 9 for different buffer conditions (different cations, different solvents...). Note that the imaging buffer becomes acidic overtime, so it is recommended to clean the imaging buffer in the channel by T0-Tris after data acquisition.

Figure 5.1: Differences between 2 smFRET channels: (A) DNA in imaging buffer vs. (B) imaging buffer without DNA confirms most of the bright spots on (A) are real DNA molecules.

Figure 5.1-A shows an image of DNA molecules immobilized on the surface captured by the CCD camera. Continuous acquisition of data generates a movie with 0.1 second resolution which actually is a stack of images captured at each 0.1 second. A FRET histogram can be obtained by averaging the first 10 images of 15 –

30 sets of movie data. The FRET efficiency is corrected for emission leakage and the variable quantum yield property of the 2 fluorophores as described in Appendix D. Background subtraction is applied as previously reported (Ha 2001). Note that some spots on acceptor channel have very weak signals which will result in negative emission intensities after background subtraction, thus leading to negative values of FRET efficiency. The zero peak in the FRET histogram is due to inactive or missing acceptor molecules and is not included in analyses. A negative control channel with only imaging buffer (step 1 – 10 but skip step 4) is used to check how clean the channel is (Figure 5.1-B). A clear contrast between two channels: one with DNA vs. one without DNA (Figure 5.1) confirms that most of the observed spots (Figure 5.1-A) are real DNA molecules, not junk molecules or impurities.

5.2. smFRET analysis of human telomeric G-quadruplexes: comparison with previously reported works

Using the constructed smFRET system, we first performed an analysis on the human telomeric G-quadruplex DNA sequence $d[(G_3TTA)_3G_3]$ and compared our data with the previously published data from TJ Ha's group (Lee et al. 2005). The consistency of all the results (see below) validates our smFRET system.

5.2.1. Observation of 3 FRET values: one unfolded and two folded states

Figure 5.2 shows the result of smFRET study of human telomeric DNA sequence $d[(G_3TTA)_3G_3]$ obtained in our system. We observed changes in the FRET efficiency histogram as the K^+ concentration increases (Figure 5.2-C). The FRET histogram at 2 mM K^+ (Figure 5.2-D) is well fitted by the sum of Gaussian functions with 3 peaks: one at low FRET value (~ 0.3), another at intermediate FRET value (~ 0.6) and the last one at high FRET value (~ 0.75). These values are well matched with the FRET values reported for one unfolded state **U** and two folded states **F1**, **F2** of this sequence at the same condition in previous study (Lee et al. 2005). At 5 mM Li^+ , most of the molecules are at unfolded state **U**. As more K^+ are added in the solution, the low FRET state **U** (~ 0.3) decreases and the high FRET state **F2** (~ 0.75) increases. At very high K^+ concentration (500 mM), the peak in FRET histogram shifted to **F2** state. These observations are also consistent with the previously reported data (Lee et al. 2005).

Figure 5.2: smFRET analysis of human telomeric G-quadruplex DNA. (A) Schematic diagram of the DNA construct (with its sequence on the right). (B) Image of the DNA immobilized on quartz surface captured by the CCD camera. (C) Single molecule FRET histograms of the DNA at different of K^+ concentrations. (D) Expanded view and Gaussian fitting of the histogram of the DNA in 2 mM K^+ , revealed 3 states named **U** (unfolded), **F1** and **F2** (folded). (E) Time trace of donor, acceptor and corresponding FRET efficiency of a single molecule at 2mM K^+ monitored over 220 seconds, showing interconversion between **U**, **F1** and **F2**.

5.2.2. Single molecule FRET time traces reveal state transitions always through unfolded (U) state

We could also monitor transitions between all three FRET states within single molecules by monitoring the emission intensities of the molecule's spots on both channels. Figure 5.2-E shows one example of the time traces of donor (green), acceptor (red) and the corresponding FRET efficiency (blue) of a single molecule at 2mM K^+ monitored over 220 seconds, reveals transitions between **F1** and **F2** states

passed through **U**. More examples of representative single molecule FRET time traces are shown in Figure 5.3. Some of the molecules can adopt single conformation in almost hundred of seconds before Cy5 is photo-bleached (Figure 5.3-D), while others can be extremely dynamic, switching between **F2** and **U** in just a few seconds (Figure 5.3-B, C, E, F). Over all molecules that have been investigated, transitions between **F1** and **F2** always have to go through **U** state. At high salt concentration (500 mM K^+), more stable **F1** and **F2** states are observed from single molecule time traces (Figure 5.4). These **F1** and **F2** states can be stable throughout the acquisition period (Figure 5.4-A,C), or transit to **U** state, for example **F1** to **U** in Figure 5.4-D and **F2** to **U** in Figure 5.4-B. We note that there is no fast switching between **F1**, **F2** and **U** observed in this high K^+ concentration as had been observed in low K^+ concentration.

These results are also consistent with previous report (Lee et al. 2005).

Figure 5.3: Examples of FRET time traces from single molecules of human telomeric G-quadruplex DNA at 2 mM K^+ .

Figure 5.4: Examples of FRET time traces from single molecules of the human telomeric sequence $d[(G_3TTA)_3G_3]$ in high salt condition: 500 mM K^+ .

5.3. Conformation and stability of parallel and antiparallel G-quadruplexes studied by smFRET

After successfully constructing and validating the single molecule FRET system, we are interested in using the system to study G-quadruplexes. We examine by ensemble and single-molecule techniques various sequences containing the fragments $d[(G_3T)_3G_3]$ and $d[(G_3T_4)_3G_3]$. These sequences have the same number of Guanines, therefore they are expected to form the same number of G-tetrad layers. The only difference between these sequences is the number of Thymine in the loops. The former one has one Thymine in each loop, while the latter one has four Thymines in each loop. The former has been shown to adopt a parallel G-quadruplex with three propeller loops (Do et al. 2011) (Figure 5.5-A), while the latter is assumed to adopt an antiparallel G-quadruplex with one diagonal and two edgewise loops (Figure 5.5-B) based on previous reports on analogous sequences (Smith et al. 1993). For single-molecule measurements, a duplex stem was added to these G-quadruplex-forming fragments. The effect of this addition on the G-quadruplex formation has been investigated by NMR spectroscopy.

Figure 5.5: Proposed schematic folding models of (A) a parallel G-quadruplex with 3 propeller loops formed by $d[(G_3T)_3G_3]$ and (B) an antiparallel G-quadruplex with one diagonal and two edgewise loops formed by $d[(G_3T_4)_3G_3]$.

5.3.1. Ensemble measurements

Figure 5.6: CD, CD and UV melting and of 2 G-quadruplex forming sequences $d[T_2(G_3T)_3G_3]$ – (red curves) and $d[T_2(G_3T_4)_3G_3]$ – (black curves) in different K^+ concentration. (A) CD spectra in 10 mM K^+ indicate parallel (red) and antiparallel (black) G-quadruplexes formation by these 2 sequences. (B) Normalized spectra from UV melting of $d[T_2(G_3T_4)_3G_3]$ – (black) in 10 mM K^+ and CD melting of $d[T_2(G_3T)_3G_3]$ – (red) in 1 mM K^+ . The fill and open symbols represent heating and cooling processes. Usually, UV melting is preferred because of its low cost and convenience. However, the 295 nm absorbance signal of the antiparallel structure shows very little change in UV melting, therefore CD melting at 295 nm is used for $d[(T_2(G_3T_4)_3G_3)]$ sequence. Note that the CD melting data of $d[T_2(G_3T)_3G_3]$ in 1 mM K^+ (figure 5.4-B) is provided by one of our group's member, Li Zhe.

Figure 5.6-A shows CD spectra of these two sequences in 10 mM K^+ buffer. As expected the CD spectrum of $d[T_2(G_3T)_3G_3]$ has a signature of a parallel G-

quadruplex conformation with a negative peak at 245 nm and a positive peak at 260 nm, while the CD spectrum of $d[T_2(G_3T_4)_3G_3]$ shows a signature of an antiparallel (2 + 2) G-quadruplex conformation with a positive peaks at 295 nm a negative peak at 260 nm.

We also investigate the stability of these 2 sequences by UV and CD melting. The melting temperature of $d[T_2(G_3T)_3G_3]$ is significantly higher than that of $d[T_2(G_3T_4)_3G_3]$. The melting temperature of $d[T_2(G_3T)_3G_3]$ at 1 mM K^+ is 75 °C (Figure 5.6-B – red curve), while the melting temperature of $d[T_2(G_3T_4)_3G_3]$ at 10 mM K^+ is only 34 °C (Figure 5.5-B – black curve).

In order to figure out whether the duplex stem would affect the formation of the G-quadruplexes, we performed NMR experiments on 2 sequences, $d[T_2(G_3T)_3G_3AGAGGTAGCATACCTCT]$ and $d[T_2(G_3T_4)_3G_3AGAGGTAGCATACCTCT]$, which consist of the G-quadruplex forming fragments, $d[T_2(G_3T)_3G_3]$ and $d[T_2(G_3T_4)_3G_3]$, and a duplex hairpin forming fragment, $d[AGAGGTAGCATACCTCT]$.

Figure 5.7: Imino-proton NMR spectra and proposed schematic folding models of (A) the duplex hairpin d[AGAGGTAGCATACCTCT]; (B) G-quadruplex forming sequences d[T₂(G₃T)₃G₃]; and (C) (G-quadruplex + hairpin) complex d[T₂(G₃T)₃G₃AGAGGTAGCATACCTCT] in 10 mM K⁺ buffer. Circles and cross are denoted for duplex and G-quadruplex peaks, respectively.

The imino proton spectrum of d[AGAGGTAGCATACCTCT] shows six peaks from 12 to 14 ppm confirming the formation of a duplex hairpin (Figure 5.7-A). The imino proton spectrum of d[T₂(G₃T)₃G₃] with 12 peaks at 10-12 ppm is similar to the previously reported spectrum of a propeller-type parallel G-quadruplex adopted by the same sequence (Do et al. 2011). For the first G-quadruplex-hairpin construct, the imino proton spectrum of the d[T₂(G₃T)₃G₃AGAGGTAGCATACCTCT] sequence (Figure 5.7-C) resembles to the sum of the spectra of the two components, G-quadruplex (Figure 5.7-B) and duplex hairpin (Figure 5.7-A), suggesting that the G-quadruplex-duplex hairpin structure (Figure 5.7-C, right side) is formed as expected. A small hump in NMR spectrum (Figure 5.7-C) indicates the existence of minor conformation(s), which are less than 5% of the population.

Figure 5.8: Imino NMR spectra and proposed schematic folding models of (A) the duplex hairpin $d[AGAGGTAGCATACCTCT]$; (B) G-quadruplex forming sequences $d[T_2(G_3T_4)_3G_3]$; and (C) (G-quadruplex + hairpin) complex $d[T_2(G_3T_4)_3G_3AGAGGTAGCATACCTCT]$ in 10 mM K^+ buffer.

The imino proton spectrum of $d[T_2(G_3T_4)_3G_3]$ shows peaks at 10-12 ppm indicating the G-quadruplex formation. An antiparallel G-quadruplex topology can be proposed for this sequence based on the previous structural results for the intramolecular G-quadruplex of $d[(G_4T_4)_3G_4]$ (Wang et al. 1995) and dimeric G-quadruplexes of $d[G_3T_4G_3]$ (Smith et al. 1994) and $d[G_4T_4G_4]$ (Smith et al. 1993). For the $d[T_2(G_3T_4)_3G_3AGAGGTAGCATACCTCT]$ sequence, besides the peaks (Figure 5.8-C) resembling to the sum of NMR spectra of two components, G-quadruplex (Figure 5.8-B) and duplex hairpin (Figure 5.8-A), there is a hump and minor peaks at 10 – 12 ppm, indicating the coexistence of different folded structures. Concomitantly, the hump and minor peaks also could indicate that the formation of the G-quadruplex might be affected by the duplex stem as it might interfere with the G-quadruplex's loops (Figure 5.8-C).

5.3.2. Single molecule FRET measurements

To further study the conformation and stability of the $d[(G_3T)_3G_3]$ and $d[(G_3T_4)_3G_3]$ sequences at the single-molecule level, we performed smFRET measurements. The hybridization of DNA constructs and experiment steps for single-molecule measurements are described in the previous section.

Figure 5.9: (A, B) Single molecule FRET histograms of 2 different G-quadruplex forming sequences: (A) $d[(G_3T)_3G_3]$ and (B) $d[(G_3T_4)_3G_3]$ in different salt conditions. (C, D) 2 representative FRET time trace of single molecules from (C) $d[(G_3T)_3G_3]$ in 5mM Li^+ and (D) $d[(G_3T_4)_3G_3]$ in 10 mM K^+ .

Figure 5.9 shows the FRET histograms obtained for $d[(G_3T)_3G_3]$ and $d[(G_3T_4)_3G_3]$ in different salt conditions. For the $d[(G_3T)_3G_3]$ sequence, the FRET histograms showing a broad band FRET value peaking at ~ 0.63 at all different salt conditions. In all cases including the case with only 5 mM Li^+ (a condition where G-quadruplexes are presumably least stable), FRET time traces of these molecules show only a very stable FRET value at ~ 0.63 without any transitions to other FRET values observed (Figure 5.9-C and Figure 5.10). We assigned this 0.63 value to the folded state of the propeller-type parallel G-quadruplex formed by this sequence. This observation is consistent with our NMR and CD melting data in ensemble condition. Since the parallel G-quadruplex formed by $d[(G_3T)_3G_3]$ is very stable, most of the molecules are

still folded even at 5 mM Li^+ condition. We should not exclude the presence of Na^+ which might remain in the channel after washing by T0-Tris (step 7th) or contamination of other cations (even at μM concentration) in our buffer. These cations might also stabilize G-quadruplex structures.

For the $d[(\text{G}_3\text{T}_4)_3\text{G}_3]$ sequence, at 5mM Li^+ , there is only one dominant peak at ~ 0.1 FRET value. As the K^+ concentration increases, this low FRET value peak decreases and gives rise to higher FRET values. However, the distribution of the FRET value at intermediate salt concentration (10 mM K^+) is very broad (Figure 5.9-B). Time trace of single molecule FRET in 10 mM K^+ reveals different FRET values (for example, 0.4 , 0.6 and 0.7 in Figure 5.9-D) and their transitions through the lowest FRET value (~ 0.1). More single-molecule time traces at this 10 mM K^+ concentration shown in Figure 5.11 confirm the present of various FRET values and their transitions always going through the lowest FRET value 0.1. This observation also suggests that different FRET values observed here are unlikely originated from the properties of the dyes (changing in orientation or blinking). Instead, it is also expected since transitions between two folded states most likely require the DNA to unfold. We therefore assign this lowest FRET value 0.1 to the unfolded conformation (U') of this sequence.

Figure 5.10: Examples of FRET time trace from single molecules of $d[(\text{G}_3\text{T})_3\text{G}_3]$ in 5 mM Li^+ .

Figure 5.11: Examples of FRET time trace from single molecules of $d[(G_3T_4)_3G_3]$ in 10 mM K^+ .

At 500 mM K^+ , the FRET histogram of $d[(G_4T_4)_3G_4]$ still shows a broad band spread across medium and high FRET values with a slightly increase of the ~ 0.8 value peak (Figure 5.12). This histogram could be fitted by Gaussian functions with at least 5 components peaking at ~ 0.4 , ~ 0.5 , ~ 0.68 , ~ 0.75 and ~ 0.85 values (Figure 5.12). However, these fitted values might represent more than 5 folded states. Therefore, to analyze the folded states of this sequence, we have to look into each molecule. In contrast with the human telomeric $d[(G_3TTA)_3G_3]$ sequence, together with a stable folded state at ~ 0.75 FRET value (Figure 5.13-C,D), we still observe the existence of various folded state as well as their transitions through unfolded (**U'**) state of $d[(G_4T_4)_3G_4]$ sequence even at this high salt condition: 500 mM K^+ (Figure 5.13-A,B). It is important to note that we did not observe any direct transitions between different folded states. Meanwhile, as reported in the previous section, only two stable FRET state **F1** and **F2** with no transition between them through **U** are observed in the human telomeric sequence at 500 mM K^+ . The comparison between these two sequences highlights the important effect of the loop length and its composition on the stability of G-quadruplexes. A modification of Adenine to Thymine and an addition of a Thymine to each loop can lead to a very complex structural polymorphism with lower stability even at high salt condition. The result found in this study should be taken into account in the design and prediction of G-quadruplex topologies.

Figure 5.12: Single molecule FRET histogram and its Gaussian fitting of $d[(G_3T_4)_3G_3]$ in 500 mM K^+ .

Figure 5.13: Examples of FRET time trace from single molecules of $d[(G_3T_4)_3G_3]$ in high salt condition: 500 mM K^+ .

5.4. Conclusion

We have successfully constructed a prism-type smFRET system based on an Olympus IX71 microscope and the agreement of our data with those reported in the literature for a human telomeric G-quadruplex forming sequence $d[(G_3TTA)_3G_3]$ has validated our smFRET system. Our smFRET analysis confirms a very stable property of a single well-defined propeller-type parallel G-quadruplex conformation formed by $d[(G_3T)_3G_3]$ at low salt concentration. Meanwhile, smFRET revealed an extreme structural polymorphism of G-quadruplexes formed by $d[(G_3T_4)_3G_3]$ at different salt conditions. Although CD spectra of this $d[(G_3T_4)_3G_3]$ sequence at the same condition showed a very clear signature of an antiparallel G-quadruplex, the structural polymorphism revealed by smFRET suggested that various folded structures might co-exist at this condition. The structural polymorphism exhibited by this sequence might be the consequence of the flexibility of the long loops (there are 4 Thymines in the each loop). We also have to include the possibility that this particular sequence might fold into other peculiar structures due to the flexibility of its long loop.

Chapter 6: Summary and future work

6.1. Summary

This thesis has been focused on ensemble and single-molecule fluorescence studies as well as excited state dynamics and stability studies of G-quadruplexes. This work has revealed number of interesting results:

We have characterized and compared fluorescence properties of various well-defined G-quadruplex structures. We showed that the increase of intrinsic fluorescence of G-rich DNA sequences when they form G-quadruplexes can be used to monitor the folding and unfolding of G-quadruplexes as a function of cations and temperature. We found that the temperature-dependent fluorescence spectra of different G-quadruplexes also exhibit characteristic patterns. Therefore, we demonstrated that the stability and possibly also the structure of G-quadruplexes can be characterized and distinguished by their intrinsic fluorescence spectra.

We have reported for the first time the observation of excimer formation by stacking of two G-quadruplex blocks based on the spectroscopic signatures of this excimer: no change in absorption, red-shifted fluorescence emission and lifetime decay in the order of nanoseconds. It is also for the first time, NMR and fluorescence spectra obtained from identical DNA sample were used to confirm and study the effect of excimer formation. We also demonstrated the ability to manipulate the excimer formation by using different DNA concentration, salt concentration, temperature, and mutated DNA flanking sequence. This effect not only could be used as a signature for detection of stacking of G-quadruplex, but also offer a unique means for studying the excited state dynamics between G-tetrad stacking layers.

To study the conformation and stability of G-quadruplex DNAs at single molecule regime, we have successfully constructed a prism-type smFRET system based on an Olympus IX71 microscope. We have performed smFRET analysis on a human telomeric G-quadruplex forming sequence $d[(G_3TTA)_3G_3]$ and the agreement of our data with those reported in the literature has validated our smFRET system. We have

combined extensively the smFRET technique and various ensemble measurements including NMR spectroscopy to study the structure, dynamics and stability of two distinct well-defined G-quadruplex folds: a parallel G-quadruplex formed by d[(G₃T)₃G₃] and an antiparallel G-quadruplex formed by d[(G₃T₄)₃G₃]. While the d[(G₃T)₃G₃] sequence forms an extremely stable parallel G-quadruplex conformation even at very low salt concentration, smFRET result of the d[(G₃T₄)₃G₃] sequence reveals its structural polymorphic property even at high salt concentration. In addition, smFRET data obtained from these sequences were compared with ones of the human telomeric d[(G₃TTA)₃G₃] sequence in order to understand the effect of the loop length and its composition to the stability of G-quadruplexes. The results obtained from these studies could be applied in the design and prediction of G-quadruplex topologies.

6.2. Perspective and Future directions

Both G-quadruplexes' structure and their interactive ligands have been studied intensively so far, but relatively little information about the mechanism of their interactions is known. Meanwhile, there are still many remaining questions regarding the effect of G-quadruplexes' structure on their excited states. Further studies will be needed to look at the mechanism of the excimer formation in ensemble study level, as well as interactions of G-quadruplexes with other molecules at single molecule level.

For examples, in ensemble study, the effect of excimer formation by different stacking configuration has not been fully investigated and understood yet. We also haven't observed the rise of the excimer state in the time-resolved spectroscopy, which is one of the critical properties of the excimer (we have only probed at 430 nm). In single molecule measurement, we would like to apply smFRET technique to study the interaction of G-quadruplexes with other molecules. While some of the works have been in progress, for future directions, we propose:

+) To study the effect of different stacking configurations of G-quadruplexes such as "5' end to 5' end" vs. "3' end to 3' end" vs. "5' end to 3' end" as well as different overlapping modes of the guanines and other bases in the stacking interface on the excimer emission in order to achieve better understanding of the mechanism of this excimer formation.

+) To study the rise of the excimer state using time-resolved spectroscopy by probing at different wavelength and optimizing the time resolution of the system in order to reveal how the excimer states are formed. Decay-associated fluorescence spectra at different K^+ concentrations and temperature, and time-resolved measurements of fluorescence anisotropy are also needed in order to reveal the interplay of different decay components. This information is critical in order to understand the correlations between the fluorescence features and the microscopic electronic structures of G-quadruplexes.

+) To study the interaction of G-quadruplexes with small molecules using PTIR smFRET technique. For examples, recently, the G-quadruplex DNA unfolding processes by the BLM helicase have only been studied in ensemble fluorescence quenching experiment (Liu et al. 2010). Unfolding processes of G-quadruplex DNA in human KRAS promoter by the UP1 protein also have just been revealed by gel electrophoresis and other ensemble measurements such as CD spectroscopy and FRET melting (Paramasivam et al. 2009). But the detail structure of the complex as well as the mechanism of their interaction is not fully understood yet. Our lab is working to solve the structure of the binding complex and by employing smFRET techniques, we can monitor the unfolding processes and their pathways overtime; therefore leads to a better understanding of the mechanism of these interactions.

BIBLIOGRAPHY

- Adrian, M., B. Heddi and A. T. Phan (2012). "NMR spectroscopy of G-quadruplexes." Methods **57**(1): 11-24.
- Akhshi, P., N. J. Mosey and G. Wu (2012). "Free-Energy landscapes of ion movement through a G-Quadruplex DNA channel." Angew. Chem. Int. Ed. **51**(12): 2850-2854.
- Alzeer, J., B. R. Vummidi, P. J. C. Roth and N. W. Luedtke (2009). "Guanidinium-modified phthalocyanines as high-affinity G-quadruplex fluorescent probes and transcriptional regulators." Angew. Chem. Int. Ed. **48**(49): 9362-9365.
- Balagurumoorthy, P. and S. K. Brahmachari (1994). "Structure and stability of human telomeric sequence." J. Biol. Chem. **269**(34): 21858-21869.
- Balasubramanian, S. and S. Neidle (2009). "G-quadruplex nucleic acids as therapeutic targets." Curr. Opin. Chem. Biol. **13**(3): 345-353.
- Balasubramanian, S., L. H. Hurley and S. Neidle (2011). "Targeting G-quadruplexes in gene promoters: a novel anticancer strategy?" Nat. Rev. Drug Discov. **10**(4): 261-275.
- Banyasz, A., I. Vaya, P. Changenet-Barret, T. Gustavsson, T. Douki and D. Markovitsi (2011). "Base pairing enhances fluorescence and favors cyclobutane dimer formation induced upon absorption of UVA radiation by DNA." J. Am. Chem. Soc. **133**(14): 5163-5165.
- Bardin, C. and J. L. Leroy (2008). "The formation pathway of tetramolecular G-quadruplexes." Nucleic Acids Res. **36**(2): 477-488.
- Bates, A. D. and A. Maxwell (1993). DNA Topology ed., IRL Press, Oxford.
- Biffi, G., D. Tannahill, J. McCafferty and S. Balasubramanian (2013). "Quantitative visualization of DNA G-quadruplex structures in human cells." Nat. Chem.: advance online publication with DOI link: 10.1038/nchem.1548.
- Bugaut, A. and S. Balasubramanian (2012). "5'-UTR RNA G-quadruplexes: translation regulation and targeting." Nucleic Acids Res. **40**(11): 4727-4741.
- Burge, S., G. N. Parkinson, P. Hazel, A. K. Todd and S. Neidle (2006). "Quadruplex DNA: sequence, topology and structure." Nucleic Acids Res. **34**(19): 5402-5415.
- Cantor, C. R., M. M. Warshaw and H. Shapiro (1970). "Oligonucleotide interactions .3. Circular dichroism studies of conformation of deoxyoligonucleotides." Biopolymers **9**(9): 1059-1077.
- Changenet-Barret, P., E. Emanuele, T. Gustavsson, R. Improta, A. B. Kotlyar, D. Markovitsi, I. Vayá, K. Zakrzewska and D. Zikich (2010). "Optical properties of Guanine nanowires: experimental and theoretical study." J. Phys. Chem. C **114**(34): 14339-14346.
- Chou, S. H., K. H. Chin and A. H. J. Wang (2005). "DNA aptamers as potential anti-HIV agents." Trends Biochem. Sci **30**(5): 231-234.
- Clark, L. B. (1994). "Electronic spectra of crystalline Guanosine: transition moment directions of the Guanine chromophore." J. Am. Chem. Soc. **116**(12): 5265-5270.
- Cohen, M. D., R. Haberkorn, E. Huler, Z. Ludmer, M. E. Michel-Beyerle, D. Rabinovich, R. Sharon, A. Warshel and V. Yakhot (1978). "Pair-states in α -perylene crystal. A theoretical study." Chem. Phys. **27**(2): 211-216.

- Crespo-Hernandez, C. E., K. de La Harpe and B. Kohler (2008). "Ground-state recovery following UV excitation is much slower in G.C - DNA duplexes and hairpins than in mononucleotides." J. Am. Chem. Soc. **130**(33): 10844-10845.
- Dao, N. T., R. Haselsberger, M. E. Michel-Beyerle and A. T. Phan (2011). "Following G-quadruplex formation by its intrinsic fluorescence." FEBS Lett. **585**(24): 3969-3977.
- Davis, J. T. (2004). "G-quartets 40 years later: From 5'-GMP to molecular biology and supramolecular chemistry." Angew. Chem. Int. Ed. **43**(6): 668-698.
- Dembska, A. and B. Juskowiak (2010). "Effect of metal cations on the fluorescence lifetimes of pyrene-labeled G-quadruplex probes." J. Photochem. Photobiol. A **212**(1): 36-42.
- Deniz, A. A., M. Dahan, J. R. Grunwell, T. J. Ha, A. E. Faulhaber, D. S. Chemla, S. Weiss and P. G. Schultz (1999). "Single-pair fluorescence resonance energy transfer on freely diffusing molecules: Observation of Forster distance dependence and subpopulations." Proc. Natl. Acad. Sci. U.S.A. **96**(7): 3670-3675.
- Do, N. Q., K. W. Lim, M. H. Teo, B. Heddi and A. T. Phan (2011). "Stacking of G-quadruplexes: NMR structure of a G-rich oligonucleotide with potential anti-HIV and anticancer activity." Nucleic Acids Res. **39**(21): 9448-9457.
- Do, N. Q. and A. T. Phan (2012). "Monomer-dimer equilibrium for the 5'-5' stacking of propeller-type parallel-stranded G-quadruplexes: NMR structural study." Chem. Eur. J. **18**(46): 14752-14759.
- Dumas, A. and N. W. Luedtke (2010). "Cation-mediated energy transfer in G-quadruplexes revealed by an internal fluorescent probe." J. Am. Chem. Soc. **132**(51): 18004-18007.
- Dumas, A. and N. W. Luedtke (2011). "Highly fluorescent guanosine mimics for folding and energy transfer studies." Nucleic Acids Res. **39**(15): 6825-6834.
- Eisinger, J. and J. Flores (1979). "Front-face fluorometry of liquid samples." Anal. Biochem. **94**(1): 15-21.
- Fegan, A., P. S. Shirude, L. M. Ying and S. Balasubramanian (2010). "Ensemble and single molecule FRET analysis of the structure and unfolding kinetics of the c-kit promoter quadruplexes." Chem. Commun. **46**(6): 946-948.
- Feigon, J., V. Sklenar, E. Wang, D. E. Gilbert, R. F. Macaya and P. Schultze (1992). "H-1-NMR spectroscopy of DNA." Meth. Enzymol. **211**: 235-253.
- Fleck, A. and H. N. Munro (1962). "Precision of ultraviolet absorption measurements in Schmidt-Thannhauser procedure for nucleic acid estimation." Biochim. Biophys. Acta **55**(5): 571-&.
- Ford, K. G. and S. Neidle (1995). "Perturbations in DNA-structure upon interaction with porphyrins revealed by chemical probes, DNA footprinting and molecular modeling." Biorg. Med. Chem. **3**(6): 671-677.
- Gabelica, V., F. Rosu, E. De Pauw, J. Lemaire, J. C. Gillet, J. C. Pouilly, F. Lecomte, G. Gregoire, J. P. Schermann and C. Desfrancois (2008). "Infrared signature of DNA G-quadruplexes in the gas phase." J. Am. Chem. Soc. **130**(6): 1810-1811.
- Gellert, M., M. N. Lipsett and D. R. Davies (1962). "Helix formation by guanylic acid." Proc. Natl. Acad. Sci. U.S.A. **48**(12): 2013-&.
- Gepshtein, R., D. Huppert, I. Lubitz, N. Amdursky and A. B. Kotlyar (2008). "Radiationless transitions of G4 wires and dGMP." J. Phys. Chem. C **112**(32): 12249-12258.
- Green, J. J., L. M. Ying, D. Klenerman and S. Balasubramanian (2003). "Kinetics of unfolding the human telomeric DNA quadruplex using a PNA trap." J. Am. Chem. Soc. **125**(13): 3763-3767.
- Green, J. J., S. Ladame, L. M. Ying, D. Klenerman and S. Balasubramanian (2006). "Investigating a quadruplex-ligand interaction by unfolding kinetics." J. Am. Chem. Soc. **128**(30): 9809-9812.
- Gurzadyan, G. G. and H. Goerner (1993). "Photolesions in DNA upon 193 nm excitation." Photochem. Photobiol. **58**(1): 71-80.

- Gustavsson, T., A. Sharonov and D. Markovitsi (2002). "Thymine, thymidine and thymidine 5'-monophosphate studied by femtosecond fluorescence upconversion spectroscopy." *Chem. Phys. Lett.* **351**(3-4): 195-200.
- Gustavsson, T., R. Improta and D. Markovitsi (2010). "DNA/RNA: building blocks of life under UV irradiation." *J. Phys. Chem. Lett.* **1**(13): 2025-2030.
- Ha, T., T. Enderle, D. F. Ogletree, D. S. Chemla, P. R. Selvin and S. Weiss (1996). "Probing the interaction between two single molecules: Fluorescence resonance energy transfer between a single donor and a single acceptor." *Proc. Natl. Acad. Sci. U.S.A.* **93**(13): 6264-6268.
- Ha, T. (2001). "Single-molecule fluorescence resonance energy transfer." *Methods* **25**(1): 78-86.
- Han, H. and L. H. Hurley (2000). "G-quadruplex DNA: a potential target for anti-cancer drug design." *Trends Pharmacol. Sci.* **21**(4): 136-142.
- Hardin, C. C., E. Henderson, T. Watson and J. K. Prosser (1991). "Monovalent cation induced structural transitions in telomeric DNAs - G-DNA folding intermediates." *Biochemistry* **30**(18): 4460-4472.
- Henderson, E., C. C. Hardin, S. K. Walk, I. Tinoco and E. H. Blackburn (1987). "Telomeric DNA oligonucleotides form novel intramolecular structures containing guanine guanine base-pairs." *Cell* **51**(6): 899-908.
- Hinterdorfer, P. and V. O. Antoine (2009). *Handbook of single-molecule biophysics*. 1st ed., Springer.
- Hu, L. Y., K. W. Lim, S. Bouaziz and A. T. Phan (2009). "Giardia telomeric sequence d(TAGGG)₄ forms two intramolecular G-quadruplexes in K⁺ solution: effect of loop length and sequence on the folding topology." *J. Am. Chem. Soc.* **131**(46): 16824-16831.
- Hua, Y., P. Changenet-Barret, R. Improta, I. Vayá, T. Gustavsson, A. B. Kotlyar, D. Zikich, P. Šket, J. Plavec and D. Markovitsi (2012). "Cation effect on the electronic excited states of guanine nanostructures studied by time-resolved fluorescence spectroscopy." *J. Phys. Chem. C* **116**(27): 14682-14689.
- Hud, N. V., P. Schultze, V. Sklenar and J. Feigon (1999). "Binding sites and dynamics of ammonium ions in a telomere repeat DNA quadruplex." *J. Mol. Biol.* **285**(1): 233-243.
- Huppert, J. L. and S. Balasubramanian (2007). "G-quadruplexes in promoters throughout the human genome." *Nucleic Acids Res.* **35**(2): 406-413.
- Ida, R. and G. Wu (2008). "Direct NMR detection of alkali metal ions bound to G-quadruplex DNA." *J. Am. Chem. Soc.* **130**(11): 3590-3602.
- Jena, P. V., P. S. Shirude, B. Okumus, K. Laxmi-Reddy, F. Godde, I. Huc, S. Balasubramanian and T. Ha (2009). "G-Quadruplex DNA Bound by a Synthetic Ligand is Highly Dynamic." *J. Am. Chem. Soc.* **131**(35): 12522-+.
- Jin, R. Z., B. L. Gaffney, C. Wang, R. A. Jones and K. J. Breslauer (1992). "Thermodynamics and structure of a DNA tetraplex - a spectroscopic and calorimetric study of the tetramolecular complexes of d(TG3T) and d(TG3T2G3T)." *Proc. Natl. Acad. Sci. U.S.A.* **89**(18): 8832-8836.
- Jing, N. J., W. Sha, Y. D. Li, W. J. Xiong and D. J. Tweardy (2005). "Rational drug design of G-quartet DNA as anti-cancer agents." *Curr. Pharm. Design* **11**(22): 2841-2854.
- Kang, C., X. H. Zhang, R. Ratliff, R. Moyzis and A. Rich (1992). "CRYSTAL-STRUCTURE OF 4-STRANDED OXYTRICHA TELOMERIC DNA." *Nature* **356**(6365): 126-131.
- Kato, Y., T. Ohyama, H. Mita and Y. Yamamoto (2005). "Dynamics and thermodynamics of dimerization of parallel G-quadruplexed DNA formed from d(TTAG(n)) (n=3-5)." *J. Am. Chem. Soc.* **127**(28): 9980-9981.
- Keeler, J. (2005). *Understanding NMR Spectroscopy*. 1st ed., Wiley.

- Kim, N., M. Piatyszek, K. Prowse, C. Harley, M. West, P. Ho, G. Coviello, W. Wright, S. Weinrich and J. Shay (1994). "Specific association of human telomerase activity with immortal cells and cancer." Science **266**(5193): 2011-2015.
- Krawczyk, S. and R. Luchowski (2007). "Electronic excited states of polynucleotides: A study by electroabsorption spectroscopy." J. Phys. Chem. B **111**(5): 1213-1221.
- Krishnan-Ghosh, Y., D. S. Liu and S. Balasubramanian (2004). "Formation of an interlocked quadruplex dimer by d(GGGT)." J. Am. Chem. Soc. **126**(35): 11009-11016.
- Kwok, W.-M., C. Ma and D. L. Phillips (2006). "Femtosecond time- and wavelength-resolved fluorescence and absorption spectroscopic study of the excited states of Adenosine and an Adenine oligomer." J. Am. Chem. Soc. **128**(36): 11894-11905.
- Kypr, J., I. Kejnovska, D. Renciuik and M. Vorlickova (2009). "Circular dichroism and conformational polymorphism of DNA." Nucleic Acids Res. **37**(6): 1713-1725.
- Lakowicz, J. R., B. Shen, Z. Gryczynski, S. D'Auria and I. Gryczynski (2001). "Intrinsic fluorescence from DNA can be enhanced by metallic particles." Biochem. Biophys. Res. Commun. **286**(5): 875-879.
- Lakowicz, J. R. (2006). Principles of fluorescence spectroscopy. 3rd ed., New York, Springer.
- Lange, A. W. and J. M. Herbert (2009). "Both intra- and interstrand charge-transfer excited states in aqueous B-DNA are present at energies comparable to, or just above, the $^1\pi\pi^*$ excitonic bright states." J. Am. Chem. Soc. **131**(11): 3913-3922.
- Lech, C. J., B. Heddi and A. T. Phan (2012). "Guanine base stacking in G-quadruplex nucleic acids." Nucleic Acids Res.: published online with DOI link: 10.1093/nar/gks1110.
- Lee, J. Y., B. Okumus, D. S. Kim and T. J. Ha (2005). "Extreme conformational diversity in human telomeric DNA." Proc. Natl. Acad. Sci. U.S.A. **102**(52): 18938-18943.
- Lee, J. Y., J. M. Yoon, H. W. Kihm and D. S. Kim (2008). "Structural diversity and extreme stability of unimolecular Oxytricha nova telomeric G-quadruplex." Biochemistry **47**(11): 3389-3396.
- Lee, J. Y. and D. S. Kim (2009). "Dramatic effect of single-base mutation on the conformational dynamics of human telomeric G-quadruplex." Nucleic Acids Res. **37**(11): 3625-3634.
- Lipps, H. J. and D. Rhodes (2009). "G-quadruplex structures: in vivo evidence and function." Trends Cell Biol. **19**(8): 414-422.
- Liu, J. Q., C. Y. Chen, Y. Xue, Y. H. Hao and Z. Tan (2010). "G-Quadruplex Hinders Translocation of BLM Helicase on DNA: A Real-Time Fluorescence Spectroscopic Unwinding Study and Comparison with Duplex Substrates." J. Am. Chem. Soc. **132**(30): 10521-10527.
- Liu, W., H. Zhu, B. Zheng, S. Cheng, Y. Fu, W. Li, T.-C. Lau and H. Liang (2012). "Kinetics and mechanism of G-quadruplex formation and conformational switch in a G-quadruplex of PS2.M induced by Pb²⁺." Nucleic Acids Res.
- Luu, K. N., A. T. Phan, V. Kuryavyi, L. Lacroix and D. J. Patel (2006). "Structure of the human telomere in K⁺ solution: An intramolecular (3+1) G-quadruplex scaffold." J. Am. Chem. Soc. **128**(30): 9963-9970.
- Maizels, N. (2006). "Dynamic roles for G4 DNA in the biology of eukaryotic cells." Nat. Struct. Mol. Biol. **13**(12): 1055-1059.
- Malcolm, H. L. (2008). Spin Dynamics: Basics of Nuclear Magnetic Resonance. 2nd ed., John Wiley & Sons Ltd.
- Marguet, S. and D. Markovitsi (2005). "Time-resolved study of thymine dimer formation." J. Am. Chem. Soc. **127**(16): 5780-5781.
- Markovitsi, D., T. Gustavsson and A. Sharonov (2004). "Cooperative effects in the photophysical properties of self-associated triguanosine diphosphates." Photochem. Photobiol. **79**(6): 526-530.

- Marsh, T. C. and E. Henderson (1994). "G-wires-self-assembly of a telomeric oligonucleotide, d(GGGGTTGGGG), into large superstructures." Biochemistry **33**(35): 10718-10724.
- Martadinata, H. and A. T. Phan (2009). "Structure of propeller-type parallel-stranded RNA G-quadruplexes, formed by human telomeric RNA sequences in K⁺ solution." J. Am. Chem. Soc. **131**(7): 2570-2578.
- Martadinata, H. and A. T. Phan (2012). "NMR structure of human telomeric RNA (TERRA): Stacking of two G-quadruplex blocks in K⁺ solution." submitted.
- Matsugami, A., K. Ouhashi, M. Kanagawa, H. Liu, S. Kanagawa, S. Uesugi and M. Katahira (2001). "An intramolecular quadruplex of (GGA)₄ triplet repeat DNA with a G:G:G:G tetrad and a G(:A):G(:A):G(:A):G heptad, and its dimeric interaction." J. Mol. Biol. **313**(2): 255-269.
- Mendez, M. A. and V. A. Szalai (2009). "Fluorescence of unmodified oligonucleotides: a tool to probe G-quadruplex DNA structure." Biopolymers **91**(10): 841-850.
- Mergny, J. L. and C. Helene (1998a). "G-quadruplex DNA: A target for drug design." Nat. Med. **4**(12): 1366-1367.
- Mergny, J. L., A. T. Phan and L. Lacroix (1998b). "Following G-quartet formation by UV-spectroscopy." FEBS Lett. **435**(1): 74-78.
- Mergny, J. L. and J. C. Maurizot (2001). "Fluorescence resonance energy transfer as a probe for G-quartet formation by a telomeric repeat." ChemBioChem **2**(2): 124-132.
- Mergny, J. L., J. Li, L. Lacroix, S. Amrane and J. B. Chaires (2005). "Thermal difference spectra: a specific signature for nucleic acid structures." Nucleic Acids Res. **33**(16): e138.
- Miannay, F. A., A. Banyasz, T. Gustavsson and D. Markovitsi (2009). "Excited states and energy transfer in G-quadruplexes." J. Phys. Chem. C **113**(27): 11760-11765.
- Middleton, C. T., K. de La Harpe, C. Su, Y. K. Law, C. E. Crespo-Hernandez and B. Kohler (2009). "DNA excited-state dynamics: from single bases to the double helix." Annu. Rev. Phys. Chem. **60**: 217-239.
- Miura, T., J. M. Benevides and G. J. Thomas (1995a). "A phase-diagram for sodium and potassium-ion control of polymorphism in telomeric DNA." J. Mol. Biol. **248**(2): 233-238.
- Miura, T. and G. J. Thomas (1995b). "Structure and dynamics of interstrand guanine association in Quadruplex telomeric DNA." Biochemistry **34**(29): 9645-9654.
- Monchaud, D. and M. P. Teulade-Fichou (2008). "A hitchhiker's guide to G-quadruplex ligands." Org. Biomol. Chem. **6**(4): 627-636.
- Moyzis, R. K., J. M. Buckingham, L. S. Cram, M. Dani, L. L. Deaven, M. D. Jones, J. Meyne, R. L. Ratliff and J. R. Wu (1988). "A highly conserved repetitive DNA sequence, (TTAGGG)_n, present at the telomeres of human chromosomes." Proc. Natl. Acad. Sci. U.S.A. **85**(18): 6622-6626.
- Mukundan, V. T., N. Q. Do and A. T. Phan (2011). "HIV-1 integrase inhibitor T30177 forms a stacked dimeric G-quadruplex structure containing bulges." Nucleic Acids Res. **39**(20): 8984-8991.
- Nadler, A., J. Strohmeier and U. Diederichsen (2011). "8-Vinyl-2'-deoxyguanosine as a fluorescent 2'-deoxyguanosine mimic for investigating DNA hybridization and topology." Angew. Chem. Int. Ed. **50**(23): 5392-5396.
- Neidle, S. (2009). "The structures of quadruplex nucleic acids and their drug complexes." Curr. Opin. Struct. Biol. **19**(3): 239-250.
- Oganesian, L., I. K. Moon, T. M. Bryan and M. B. Jarstfer (2006). "Extension of G-quadruplex DNA by ciliate telomerase." EMBO J **25**(5): 1148-1159.
- Okamoto, K., Y. Sannohe, T. Mashimo, H. Sugiyama and M. Terazima (2008). "G-quadruplex structures of human telomere DNA examined by single molecule FRET and BrG-substitution." Biorg. Med. Chem. **16**(14): 6873-6879.

- Paramasivam, M., A. Membrino, S. Cogoi, H. Fukuda, H. Nakagama and L. E. Xodo (2009). "Protein hnRNP A1 and its derivative Up1 unfold quadruplex DNA in the human KRAS promoter: implications for transcription." *Nucleic Acids Res.* **37**(9): 2841-2853.
- Patel, D. J. and A. E. Tonelli (1974). "Assignment of the proton nmr chemical shifts of the T-N₃H and G-N₁H proton resonances in isolated AT and GC Watson-Crick base pairs in double-stranded deoxy oligonucleotides in aqueous solution." *Biopolymers* **13**(10): 1943-1964.
- Patel, D. J., A. T. Phan and V. Kuryavyi (2007). "Human telomere, oncogenic promoter and 5'-UTR G-quadruplexes: Diverse higher order DNA and RNA targets for cancer therapeutics." *Nucleic Acids Res.* **35**(22): 7429-7455.
- Pecourt, J. M. L., J. Peon and B. Kohler (2000). "Ultrafast internal conversion of electronically excited RNA and DNA nucleosides in water." *J. Am. Chem. Soc.* **122**(38): 9348-9349.
- Pedersen, E. B., J. T. Nielsen, C. Nielsen and V. V. Filichev (2011). "Enhanced anti-HIV-1 activity of G-quadruplexes comprising locked nucleic acids and intercalating nucleic acids." *Nucleic Acids Res.* **39**(6): 2470-2481.
- Peon, J. and A. H. Zewail (2001). "DNA/RNA nucleotides and nucleosides: direct measurement of excited-state lifetimes by femtosecond fluorescence up-conversion." *Chem. Phys. Lett.* **348**(3-4): 255-262.
- Phan, A. T. and D. J. Patel (2003). "Two-repeat human telomeric d(TAGGGTTAGGGT) sequence forms interconverting parallel and antiparallel G-quadruplexes in solution: Distinct topologies, thermodynamic properties, and folding/unfolding kinetics." *J. Am. Chem. Soc.* **125**(49): 15021-15027.
- Phan, A. T., Y. S. Modi and D. J. Patel (2004). "Two-repeat Tetrahymena telomeric d(TGGGGTTGGGGT) sequence interconverts between asymmetric dimeric G-quadruplexes in solution." *J. Mol. Biol.* **338**(1): 93-102.
- Phan, A. T., V. Kuryavyi, J. B. Ma, A. Faure, M. L. Andreola and D. J. Patel (2005). "An interlocked dimeric parallel-stranded DNA quadruplex: A potent inhibitor of HIV-1 integrase." *Proc. Natl. Acad. Sci. U.S.A.* **102**(3): 634-639.
- Phan, A. T. (2010). "Human telomeric G-quadruplex: structures of DNA and RNA sequences." *FEBS J.* **277**(5): 1107-1117.
- Plateau, P. and M. Guéron (1982). "Exchangeable proton NMR without base-line distortion, using new strong-pulse sequences." *J. Am. Chem. Soc.* **104**(25): 7310-7311.
- Plessow, R., A. Brockhinke, W. Eimer and K. Kohse-Hoinghaus (2000). "Intrinsic time- and wavelength-resolved fluorescence of oligonucleotides: A systematic investigation using a novel picosecond laser approach." *J. Phys. Chem. B* **104**(15): 3695-3704.
- Randazzo, A., G. Spada and M. Silva (2012). Circular Dichroism of Quadruplex Structures, Topics in Current Chemistry, Springer Berlin / Heidelberg: 1-20.
- Rasnik, I., S. A. McKinney and T. Ha (2006). "Nonblinking and long-lasting single-molecule fluorescence imaging." *Nat Meth* **3**(11): 891-893.
- Rhoades, E., E. Gussakovsky and G. Haran (2003). "Watching proteins fold one molecule at a time." *Proc. Natl. Acad. Sci. U.S.A.* **100**(6): 3197-3202.
- Rich, A. and M. Kasha (1960). "The $n \rightarrow \pi^*$ transition in nucleic acids and polynucleotides." *J. Am. Chem. Soc.* **82**(23): 6197-6199.
- Roy, R., S. Hohng and T. Ha (2008). "A practical guide to single-molecule FRET." *Nat. Methods* **5**(6): 507-516.
- Sabanayagam, C. R., J. S. Eid and A. Meller (2005). "Using fluorescence resonance energy transfer to measure distances along individual DNA molecules: Corrections due to nonideal transfer." *J. Chem. Phys.* **122**(6).
- Santoro, F., V. Barone and R. Improta (2009). "Excited states decay of the A-T DNA: a PCM/TD-DFT study in aqueous solution of the (9-Methyl-adenine)₂-(1-methyl-thymine)₂ stacked tetramer." *J. Am. Chem. Soc.* **131**(42): 15232-15245.

- Schuler, B., E. A. Lipman and W. A. Eaton (2002). "Probing the free-energy surface for protein folding with single-molecule fluorescence spectroscopy." Nature **419**(6908): 743-747.
- Schultze, P., N. V. Hud, F. W. Smith and J. Feigon (1999). "The effect of sodium, potassium and ammonium ions on the conformation of the dimeric quadruplex formed by the *Oxytricha nova* telomere repeat oligonucleotide d(G₄T₄G₄)." Nucleic Acids Res. **27**(15): 3018-3028.
- Schwalb, N. K. and F. Temps (2008). "Base sequence and higher-order structure induce the complex excited-state dynamics in DNA." Science **322**(5899): 243-245.
- Selvin, P. and T. J. Ha (2007). Single Molecule Techniques: a Laboratory Manual. 1st ed., New York, Cold Spring Harbor Laboratory Press, Cold Spring Harbor.
- Sen, D. and W. Gilbert (1988). "Formation of parallel 4-stranded complexes by guanine-rich motifs in DNA and its implications for meiosis." Nature **334**(6180): 364-366.
- Shim, J. W., Q. Tan and L.-Q. Gu (2009). "Single-molecule detection of folding and unfolding of the G-quadruplex aptamer in a nanopore nanocavity." Nucleic Acids Res. **37**(3): 972-982.
- Siegmund, K., P. Daublain, Q. Wang, A. Trifonov, T. Fiebig and F. D. Lewis (2009). "Structure and photoinduced electron transfer in DNA hairpin conjugates possessing a tethered 5'-pyrenecarboxamide." J. Phys. Chem. B **113**(50): 16276-16284.
- Simonsson, T. (2001). "G-quadruplex DNA structures - Variations on a theme." Biol. Chem. **382**(4): 621-628.
- Sket, P. and J. Plavec (2010). "Tetramolecular DNA Quadruplexes in Solution: Insights into Structural Diversity and Cation Movement." J. Am. Chem. Soc. **132**(36): 12724-12732.
- Smargiasso, N., F. Rosu, W. Hsia, P. Colson, E. S. Baker, M. T. Bowers, E. De Pauw and V. Gabelica (2008). "G-quadruplex DNA assemblies: Loop length, cation identity, and multimer formation." J. Am. Chem. Soc. **130**(31): 10208-10216.
- Smith, F. W. and J. Feigon (1992). "Quadruplex structure of *Oxytricha* telomeric DNA oligonucleotides." Nature **356**(6365): 164-168.
- Smith, F. W. and J. Feigon (1993). "Strand orientation in the DNA quadruplex formed from the *Oxytricha* telomere repeat oligonucleotide d(G₄T₄G₄) in solution." Biochemistry **32**(33): 8682-8692.
- Smith, F. W., F. W. Lau and J. Feigon (1994). "d(G₃T₄G₃) forms an asymmetric diagonally looped dimeric quadruplex with guanosine 5'-syn-syn-anti and 5'-syn-anti-anti n-glycosidic conformations." Proc. Natl. Acad. Sci. U.S.A. **91**(22): 10546-10550.
- Stuhldreier, M. C., C. Schüler, J. Kleber and F. Temps (2010). "Femtosecond fluorescence measurements of the Adenine dinucleotide: direct observation of the excimer state." International Conference on Ultrafast Phenomena, Optical Society of America, Colorado, July: E8.
- Su, C., C. T. Middleton and B. Kohler (2012). "Base-stacking disorder and excited-state dynamics in single-stranded Adenine homo-oligonucleotides." J. Phys. Chem. B **116**(34): 10266-10274.
- Sundquist, W. I. and A. Klug (1989). "Telomeric DNA dimerizes by formation of guanine tetrads between hairpin loops." Nature **342**(6251): 825-829.
- Takaya, T., C. Su, K. de La Harpe, C. E. Crespo-Hernandez and B. Kohler (2008). "UV excitation of single DNA and RNA strands produces high yields of exciplex states between two stacked bases." Proc. Natl. Acad. Sci. U.S.A. **105**(30): 10285-10290.
- Todd, A. K., M. Johnston and S. Neidle (2005). "Highly prevalent putative quadruplex sequence motifs in human DNA." Nucleic Acids Res. **33**(9): 2901-2907.

- Vayá, I., T. Gustavsson, F. o. Miannay, T. Douki and D. Markovitsi (2010). "Fluorescence of natural DNA: from the femtosecond to the nanosecond time scales." J. Am. Chem. Soc. **132**(34): 11834-11835.
- Vayá, I., J. Brazard, T. Gustavsson and D. Markovitsi (2012a). "Electronically excited states of DNA oligonucleotides with disordered base sequences studied by fluorescence spectroscopy." Photochem. Photobiol. Sci. **11**(11): 1767-1773.
- Vayá, I., T. Gustavsson, T. Douki, Y. Berlin and D. Markovitsi (2012b). "Electronic: excitation energy transfer between nucleobases of natural DNA." J. Am. Chem. Soc. **134**(28): 11366-11368.
- Venkatesan, N., Y. J. Seo and B. H. Kim (2008). "Quencher-free molecular beacons: a new strategy in fluorescence based nucleic acid analysis." Chem. Soc. Rev. **37**(4): 648-663.
- Vigny, P. and J. P. Ballini (1977). Excited states of nucleic acids at 300 K and electronic energy transfer. Excited States in Organic Chemistry and Biochemistry. B. Pullman and N. Goldblum, Reidel, Dordrecht: 1-13.
- von Feilitzsch, T., J. Tuma, H. Neubauer, L. Verdier, R. Haselsberger, R. Feick, G. Gurzadyan, A. A. Voityuk, C. Griesinger and M. E. Michel-Beyerle (2008). "Chromophore/DNA interactions: femto- to nanosecond spectroscopy, NMR structure, and electron transfer theory." J. Phys. Chem. B **112**(3): 973-989.
- Von freydorf, E., J. Kinder and M. E. Michel-Beyerle (1978). "Low-temperature fluorescence of perylene crystals." Chem. Phys. **27**(2): 199-209.
- Wang, Y. and D. J. Patel (1992). "Guanine residues in d(T₂AG₃) and d(T₂G₄) form parallel-stranded potassium cation stabilized G-quadruplexes with anti glycosidic torsion angles in solution." Biochemistry **31**(35): 8112-8119.
- Wang, Y. and D. J. Patel (1995). "Solution Structure of the Oxytricha Telomeric Repeat d[G₄(T₄G₄)₃] G-tetraplex." J. Mol. Biol. **251**(1): 76-94.
- Watson, J. D. and F. H. C. Crick (1953). "Molecular structure of nucleic acids - a structure for deoxyribose nucleic acid." Nature **171**(4356): 737-738.
- Williamson, J. R., M. K. Raghuraman and T. R. Cech (1989). "Mono-valent cation induced structure of telomeric DNA - the G-quartet model." Cell **59**(5): 871-880.
- Williamson, J. R. (1994). "G-quartet structures in telomeric DNA." Annu. Rev. Biophys. Biomol. Struct. **23**: 703-730.
- Yang, P., A. De Cian, M. P. Teulade-Fichou, J. L. Mergny and D. Monchaud (2009). "Engineering bisquinolinium/thiazole orange conjugates for fluorescent sensing of G-quadruplex DNA." Angew. Chem. Int. Ed. **48**(12): 2188-2191.
- Ying, L. M., J. J. Green, H. T. Li, D. Klenerman and S. Balasubramanian (2003). "Studies on the structure and dynamics of the human telomeric G quadruplex by single-molecule fluorescence resonance energy transfer." Proc. Natl. Acad. Sci. U.S.A. **100**(25): 14629-14634.
- Yoo, H., J. Yang, A. Yousef, M. R. Wasielewski and D. Kim (2010). "Excimer formation dynamics of intramolecular pi-stacked perylenediimides probed by single-molecule fluorescence spectroscopy." J. Am. Chem. Soc. **132**(11): 3939-3944.
- Zahler, A. M., J. R. Williamson, T. R. Cech and D. M. Prescott (1991). "Inhibition of telomerase by G-quartet DNA structures." Nature **350**(6320): 718-720.
- Zhuang, X., L. E. Bartley, H. P. Babcock, R. Russell, T. Ha, D. Herschlag and S. Chu (2000). "A Single-Molecule Study of RNA Catalysis and Folding." Science **288**(5473): 2048-2051.

APPENDICES

A. Quartz slides cleaning protocol for smFRET measurement

A1. Starting with old slides

1. Because quartz slides are quite expensive and it is time-consuming to drill new slides, it is recommended to reuse the old slides. First, remove the tape and coverslip from the slide. There are many ways to do it. The one which is least toxic and more time-saving is microwaving the slides for 1-2 minutes in a beaker filled with water. Once the water boils, the glue should turn a yellowish color. While it is hot, the coverslip and glue are easily to be removed.
2. The slide should be relatively clean, run across it with a fresh razor blade and rinse and wipe with acetone.
3. (Optional) Use cleaning jars filled with acetone and sonicate them for 15 minutes.
4. Use cleaning jars filled with 1-2% Hellmanex and deionized water and sonicate them for 20 minutes. This is the mixture we usually use to clean quartz cuvettes after annealing experiments (these experiments used parafilm oil).
5. Follow steps 3-9 in **A2**.

A2. Starting with new slides

1. Use a template slide to mark the holes on the new slides that needed to be drill. Figure A1-A show picture of a template slide for 5 channels. (We can distinguish quartz slides with glass slides by looking in their rear side. Quartz slides have white color while glass slides have light green color. If we make a mistake and take glass slides, when we measure sm-FRET, the Cy5 channel will look extremely bright and we can't see any molecules.)
2. First, start drilling in the little puddle above the marker point. We usually put a smallish drop of water over the marked holes and a lot of water around the

drilling wood stand to cool the drill bit (Figure A1-B). Then drill the hole with a speed between 6000-9000rpm. We'll feel the drill bit punch through in less than a minute if the bit is good.

3. Rinse slides with deionized-H₂O and place them into the white plastic slide cleaning containers.
4. Sonicate the plastic slide cleaning jars filled with 1M KOH for 20-30 minutes. At the same time fill another plastic slide cleaning jar with glass coverslips and sonicate them simultaneously for 20-30min.
5. Pour out the KOH in the cleaning jars and fill them with deionized-H₂O.
6. Sonicate the plastic slide cleaning jars with deionized-H₂O for 20 minutes and pour out the deionized-H₂O.
7. Blow these slides dry with nitrogen gas flow.
8. Burn these slides for ~30 seconds with the torch (~5 seconds for a glass coverslip) (Figure A1-C) to remove any organic molecules on their surface.
9. Cool the slides off with nitrogen on their backside and place them in a methanol filled jar (see B2). These slides are ready for PEG coating.

Figure A1: (A) Picture of a template slide for 5 channels. (B) Slide drilling. (C) Slide burning to remove any organic molecules on their surface.

B. PEG coating protocol for immobilization of DNA on quartz surface for smFRET measurement

B1. Aminosilanization

1. Rinse Amino-jars with Methanol and fill with Methanol (do not make these jars contact with water because water will stop the reaction).
2. Thaw Aminosilane bottle. If we open the bottle when it is still cold, there will be water vapor in the bottle and it is bad for the reaction.
3. Rinse and sonicate slide-cleaning jars with de-ionized water, then store with de-ionized water.
4. Prepare aminosilanization mixture as in the table below, mix well with glass pipette tip and replace methanol with the mixture, incubate for 10 min.

No. of jars	Methanol (ml)	Acetic Acid (ml)	Aminosilane (ml)
2	100	5	1
3	150	7.5	1.5
4	200	10	2

5. Replace aminosilanization mixture in the jars with methanol, rinse each slide or coverslip with (i) deionized water; (ii) methanol; then (iii) dry them by blowing with nitrogen and lace in PEG assembly boxes.
6. Amino-jars storage: Rinse and sonicate with methanol for 5 minutes then store dry.

B2. PEG coating

1. Prepare PEG prep box by filling with de-ionized water and blotting any excess water on the surface.
2. Thaw biotin-PEG and mPEG bottles.
3. Prepare PEG buffer: 84 mg NaHCO_3 in 10 mL de-ionized.
4. Weigh out biotin-PEG and mPEG, shake well in a 1.5 mL tube, mix thoroughly by tapping. Add the PEG buffer in, shake and flip thoroughly to dissolve (DO NOT vortex). The appropriate amount of biotin-PEG and mPEG depend on the number of slides as in the table below:

Slides	bPEG (mg)	mPEG (mg)	PEG buffer (uL)
1	0.33	17	70
5	1.6	90	350
8	2.5	140	560
14	4.5	240	980

5. Centrifuge the PEG mixture at 10,000 rpm for 1 minute. Apply 75 μ L PEG per slide and then place coverslip on each.
6. Check for sliding after 15 minutes and incubate for at least 3 hours (or overnight) in a dark, stable place.

B3. Disassembly and storage of PEG slides

1. First, write a number “2” on one of the top corner of the non-PEG slide’s face. This is for distinguish which face has PEG on, since they both look the same.
2. Disassemble and rinse each slides thoroughly with de-ionized water.
3. Dry with nitrogen and store each in a separate tube.
4. Keep in -30 °C fridge. Thaw the slides before use.

C. Chamber fabrication protocol for smFRET measurement

1. Thaw slide.
2. Stick the double side tape on a flat and clean surface. Use a razor to cut the double side tape to a desired width for 5 channels slide. Lay down the double side tape to the quartz slide to make 5 channels (Figure A2-A).
3. Rinse a cover slip well with de-ionized water, acetone then de-ionized water. Quickly burn it (~ 5 seconds) and dry with nitrogen gas. Lay down the cover slip on top of the double slide tape.
4. Use a 200 μ L pipette (yellow) tip to press on the coverslip side where the tape is sandwiched. This provides a very tight seal.
5. Cut the slide-coverslip chamber out from the pipette box with a razor (Figure A2-B).
6. Mix two blobs from 5-minutes-epoxy tube very well in roughly equal volume with a yellow pipette tip. If not, the epoxy will take a very long time to dry. Seal the side ends of the chamber with the epoxy (Figure A2-C).
7. Solution can be floated into the chamber by pressing solution from a pipette tip at one end and taping a paper at the other end (capillary effect).

Figure A2: (A) Picture of double side tapes stuck on the quartz slide to make 5 channels. (B+C) The slide's chamber after (B) cutting the spare tapes and (C) sealing with epoxy.

D. Procedure of correcting FRET efficiency in single molecule time trace

In a single molecule FRET time trace, FRET efficiency simply is defined by:

$$E = I_A / (I_A + I_D) \quad (\text{Eq. A.1})$$

with I_D and I_A are the fluorescence intensity of donor and acceptor channel. However, because of the presence of emission leakage from donor to acceptor channel (due to overlap of donor emission with the detection window of the acceptor channel) and the variety of fluorophores' quantum yields, the FRET efficiency usually is corrected by a method introduced by Meller's group (Sabanayagam et al. 2005), followed an equation:

$$E_{corr} = (I_A - \beta I_D) / (I_A + \gamma I_D) \quad (\text{Eq. A.2})$$

with β is the leakage of donor to acceptor channel and γ is the factor included quantum yields of the fluorophore and the detection efficiency of 2 channels. The β factor depends on the construction of smFRET system (due to different filter set or optical alignment), but it should be the same for the molecules in one measurement. The γ factor not only depends on the dyes we used but also varies from molecule to molecule and thus ideally, has to be corrected for each individual molecule. For example, in a particular duplex, the (TMR+Cy5) pair has much narrower ($\beta + \gamma$) distribution compared with (Cy3+Cy5) pair, which make it a better dye pair to use in this case in order to get a more precise FRET value (Figure 3 of (Sabanayagam et al. 2005)).

Figure A3: Schematic of an example of single molecule FRET time trace in an acceptor photobleaching event.

In order to correct the FRET efficiency, let's look at an acceptor photobleaching event (Figure A3). After acceptor intensity photobleached from I_A to I_A' , the donor intensity increased from I_D to I_D' (because it didn't have to transfer energy to the acceptor anymore). The two factors then can be calculated follow these two equations:

$$\beta = I_A' / I_D' \quad (\text{Eq. A.3})$$

$$\gamma = (I_A - I_A') / (I_D' - I_D) \quad (\text{Eq. A.4})$$

The corrected FRET value from then can be calculate from I_A and I_D values in its single molecule time trace by Equation A.2.