# Conformational dynamics of G-quadruplex DNA probed by time-resolved circular dichroism

Marco Schmid

NNT : 2017SACLX107

Thèse de doctorat

# Conformational dynamics of G-quadruplex DNA probed by time-resolved circular dichroism

Thèse de doctorat de l'Université Paris-Saclay
préparée à Ecole Polytechnique

École doctorale : Interfaces (ED 573)
Spécialité de doctorat: Physique

Thèse présentée et soutenue à l'Ecole Polytechnique, le 13 décembre 2017

**Marco Schmid**

Composition du Jury :

| | |
|---|---|
| Valérie Gabelica<br>IECB, Université de Bordeaux | Rapporteur |
| Martin Volk<br>Department of Chemistry, University of Liverpool | Rapporteur |
| Stefan Haacke<br>IPCPM, Université de Strasbourg | Président du jury |
| Frank Wien<br>DISCO, Synchrotron Soleil | Examinateur |
| François Hache<br>LOB, Ecole Polytechnique | Directeur de thèse |
| Pascale Changenet-Barret<br>LOB, Ecole Polytechnique | Co-Directrice de thèse |

Es ist nicht das Wissen, sondern
das Lernen, nicht das Besitzen,
sondern das Erwerben, nicht das
Dasein, sondern das Hinkommen,
was den größten Genuß gewährt.

---

*Carl Friedrich Gauß*

# Acknowledgment

A dissertation is in general of a long and stony with many obstacles. Therefore, I am grateful for having two great supervisors who gave me a lot of support and advice. On that note, I want to thank sincerely to Pascale Changenet-Barret and François Hache for those three rewarding years. I would like to also give special thanks to my jury: my two reviewers Valérie Gabelica and Martin Volk as well as our collaborator and member of the jury Frank Wien and the jury's president Stefan Haacke. Their reports and questions have helped us to deepen the understanding of our research.

In order to mount our experiments we had to add many modifications, so without Xavier Solinas and Jean-Marc Sintes our eventual success would not have been possible. Equally, I want to thank Christelle Français and Laure Lachapelle who, as our secretaries, did all the indispensable administrative work during my stay.

I also had fruitful discussions in my first year with Vincent Kemlin who proposed to measure temperature changes via a change of the refractive index of water. Additionally, as an pH indicator to monitor temperature changes, Hubert Becker proposed phenol red, which turned out to be applicable at 293 nm. This facilitated alignment and temperature control significantly. He further helped me with Mayla to perform a gel electrophoresis and introduced me kindly to it.

Three years of working on a interesting scientific project is quite joyful. However, without human relationships is is not completely satisfying. Therefore, I have been quite lucky to perform my thesis in the LOB. Besides the helpfulness of permanent researchers as well as students, there have been several get-togethers, evenings and daily visits. Added to this, we had many discussions at lunch about politics and science.

In this sense, I want to thank my collegues Lipsa, Olga, Ravi, Joséphine, Maila, Thuy, Xiujun, Pascal, Max and all the other students. Additionally I also want to give thanks to all the other members of the LOB for this rewarding and mind-opening three years' journey.

My parents have supported me not only during my studies but throughout my whole life without wanting anything back. For that, words are not enough to express my thankfulness. Also I want to thank my brother and sister and the rest of the family.

# Résumé en français

Les quadruplexes de guanines (G4) sont des structures secondaires de l'ADN qui consistent en l'assemblage de tétrades de guanines reliées par des liaisons hydrogène Hoogsteen. La formation des G4 nécessite généralement la présence de cations métalliques (typiquement $Na^+$ ou $K^+$) qui stabilisent la structure. Les G4 sont des structures très polymorphes dont la topologie dépend non seulement de la séquence d'ADN, mais aussi de la nature des cations métalliques et de leur environnement. Ils peuvent être formés à partir d'un, deux, trois ou quatre brins d'ADN. Les G4 formés à partir d'un seul brin d'ADN (i.e. monomoléculaires) sont les structures les plus pertinentes biologiquement. Ces dernières décennies, l'intérêt pour ces structures monomoléculaires a fortement augmenté avec l'identification dans divers génomes d'un grand nombre de séquences riches en guanine capables d'adopter une conformation G4 *in vitro*, dans des conditions biologiques. Longtemps contestées, il existe aujourd'hui un ensemble de preuves expérimentales qui attestent de l'existence des G4 *in vivo* et de leurs implications dans d'importants processus cellulaires, comme la maintenance des télomères, la réplication, la transcription ou la translation des gènes. Toutefois, malgré le nombre croissant des études des G4, très peu d'entre elles ont abordé les aspects dynamiques de leur formation. La plupart de ces études qui se sont focalisées sur des séquences télomèriques humaines très polymorphes ont conduit à des images contradictoires de leurs mécanismes de repliement et dépliement.

Dans le cadre de ma thèse, nous avons choisi d'étudier la dynamique de repliement et de dépliement de plusieurs séquences G4 modèles, constituées d'une vingtaine de base, au moyen d'expériences pompe-sonde de saut de température (« T-jump ») couplées à une détection de dichroïsme circulaire (CD) résolu en temps. Le CD, qui est la différence d'absorption d'une onde polarisée circulairement gauche et d'une onde polarisée circulairement droite, est particulièrement sensible à l'arrangement des guanines dans les G4. Il offre la possibilité de sonder la dynamique des changements structuraux de séquences d'ADN non-modifiées, sur une très large fenêtre temporelle. Afin d'initier le dépliement de l'ADN, nous avons utilisé des expériences de « T-jump

» qui consistent à élever brusquement la température de l'échantillon, puis à sonder le retour à l'équilibre à cette nouvelle température.

Toutefois les expériences de « T-jump » conventionnelles, où le changement de température de l'échantillon est généralement induit par l'absorption par le solvant d'un faisceau de lumière infra-rouge pulsé nanoseconde (1,4 $\mu$m pour $H_2O$), ne permettent pas de mesurer les dynamiques de repliement/dépliement des G4 au-delà de quelques millisecondes. Afin d'élargir notre fenêtre d'observation jusqu'aux secondes pour capturer l'intégralité de ces dynamiques, nous avons remplacé notre source laser pulsée IR nanoseconde par une diode laser IR continue (3 W @ 1,48 $\mu$m) et un obturateur rapide pour contrôler le temps de chauffage des échantillons. Pour mesurer la faible variation des signaux CD (i.e. $\Delta DO \ll 10^{-4}$) induite par l'élévation de température des échantillons d'ADN, nous avons transposé le principe de mesure d'un spectropolarimètre statique sur notre montage pompe-sonde. Pour cela, la polarisation de la sonde générée par un oscillateur Titane saphir femtoseconde accordable est alternativement polarisée circulairement gauche et droite à la fréquence de 50 kHz à l'aide d'un modulateur photo-élastique (PEM) couplé à un amplificateur lock-in. Grâce à la modulation rapide de la polarisation de la sonde, nous avons pu ainsi réduire considérablement le bruit de nos mesures ($\Delta DO = 10^{-5}$). Cependant quelques cycles du PEM sont nécessaires pour extraire un signal CD de l'amplificateur lock-in, limitant ainsi la résolution temporelle de notre détection à quelques millisecondes.

Grâce à ce nouveau montage expérimental nous avons mesuré les dynamiques de dénaturation et de renaturation de séquences G4 présentant deux topologies distinctes, parallèle et antiparallèle. Du fait de leur grand polymorphisme, nous nous avons choisi des séquences bien caractérisées, telles que les séquences télomèriques Tel21 (*5'-GGG(TTAGGG)$_3$-3'*) et Tel22 (*5'-AGGG(TTAGGG)$_3$-3'*), la séquence de l'aptamère de la thrombine TBA G4 (*5'-GGTTGGTGTGGTTGG-3'*) et la séquence promotrice oncogène c-MYC (*5'-TGAGGGTGGGTAGGGTGGGTAA-3'*).

Préalablement aux études résolues en temps, en collaboration avec l'équipe DISCO au synchrotron SOLEIL, nous avons mesuré les spectres CD statiques (SRCD) de chacune de ces séquences G4, à différentes températures. Dans la région spectrale en dessous de 260 nm, les signaux SRCD qui dépendent fortement de la séquence d'ADN reflètent majoritairement la contribution des interactions intra-brins des bases. En revanche, aux longueurs d'onde plus grandes ($\lambda > 260$ nm), les signaux SRCD qui

dépendent peu de la séquence d'ADN sont caractéristiques de l'arrangement des guanines dans les structures G4.

Les études de CD résolues en temps des séquences G4 ont révélé, dans tous les cas, des dynamiques de dénaturation et de renaturation de l'ADN biexponentielles, avec des constantes de temps de quelques centaines de millisecondes et quelques secondes qui dépendent fortement de l'amplitude du saut de température. Nous avons également observé des cinétiques de dénaturation et de renaturation de l'ADN très différentes qui suggèrent l'existence de voies bien distinctes de repliement et de dépliement. Par ailleurs, l'étude des effets de la concentration des cations sur la cinétique de dénaturation et de renaturation de Tel21 a mis en évidence leur rôle important à la fois dans les propriétés thermodynamiques et dynamiques des G4. L'ensemble de ces tendances a pu être reproduit à l'aide d'un modèle empirique de diffusion de la population le long d'une surface de potentiel en forme d'entonnoir comparable à celui proposé pour la formation de structures simples double-brin de l'ADN. Toutefois les cinétiques de formation des G4 mesurées sont 2 à 3 ordres de grandeur plus lentes, une différence qui pourrait s'expliquer par les fortes interactions intra-brin des guanines qui doivent être surmontées pour que l'ADN se replie.

Enfin, nous avons mené une étude comparée des dynamiques de dénaturation et renaturation de la séquence TBA G4 avec une séquence tronquée TBA G3 (*5'-GGTTGGTGTGG-3'*) connue pour former une structure triplex (G3). Cette étude a révélé des vitesses de repliement et dépliement du G3 4 fois plus lentes que celles du G4. Ces toutes premières études dynamique d'une structure G3 semblent donc exclure l'hypothèse généralement avancée dans la littérature de la formation d'un intermédiaire triplex de courte durée de vie au cours du repliement des G4 antiparallèles.

# Contents

| Abbreviation | Sequence | G4 structure obtained by NMR with cation: | cation type |
|---|---|---|---|
| human telomeric sequences | | | |
| **Tel21** | *5'-GGGTTAGGGTTAGGGTTAGGG-3'* | mixture | $Na^+$ |
| Tel21 | *5'-GGGTTAGGGTTAGGGTTAGGG-3'* | mixture | $K^+$ |
| **Tel22** | *5'-AGGGTTAGGGTTAGGGTTAGGG-3'* | antiparallel basket-type [1] | $Na^+$ |
| Tel22 | *5'-AGGGTTAGGGTTAGGGTTAGGG-3'* | mixture | $K^+$ |
| 2GKU | *5'-TTGGGTTAGGGTTAGGGTTAGGGA-3'* | hybrid-type form 1 [2] | $K^+$ |
| 22CTA | *5'-AGGGCTAGGGCTAGGGCTAGGG-3'* | antiparallel chair-type [3] | $K^+$ |
| Tel23 | *5'-TAGGGTTAGGGTTAGGGTTAGGG-3'* | hybrid-type form 1 [4] | $K^+$ |
| Tel24 | *5'-TTAGGGTTAGGGTTAGGGTTAGGG-3'* | mixture | $Na^+$ |
| Tel26 | *5'-TTGGGTTAGGGTTAGGGTTAGGGTTT-3'* | mixture | $K^+$ |
| 26TTA | *5'-TTAGGGTTAGGGTTAGGGTTAGGGTT-3'* | hybrid-type form 2 [5] | $K^+$ |
| Tel27 | *5'-TTAGGGTTAGGGTTAGGGTTAGGGTTA-3'* | mixture | $Na^+$ |
| other structures | | | |
| **TBA G4** | *5'-GGTTGGTGTGGTTGG-3'* | antiparallel chair-type [6] | $K^+$ |
| **TBA G3** | *5'-GGTTGGTGTGG-3'* | triplex structure [7] | $K^+$ |
| **c-MYC** | *5'-TGAGGGTGGGTAGGGTGGGTAA-3'* | all-parallel [8] | $K^+$ |
| c-MYC2 | *5'-TGAGGGTGGGGAGGGTGGGTAA-3'* | all-parallel [9] | $K^+$ |
| hTERT | *5'-GGGGCTGGGCCGGGGACCCGGGAGGGGTCACGGGGCGGGG-3'* | ? | $K^+$ |

Table 1: Abbreviation of G4-forming sequences used in the present work. Guanines that participate in the G-quartet are colored blue. Underlined abbrevations are structures, whose CD spectra were studied in the present work. Bold are those whose dynamics were also studied.

# Chapter 1

# Introduction

Nature is beautiful. This is reflected in its simplicity as well as in its complexity. Based on a simple concept of merely four nucleobases, Nature is capable to construct an enormous diversity of living beings. Even by using only guanylic acid, Gellert *et al.* found a non-canonical secondary DNA structure that forms *in situ* containing only Guanines in its center: the Guanine quadruplex (G4) [10] (see Fig. 1.1). Cation coordination in the center tremendously increase G4 stability [11].

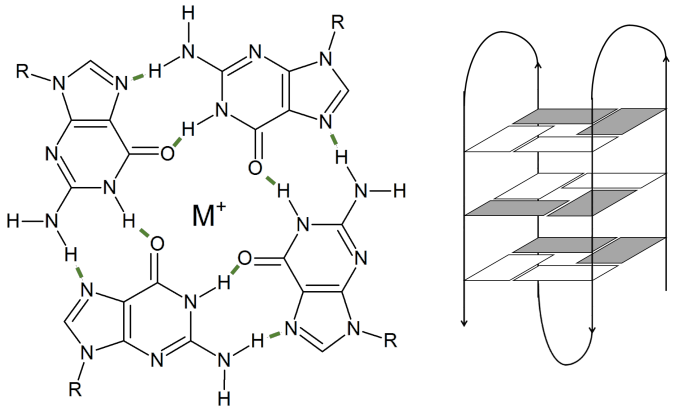


Figure 1.1: Left: G4 quartet. Right: Structure of Tel22 with $Na^+$.

Since the determination of the DNA double helix [12], it has been commonsense that only the double helix forms *in vivo*. However, advances in recent years have proven that DNA can adopt a variety of structures. Among them G-quadruplexes (G4) are non-canonical structures that result from the hydrophobic stacking of guanine-tetrads stabilized by cations (typically $Na^+$ and $K^+$). Recently, the interest in G4 has grown considerably, due to the identification in various eukaryotic and prokaryotic genomes, of many guanine-rich regions capable of forming such structures in solution, in vitro. There is now a body of experimental evidence for the occurrence of G4 in vivo and their participation in several biological functions related to their

folding mechanism, such as chromosomal integrity, gene replication, transcription and translation [13, 14].

Notably, the interest in G4 has been amplified by the understanding of telomere maintenance and the discovery of telomerase [15]. In humans, the telomeric sequence is rich in guanine and thus has the potential to form G4: $(GGGTTA_n)$ [16, 17]. Telomerase is an enzyme that maintains the length of telomeres [18]. In approximately 90% of all malignant tumors, telomerase has been detected and inhibits tumor cell to undergo senescence due to telomere shortening [19]. Consequently, telomerase and human telomeres have become good potential target structures for novel drugs in cancer treatment [20, 21]. The telomeric overhang has to be bound to proteins so that telomerase can interact [22]. Therefore, by adding ligands that favor G4 formation in this overhang, telomerase cannot interact and telomere maintenance is impeded.

Besides the design of ligands targeting G4, most of the experimental studies carried out so far on G4, have been focused on the characterization of their 3D structure and stability highlighting their large conformational diversity. Although the literature devoted to G4 has exploded considerably, experimental studies of their folding remain scarce and are mainly focused on the human telomere sequence. Elucidating G4 folding mechanisms is however a crucial issue for understanding their biological functions. Despite recent experimental and theoretical efforts, there is not yet a clear picture these mechanisms [20–22]. One of the reasons may be due to fact that most of the studies of G4 folding have focused on the highly polymorphic human telomeric sequence, due to its biological pertinence [23, 24]. On the other hand, most of the studies of G4 folding have yielded a limited structural information with time resolution generally limited to milliseconds or seconds.

Laser-induced temperature jump is an established method to study unfolding dynamics of proteins [25, 26]. However, it only allows to monitor dynamics up to several ms *in situ* [27].

Laser-induced temperature jump (T-jump) is an established method to study the folding dynamics of proteins and to a lesser extend DNA hairpin [26, 28]. Combined with various optical detections, it provides the opportunity to probe conformational changes of biomolecules on shorter time scale than single-molecule spectroscopy or rapid mixing techniques often used to study G4 folding. The principle of T-jump is to heat up the solvent with an infrared laser pulse and to monitor the resulting DNA denaturation as a function of time. However most of experimental studies performed on DNA have used fluorescence or FRET detection that requires the use of covalently attached fluorescent labels that may affect the thermodynamic and the mechanical properties of G4. In contrast, circular dichroism can provide information on the

secondary structure of biomolecules [29].

Recently, our laboratory has implemented a T-jump set-up combined with time-resolved CD detection [28], for measuring the folding dynamics of biomolecules within the time window of ca. 10 ns and 1 ms. In the frame of my PhD work, we have extended this method to probe slower dynamics on the time window of a few ten milliseconds up to the seconds. With this new setup, both denaturation and consecutive renaturation of DNA can be measured. In combination with our conventional T-jump, it becomes therefore possible to measure the thermal-induced changes of DNA secondary structure over a time-scale of more than 8 orders of magnitude.

This manuscript is organized as follows:

- Chapter 2 addresses the state of art of the knowledge on G4 focusing on the dynamical studies of their folding.
- Chapter 3 is dedicated to the description of the principles and the implementation of the used experimental techniques.
- Chapter 4 describes the static synchrotron radiation CD spectra of various G4-forming sequences displaying distinct topologies (antiparallel, hybrid and parallel) measured in collaboration with DISCO/SOLEIL.
- Chapter 5 discusses the results of our time-resolved studies of human telomeric G4-forming sequences (Tel21 and Tel22) in $Na^+$ phosphate buffer.
- Chapter 6 presents the results of the shortest forming G4 sequence: the thrombin binding aptamer G4 (TBA G4) [6] as well as on its related triplex structure (TBA G3) [7]. Finally, the dynamics of a modified G4-forming sequence of the oncogene promoter sequence, c-MYC, have been investigated and are presented here.
- Chapter 7 gives the overall conclusion of this work with some perspectives.

# Chapter 2

# An overview of G-quadruplex structures

Such as we are made of, such we be.

*William Shakespeare*

## Contents

## 2.1 What defines a G-quadruplex

The double helical structure of B-DNA was determined in 1953 [12]. Various alternative structures have been discovered since. One of these structures is the so-called guanine quadruplex (G4). It seems to be the most biological relevant structure after the B-form due to its high stability at neutral pH [13, 30–32]. G4 consist of a DNA sequence that is abounding in guanine bases. In a G4 core, a planar tetrad made of 4 guanine bases is found as well as cations such as $K^+$ and $Na^+$ [1]. It is called a G-quadruplex when there are at least two tetrads stacked on each other [35]. In these tetrads the guanines are connected to each other by Hoogsteen-hydrogen bonds (see Fig. 1.1). Due to the negative dipolar charge of the oxygen atoms pointing to the tetrad center, a cation is necessary to compensate these negative charges and stabilize the G4 structure.

Formation of G4 was first discovered in 1962 [10] in a concentrated solution of guanosine monophosphate (GMP) that formed a gel over night. In that case, the G4 structure is made of guanine monomers. G4 can also form from association of different DNA strands. Tetramolecular, trimolecular as well as bimolecular and monomolecular G4 are possible [36]. Typical intramolecular G4 forming sequences are as follows: $G_{\geq 3}N_xG_{\geq 3}N_xG_{\geq 3}N_xG_{\geq 3}$ with $x$ between 1 and 7 and $N$ being any base. Longer connections destabilize the G4 structure [37]. My work, however, focuses only on intramolecular G4, e.g. monomolecular G4. In that case, guanines that form the quartets are connected directly by a negatively charged phosphate backbone in a way that adjacent bases do not participate on the same quartet, but are stacked on each other [38]. In human telomeric sequences (*5'-(GGGTTA)$_n$-3'*) three quartets form a G4. The stacked quartets are slightly twisted [39] and are spatially separated by ca. 3.25 Å [40].

In the example of the human telomeric G4 sequence Tel22 with $Na^+$ [1] three quartets are formed (see Fig. 1.1 right). The four *GGG* run are connected by three loops. The length of the loops are known to influence the relative orientation of the *GGG* runs, which can be all parallel (all in same direction), antiparallel (all in opposite direction) or hybrid (one in opposite direction). Besides the primary DNA sequence, the environment can also influence the relative direction of the *GGG* runs, such as the nature of cations, the presence of cosolvents or crowding agents.

[1] Exeptions exist in which guanines are replaced by cytosines [33] or some guanines are missing [34]

## 2.2 Polymorphism of G-Quadruplexes

G4 can form a vast variety of structures. For a better understanding, we discuss here the features that allow to discriminate different G4 structures. The human telomeric G4 have been the most studied ones up to date [41]. These sequences exhibit the largest structural polymorphism [42]. All different human telomeric G4 structures found up to now are depicted in Fig. 2.1.

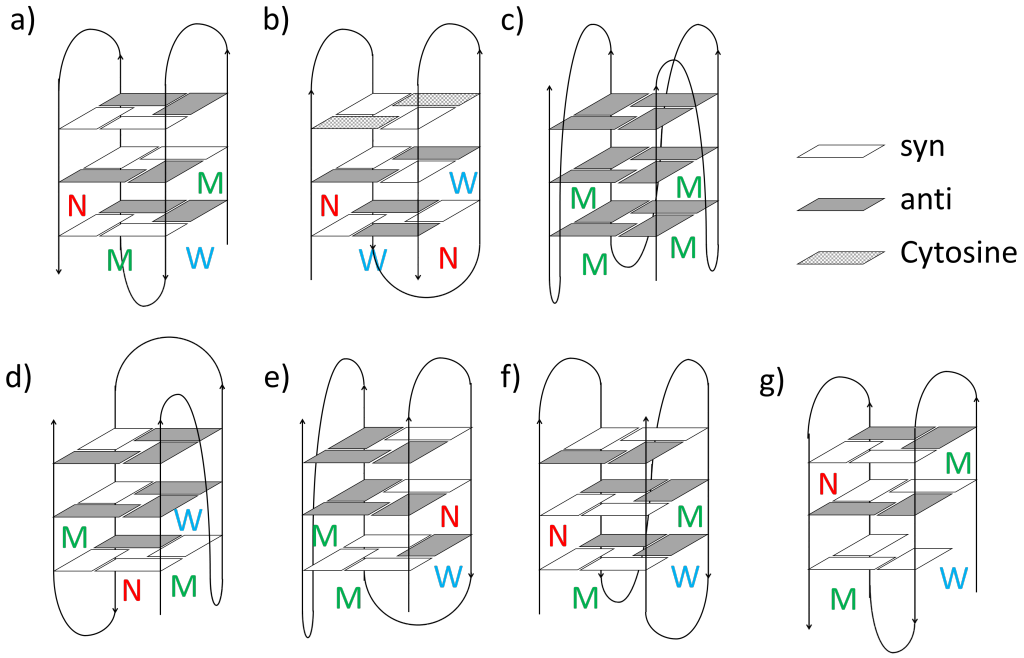


Figure 2.1: All characterized telomeric G4 topologies. (a) Antiparallel basket-type form [1]. (b) antiparallel chair-type structure [3]. (c) all-parallel G4, propeller-type form [43]. (d) (3+1) [4, 44]. (e) (3+1) form 2 (hybrid2) [4]. (f) antiparallel G4, sequence variant *d[(TTAGGG)$_3$TTA($^{Br}$G)GGTTA]* [45]. (g) basket-type with only two quartets observed [33]. *anti* guanines are colored grey and *syn* guanines are colored white. M, N and W indicate medium, narrow and wide grooves, respectively. The strand direction goes from the 5' to the 3' end.

The first crucial factor that narrows significantly the number of possible G4 conformations is the primary DNA sequence. For example, a sequence that contains only one *GGG* run forms tetrameric G4 [36]. The second important feature that enhances G4 polymorphism is the *syn* and *anti* orientations of the glycosidic bond angles of guanosines. Fig. 2.2 shows the two orientation angles. As indicated in Fig. 2.2, guanosines with *syn* and *anti* orientation angle are colored in white and in grey, respectively. These colors will consequently be maintained throughout the manuscript. In Fig. 2.2 bottom left, a rotation sense is depicted. Looking from the top of the

G4 (indicated by the 5' end) to the G4 one can see that the twist of the quartets is anti-clockwise (e.g. right-handed). This is valid for all G4 with only one exception (see Phan *et al.* [46]). As previously mentioned, the relative strand orientations of the *GGG* runs in the G4 can be either antiparallel (Fig. 2.1(a,b,f,g)) [1], all-parallel (c) or hybrid (d,e) [47].

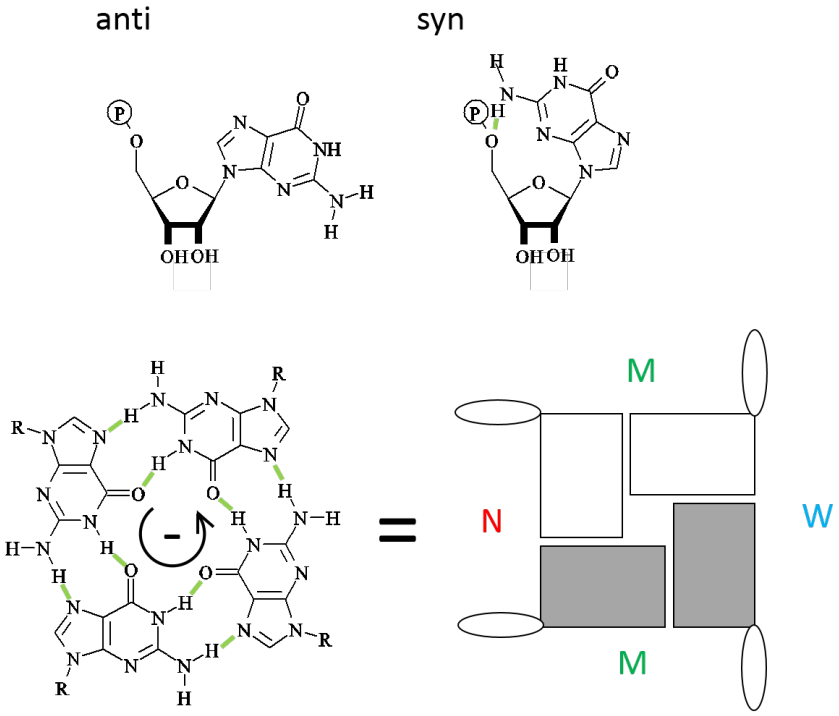


Figure 2.2: Top: Guanosine monophosphate with an *anti* and *syn* glycosidic angle. Bottom left: G4 quartet with anitclockwise rotation (-) from the donor *NH* to the acceptor *C=O* hydrogen bonds. Bottom right. Illustration of different groove sizes. *syn* guanines appear white and *anti* guanines in grey. M, N and W indicate medium, narrow and wide grooves, respectively.

A third feature of G4 structures is the loop orientation, which can be diagonal, lateral (edgewise) and external (propeller) [36]. All quadruplex structures have four grooves, whose dimensions depend on the overall topology and the nature of the loops. G4 grooves are classified according to their size: Narrow (N), medium (M) and wide (W).

### 2.2.1 Effect of Loop length

The number of nucleotides between G runs is known to determine partly the folding topology of G4. A sequence $G_3T_1G_3T_1G_3T_1G_3$ with solely one nucleoside between the G runs forms the parallel propeller-type G4, most likely independent of flanking loops [48]. Bugaut and Balasubramanian studied a number of sequences containing only G and T with loops containing only T. By varying the loop lengths, they observed some trends: parallel G4 are the most stable structure, when there is only one nucleoside in their loops [49]. In general, an increase of the loop length decreases the G4 stability up to 5 nucleosides. Longer loops do not significantly destabilize the G4 structure compared to 5 thymines [49]. However, this cannot be used as a general statement for all sequences. The Mergny group performed a more general essay, in which they used a sequence $G_3L_aG_3L_bG_3L_cG_3$ [50]. As loop bases $L_{a,b,c}$ they chose T or TTA loops. Salt conditions were either 100 mM $K^+$ or $Na^+$. An increase of loop length showed for both cations a decrease of the melting temperature $T_m$. It is, however, more pronounced and more linear for $K^+$ than for $Na^+$.

Additionally, some sequences with loop lengths of one to three nucleotides can even form G4 in the presence of the complementary DNA strand, which reflects their stability [51–53]. A G4 formed by the c-MYC promoter sequence in 100 mM $KCl$ remains stable even in the presence of its complementary DNA strand up to at least 70°C [54]. It has two loops with one nucleotide and one loop with two nucleotides and forms a parallel structure. It has been shown in some cases that for two linking loops containing only one nucleoside, the length of the third one does not influence the G4 topology in potassium. In that case only parallel quadruplexes are formed [49, 50, 55]. With sodium there can be more diversity [50]. Quadruplexes with loop lengths of three nucleosides do not fold only in one specific topology. They rather form a mixture of parallel and antiparallel conformations, depending on their sequence, the nature of the cations in solution and the presesence of cosolvent [56]. However, parallel quadruplexes are energetically more favorable, even for loop lengths of three nucleosides [56].

### 2.2.2 Effect of Cations

#### 2.2.2.1 Specific Binding

There are several factors that contribute to G4 stabilization: stacking interactions [57] between the nucleosides, hydrogen bonding [53, 58], solvation and cation binding [59–61]. Without cations, G4 do practically not form. Cations located inside the G4 core play a key role in the G4 stability and are classified as specific binding cation.

Those cations can be located inside the G-tetrad plane or between two G-tetrads.

A cation located in the plane of a G-quartet has 4 oxygen binding sides, one from each Guanine. When it is located between two planes, there are even 8 oxygen atoms available for coordination. Due to the nature of the cavity, no anions enter. Hence, they have a weak influence on G4. Gu *et al.* have stated that according to their *ab inito* calculations the coordination of the cation contributes more to the G4 stability than the Hoogsteen hydrogen bonds between the guanines [62].

In literature it is often stated that $Na^+$ is located inside the plane of G-tetrads because it is small enough. In contrast, $K^+$ and $NH_4^+$ for example are too large and are thus located between the G-tetrads of the G4 core [63]. In fact, there are other parameters that contribute to the exact location of cations in the G4 core [61]. For instance, the presence of several cations inside the G4 core induces ionic repulsion (example with $Na^+$ see [64]). Furthermore, the location of cations inside the G4 core also depends on the presence of the cations in the phosphate backbone or on top of the terminal G-quartet.

#### 2.2.2.2 Non-specific binding and ionic strength

Once all the cavities are filled with cations, no more cations go inside the G4 core. However, the negatively charged phosphor backbone allows more ions to be coordinated in a non-specific manner outside the G4 core. Gray *et al.* proposed that Tel22 can bind up to five $K^+$ in a non-specific manner [65]. A study on c-MYC has focused on the effect of $K^+$ concentration on G4 stability. It has been shown that addition of cations once the G4 is filled, leads solely to non-specific binding. The increase of the concentration of cations results in a stabilization of the quadruplex structure, reflected by an increase in the melting point $T_m$ [66].

#### 2.2.2.3 G4 stability according to cation type

Physiological relevance is especially assigned to $K^+$, due to its high intracellular concentration [67], but also to $Na^+$. The melting temperature ($T_m$) for G4 is in general higher with $K^+$ than with $Na^+$ [68, 69]. Since specific binding inside the G4 core necessitates a complete dehydration of cations, this effect is due to their sizes and their energies of hydration. $K^+$ is rather coordinated to 8 oxygen atoms, whereas $Na^+$ coordinates to 4 oxygen atoms. Hence, these different coordinations and G4 oxygen-ion distances are partly responsible for the difference in the free energy between $Na^+$ and $K^+$ [69]. Hydration energies for monovalent ions depend roughly on their ionic radii. Larger radius means smaller hydration energy. The difference between the free

energy distributions of coordination and hydration determines the selectivity of G4 to a specific cation [68, 70–72]. Gray *et al.* have shown that by adding $K^+$ to an already formed human telomeric G4 (Tel22) with $Na^+$, $K^+$ replaces $Na^+$ and changes the G4 structure [73].

G4 do not only show different stabilities in the presence of different cations, but they also form different structures [74]. Therefore, the melting temperatures of G4 with $K^+$ and $Na^+$ cannot be compared directly. In the case of Tel22, two hybrid structures form in the presence of $K^+$, whereas merely one antiparallel (basket-type) structure forms in the presence of $Na^+$ [73]. Interestingly, NMR measurements have shown that cations in the G4 core can exchange with the bulk [75]. $^{23}Na$ NMR measurements have shown that it takes 180 $\mu$s at 20°C for a sodium cation to be exchanged by another one in a dimeric G4 made of the sequence $d(G_4T_4G_4)$. This suggests that no quadruplex opening or conformational change is needed, since bimolecular G4 folding dynamics are known to be slower [76].

In addition to $Na^+$ and $K^+$ there are several other ions whose presences induce quadruplex formation: $Rb^+$, $Cs^+$ [60, 77],$NH_4^+$ [78, 79] and $Tl^+$ [80]. The work of Wong *et al.* revealed the following order for G4 stability: $K^+ > NH_4^+ > Rb^+ > Na^+ > Cs^+ > Li^+$, where $K^+$ yields the most stable G4 and $Li^+$ the weakest one [81]. Subsequent studies revealed a slightly different order: $K^+ > NH_4^+ > Na^+ > Rb^+ > Li^+ \geq Cs^+$ [59–61], as illustrated in Fig. 2.3. As one can see, between $K^+$ and $Na^+$ decreasing the cation size exhibits the same tendency as decreasing stabilization effect. $Li^+$ and $Cs^+$ are too small and too big, respectively, to fit in the G4 cavity [36].

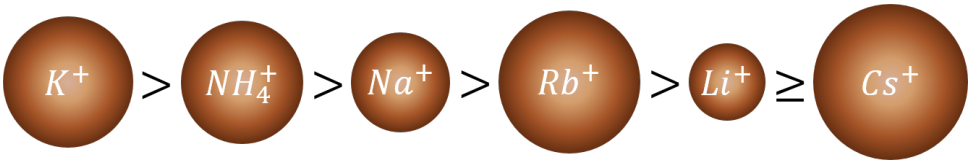


Figure 2.3: Monovalent ions with decreasing stability effect on G4 from left to right. $K^+$ being the most stabilizing one, $Cs^+$ is the least stabilizing one according to [61]. Relative ion radii are reflected by size of spheres.

Not only with monovalent cations G4 can form, but also with divalent cations as $Sr^{2+}$ [60, 82]. As well as for monovalent cations, a similar trend in quadruplex stability is observed: $Sr^{2+} >> Ba^{2+} > Ca^{2+} > Mg^{2+}$. However, it seems that there is a more complex effect of bivalent cations on G4 stability. In the case of Tel24 $Sr^{2+}$ stabilizes G4 even better than $K^+$, though they have similar ionic radii [36]. Other

interesting effects, such as destabilizing effects of quadruplex structures by bivalent cations have been reported [36, 66, 83]. Since in the present work only monovalent cations have been investigated, these effects are not discussed here.

### 2.2.3 Effect of Cosolutes and Molecular Crowding

Human telomeric quadruplexes can adopt a variety of different structures depending on their sequence. In $K^+$ solution four structures have been reported so far [84] (see Fig. 2.1(b,d,e,g)). A change from solution to solid state, e.g. G4 crystals, the very same sequence Tel22 with the same amount of cations changes from two coexisting hybrid to a sole all-parallel structure (Fig. 2.1). Since in crystals the relative distance between two G4 is reduced, one can think that this structural change is due to molecular crowding. Heddi and Phan have found that human telomeric G4 with $K^+$ adopts the all-parallel topology under molecular crowding conditions [84]. This is an important result, since crowding effects are expected to reflect *in vivo*-like environment [85]. Crowding conditions result in a reduction of G4 hydration [85]. Under such conditions the G4 structure with the minimum hydration energy should prevail. According to the work of Heddi and Phan this is the propeller-type quadruplex [84]. A commonly used cosolute is polyethylene glycol (PEG) [36, 86]. It is considered to not interact with the quadruplex. However, a study of the group of Chaires suggests that PEG rather binds to the grooves of propeller-type quadruplexes and does not act as a crowding agent [87]. Other used cosolutes are ethanol, Ficoll 400, acetonitrile or DMSO [84]. It was shown that for all these agents the same parallel quadruplex structure formed out of different other structures [84], which is in accordance with the suggestions of Heddi and Phan.

## 2.3 Biological Pertinence and Applications of G4

It is extremely difficult to detect G4 *in vivo*, in part because reagents to detect G4 have been laking [32]. Nevertheless, advances in recent years have shown accumulated evidence that G4 also form *in vivo* [13, 30–32, 88]. Some remaining questions are: where would they appear and what would be their exact role. An algorithm developed by Balasubramanian's group has scanned the human genome in search of potential quadruplex forming sequences. In their work of 2005 they have considered only sequences that can form 3 or more tetrads [37]. The following sequence motif was used: $G_{\geq 3}N_xG_{\geq 3}N_xG_{\geq 3}N_xG_{\geq 3}$. With those conditions, a number of over 350,000 potential G4 forming sequences have been identified. These motifs are found at very specific regions of the genome, such as telomeres or promoter regions.

Human telomeres consist of thousands of DNA repeats (TTAGGG) and have a single-stranded 3'-tail. Short telomeric repeats *in vivo* form quadruplexes (see e.g. [89]). Furthermore, telomeres play an important role in cancer cells and aging. Consequently, they are considered as potential targets for the development of anti-cancer therapy [90].

There are some speculations about the possible role of G4 within the genome. G4 in telomeres might be there to protect the genome. G4 could also help telomeres to form DNA t-loops more easily [30]. Furthermore, it is possible that G4 form during the replication process when duplex DNA is unwound. The fact that a couple of helicases exist to unwind quadruplexes supports this suggestion [91].

Interestingly, many promoters of oncogenes also contain G4 forming motifs: c-KIT, wt c-MYC, k-RAS, B-RAF, BCL-220 and RET [90].

### 2.3.1 G4 Ligands

Due to the potential biological relevance of G4 there is a big field of research dedicated to the design of G4 ligands. This approach to target specific DNA regions by either stabilizing G4 or inducing their formation provides a strategy for elaboration of new disease treatments. Special interest has been dedicated to the inhibition of telomerase and promoter regions of some oncogenes to pave the way for effective anti-cancer targets [90]. Another target can be potential RNA G4 motifs of ribosomal viruses as HIV [92].

In promoters, ligand binding is expected to lead to an inhibition and to down-regulate the expression of the targeted gene [93]. Telomerase and telomere were the first targets for small ligand molecules [94]. However, up to now no ligands have been reported that are 100% specific to one particular G4 structure [90]. Under physiological conditions, it seems furthermore almost impossible to specifically target only one specific G4. Single G4 may be not the best candidates. Some studies show that higher order G4 structures have very particular surfaces to provide a higher ligand specificity [95].

Currently, more than 1000 G4-binding ligands have been reported. They are classified in two main groups so far: i) hetero-aromatic polycyclic ring systems, and ii) macrocycles with linked hetero-aromatic compounds [101]. One specific binding side of a G4 is its planar surface G-quartet. This is still the most explored one, but side grooves can also serve to increase ligand specificity. Theoretical studies have suggested that groove dimensions and loop sizes are crucial factors for the design of selective ligands [102].

In Fig. 2.4 four different examples of small G4 binding agents are depicted. The

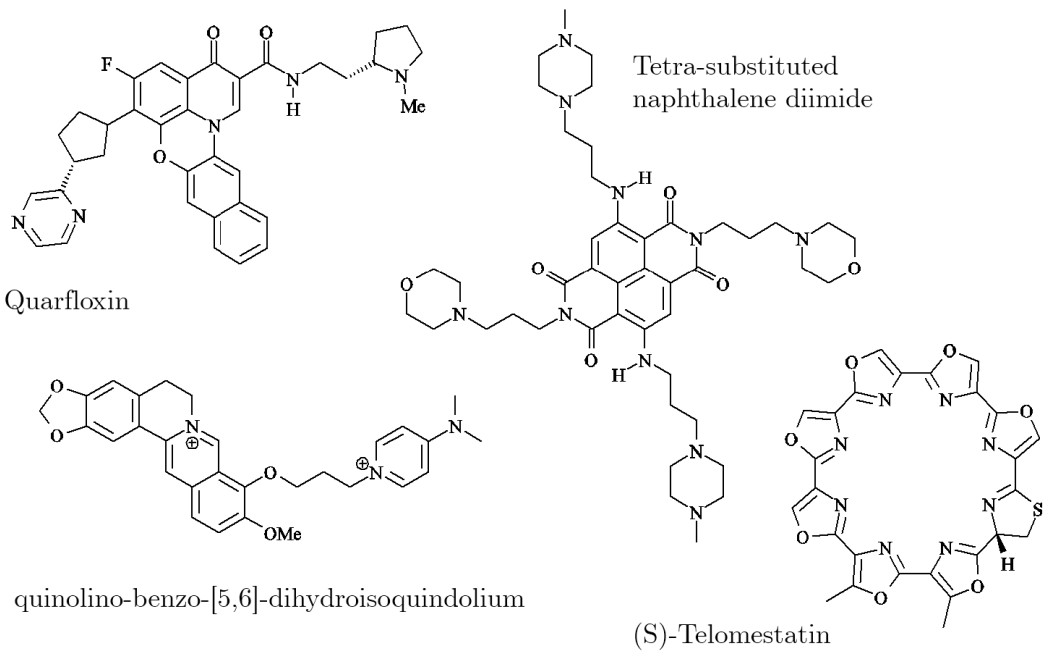


Figure 2.4: Examples of G4 ligands. Quarfloxin, the first G4-binding molecule that underwent clinical trials [96]. Tetra-substituted naphthalene diimide, containing two different side-chaines to increase G4-binding affinity [97]. Five-ring quinolino-benzo-[5, 6]-dihydroisoquindolium, having a high specifity to discriminate between c-MYC and telomeric G4 [98]. Telomestatin, cyclic G4 ligand with a high selectivity and affinity for G4 [99, 100]

.

first one, quarfloxin is the only compound used for clinical trials so far [96]. It consists of a heteroaromatic planar core with two distinct side chains. The tetra-substituted naphthaline diimide illustrates another type of highly potent G4 binding agents [97]. The third example is a five-ring quinolino-benzo-[5,6]-dihydroisoquindolium. Its core consists of a positively charged heteroaromatic structure and in its sole side chain it exhibits a charged pyridine ring to increase water solubility [90]. This compound preferentially binds to wt c-MYC (parallel G4) vs. telomeric G4 (antiparallel G4) [98]. The last example is telomestatin, a macrocycle without a polycyclic core. It displays a high G4 selectivity and affinity over duplex DNA [99, 100]. For example, it stabilizes specifically anti-parallel human telomeric G4 [103]. But it has a very low aqueous solubility. In addition, its demanding isolation procedures and current synthetic approaches are reasons why it has not been developed further as a therapeutic agent [36].

### 2.3.2 G-Quadruplexes as nanomachines and -devices

Beside the interest of G4 in therapeutics, their polymorphism and self-assembly properties can serve to create nanodevices. A few examples of G4 applications are presented below.

For instance, functionalization of G4 has been performed by the group of Fettinger (e.g. [104]) who created lipophilic G4 by using modified lipophilic guanosine monomeres. The phosphate group has been replaced by a nonpolar silicon residue and the hydroxy groups have been protected with an acetalic nonpolar residue. Hence, these nanostructures are stable in nonpolar solvents. Still needing a cation to form, they can extract salts from water into nonpolar solvents [40]. While anions have little influence on the formation of G4 made of oligonucleotides, the group of Davis showed that they influence the properties of lipophilic G4 in organic solvents [40]. By increasing the basicity of anions, the rearrangement of G4 octamers can also be halted [105]. Hence, anions could be used to control regioselectivity of G4 assemblies. Additionally, lipophilic guanosines and folate analogues form dynamic crystals, that can be controlled by changing parameters as ions, solvents or temperature [40]. Folate derivatives of guanines, for example, form a smectic phase. Introducing cations changes the assembly to a columnar discotic phase similar to abreast columns of stacked guanine quartets [106]. Based on NMR studies that show that $K^+$ are only bound for milliseconds to G4 [107], oligomers or polymers of stacked guanine quadruplexes can function as ion channels. In this respect, lypophilic G4 could be used as transmembrane ion channels [40].

On the other hand, G4 built out of monomers can be used as self-assembly ionophores [40]. Ion-selectivity can be achieved by influencing the non-covalent bonds. This is a big advantage in terms of synthesis, since established ionophores such as crown ethers and calixarenes demand a considerable synthetic effort to have ion-selectiviy [108, 109].

Nanometer assemblies of G-rich DNA have been discovered by the groups of Henderson and Sheardy [110]. Marsh and Henderson established the term "G-wires" to describe these parallel-stranded superstructures (see Fig. 2.5). This term is further justified, since G4 is capable of transporting electric current [111]. In this respect, G4 assemblies are interesting candidates, since guanine has the lowest potential of all ncucleic bases. "G-wires" can also be used as a cation channel for $Na^+$, $K^+$ and $NH_4^+$ with the best conduction for $Na^+$ [112].

According to Mergny's group, switching between double-stranded duplex and quadruplex can be considered as a nanomolecular machine [113]. Conformational changes of a G-rich 21-mer DNA have been studied. An alternating switching be-

tween duplex and G4 was achieved by adding C-rich strands to unfold quadruplexes and G-rich strands to fold G4 again. In a more recent work a combination of i-motif (cytosine analogue to G4) and G4 was synthesized in one strand [114]. Formation of the two structures can be controlled independently by changing the pH and the concentration of cations in solution. At acid pH without cations, solely the i-motif forms, while basic pH with cations mostly G4 is present. Many more applications do already exist. The potential of further guanine-based nanodevices is paramount and depends on the new G4 structures that will be discovered in the future.

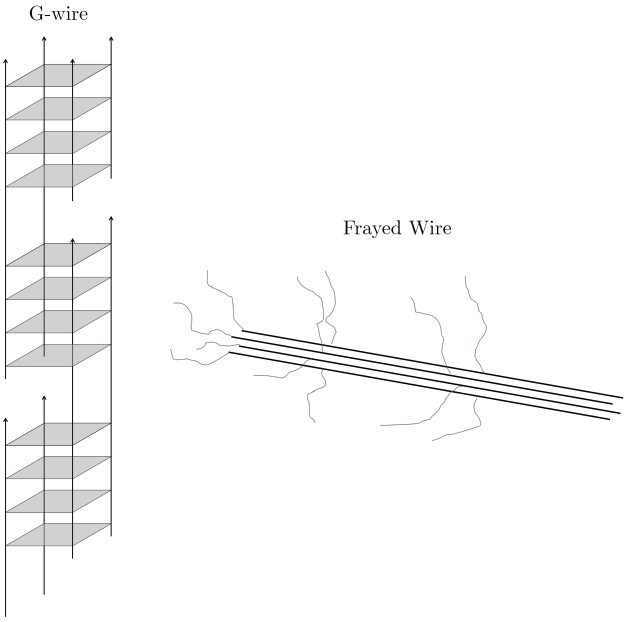


Figure 2.5: Scheme of a G-wire formed out of $d(G_4T_2G_4)$ and a frayed wire formed out of $d(G_{15}A_{15})$ according to [40].

## 2.4 State of Art of the Investigations on G4 Folding

### 2.4.1 Structure characterization of G4

#### 2.4.1.1 UV & CD melting experiments

The simplest way to verify G4 formation is to monitor the spectral changes induced on the absorption or the circular dichroism spectra of DNA as a function of the temperature. Upon G4 formation, a hyperchromism is generally observed around 295 nm concomitant with a hypochromism around 260 nm in the absorption spectrum of DNA. The melting temperature ($T_m$) corresponds to the temperature at

which half of the population is folded and half is unfolded ($K_{eq} = 1$). By taking the first derivative of the melting curve, one can simply determine $T_m$. Further analysis assuming a folding model allows the determination of thermodynamic parameters of the reaction: the Gibbs free-enthalpy ($\Delta G^0$), the enthalpy ($\Delta H^0$) and the standard entropy ($\Delta S^0$) [115, 116].

In addition to absorption and CD, formation of G4 can be also investigated by fluorescence either by FRET or by substitution of certain bases in the G4 loops by a fluorescent base such as 2-aminopurine (2-AP) [117]. While absorption and CD yield global information of the secondary DNA structure, FRET provides insight into the distance between the 5' and 3' endings of the G4 strand. This allows to distinguish between the relative positions of the first and last *GGG* run [23]. On the other hand, fluorescence of 2-AP monitors the stacking in the different G4 loops in the G4 [23]. As a matter of fact, fluorescence melting provides a better sensitivity than UV absorption and CD melting.

#### 2.4.1.2 Gel Electrophoresis

Polyacrylamide gel electrophoresis (PAGE) is a method to separate molecules according to their weight [118]. Therein, molecules in a porous polymer gel move due to an applied electric field. Their mobility depends on their molecular weight, i.e. number of nucleotides. By adding a denaturating agent, one can verify the sequence's length and its purity. Additionally, in a nondenaturating environment, the DNA secondary structure remains intact. Therefore, it allows to distinguish between unimolecular G4 and higher order G4 [119].

#### 2.4.1.3 X-ray Crystallography

X-ray crystallography provides information about the molecular structure of G4 crystals [43]. In 2002 the first crystal structure of the human telomeric sequence Tel22 in the presence of $K^+$ has been determined [43]. Tel22 adopts an all-parallel topology in crystals. With X-ray crystallography, positions of most of all atoms of G4 can be determined e.g. the absolute configuration. Furthermore, parameters such as distance between two G4 quartets, the tilt angles and the cation positions and distances can be determined [43]. Despite the fact that X-ray crystallography is a powerful tool to determine the absolute configuration of G4, X-ray radiations can easily destroy DNA [120]. Additionally, structure determination via X-ray scattering needs monocrystals, whose growth is a long and demanding process (see e.g. [43]).

#### 2.4.1.4 NMR

In contrast to crystallography, nuclear magnetic resonance spectroscopy (NMR) enables structure determination in solution, which is closer to *in vivo* conditions [1]. Cross peaks of the nuclear Overhauser effect (NOE) reveal information about the spatial distance between H protons. From that structural information is derived. The general procedure to determine the full structure is as follows [1, 3, 4, 7, 33, 45]: Spatial proximity of atoms yields a first idea of G4 structure. Consecutively, it is used as an input to perform structure calculations. NMR analysis reveals important parameters such as tilt, loop orientation, *syn*, *anti* orientation of the glycosidic bond angles of guanines and yields information about spatial distances [1, 3, 4, 7, 33, 45].

#### 2.4.1.5 CD spectroscopy

Crystallography and NMR spectroscopy are used for a precise structure analysis. However, these experiments are time-consuming and require a large amount of DNA. Circular dichroism (CD) spectroscopy provides a fast method to check G4 structures and requires a small amount of DNA. CD spectroscopy, which is very sensitive to the guanine arrangement in the G4 core, allows to distinguish between parallel, antiparallel and hybrid topologies, as shown in Fig. 2.6. The origin of the CD signals in G4 will be discussed in Sec. 3.2.5.

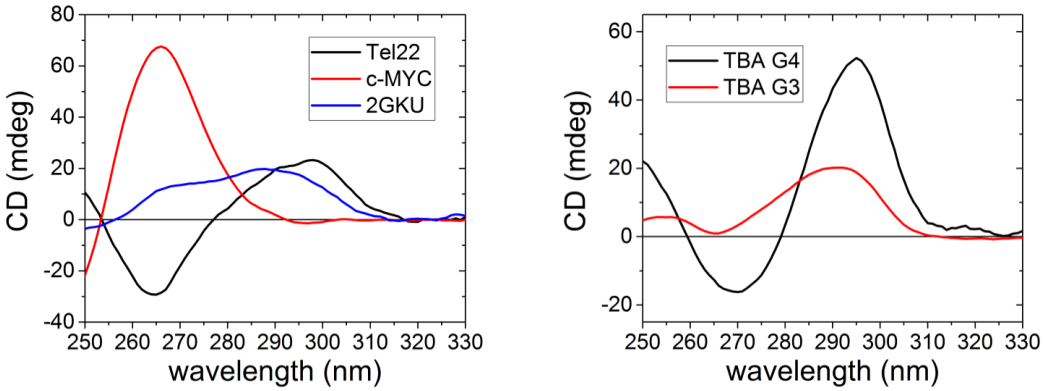


Figure 2.6: Left: CD spectrum of an antiparallel basket-type G4 (Tel22 with $Na^+$), all-parallel G4 (c-MYC with $Na^+$) and hybrid1 G4 (2GKU with $K^+$). Right: TBA quadruplex (G4 with $K^+$) and triplex (G3 with $K^+$), respectively.

In a parallel G4 (c-MYC), in which all glycosidic bond angles of guanines are in *anti*, an intense peak is observed at 265 nm [36]. In antiparallel G4 (Tel22), in which the glycosidic bond angles of guanines exhibit both *syn* and *anti* orientation, one positive peak at 295 nm and a negative one at 265 nm are observed [36]. Hybrid G4

structures (e.g. 2GKU) are mixtures between antiparallel and parallel topologies and display two positive peaks centered around 270 nm and 290 nm [121].

## 2.4.2 Studies of G4 folding/unfolding pathways

### 2.4.2.1 Rapid mixing experiments

The folding/unfolding processes of biomolecules are known to be complex mechanisms involving multidimensional energy landscapes reflecting all possible conformations. From the experimental point of view, it is quite difficult to determine these energy landscapes, since the perturbation that is used to trigger folding or unfolding processes may also influence its shape [22]. One way to study folding is to use rapid-mixing techniques. In the absence of cations G4 structure hardly form [21]. Therefore, G4 formation (folding) can be induced by a rapid adjunction of cations in an aqueous solution of unfolded DNA [2, 23, 73, 122–124]. Another method, less used, consists to add a solution of the complementary DNA strand to a solution containing the folded G4, in order to induce a destabilization of the G4 structure for measuring the unfolding dynamics [125, 126]. The group of Chaires has monitored cation-induced folding kinetics of several human telomeric sequences by stopped-flow combined with absorption, CD detection, FRET and 2-AP fluorescence [23, 73, 122].

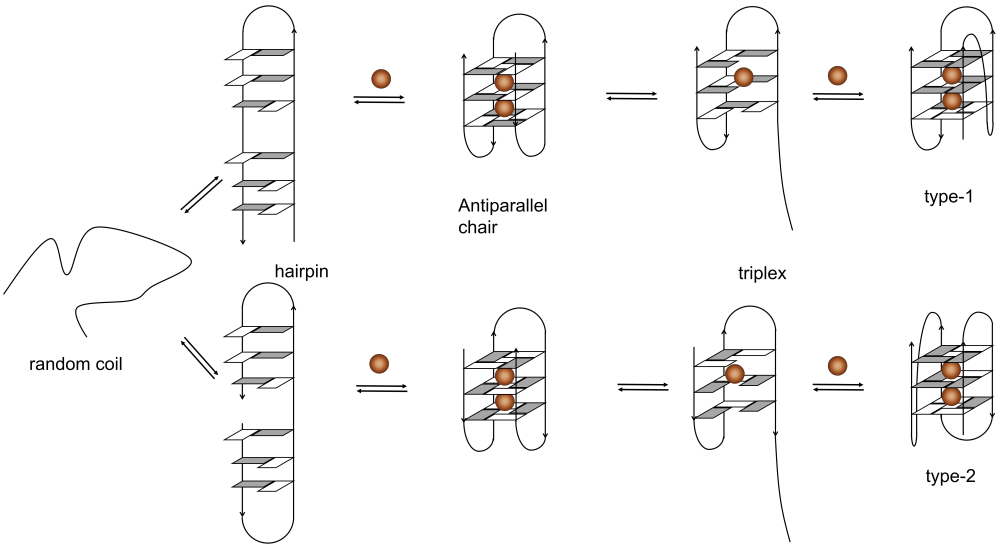


Figure 2.7: Proposed folding pathway for Tel22 in $K^+$ [23, 127]. A random coil undergoes transition to a hairpin structure, followed by the formation of antiparallel chair-type G4. A consecutive triplex formation allows the folding to the corresponding hybrid structure. $K^+$ is indicated by brown spheres.

For instance, their recent study of Tel22 in the presence of 25 mM $K^+$ has re-

vealed multiphasic CD changes spanning the time scale of hundred milliseconds to a few hours [23]. Under these conditions, Tel22 is known to form two structures, hybrid1 and hybrid2. In order to explain the observed multiphasic CD changes, Gray *et al.* have proposed a four-step sequential folding mechanism ($U \rightleftharpoons I1 \rightleftharpoons I2 \rightleftharpoons I3 \rightleftharpoons F$, see Fig. 2.7) leading to the formation of the equilibrium of the two hybrid conformations [23]. From the singular value decomposition (SVD) analysis of the time-resolved CD spectra, they propose the formation of an antiparallel topology ($I2$) in ca. 100 ms from a short-lived hairpin structure (not observed here), followed by the formation of a triplex ($I3$) in ca. 4000 s which rapidly form the final hybrid conformation in ca. 800 s, as shown in Fig. 2.7.

In contrast to the group of Chaires, very recent studies of Tel22 by stopped-flow combined with FRET detection have revealed faster folding kinetics, upon addition of 100 mM $K^+$ [124]. The observed biphasic fluorescence decays (200 ms and 1.5 s) have been suggested to arise from the existence of multiple folding pathways rather than a sequential folding process. Interestingly, biphasic folding kinetics have also been reported previously by the group of Balabrasumanian for the human telomeric sequence Tel21 upon the addition of 90 mM $K^+$ [123]. In that case, a sequential folding model involving the formation of one triplex intermediate preceded by a fast pre-equilibration step with the binding of a single $K^+$ has been proposed as the most plausible model (see Fig. 2.8). A similar model has been also proposed for RNA G4 and a non-telomeric G4 [123].

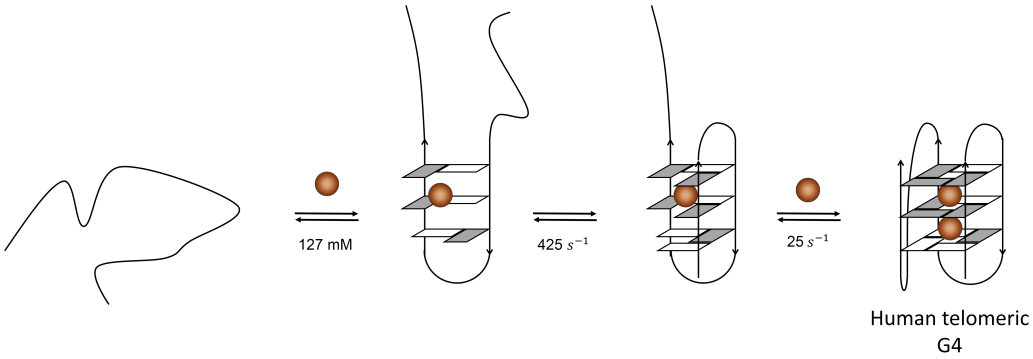


Figure 2.8: Proposed folding pathway for Tel22 in $K^+$ according to the group of Balasubramanian [123]. A random coil undergoes transition to a hairpin structure, followed by the formation of a triplex formation allowing the folding to the corresponding hybrid structure. $K^+$ is indicated by brown spheres.

Besides rapid-mixing technique, one recent study using an ultrafast microfluidic mixer has highlighted the existence of events on the microsecond time scale in the

folding process of the G4 sequence Tel24 [128]. Such fast events would be consistent with the idea of the formation of a hairpin structure as the first step of the G4 folding process.

On the other hand, the group of Schwalbe has performed slow-mixing experiments combined with NMR on the 2GKU sequence with $K^+$. This high-resolution method allowed them to discriminate the absolute structures of DNA involved in the folding mechanism. They have found that it takes several hours at 25°C for the G4 to reach its final hybrid1 conformation. Several folding pathways have been identified involving the formation of several misfolded conformations, such as the hybrid2 conformation and a partially unfolded conformation. These observations are consistent with a multipathway model of folding.

#### 2.4.2.2 Single molecule FRET studies

Besides rapid-mixing techniques, single-molecule approaches have been used to probe the folding mechanism of G4. In contrast to ensemble experiments, single-molecule FRET microscopy allows to probe metastable and heterogeneous states of one molecule at a time.

Single molecule measurements are typically implemented as depicted in Fig. 2.9. A single DNA strand containing a FRET marker (Dye1) and a G4 forming sequence is attached to a glass surface. A complementary strand with a second FRET marker (dye2) is then added in order to form a double helix as shown in Fig. 2.9. Formation of the G4 structure is induced by addition of cations. G4 formation results in an increase of the FRET efficiency.

The earliest studies of Balasubramanian carried out on Tel21 in the presence of 100 mM $Na^+$ or $K^+$ have shown the existence of two folded ensembles, they attributed to two distinct G4 structures in both cases: an antiparallel and a parallel topologies, respectively [129]. In contrast, the subsequent study of Lee *et al.* [130] of Tel21 with various concentrations of $K^+$ has highlighted the presence of an unfolded conformation in addition to the two folded conformations. Each folded topologies have been found to exhibit long (minutes) and short (seconds) lifetimes. To explain these results, they proposed a complex kinetics scheme involving two parallel folding pathways.

The studies of Okamoto *et al.* of Tel22 with 100 mM $K^+$ have revealed the presence of two major conformations, attributed to an unfolded one and a folded one, respectively. By modifying certain guanines via bromide substitution, they have proposed the existence of a triplex topology as a possible intermediate in the folding mechanism of Tel22 [132]. Subsequent studies of Tel22 and two other human telomeric

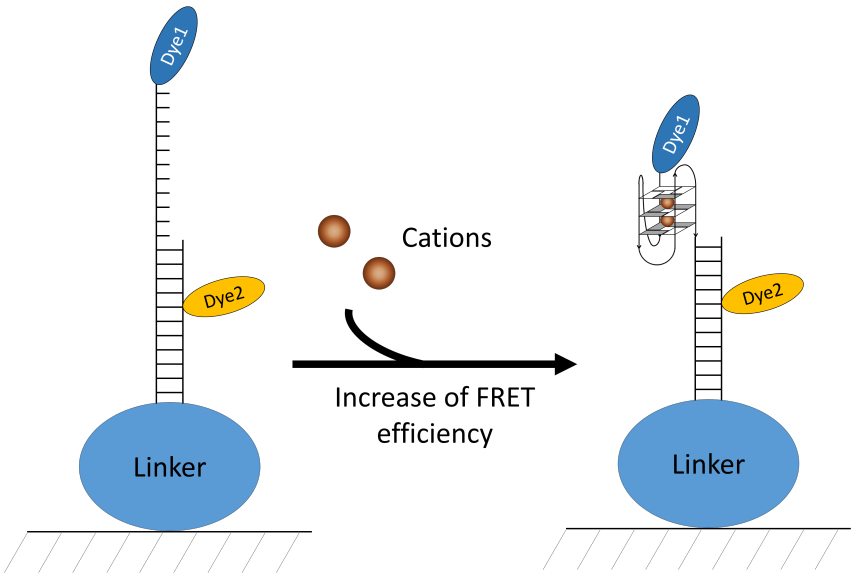


Figure 2.9: Schematic representation of a single molecule FRET experiment on G4. A DNA strand containing the G4 forming sequence at its end is attached to a linker. To assure that there is no interaction of the G4 forming sequence with the surface a complementary strand is added in a way that the G4 sequence is attached to a DNA double helix. One dye (Dye1) is connected to the 5' end of the G4 sequence whereas Dye2 is attached to the complementary strand. By adding cations G4 topologies form and the distance between the two dyes decreases which increases FRET efficiency.

sequences, Tel23 and 2GKU with 100 mM $Na^+$ or $K^+$ have clearly shown the impact of the protocol of G4 preparation on the observed conformations [133]. These studies of Long and Stones may explain the different distributions of conformations that have been observed in different experiments performed on similar sequences. The overall results of Long and Stones have been found to be consistent with a multi-pathway folding model.

Very recently, the group of Birkedal has published two studies of Tel22 in the presence of 100 mM $Na^+$, $Li^+$ and $K^+$ [131]. Despite a similar protocol of preparation than Long and Stone (i.e. a slow annealing process), they have identified four different conformations (one unfolded and three folded) of Tel22 in the presence of $Na^+$. Their populations were found to be strongly dependent on the concentration of cations. With $Li^+$, a sole conformation has been observed, attributed to an unfolded structure, as expected. With $K^+$, the existence of several conformations of Tel22 has been observed. Their thermodynamic and kinetic properties have been attributed to a multi-pathway folding model involving several marginally stable conformations (possible misfolded states) as folding intermediates (see Fig. 2.10).

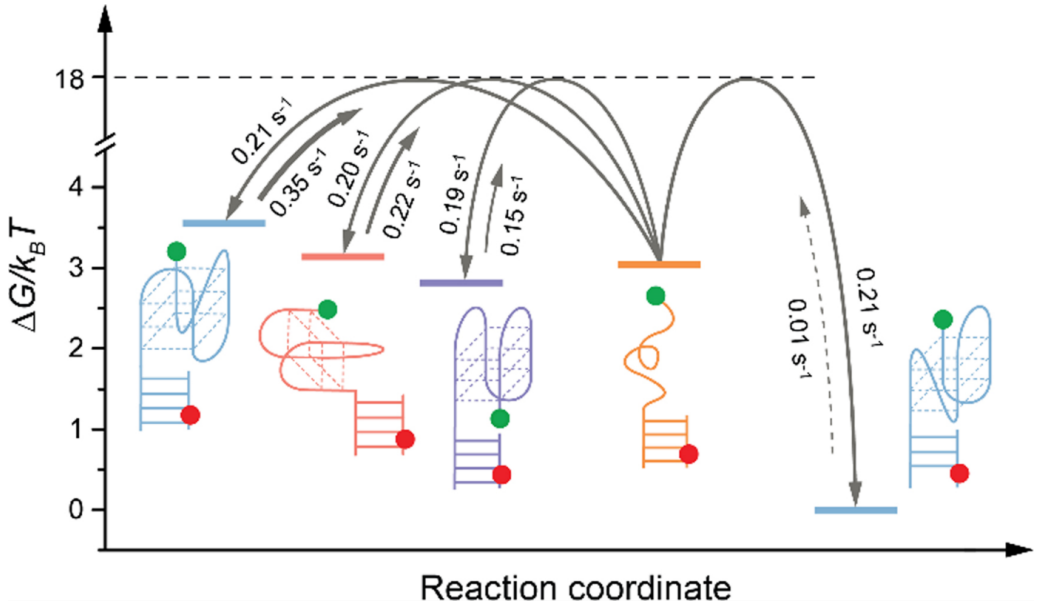


Figure 2.10: Figure taken from [131]. Different possible intermediates and folded structure (horizontal lines) and their stabilities and folding rates from unfolded species (dotted line).

#### 2.4.2.3 Single-molecule force spectroscopy

Additional single molecule studies have utilized optical and magnetic tweezers to study the folding mechanisms of G4 [134–137]. Fig. 2.11 depicts the principle of an optical tweezers experiment. Typically, a DNA strand contains beads on both ends that are trapped via lasers. The G4 forming sequence is placed in the center of the DNA strand. After G4 formation, one can elongate the distance between the two laser foci to apply a stretching force on the quadruplex. This results in G4 unfolding which can be measured as a function of extension for a given applied force [134].

Sugiyama and coll. have performed several studies on different human telomeric G4 sequences [134, 138]. Notably, they have compared the mechanical properties of Tel27, with that of the truncated human telomeric sequence, $TTA(GGGTTA)_3$ in the presence of 100 mM $Na^+$. They have observed a thermodynamically and mechanically stable species for the truncated sequence that they attributed to a triplex scaffold. Interestingly, the study of Tel27-mer has revealed the existence of two distinct conformations, one having an extension comparable to that of the putative triplex species. In subsequent works, Sugiyama and coll. have shown the formation of misfolded conformations as the number of TTAGGG repeats increases (4 vs 8). However, the number of misfolded conformations has been found to decrease if the length of the loops is increased.

In the meantime, Yu *et al.* have performed force spectroscopy on the human

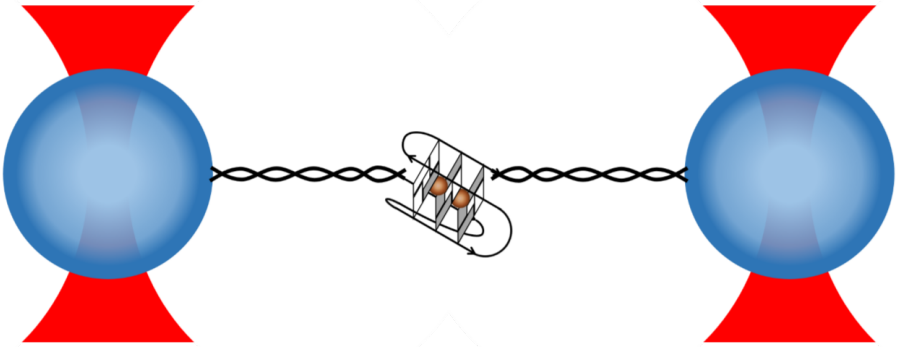

Figure 2.11: Scheme of an optical tweezers experiment according to [134]. Blue spheres reflect beads that are trapped in laser foci (red profile). DNA is attached to the beads depicted as black double helices. In the center of the DNA double strand a quadruplex forming sequence is employed. This is shown by a type-1 forming G4.

telomeric sequence Tel27 with 100 mM $K^+$ [135]. By substituting one thymine of the G4 loops by one uridine, they have built 6 different DNA constructs in which ds-DNA is attached at different locations of the G4 sequence. They found that the G4 unfolding trajectories of these 6 DNA constructs might follow a funneled energy landscape. In another study [139], Yu et al. have identified several conformations in the mechanical unfolding/refolding of a longer G4 sequence, the 41-mer h-TERT in the presence of 100 mM $K^+$. Kinetics and thermodynamics properties of these conformations have been attributed to a multi-pathway folding mechanism involving both sequential and cooperative pathways. No evidence of the existence of a triplex intermediate has been found in these experiments.

The use of magnetic tweezers is another possibility to study mechanically induced folding/unfolding of G4. Its principle is similar to that of optical tweezers. However, a slightly different experimental setup is needed. One end of the ds-DNA is attached to a surface, whereas the other end is linked to a magnetic bead [136, 137]. This magnetic bead is trapped by a magnetic field. Finely tuned extension of DNA is realized by slightly moving the magnets that trap the bead [136]. Li *et al.* performed magnetic tweezer studies on a repeat of the human telomeric sequence Tel21. To reduce polymorphism they have chosen to work with 100 mM $Na^+$ solution [136]. By finely adjusting the force exerted on DNA, they have observed several conformations in the mechanical unfolding/refolding of Tel21, they have attributed it to a sequential mechanism involving the formation of a triplex conformation.

Another sequence, Tel26, has been studied with magnetic tweezers by You *et al.* in the presence of 100 mM $K^+$. In these conditions, Tel26 may form a mixture of

hybrid1 and hybrid2 conformations in solution. Measurements by force spectroscopy have suggested a sequential folding mechanism involving the formation of a short-lived intermediate with a lifetime of 5 seconds and a subsequent long-lived intermediate with a lifetime of tens of seconds (comparable to [137]).

#### 2.4.2.4 Other experimental approaches

A thorough work has been published by the group of Gabelica [21]. They have used electrospray mass spectrometry in order to characterize the folding pathways of various G4 sequences, Tel21, Tel21-T, Tel22, 22CTA, Tel23, 2GKU and c-MYC2. Titration experiments showed that for low concentration of cations, the short sequences (Tel21, Tel22, Tel21-T and 22CTA) preferentially form an antiparallel G4 structure with 2 G-quartets and one $K^+$. For the longer sequences ($>$ 22 bases), $K^+$-binding cooperativity has been found to increase. Kinetics of $K^+$ binding of the 3 human telomeric sequences, Tel23, 2GKU and 26TTA, in addition to the sequence c-MYC2 have been found to be consistent with a multiple branching pathways as illustrated in Fig. 2.12. The antiparallel 2-G-quartet structure observed for low cation concentration has been suggested to be one misfolded intermediate of G4 folding (an off-pathway structure). Despite the fact that there is no direct evidence, the formation of a hairpin and a G-triplex structure often invoked as G4 intermediates of the folding/unfolding mechanism of G4 are not excluded [21].

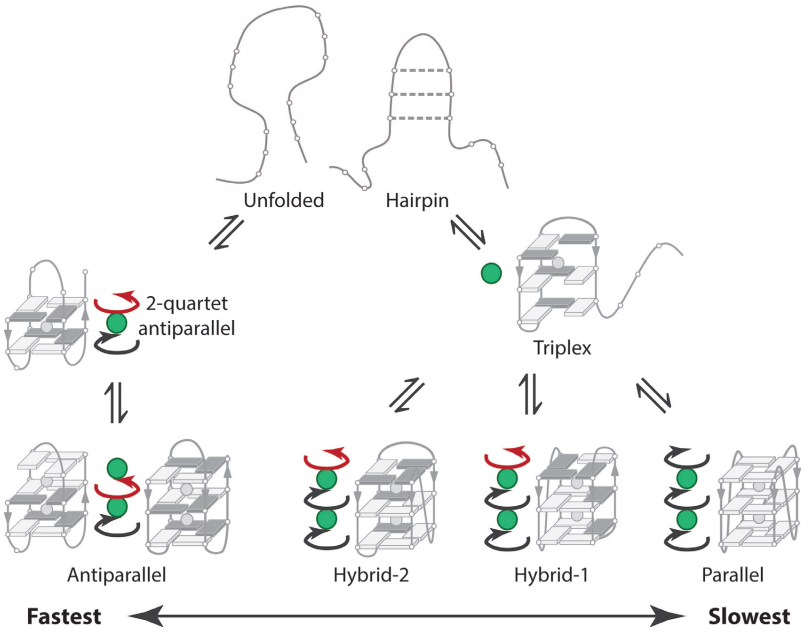


Figure 2.12: Scheme of possible folding pathways taken from [21].

In addition to mass spectrometry, temperature-jump (T-jump) experiments com-

bined with time-resolved absorption detection has been performed on Tel2, by the group of García. T-jump experiments measure conformational changes induced by a fast temperature rise, as will be explained in detail in the next chapter. Typical T-jump experiments explore the time-range of some nanoseconds up to 1 ms [27]. First T-jump measurments on Tel22 in presence of 150 mM $Na^+$ or $K^+$ have revealed kinetics in the time scale of a few ten microseconds. Despite the fact that these measurements have been done at low temperature (i.e. 22.5°C), far from the melting temperature, they have observed unexpected significant changes in the absorption of DNA after a T-jump of 2.5°C that has been attributed to the transition from the G4 structure to a triplex one.

#### 2.4.2.5 Theoretical Approaches

Many experimental approaches require the modification of the G4 forming sequence in order to monitor the conformational changes during folding/unfolding (FRET, 2-AP, optical/magnetic tweezers). These modifications are suspected to alter significantly the folding landscapes of G4. In fact, most of the experimental studies have revealed complex G4 folding mechanisms that involve several intermediates. While the nature of these species remains elusive, it comes out that their formation and their kinetics do not only depend on the primary sequence of DNA but also on the environment (cations, proximity of duplex DNA, ionic strength, ...). In this context, it is difficult to compare experimental results obtained in different conditions. In addition, most of the experimental approaches used up to now have a limited time-resolution (in general millisecond or seconds) and/or lead to a limited structural information. In order to get an atomistic picture of G4 folding, advanced molecular dynamics (MD) simulations have been recently developed, notably by the group of Šponer [140–142]. Like experimental approaches, MD also have to face several drawbacks correlated, for instance, to the time-scale of the simulations and the sampling methods. A review explaining the basic principles and limitation of MD methods has been published in 2017 [20]. In 2007, Mashimo and Sugiyama have proposed a folding mechanism involving the formation of a G-triplex intermediate from a G-hairpin conformation for Tel22 in the presence of $K^+$ [127]. With a combination of *ab initio* calculations and MD simulations, they have shown that the formation of a G-hairpin and then a G-triplex structure are energetically favorable intermediates in the formation of hybrid1 and hybrid2 topologies of Tel22 with $K^+$. *GG* base pairs have been found to have comparable energetic stability than AT base pairs but are less stable than *GC* base pairs. Their calculations have also shown that G-quartets are more stable than *GG* base pairs. The folding process has been found to be facilitated by $K^+$ binding

at each step of the process.

The group of Šponer has also performed a series of (MD) simulations on G4 folding and unfolding dynamics. Their works have suggested that the formation of antiparallel G-hairpin conformations may take place at all stages of the folding process of G4 [142]. The existence of various antiparallel structures has been taken as an indication of a multi-pathway nature of G4 folding. In contrast to antiparallel hairpins parallel ones have been found to be very unstable and may not form. It is noteworthy that a stable G-hairpin structure resulting in a complex topology that includes a chain reversal arrangment of the backbone in the continuous G-tract has been recently observed for the *5'-GTGTGGGTGTG-3'* sequence [143]. Other MD simulations of the Šponer group have shown that antiparallel G-triplex conformations may have a sufficient lifetime (i.e. a few tens of microseconds) to be a G4 folding intermediate. At this point, it is noteworthy that an antiparallel G-triplex conformation (TBA G3) has also been predicted from metadynamics simulations of the shortest G4 forming sequence, TBA G4, by Limongelli *et al.* [7]. Such a G-triplex structure has been then observed in solution by NMR spectroscopy from the 3'-end truncated TBA sequence [144].

In addition to putative G-triplex intermediates, recent simulations of Šponer and coll. [140, 145] as well as experimental approaches [2, 21] have pointed to the importance of misfolded structures in G4 folding, such as those with shifted strands with respect to the final G4 structure and a reduced number of tetrads. These structures, that have to unfold to form the final G4 structure, are expected to have a much longer lifetime and to be more stable than G-triplex structures. The existence of such misfolded structures indicates extremely rugged G4 folding energy landscapes. The natural kinetic partitioning of G4 energy landscapes has been attributed to the propensity of G-rich sequences to populate a large number of *syn-anti* patterns in their G-tracts [20].

### 2.4.3 Conclusion

Several experiments and simulations exploring the folding pathways of G4 have been performed. Most of these studies have focused on human telomeric G4 sequences. However, due to their high polymorphism resulting in part from their extreme sensitivity to the environment (solvent, cations, ...), it is difficult to compare the results of different experiments and simulations carried out under different conditions.

All studies of G4 folding have revealed complex mechanisms, but there is not yet a unified picture of these mechanisms. While ensemble experiments are generally in favor of a sequential folding mechanism, single molecule experiments and simulations

have often suggested a multi-pathway folding mechanism of G-rich sequences. However, most of all experimental approaches of G4 folding have required modifications of oligonucleotides that may alter both thermodynamics and kinetics properties of DNA. Notably FRET detection necessitates the adjunction of fluorescent probes at the end of oligonucleotides and single-molecule approaches require immobilization of DNA on a surface with the attachment of G4 forming sequences to long dublex DNA strands.

Several recurrent intermediates in G4 folding process have been proposed, such as G-hairpins and G-triplexes. However there is not yet direct experimental evidence of their formation. In addition, the slow folding kinetics spanning the time scale of a few milliseconds up to seconds or minutes suggest an extremely multi-pathway folding mechanism involving the formation of several misfolded G4 conformations [20, 22].

Regarding the fact that even slight modifications may alter kinetics [21], it is important to find experimental methods that allow to probe DNA conformation without any modifications of oligonucleotides. In this respect UV absorption and CD offers such conditions. CD is known to be very sensitive to the orientations of the guanines in the G4 core and therefore to the G4 topology [36]. This optical technique in combination with T-jump experiments provides the opportunity to probe the conformational changes of DNA with an ultimate time-resolution of a few picoseconds [146]. Such time-resolved melting studies of DNA allow to investigate the G4 folding landscapes starting from G4 folded state. In the frame of my work, by using time-resolved CD, I have carried out comparative studies of several G4 forming sequences under a uniform set of experimental conditions. These experiments are expected to provide new insights in G4 folding mechanisms. The underlying principles and techniques I used in the present work are presented in the following chapter.

# Chapter 3

# Techniques and methods

## Contents

## 3.1 Preface

Molecular structures and events are both too small and too fast to be seen and observed with our eyes. Light interaction with matter takes place, however, and allows us by carefully choosing a method to study desired effects and to increase our insight into this nanoworld of guanine quadruplexes. Chirooptical techniques have been developed in the second half of the 20th century [147] to determine the absolute configuration of chiral molecules [148]. In particular, CD is extremely sensitive to stereochemical variations and has thus become an important technique to study slight conformational changes and interactions [149].

This chapter focuses on circular dichroism (CD) and techniques, that we have used in this work. G4 folding (resp. unfolding) in which a secondary structure forms (resp. disappears) can be monitored by changes in CD. The first section treats circular dichroism, its definition and its origin in G4 structures. After giving the definition of CD, I will describe all the techniques that we have used to probe the conformational changes of G4.

## 3.2 Circular Dichroism

In this work, we have utilized CD that can be measured by the difference in absorption of left and right circularly polarized light [29]. Left and right circular polarization states are simple mirror images of each other and cannot be superimposed. Hence, difference in absorption yields information about the three dimensional molecular structure that cannot be superimposed with its mirror image. Such molecules or supramolecular structures are called chiral (from Greek $\chi\epsilon\iota\rho$ meaning hand). In contrast to absorption, CD is extremely sensitive to slight conformational changes in molecules [150].

Most biomolecules are chiral. This includes proteins, sugars and DNA. Not only double helix DNA is chiral, but also G4. Thus, they exhibit CD signals.

### 3.2.1 Definitions

There are different definitions for quantifying CD and thus different interconversions are necessary [29].

#### 3.2.1.1 Difference in absorption coefficients

CD can be defined by Euler's form of the Beer-Lambert law (Eq. 3.1). This is the one we have used for the time-resolved CD studies:

$$CD = (\alpha_L - \alpha_R)L = ln\left(\frac{I_R}{I_L}\right) \tag{3.1}$$

with $\alpha_{L,R}$ the absorption coefficient for left and right circular polarizations. $I_{L_R}$ intensity of the transmitted light for left and right circular polarization. By defining CD as the difference in absorption coefficients ($\alpha_{L,R}$) times pathlength ($L$), it depends linearly on the concentration, since $\alpha$ is directly proportional to sample concentration. Furthermore, that definition is unitless. CD signals are in general quite small ($10^{-2} - 10^{-4}$). Thus, taking only into account first order terms of a Taylor series of the natural logarithm $ln$, a good approximation of CD is obtained by simply taking the difference of transmitted intensity. Eq. 3.1 facilitates to

$$CD = (\alpha_L - \alpha_R)L = 2\frac{I_R - I_L}{I_R + I_L} \tag{3.2}$$

#### 3.2.1.2 Ellipticity

Due to historical reasons, another definition of CD can be found, denoted as ellipticity [29].

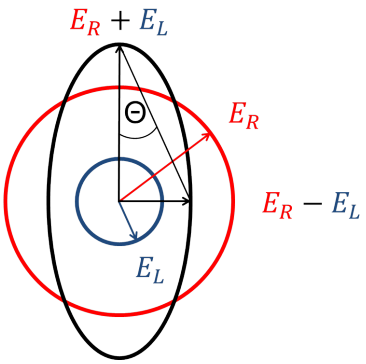


Figure 3.1: Electric field components of left and right circularly polarized light after transmission ($E_L$ and $E_R$, respectively). Black ellipse denotes constructed electric field out of both circular polarized electric field components. Angle $\theta$ describes angle between both perpendicular field components of the ellipse.

Static CD spectrometer with a photoelastic modulator (PEM) as modulator need the $ln$ approximation (see Eq. 3.2), since the Lock-in amplifier measures $I_R - I_L$. However, due to historical reasons CD spectrometer convert these CD signals from $(\alpha_L - \alpha_R)L$ into mdeg, the unit of ellipticity $\theta$. Different absorption of left and right circularly polarized light yields less intense light for one polarization state compared to the other. An elliptical polarization state is created out of the different electric field amplitudes. This is realized in a way that the major axis is defined as the sum of both electric fields for left and right circularly polarized light after the sample ($E_L$

and $E_R$, respectively). The minor axis is defined as $E_R - E_L$. As shown in Fig. 3.1, $\theta$ describes the angle between major and minor axis of the constructed polarization state. It is defined as follows

$$\tan\theta = \frac{E_R - E_L}{E_R + E_L} \tag{3.3}$$

Thus, if the sample is not chiral and no difference in absorption occurs ($E_R = E_L$), the reconstructed ellipse is a line and $\theta = 0°$. Equally, if one circular polarization state is completely absorbed, the ellipse becomes circular and $\theta = 45°$. Since effects of CD are small, $\theta$ is expressed in mdeg (0.001°).

Conversion between $\theta$ and CD is done according to Eq. 3.4.

$$\theta[mdeg] = 32980/2.3 \cdot (\alpha_L - \alpha_R)L \tag{3.4}$$

### 3.2.2 Artifacts

Since CD signals are small, they are prone to artifacts. Hence, this part discusses possible artifacts with respect to the present work.
The first case is when both polarization states are symmetric but not perfectly circular. With transmitted intensities of left and right elliptically polarized light, the measured CD signal writes as

$$CD^{measured} = 2\frac{I_R^{ell} - I_L^{ell}}{I_R^{ell} + I_L^{ell}} = 2\frac{I_R - I_L}{2I^{lin} + I_R + I_L} \tag{3.5}$$

where $CD^{measured}$ is the measured CD signal, $I^{ell}$ describes imperfectly circular polarization intensities ($I^{ell} = I^{lin} + I$). Herein, $I^{lin}$ is the linear contribution of elliptically polarized light. $I$ is the contribution of the circular polarization component. Due to symmetry assumption, $I^{lin}$ is equal for both, left and right, polarization states. Therefore, it only contributes to the denominator. This means that a symmetric deviation from circular polarization results in a smaller signal. Consequently, in that case, the amplitude of CD signals is not correct. Since this artifact may differ at each wavelength, it induces a distortion of the CD spectra. For time-resolved measurements at a single wavelength, it does not influence the measured dynamics.

A more complicated artifact occurs when both polarization states are unsymmetrical. If we assume $I_L$ to be perfectly polarized light, whereas $I_R^{ell}$ is elliptically polarized, Eq. 3.5 becomes

$$CD^{measured} = 2\frac{I_R^{ell} - I_L}{I_R^{ell} + I_L} = 2\frac{I_R^{lin} + I_R - I_L}{I_R^{lin} + I_R + I_L} \tag{3.6}$$

In that case, no clear statement can be made about the artifact's impact on measured CD. If $I_R^{lin}$ is large, it can even change the sign of CD. Furthermore, it is very difficult to correct it mathematically. Hence, if possible, asymmetric polarization states should be avoided. By looking at CD changes, i.e. relative CD signals, either for static melting curves or time-resolved dynamics, Eq. 3.6 writes

$$\Delta CD^{measured} = CD_{final} - CD_{initial} = 2\frac{I_R^{lin} + I_R^f - I_L^f}{I_R^{lin} + I_R^f + I_L^f} - 2\frac{I_R^{lin} + I_R^i - I_L^i}{I_R^{lin} + I_R^i + I_L^i} \tag{3.7}$$

where $I^f$ and $I^i$ describe final and initial transmitted intensity, respectively. Since $I_R + I_L \gg I_R - I_L$, the denominators for initial and final CD can be approximated to be roughly equal with respect to numerators (meaning $I_R^f + I_L^f \approx I_R^i + I_L^i$). Thus, Eq. 3.7 becomes

$$\Delta CD^{measured} = CD_{final} - CD_{initial} = 2\frac{(I_R^f - I_L^f) - (I_R^i - I_L^i)}{I_R^{lin} + I_R + I_L} \tag{3.8}$$

Looking at this approximation, it is clear that $I_R^{lin}$ disappears in the nominator for small CD changes ($I^{lin} \gg I_R - I_L$). Note that this approximation is only valid if $I_R^{lin}$ is sufficiently small.

The described possible artifacts originate from the light source. There are other possible artifacts. For instance, CD measurements can be influenced by linear birefringence (LB) occurring from the sample cell. Furthermore, in pump probe experiments this effect can be generated by the pump beam, if it changes the refractive index [151]. In solution, this change can be due to molecular orientation induced by the pump. It is however only significant at sub-nanosecond time-scales, since thermal rotation destroys pump-induced LB. In static CD measurements, LB occurring from sample cell and optics is easily removed by subtracting a baseline.

Linear dichroism (LD) is another possible artifact. It occurs when material transmittance depends on the incident orientation angle of linearly polarized light. It is compensated too by subtracting a baseline or measuring relative CD signals. Nevertheless, both artifacts can become problematic, if there is coupling between LB and LD. This was illustrated by the group of Kliger [152]. Its relative influence on CD is shown in Eq. 3.9.

$$\frac{Art_{LD,LB}}{CD} = \frac{2 \cdot LD \cdot LB \cdot \sin 2\chi}{CD} \tag{3.9}$$

where $\chi$ denotes the angle between the main axis of both artifacts. If they are parallel, no coupled artifact occurs. At 45° it reaches its maximum.

### 3.2.3 Theoretical Description

Here, a general mathematical description of CD is provided.

CD is connected to absorption. This means that if there is no absorption, there is no CD, either. In order to understand this, we will look at dipole strength $D$, which is the probability of an electronic transition [29]:

$$D = 6909\frac{hc}{8}\pi^3 N_0 \int \frac{\beta}{\lambda} d\lambda \tag{3.10}$$

where h and c are Planck's constant and the speed of light, respectively. $N_0$ is Avogadro's constant and $\lambda$ is the wavelength. A parameter $\beta$ is introduced to quantify the intensity of absorption. Since absorption bands are not mere lines and have a natural broadening, integration over a peak is necessary. In analogy to absorption, CD transitions are described by rotational strength $R$:

$$R = 6909\frac{hc}{32}\pi^3 N_0 \int \frac{\Delta\beta}{\lambda} d\lambda \tag{3.11}$$

where $\Delta\beta$ is the difference in absorption of left and right circularly polarized light. Eqs. 3.10 and 3.11 enable to extract both values ($D$ and $R$) out of measured spectra, meaning that they are merely empirical definitions [29]. Corresponding to these integrated intensities, the theoretical definitions for both $D$ and $R$ are as follows:

$$D = \vec{\mu}_{0i} \cdot \vec{\mu}_{0i} \tag{3.12}$$

$$R = Im\{\vec{\mu}_{0i} \cdot \vec{m}_{i0}\} \tag{3.13}$$

where $Im$ means that only the imaginary part of the dot product in the curly brackets has to be considered. $\vec{\mu}$ and $\vec{m}$ are the respective electric transition dipole moment and the magnetic transition dipole moment. The former term is defined as

$$\vec{\mu}_{0i} = (0|\mu|i) = e(0|\Sigma_j \vec{r}_j|i) \tag{3.14}$$

where $\mu$ is the dipole moment operator, $e$ is the electric charge and $\vec{r}_j$ are the electronic positions. The dipole strength ($D$) thus depends on the squared dipole moment ($|\vec{\mu_{0i}}|^2$). Rotational strength $R$ is more complicated. It depends not only on the electric dipole moment, but also on the magnetic dipole transition moment, which writes as follows:

$$\vec{m}_{a0} = (a|m|0) = (e/2mc)(a|\Sigma_j \vec{r}_j \times \vec{p}_j|0) = -i(e\hbar/2mc)(a|\Sigma_j \vec{r}_j \times \nabla_j|0) \tag{3.15}$$

where $\hbar$ is the reduced Planck's constant, $\vec{p}_j$ is the linear momentum of electron $j$. $\nabla_j$ describes the gradient operator for electron $j$.

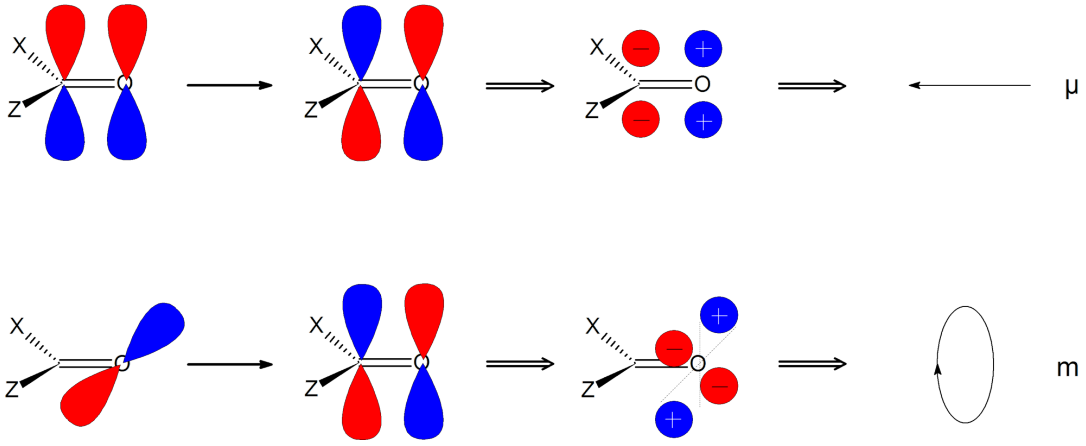


Figure 3.2: Transitions of carbonylic CO double bound. Top: Transition from bonding $\pi$ orbital to antibonding $\pi$ orbital (indicated via simple arrow). Orbitals are shown as linear combination of atomic orbitals. Red and blue indicate opposite signs of wavefunction. The product of both participating orbitals is shown after the first double arrow. The electric dipole moment is indicated via $\mu$ and an arrow pointing to the left. Bottom: Transition from non-bonding lone pair orbital to antibonding $\pi$ orbital. Its product is shown after the double arrrow. A rotational electric charge redistribution during excitation is indicated via circular arrow, meaning magnetic transient dipole moment $m$. X and Z are residual groups. This graph was redrawn according to [29].

For a better understanding the electronic dipole transition moment $\vec{\mu}$ can be considered as a linear charge displacement during excitation between ground and excited state. In analogy to that, the magnetic dipole transition moment can be considered as a rotational charge displacement during excitation [29]. This is similar to Maxwell's third equation. The picture one should have in mind is that a linear movement of electric charge does not provoke a magnetic field, whereas a rotatory movement of electric charge indeed provokes it. The origins of electric and magnetic transition moments are depicted in Fig. 3.2 for a CO transition. In the top part, a $\pi\pi^*$ transition from a $\pi$ orbital of a cabonylic double bond to an energetically higher $\pi^*$ orbital is shown. The bottom graph describes another transition from a lone pair orbital to the same excited $\pi^*$ orbital as in the top part. Note that the orientation in space of orbitals is due to its linear momentum $\vec{p}_j$. Hence, this figure illustrates why the $\pi\pi^*$ transition cannot have a CD signal, since there is no angular momentum during the transition. In the second case, there is a rotational charge redistribution combined with a linear motion of the charge: Most electron density of the non-binding orbital in the ground state is located around the oxygen atom, whereas

the excited state has an equal distribution between both carbon and oxygen atom. Consequently, a helical charge distribution takes place. Having neighboring residues with an optical activity, this transition does not absorb left and right polarized light equally, thus becoming chiral. An example is shown later on in this chapter with R-3-Methylcyclohexanone (see Sec. 3.6.2). CD can also occur from non-chiral transitions that bevome chiral by dipole-dipole coupling [29, 153]. This is the case for CD signals originating from the secondary structures of biomolecules including G4.

### 3.2.4 Dipole-Dipole Coupling

De Voe developed a method to predict absorption and CD spectra that result from dipole-dipole coupling with a complete classical approach [154, 155]. Therein, dipole moments are considered as a set of one-dimensional oscillators that are interacting with an applied electromagnetic field and each other. An oscillator describes an electronic transition. It is characterized by the complex tensor $\alpha(\nu)$

$$\alpha_i(\nu) = [\alpha'_i(\nu) - i\alpha"_i(\nu)]\vec{e}_i\vec{e}_i \tag{3.16}$$

where $e_i$ is the unit vector in the polarization direction of transition $i$. The imaginary part of Eq. 3.16 is proportional to the extinction coefficient of the absorption band:

$$\alpha"_i(\nu) = -(6909\lambda/8\pi^2 N_0)\epsilon_i(\nu) \tag{3.17}$$

In order to obtain the real part of Eq. 3.16, one can derive it by using the Kramers-Kronig relationship to the imaginary part:

$$\alpha"_i(\nu) = -(2/p)P\int_0^\infty \nu'\alpha"_i d\nu'/(\nu'^2 - \nu^2) \tag{3.18}$$

where $P$ indicates the Cauchy principal value, which means that the point at $\nu' = \nu$ is omitted. The electric moment induced in oscillation $i$ writes as

$$\vec{\mu}_i^{ind} = \alpha_i\vec{E}_i^{eff} \tag{3.19}$$

where $\vec{E}_i^{eff}$ is the effective electric field at oscillation $i$ and has two components:

$$\vec{E}_i^{eff} = \vec{E}_i^{ext} - \sum_{j\neq i} T_{ij}\cdot\vec{\mu}_j^{ind} \tag{3.20}$$

where $\vec{E}_i^{ext}$ is the externally applied field on the oscillation $i$. The second term describes the field at $i$ created by the other oscillators $j$. $T$ is a dipole interaction tensor and depends inversely on the cubic distance between two oscillators ($T_{ij} \propto$

$1/R_{ij}^3$). By inserting Eq. 3.20 into Eq. 3.19, a set of $N$ linear equations is obtained of the form

$$\vec{\mu}_j^{ind} = \alpha_j(\vec{E}_i^{ext} - \sum_{j \neq i} T_{ij} T_{ij} \cdot \vec{\nu}_j^{ind}) \tag{3.21}$$

In matrix representation, it is as follows:

$$G\vec{\mu}_j^{ind} = E^{ext} \tag{3.22}$$

where $G$ is a $3N \times 3N$ matrix. The induced dipole moments write as

$$\mu^{ind} = G^{-1}E^{ext} = AE^{ext} \tag{3.23}$$

Those induced dipole moments are equivalent to electric dipole transition moments in quantum mechanics. Their moments about the origin $r_i \times \mu_i^{ind}$ correspond to the magnetic transition dipole moments. Therefore, one can derive absorption and CD spectra due to dipole-dipole coupling as follows:

$$\epsilon(\nu) = -(8\pi^2 N_0/6909\lambda) \sum_i \sum_j Im(A_{ij}) \vec{e}_i \cdot \vec{e}_j \tag{3.24}$$

$$\Delta\epsilon(\nu) = -(16\pi^3 N_0/6909\lambda^2) \sum_i \sum_j Im(A_{ij}) R_{ij} \cdot \vec{e}_i \times \vec{e}_j \tag{3.25}$$

### 3.2.5 Origin of CD in G Quadruplexes

In the case of G4, some CD bands originate from guanine transition dipole moment coupling within a G4 quartet and between different quartets [153].

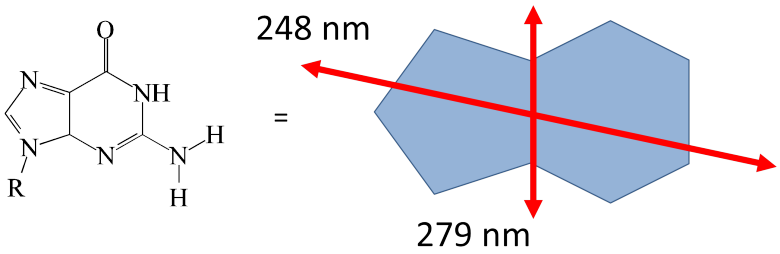


Figure 3.3: Schematic representation of a Guanine base with its two most relevant electronic transition moments (red arrows). The size of the arrows represent qualitatively their magnitude.

A guanine base has two distinct absorption bands in the region between 240 and 290 nm. This is due to two $\pi\pi^*$ transitions around 279 nm and 248 nm [156, 157]. They are schematically shown in Fig. 3.3. The transition with a smaller dipolar strength occurs at 279 nm and is roughly located in between the two rings. The

larger one at 249 nm is roughly perpendicular to it. In Fig. 3.4 orientations between transition dipole moments of different guanine bases are depicted. Coupling takes place between all guanine bases of one quartet and even between different quartets. Therefore, as we will see later, G4 with two quartets have smaller CD intensities than G4 with three or more quartets. Furthermore, the two guanine bases shown in Fig. 3.4 indicate that a coupling of these transition moments results in a rotational charge redistribution during excitation. Consequently, these two transitions exhibit a CD.

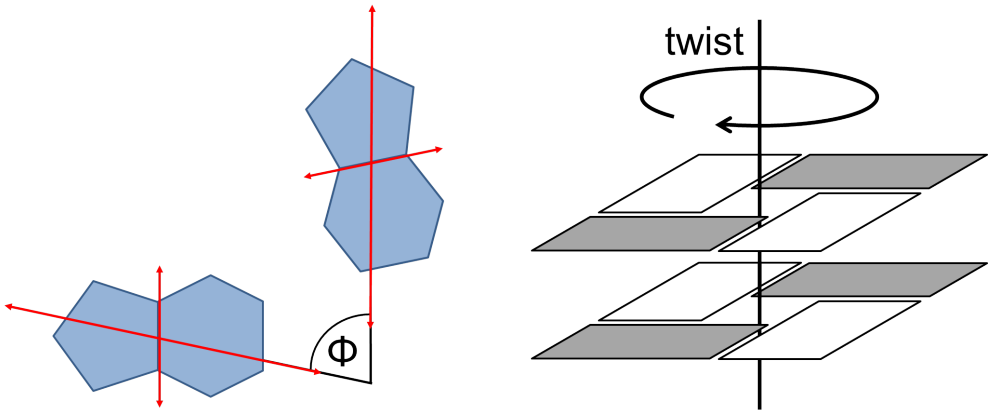


Figure 3.4: Left: Orientation between major transition dipole moments at 249 nm of two Guanines inside one G quartet, indicated by angle Φ. For transition dipole moments at 279 nm the principle is analogous and therefore not shown. Right: Relative twist of two G quartets. It influences coupling of transition dipole moments between quartets

Since guanine base orientations are crucial for transition dipole moment coupling, information about secondary structure can be extracted from CD spectra. In the case of an all-parallel G4, all guanine bases are in *anti* orientation. The first theoretical explanation was undertaken by Spada *et al.* [158] for an all-parallel G4. They successively reproduced two peaks of poly(G) at 240 nm and 260 nm by coupling of the aforementioned transition dipole moments. In that calculation only two quartets were considered without taking the sugar-phosphate backbone into account. A simplified model that describes head-to-tail orientations of two adjacent guanine transition dipole moments at around 249 nm splits it up in a positive peak at around 260 nm and a negative one at around 240 nm as observed for all-parallel G4. Consequently, in anti-parallel G4 head-to-head orientation occurs. Therein, an alternation between *syn* and *anti* shows up. This results in a positive peak at around 240 nm and a negative one at 260 nm for the same transition moment at 249 nm. Indeed, this is in agreement with experimental data and quantum mechanical calculations

[158, 159]. In addition to those two peaks, a third positive one appears at around 290 nm. Hypothesis supported by calculations states that this peaks results from head-to-head stacking of G quartets [153]. In general, it is highly likely that the orientation of stacked quartets is the determining factor for G4 CD spectra. The relative orientation of the backbone strands does probably not contribute significantly in the spectral region between 250 nm and 320 nm [150].

Besides all-parallel and antiparallel G4, mixtures of both exist: so-called hybrid-type G4 (see also Sec. 2.2). Those G4 can be easily distinguished from the other two by their CD spectra. Since they show features of the two other types, their CD spectra are also mixtures of the two other ones. In general, they show a negative peak at around 240 nm and a positive one at around 260 nm, equal to the all-parallel G4. Additionally, they have a positive peak around 290 nm as antiparallel G4.

## 3.3 Temperature jump (T-jump)

Stopped-flow methods and other mixing techniques have long been performed to study biomolecular dynamics [160] including G4 (see Sec. 2.4.2.1). Time-resolution for stopped-flow can go down to several milliseconds. An alternative to reach shorter time resolution is the use of fast sample heating by a laser pulse [27]. This can be done by optically exciting either a chromophore or water (IR pulse). The energy is then rapidly transformed into heat. The principle of the experiment is shown in Fig. 3.5.

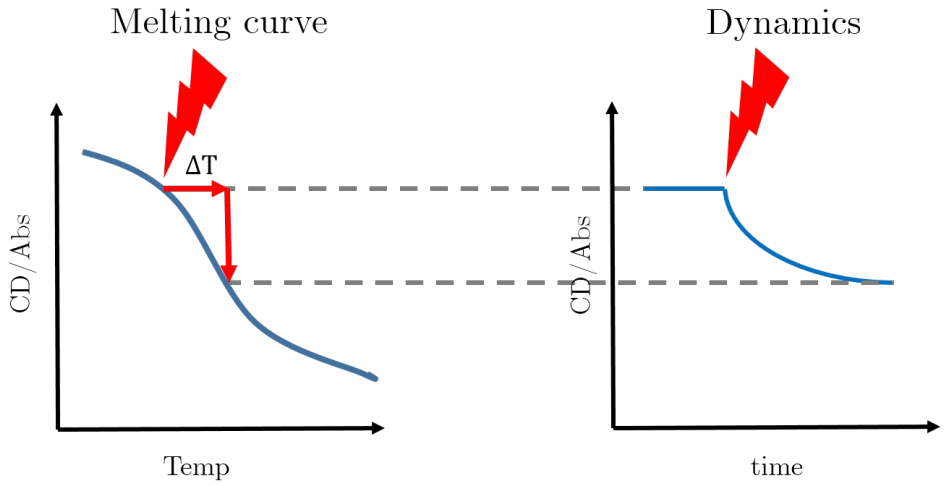


Figure 3.5: Principle of a T-jump experiment. Left: melting curve, CD or Absorbance (Abs) changes as a function of temperature. Red flash indicated heating laser (T-jump). Instantaneous temperature change (1st red arrow), followed by CD/Abs change (2nd red arrow). Right: T-jump induced CD/Abs change as a function of time.

On the left a scheme of a melting curve of a biomolecule is monitored by CD or absorbance (Abs). It reflects the population ratio of unfolded and folded population. At low temperature, normally the ratio folded/unfolded goes to infinity, at high temperature it is equal to 0. At the melting point $T_m$ this ratio is 1. Note that an equilibrium melting curve does not contain any dynamical information. In a T-jump experiment a laser pulse changes the temperature instantaneously by ultrafast heating (red arrow parallel to abscissa, $\Delta$T). The population is out of equilibrium and changes in order to reach equilibrium in the new conditions. This is shown in Fig. 3.5 right.

Heating a volume by a laser pulse results in a Gaussian-shaped heating profile. It is thus far from being homogeneous. Therefore, the probe laser has to be significantly

smaller than the pump (= heating) laser for probing a homogeneous heated volume. In x and y direction (perpendicular to laser propagation direction z), the heated profile corresponds to the beam profile in x and y direction when heating occurs. This profile broadens according to Fourier's law of heat transfer. In the z direction, there is a heat profile according the Beer-Lambert law of absorption (also directly after heating). This law is as follows:

$$I_L = I_0 e^{-\alpha L} \tag{3.26}$$

where $I_0$ is the initial light intensity, $I_L$ the transmitted light intensity after a pathlength $L$ through the medium. Absorption coefficient is denoted as $\alpha$.

The amount of light absorbed by a given pathlength interval $dL = L_{i+1} - L_i$ is proportional to the initial intensity $I_i$ (amount absorbed $= I_i - I_{i+1}$). $I_i$ or $I_L$ decreases according to Eq. 3.26. Since the temperature change of the solution for a given $dL$ is proportional to the absorption, Eq. 3.26 becomes as follows for temperature changes along the z direction:

$$\Delta T_L = \Delta T_0 \cdot e^{-\alpha L} \tag{3.27}$$

Because light intensity is maximal at the beginning, the T-jump at $L = 0$ is maximal: $T_0$. In Fig. 3.6 (left) the temperature profile of an average T-jump of 5K in a cuvette of 100 $\mu$m pathlength is depicted. The temperature change ranges from 7.9K at $L = 0$ to 2.9K at $L = 100\ \mu m$. Smaller pathlengths yield higher homogeneity, whereas longer pathlengths decrease it. A way to increase homogeneity is by redirecting the transmitted heating beam through the sample [27, 161].

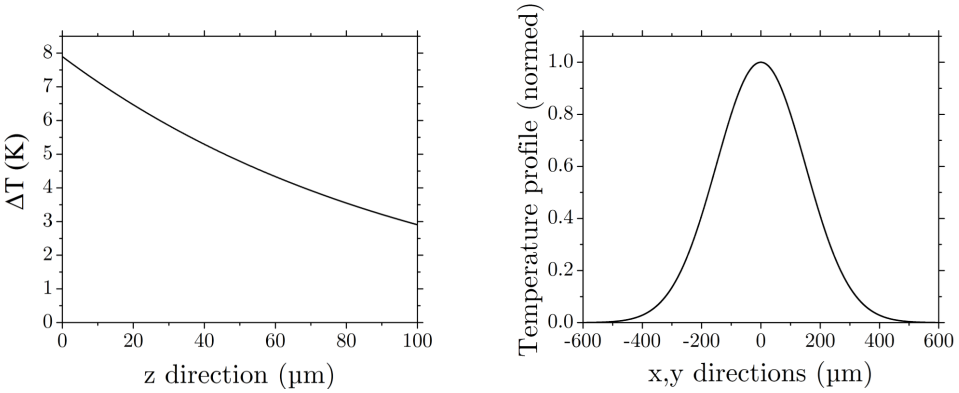


Figure 3.6: Temperature profiles directly after heating. Left: z-direction. Temperature distribution for an observed $\Delta$T of 5K in a cuvette of 100 $\mu$m. Right: distribution in x and y direction for a symmetric gaussian beam profile with FWHM = 300 $\mu$m.

In order to increase homogeneity of the heating volume, one can also try to flatten the peak of the transverse heating laser profile [27]. Another possibility could be to manipulate the beam in z direction. Strong focusing increases the photon density from $L = 0$ to $L = 100\ \mu m$. Thus, one should be able to partly compensate the intensity decrease through absorption.

### 3.3.1 Direct heating vs. indirect heating

Heating the solvent via a chromophore allows fast T-jumps within the picosecond time regime [27]. However, high dye concentrations are necessary and spectra of chromophore and the studied compound could overlap. If it is guaranteed that there is no interaction between the dye and the studied biomolecule, indirect heating can be the method of choice. Most of the heating chromophores contain large aromatic rings. Since G4 tetrads exhibit large planar $\pi$-conjugated surfaces [36] these chromophores are expected to interact with DNA.

A different way to heat the solvent is direct excitation of water molecules. This was used in the present work and performed on a $ns$ time scale with a pulsed Nd:YAG laser (Quanta-Ray INDI, Spectra-Physics) and a BBO-OPO ($\beta$-barium borate optical parametric oscillator, OPO-C-532, GWU-Lasertechnik). The second-harmonic of the Nd:YAG laser at 532 nm was converted to 1485 nm by the OPO which is close to the absorption peak at 1.51 $\mu$m of water [162]. This allowed us to obtain T-jumps between 4K and 7K.

### 3.3.2 Nanosecond T-jump

Time-resolution of T-jump experiments depends on the duration of the IR heating laser and the cooling time. Nölting showed that the heating laser energy is transferred to the environment within tens of picoseconds [163]. Therefore, the nanosecond laser we used limits the time-resolution simply by its duration (a couple of nanoseconds (see also [161]). The duration of heating is limited by thermal diffusion. Most of the heat energy departs through the glass windows of the cuvette, due to its large thermal diffusivity compared to water [27]. Quartz cuvettes are used instead of $CaF_2$ ones because of its smaller heat capacity [161]. It was previously shown that the temperature stays stable for around 1 ms. After 20 ms only half of the heat was maintained in the probed volume. Duration of heating depends also on the difference in full width of half maxima (FWHM) of probe and pump laser.

### 3.3.3 Millisecond T-jump

Nanosecond T-jump experiments allow to study a time frame from a couple of nanoseconds up to 1 ms. However, G4 folding may occur in the millisecond up to the second time-scale and therefore, an increase of the observation time window is necessary. If continuous heating is applied, an equilibrium between cooling and heating is expected to set in. However, simulations showed that there is a fast temperature change (around 200 ms), but the temperature still continues to increase (see Fig. 3.7). For the simulations a program of the group of Martin Volk was used that has been used for nanosecond T-jump simulations by L. Mendonca in our group [161]. It was modified to simulate continuous heating by adding heat before each time step in the program. Parameters chosen for this simulation are as follows: FWHM(probe) = 80 $\mu$m, FWHM(pump) = 700 $\mu$m, l(cuvette) = 100 $\mu$m, time step $\Delta$t = 0.05 $\mu$s, time window = 500 ms. The cell window is quartz in all our experiments. Therefore, the corresponding parameters were used for the simulations (themral conductivity: 1.4 $WK^{-1}m^{-1}$, volumetric heat capacity: $1.4 \cdot 10^6$ $Jm^{-3}K^{-1}$, thermal. These parameters are as the experimental ones (see Sec. 3.7). Different $\Delta T$ do not influence the dynamics of the simulations.

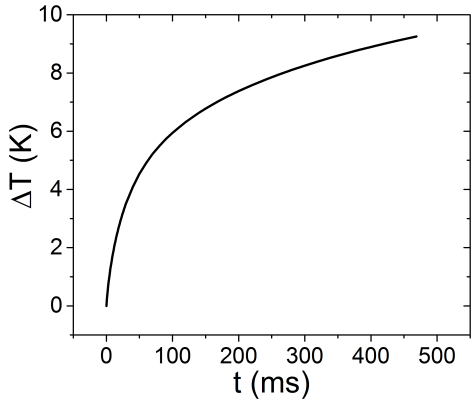


Figure 3.7: Simulation of temperature change measured by probe with a continuous wave heating laser for a $\Delta T = 10K$. Fast rising time is around 22 ms, slow rising time is around 230 ms.

However, the exact evolution of the transverse as well as the longitudinal temperature profile during heating is not known in that case. One can assume that the initial beam profile in z direction might be a bit more homogeneous than in the case of ns T-jump, since hotter parts cool down faster and heat transfer from hotter to colder regions should occur. The same behavior might also be valid for the x and y directions. An initial gaussian profile most probably broadens during

heating. Consequently, the heating area should be homogeneous in the configuration.

In analogy to heat energy flux diffusion, particle diffusion also takes place during the heating process. This is especially crucial for a continuous heating, since probing time scales can range into the time scale of particle diffusion.

Diffusion constants ($D_t$) of some G4 are listed in Table 3.1. These values were taken from [164]. There is no indication at which temperature these values were determined. However, it was most probably at room temperature ($20°C$ or $25°C$). This is important, since diffusion coefficients vary with temperature according to Stokes-Einstein equation (Eq. 3.28) for globular approximation of diffusing objects [165].

$$D = \frac{k_B T}{6\pi\eta r} \tag{3.28}$$

where $D$ signify the diffusion constant/coefficient, $k_B$ is the Boltzmann's constant and T is the temperature. $\eta$ denotes the solvent viscosity, $r$ stands for the radius of the diffusing object.

| Structure | ion | $D_t$ $(10^{-6} cm^2/s)$ |
|---|---|---|
| Tel22 | $Na^+$ | 1.32 |
| Tel22 | $K^+$ | 1.47 |
| 2GKU | $K^+$ | 1.40 |

Table 3.1

According to Table. 3.1 diffusion coefficients of the studied G4 are in the range of 100 to 200 $\mu m^2$ per second at room temperature. These values increase with decreasing solvent viscosity $\eta$ (comp. Eq. 3.28). It decreases by a factor 3 between 20 and 90°C [166]. Hence, G4 diffusion increases up to 300 or 600 $\mu m^2$ per second at 90°C.

During our experiments, the FWHM for the IR laser was measured to be 700 $\mu m$ and 1200 $\mu m$ in vertical and horizontal direction, respectively. Taking the smaller value, a heated surface of at least $3.8 \cdot 10^5$ $\mu m^2$ at FWHM occurs. Our probe pulse has a FWHM of 80 $\mu m$, i.e. a surface of ca. 5000 $\mu m^2$. Considering the diffusion coefficients of G4, it takes at least 8 s for the whole probed volume to be exchanged (600 $\mu m^2$ at 90 °C, probe surface : $5 \cdot 10^3$ $\mu m^2$). Since the probe beam is much smaller than the heating beam, we do not expect interference of particle diffusion with G4 dynamics at the time scale of our experiments ($t \leq 5$ $s$).

### 3.3.4 Probing Temperature Changes in situ

#### 3.3.4.1 Bromothymol Blue

pH indicators have been used for measuring the sample temperature in T-jump experiments [161]. Their absorption changes with the temperature in buffer solutions whose pH is temperature sensitive. For instance, in Tris buffer within the pH range between 7.5 and 9, the pH changes linearly with temperature. As pH indicator, we have used bromothymol blue [161]. Its absorbance strongly changes in the spectral range between 580 nm and 640 nm with the temperature, as illustrated in Fig. 3.8. Changes exhibit a linear dependence with temperature at 633 nm as follows:

$$\frac{\Delta Abs}{\Delta T} = -0.006\frac{1}{K} \tag{3.29}$$

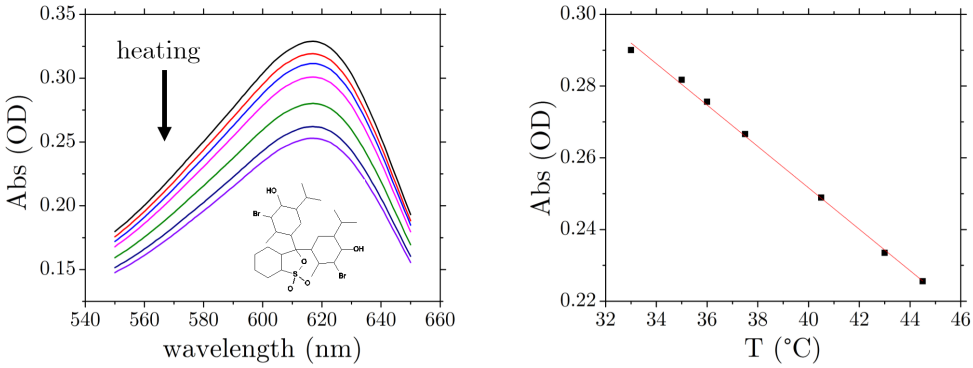


Figure 3.8: Left: Spectrum of bromothymol blue between 550 and 650 nm at different temperatures betweent 33°C to 45°C. Molecular structure is depicted in the downer right corner. Right: Absorbance at 633 nm, measured at different temperatures (black dots) in a 100 $\mu m$ optical path cell. Red line corresponds to the fit of the experimental data points with Eq. 3.29.

Heating and cooling had been thoroughly studied with bromothymol blue in ns T-jump experiments by Mendonca in our group [161]. In the present work, it has been used to verify a T-jump with a continuous wave laser diode. Fig. 3.9 shows a measurement with a CW laser diode for a final $\Delta T = 12$ K and an irradiation period of 1.8 s. Note, that they heating dynamics do not depend significantly on the applied $\Delta T$. Two exponential fit functions were necessary to fit the data accurately. A fast change of about 230 ms is followed by a slower one of ca. 1200 ms. Cooling process is similar: fast cooling (ca. 160 ms) followed by a slower one in the range of 1000 ms. In both cases the amplitude of the fast dynamics is slightly larger than that of the slow component. The first time constant is comparable to our simulations (200 ms vs.

230 ms, see Sec. 3.3.3). However, there is no such slow temperature changes predicted by the calculations. Therefore, one can think first, that heating (and cooling) is more complicated than the simulations suggest. Second, it is not excluded that dynamics arising from the indicator itself could contribute, too. From the results of Chapter 6 and 7, we will see that there are also biexponential denaturation/renaturation dynamics for G4. However, in the case of TBA G3 and Tel21 with 25 mM $Na^+$, those are almost completely monoexponential, which indicates that the observed slow dynamics of bromothymol is perhaps not related to the heating. Besides, dynamics of Tel21 have been observed with time constants of less than 100 ms (not shown). In order to answer this question sufficiently, a different temperature probe is needed. This will be part of a future work. For the sake of simplicity, the results of Chapter 6 and 7 are treated as if there was only a fast heating. We observed that the profiles obtained with bromothymol did not vary with the initial temperature, neither with $\Delta T$. Therefore, all observed tendencies in Chapter 6 and 7 are definitely real, only its absolute values could be influenced by a slow second heating.

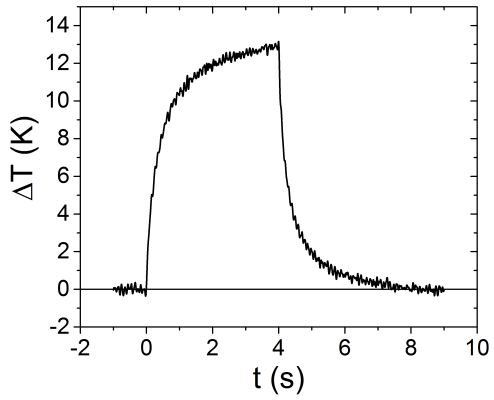


Figure 3.9: T-jump induced by a CW laser diode monitored via bromothymol blue in Tris buffer. Irradiation happened between 0 and 4 s. Two exponential fit functions have let to the following time constants: heating $t_1 = 230\ ms, t_2 = 1200\ ms$; cooling $t_1 = 160\ ms, t_2 = 1000\ ms$.

#### 3.3.4.2 Phenol Red

Bromothymol blue is convenient to detect temperature changes with a HeNe laser at 633 nm. However, our experiments have been performed at 293 nm and 265 nm. Consequently, we have searched for another indicator that changes its absorbance in that spectral range. Phenol red (see Fig. 3.10) meets this criterium, as shown in Fig. 3.10. Its absorption change with temperature at 293 nm follows

$$\frac{\Delta Abs}{\Delta T} = -0.0013\frac{1}{K}. \tag{3.30}$$

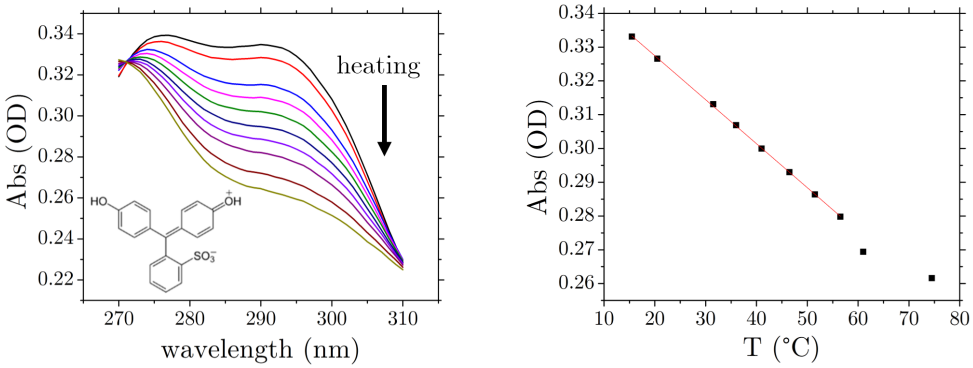


Figure 3.10: Left: Spectrum of phenol red between 270 and 310 nm at different temperatures betweent 10°C to 76°C. Molecular structure is depicted in the downer left corner. Right: Absorbance at 293 nm at different temperatures (black dots) and fit function between 10°C and 55°C in red. c(Tris)=100 mM, pH(20 °C)=8.3.

In the present examples, bromothymol blue and phenol red have a comparable absorbance at the wavelength of interest (632.8 nm and 293 nm, respectively). However, absorbance changes with temperature are smaller by a factor of around 4 for phenol red (compare Eq. 3.29 and Eq. 3.30).

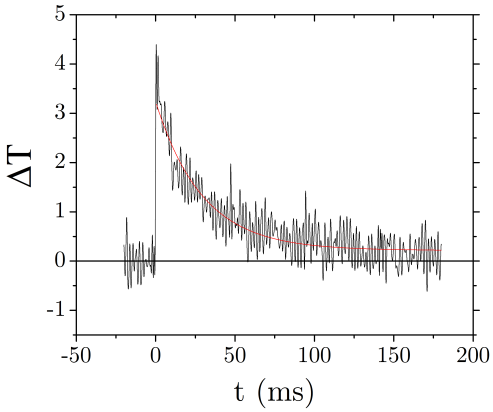


Figure 3.11: A rough $\Delta T$ of 4 K with a nanosecond T-jump monitored by phenol red at 293 nm. FWHM(probe) = 80 $\mu$m, FWHM(pump) = 300 $\mu$m. Time constant for cooling is around 30 ms. Red line indicates fit.

Note that the solution of phenol red has also been used to optimize the alignment of probe and pump beams, due to its strong absorbance changes at 293 nm. Fig. 3.11

shows the absorbance change of phenol red after a ns T-jump of 4K. We observe a cooling in ca. 30 ms.

#### 3.3.4.3 Refractive Index of Water

Temperature changes can, in principle, also be measured by the change in the refractive index of water as a function of temperature. According to Eq. 3.3.2 the refractive index of water at 633 nm exhibits the following temperature dependence [167].

$$n(T) = 1.333 - 4.5 \cdot 10^{-5} \cdot T/(^{\circ}C) - 1.2 \cdot 10^{-6} \cdot T^2/(^{\circ}C)^2 \tag{3.31}$$

where $n$ is the refractive index, T is temperature in °C. The change of refractive index with temperature results in a different speed of light. Eq. 3.32 shows the difference in the speed of light introduced by a T-jump ($\Delta c^{\Delta T}$):

$$\Delta c^{\Delta T} = c_{Tin+\Delta T} - c_{Tin} = c_0 \left( \frac{\Delta n}{n(n+\Delta n)} \right) \tag{3.32}$$

where $c_i$ are the velocity of light in vacuum $c_0$ and before and after T-jump: $c_{Tin}$ and $c_{Tin+\Delta T}$, respectively.

A small change of $\Delta n = n_{Tin+\Delta T} - n_{Tin}$ introduced by $\Delta T$ only affects light while passing through the cuvette (pathlength $L = 100\ \mu m$). Therefore, the resulting time delay $\Delta t$ is as follows:

$$\Delta t = \frac{L}{\Delta c^{\Delta T}} \tag{3.33}$$

Fig. 3.12 illustrates the scheme of the experimental setup for measuring refractive index change of water in a T-jump experiment. A continuous heating laser, triggered by a shutter is focused inside the sample (green beam). A Mach-Zehnder interferometer serves to measure changes in refractive index due to temperature changes. For that the HeNe laser probe beam is separated in two by a beam splitter ($BS1$). One beam passes through the sample, the second one through air. We measure their interferences behind the sample with a photomultiplier tube (PMT) connected to an oscilloscope (Osc).

The measured signal is the integrated square of the interference given by Eq. 3.34.

$$I = \int_0^{t_f} \Big( \sin(\omega t) + \sin(\omega(t+\Delta t) \Big)^2 dt \tag{3.34}$$

where $I$ is the measured signal, $\omega = 2\pi c/\lambda_{HeNe}$ and $\Delta t$ is the time shift introduced by a T-jump in the sample. $t_f$ is the time period of one oscillation ($t_f = 2\pi/\omega$). After integration, Eq. 3.34 becomes

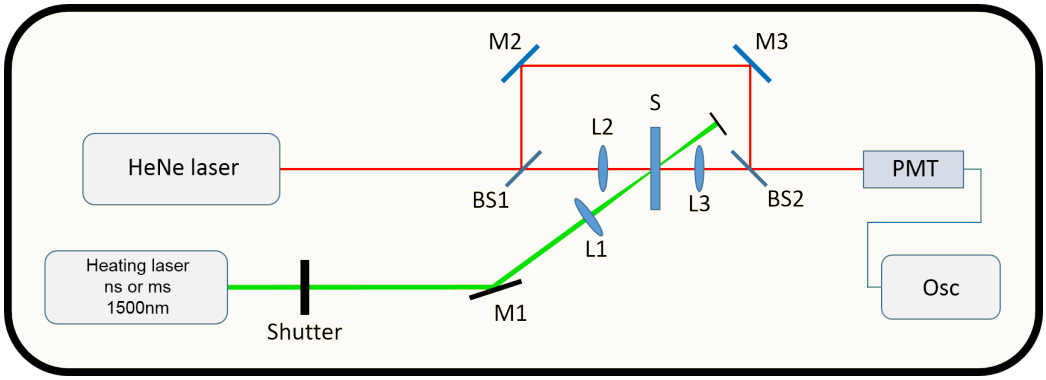


Figure 3.12: Principle of a Mach-Zehnder interferometer applied to measure changes of the refrective index of water induced by a laser heating source. $BS1$ and $BS2$ signify beamsplitters. Mirrors and lenses are denoted by $M$ and $L$, respectively. $S$ stands for sample. $PMT$ and $Osc$ indicates photomultiplier tube and oscilloscope, respectively.

$$I = 2t_f \cdot \cos^2\left(\frac{L(n(T)-1)}{4\pi\lambda_{HeNe}}\right) \tag{3.35}$$

which shows that the measured signal solely depends on the index of refraction and thus on the temperature. Note that the phase difference between reference beam and beam through cuvette before heating is set to be 0 in that equation. In an experiment, it is advised to have half of the maximum intensity to have a maximum signal change.

A $\Delta T$ of 1K at 60°C corresponds to a $\Delta n$ of around $2 \cdot 10^{-4}$. With a pathlength of $L = 100\mu m$ a $\Delta t$ of 3 ns is obtained.

Unfortunately, experimental realization has shown to be too challenging for us due to large fluctuations in the measured signals. A maximum shortening of the beam paths, protection again air circulation as well as the use of a damped laser table could not sufficiently reduce these fluctuations. These measurements may require a more stable laser source. This could be part of subsequent works.

## 3.4 CD Spectrometer with Temperature Automation

A home-built circular dichroism (CD) spectrometer has been constructed by François Hache's former PhD student Lucille Mendonca [161], allowing to measure CD spectra from 215 to 500 nm at room temperature.

Modifications and improvements have been achieved during my thesis. First of all, its acquisition time was decreased by around 30% via software modifications. Second, automated measurements of the CD spectra as a function of temperature has been implemented during my thesis (see Sec. SI:7).

This section describes the principle of our home-built CD spectrometer and its implementation for time-resolved CD measurements.

### 3.4.1 Setup, Principle and Implementation

Our home-built setup employs the same principle as commercial CD spectrometers [29]. Fig. 3.13 depicts the setup to scale. A deuterium lamp serves as white-light source (Lamp). The emitted light passes by a monochromator to select the desired wavelength. Then, two mirrors redirect the light before it passes a linear polarizer (LP). This is rotated by 45° with respect to the optical table plane. Consecutively, a photoelastic modulator (PEM) oriented at 90° with respect to the optical table plane is mounted. It modulates the polarization of the incoming beam alternatively between left and right circularly polarized light. The modulated light is sent to the sample in a temperature controlled sample holder (Qpod) before being detected by a photomultiplier tube ($PMT$). A lock-in amplifier extracts the CD signal with the aid of a PEM working at the frequency of 50 kHz. The sample holder ($Qpod$) is used to adjust the temperature.

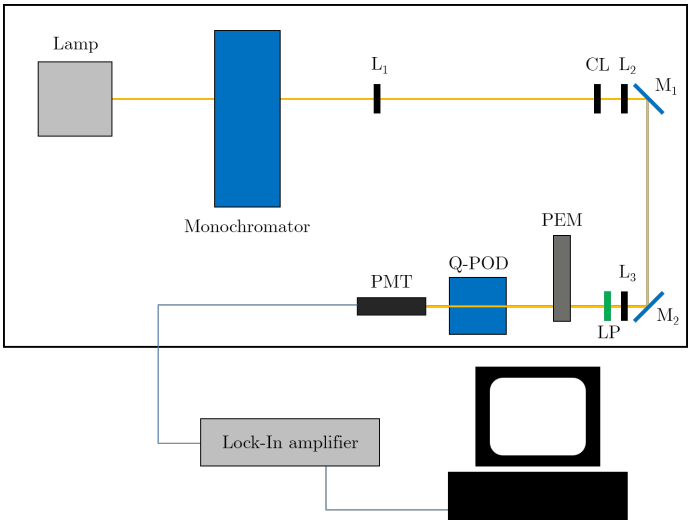


Figure 3.13: CD spectrometer to scale. Lenses are indicated as $L_i$, the cylindrical one as $CL$. Mirrors are depicted as $M_i$. PEM and PMT stand for photoelastic modulator and photomultiplier tube, respectively. Linear polarizer is abbreviated as $LP$.

Lenses are mounted to avoid strong divergence of the lamp light ($L_1$ and $L_2$). A cylindrical lens is added to reduce the elliptic beam profile induced by the monochromator slit. A precise control of the probe light diameter will be crucial for implementing the present for time-resolved CD measurements (see Sec. 3.17). Lens $L_3$ is mounted prior to the linear polarizer ($LP$) to ensure that there are no depolarizing optics between modulation ($LP$, $PEM$), sample ($Qpod$) and detector ($PMT$). A program written in labVIEW$^{TM}$ drives the apparatus. It controls the monochromator's shutter and grades, the amplitude of sonic waves of the PEM as well as the temperature of the Qpod and the voltage of the PMT.

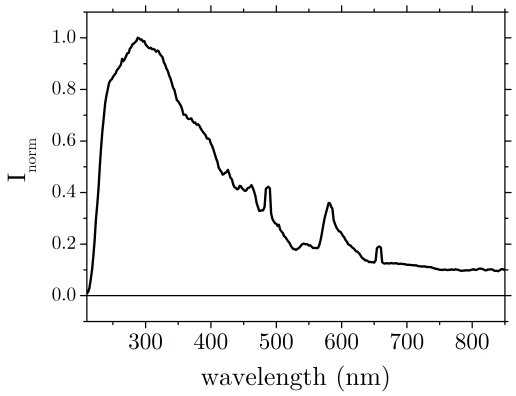


Figure 3.14: Normalized intensity spectrum of deuterium lamp between 205 and 850 nm.

A spectrum of our light source is depicted in Fig. 3.14. An intensity maximum occurs at around 300 nm. It decreases rapidly below 250 nm. This is probably due to absorbing optics.

### 3.4.2 Signal processing

CD is measured as the difference of absorbance of left and right circularly polarized light, as shown in Eq. 3.36.

$$Abs_L - Abs_R = -ln\Big(\frac{I^L}{I_0}\Big) + ln\Big(\frac{I^R}{I_0}\Big) = -ln\Big(\frac{I^L}{I^R}\Big) \tag{3.36}$$

where $Abs_{L(R)}$ is the absorbance for left (right) circularly polarized light. $I_0$ and $I$ are the transmitted and initial light intensity, respectively. A lock-in amplifier is used to detect the amplitude of the variation of the modulated light at 50 kHz. CD is calculated according to Eq: 3.37:

$$CD(absorbance) = Abs_L - Abs_R = \frac{2\sqrt{2} \cdot LI}{Signal(PMT)} \tag{3.37}$$

where $2\sqrt{2} \cdot LI$ is the lock-in signal. The lock-in assumes the signal to be rectangular, whereas it is sinusoidal. The factor of $2\sqrt{2}$ corrects this. It is normalized by the average transmitted intensity ($Signal(PMT)$). Eq. 3.37 is an approximation of Eq. 3.36 with $I^L - I^R \ll I_{R,L}$. Hence, the natural logarithm can be approximated to a first order series expansion $-ln\left(\frac{I_L}{I_R}\right) \approx I_R - I_L$, which is equal to $2\sqrt{2} \cdot LI$.

Due to the fact that sample absorbance varies with wavelength the PMT voltage is adjusted in order to ensure a constant gain(PMT) during measurements of CD spectra.

To obtain quantitative CD spectra the polarization states generated by the PEM have been characterized. For this, we added an analyzer oriented at 90° with respect to the polarization angle of $LP$, at the sample position. Ideally left and right circularly polarized light should exhibit the same amplitude for both circular polarizations. In this case, the amplitude of the transmitted signal should be half of the amplitude of the incident light. The asymmetry between both polarization states ($\Delta I$) can be measured at 50 kHz:

$$\Delta I = \frac{2|I_L - I_R|}{I_L + I_R} \tag{3.38}$$

where $I$ is the detected intensity for each respective polarization. The results are shown in Fig. 3.15. From 220 nm to 250 nm, $\Delta I$ increases. Between 260 nm and 350 nm, it decreases linearly with a slope of $1.7 \cdot 10^{-4}$ nm. Then above 350 nm, $\Delta I$ continues to decrease with a slightly different slope. At 630 nm, $\Delta I$ is found to be equal to zero, due to the supplier calibration at this wavelength.

$\Delta I$ provides a calibration curve for correcting the CD spectra from asymmetry in the circular polarization of the light. To verify the validity of our correction, we have measured the CD spectrum of one Tel21 solution with our home-built CD spectrometer at the beamline DISCO/SOLEIL. The comparison of the two spectra illustrated in Fig. 3.16 reveals differences in the relative intensity of the CD peaks. Nevertheless, their position is found to be correct. This provides evidence that our calibration method is not sufficient to allow a correct estimation of the relative intensity of the CD peaks. However, we can obtain a good qualitative measurement of CD spectra. Further investigations for an accurate PEM calibration are necessary for quantitative CD measurements.

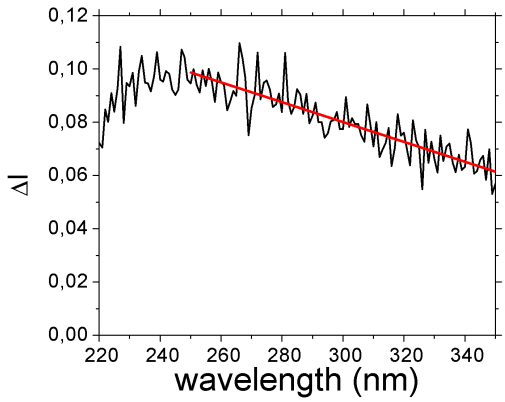


Figure 3.15: Lock-In signal at 50 kHz between wavelengths between 220 and 350 nm. It reflects the asymmetry ($\Delta I$) between left and right circular polarization.

### 3.4.3 CD spectrometer for T-jump measurements

The duration of measurements depends both on the integration time of the lock-in amplifier and the frequency of our PEM. Our PEM works at a frequency of 50 kHz, i.e. a cycle takes 50 $\mu s$. During 1 ms 50 cycles are integrated. Since the folding dynamics of G4 are of the order of ten to hundred ms, one can use our spectrometer to measure the time-resolved CD spectra of G4 with a time-resolution of 1 ms. This decrease of integration time results in more noisier CD signals. By averaging of several hundreds of curves, this additional noise can be compensated.

In order to implement this idea, an IR laser diode was added (*BrightSolutions*, BDL 5- 1470 F6) to our spectrometer as shown in Fig. 3.17. In that case, the IR laser beam is sent to the sample from the opposite side than the probe beam, due to a lack of space. A photodiode ($PD$) for detecting the IR light after the sample has been added and triggers the registration of the CD signals.

First of all, it must be ensured that our T-jump source heats an area of the sample larger than the probe diameter (see Sec. 3.3.3). This is more complicated with a lamp than with a laser source. Besides, its beam profile is quite stretched due to the monochromator slit. Therefore, a cylindrical lens ($f = 100\ mm$) is added to obtain a circular beam profile (see Fig. 3.13). In Fig. 3.18 horizontal and vertical beam profiles are depicted. Deuterium lamp beam could be focused to FWHM$_x$(probe) = 220 $\mu$m in the horizontal plane, which is significantly smaller compared to its vertical FWHM$_y$(probe) = 1.61 mm. The laser diode was focused to a FWHM$_{x,y}$(pump) = 3 mm to avoid interference of particle diffusion with time-resolved CD measurements on a time window of 5 s.

First measuremnts of Tel21 with $Na^+$ after a T-jump of 9°C from an initial tem-

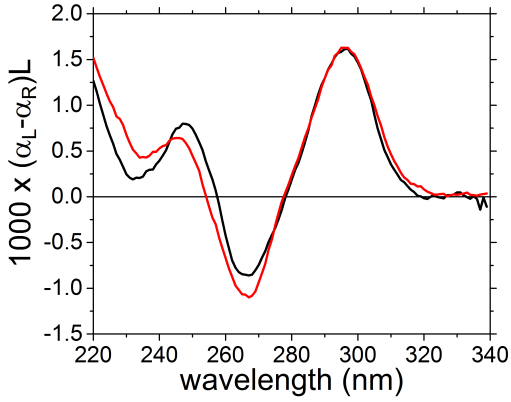


Figure 3.16: Comparison of theCD spectra of Tel21 measured at SOLEIL (black) with our spectrometer (red). The latter has been corrected by the addition of curve of Fig. 3.15 and normalized with the peak at 293 nm (multiplication of the red spectrum by 2.2).

perature of 57°C are depicted in Fig. 3.19 (above: absorbance, below CD). Denaturation was initiated with the IR CW laser diode at time 0 during 4 s. After turning off the laser diode renaturation is measured between 4 s and 9 s. Four peaks are observed in the transient absorption spectrum, two negative ones at 295 and 215 nm and two positive ones around 270 nm and 245 nm, respectively. Similar results are observed for TRCD spectra (Fig. 3.19, below). CD and absorption changes behave similarly and show the same dynamics from 320 nm to 215 nm. Observed time constants for denaturation and renaturation are found to be comparable to those obtained with the single-wavelength technique, which will be described in Sec. 3.7.

In conclusion, it is possible to use our static CD spectrometer to measure the unfolding/olding dynamics of DNA with a time-resolution of 1 ms. CD and absorption changes are recorded at the same time yielding a deeper insight into molecular dynamics. The biggest drawback of our setup comes from the acquisition time which takes more than 1 day. This is due to the measurement of th CD signal for each wavelength. On the other hand, a decrease of the probe diameter should be considered to improve both $\Delta T$ and the temperature homogeneity. Furthermore, pump and probe sources should enter the sample from the same side to increase the observed T-jump. However, our measurements of Tel21 showed clearly no peak shift upon denaturation and renaturation of DNA, which means that single-wavelength experiments are meaningful.

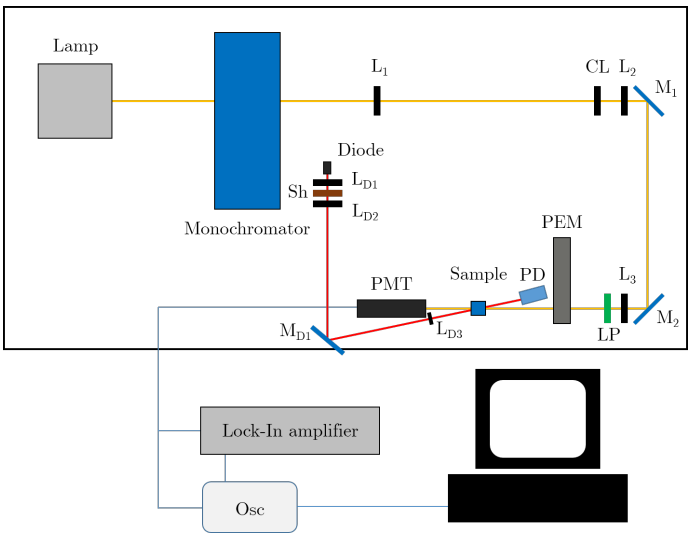


Figure 3.17: Scheme of CD spectrometer to scale. Supplementary optics to Fig. 3.13 are indicated as follows: Lenses are indicated as $L_{Di}$. The mirror is depicted as $M_{D1}$, the IR laser diode is indicated as *Diode*. A shutter and a photo diode are appreviated as *Sh* and *PD*, respectively. An oscilloscope *Osc* was added. The Q-POD was replaced by a sample cell.

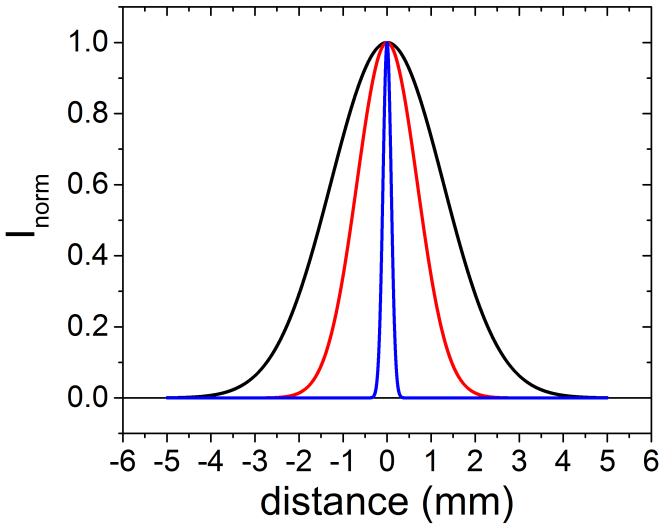


Figure 3.18: Reconstructed beam profiles from knife-edge scans in the focal point. Black: laser diode (heating laser) horizontal and vertical are the same (FWHM = 3 $mm$). Red and blue: vertical (FWHM$_x$ = 1.61 $mm$) and horizontal (FWHM$_y$ = 0.22 $mm$) lamp profile, respectively.

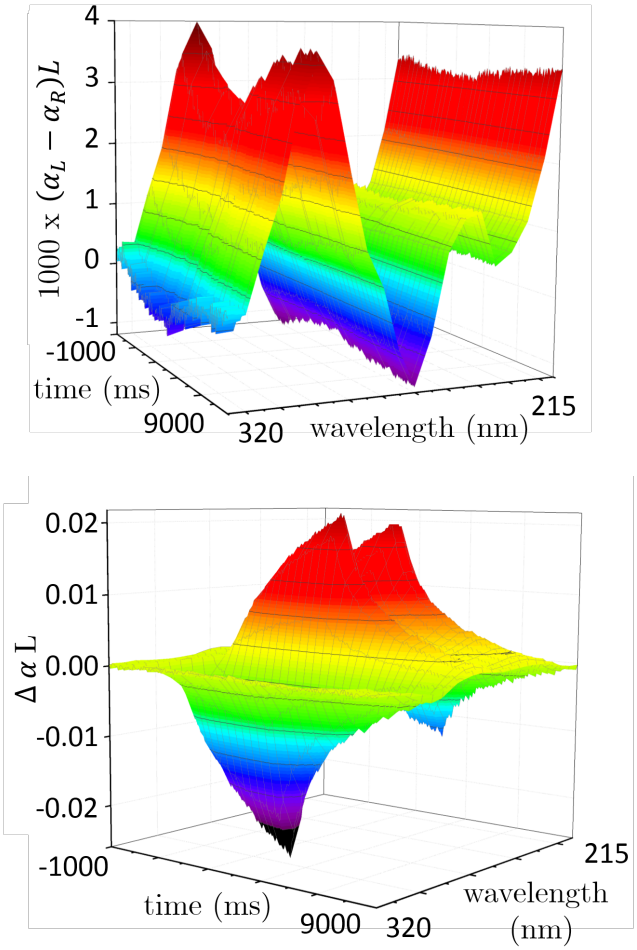


Figure 3.19: Relative absorption changes (above) and CD changes (below) of Tel21 in 150 mM $Na^+$ (c=0.45 mM) at 57°C between 320 and 215 nm. A T-jump of ca. 9°C heated arrived at time = 0. Renaturation of DNA starts after 4 s. Absorption changes are expressed as the change in the absorption coefficient times pathlength.

## 3.5 Combination of T-jump with TRCD using a Pockel's cell

Time-resolved CD measurements (TRCD) typically use the difference in left and right circularly polarized light to detect CD [168–171]. Pockels cells that are electronically adjustable wave plates [172] have been used in several TRCD experiments [169, 170].

Here, a combination of T-jump with a TRCD detection is illustrated in Fig. 3.20. The setup implemented by L. Mendonca during her thesis [161].

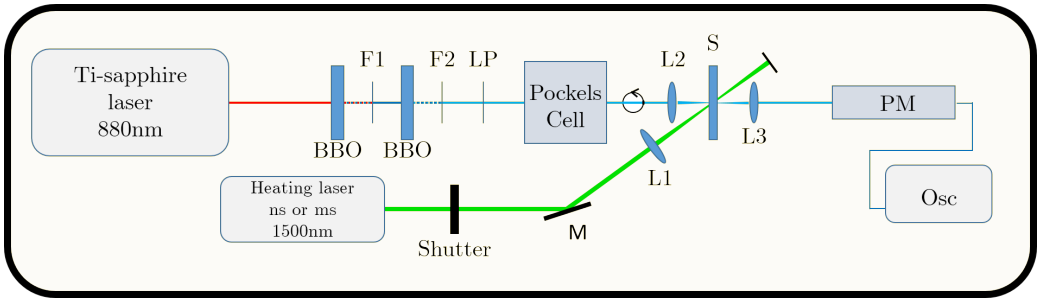


Figure 3.20: Scheme of our TRCD setup with a Pockels cell for modulating the circularly polarized probe light. BBO stands for $\beta$-Barium borate crystals. Filters are indicated as $F$, lenses as $L$, mirror as $M$, linear polarizer as $LP$ and the sample as $S$. Photomultiplier tube ($PMT$) is connected to an oscilloscope ($Osc$).

#### 3.5.0.1 Heating path

The heating source (green line) is focused inside the sample and blocked afterwards. Its reflection off a lens is sent to a photodiode and used a trigger for the oscilloscope (not shown). A ns 27 Hz Nd:YAG laser or a CW IR laser diode are used as heating pump sources. Switching between the two heating modes is realized by changing only one mirror (not shown) and $L1$. A shutter is used to adjust the exposure time with the CW laser diode or to reduce the repetition rate of the ns laser. The observed G4 unfolding dynamics are in the regime of a few hundred milliseconds. Therefore, refolding is not completed between two consecutive heating ns pulses. As shown in Fig. 3.20, pump and probe pulses enter the sample from the same side.

#### 3.5.0.2 Probe modulation via Pockels cell

The THG of a Ti-sapphire laser source is used as the probe, implemented in a non-colinear configuration. This allows to separate easily the THG from SHG and the fundamental with a pinhole. It is possible to use our Ti-sapphire laser between 795

and 900 nm. Hence, wavelengths between 265 and 300 nm can be produced in this configuration. Typically, experiments on G4 have been performed at 293 nm and 265 nm.

A linear polarizer ($LP$) in front of the Pockels cell insures the linear polarization of the probe light. Operation voltages at $\pm$500 V modulate light to left and right circular polarization, respectively. Response time of our Pockels cell is slow compared to the repetition rate of the ns heating laser. Therefore, a series of the kinetic trace is measured at one circular polarization state, followed by another series at the other polarization. This cycle was repeated several times to extract the CD signals [169]. Changes in CD in the order of $10^{-3}$ are detectable with this method [169]. This technique has however several drawbacks. Intensity fluctuations of the probe are not corrected and long acquisition times are necessary. Furthermore, left and right measurement are separated by several minutes. This limits the resolution of the measured CD signals. An effort to overcome these challenges is illustrated in the next section.

## 3.6 A new TRCD setup: one-pulse TRCD

A combination of T-jump with TRCD by modulating the probe via a Pockels cell requires highly concentrated samples to maximize the change in their CD induced by T-jump. For the proteins studied by Mendonca et al. [161] the use of high concentrations was possible. However, G4 tend to form dimers at concentrations above 1 mM (in strands). Hence, the measurement their small CD changes requires a more sensitive detection method. To this end, we have tested two different setups for extracting the TRCD signals from a unique probe pulse. This is expected to reduce the noise significantly. Laser intensity variations are compensated and also fluctuations inside the sample and optics.

### 3.6.1 One-pulse experiment

Linearly polarized light can be described as a combination of left and right circularly polarized light. Thus, if one can extract the absorption changes of both circular polarization states from a linearly polarized state, one obtains a CD signal. This principle is depicted in Fig. 3.21(a). A linearly polarized light (after polarizer P) passes through the sample. A quarter-wave plate ($\lambda/4$) at 45° transforms the light in a superposition of two linearly polarized states. Consecutively, a Wollaston prism (PS) separates both linearly polarized states. Each linear polarization state corresponds to one circular polarization state. Hence a direct comparison of both detectors should

reveal CD signals. To understand that, a definition of optical rotation and CD is introduced. Optical rotation is defined as $2\theta = 2\pi(n_L - n_R)L/\lambda$. CD is defined as $CD = 4\gamma = (\alpha_L - \alpha_R)L$, with $n_L,\ n_R,\ \alpha_L,\ \alpha_R$ being the refractive indices and absorption coefficents of left and right circularly polarized light, respectively. What happens in the one pulse setup is illustrated by Jones calculus [173].

$$E^{out} = M_{\lambda/4} \cdot M_{Sample} \cdot E^{in} \tag{3.39}$$

where $E^{out}$ and $E^{in}$ describe the electric field after and before entering the setup. $E^{in} = \begin{bmatrix} 1 & 0 \end{bmatrix}$ with (x,y) describing the laboratory frame for vertical and horizontal axis, respectively. Matrices $M_{\lambda/4}$ and $M_{Sample}$ signify matrix of quarter-wave plate and sample, respectively. An optical sample interacts with light as described in $M_{Sample}$.

$$M_{Sample} = e^{-\frac{\alpha L}{2}} \begin{bmatrix} 1 & i(\gamma + i\theta \\ -i(\gamma + i\theta) & 1 \end{bmatrix} \tag{3.40}$$

where $\alpha = (\alpha_L + \alpha_R)/2$ describes the mean absorption coefficient and L the pathlength through the sample. The effect of the quarter-wave plate is contained in $M_{\lambda/4}$

$$M_{\lambda/4} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & i \\ i & 1 \end{bmatrix} \tag{3.41}$$

Considering that effects of optical activity are quite small and thus $\gamma$ and $\theta$ are very small too, only lowest non-zero order terms are kept. Consequently, $E^{out}$ becomes

$$E^{out} = \frac{e^{-\frac{\alpha L}{2}}}{\sqrt{2}} \begin{bmatrix} 1 + (\gamma + i\theta) \\ i(1 - (\gamma + i\theta)) \end{bmatrix} \tag{3.42}$$

These two linear polarization are separated via a Wollaston prism. Regarding only first-order terms of Eq. 3.42, intensities show to be $I_{x,y} \propto 1 \pm 2\gamma$. According to Eq. 3.37 in Sec. 3.13, normalized CD writes as follows

$$CD = 2\frac{I_x - I_y}{I_x + I_y} = 4\gamma. \tag{3.43}$$

Thus, a simple difference in intensity of the two PMT should reveal CD signals and changes. However, this is only valid for an ideal configuration. Often, artifacts due to linear birefringence changes due to the heating laser arise in the sample cell. This is described in Eq. 3.44.

$$M_{LB} = \begin{bmatrix} 1 + ib\cos 2\delta & ib\sin 2\delta \\ ib\sin 2\delta & 1 - ib\cos 2\delta) \end{bmatrix} \tag{3.44}$$

where b is the retardation phase, $\delta$ is the orientation oft the fast axis. At first order our measured signal becomes $Signal = 4(\gamma - b \sin 2\delta)$, which shows that there is no easy way to separate CD from linear birefringence. And indeed, a large birefringence was observed when the sample was heated by our pump laser. Slight changes in alignment could even completely invert measured signals of the PMTs. In conclusion, a linear pulse before passing the sample did not allow CD measurements due to the strong influence of linear birefringence.

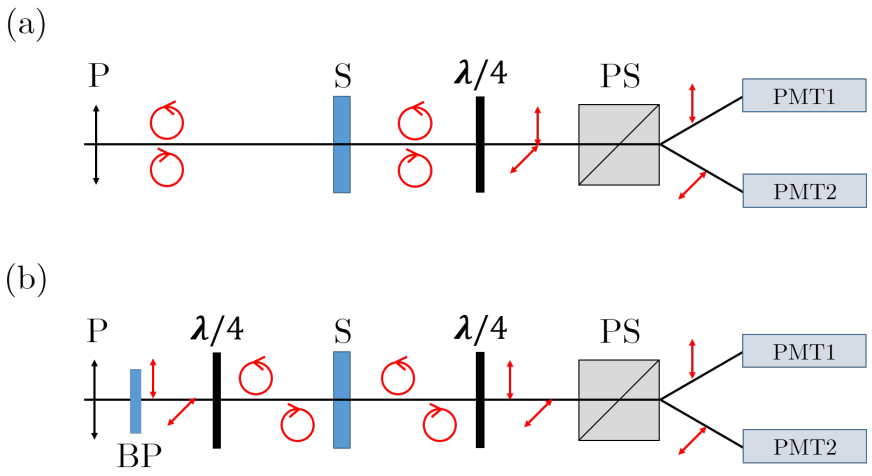


Figure 3.21: (a) Principle of CD measurement via one pulse. (b) Principle of two sub-pulses to measure CD. P denotes polarizer, S means sample, $\lambda/4$ denotes quarter-wave plate, PS: polarization splitter and BP is birefringent plate. Photomultiplies are abbreviated as PMT.

### 3.6.2 Two-pulse experiment

In order to overcome the linear birefringence, we came up with a more complicated setup that is illustrated in Fig. 3.21(b). Here, a birefringent BBO crystal (BP) was mounted after the polarizer with its fast axis at 45° with respect to light polarization. Consequently, the light component parallel to the fast axis travels faster through the material than the perpendicular component. Hence, if the material (BP) is thick enough, a separation in space of both perpendicular components takes places. A quarter-wave plate (QWP) was added in between BP and sample S with its optical axis parallel to the axis of P. Both perpendicular polarized pulses are oriented at ±45° with respect to QWP. Consequently, QWP transforms both linear in opposite circular polarization states. After passing through sample and consecutively being separated, both states are detected independently by two PMTs. The separation

of both circular polarization states after the sample was performed equally in both Fig. 3.21(a) and (b).

The main advantage of having circular polarization states before entering the sample reduces contribution of artifact. Electric input field (before sample, already circularly polarized) can be expressed as $E^{in}_{L,R} = E_0/\sqrt{2}\begin{bmatrix} 1 & \pm i \end{bmatrix}$, where +, (-) sign refers to L, (R) circular polarization. By applying Eq. 3.42, after the second QWP one obtains $E^{out}_L = iE_0 e^{-\frac{\alpha L}{2}}\begin{bmatrix} 0 & 1-(\gamma+i\theta) \end{bmatrix}$ and $E^{out}_R = iE_0 e^{-\frac{\alpha L}{2}}\begin{bmatrix} 1+(\gamma+i\theta) & 0 \end{bmatrix}$ for left and right circularly polarized light, respectively. This yields again $Signal = CD = 4\gamma$. Adding linear birefringence as in the previous section in Eq. 3.44, output fields become

$$E^{out}_L = iE_0 e^{-\frac{\alpha L}{2}}\begin{bmatrix} b\cos 2\delta + ib\sin 2\delta & 1-(\gamma+i\theta) \end{bmatrix} \tag{3.45}$$

and

$$E^{out}_L = iE_0 e^{-\frac{\alpha L}{2}}\begin{bmatrix} 1+(\gamma+i\theta) & -b\cos 2\delta + ib\sin 2\delta \end{bmatrix} \tag{3.46}$$

If we neglect second order terms, the difference in intensity of left and right circularly polarized light again yields $4\gamma$, resulting in a compensation of the first order term of linear birefringence. Hence, the measured signal is free of this artifact. This is valid as long as it is small enough so that the first order approach for Eq. 3.44 is sufficient.

Since this setup uses one pulse to create left and right circular polarization state, it is self-referencing. Thus not only dynamics should be measurable but also static CD spectra. For G quadruplexes it proved to be not sensitive enough for static CD. Consequently, we found R-3-methylcyclohexanone which has a strong CD signal at 290 nm $(\alpha_L - \alpha_R)L = 6 \cdot 10^{-2}$. Besides a test at 293 nm different wavelengths in the range of our Ti-sapphire were investigated as depicted in Fig. 3.22. It shows that indeed static CD spectra can be measured. However, for signals smaller than $10^{-2}$ it is not possible. Static state CD with two-pulse setup was performed as a proof of principle.

Unfortunately, many challenges have not been overcome so far. Probably the most crucial one is the polarization control. As a quarter-wave plate we used Babinet-Soleil compensators. They proved to not exhibit perfect extinction when mounted between cross polarizers. This suggests that some birefringence is added and also sufficiently circular polarization states were probably not obtained. A substitution with adjustable quarter-wave plates also did not reveal accurate polarization states, since they were sensitive to the incident linear polarization angle. Hence, two perpendicular polarized pulses did not interact equally with them.

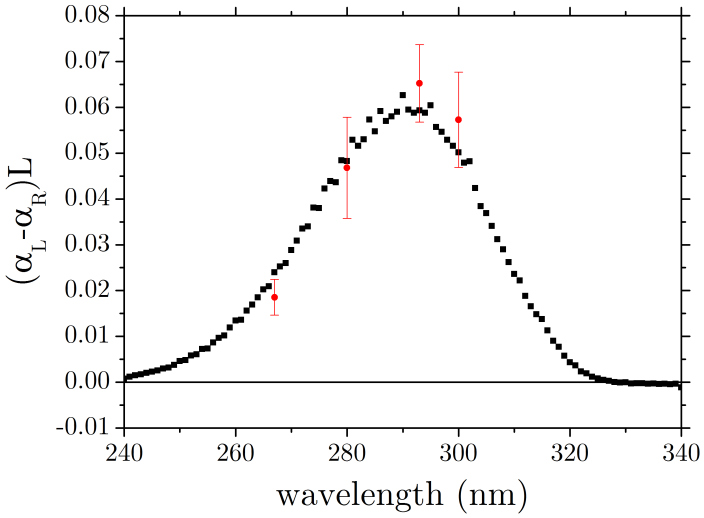


Figure 3.22: CD spectrum of R-3-methylcyclohexanone (black) measured with our CD spectrometer (values corrected according to Sec. 3.4.2). Red dots: Points measured with two-pulse setup.

During the trials we used only photomultipier tubes. First of all, it is difficult to ensure them to have the same gain. This ought not to be a problem, since as long as they behave linearly, relative signal changes are detectable. Being a strong signal-amplification device, their output signals were generally quite noisy. Available photodiodes were not sensitive enough at 293 nm to improve that.

### 3.6.3 Outlook

Despite many efforts, some challenges have been remaining. In this section, conceivable improvements and steps are discussed to overcome them.

Achromatic quarter-wave plates from Thorlabs were purchased that show promising polarization properties. Consecutive tests have to be performed to see its dependency on the incident angle between linear polarized light and the fast axis of quarter-wave plate. We resume the proper light manipulation as the most crucial limitation of the two-pulse setup.

There are sensitive photodiodes in the ultraviolet spectral region. In a consecutive step, these should replace photomultiplier tubes. Since, compared to PMT, photodiodes are low amplifying devices, such a replacement would reduce noise.

The oscilloscope we have been using (Tektronix TDS 3052B) has a good time

resolution (down to nanoseconds). However, its vertical axis contains only 65534 pixels. Therefore, one should choose the smallest possible scale to increase resolution. For a signal at 400 mV and a scale of 100 mV (±500 mV in total) changes of not smaller than 0.015 mV are detectable. This corresponds to a relative change of around $4 \cdot 10^{-5}$. Thus, subtraction of absolute values to obtain CD signals are limited by that value. In order to obtain higher resolution, signals should be read out by the computer directly. The oscilloscope has been used due to its conveniently easy time scaling. Once the method is improved in terms of resolution, changing the signal processing should be considered.

During a small post-doc period of two months after my thesis, I will focus on these points and try to overcome previous challenges. Furthermore, attempts of this technique for fs pump-probe experiments are ongoing. The aim is to see whether the actual incapacity is due to optics and instruments or whether it is perhaps due to intrinsic limitations as artifacts due to heating. The results of CD dynamics presented in Chapters 5 and 6 were measured exclusively with the technique described in the following section.

## 3.7 Combination of T-jump with TRCD by using PEM

The first two years of my thesis have been focused on the implementation of a TRCD setup to measure the folding dymanics of G4. For technical reasons, the two-pulse technique does not allow to measure the small CD signals of G4 and still need some improvements. Under the advice of John Sutherland, a recognized expert of CD spectroscopy using PEM, I undertook the implementation of a T-jump setup with a TRCD detection using a PEM. In 2000, Akiyama *et al.* already implemented a PEM to study stopped-flow mixing dynamics with a TRCD detection having a time-resolution of 400 $\mu s$ [174]. However, such a method was not applied in a pump-probe experiment up to now to the best of our knowledge. This idea has been proved to be able to detect CD changes of less than $10^{-5}$. Its final realization is depicted in Fig. 3.23. In comparison with our previous setup with the Pockels cell, only small changes in the probe beam path have been done. A long-focal lens ($L2$) in front of the PEM has been mounted to avoid the presence of optics between modulator, sample and detector that may effect the probe polarization. This setup is analogous to our static CD spectrometer and the same PEM and IR laser diode have been used (see Sec. 3.4). We found that a minimum integration time of 1 ms on the lock-in amplifier

and an average over 256 CD kinetic traces are necessary to reach a precision of $10^{-5}$ in the TRCD measurements. It is noteworthy, that under these conditions, our probe source which is a 82 MHz Ti:sapphire laser can be considered as a continuous light source.

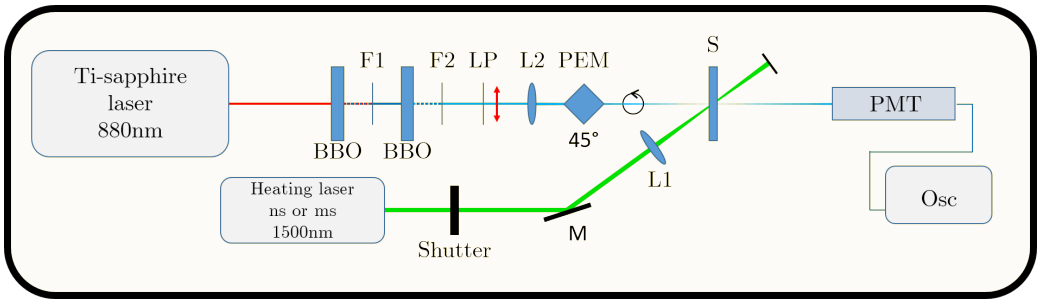


Figure 3.23: Scheme of TRCD setup with PEM to generate circularly polarized light. BBO stands for $\beta$-Barium borate crystals. Filters are indicated as $F$, lenses as $L$, mirror as $M$, linear polarizer as $LP$ and the sample as $S$. Photomultiplier tube ($PMT$) is connected to an oscilloscope ($Osc$).

The setup that is illustrated in Fig. 3.23 contains many advantages. First of all, it uses only one detector and equally one beam path for the probe pulses. The use of a lock-in amplifier significantly reduces noise of the CD signal. Furthermore, the lock-in signal is transferred to the oscilloscope on a separate channel than the one for absorbance. Since no comparison of two detectors is necessary to extract CD signal, the pixel limitation of our oscilloscope is negligible.

The parameters of the setup used in the experiments shown in Chapter 6 and 7 are as follows: FWHM(probe) = 80 $\mu$m for both, nanosecond and millisecond experiments. FWHM(pump,ns) = 300 $mu$m, FWHM$_{vert}$(pump,ms) = 700 $\mu$m, FWHM$_{hor}$(pump,ms) = 1200 $\mu$m. For all experiments, quartz cells with l = 100 $\mu$m have been used.

#### 3.7.0.1 Possible Artefact

We have observed CD dynamics of G4 to be identical with absorption dynamics. However, in a few acquisitions, we have observed a different behavior between CD and absorption as illustrated in Fig. 3.24.

The curves depicted in Fig. 3.24, are clearly not a real signal but arise from an artifact whose exact origin has not been completely understood yet. However, we have been able to avoid it. The artifact's amplitude has been found to increase with increasing signal change. By reducing the gain voltage of the PMT, we have found that this artifact can be eliminated. Another way to avoid it is to exchange

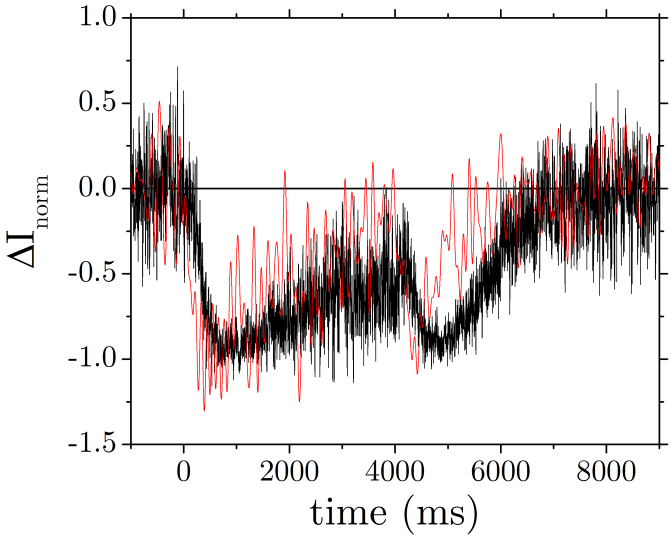


Figure 3.24: Normalized difference between absorption and CD dynamics (black) due to an artifact (red). This CD artifact was measured for c-MYC at 293 nm at which only absorption changes take place and no CD changes. Irradiation of CW pump laser took place during the first 4 s followed by a non-irradiation between 4 s and 9 s.

the impedance that transforms the PMT's current in voltage. An empirical value of 4.7 kΩ has been found to be the highest applicable one. Alternatively, a simple reduction of the probe intensity has been found to eliminate this artifact. However, this is not recommended, since it decreases signal-to-noise ratio.

Equipped with knowledge about techniques, the next chapters are dedicated to the studies of G4.

# Chapter 4

# Steady-state CD spectra

### 4.0.1 Preface

CD spectroscopy has been proven to be very sensitive to the guanine arrangement in G4 structure. CD has been widely used to distinguish between the three G4 topologies, parallel, anti-parallel and hydrid. Both empirical models and molecular modeling have shown that in the spectral region between 240 - 300 nm CD signals arise mainly from the excitonic coupling between the two lowest $\pi\pi^*$ transitions of guanines [175–177]. There is only a poor understanding of the origin of the CD at lower wavelengths in the VUV spectral region. Two previous studies of synchrotron radiation circular dichroism (SRCD) have however suggested that the CD signals in this spectral region may stem from intrastrand interaction of the bases and do not reflect the G4 topology [178, 179]. Furthermore, in that region contribution of the chiral sugar-phosphate backbone cannot be ignored. To obtain additional structural information on the G4 structures, we have performed a series of SRCD measurements at DISCO/SOLEIL on several G4 forming sequences. SRCD measurements in comparison with CD measurements with conventional spectropolarimeter allow to access to VUV spectral region down to 180 nm.

## 4.1 Sample preparation

This section focuses on the sample preparation and the selected G4 forming sequences studied for SRCD measurements and time-resolved CD experiments. G4 forming sequences have been purchased from Integrated DNA Technologies (IDT), on 5 to 10 $\mu molar$ scale with standard desalting. Sample preparation has been accomplished by dissolving DNA in 10 mM phosphate buffer (pH 7) with the appropriate cation ($Na^+$ or $K^+$). $pK_a$ of phosphate buffers are known to be weakly temperature depen-

dent ($\Delta pH = +0.015$ per K), which allow to perform T-jump measurements without a significant change of the sample pH. DNA concentration has been adjusted to be $\leq 1$ mM (in strands), in order to avoid the formation of higher order G4 structures. Solutions have been heated in a dry bath (ThermoStatTM C, Eppendorf) to 95°C during 5 min and then cooled down to 4°C at a rate of 1 K/min. In order to verify the formation of the G4 structure, UV melting curves have been recorded at 295 nm and 260 nm with a CARY 100 Biomelt spectrometer (Agilent) with a temperature gradient of 1 K/min. All the absorption spectra have been measured in a 1 cm optical path quartz cell. G4 structures can also self-associate to form higher order structures (dimers or multimers) that may hinder measurements of the folding process of G4 monomers. The absence of G4 multimers has been verified by UV melting experiments prior time-resolved measurements and SRCD measurments. In this respect, melting curves have been recorded both for increasing and decreasing temperature gradients. Since multimer formation is very slow, heating and cooling melting curves do not superimpose in the presence of G4 multimers. CD spectra of G4 as a function of temperature have been recorded with a home-made spectropolarimeter and at SOLEIL synchrotron (DISCO line) in collaboration with Frank Wien. SRCD measurements have been performed in 100 $\mu m$ optical path $CaF_2$ cells. For these measurements, fluoride salts instead of chloride salts have been used to adjust the cation concentration. In contrast to chloride salts, fluoride salts have a very low absorbance in the far UV range, which allows SRCD measurements down to 180 nm. Prior to these measurements, we have performed several UV melting measurements providing evidence that $F^-$ has a negligible impact on the G4 stability with respect to $Cl^-$. Temperature dependent SRCD spectra presented here have been normalized to a G4 concentration of 1 mM (in strands) and 100 $\mu m$ path length. The concentrations used for SRCD measurements are given in SI:Table 7.1.

## 4.2 Studied G4 forming sequences

We have previously seen that G4 folding are complex processes. This may even be exacerbated by the inherent polymorphism of guanine-rich DNA sequences. In order to get insight on G4 folding mechanisms, we have chosen to work with short and low-polymorphic G4 forming sequences, with well-characterized structures in solution, as far as possible. The chosen sequences presented below represent the three G4 topologies: antiparallel, hybrid and parallel.

### 4.2.1 Human telomeric sequences

According to the importance of human telomeric G4, static and time-resolved CD measurement have been performed on two of these sequences, Tel21 and Tel22. In order to reduce their polymorphism, their folding kinetics have been measured in the presence of $Na^+$. Tel21 is the shortest human telomeric sequence thought to form mainly an intramolecular antiparallel G4 topology with $Na^+$. Since recent study of Tel21 by NMR has suggested that that this sequence may adopt several conformations in solution in the presence of $Na^+$ [45], we have performed a comparative study of Tel21 with Tel22. Tel22 is known from NMR studies to adopt one main antiparallel basket conformation with $Na^+$ [1].

Since the investigation of native human telomere sequences has proven difficult in the presence of $K^+$ due to their high polymorphism, we have studied by SRCD spectroscopy two modified model sequences (2GKU and 22CTA). These two sequences are known to adopt single well-defined hybrid1 and antiparallel conformations, respectively, in solution with $K^+$.

### 4.2.2 Thrombin-binding aptamer sequence, TBA

In order to limit the conformational space of G4 folding, my work has also focused on the thrombin aptamer sequence, TBA G4. This sequence that inhibits thrombin-catalyzed fibrin clot formation *in vivo* [180] forms a G4 in solution [6]. In contrast to human telomeric G4 it only contains two G-quartets. TBA G4 adopts an antiparallel chair-topology G4 in the presence of $K^+$ [181]. TBA G4 is the shortest intramolecular G4 structure and thus has been often used as a model for simulations. In addition to that, I have studied truncated TBA sequence, TBA G3 is known to form a stable G-triplex in solution in the presence of $K^+$ [7]. Since the formation of G-triplex has been often proposed to be one intermediate in the folding pathway of G4 [67, 182], it is particularly interesting to compare the folding kinetics of this triplex with that of G4.

### 4.2.3 Modified oncogene promoter sequence c-MYC

Interestingly, by scanning the human genome for potential G4 forming sequences Balasubramanian found that many promoter regions of oncogenes are enriched in guanines [37]. One of those regions is the c-MYC promoter region. A modification of the wild-type sequence is presented in this work, denoted as c-MYC. This sequence is known to fold into one well-defined G4 structure, an all-parallel propeller topology with both $Na^+$ and $K^+$ [8]. SRCD and time-resolved CD measurements, in that

case, have been performed with $Na^+$ since, with $K^+$, the G4 structure is too stable to be denaturated thermically.

## 4.3 Static SRCD Spectra of G4

### 4.3.1 Melting curve - fit procedure

Melting temperatures and enthalpies have been determined as follows, according to Krejtschi and Hauser [183]. Assuming a two-state folding process, SRCD melting curves as well as UV melting curves can be described with a temperature-dependent function as

$$y(T) = (Y_1 + m_1 \cdot T) + (Y_2 + m_2 \cdot T) \cdot \frac{K_{eq}(T)}{1 + K_{eq}(T)} \tag{4.1}$$

where $Y$ and $m$ are fitting parameters. $T$ is the absolute temperature in kelvin and $K_{eq}$ is the equilibrium constant described with Eq. 4.2.

$$K_{eq}(T) = e^{-\frac{\Delta G^0(T)}{RT}} \tag{4.2}$$

where $\Delta G^0(T)$ is the free reaction energy under standard conditions, the gas constant is $R = 8.3134\ J/(molK)$. Equation 4.3 describes $\Delta G^0(T)$:

$$\Delta G^0(T) = \Delta H_m + \Delta C_m \cdot (T - T_m) - T \cdot \left( \frac{\Delta H_m}{T_m} + \Delta C_m \cdot \ln\left(\frac{T}{T_m}\right)\right) \tag{4.3}$$

where $T_m$ is the melting point, $\Delta H_m$ is the enthalphy change and $\Delta C_m$ is the heat capacity change at $T_m$. This method does not allow to determine $\Delta C_m$ values, since they are quite small. Henceforward, it will be assumed to be zero, leading Eq. 4.3 to become

$$\Delta G^0(T) = \left(1 - \frac{T}{T_m}\right) \cdot \Delta H_m \tag{4.4}$$

and thus

$$K_{eq}(t) = e^{-\left(\frac{1}{T} - \frac{1}{T_m}\right)\frac{\Delta H_m}{R}} \tag{4.5}$$

Parameters can vary a lot by only fitting the function of Eq. 4.4 to a single wavelength melting curve. Therefore, melting points were determined by a sigmoidal Boltzmann fit and then fixed in Eq. 4.3. The two lines included in Eq. 4.1, are also neglected. This is necessary for the fit function to converge. Hence, the final fit function reads:

$$y(T) = y0 + \frac{e^{-\left(\frac{1}{T}-\frac{1}{T_m}\right)\cdot\Delta H_m}}{1+e^{-\left(\frac{1}{T}-\frac{1}{T_m}\right)\cdot\Delta H_m}} \tag{4.6}$$

where $y0$ describes an offset.

For a detailed analysis of the temperature-dependent SRCD spectra, we have performed a global analysis by singular value decomposition (SVD) [184] with MATLAB$^{TM}$. SVD can reveal important information by extracting small changes from the data that could not be seen directly. The set of SRCD spectra measured at different temperatures can be converted into one matrix $D$ whose elements are the CD signals as a function of wavelength (rows) and temperature (columns). SVD allows to decompose the matrix $D$ into the product of three matrices ($U$, $S$, $V$), as follows:

$$D = U \cdot S \cdot V^T \tag{4.7}$$

where $S$ is the diagonalized matrix of singular values. Matrix $U$ contains the basis spectra, matrix $V$ contains the associated temperature-dependent vectors. In other words, $U$ contains the spectra of the species contributing to the SRCD spectra and $V$ contains the associated melting curves. In order to determine the minimum number of species involved in the spectral chages of the SRCD spectra, we have first examined the relative magnitude of the singular values. In the respect, we have defined the relative variance ($Var$) of each singular value $S_i$ as [185]:

$$Var = \frac{S_i^2}{\sum_i S_i^2} \tag{4.8}$$

If only two singular values contributes to $> 0.99$ of the relative variance, a two-state model can be considered as valid [185]. In a second step, we have also inspected the basis spectra and the amplitude vectors for judging the validity of the first selection criterion. By reconstructing matrix $D$ with the relevant singular values, one gets rid of most of the noise [184].

## 4.4 Antiparallel human telomeric G4

### 4.4.1 Tel21 with Sodium

Fig. 4.1 shows CD spectra of Tel21 with $Na^+$ at different temperatures. The positive peaks at 295 nm and 245 nm as well as the negative one at 265 nm are characteristic of an antiparallel structure. When temperature is increased, the peaks at 295 nm and 265 nm almost fully disappear upon G4 unfolding.

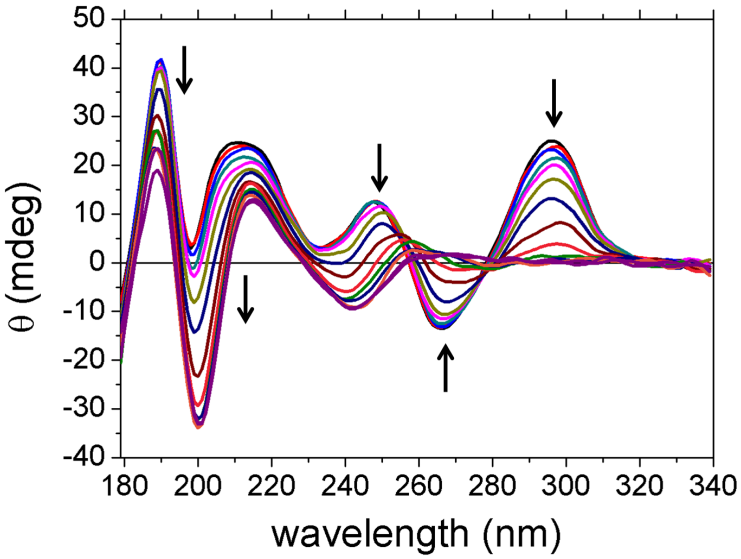


Figure 4.1: Melting spectrum of Tel21 in $NaF$ solution (150 mM $Na^+$). Spectral range: 179 - 339 nm. Arrows indicate increasing temperature. Spectra were taken every 5K between 30°C and 95°C. Shown specta were smoothed (11pts) with the Savitzky-Golan method in OriginPro$^{TM}$. All following shown curves underwent the same smoothing procedure.

In contrast, at lower wavelengths below ca. 260 nm, one observes complicated spectral changes (i.e. amplitude change of the CD signals concomitant with a shift of some peaks) when the temperature is increased. It worthnoting that previous SRCD studies of several G4 structures have suggested that this spectral domain mainly reflect intra-strand interactions between the bases rather than the guanine arrangement in the G4 core. After G4 dissociation at 95°C, two positive and two negative peaks are observed at 190 nm, 210 nm, 200 nm and 245 nm, respectively. Remarkably, these spectral features are similar to those of dGMP and TMP monomers in solution indicating weak interactions between the bases at this temperature [186].

Fig 4.2 illustrates the melting curves at major peak changes, as a function of the

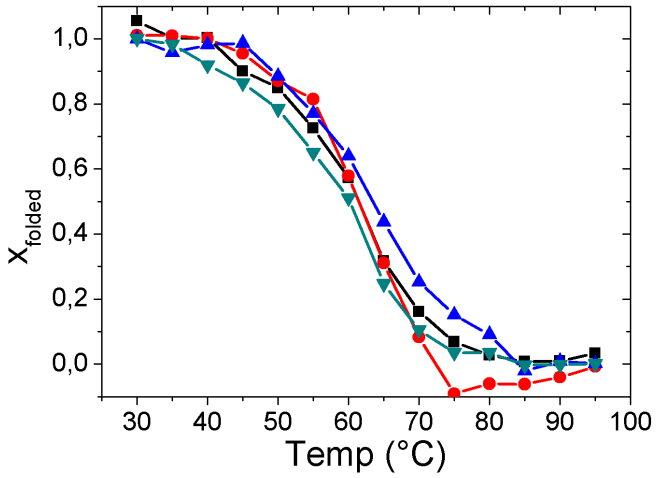


Figure 4.2: Melting curves of Tel21 before SVD analysis at all major peaks (except the one at 189 nm, which is discussed at the end with other G4 structures): 295 nm (black), 265 nm (red), 245 nm (blue) and 201 nm (cyan).

folded species $x_{folded}$ with $x_{folded} + x_{unfolded} = 1$. $x_{folded} = 1$ means all DNA strands are folded into G4 and $x_{folded} = 0$ means all strands are unfolded. The four melting curves exhibit similar behavior but cannot be superimposed. Melting temperatures extracted from sigmoidal fits are listed in Table 4.1. Slight differences of a few degrees are observed between the four $T_m$. Enthalpies are found to be similar, excepted at 265 nm where a larger value is obtained.

| Wavelength (nm) | $T_m$ (°C) | $\Delta H(kJ/mol)$ |
|---|---|---|
| 295 | 60.4 ± 0.6 | -160 ± 10 |
| 265 | 61.9 ± 0.8 | -221 ± 32 |
| 245 | 63.6 ± 0.8 | -143 ± 12 |
| 201 | 59.3 ± 0.4 | -148 ± 13 |

Table 4.1

SVD analysis has been performed on Tel21 SRCD spectra. Singular values are plotted as the relative variance $Var_i$ in Fig. 4.3. The first two singular values (0.758 and 0.237) are found to contribute 99.5% of the total variance. Fig. 4.4 shows the basis spectra and their amplitude vectors associated to the three more significant singular values together with the reconstructed SRCD spectra with the first two singular values. Interestingly, the first basis spectrum is similar to that of the folded anti-parrallel G4 measured at 30°C, while the second one displays the characteristic

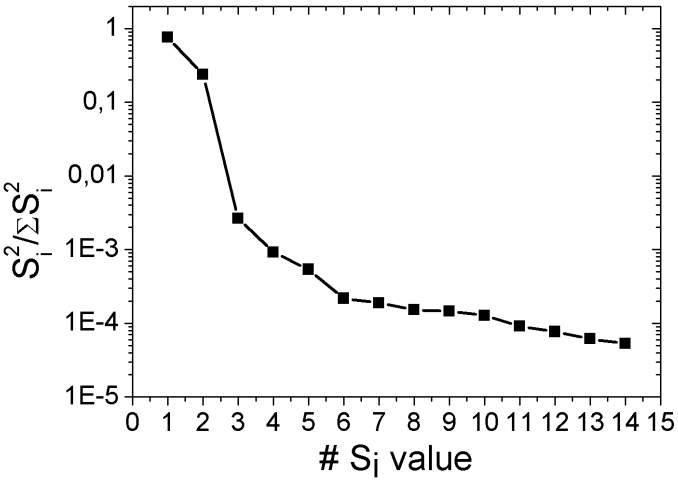


Figure 4.3: Squared singular values normalized by the sum of all squared singular values arranged in decreasing magnitude. $\frac{S_i^2}{\Sigma S_i^2} = Var$

spectral features of the unfolded DNA strand measured at 95°C. As expected in a simple two-state model (folded state $\rightleftharpoons$ unfolded state), the two associated amplitude vectors exhibit the same sigmoidal variation. Their global fit leads to $T_m = 61.1 \pm 0.4$°C and $\Delta H^0 = -147 \pm 9\ kJ/mol$. These values are comparable to that extracted from the individual fits (Table 4.1).

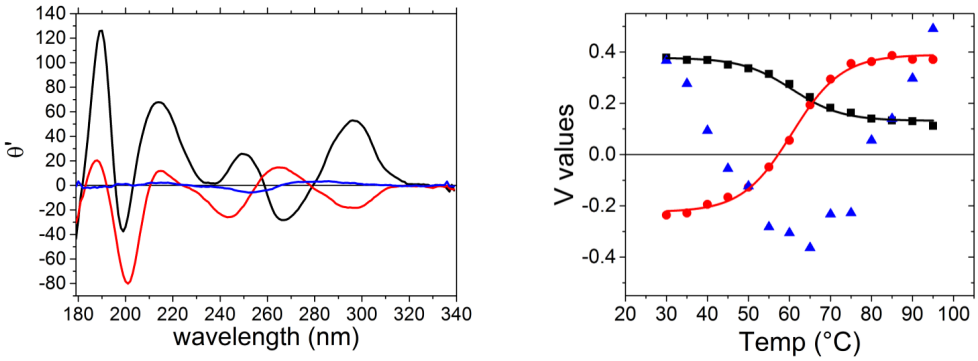


Figure 4.4: Results of SVD analysis. Left graph shows the basic spectra of the first (black), second (red) and third (blue) singular value The unit $\theta'$ is proportional to $\theta$: $\theta' = \theta \cdot V_{value}(T)$. Right: corresponding V vectors. Color code is maintained.

As illustrated on Fig. 4.5, the two first singular values are sufficient to reproduce most of the observed spectral changes of Tel21 SRCD spectra. However, one can see that at 260 nm there are slight differences. A careful inspection of SVD matrixes provides evidence that this stems from the neglected third singular value. Its basis

spectrum and amplitude vector exhibit spectral features and variation that are clearly distinct from noise. Such behavior provides evidence that an additional species should be taken into account in the global analysis of the SRCD spectra of Tel21.

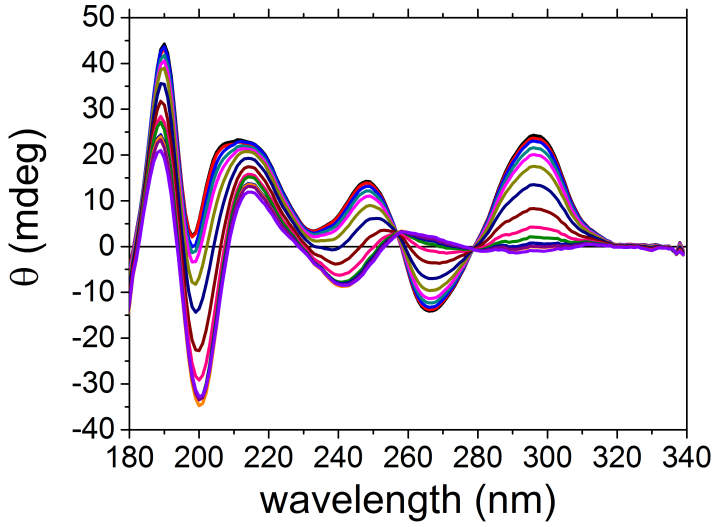


Figure 4.5: Reconstructed CD melting curves of Tel21 with $Na^+$ by using only $S_1$ and $S_2$.

Several hypotheses can be envisaged with regard to this behavior:

- the first one involves the existence of several different folded conformations at room temperatures. This is consistent with NMR study of Tel21 with $Na^+$ that has suggested the presence of different conformations at room temperature [45]. As a matter of fact, the third basis spectrum displays a negative peak at 250 nm and a positive one at 280 nm that may be related to a particular structure. But its bell-shaped temperature dependence cannot not be properly fitted by sigmoidal functions. Therefore, it has not been possible to associate a melting temperature to this putative conformation.

- the second hypothesis involves the detection of a folding intermediate formed in the course of the thermal denaturation of Tel21. Such possibility has been previously considered for Tel22 with $K^+$ by the group of Chaires [23, 187]. Such a scenario can be confirmed by time-resolved melting experiments. It will be discussed further in Chapter 5.

- finally, one might argue that the contribution of this third species could be an artifact resulting from hypothetical continuous spectral shifts or shape changes during thermal unfolding of Tel21, which would indeed not be properly captured

by our global analysis. It is noteworthy that G4 unfolding involves G-quartet dissociation concomitant to a change of the loop conformation. As unfolding proceeds, the base stacking inside the G4 core is expected to decrease while that inside the loops increases. These intricated processes may lead to complex spectral CD changes that cannot properly distinguished by SVD.

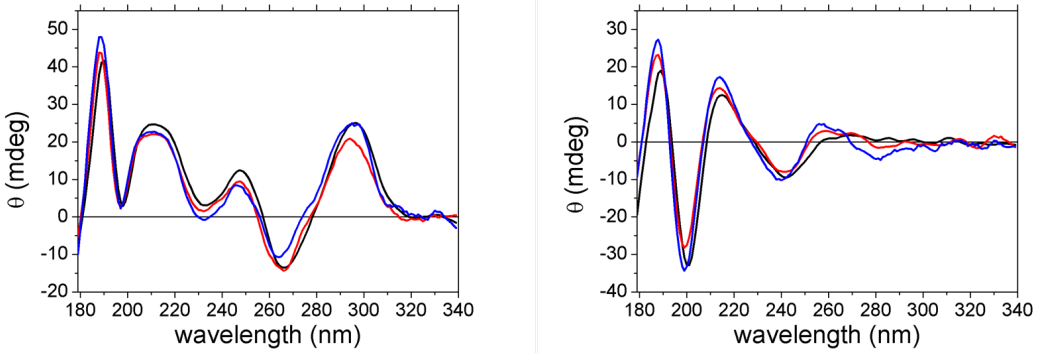


Figure 4.6: CD spectra of Tel21 at different $Na^+$ concentrations. Black: 150 mM, red: 90 mM, blue: 25 mM. Curves in the left were taken at 30°C (25°C for 25mM) Curves on the right at high temperature (25 mM: 70 °C, 90 mM: 75°C, 150 mM: 95°C.

In order to disentangle the origin of the complex spectral changes of Tel21, we have undertaken a study at various $Na^+$ concentrations. As illustrated by Fig. 4.6, the SRCD spectra of Tel21 measured for three concentrations of cations exhibit a similar behavior (both in shape and in intensity) that is further confirmed by SVD analysis (see Fig. SI:7.1).

| $[Na^+]$ (mM) | $T_m1$ (°C) | $\Delta H(kJ/mol)$ (at 295 nm) | $T_m2$ (°C) | $\Delta H2(kJ/mol)$ (at 295 nm) |
|---|---|---|---|---|
| 25 | 47.4 ± 0.5 | -143 ± 50 | - | - |
| 90 | 54.2 ± 0.7 | -145 ± 52 | - | - |
| 150 | 60.4 ± 0.7 | -160 ± 10 | - | - |
| 500 | 71.4 ± 0.1 | -160 ± 5 | 38.0 ± 0.4 | -80 ± 1 |
| 900 | 74.5 ± 0.1 | -147 ± 4 | 48.0 ± 3.3 | -52 ± 20 |

Table 4.2

This provides evidence that, at these concentrations, the G4 topology is barely affected suggesting that all the binding sites of inside the G4 core are filled. In the

meantime, we have observed an increase of $T_m$ with the concentration of cation that is due to an increase of G4 stability. Such a trend is consistent with non-specific cation binding outside the G4 core, which stabilizes the DNA backbone. The values of $T_m$ and the associated enthalpies extracted from the sigmoidal fits of CD or UV melting curves at 295 nm are given in the Table 4.2.

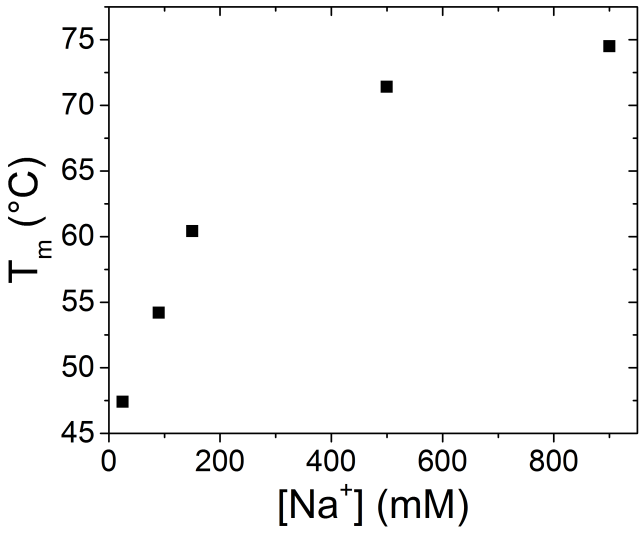


Figure 4.7: $T_m$ at 295 nm at different sodium concentrations. $T_m$ at 25, 90 and 150 mM were extracted from CD melting curves. $T_m$ at 500 and 900 mM were determined from UV melting curves.

It is worth noting that the melting curves measured at concentrations of cations higher than 150 mM exhibit clearly two transitions (Fig. SI:7.2), whose origin has not been clarified yet. However, one can see in Fig. 4.7 that variation of $T_m$ as a function of concentration of cations significantly diminishes above 150 mM. This suggests that most of the external binding sites are occupied at this concentration.

### 4.4.2 Tel22 with Sodium

In order to check the hypothesis of an equilibrium between several distinct conformations of Tel21 in presence of $Na^+$, we have performed a comparative study with Tel22, under similar conditions. It is known since 1993 that Tel22 forms only one well-defined antiparallel basket G4 structure with sodium [1]. SRCD spectra of Tel22 are depicted in Fig. 4.8. They are qualitatively similar to those of Tel21. The main difference relies on the intensity of the negative peak at 265 nm that is more intense for Tel22. Results of SVD analysis are illustrated in Fig. 4.9.

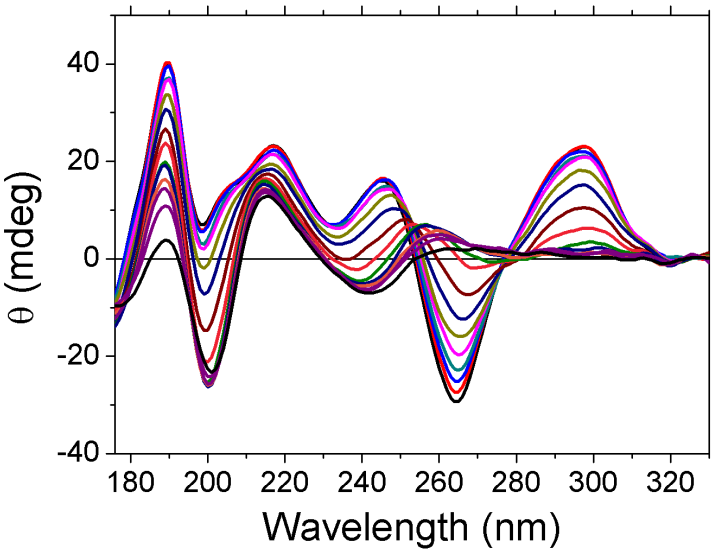


Figure 4.8: CD spectra of Tel22 in 150 mM $Na^+$. Spectral range is 176-330 nm. CD spectra were recorded from 25°C to 95°C every 5K. Black curve: 35°C, dark green curve: 95°C.

Like for Tel21, the two first singular values (0.80 and 0.19) are found to contribute 99% of the total variance. The two first basis spectra display the spectral features of the folded antiparallel G4 structure and the unfolded DNA strand, respectively. The global fit of the two first amplitude vectors leads to $T_m = 57.1 \pm 0.7$°C and $\Delta H^0 = -101 \pm 13\ kJ/mol$. Despite the reconstructed SRCD spectra reproduce the main spectral changes of Tel22 SRCD spectra, one can see notable differences in the spectral region around 260 nm and 216 nm. Inspection of SVD matrices associated to the third and fourth singular values reveals non-negligible spectral variations. Contrary to our expectations, the study of Tel22 in the presence of $Na^+$ reveals complex spectral changes as in the case of Tel21 despite its lower polymorphism. One can conclude that SRCD spectroscopy in the case of Tel21 cannot distinguish

the different conformations that might be present at the thermal equilibrium.

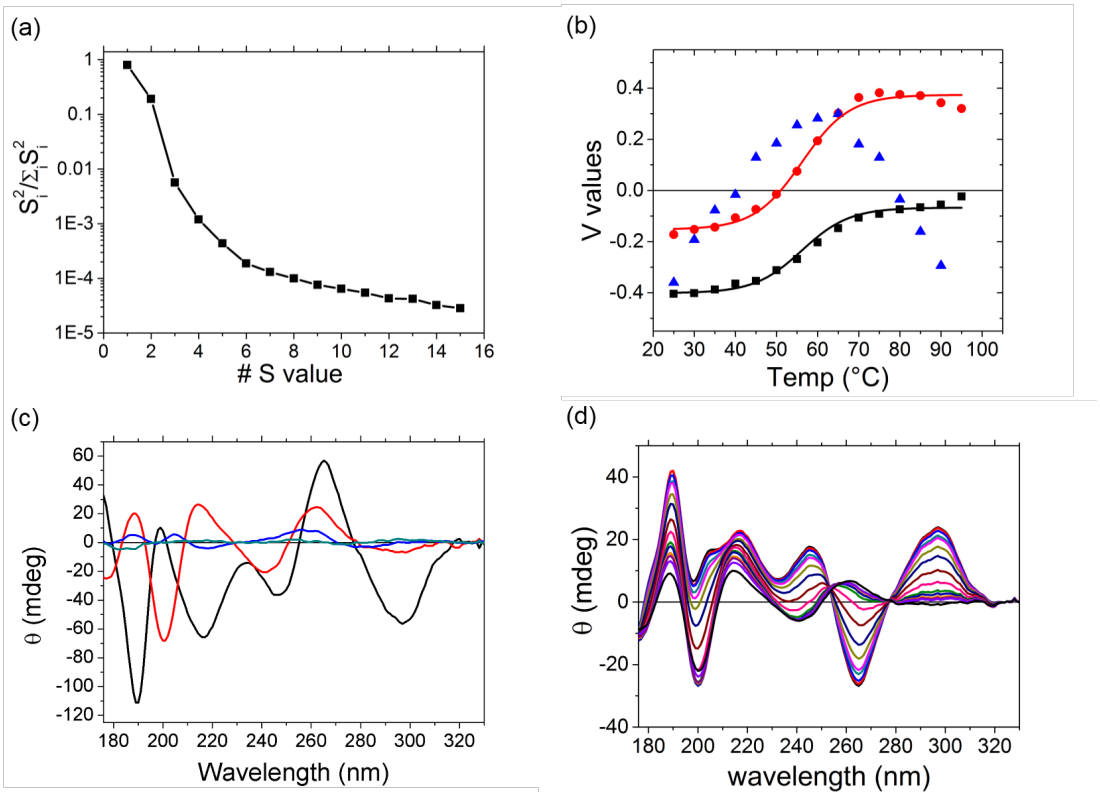


Figure 4.9: Results of SVD analysis of Tel22 with $Na^+$. Graph (a) ($Var$) Squared singular values normalized by the sum of all squared singular values arranged in decreasing magnitude. Graph (b): V vectors to the first three singular values of first (black), second (red) and third (blue) singular value with Boltzmann fits. Graph (c) shows the basic spectra of the first four S values (color code is as in (b)). Graph (d) reconstructed CD melting curves by using only $S_1$ and $S_2$.

### 4.4.3 22CTA with Potassium

22CTA is known to fold into an antiparallel chair-type G4 with two G-quartets and a third one containing $C \cdot G$ Watson-crick base pairs [3]. SRCD spectra of 22CTA as a function of temperature are given in Fig. 4.10.

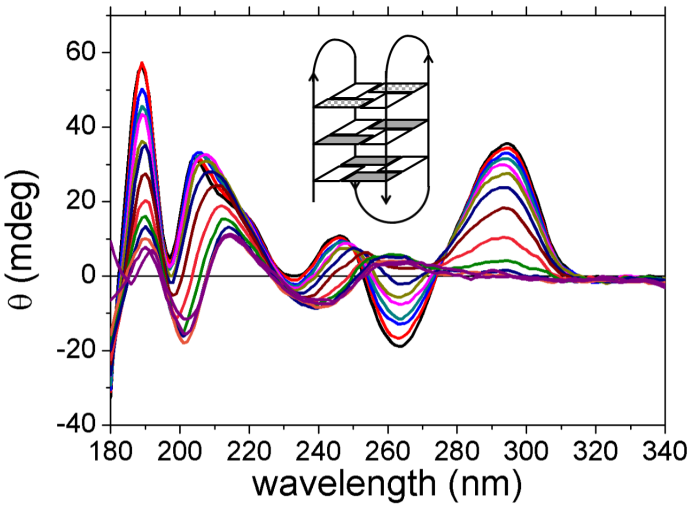


Figure 4.10: CD spectra of 22CTA in 150 mM $K^+$ phosphate buffer. Spectra were recorded between 180 and 340 nm in the temperature range of 30°C to 95°C every 5°C. Depicted G4 structure contains two Cytosines indicated by dotted rectangles.

The spectral region above 240 nm shows clearly the antiparallel character of 22CTA. Individual fits of the melting curves at the major CD peaks revealed $T_m = 63.2 \pm 2.3$°C and $\Delta H^0 = -140 \pm 20\ kJ/mol$. Only the peak at 263 nm showed a different behavior with $T_m = 51.6 \pm 2.3$°C, $\Delta H^0 = -119 \pm 12\ kJ/mol$. This might be related to the presence of the third quartet containing $C \cdot G$ base pairs and the participation of cytosines. SVD analysis has been performed to see whether more details can be extracted from the SRCD spectra of 22CTA (see Fig. 4.11). The first two singluar values ($Var_1 = 0.900$ and $Var_2 = 0.080$) account for 98% of the total variance. Global fitting of the two first vectors yields $T_m = 65 \pm 1$°C and $\Delta H^0 = -86 \pm 12\ kJ/mol$. $\Delta H^0$ is smaller than that of the individual fits, since it is the overall value including peaks at 180 nm and 263 nm. Analysis of the basis spectra show that the third singular value cannot be neglected. This is also visible in the reconstructed SRCD spectra with the first two singular values, which do not completely reproduce the experimental SRCD spectra of Fig. 4.10, especially around 260 nm and 210 nm.

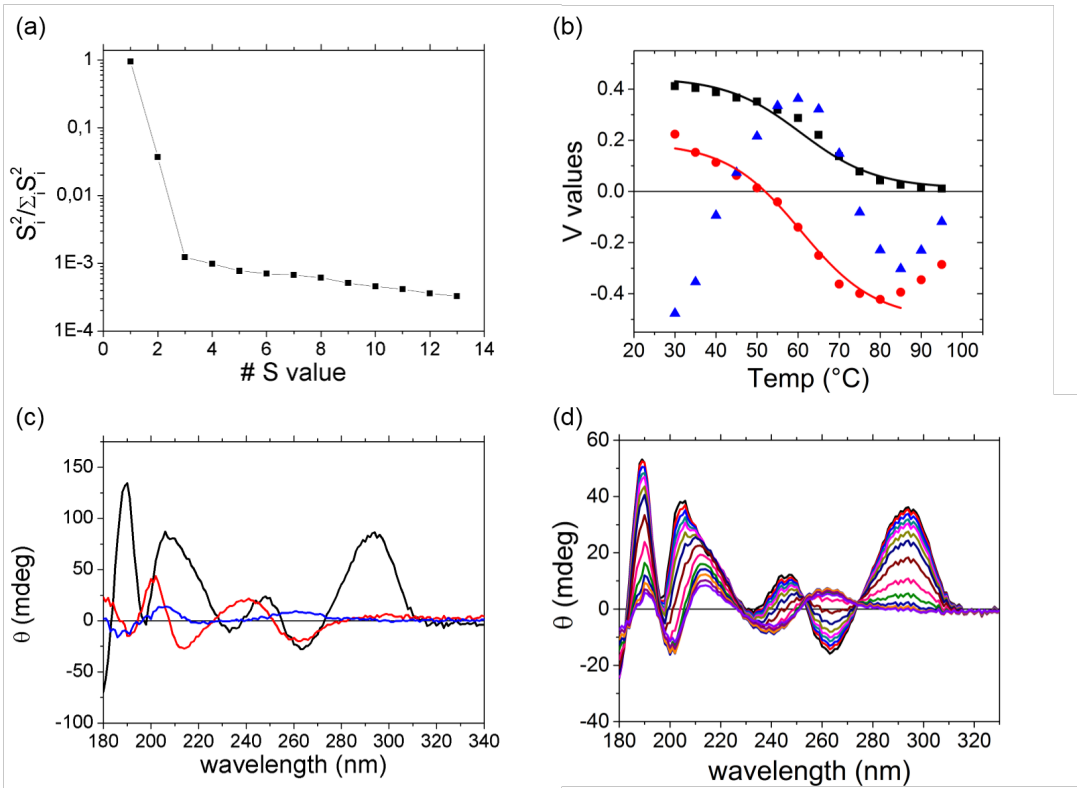


Figure 4.11: Results of SVD analysis of 22CTA with $K^+$. Graph (a) ($Var$) Squared singular values normalized by the sum of all squared singular values arranged in decreasing magnitude. Graph (b): V vectors to the first three singular values of first (black), second (red) and third (blue) singular value with fits. Graph (c) shows the basic spectra of the first three S values (color code is as in (b)). Graph (d) reconstructed CD melting curves by using only $S_1$ and $S_2$.

## 4.5 hybrid-type human telomeric G4

### 4.5.1 Tel22 with Potassium

The SRCD spectra depicted in Fig. 4.12 clearly show the hybrid-type G4 nature of Tel22 with $K^+$. They exhibit positive peaks at 293 nm and 273 nm, but none at 245 nm.

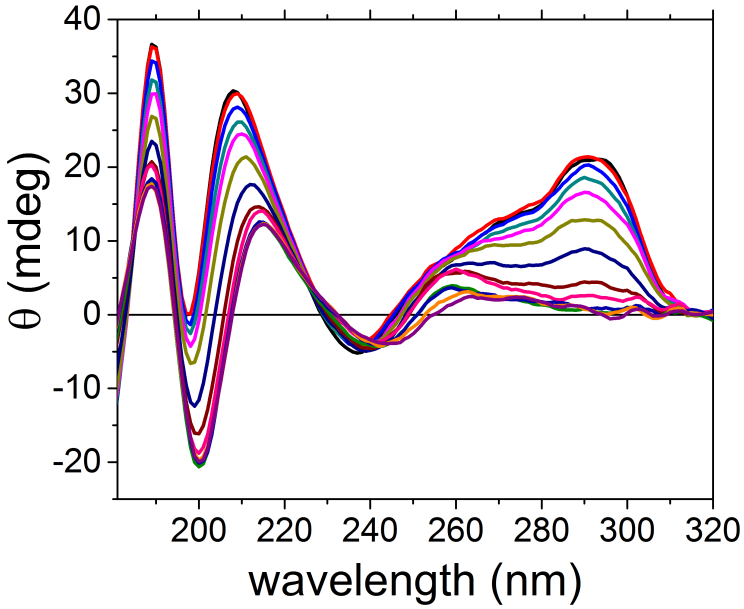


Figure 4.12: CD spectra of Tel22 in 100 mM $K^+$. Spectral range is 178-320 nm. CD spectra were recorded from 35°C to 95°C every 5°C. Black curve: 35°C, dark green curve: 95°C.

The blue part of the SRCD spectra ($\leq$ 230 nm) is similar to that of Tel21 and Tel22 with $Na^+$. The fit of the melting curves at the major peaks yielded $T_m = 61.6 \pm 0.5$°C and $\Delta H^0 = -146 \pm 6\ kJ/mol$. Despite the presence of a mixture of hybrid forms at room temperature, all melting curves were found to superimpose. Additionally, SVD analysis yields two singular values contributing to 99.7% to the SRCD spectra ($Var_1 = 0.868$ and $Var_2 = 0.129$, see Fig. 4.13). This is significantly larger than in the case of Tel21 and Tel22 with $Na^+$. Global analysis of the first two $V$ vectors revealed $T_m = 62.3 \pm 0.3$°C and $\Delta H^0 = -165 \pm 10\ kJ/mol$, in agreement with the values of single wavelength fits. Coherently, the reconstructed SRCD spectra with the first two singular values (Fig. 4.13(d)) completely reproduces the experimental SRCD spectra of Fig. 4.12.

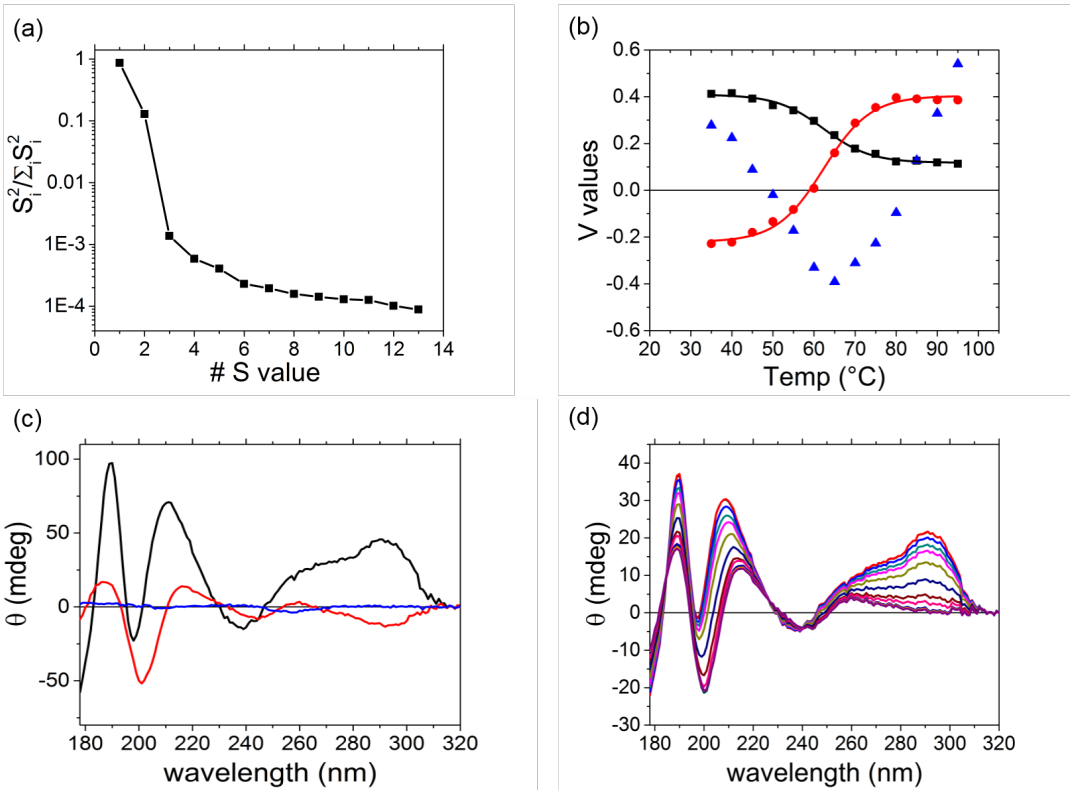


Figure 4.13: Results of SVD analysis of Tel22 with $K^+$. Graph (a) ($Var$) Squared singular values normalized by the sum of all squared singular values arranged in decreasing magnitude. Graph (b): V vectors to the first three singular values of first (black), second (red) and third (blue) singular value with fits. Graph (c) shows the basic spectra of the first four S values (color code is as in (b)). Graph (d) reconstructed CD melting curves by using only $S_1$ and $S_2$.

### 4.5.2 2GKU with Potassium

In addition to Tel22 with $K^+$, we have studied 2GKU in order to identify the spectral signatures of the hybrid1 topology [44]. Fig. 4.14 shows the SRCD spectra of 2GKU at different temperatures. SRCD spectra of 2GKU display the same peaks as those of Tel22 with $K^+$. However, their relative intensities are different. The melting curves at the major peaks are found to superimpose and yield $T_m = 68.0 \pm 0.4$°C and $\Delta H^0 = -181 \pm 13\ kJ/mol$. These values are larger than those of Tel22 with $K^+$. 99.3% of the total variance is obtained by considering only the first two singular values ($Var_1 = 0.956$, $Var_2 = 0.037$), see Fig. 4.15. Coherently, global fits of $V$ vectors (see (b)) yielded $T_m = 68.0 \pm 0.4$°C and $\Delta H^0$ of $-225 \pm 20\ kJ/mol$, which are comparable to those determined by individual fits of the melting curves at the CD peaks. Reconstructed SRCD spectra (Fig. 4.15(d)) did not reveal a significant difference with the experimental one.

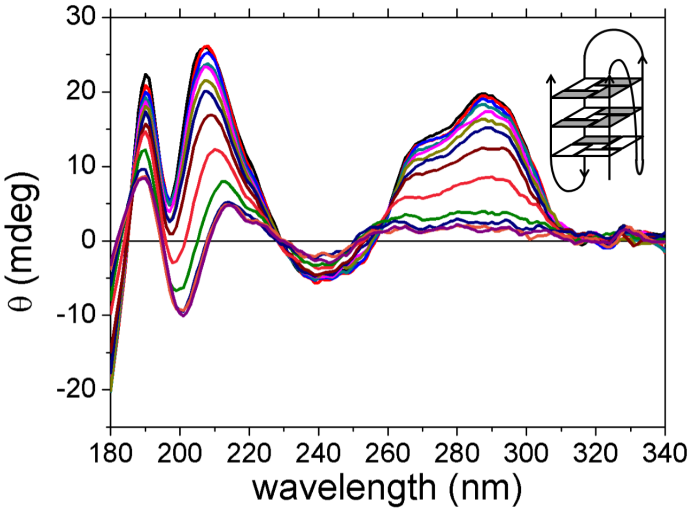


Figure 4.14: CD spectra between 180 and 340 nm of 2GKU in 150 mM $K^+$ phosphate buffer. Spectra were recorded every 5K in the range of 30°C to 90°C.

As a matter of fact, our present study did not reveal the presence of possible stable parallel G4-shaped intermediates during melting of Tel22 and 2GKU with $K^+$ that have been found by the group of Chaires [23]. This difference may result from our different experimental conditions such as salt, buffer and strand concentrations.

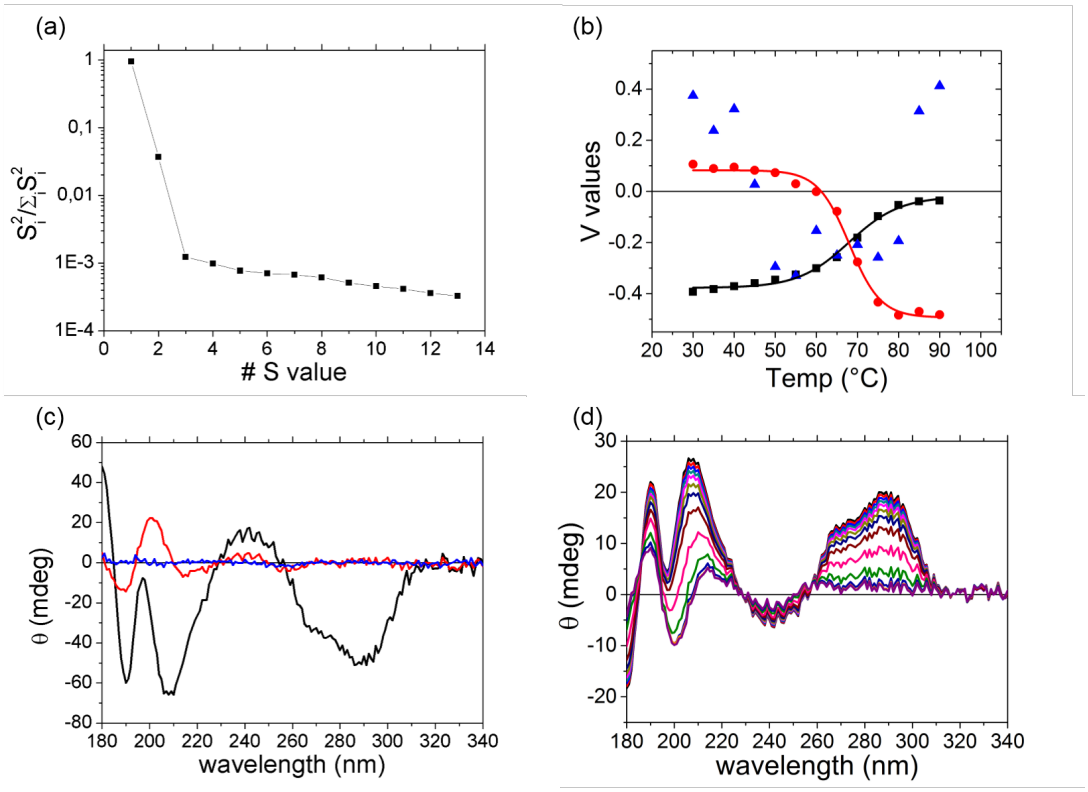


Figure 4.15: Results of SVD analysis of 2GKU with $K^+$. Graph (a) ($Var$) Squared singular values normalized by the sum of all squared singular values arranged in decreasing magnitude. Graph (b): V vectors to the first three singular values of first (black), second (red) and third (blue) singular value with fits. Graph (c) shows the basic spectra of the first four S values (color code is as in (b)). Graph (d) reconstructed CD melting curves by using only $S_1$ and $S_2$.

## 4.6 Thrombine binding aptamer

### 4.6.1 TBA G4 with Potassium

SRCD of TBA shown in Fig. 4.16 exhibit positive peaks at 293 nm and 248 nm, respectively, and a negative one at 263 nm. They are characteristic of the antiparallel topology of TBA G4. At high temperatures, the spectrum of the unfolded structure is different to those of Tel21, Tel22 and 2GKU. This is due to the absence of adenine bases (see [186]).

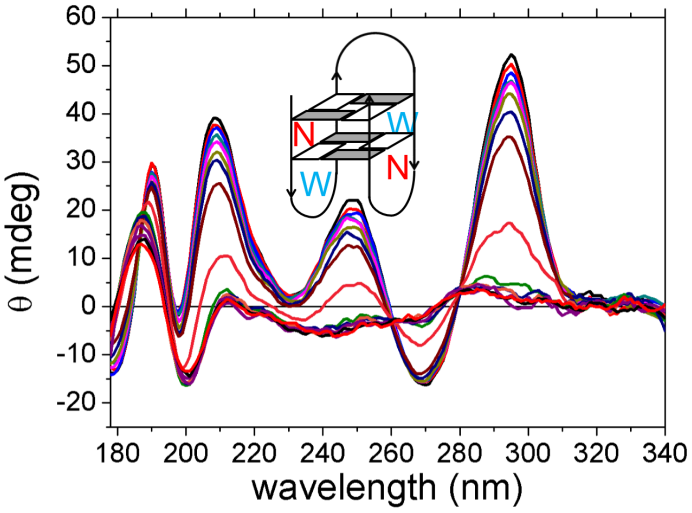


Figure 4.16: CD spectra of TBA 15mer in 150 mM $K^+$ phosphate buffer. Spectra were recorded between 178 and 340 nm in the temperature range of 10°C to 95°C every 5°C. Two temperatures are missing, though: 55°C and 65 °C. G4 structure is that of TBA 15mer. Guanine in *syn* and *anti* orientation are colored white and grey, respectively.

SVD analysis illustrated in Fig. 4.17 shows that two singular values ($Var_1 = 0.912$, $Var_2 = 0.080$) yield 99.2% of the overall variance. Global fitting of the first two $V$ vectors leads to $T_m = 48.3 \pm 0.8$°C and $\Delta H^0 = -144 \pm 16\ kJ/mol$ in agreement with the independent fits of the CD peaks. The reconstructed SRCD spectra with these two singular values correctly reproduce the experimental spectra (see Fig. 4.16).

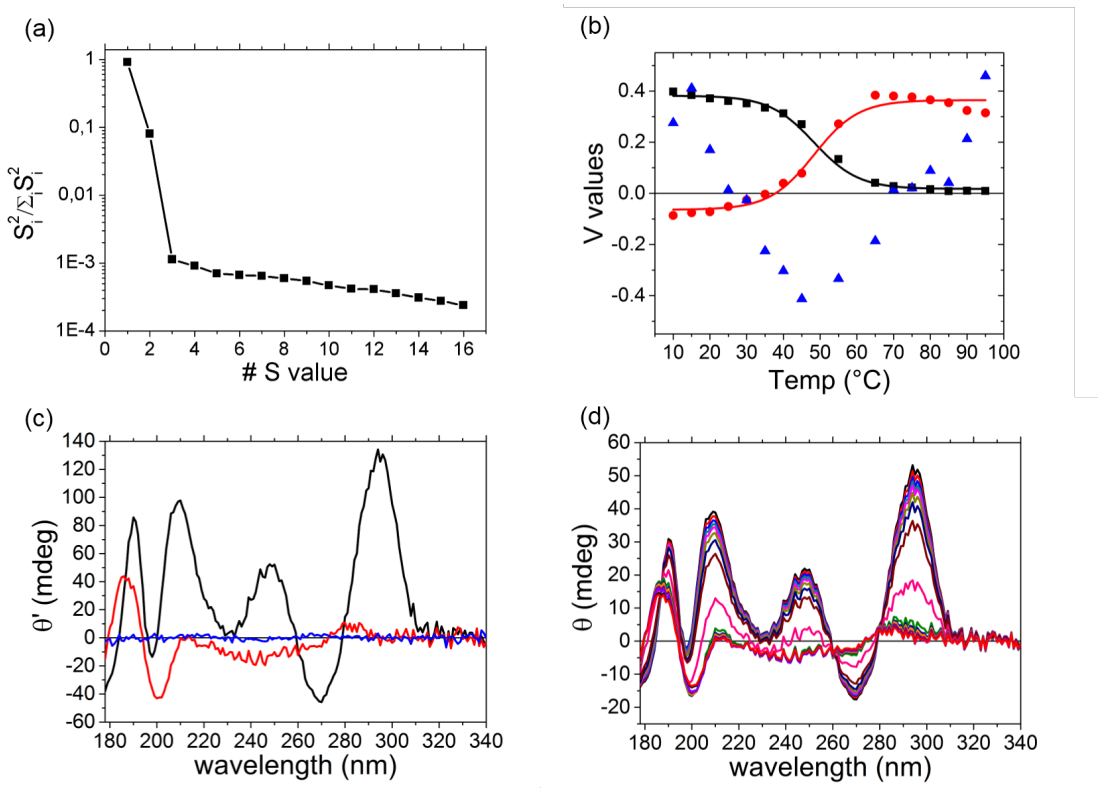


Figure 4.17: Results of SVD analysis of TBAs G3 with $K^+$. Graph (a) ($Var$) Squared singular values normalized by the sum of all squared singular values arranged in decreasing magnitude. Graph (b): V vectors to the first three singular values of first (black), second (red) and third (blue) singular value with fits. Graph (c) shows the basic spectra of the first four S values (color code is as in (b)). Graph (d) reconstructed CD melting curves by using only $S_1$ and $S_2$.

### 4.6.2 TBA G3

SRCD spectra of TBA G3 at different temperatures is given in Fig. 4.18. The spectrum of the folded structure is different from that of TBA G4 in the spectral region between 240 nm and 320 nm. In contrast to that, at lower wavelengths (> 240 nm) it looks similar to that of TBA G4. This seems to confirm that this spectral region is mainly sensitive to intrastrand interactions of the bases rather than the guanine arrangement. At high temperature, the spectra of TBA G3 are similar to those of TBA G4 (see Fig. 4.18).

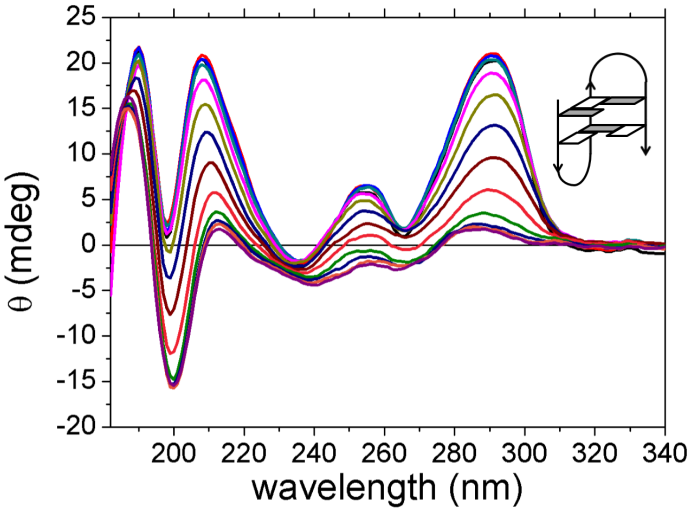


Figure 4.18: CD spectra of TBA 15mer in 150 mM $K^+$ phosphate buffer. Spectra were recorded between 183 nm and 340 nm in the temperature range of 10°C to 70°C every 5°C. Shown structure is that of TBA triplex.

SVD analysis of the SRCD spectra is shown in Fig. 4.19. The first two singular values ($Var_1 = 0.847$ and $Var_2 = 0.141$) account for 98.8% of the spectral changes with temperature. Global fitting of the first two $V$ vectors gives $T_m = 43.5 \pm 0.3$°C and $\Delta H^0 = -135 \pm 7 kJ/mol$ (see Fig. 4.6.2(b)). The reconstruction of the SRCD spectra with the first two singluar values allow to reproduce the complete experimental SRCD melting spectra (see Fig. 4.6.2(d)).

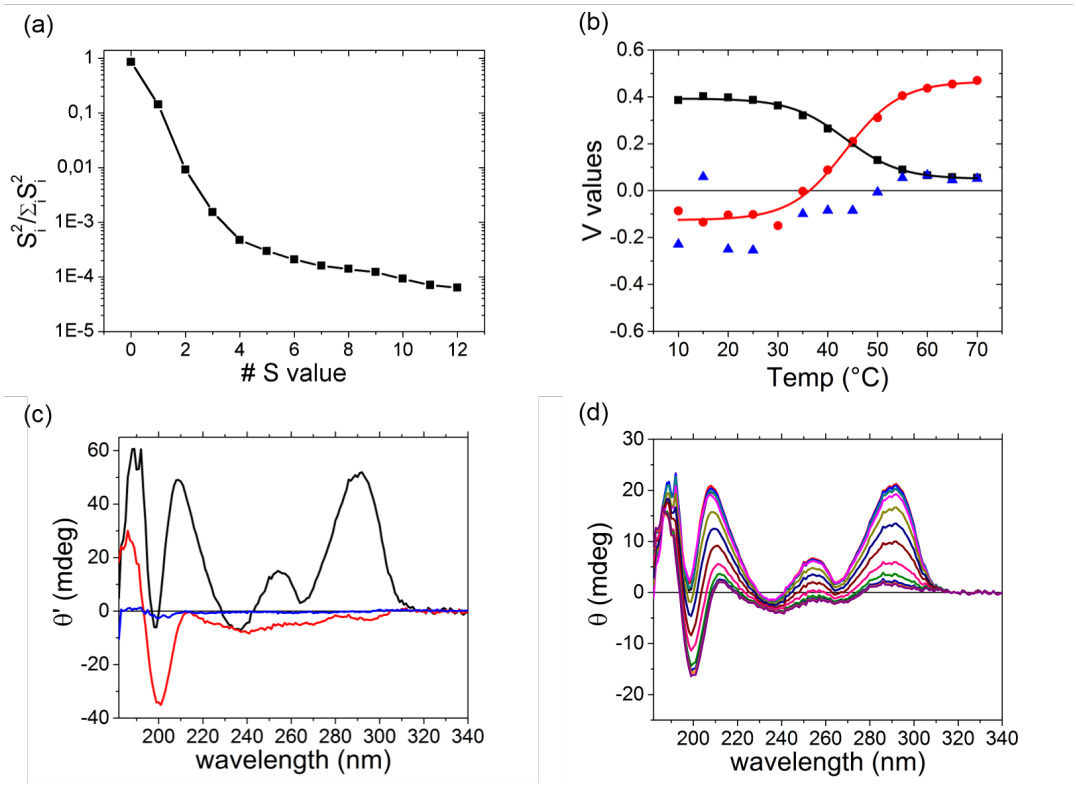


Figure 4.19: Results of SVD analysis of TBA G3 with $K^+$. Graph (a) ($Var$) Squared singular values normalized by the sum of all squared singular values arranged in decreasing magnitude. Graph (b): V vectors to the first three singular values of first (black), second (red) and third (blue) singular value with fits. Graph (c) shows the basic spectra of the first four S values (color code is as in (b)). Graph (d) reconstructed CD melting curves by using only $S_1$ and $S_2$.

## 4.7 c-MYC

Finally, a representative for the all-parallel G4 topology has been investigated: c-MYC. Fig. 4.20 shows its SRCD spectra at different temperatures. At room temperature, they exhibit a positive peak at 265 nm and a negative one at 245 nm that are characteristic for all-parallel G4 (see Sec. 3.2.5). Interestingly, at 189 nm a negative peak is observed. This contrasts with all other studied structures. From a previous study of the CD spectra of isolated nucleosides in solution, we can attribute this band to the thymine nucleosides [186]. Its negative sign may reflect a different arrangement of the thymines in this sequence in comparison to the other structures. Further investigations are however necessary to verify this assumption.

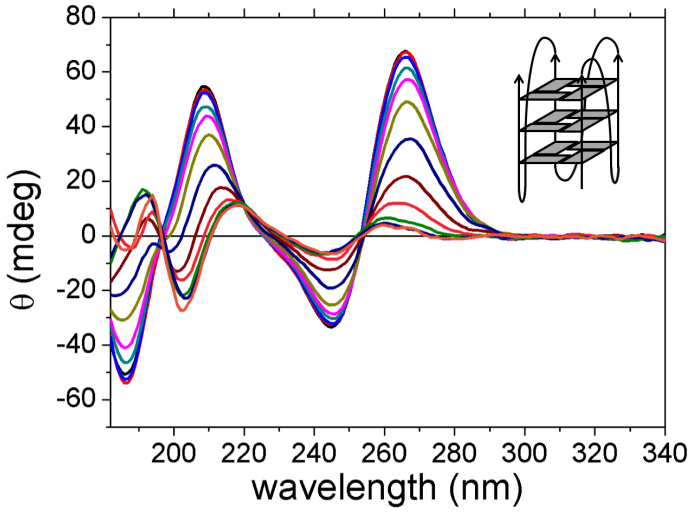


Figure 4.20: CD spectra of c-MYC in 150 mM $Na^+$ phosphate buffer. Spectra were recorded between 181 nm and 340 nm in the temperature range of 25°C to 80°C every 5°C. Shown structure is that of all-parallel c-MYC.

SVD analysis shows that only two singular values necessary ($Var_1 = 0.948$ and $Var_2 = 0.045$) yield 99.3% of the overall variance (see Fig. 4.21). Global fitting of the two $V$ vectors yields similar values than independent fitting of the melting curves with $T_m = 54.4 \pm 0.5$°C and $\Delta H^0 = -150 \pm 8\ kJ/mol$. Consistently, two singular values are sufficient to reproduce the overall CD melting curves (see Fig. 4.21).

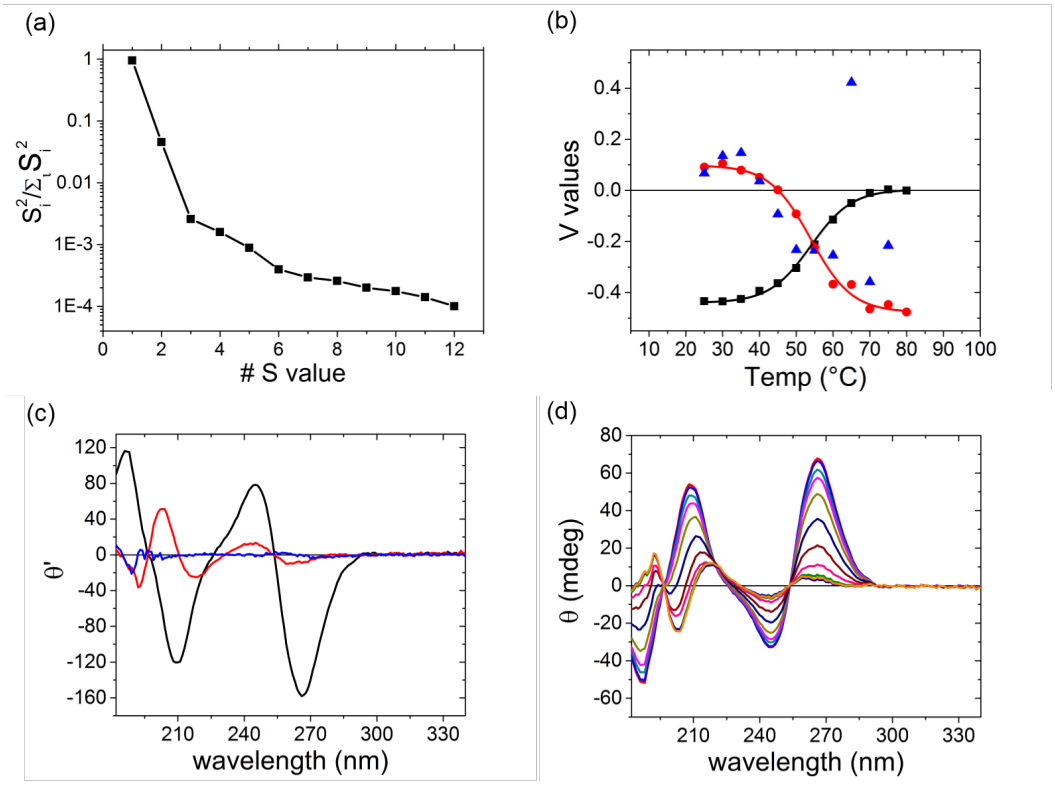


Figure 4.21: Results of SVD analysis of c-MYC with $Na^+$. Graph (a) ($Var$) Squared singular values normalized by the sum of all squared singular values arranged in decreasing magnitude. Graph (b): V vectors to the first three singular values of first (black), second (red) and third (blue) singular value with fits. Graph (c) shows the basic spectra of the first four S values (color code is as in (b)). Graph (d) reconstructed CD melting curves by using only $S_1$ and $S_2$.

## 4.8 Conclusion

By comparing the SRCD spectra of different G4 structures, we have observed in all cases that more than 98% of the spectral changes induced by temperature increase are consistent with the contribution of only two species, a folded one and an unfolded one. However, a careful analysis of the SRCD spectra has revealed subtle differences between the different sequences in the light of a SVD analysis

In particular, we have observed when a single conformation is present at room temperature (Tel22 with $Na^+$, 2GKU, 22CTA, TBA G4, c-MYC) the first $V$ vector always becomes zero at high temperatures. Contrarily, when a mixture of G4 structures exists at room temperature, such as for Tel21 and Tel22 with $K^+$ we have noticed that the first $V$ vector behave distincly. This peculiar behavior may be an indication that one sole basis spectrum/vector does not suffice to capture the spectral changes of these distinct formations.

However, SVD analysis did not allow us to extract their spectra. This could be due to the fact that these conformations have quite similar melting temperatures.

At high temperature ($\gg T_m$) we have observed that SRCD spectra of all G4 become similar to that of the isolated nucleosides in solution. This provides clear evidence that most of the interactions between the bases have been disrupted (intra- and interstrand interactions) at these temperatures. Therefore, these CD signals which span the spectral region between 180 nm and 280 nm stem from the intrinsic chirality of the nucleobases.

On the other hand, those static CD studies show that equilibrium measurements at different temperatures cannot provide information on the folding/unfolding pathways of G4. To this end, we undertook time-resolved CD studies on different G4 forming sequences representing distinct topologies. Their results are illustrated in the following chapters.

# Chapter 5

# Folding dynamics of Tel21 & Tel22 with Sodium

> Our life is frittered away by detail.
> Simplify, simplify.
>
> *Henry D. Thoreau*

## Contents

## 5.1 Preface

As discussed in Chapter 2, there is still a limited knowledge about G4 folding (for recent review see [20]). For better understanding, we have performed time-resolved thermal studies with a T-Jump setup combined with a TRCD/time-resolved absorption detection on the time-scale of a few microseconds up to seconds. This chapter focuses on the results of Tel21 with $Na^+$ at two different wavelengths (293 nm and 265 nm). Tel21 is the shortest human telomeric G4 sequence, but it is suspected to form a mixture of G4 topologies with $Na^+$ [45]. Therefore, we have performed a comparative study of its dynamics with those of Tel22 with $Na^+$ (150 mM), which is known to adopt a well-defined antiparallel basket topology [1].

In order to get a deeper comprehension of the folding mechanism of Tel21, we undertook a study of the effect of the $Na^+$ concentration on the folding/unfolding kinetics, illustrated in Sec. 5.3. Additionally, time-resolved absorption on the $\mu$s time-scale has also been carried out on Tel21. Results are presented in Sec. 5.4.

Furthermore, absorption dynamics on the $\mu s$ time scale has also been investigated and are presented in Sec. 5.4.

## 5.2 Dynamics on the millisecond time-scale

### 5.2.1 Results of Tel21 with Sodium

The absorption and CD spectra of the folded (15°C) and unfolded (95°C) species of Tel21 are depicted in Fig. 5.1.

As discussed earlier, we observe a dramatic change in CD which denotes the G4 denaturation. We also observe a significant change in absorption. At 293 nm, absorption decreases with increasing temperature, while conversely, at 265 nm it increases. Because absorption is around 9 times higher than at 293 nm, TRCD measurements at 265 nm are noisier than at 293 nm and thus experiments are more difficult. Due to the magnitude of its absolute value, thermal changes in absorption are more easily measured compared to CD.

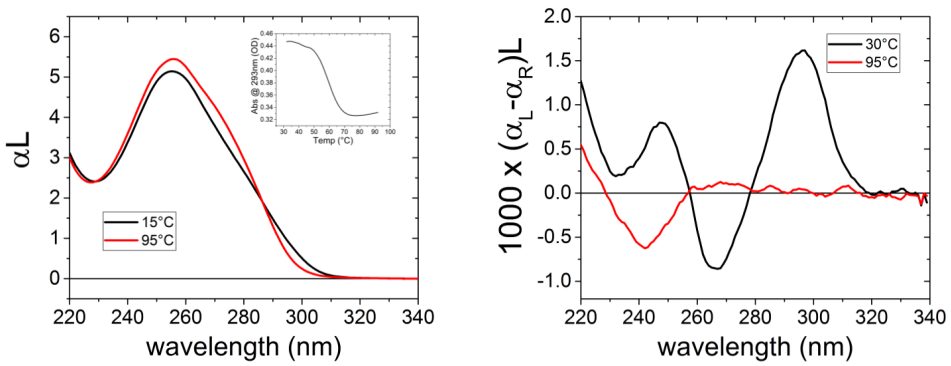


Figure 5.1: Comparison of the absorption and CD spectra of Tel21 with $Na^+$ (150 mM in total, whereof 120 mM due to NaF) at low and high temperature. Buffer conditions were as before: 10 mM phosphate buffer at pH = 7. Pathlength = 100 $\mu$m. Left: Absorption spectrum at 15°C (black) and 95°C (red), inset: melting curve at 293 nm. Right: CD at 30°C (black) and 95°C (red).

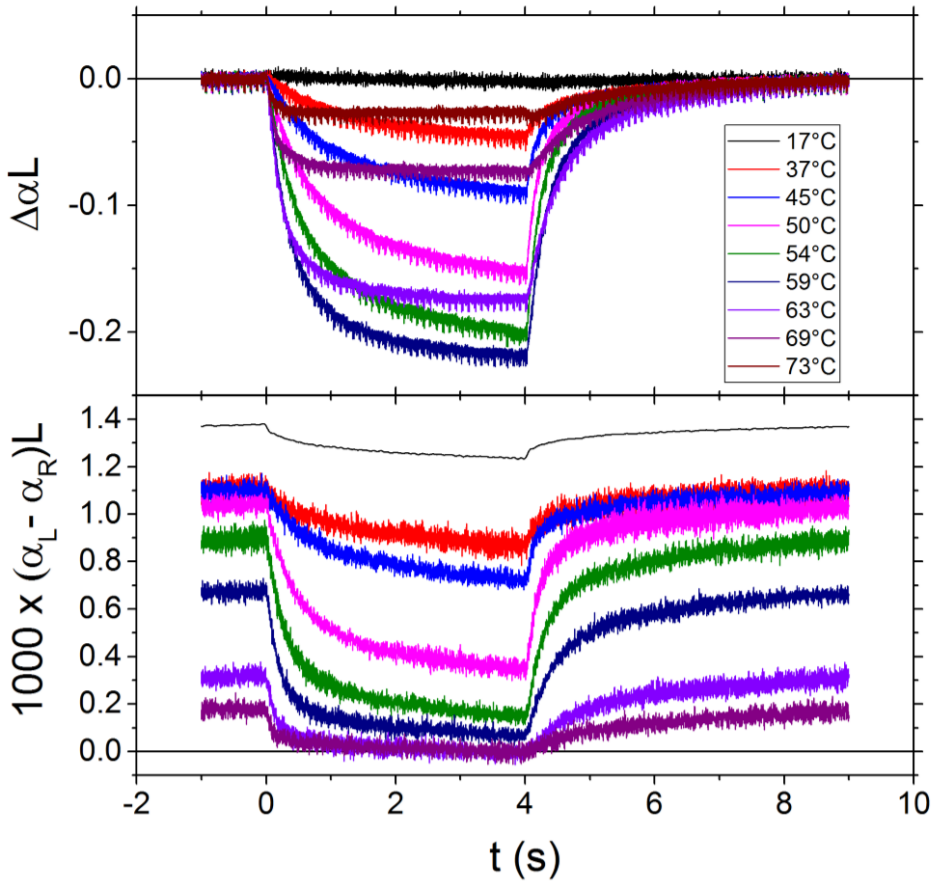


Figure 5.2: Upper graphs: Dynamics of Tel21 at 293 nm at different initial temperatures $T_{in}$ (small box to the right) after a T-jump of $\Delta T = 10$°C. Upper graph: absorption, lower graph: CD. Denaturation occurred during the first 4 s and renaturation between 4 to 9 s.

Dynamical studies on the ms time-scale were performed with a CW heating laser diode as described in Sec. 3.3.3: FWHM(probe) = 80 $\mu$m, FWHM$_{vert}$(pump) = 700 $\mu$m, FWHM$_{hor}$(pump) = 1200 $\mu$m. This is valid for all the following millisecond T-jump experiments, including Chapter 6. As mentioned before, all experiments have been done in a quartz cell with l = 100 $\mu$m. Results at 293 nm are shown in Fig. 5.2. The applied heating-cycle is as follows: The sample has been irradiated from time 0 up to 4 s (heating). After 4 s, the heating irradiation has been halted during 6.2 s (cooling). The sample temperature change is not completely clear up to now, especially the slow increase of temperature (see Sec. 3.3.4.1). It should influence the absolute values of the slow dynamics of G4. Further investigations are needed in order to quantify this. For the sake of simplicity we will treat the data as if there was only one fast heating, knowing that there is a slow one. Measurements with bromothymol as well as simulations have shown that there is no dependence of the heating dynamics on the initial temperature and $\Delta T$. Therefore, as we will see, the observed tendencies are real even though there rests an uncertainty about the absolute values.

The absorption curves in Fig. 5.2(upper graph) show dynamical changes provoked by a T-Jump of $\Delta T$ = 10°C. These dynamics were recorded at initial temperatures varying between 17°C and 73°C. At 17°C almost no change in absorption takes place, since the temperature is still far from the melting point $T_m$ = 62°C. In contrast, CD dynamics are already observed at this temperature. This illustrates that CD is more sensitive than absorption to the G4 denaturation. It is noteworthy that the CD curve at 17°C, which exhibits low noise has been averaged over 5120 heating-cooling cycles, while the other curves have been only averaged over 256 cycles.

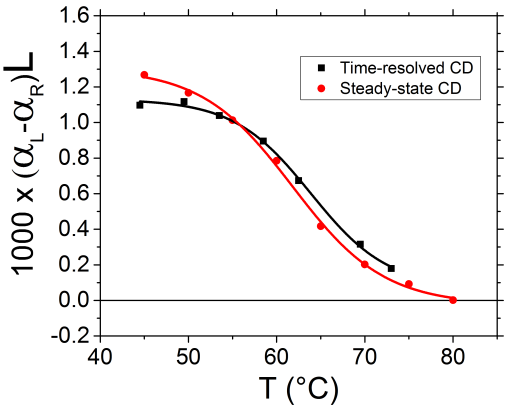


Figure 5.3: Comparison of TRCD signals prior to T-Jump (black dots) with the CD melting curve obtained at SOLEIL at 293 nm (red dots). Lines indicate fits.

A comparison of the TRCD signal before T-Jump revealed a similar shape than that of the static state melting curve measured at SOLEIL, as illustrated in Fig. 5.3. The initial values (black dots) are compared to the ones measured at SOLEIL. The figure shows that both static CD and TRCD measurements yield the same sigmoidal profile. Their first derivative are gaussian like. By going from an infinitesimal $dT$ to a larger $\Delta T$ as provided by the T-jumps, this profile should also be found in the amplitude changes induced by the T-Jump with temperature. The curves depicted in Fig. 5.4 represent the maximum absorption and changes for a given initial temperature $T_{in}$. They could be well fitted with a gaussian function with a maximum at 57°C.

A larger $\Delta T$ results in larger absorption and CD changes. Thus, by comparing these changes with those measured by steady state spectroscopy, one can derive $\Delta T$.

However, there is a more elegant and faster way to determine $\Delta T$. For this, we have to imagine at which temperature $T_{in}$ the maximum change of absorption and CD induced by $\Delta T$ should be expected. This occurs when the temperature in between $T_{in}$ and $T_{in} + \Delta T$ is the melting point $T_m$. Hence, $\Delta T$ can be determined as follows:

$$\Delta T = 2 \cdot (T_m - T_{in}[\Delta Abs(max) \; or \; CD(max)]) \tag{5.1}$$

Maximum changes of both absorption and CD are found at 57°C according to the gaussian fits in Fig. 5.4. This value is 5°C below the melting point ($T_m$ = 62°C). Therefore, $\Delta T$ for the presented curves is 10°C.

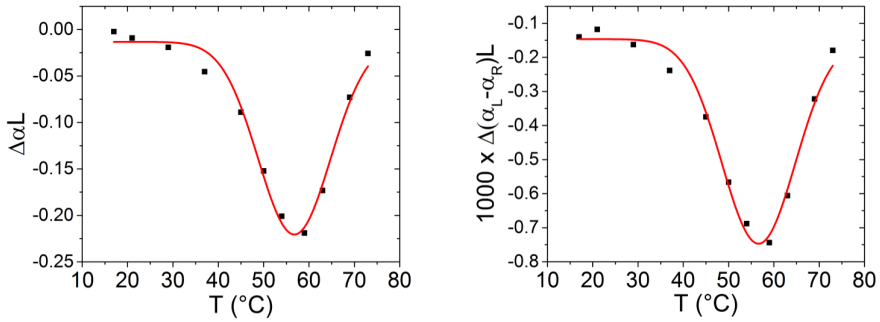


Figure 5.4: Absolute values of absorption (left graph) and CD changes (right graph). Both show a roughly gaussian-shaped profile (red fit function). Both curves revealed a maximum change at 57°C which corresponds to $\Delta T$ = 10°C.

Besides these results for a T-jump of 10°C, we have measured Tel21 dynamics for different $\Delta T$ at a given initial temperature of $T_{in}$ = 52°C. These measurements are shown in Fig. 5.5.

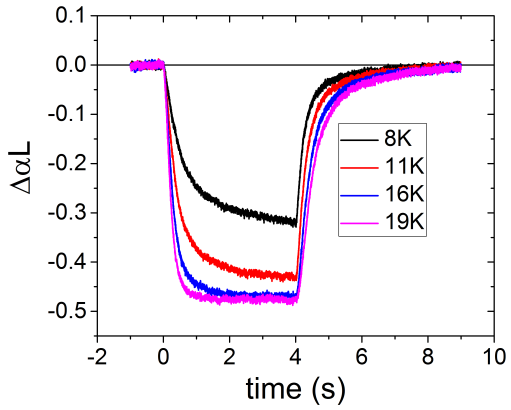


Figure 5.5: Absorption dynamics of Tel21 at $T_{in} = 52°C$ for different $\Delta T$.

One can clearly see that for small $\Delta T$ (black curve), denaturation is significantly slower than for large $\Delta T$. This is also true for renaturation. Actually, it is difficult to compare the denaturation kinetics because for denaturation the relevant temperature is $T_{in} + \Delta T$ and thus varies with $\Delta T$. In the case of renaturation the relevant temperature is $T_{in}$ and is thus the same for all measured kinetics. This feature will be discussed in the following. In addition to the studies at 293 nm, measurements at 265 nm have been performed. They revealed the same dynamics for denaturation and renaturation as at 293 nm. A comparison of absorption dynamics at $T_{in} = 50°C$ after a T-jump of $\Delta T = 10°C$ is shown in Fig. 5.6. A global fit with a biexponential function shows clearly that both dynamics are equal.

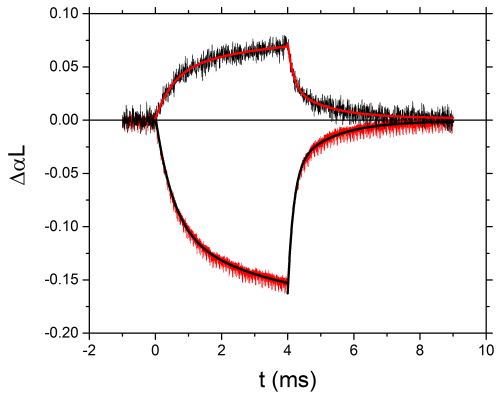


Figure 5.6: Comparison of absorption dynamics at $T_{in} = 50°C$ at 293 nm (black curve) and 265 nm (red curve) after a T-jump of $\Delta T = 10°C$. Global fits with shared time constants for dynamics at both wavelengths were performed with a biexponential fit function.

#### 5.2.1.1 Multiexponential fitting of results

Absorption and CD kinetics exhibit clearly a non-exponential behavior. We thus tried to fit them with a stretched exponential function and a sum of exponential functions. An example of these fits is highlighted by Fig. 5.7 of Tel21 denaturation at 59°C after a T-jump of 10°C. It provides a generic example of the dynamics measured for different initial temperatures, in absorption or in CD.

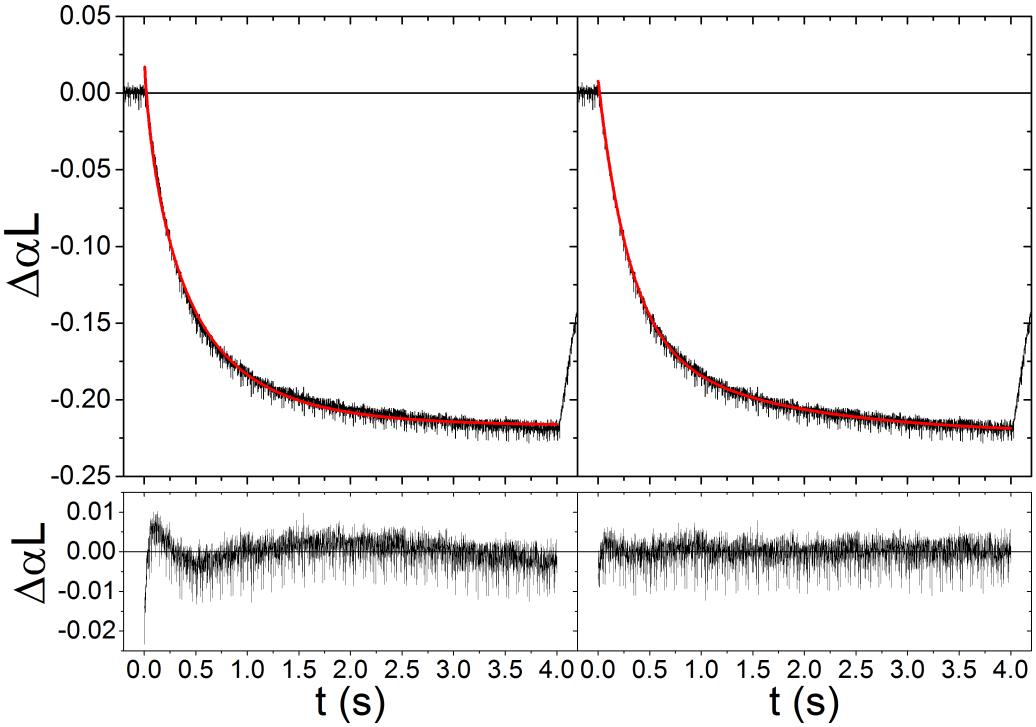


Figure 5.7: Comparison of the different used fit functions for Tel21 denaturation dynamics at $T_{in} = 59°C$ monitored with absorption. Left: stretched exponential fit (according to Eq. 5.2), right: biexponential fit (according to Eq. 5.3). The graphs below show the corresponding residue values.

The used fit function are described below:

$$y_\beta(t) = A \cdot e^{(-(t-t_0)/\tau)^\beta} + y_0 \tag{5.2}$$

$$y(t) = A_1 \cdot e^{-(t-t_0)/\tau_1} + A_2 \cdot e^{-(t-t_0)/\tau_2} + y_0 \tag{5.3}$$

where $A$ is the amplitude, $t_0$ and $y_0$ are the offset in time and absorption, respectively. $\tau$ describes the time constant and $\beta$ in Eq. 5.2 is the exponential factor, characteristic for streched exponential functions. Fig. 5.7 points out that the fit with the stretched

exponential function does not properly describe Tel21 dynamics (left graph). Contrarily, a sum of two exponential functions is found to fit properly the dynamics (right graphs).

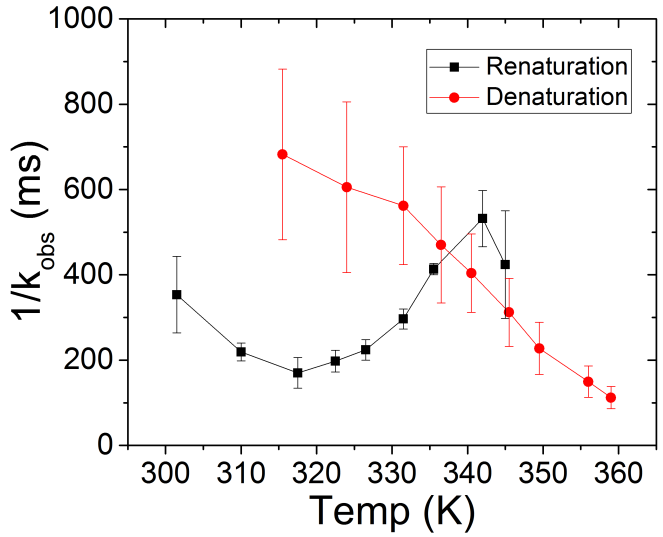


Figure 5.8: Extracted time constants $1/k_{1,obs} = \tau_{1,obs}$ of the curves shown in Fig. 5.2. Red dots and black dots are the values for denaturation and renaturation, respectively. The large error bars result from the interdependence of $1/k_{1,obs}$ and $1/k_{2,obs}$.

Biexponential fits of Tel21 dynamics yield two time constants $\tau_{1obs}$ and $\tau_{2,obs}$. Since absorption curves exhibit less noise, fitting was in general performed only on absorption curves. A global fit with shared time constants for absorption and CD curve did not decrease the error bars. Therefore, fitting was only performed on absorption dynamics. This is maintained throughout the work including Chapter 6. The plotted time constants in Fig 5.8 illustrate the faster ones ($\tau_{1obs}$). The second time constant $\tau_{2,obs}$ is in the range of around 2 s. In order to relate denaturation and renaturation dynamics, one has to keep in mind that for denaturation the final temperature is $T_{in} + \Delta T$ and for renaturation it is $T_{in}$. By respecting that, we found that for all different series the measured denaturation and renaturation curves intersect at $T_m$.

Biexponential decays of denaturation and renaturation of Tel21 reveal that a simple two-state mechanism is not valid. The question is to know whether this biexponential behavior arises from a multi-pathway scenario or a sequential one. Consequently, different data treatments are described in the following sections to get a better understanding of the Tel21 folding mechanism.

#### 5.2.1.2 Data Treatment of biexponential fitting

Since T-jump experiments rely on the return to equilibrium of a system after a rapid change of the temperature (heating or cooling), it can provide a measure of both the folding and unfolding rates. Under the crude assumption of a two-state equilibrium model, the observed relaxation times of denaturation and renaturation ($k_{obs}$) are the sum of the folding rate $k_f$ and the unfolding rate $k_u$. For each initial temperature, we can calculate those rates according to the following relationships:

$$k_{obs} = k_u + k_f \tag{5.4}$$

The equilibrium constant $K_{eq}$ is determined from the static CD melting curve (Chapter 4) with equation

$$K_{eq} = \frac{k_f}{k_u} = e^{(T^{-1} - T_m^{-1})\Delta H^0/R} \tag{5.5}$$

where $\Delta H^0$ is the enthalpy of the reaction and $T_m$ is the melting temperature. We can calculate $k_f$ and $k_u$, the rate constants for folding and unfolding, respectively by using:

$$k_{obs} = k_u \cdot (K_{eq} + 1) \tag{5.6}$$

thus

$$k_f = \frac{k_{obs}}{1/K_{eq} + 1} \tag{5.7}$$

and

$$k_u = \frac{k_{obs}}{K_{eq} + 1} \tag{5.8}$$

Rate constants $k_f$ and $k_u$ were then determined for each temperature with Eq. 5.7 and Eq. 5.8, respectively. Subsequently, the values of $ln(k(T))$ were plotted against $1/T$ to determine the activation energies of folding and unfolding. Such Arrhenius plots are shown in Fig. 5.9. This treatment could not be done with $k_{2obs}$, since it is too noisy. Arrhenius plots associated to $k_{1obs}$ reveal deviations from Arrhenius behavior.

The unfolding rate display an Arrhenius-like behavior associated with a positive activation energy of ca. 159 $kJ/mol$ (38.4 $kcal/mol$) above $1/T = 0.003\ K^{-1}$. In stark contrast, the folding rate displays a non-Arrhenius behavior. At low temperatures (between 0.0031 and 0.0035 $K^{-1}$) the folding reaction is barrierless. At temperatures above $T_m$ the folding reaction exhibits a negative activation energy $\Delta E$. Such trends have been found for all our measurements.

Furthermore, measurements at fixed $T_{in}$ with varying $\Delta T$ represented in Fig. 5.5 yield increasing time constants for renaturation as depicted in Fig. 5.10. This feature is in contradiction with a simple two-state model. In summary, the ensemble

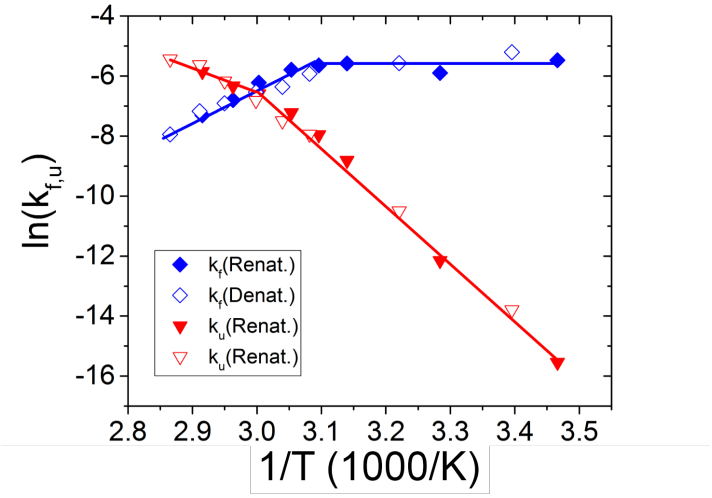


Figure 5.9: Arrhenius plot of $ln(k(T))$ against $1/T$. Values extracted from folding are depicted with filled dots, values from unfolding are shown with empty dots. Values of $k_f$ are blue, values of $k_u$ are red. Straight lines indicate tendencies.

of our results on ms time-scale show that a two-state model is not valid for the unfolding/refolding of Tel21 in agreement with recent literature on G4 [20].

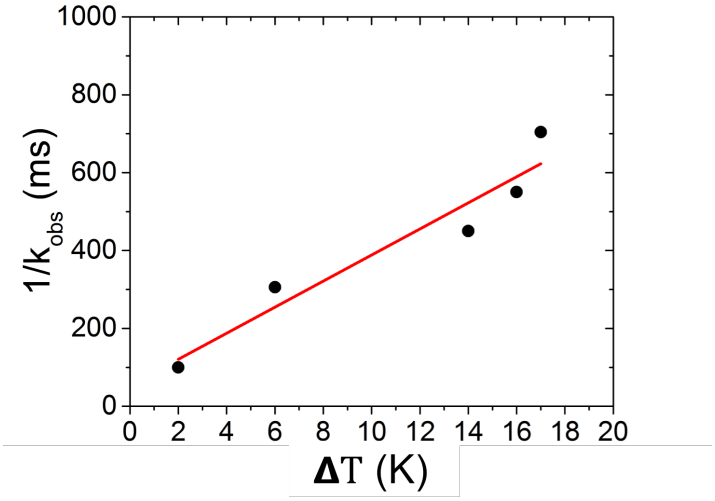


Figure 5.10: Extracted time constants $\tau_{1,obs}$ of the curves shown in Fig. 5.5, $T_{in} = 52°C$. A linear regression yielded $\tau_{1,obs}/\Delta T = 33 \pm 6 ms/K$

### 5.2.2 First model: 1D diffusional folding model

In order to obtain an indicative picture of the folding mechanism of Tel21, we have first used an empirical model consisting in a free diffusion along a one-dimensional free-energy landscape without significant barriers. This energy surface is meant to describe the balance between enthalpic effects favoring the folded structure and entropic effects favoring the unfolded one.

To develop an empirical folding model from our experimental data, some assumptions have been made. Folding is described with one effective folding coordinate N, accounting for the number of the guanines in the G4. N=1 and N=12 corresponds to the fully unfolded and folded structure, respectively. A fixed folding order has been assumed to simplify the problem, meaning folding happens from the 5' end consecutively to the 3'. The stabilizing effect of cation coordination has been neglected, since it partly compensates the repulsion of oxygen atoms, which is also not considered in our model. Therefore, in the present model, folding enthalpy comes from two contributions: Hoogsteen H-bond energy and the base stacking. Accordingly, the Gibbs free energy was determined by Eq. 5.9

$$\Delta G^0 = n_{hb}\Delta H^0_{hb} + n_{st}\Delta H^0_{st} - n_S T\Delta S^0 \tag{5.9}$$

where $n_{hb}$, $n_{st}$ and $n_S$ are the number of hydrogen-bonded guanines, the number of stacked guanines and the number of free guanines within the G4 scaffold, respectively. The last term accounts for the entropy gain of unfolded guanines. Corresponding parameters are listed in Table 5.1. $\Delta H^0_{hb}$, $\Delta H^0_{st}$ and $\Delta S^0$ describe the respective associated enthalpies and entropy. Evaluation of those values is made in order to reproduce the experimental fusion curve as depicted in Fig. 5.11 and yields a $\Delta H_{hb} = 3\ kJ/mol$ (0.7 $kcal/mol$), $\Delta H_{st} = 4.2\ kJ/mol$ (1.0 $kcal/mol$) and $\Delta S = 32.6\ J/mol$ (7.8 $cal/mol/K$).

With Eq. 5.9 free energy landscapes were calculated as a function of the folding parameter N in the temperature range of 270-370 K. They are depicted in Fig. 5.12. Note that our curves resemble very much to the energy curves found by Ansari *et al.* for DNA hairpins [188].

A consecutive step now is to see whether the dynamics of denaturation and renaturation of Tel21 can be reproduced qualitatively from that model. For this we have only considered $k_{1obs}$ for different initial temperatures (see Fig. 5.5). The temporal evolution of the non-equilibrated distribution of population, after a T-jump and consecutive cooling to lower temperature has been calculated according to a derivation

| N | # H-bonds: $n_{hb}$ | # stacking: $n_{st}$ | # entropy: $n_S$ |
|---|---|---|---|
| 1 | 0/0 | 0/0 | 12/10.0 |
| 2 | 0/0 | 1/1.1 | 10/9.5 |
| 3 | 0/0.2 | 2/1.7 | 9/9.0 |
| 4 | 2/1.8 | 2/2.3 | 8/8.1 |
| 5 | 4/4.0 | 3/3.3 | 7/7.0 |
| 6 | 6/6.2 | 4/3.6 | 6/6.0 |
| 7 | 8/8.0 | 4/4.4 | 5/5.0 |
| 8 | 10/9.8 | 5/5.0 | 4/4.0 |
| 9 | 12/12.4 | 6/5.7 | 4/4.0 |
| 10 | 16/15.8 | 6/6.2 | 2/2.0 |
| 11 | 20/20.0 | 7/6.9 | 1/1.0 |
| 12 | 24/24.0 | 8/8.0 | 0/0 |

Table 5.1: Values of $n_{hb}$, $n_{st}$ and $n_S$ as a function of the folding parameter N. The first value (on the left of /) has been estimated from the G4 structure by assuming that folding occurs sequentially from N=1 to N=12. Second values (on the right of /) were used for the calculations in order to obtain (i) a smoother variation of these parameters with N and (ii) a better adjustment of the experimental melting curve.

of the Smoluchowski diffusion equation proposed by Hamm *et al.* [189]. In their work, dynamics of a system containing many discrete microstates is described by the following rate equation:

$$\dot{\rho} = K \cdot \rho \tag{5.10}$$

where $\rho$ is the population density of a given structure and $K$ is a matrix containing the rate constants for all microstate transitions $k_i$ between two steps N and N+1 for both folding and unfolding. In our proposed one-dimensional case, the matrix K writes

$$K = \begin{pmatrix} -k_f^I & +k_u^I & & 0 \\ +k_f^I & -k_u^{II} - k_u^I & +k_u^{II} & \\ & +k_f^{II} & -k_f^{III} - k_u^{III} & \ddots \\ 0 & & \ddots & \ddots \end{pmatrix} \tag{5.11}$$

where $k_u^I$ describes the rate constant by going from N=1 to N=2. $k_f^I$ represents the rate constant from N=2 to N=1. Eq. 5.10 and Eq. 5.11 correspond to a discretized

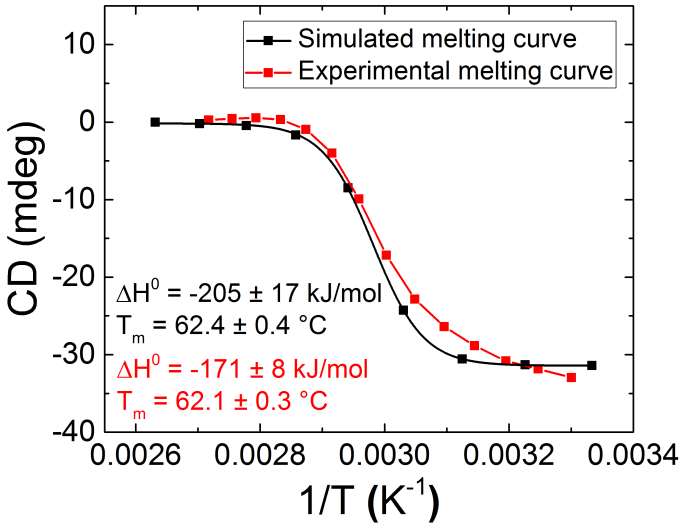


Figure 5.11: Comparison of simulated and experimental melting curves of Tel21. Comparable melting temperatures $T_m$ and enthalpies $\Delta H^0$ were obtained for both. Parameters are depicted in the graph.

version of a continuous Smoluchowski equation

$$\frac{\partial \rho}{\partial t} = D\Big[\frac{\partial}{\partial x}\Big(\frac{\partial G}{\partial x}\rho\Big) + \frac{\partial^2}{\partial x^2}\rho\Big] \tag{5.12}$$

The diffusion constant $D$ in Eq. 5.10 is connected to Eq. 5.11 by:

$$k_{(f)}^{initial} = k^{(0)} e^{-\frac{G_{i+1}-G_i}{2k_B T}} \tag{5.13}$$

$$k_{(f)}^{final} = k^{(0)} e^{+\frac{G_{i+1}-G_i}{2k_B T}} \tag{5.14}$$

with $k^{(0)} = D/\Delta x^2$. It is assumed that $D$ is independent of position $x$ in Eq. 5.12. We do not know the diffusion constant on that surface. Therefore, the $k^{(0)}$ are completely arbitrary and so are the $k$ values of Eq. 5.11. Eq. 5.10 is solved by Eq. 5.15.

$$\rho(t) = e^{Kt}\rho^{(0)} = Qe^{\Lambda t}Q^{-1}\rho^{(0)} \tag{5.15}$$

where $\Lambda = diag(\lambda_0, \lambda_1, \cdots, \lambda_{n-1})$ is the diagonal matrix of $K$ which contains its eigenvalues. $Q$ is a transformation matrix with the corresponding eigenvectors. The first eigenvalue $\lambda_0$ is equal to zero, due to particle conservation. It corresponds to the equilibrium distribution. The other eigenvalues are all negative [189]. In order to calculate the dynamics and to reproduce a denaturation (resp. renaturation) curve at a given temperature, one has to use the initial population at $T_{in}$ (resp. $T_{in} + \Delta T$) for $\rho_0$ and calculate the evolution of $\rho$ for the temperature $T_{in} + \Delta T$ (resp. $T_{in}$).

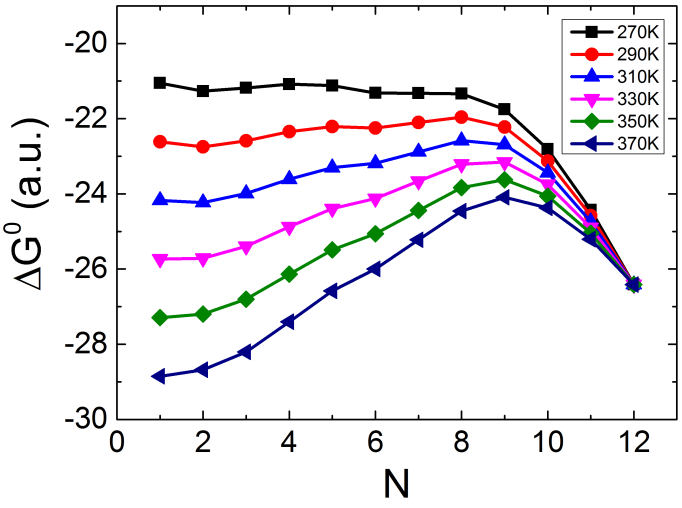


Figure 5.12: Calculated Gibbs free energy landscapes in arbitrary units as a function of folding parameter N for different temperatures between 270 K and 370 K.

Selected curves are shown in Fig. 5.13(left). The time scale being arbitrary, since we cannot determine $k^{(0)}$ quantitatively, the calculated curves can be compared only qualitatively to the experimental results.

In Fig. 5.13, denaturation is slower than renaturation at low temperature ( black, 47°C). With an initial temperature of 57°C and a $\Delta T = 10$°C the red curve reflects dynamics around $T_m$. Thus, as expected, denaturation and renaturation exhibit equal rate constants. Curves at $T_{in} > T_m$ show faster denaturation than renaturation (see blue curve), in agreement with the experiments. More importantly, the Arrhenius curves derived from the model, displayed in Fig. 5.13, behave very similarly to the experimental ones, with negative activation energy for the folding process and a change of slope around $T_m$ for both curves. This simple calculation suggests that there are no distinct intermediates in the folding/unfolding process in G4's in agreement with a downhill folding model.

At this point we have not discussed the second time constants. Additionally, the slow dynamics of G4 compared to DNA hairpins [188, 190, 191] can only be explained by adding a roughness to the energy landscape as it is more likely that the landscape contains a manifold of energy traps. Consequently, in a consecutive step, a different data treatment is introduced that allows statements of the ruggedness of the potential energy landscape.

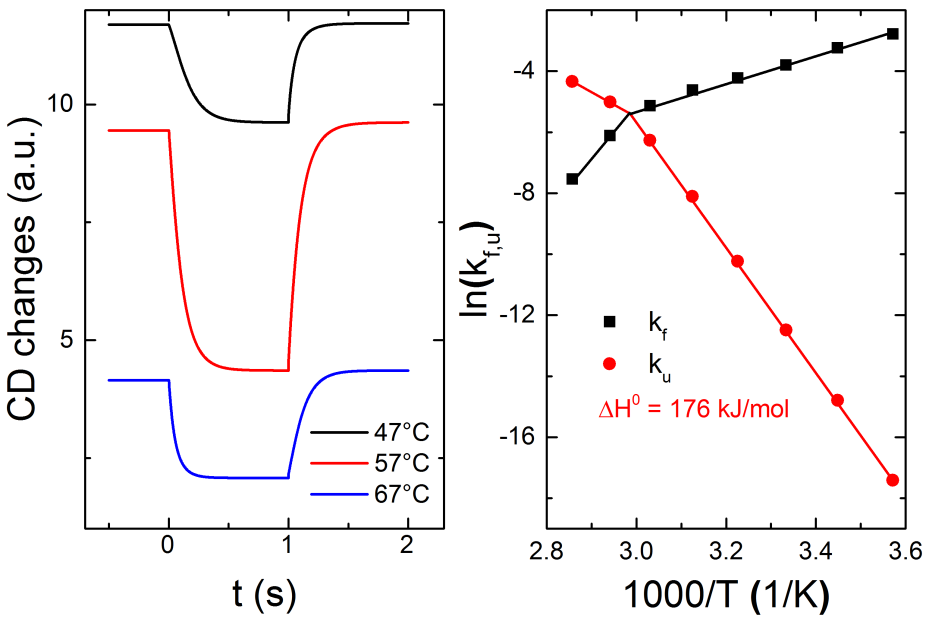


Figure 5.13: Left: Simulation of the time-dependence of CD changes of Tel21 after a T-jump ($\Delta T = 10°C$) according to Eq. 5.15. Curves for three different temperatures are shown. Right: Arrhenius plot for folding and unfolding rates determined from the simulated denaturation/renaturation dynamics.

### 5.2.3 Extension of our model: rugged free-energy folding landscapes

In the first model only $k_1$ was considered. Therein, only a part of the whole dynamical picture of G4 folding/unfolding can be obtained and additional information of longer times is neglected. In order to investigate longer times, we follow Volk's treatm ent and introduce an instantaneous, time-dependent rate constant, $k_{inst}$ [192]. As we will see, the experimental $k_{inst}$ are not constant, but decrease with time. To explain this feature we invoke the ruggedness of the energy landscape, since many traps can be present, corresponding to GG stacking, hydrogen-bond base pairs, haripin structures [142] or structures containing 1 or 2 G-quartets [20]. The magnitude of the energetic barriers of those traps are most likely in the order of several $k_BT$ and can be higher for misfolded G4.

#### 5.2.3.1 Data treatment with $k_{inst}$

We determine the instantaneous rate constant $k_{inst}(t)$ as Eq. 5.16 [192].

$$k_{inst}(t) = -\frac{dc(t)/dt}{c(t)} = -\frac{dA(t)/dt}{A(t)} = -\frac{dCD(t)/dt}{CD(t)} \tag{5.16}$$

where $c(t)$ is the remaining initial species concentration. $A(t)$ and $CD(t)$ are the corresponding absorbance and CD signals, respectively. For monoexponential dynamics, $k_{inst}(t)$ is constant.

Denaturation/renaturation dynamics occur because the initial G4 distribution is out-of-equilibrium with respect to the new conditions. We measure the time it takes to reach equilibrium and we suppose that the underlying dynamics result from the evolution of G4 on a rugged energy landscape. This landscape should be full of random traps and we suppose that each G4 molecule passes one trap before relaxing to equilibrium. With an activation energy $E_n$ of trap $n$ we obtain

$$k_n = k_0 \cdot e^{\frac{E_n}{k_B T}} \tag{5.17}$$

where $k_0$ is the maximum rate constant. This value $k_0$ would be measured if the energy landscape was without these traps. Subsequently, we suppose the energy traps distribution to follow Gaussian statistics. The probability is given by

$$G_n = N \cdot e^{-\frac{E_n^2}{2\sigma^2}} \tag{5.18}$$

where N is a normalization constant. This distribution supposes that the majority of traps has only a small barrier. We verified that by relaxing this restriction and the

fitting results did not change significantly. In Eq. 5.18 the standard deviation $\sigma$ is the most important factor, for it characterizes the width of the distribution of energy traps. Therefore, it yields a description of the landscape ruggedness. The measured signal is an accumulation of all these individual interactions. The probability for the time period ($t$) of an interaction is given by $p_n = e^{-k_n t}$. Hence, the transient signal ($A$ or $CD = S$) reflecting the population of G4 before equilibrium writes as

$$S(t) = \sum_n p_n(t) G_n \tag{5.19}$$

Thus, the instantaneous rate constant writes as

$$k_{inst}(t) = \int_0^{+\infty} k_n p_n(t) G_n dE_n / \int_0^{+\infty} p_n(t) G_n dE_n \tag{5.20}$$

For the sake of simplicity, we replaced the discrete summations by integrals, assuming a very large number of traps.

This formula is used to fit the experimental data. A smooth procedure with a progressively increasing time window was applied on the data ($\pm$ 10% of the central point) and $k_{inst}$ was calculated.

We will first focus on the experiment with varying $\Delta T$ with the same final temperature (see Fig. 5.5). The results are illustrated in Fig. 5.14.

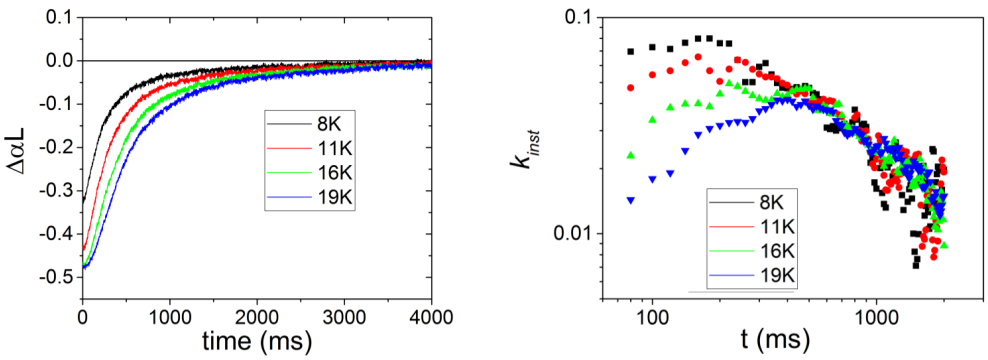


Figure 5.14: Left: Renaturation curves of different $\Delta T$ (see in-set). $T_{in} = 53°C$. Right: corresponding $k_{inst}$.

As this figure indicates, values of $k_{inst}$ are the same for each $\Delta T$. These four fits yield indeed equal results: $k_0 = 10\ s^{-1}$ and $\sigma = 1.4$. However, at short time-scales $k_{inst}$ is more erratic. This error results even in a decrease of $k_{inst}$ at short times with increasing $\Delta T$. By looking at the absorption curves it becomes clear that this is due to the shape of the curves. One sees a lag time at the beginning for large $\Delta T$. The

fact that all four curves yield the same values is as expected and in good agreement with our model. They all start at different temperatures but are out of equilibrium for the same final temperature ($T_{in} = 52°C$). Therefore, all four curves describe relaxation to an equilibrium on the same energy landscape. These results strongly indicate that data interpretation with an instantaneous rate constant is meaningful. It yields relevant information of the energy profile of our G4 samples. For longer times, the G4 are quite far from equilibrium where there are many energetic traps that slow down the relaxation process. Such a behavior has been observed by Volk *et al.* in peptide diffusion [192]. In that work, peptide motion was slowed down because of many trapping states. Concerning G4, it is somehow different, since we do not study diffusion in real space, but on an energy landscape. However, there is much similarity between these two diffusion models.

In a next step, we will look at the results with a constant $\Delta T$ but varying $T_{in}$ (see Fig. 5.2). Examples of the fits with $k_{inst}$ are depicted in Fig. 5.15.

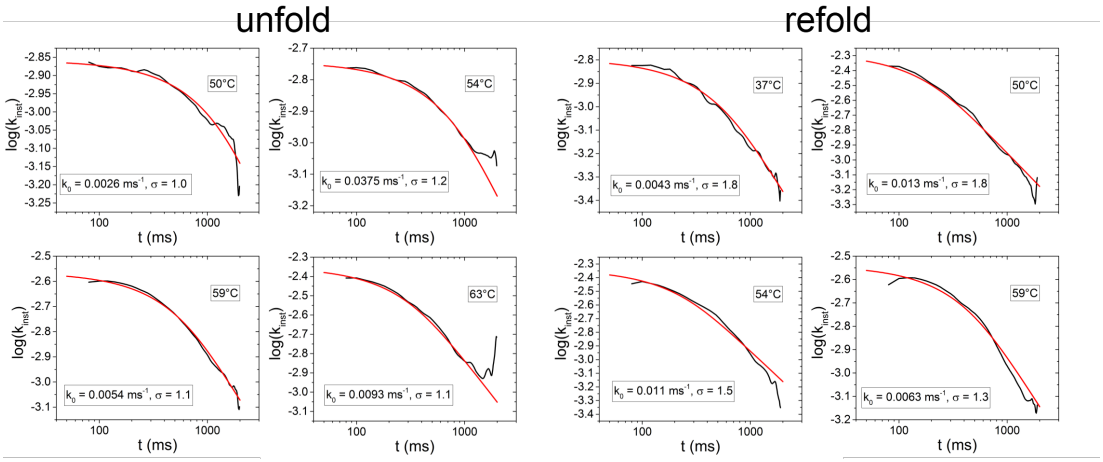


Figure 5.15: Selected $k_{inst}$ plotted against $t$. Indicated temperatures are the final ones: unfold: $T_{in} + \Delta T$, refold: $T_{in}$.

Curves for higher and lower $T_{in}$ are noisier, since the amplitude change is smaller. However, they yield reliable parameters with a larger uncertainty. Extracted values of $k_0$ and $\sigma$ are illustrated in Fig. 5.16.

Interestingly, for $k_0$ we observe the same trend as with the biexponential fit procedure for $k_1$. Equally, the corresponding Arrhenius plot yields the same information as before: No activation barrier between the unfolded and folded state and a non-Arrhenius behavior at higher temperatures. However, absolute values are different. Indeed, $k_0$ represents an ideal rate constant on a smooth energy landscape, which is not directly accessible in our experiments. Furthermore, folding and unfolding rates

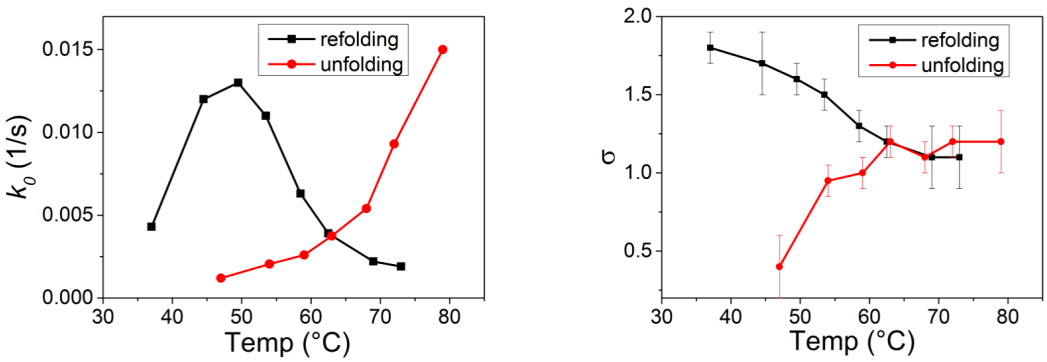


Figure 5.16: Values of extracted $k_0$ (left) and $\sigma$ (right) at different final temperatures.

are equal at $T_m$, as expected. They have a value of 4.8 $s^{-1}$ (= 1/(260 $ms$)). This is still too large to be a movement on a smooth energy landscape. It is probably the consequence of some shallow trapping states that are not included in our Gaussian trap distribution, but which appear like a viscous effect and result in a decrease of $k_0$.

A clearly different behavior of $\sigma$ is observed for denaturation and renaturation. This difference indicates again that a simple two-state model is not valid and shows that folding and unfolding do not follow the same pathway.

Additionally, the difference between denaturation and renaturation for a given final temperature is the initial population distribution. For denaturation the initial fraction of unfolded species is smaller than the final one. For renaturation it is the opposite. Therefore, denaturation is dominated by unfolding dynamics and renaturation by folding dynamics.

The completely folded G4 can only unfold, whereas partly unfolded G4 can fold either into the completely folded structure or in another partly unfolded G4. Therefore, the entropy is larger in the latter case and could explain the tendency at low temperatures that $\sigma$ is smaller for unfolding and larger for folding compared to the value at higher temperatures (Fig. 5.16 right). This is also supported by looking at the curves in Fig. 5.15. For unfolding curves we see clearly that it is flat at short times, whereas refolding curves of $k_{inst}$ show a larger time dependence for short times.

Šponer *et al.* pointed out that the slow dynamics could also result from movement between the energy wells of multiple well-folded G4 [20]. In the case of Tel21, this is possible. Therefore, we performed the same T-jump measurements with Tel22($Na^+$), which forms mainly one folded G4, in order to compare the dynamics and verify the aforementioned idea of Šponer *et al.*

### 5.2.4 Comparison with Tel22 with Sodium

The dynamics of Tel22 have been studied in 150 mM $Na^+$ phosphate buffer to be in the same conditions as Tel21.

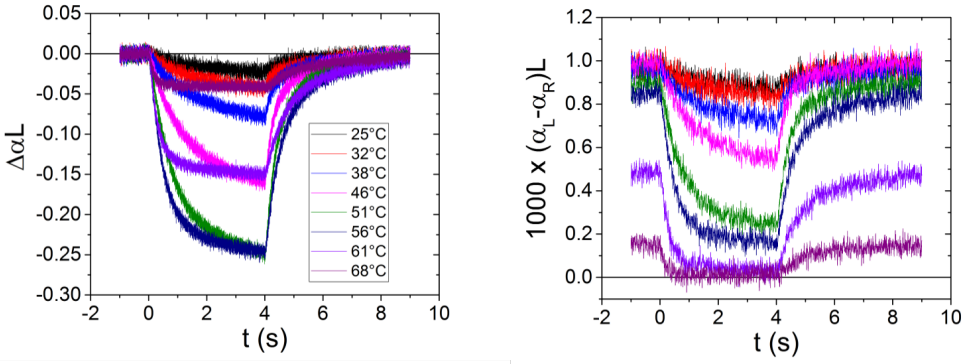


Figure 5.17: Absorption (left) and CD (right) dynamics of Tel22($Na^+$) at 293 nm at different initial temperatures (small box in the left graph) with a T-jump of $\Delta T = 10°C$. CD curves were smoothed with the Savitzky-Golay method (20pts). Denaturation occurred during the first 4 s, Renaturation between second 4 to 9.

Fig. 5.17 shows absorption and CD dynamics of Tel22. They exhibit the same dynamics for a given initial temperature. In analogy to Tel21 denaturation is slower than renaturation at low temperature (black curve). With increasing temperature this tendency changes and denaturation becomes faster than renaturation (violet curve).

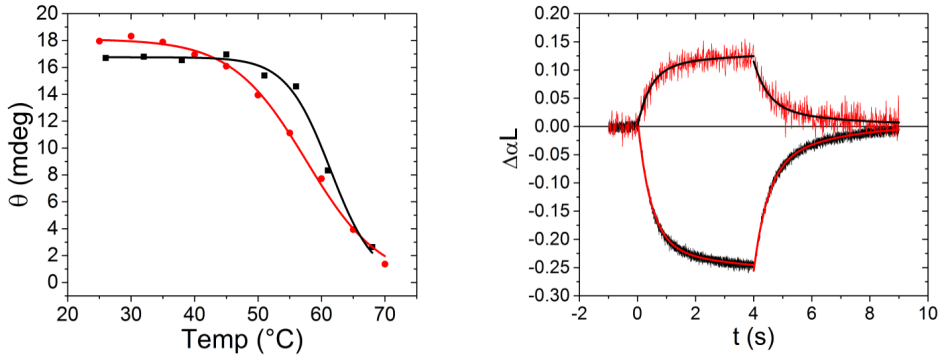


Figure 5.18: Left: Comparison of the average of the first 990 values prior to heating of dynamics measurements (black dots) with CD melting curve obtained at Soleil (red dots). Lines indicate fits. Right: Comparison of absorption dynamics at 293 nm (black) and 265 nm (red) at $T_{in} = 56°C$ with $\Delta T = 10°C$. Lines indicate global fits.

In addition to that, the initial CD values correspond to its melting profile (see

Fig. 5.18 left). Furthermore, a comparison of the dynamics of Tel22 at 293 nm and 265 nm illustrates that they are also equal at both wavelengths (see Fig. 5.18 right).

Finally, for a given $T_{in}$ different $\Delta T$ were implemented, yielding the curves depicted in Fig. 5.19. On the right, one clearly sees a linear dependence of $\tau = 1/k_1^{obs}$ with temperature. Compared to Tel21 in which a value of $\tau = 1/k_1^{obs} = 27 \pm 2\ ms/K$ was measured at $T_{in} = 52°C$, Tel22 exhibits a stronger temperature dependency. This might be simply due to the fact that $T_{in}$ is close to $T_m$ in the case of Tel22.

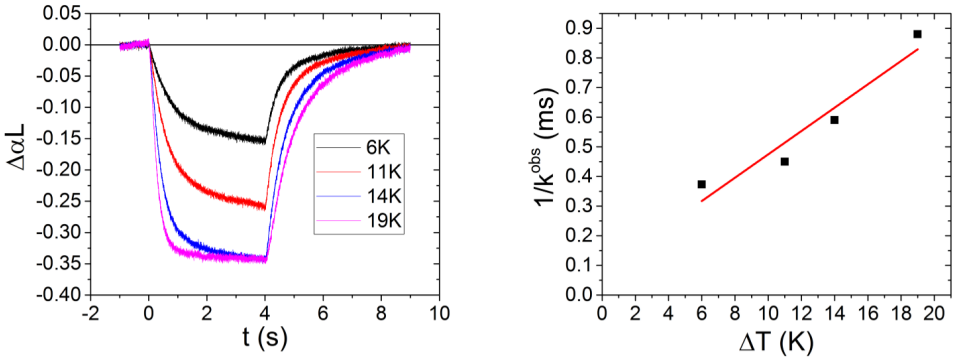


Figure 5.19: Left: Absorption dynamics of Tel22 at $T_{in} = 56°C$ for different $\Delta T$. It increases linearly with the pump power of the CW pulse. Right: First time constants $\tau_1^{denat} = 1/k_1^{denat}$ plotted against $\Delta T$. This yields a tendency of $\tau/\Delta T = 39 \pm 8 ms/K$.

The Arrhenius curves obtained with the extracted rate constants of folding and unfolding are shown in Fig. 5.20. The red triangles represent the unfolding and yielded a positive energy barrier as was found for Tel21($Na^+$). Renaturation dynamics decrease with temperature. Within the accuracy of our experiments, we can conclude that there is no difference between Tel21($Na^+$) and Tel22($Na^+$). A data treatment with $k_{inst}$ yield comparable $k_0$ and $\sigma$ values as depicted in Fig. 5.21. In both cases (Tel21 and Tel22), $k_0$ at $T_m$ is around 0.003 $ms^{-1}$. This is as expected, since in both cases a $\Delta T = 10°C$ was applied. The values of $\sigma$ are again equal for denaturation and renaturation after $T_m$. At lower temperatures, we have the same tendency as for Tel21: the ruggedness of the potential landscape that is experienced during denaturation is small and increases with temperature. For renaturation, it is the opposite. Therefore, this interpretation given previously accounts for both scenarios: mixture of folded G4 (Tel21) and a single initial G4 (Tel22). Even though the scenario of an interconversion between both folded G4 topologies in Tel21 cannot be excluded, it is more likely that other energy traps are responsible for the slow dynamics in both human telomeric sequences.

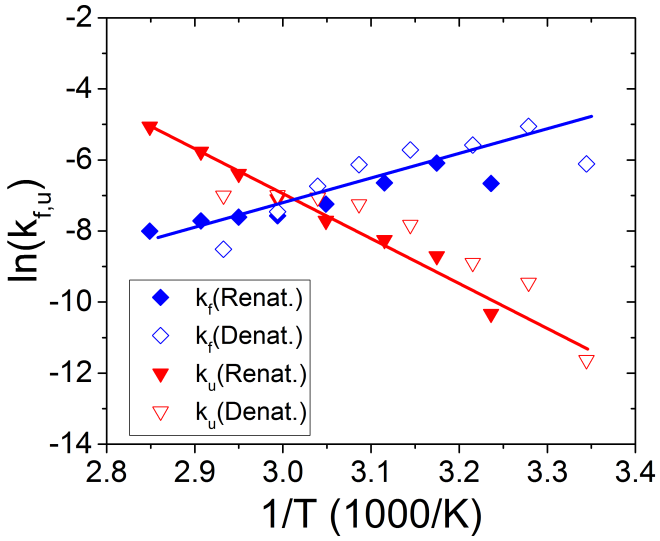


Figure 5.20: Arrhenius plot of $ln(k(T))$ against $1/T$. Values extracted from folding are depicted with filled dots, values from unfolding are shown with empty dots. Values of $k_f$ are blue, values of $k_u$ are red. Straight lines indicate tendencies.

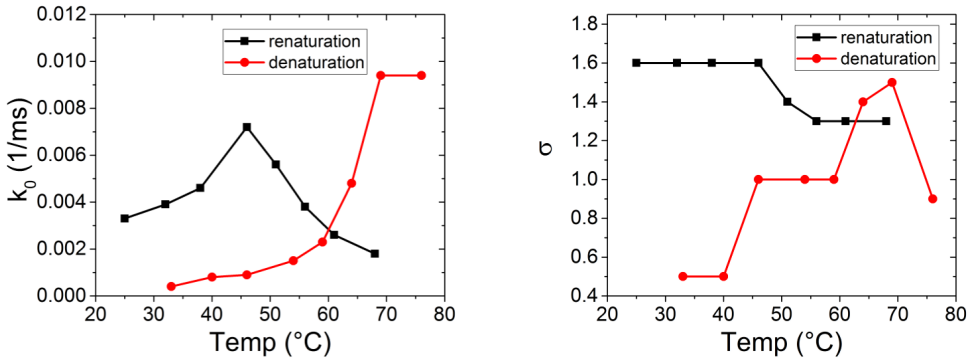


Figure 5.21: Values of extracted $k_0$ (left) and $\sigma$ (right) at different final temperatures. On the right, error bars are not indicated, but they are in the same order as for Tel21.

### 5.2.5 Conclusion

Whatever the data treatments, we found non-Arrhenius behavior for refolding. This is a strong indication of a multi-pathway folding process independent of the treatment procedure. At this point it is interesting to compare our results with the works published in literature. Most of the works on human telomeric G4 dynamics investigated G4 with $K^+$ [2, 20, 21, 23, 131] but a few groups have also performed studies on G4 with $Na^+$ [24, 122, 124]. We will now examine our results in view of the conclusions drawn by these previous works.

The group of Chaires performed a cation jump experiment on Tel22 with $Na^+$. They used UV changes of the peak at 295 nm to monitor folding. Interestingly, after cation injection, they also observed complex dynamics, that could not be fitted with a monoexponential function. They clearly obtained multiexponential dynamics, that could not be sufficiently fitted even by a triexponential function. The fastest one was in a couple of hundreds of milliseconds, whereas the second dynamic took 20-25 s to complete. The order of magnitude of these dynamics is in agreement with our results. Since our time window is not that large, we do not need a third time constant to fit. Our times are also in agreement with another work on Tel22 with $Na^+$ in which formation is induced by rapid-mixing and measured with FRET [124]. Therein, the found a fast folding time constant in the hundreds of ms range and a slower one of around 2 $s$. Gray and Chaires interpret their results as a sequential G4 forming with two well-defined intermediates.

Alternatively, one could claim that the second time constants are due to parallel unfolding/refolding routes. This is probably not due to the initial mixture of G4, since their dynamics should be comparable and not as different as $\tau_1$ and $\tau_2$. It could be that a fraction of folded G4 reaches a deep energy well of a misfolded structure during unfolding. That scenario explains the longer time constants of renaturation, since this fraction would have to overcome a large energy barrier in the case of renaturation. However, it does not explain the long denaturation times, since the initial structures are well-folded G4, assured by proper annealing of the sample prior to dynamics measurements.

A consecutive work on Tel22 with $K^+$ TRCD with stopped flow of the group of Chaires showed that there can be some early, probably misfolded structures [23]. The group of Gabelica showed that there are indeed misfolded structures with $K^+$ consisting of antiparallel 2-layer G4, which can explain the observation of the group of Chaires. The conversion, however, was in the time range of minutes to hours. It is true, that those misfolded structures cannot be distinguished by absolute CD

signals in our case, since they are all antiparallel and have similar CD spectra. Its dynamics should reveal potential intermediates/misfolded structures, though. There is a slow event in all of our dynamics in the order of several seconds. It cannot be fitted accurately due to our small time window. At any rate, it is not in the order of minutes. In contrast to stopped-flow experiments in which there is no cation at the beginning, the G4 structures in our experiments always see an excess of cation. Therefore, it is possible that those misfolded structures form preferentially out of an ensemble of random coils and not when cations are already present and the injection of cations could change the energy landscape tremendously. Thus, it is not surprising that their observed dynamics are quite slow compared to ours. Misfolded two-layer quartets have only been directly observed with $K^+$ [21]. It is possible that they are not stable enough with $Na^+$. The group of Birkedal, however, observed different interconvertable FRET efficiency peaks of Tel22($Na^+$) in a single molecule FRET study [24]. One of those has been attributed to a two-layer G4. Additionally, they show that their results could be explained by a sequential multi-pathway folding and an interconversion takes place during their time window of 70 s. This is surprising, since it has been clearly shown by NMR, that Tel22($Na^+$) forms a sole structure in aqueous solution [1, 45]. The reason for that might be the attachment to a DNA double strand and the FRET chromophores. The conditions of our experiments are as in NMR experiments and our dynamics do not show thermodynamically sufficiently stable intermediates to be detected.

Noer *et al.* further state that early dynamics could include triplex formation, but this is not accessible due to their time resolution of 200 ms. In our case, we also have a dead time due to heating that is around 50 ms. However, if a triplex structure forms for Tel21($Na^+$) and/or Tel22($Na^+$) in the early unfolding events, it should be visible, since we studied dynamics at two different wavelengths. As seen in Chapter 4 the stable TBA triplex structure [7] shows a clearly different CD spectrum than the one of TBA G4. This suggests that triplex CD spectra should be different from quadruplex CD spectra. In our case, we would obtain a difference of CD dynamics between the curves at 265 nm and 295 nm, since it is highly unlikely that a triplex intermediate has the same CD peak ratio of those two wavelengths as the corresponding G4 structure. This is supported by remembering the different origin of the peaks at 293 nm, which is due to GG dipole coupling between different G-quartets, whereas the peak at 265 nm results from GG coupling within a G-quartet. This is no direct proof that there is no triplex intermediate in the folding pathway of human telomeric G4, but this highly excludes it. Coherently, a multi-pathway folding model does not include one well-defined intermediate, either.

Most results in literature suggest that G4 folding undergoes multi-pathway folding on a rugged potential landscape [20]. Additionally, derived mostly from stopped-flow experiments, this landscape contains many deep energy traps of misfolded structures [2, 20, 22, 23, 133, 137]. Concerning our results, there are three possibilities: i) Intermediates/misfolded structures do not exist for Tel21 with $Na^+$ and Tel22 with $Na^+$, ii) in our measurements, the G4 are never far enough from the energy well of the well-folded structure to see misfolded structures, iii) the biexponential fit is meaningful and the longer time constants do describe these dynamics. This leads to the representation of a potential folding landscape of Fig. 5.22. The third possibility (iii) is unlikely, since there is no reason why these slow dynamics should be also visible during denaturation. This could be possible, if folded and misfolded structure were in an equilibrium. However, NMR shows that there is only one dominant G4 topology for Tel22 with $Na^+$ [1, 45]. Structural conversions on the second time-scale should be visible in NMR by peak broadening. Therefore, the picture that can be derived from our data only includes the deep energy well in Fig 5.22. It would be interesting to see, if Tel21 and/or Tel22 form stable 2-layer quadruplexes in $Na^+$.

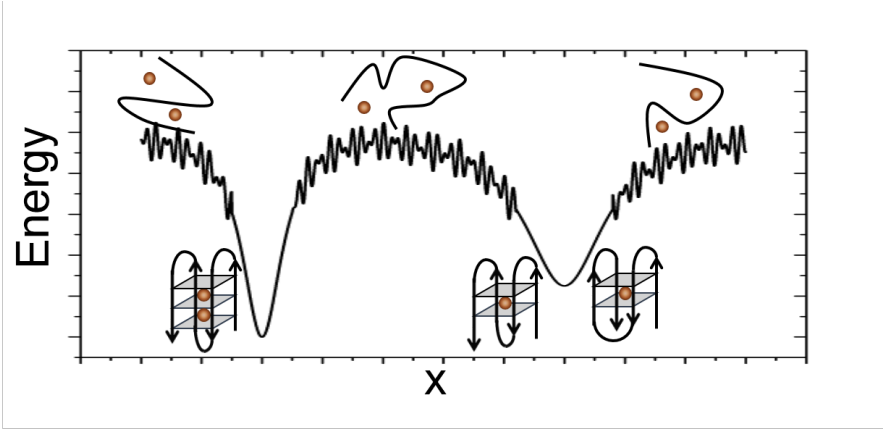


Figure 5.22: Schematic representation of potential landscape with one coordinate x. Deep wells indicate folded and misfolded G4 structures. Small wells (rough landscape) are energy traps

Since at low temperatures, denaturation has a ruggedness with a $\sigma$ close to 0, there are no significant energy traps depicted around the minimum of the well. The number of traps increases constantly above 50°C for Tel21($Na^+$), meaning the structure gets more and more unfolded and thus has more degrees of freedom that lead to energetic traps. One could imagine stacked $GG$ pairs as energetic traps. Mak calculated it to be $-12.5\ kJ/mol$ per pair [193]. Also, H-bonded $GG$ pairs are possible. Additionally, one layer G4 could be another reason for energetic traps. Also two-layer quadruplexes

that are close to the well-folded configuration but form a trap by sliding one of its arms [24, 122, 194].

Unfolding and folding clearly show multi-pathway behavior. However, some speculations are made about how those pathways could look like. The driving force of G4 formation is clearly the G-quartet [20]. Therefore, it is possible to imagine that the first step during unfolding is the destruction of a quartet. For that, many possibilities exist to remove one guanine or more. It is thus in accordance with a multi-pathway unfolding and those structures are close to the folded one and are thermodynamically not stable, since they would directly fold into the completely folded G4. In a consecutive step, the second quartet could be destroyed. Since monolayer quadruplexes are probably no stable enough to have long life-times, they would rapidly unfold.

Besides those energy traps, we still observe very small values for $k_0$. Folding dynamics of small proteins are a 100 to 1000-fold faster than DNA hairpins [191]. Wallace *et al.* found DNA stacking to be the main reason, since it increases the ruggedness of the energy landscape [191]. Therefore, it could be that there is an underlying additional ruggedness in G4 that cannot be extracted with our data treatment. In order to better understand the nature of the energy traps and the folding/unfolding dynamics in general, we investigated the effects of cation concentration on G4 dynamics.

## 5.3 Effect of Sodium concentrations on dynamics

Since cations are crucial for G4 formation, different cations types and concentrations probably lead to a different energy folding landscape. Gray and Chaires showed that Tel22($Na^+$) is completely folded at cation concentrations above 25 mM [122] and we expect similar behavior for Tel21($Na^+$). In order to gain save information on the effect of cations, we performed several dynamical studies on Tel21 in 25 mM sodium and in 900 mM. G4 stability increases with cation concentrations as has been seen in the last chapter by an increase of $T_m$: $T_m(25\ mM) \approx 47°C$, $T_m(150\ mM) \approx 60°C$, $T_m(900\ mM) \approx 75°C$.

### 5.3.1 Tel21 in 25 mM sodium solution

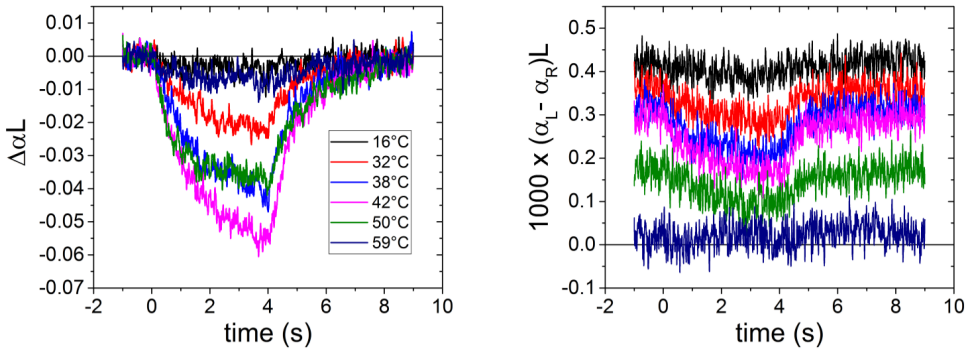


Figure 5.23: Left: Selected absorption curves at different temperatures. Right: selected CD curves. Curves were smoothed with the Savitzky-Golay method (40pts). Curves were recorded at 293 nm. Color codes are equal for both graphs. Heating started at time 0 and heating laser was stopped at t = 4 s. $\Delta T = 9°$.

The buffer solution was prepared with 8.6 mM $Na_2HPO_4$ and 8.6 mM $NaH_2PO_4$ in order to obtain a $Na^+$ concentration of 25 mM. Selected curves of absorption and CD dynamics of Tel21 are depicted in Fig. 5.23. Note that DNA strand concentration is decreased by a factor of 2 compared to the previous series. Therefore, absorption and CD signal are reduced by a factor of 2. Consistent with the previous curves denaturation becomes faster with temperature. However, for renaturation, dynamics appear less dependent on temperature. This is illustrated in Fig. 5.24(left). Since the absorption as well as CD dynamics are less complex compared to those of 150 mM, it was sufficient to use only one exponential function to fit the data. Therefore, the depicted time constants are in the time-range of seconds.

The Arrhenius-plot in the right graph of Fig. 5.24 shows a negative slope for

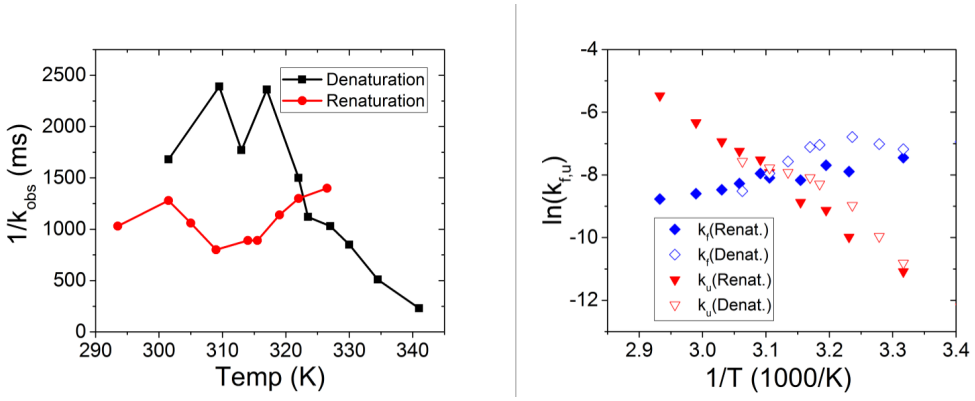


Figure 5.24: Left: Time constants $\tau_1 = 1/k_{obs}$ plotted against T. Right: Arrhenius plot of $ln(k(T))$ against $1/T$. Values extracted from folding are depicted with filled dots, values from unfolding are shown with empty dots. Values of $k_f$ are blue, values of $k_u$ are red.

unfolding rates, corresponding to an energy barrier of $E_A = 114 \pm 6\ kJ/mol$. This value is similar to the one obtained for Tel21 with 150 mM $Na^+$. A slight decrease of refolding rates with temperature is observed. However, those rates do not depend as much on the temperature as for Tel21 and Tel22 with 150 mM $Na^+$. Data treatment with $k_{inst}$ yields the results depicted in Fig 5.25. The right graph shows that $\sigma$ is constant for both denaturation and renaturation at around $\sigma \approx 1.2$. However, the behavior of $\sigma$ at low temperatures is not very clear. But it may be due to the large uncertainty of the measurements which results from the very weak change in absorption and CD (Fig. 5.24) The reduced complexity of the dynamics that are close to a monoexponential behavior illustrates that there are fewer energy traps at low temperature for renaturation compared to Tel21 with 150 mM $Na^+$.

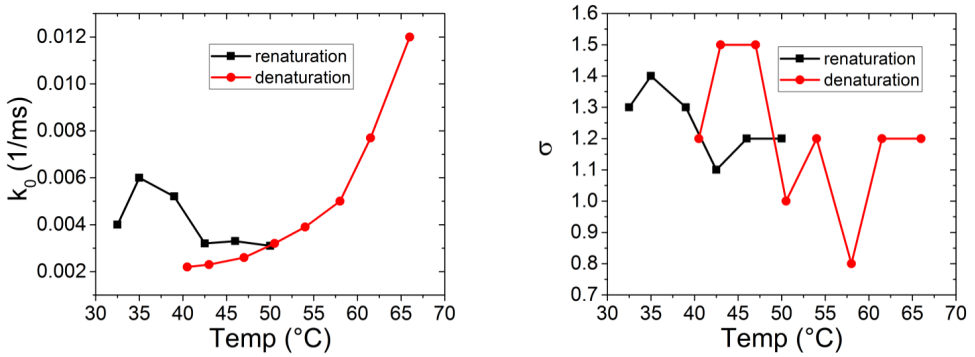


Figure 5.25: Values of extracted $k_0$ (left) and $\sigma$ (right) at different final temperatures.

Those findings indicate that a multi-pathway folding on a potential landscape is

influenced by traps that are stabilized by cations. In order to verify the tendency that some traps are stabilized with $Na^+$, we performed consecutive measurements in 900 mM cation solution.

### 5.3.2 Tel21 in 900 mM sodium solution

The denaturation and renaturation dynamics of Tel21 in a 900 mM $Na^+$ phosphate solution are depicted in Fig. 5.26. Denaturation and renaturation dynamics accelerate and slow down with temperature, respectively.

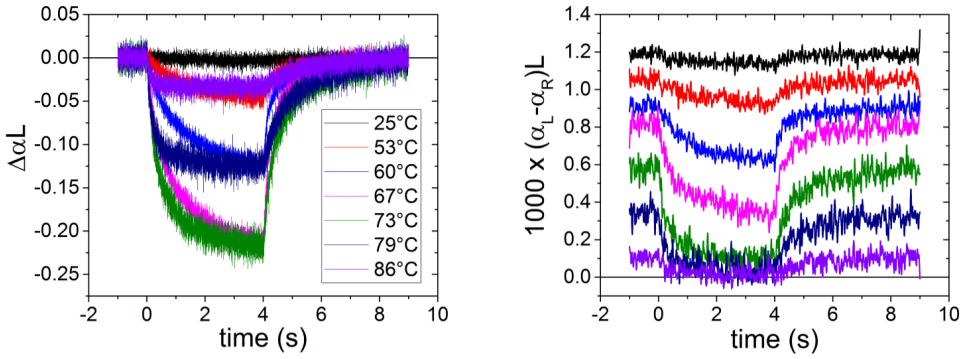


Figure 5.26: Left: Selected absorption curves at different temperatures. Left: selected CD curves. CD curves were smoothed with the Savitzky-Golay method: 60pts. Curves were recorded at 293 nm. Color codes are equal for both graphs. Heating started at time 0 and heating laser was stopped at t = 4 s. $\Delta T = 10°C$.

Additionally, the curves at 900 mM exhibit a clear biexponential behavior. It is more pronounced than at 150 mM, in a way that we could extract the long time constants ($\tau_2$) that show the same tendency as $\tau_1$ (see Fig. 5.27). The time constants $\tau_1$ are in the order of magnitude as for Tel21 and Tel22 in 150 mM $Na^+$ concentration. The middle graph of this figure shows again a activation barrier for unfolding of $E_A = 86 \pm 17\ kJ/mol$. This is also similar to the ones observed before. Correspondingly, the second time rates yielded a similar value $E_A = 103 \pm 17\ kJ/mol$ (see right graph). However, fast and slow time constants show similar temperature dependencies, which is more pronounced compared to Tel21 in 25 mM and 150 mM $Na^+$ concentration. Fig. 5.28 depicts the extracted parameters from a data treatment with $k_{inst}$. The right graph shows larger values of $\sigma$ for renaturation as was observed for Tel21 with 150 mM $Na^+$ solution. Hence, there is clearly a tendency of the number of energy traps (or their depths) to increase with increasing cation concentration. Interestingly, at all three concentrations, we find similar values of $k_0$. This means that it is not the shape of the potential energy landscape which changes, but its ruggedness. Therefore,

we can conclude that energy traps are stabilized by cations. This is not surprising, since G-quartets need cations to stabilize and so does many thinkable energy traps of G4, as stacked GG pairs, G-hairpins, H-bound GG base pairs, etc.

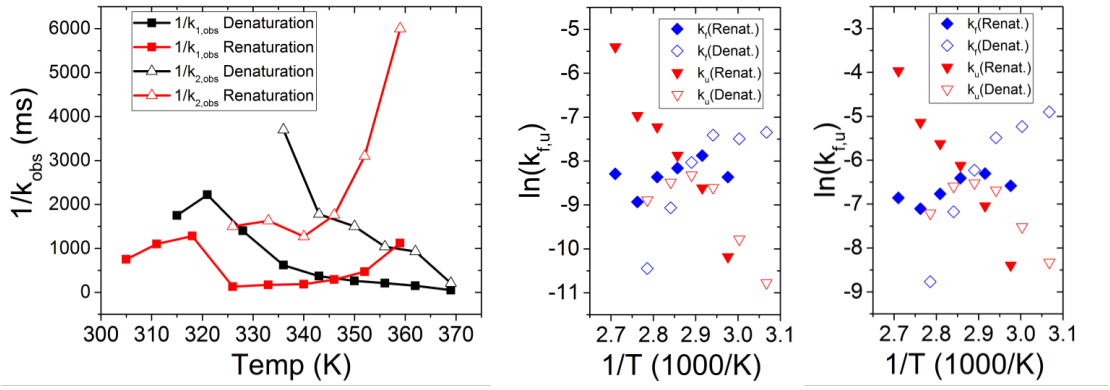


Figure 5.27: Left: time constants $\tau = 1/k_{obs}$ plotted against T. Middle and right: Arrhenius plot of $ln(k(T))$ against $1/T$ of $k_1$ and $k_2$, respectively. Values extracted from folding are depicted with filled dots, values from unfolding are shown with empty dots. Values of $k_f$ are blue, values of $k_u$ are red.

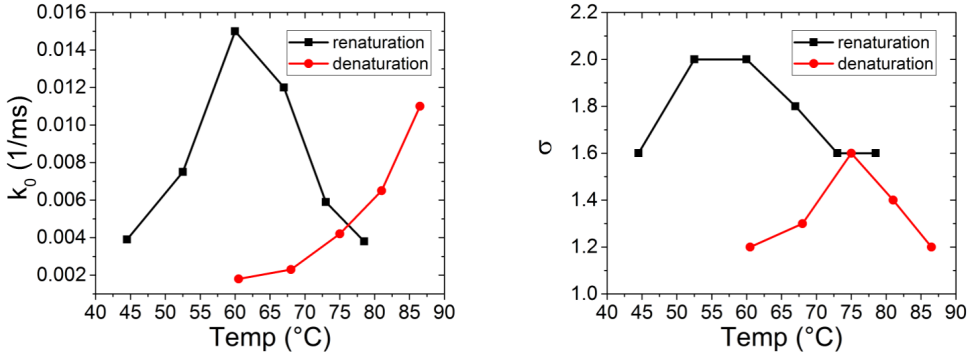


Figure 5.28: Values of extracted $k_0$ (left) and $\sigma$ (right) at different final temperatures.

It is interesting to notice that the amplitude of the longer times during the fit increases with increasing cation concentration. In the case of 900 mM, the amplitudes of both exponentials are almost equal. Additionally, we do not see a significant effect on denaturation and renaturation rate constants by varying the cation concentration.

In the conclusion of Tel21 in 150 mM $Na^+$ concentration, we have discussed the idea of a biexponential behavior, in which the long time constants result from a misfolded structure. This cannot be excluded and further investigations on other structures are necessary.

## 5.4 Folding dynamics of Tel21 on microsecond time-scale

In the perspective of G-quartets to be the driving force of G4 folding [20], triplexes are highly improbable. Nevertheless, T-jump experiments have been performed with a ns heating laser (compare Sec. 3.3.2) in order to explore $\mu s$ up to the first ms of unfolding to see whether early dynamics take place. Up to now, it was not possible for us to obtain TRCD curves of G4 below ms time resolution. Typical T-jumps of ca. 5°C induce a maximum CD change of $(\alpha_L - \alpha_R)L \approx 0.0003$. Additionally, since CD dynamics of Tel21 were found to be in the hundreds of millisecond time range, the observable change in $ns$ T-jump experiments is even smaller. Therefore, we performed it only for absorption.

Except for a different heating source, the same setup was used for the ns laser pulse T-jump as for the CW laser diode T-jump experiments. Samples were taken from the same solution of Tel21 that was used for ms experiments. This allows us to directly compare the ms and $\mu$s results.

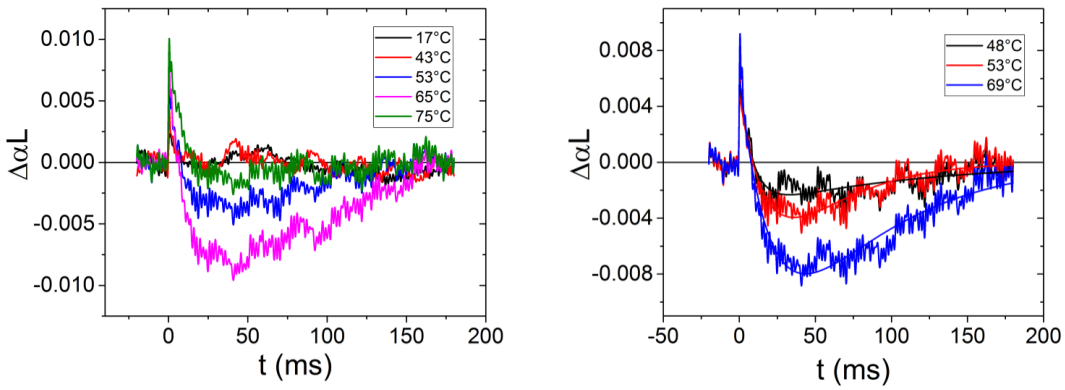


Figure 5.29: Left: Selected curves of denaturation/renaturation dynamics of Tel21 at 293 nm initiated by a ns T-jump ($\Delta T$=7K). FWHM(probe) = 80 $\mu$m, FWHM(pump) = 300 $\mu$m, quartz cuvette, l = 100 $\mu$m. Right: Selected curves with fit according to Eq. 5.21.

In Fig. 5.29 selected curves of a ns T-jump series are depicted. There is a fast increase of absorption, followed by a slow decrease and consecutive increase back to the initial absorption value. These results are counter-intuitive at first glance. Absorption should decrease during unfolding. After 1 ms the cooling become significant and after around 50 ms the whole sample has cooled down to its initial temperature. However, there are some slow dynamics.

It seems that a fast heating laser triggers some G4 denaturation that slows down with cooling sample until it renaturates. In order to fit the data some assumptions have been made leading to the fit function described below.

#### 5.4.0.1 Fitting function

We fitted our data according to Fig. 5.21. (For the full derivation see SI 7).

$$Abs = -K \cdot \left( e^{-\gamma(t-t_0)} - e^{\Gamma(t-t_0)} \right) + y_0 \cdot e^{\Gamma(t-t_0)} \tag{5.21}$$

where $K$ is the amplitude of two exponential functions. The first term is related to the renaturation in which $\gamma$ describes the time constant of renaturation. The second term is related to the cooling (opposite sign of exponential function), its time constant is $\Gamma$. It actually also contains denaturation, but this is neglected since it is significantly slower with respect to the cooling and thus the cooling dynamics are dominant. The last term $y_0 \cdot e^{\Gamma(t-t_0)}$ is used to describe the fast heating and slow cooling of the initial peak in Fig. 5.29 which we think is an artifact, possibly related to some unfolded random coil movement, since it did not occur during each experiment. Further studies are necessary to better understand this rising peak.

In Fig. 5.29 the right graph shows selected curves of the same series as in the left graph with fitted functions according to Eq. 5.21. One can see that the fits work quite well for all three selected curves. All different temperatures of this series are listed in Table 5.2. Curves at temperatures without slow dynamics (17°C-43°C, 82°C) were fitted with a simple exponential. For the other cases, the cooling time constant was fixed to $1/\Gamma = 10\ ms$. The revealed initial peaks $y_0$ show an increasing tendency that is similar to the $K$ values. The latter ones describe the amplitude of the theoretical absorption change without cooling. Both ($y_0$ and $K$) increase with temperature before reaching a plateau at around 70°C ($y_0$) and $T_m$ = 61°C ($K$), respectively. This difference indicate that $K$ is strongly related to the G4 denaturation, whereas $y_0$ is less related.

The denaturation constants derived from these fits show, however, no clear tendency. By ignoring the first determined value at 48°C a tendency to increasing time constants ($1/\gamma$) is observable. Due to the large noise, an uncertainty about the accuracy of the determined values remain. However, the observed renaturation times in the order of a 100 ms are in agreement with the ms experiments for $\Delta T$ of a few degrees (see Fig. 5.10).

| T (°C) | $y_0$ ($10^{-3}\Delta\alpha$L) | $1/\Gamma$ (ms) | K ($10^{-3}\Delta\alpha$L) | $1/\gamma$ (ms) |
|---|---|---|---|---|
| 17 | 2±1 | 4.4±0.1 | - | - |
| 27 | 3±1 | 8.5±0.1 | - | - |
| 32 | 4±1 | 6.3±0.1 | - | - |
| 43 | 4±1 | - | - | - |
| 48 | 5±1 | 10 (fixed) | 3.7±0.1 | 100±2 |
| 53 | 5±1 | 10 (fixed) | 7.5±0.1 | 60.0±0.5 |
| 59 | 6±1 | 10 (fixed) | 11.5±0.1 | 91.0±0.5 |
| 65 | 8±1 | 10 (fixed) | 13.9±0.1 | 81.0±1 |
| 69 | 9±1 | 10 (fixed) | 11.7±0.1 | 104.4±0.5 |
| 75 | 10±1 | 10 (fixed) | 11.7±0.1 | 104±1 |
| 82 | 9.2±1 | 11.6±0.1 | - | - |

Table 5.2

### 5.4.1 Conclusion

There is no fast G4 unfolding dynamics in the $\mu s$ time-range to 1 ms for Tel21 in 150 mM $Na^+$ buffer. Therefore, it does not support the idea of a fast-forming triplex intermediate during denaturation. In comparison to the results of the group of García, which performed a T-jump experiment on Tel22 with $Na^+$ [195], we do not see any dynamics during the first 100 $\mu s$, except the quasi-instantaneous rising peak. They attributed their observed change to a triplex structure. However, in all of our experiments on $\mu s$ time-scale, we have never found a similar behavior. Actually, no changes except the aforementioned peak take place at room temperature in our experiments in contrast to theirs. Therefore, we cannot confirm their claim.

The observed times for renaturation are in agreement with those found with continuous heating. In comparison of the results with continuous heating, we see dynamics already at low temperatures, we do not see any denaturation below 43°C here. This might be due to the small $\Delta T$, which also explains the fast renaturation dynamics.

# Chapter 6

# Dynamics of TBA G4 and G3 & c-MYC

> Dans les sciences naturelles, et surtout dans la chimie, les généralités doivent résulter de la connaissance minutieuses de chaque fait, et non la précéder.
>
> *Joseph L. Gay-Lussac*

## Contents

## 6.1 Preface

In addition to the studies of human telomeric antiparallel G4, another antiparallel G4 has been studied: thrombin binding aptamer TBA G4 with $K^+$. This short DNA structure forms a two-layer G4. Additionally, it is the only sequence so far that, by removing one *GG* run, forms a stable triplex G3. Both, G4 (Sec. 6.2) and G3 (Sec. 6.3) have been studied and results are presented in this chapter.

The previous section reveals the underlying folding dynamics of an all-parallel quadruplex c-MYC (Sec. 6.6).

## 6.2 TBA Quadruplex with Potassium

TBA G4 forms only one topology with $K^+$ [6]. It is a 15mer containing only Thymine and Guanine bases. Static CD melting spectra of this antiparallel chair-type G4 in 150 mM $K^+$ phosphate buffer showed no indication for intermediates Sec. 4.6.1, meaning that there is at least no thermodynamically stable one. In analogy to Tel21 and Tel22, denaturation and renaturation dynamics were studied in order to find transient intermediates. Dynamical studies have been performed on TBA G4 in a 150 mM $K^+$ phosphate buffer.

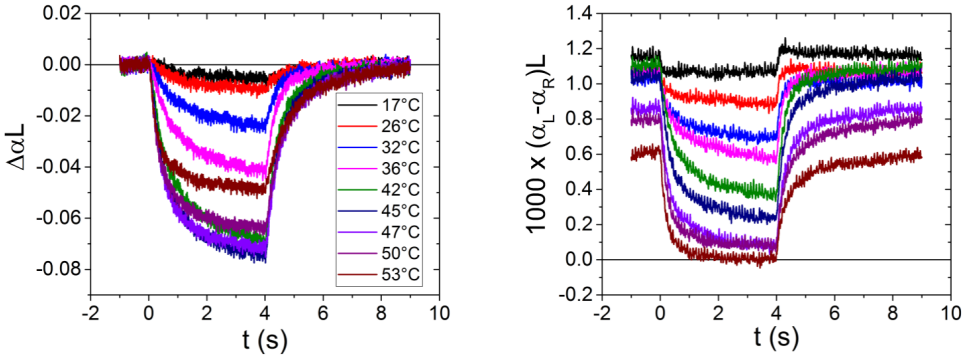


Figure 6.1: Absorption (left) and CD (right) curves at different temperatures. Curves were smoothed with the Savitzky-Golay method (40pts). Curves were recorded at 293 nm. Color codes are equal for both graphs. Heating started at time 0 and heating laser was stopped at t = 4 s. $\Delta T = 10°C$.

TBA G4 exhibits a large positive CD peak at 293 nm and a small negative one at 265 nm. Therefore, dynamical studies have been performed at these two wavelengths with emphasis put on 293 nm. Fig. 6.1 shows a series of denaturation and renaturation dynamics of 293 nm for TBA G4. $\Delta T$ of 10°C was applied and absorption and CD

changes were maximum at 47°C. Similarly to Tel21 with $Na^+$ and Tel22 with $Na^+$ denaturation is quite slow at low temperatures (see left: black & red curve) and increases with temperature. Renaturation shows an inverse behavior and slows down with increasing temperature.

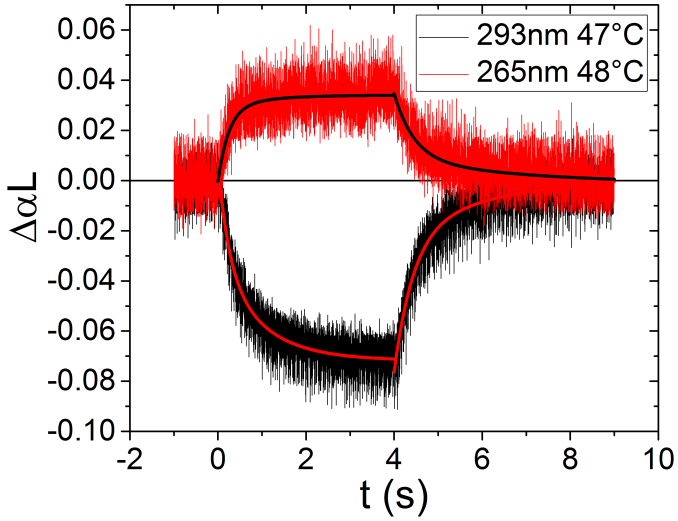


Figure 6.2: Comparison of dynamics at 293 nm ($T_{in}$ = 47°C, black) and at 265 nm ($T_{in}$ = 48°C, red). Fit functions are shown with global time constants. Heating started at time 0 and heating laser was stopped at t = 4 s. $\Delta T$ = 10K.

In analogy to Tel21($Na^+$) and Tel22($Na^+$) absorption decreases at 293 nm and increases at 265 nm during unfolding as illustrated in Fig. 6.2. Similar initial temperatures close to $T_m$ are depicted in that graph to illustrate equal dynamics at both wavelengths.

We have also measured renaturation dynamics at a given temperature $T_{in}$ = 47°C with different $\Delta T$. This is illustrated in Fig. 6.3. Complete denaturation was achieved for $\Delta T > 14$°C. Curves at that $\Delta T$ and higher $\Delta T$ show clearly that denaturation accelerates with increasing $\Delta T$. However, renaturation (between 4 to 9 s) slows down with increasing $\Delta T$. A biexponential fit was applied to the renaturation dynamics and its first time constants are plotted in Fig. 6.3 right. Renaturation exhibits a linear decrease in velocity with $\Delta T$.

Data treatment has been performed as shown in the previous chapter. A biexponential fit was first applied, then the first time constants were treated and plotted as shown before. Similar behavior as for Tel21 with $Na^+$ and Tel22 with $Na^+$ has been observed: Unfolding shows Arrhenius-like behavior with an activation barrier $E_A$ of $74 \pm 6$ $kJ/mol$. Folding velocity decreases with temperature over the whole

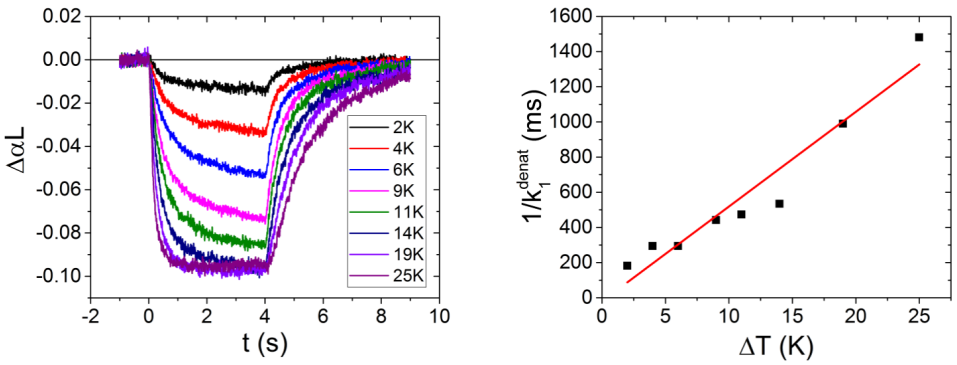


Figure 6.3: Left: Absorption curves at $T_in = 47°C$. Different T jumps $\Delta T$ from 2 to 25°C. Curves were smoothed with the Savitzky-Golay method (30pts). Heating started at time 0 and heating laser was stopped at t = 4 s. Right: first time constants of biexponential fit of renaturation ($\tau_{1,obs}^{renat} = 1/k_1^{renat}$) for different $\Delta T$. $\tau/\Delta T = 54 \pm 6ms/K$.

investigated temperature range. Thus, a clear non-Arrhenius-like behavior has been observed. Additionally, TBA G4 does not show any activation barrier by going from a partly denaturated state to its partly renaturated state. This corresponds to the results observed for Tel21 and Tel22. Coherently, the folding of TBA G4 is also in the 100s of millisecond time-range. This means it cannot be simple downhill-folding, either. As was mentioned in the case of Tel21 and Tel22, many propositions exist in literature [20]. A picture that has been emerging is that of a multi-pathway folding on a rugged surface with profound energetic wells attributed to misfolded structures. Also for TBA G4, we probably mostly explore the potential well of the folded structure. Results of the data treatment with $k_{inst}$ are depicted in Fig 6.5. In the left graph, values of $k_0$ are plotted against final temperatures. They showed the same behavior as $k_1$ determined by the biexponential fit procedure. The parameter for energy trap distribution $\sigma$ is shown in the right graph. For renaturation, a constant value of around $\sigma \approx 1.2$ was found for all temperatures. In contrast to that, $\sigma$ for denaturation is 0 at low temperatures and increases to the same value as the one for folding at $T > T_m$. Similar behavior was found for $\sigma$ of unfolding in Tel21 with 150 mM $Na^+$ concentration. The one for folding, however, decreased in the case of Tel21 with increasing temperature. In total, a larger $\sigma$ was observed in the case of Tel21 at low temperatures. This is expected, since Tel21 contains a longer sequence and three G quartets and thus has a larger number of energy traps. The fact that $\sigma$ (renaturation) is constant, whereas in Tel21 it decreases, is interesting. It means that the number of energy traps stays constant in the case of TBA throughout its

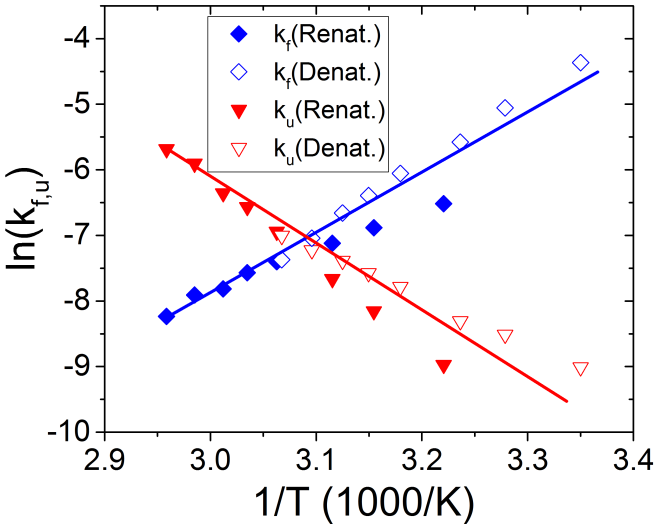


Figure 6.4: Arrhenius plot of $ln(k(T))$ against $1/T$. Values extracted from folding are depicted with filled dots, values from unfolding are shown with empty dots. Values of $k_f$ are blue, values of $k_u$ are red. Straight lines indicate tendencies.

refolding, whereas Tel21 has more at low temperatures.

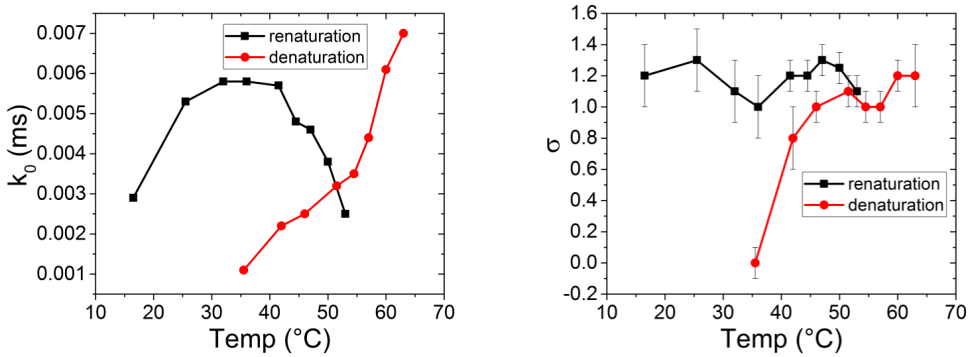


Figure 6.5: Results from data treatment via $k_{inst}$. Right: Values of $k_0$ (1/ms), left: values of $\sigma$.

Denaturation of TBA G4 at low temperature exhibits a quite small $\sigma$. This shows that at low temperatures the folded structure prevails and it can only unfold. Therefore, it is not slowed down by alternative unfolding pathways containing traps. At higher temperatures, the number of energy traps is equal for both, denaturation and renaturation. This is in agreement with the calculations of Zeng *et al.* showing that there are more transitions between partly unfolded states in the unfolded-like region than in the folded like [196]. It thus indicates that there are less traps at low temper-

ature and their number increases with unfolding of TBA. Additionally, they observed a manifold of different folding/unfolding pathways, in which transitions without intermediates exist. Other pathways included intermediates, such as a triplex. It is however not clear to what extent they are stable, i.e. observable in experiments.
In summary, TBA G4 results are quite similar to those of Tel21 and Tel22, indicating that there is no sequential mechanism and shows that TBA G4's underlying folding mechanism is rather multi-pathway-like. Additionally, the slow dynamics that have been found for TBA G4 are comparable to those of Tel21 and Tel22. Coherently, data treatment with $k_{inst}$ cannot explain the slow dynamics of $k_0$, either. Besides small traps that cannot be extracted with that method, misfolded structures are possible candidates to explain the slow dynamics [20]. In the case of 3-layer G4 possible candidates are 2-quartet G4 [21] or misfolded 3-quartet G4 [2]. In contrast to Tel21 and Tel22, TBA G4 has less possibilities to form misfolded structures, due to its shorter sequence. Hence, the only possible misfolded structure would be another 2-layer G4. However, this should be visible by monitoring CD dynamics at 293 nm and 265 nm, since all-parallel and hybrid-type G4 have clearly different CD peaks. If it exists, it is most likely an antiparallel basket-type G4.

The question, whether a triplex is a possible misfolded structure or a well-defined intermediate can be partly answered here. By comparing equal dynamics at 293 nm and 265 nm of TBA G4, we came to the same conclusion as with Tel21 and Tel22. There is most likely no fast-folding intermediate, since it should be visible in a difference of CD changes at both wavelengths even when it happens during the dead time of our measurements (50 ms). In order to be sure that there is no triplex intermediate and to answer the question whether it could be a stable misfolded structure, the next section describes the dynamics of TBA G3 and compares it to those of TBA G4.

## 6.3 TBA Triplex with Potassium & Comparison with G4

The triplex strucutre TBA G3 is investigated in 150 mM $K^+$ (same buffer as was used for TBA G4). Compared to TBA G4, the sequence is similar, but there is a *TTGG* run missing at the 3' end [7].

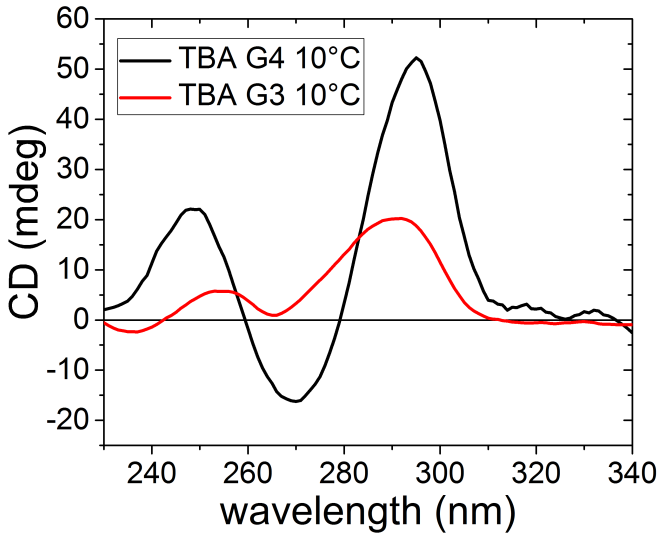


Figure 6.6: CD spectra of TBA G4 and G3 at 10° between 230 nm and 320 nm. Spectra are normalized to 1 mM DNA and 100 $\mu$m cell length.

In order to remember the differences in CD spectra of TBA G4 and TBA G3, they are depicted in Fig. 6.6. Peaks at 295 nm and 245 nm for TBA G4 are shifted in the case of TBA G3: 290 nm and 255 nm, respectively. TBA G3 has no peak at 265 nm, whereas TBA G4 has a negative one. Thus, during the denaturation and renaturation of TBA G4 this triplex should be visible by monitoring CD changes at these two wavelengths. Denaturation and renaturation dynamics of TBA G3 were studied in order to verify, whether it could be an intermediate or not. It is however not an absolute argument to exclude G3, since a possible triplex intermediate of TBA G4 could have a different CD spectrum than TBA G3.

Absorption and CD curves of TBA G3 dynamics are illustrated in Fig. 6.7. Due to its smaller signal in absorption as well as in CD, longer measurement times were necessary. Consequently, the level of the initial CD signal before the T-jump varied and did not yield its typical melting curve shape. Therefore, CD dynamics are normalized. Coherently, absorption curves in the left graph show similar behavior as TBA G4: Absorption as well as CD decreases during unfolding and the induced

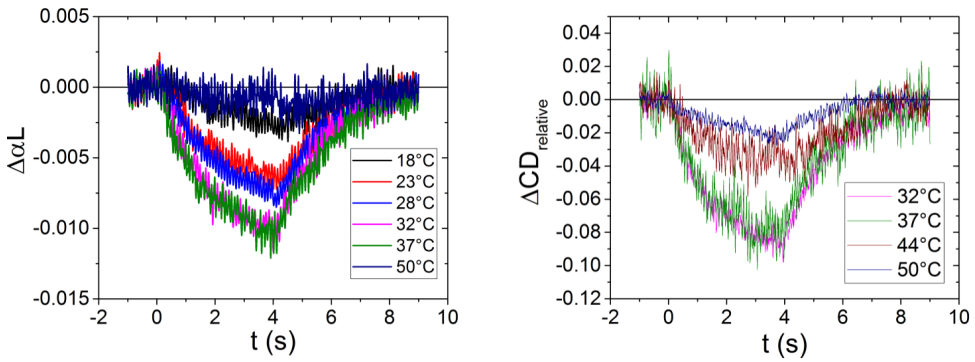


Figure 6.7: Left: Selected absorption curves at different temperatures. Right: CD curves. Curves were smoothed with the Savitzky-Golay method (60pts). Curves were recorded at 293 nm. Color codes are equal for both graphs. Heating started at time 0 and heating laser was stopped at t = 4 s. $\Delta T = 16$°C.

change increases until 37°C and decreases after $T_m$ (see curve at 50°C). In analogy to G4 dynamics, TBA G3 denaturation increases with temperature and renaturation becomes slower. There is, however, a remarkable difference. Both denaturation and renaturation dynamics are monoexponential. This is clearly different than for G4, indicating that the underlying mechanism is less complicated. The Arrhenius plot of Fig. 6.8 represents an Arrhenius-like behavior for unfolding ($E_A = 57 \pm 7\ kJ/mol$). Refolding is also non-Arrhenius-like, similar to that of TBA G4.

This non-Arrhenius-like behavior has also been found for curves at a given $T_{in} =$ 39°C with varying $\Delta T$ (see Fig. 6.9). On the left, absorption curves for $\Delta T$ between 2°C and 22°C are shown. Extracted renaturation time constants via monoexponential fit are plotted in the right graph. A clear linear behavior of increasing renaturation time constants with temperature is visible. It increases by around 150 $ms$/K. In comparison to that, TBA G4 showed around 50 $ms$/K for $\tau_1$. In order to compare these values, we fitted the curves of TBA G4 for different $\Delta T$ with a monoexponential function, which yields less accurate fits, but a value of also around 50 $ms$/K. Therefore, TBA G4 and G3 renaturation dynamics can be compared directly, not least because in both cases $T_{in}$ were 4°C under $T_m$.

Consequently, G triplex formation velocity is reduced by a factor of 3 compared to that of TBA G4. Not only formation or renaturation is slower in the case of TBA G3, the same effect is observed during denaturation. This is illustrated in Fig. 6.10 in which both curves were normalized to a maximum change of 1.

According to the rugged energy landscape proposed for Tel21 and Tel22 as well as for TBA G4, the number of potential wells and their depths depend on the number

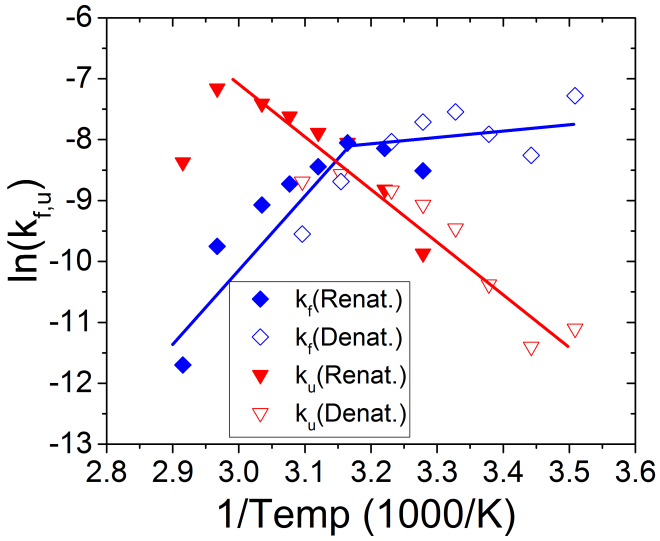


Figure 6.8: Arrhenius plot of $ln(k(T))$ against $1/T$. Values extracted from folding are depicted with filled dots, values from unfolding are shown with empty dots. Values of $k_f$ are blue, values of $k_u$ are red. Straight lines indicate tendencies.

of stacked Guanines, hydrogen bonds and potential mono-layer G-quartets. Since it is supposed that G-quartet formation is the driving force of G4 folding [20], it is not surprising that the formation of a G-triplex structure is slower than a G4 structure.

A treatment with $k_{inst}$ does not reveal more information, since $k_{inst}$ is time-independent in the case of TBA G3 and therefore no $\sigma$ can be extracted. The energy well of the folded G-triplex is probably not profound in a way that misfolded structures are equally stable. This could explain the long unfolding and refolding times. It is possible that the energy traps found for G4 with $k_{inst}$ do not exist for G3 and are thus related to the G-quartets. It could also be that due to the slow formation of G3, their influence is dominated by the slow formation dynamics.

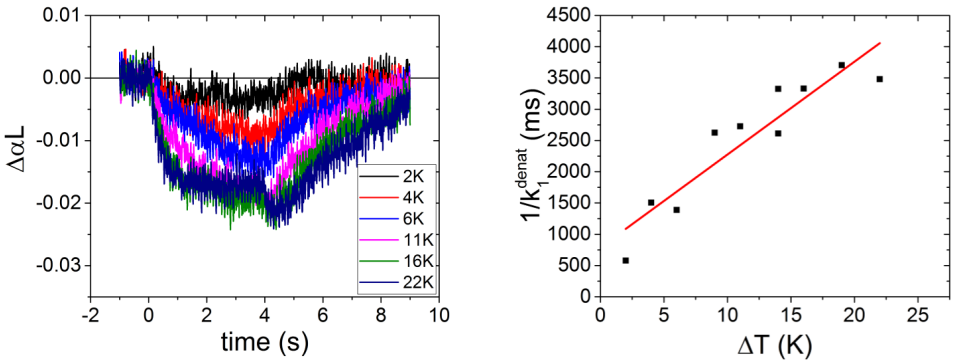


Figure 6.9: Left: Selected absorption curves at $T_{in}$ = 39°C. Different T jumps $\Delta T$ from 2 to 23K. Curves were smoothed with the Savitzky-Golay method (40pts). Heating started at time 0 and heating laser was stopped at t = 4 s. Right: time constants of monoexponential fit of renaturation ($\tau_{1,obs}^{renat} = 1/k_1^{renat}$) for different $\Delta T$. $\tau/\Delta T = 150 \pm 20\ ms/°C$.

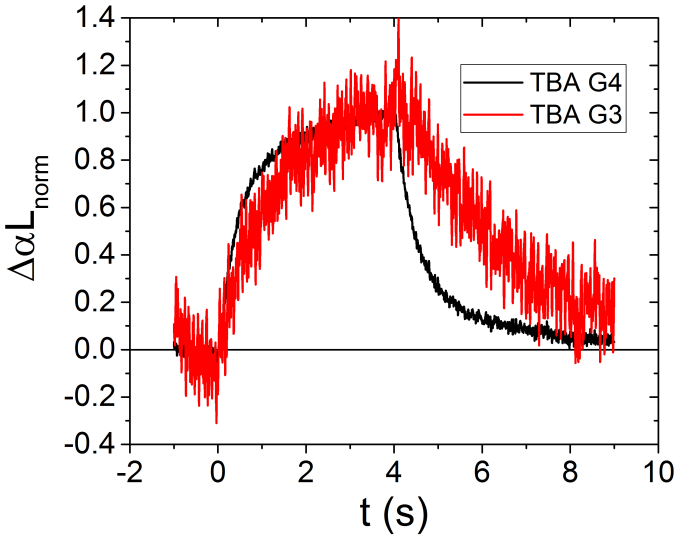


Figure 6.10: Normalized absorption curve of TBA G4 (black) and TBA G3 4°C below $T_m$ at 47 and 39°C, respectively. $\Delta T$ = 9°C. Curves were smoothed with the Savitzky-Golay method (60pts). Heating started at time 0 and heating laser was stopped at t = 4 s.

## 6.4 TBA G4 and G3 on microsecond time-scale

In order to assure that there are not fast dynamics during denaturation, we performed ns T-jump experiments for both TBA G4 and TBA G3.

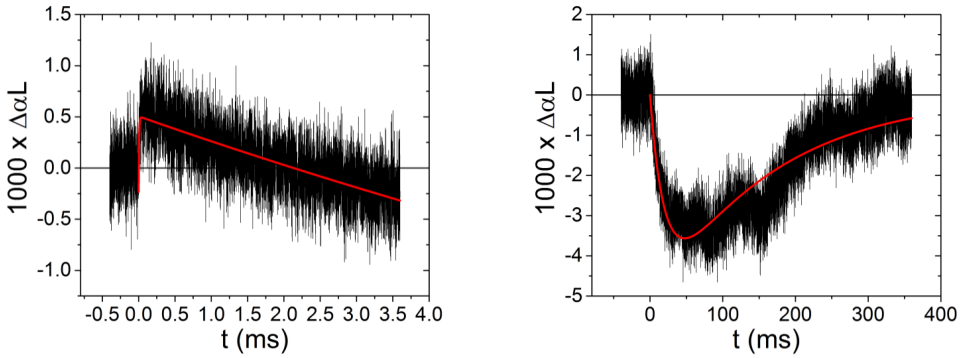


Figure 6.11: Dynamics of ns T-jump experiment on TBA G4. FWHM(probe) = 80 $\mu$m, FWHM(pump) = 300 $\mu$m, quartz cuvette, l = 100 $\mu$m. Left: First couple of ms, time resolution: 40 $\mu s$. Fit curve is in red. Decreasing time constant was fixed at 20 ms. Rising time constant was determined to be 6 $\mu s$. Right: First 360 ms after. First time constant was fixed at 20 ms. Relaxation time yielded 160 ms. $T_{in} = 49°$C, $\Delta T = 6°$C.

The left graph in Fig. 6.11 illustrates an instantaneous increase of absorption similar to Tel21, followed be a relatively slow decrease. In the right graph, a larger time window is shown, which contains the denaturation followed by renaturation. It is not different from Tel21 and does also not contain any fast $\mu s$ dynamics.

In the case of TBA G3, absorption changes are slower and smaller in amplitude (Fig. 6.12). An instantaneous increase of absorption is also visible as for TBA G4. Their amplitudes are equal, indicating that it is an artifact. The curve in the right graph of TBA G3 contains a lot of noise, which is due to its small signal change. Its renaturation is again clearly smaller than for TBA G4, which is in agreement with ms experiments.

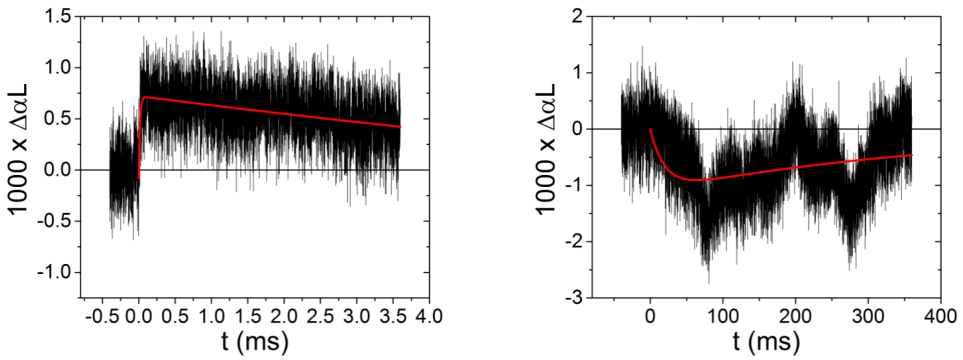


Figure 6.12: Dynamics of ns T-jump experiment on TBA G3. FWHM(probe) = 80 $\mu$m, FWHM(pump) = 300 $\mu$m, quartz cuvette, l = 100 $\mu$m. Left: First couple of ms, time resolution: 40 $\mu s$. Fit curve is in red. Decreasing time constant was fixed at 20 ms. Rising time constant was determined to be 16 $\mu s$. Right: First 360 ms after. First time constant was fixed at 20 ms. Relaxation time yielded 410 ms. $T_{in} = 40°C$, $\Delta T = 6°C$.

## 6.5 Conclusion

Comparisons with literature suggests that the impetus of G4 folding is not a G-hairpin or a G-triplex, but rather the G4 quartet [20]. By looking only on the number of hydrogen bonds, a hairpin of a three-layer G4 as Tel21 has 12 hydrogen bonds, a triplex also 12 and the final G4 24. In that perspective a two-layer misfolded structure has 16 bonds, meaning 4 more than a triplex/hairpin. However, in order to validate this sufficiently, the stabilizing effects of stacking and cation stabilization for those potential intermediates has to be known. The group of Šponer investigated this a lot and found a manifold of different hairpin topologies that could occur in the early processes of G4 folding [142]. Contrarily, they were not able to simulate stable triplex structures [20, 141] even though a stable structure has been found [7]. In summary, Šponer's work suggest that there is no specific hairpin or triplex structure in favor over the others, meaning a single sequential mechanism is unlikely.

Our results on the TBA G3 provide a crucial attribution to this discussion. The fact that TBA G3 takes significantly longer to form than TBA G4, makes it quite unlikely to be a distinct intermediate in the folding pathway of TBA G4. Furthermore, the results of TBA G4 show clearly that it is a multi-pathway folding. Regarding the evolution of CD signals at two wavelengths, no difference is visible. Since a G-triplex should clearly have a different CD shape, this highly indicates that a unfolding/folding pathway including a G-triplex is not at all favored. Additionally, it is the only characterized G3 structure so far, meaning that it is probably one of the

more stable G3 structures. Therefore, it is highly unlikely that a triplex is a misfolded structure in the case of TBA G4 at least and probably in general.

Comparisons of $\mu$s dynamics showed that there were no fast denaturation dynamics in the case of TBA G4 and TBA G3. This further excludes a fast-forming intermediate during denaturation.

## 6.6 c-MYC with Sodium

The studies on different antiparallel G4 show multi-pathway unfolding and folding. CD melting curves of c-MYC indicate that the mechanism might be different for all-parallel G4 (see Chapter 4, Sec. 4.7).

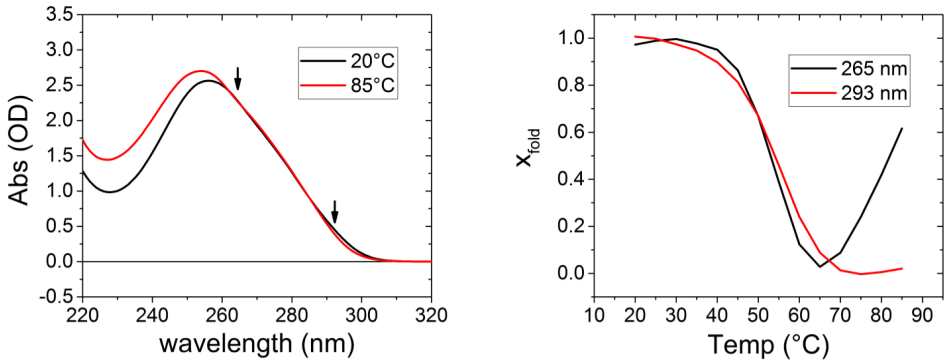


Figure 6.13: Left: Absorption spectrum at 20°C (black) and 85°C (red). Arrows indicate investigated wavelengths. Right: melting curves at the corresponding wavelengths.

Therefore, we investigated the denaturation and renaturation dynamics of c-MYC in order to get a more global picture of G4 folding. The used c-MYC derivate has been studied in the same buffer as for CD melting essays: 150 mM $Na^+$ phosphate buffer, since with $K^+$ it is extremely stable it cannot even be completely denaturated at 100°C [197]. The reason why we chose that derivative and not the wild-type one is that in our case it forms only one structure, whereas the wild-type forms at least two in $K^+$ [198].

Fig. 6.13 shows that absorption changes above 220 nm can again be investigated at 293 nm and 265 nm. On the other hand, there is no CD signal at 293 nm and a strong positive one at 265 nm (see Chapter 4, Sec. 4.7). Furthermore, the relative absorption change at 265 nm is very small between folded and unfolded species. Therefore, it is more convenient to measure absorption dynamics at 293 nm and CD dynamics at 265 nm. Results are depicted in Fig. 6.14. Those results are obtained from different series, but one can still see that dynamics are comparable and the same tendencies as for Tel21, Tel22 and TBA G4 are visible. Denaturation accelerates with increasing $T_{in}$, whereas renaturation slows down in both cases.

In analogy to the previous studied G4, c-MYC also exhibits biexponential behavior. Coherently, a biexponential fit yields similar tendencies for $\tau_1$ in the case of c-MYC (see Fig. 6.15). Unfolding rate constants increase with temperature and yield

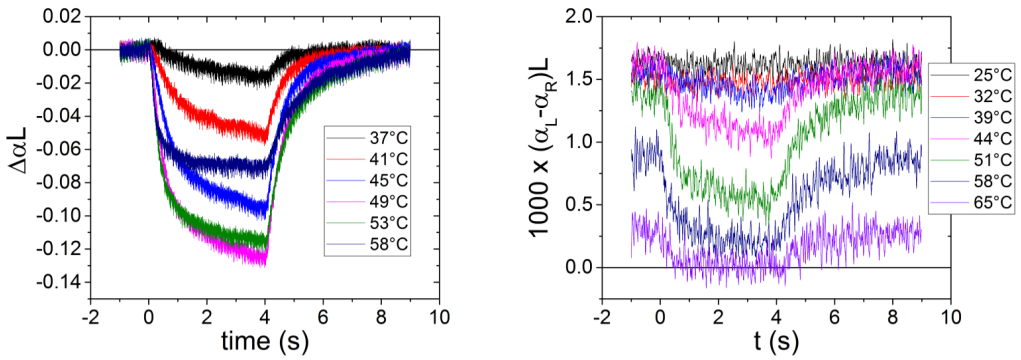


Figure 6.14: Left: Absorption dynamics at 293 nm at different temperatures. Right: CD changes at 265 nm at different temperatures (CD curves were smoothed: 60pts). $\Delta T = 8K$ in both cases.

an activation barrier of $E_A = 112 \pm 15\ kJ/mol$. Refolding rate constants showed no temperature dependence until $T_m$ and consecutively decreased with temperature.

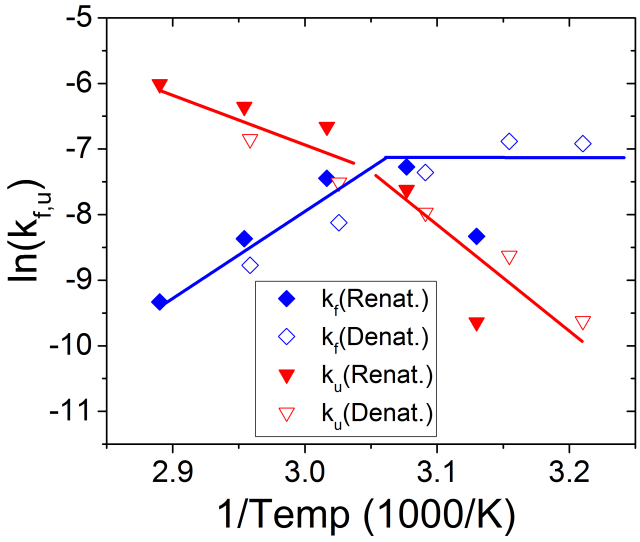


Figure 6.15: Arrhenius plots of the extracted $\tau = 1/k_{obs}$ from the c-MYC CD dynamics at 265 nm.

Data treatment with $k_{inst}$ revealed renaturation and denaturation values at $T_m$ of 0.003 $ms^{-1}$ (see Fig. 6.16(left)), which is pretty close to the values found for Tel21, Tel22 and TBA G4. This means that there is a strong similarity between all investigated structures concerning the fast time constants. It seems to be independent of sequence, cation and G4 topology. However, further investigations are necessary in order to confirm this observation.

The main difference lies in the $\sigma$ values. c-MYC exhibits constant $\sigma$ values for renaturation and denaturation Fig 6.16. This behavior has also been found for TBA G4 renaturation, but not for Tel21 and Tel22. The fact that the $\sigma$ of denaturation are also constant is unique for these investigated structures.

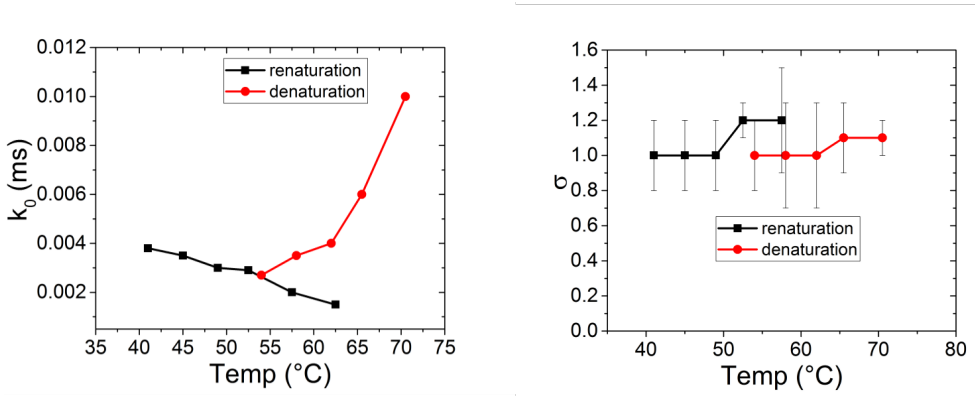


Figure 6.16: Results from data treatment via $k_{inst}$. Right: Values of $k_0$, left: values of $\sigma$.

In literature, dynamical and static studies have been performed on c-MYC (wild-type) and other derivatives with $K^+$, which forms more than one structure [54, 197, 199, 200]. There is hardly any investigation on dynamics of the used c-MYC sequence in this work. However, some interesting observations and comparisons can be made

Force jump experiments of the wild time c-MYC with $K^+$ [197] showed a 1000-fold larger rate constant compared to human telomeric G4 with $K^+$ [135, 201, 202]. Contrarily, our experiments showed a $k_0$ value of c-MYC with $Na^+$ similar to those of the telomeric G4 ($k_0 \approx 0.003\ ms^{-1}$). The wild-type c-MYC contains three $GGGG$ runs and can thus form perhaps a more stable G4 than our modified c-MYC. The sequence used in the present work exhibits a $T_m = 54°C$, which is smaller than the one of Tel21 and Tel22. This could explain the similar dynamics found in our work. Furthermore, mechanical unfolding restricts the energy landscape to a given pathway [20]. Therefore, it is possible that mechanical unfolding forces a given pathway in which wild-type c-MYC($K^+$) unfolds significantly slower than human telomeric G4 with $K^+$.

The group of Mergny studied unfolding of a c-MYC sequence containing only the G4 forming core of our sequence by adding the complementary DNA strand [54]. Their determined activation energy is around 30 $kJ/mol$ at 1.5 mM $K^+$ concentration. This is significantly smaller than our value of c-MYC: $E_A \approx 112\ kJ/mol$ at 150 mM $Na^+$ concentration, because they had to work at such a small $K^+$ concentration in order to be able to unfold c-MYC G4.

In contrast to anti-parallel and hybrid-type G4, all-parallel G4 could only start from a G-hairpin with a chain-reversal loop, which seems to be highly unlikely. The unfolding mechanism of all-parallel G4 has been studied with simulations by the group of Šponer [140]. However, they performed it on Tel22 in which the loops have three nucleotides and behave probably differently than those of c-MYC. Therein, a manifold of possible strand slippages results in the final folded all-parallel G4. However, this takes place on a submicrosecond time-scale, which does not explain the long unfolding/folding times of our experiments.

### 6.6.1 Nanosecond Dynamics

Coherently with the antiparallel studied G4, we have not observed any fast dynamics for c-MYC during a ns T-jump. Two curves at different initial temperatures are depicted in Fig. 6.17. The fast rising peak with opposite sign has also been found for c-MYC.

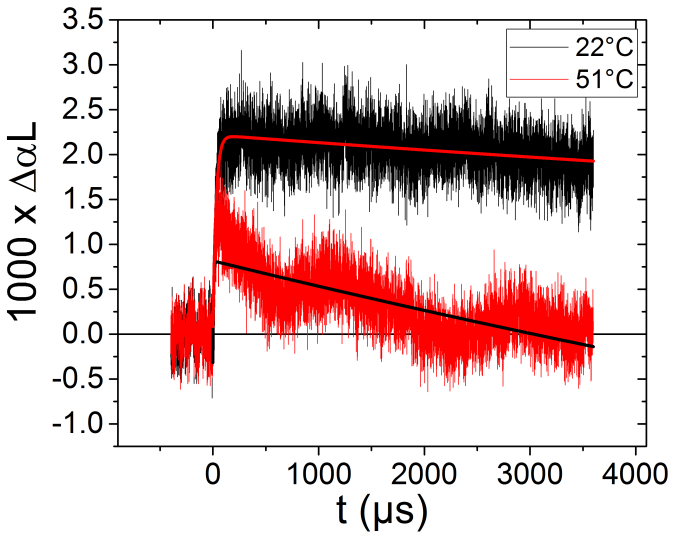


Figure 6.17: Dynamics of ns T-jump experiment on TBA G4 at 293 nm. FWHM(probe) = 80 $\mu$m, FWHM(pump) = 300 $\mu$m, quartz cuvette, l = 100 $\mu$m. Left: First couple of ms, time resolution: 40 $\mu s$. Fit curves are in opposing colors. Decreasing time constant was fixed at 20 ms. Fast rising yields time constants of $\tau = 30$ $\mu s$ (22°C) and $\tau = 6$ $\mu s$ (51°C). $\Delta T = 6$°C.

#### 6.6.1.1 Conclusion

Conclusively, c-MYC also exhibits dynamics that show a multi-pathway folding on a rugged energy landscape, in which no profound energy traps were necessary in order to fit the data.

Furthermore, our results of c-MYC show that G4 dynamics in general are related to its stability, since we find similar $k_0$ constants for Tel21 and Tel22, which have also similar $T_m$ and $\Delta H$ values. This is also true for TBA G4, which has a $T_m$ value close to that of c-MYC and a similar $\Delta H$.

In contrast to that, c-MYC does not show any temperature-dependence of its ruggedness, which is clearly different from Tel21 and Tel22. This indicates that unfolding and folding of c-MYC contains different traps and/or follow different pathways.

One significant difference of c-MYC compared to the other investigated structures is the absorption behavior at 265 nm. This is discussed in the following

### 6.6.2 Absorption Dynamics of c-MYC at 265 nm

The absorption dynamics of c-MYC at 265 nm were not discussed before, due to its particular behavior. This is reflected in its melting curves depicted in Fig. 6.18. Cooling exhibited a normal melting profile, whereas during heating, absorption increases up to $T_m$ after which it showed the same profile as during cooling. This observation gave us some headache, since it has also shown up in dynamical studies (see Fig. 6.19). A page essay showed that there are no dimers and it is not due to dimer dynamics. It occurs most likely due to a spectral shift, because 265 nm is close to spectral points at which absorption change inverses with temperature (see right graph in Fig. 6.18).

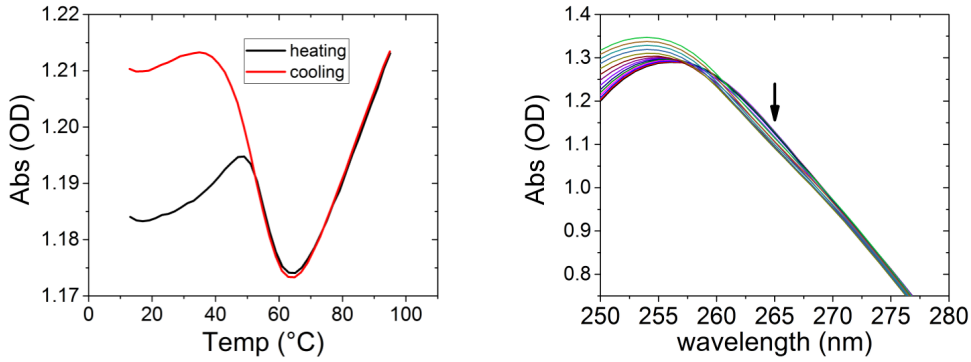


Figure 6.18: Left: absorption changes of c-MYC with $Na^+$ at 265 nm. Right: spectral changes around 265 nm. Arrow indicates rising temperature. No uncommon behavior has been found in the right graph.

The dynamics exhibited an increase in absorption at low temperatures and are superimposed (left graph in Fig. 6.19, 12°C to 32°C). At 39°C, absorption still increased but one can see some underlying dynamics that show a decrease in absorption. This decrease became dominant for higher temperatures, while the dynamics with increasing absorption were still participating. The left graph shows the same curves, which were subtracted by the curve at 12°C. One can clearly see that this almost compensates the effect of the dynamics with increasing absorption, meaning that those dynamics do not change with temperature.

Consecutive investigations will study the dynamics of different c-MYC derivatives (without flanking bases, different nucleotides within the loops) in order to understand the very origin of this behavior, which is probably due to a spectral shift.

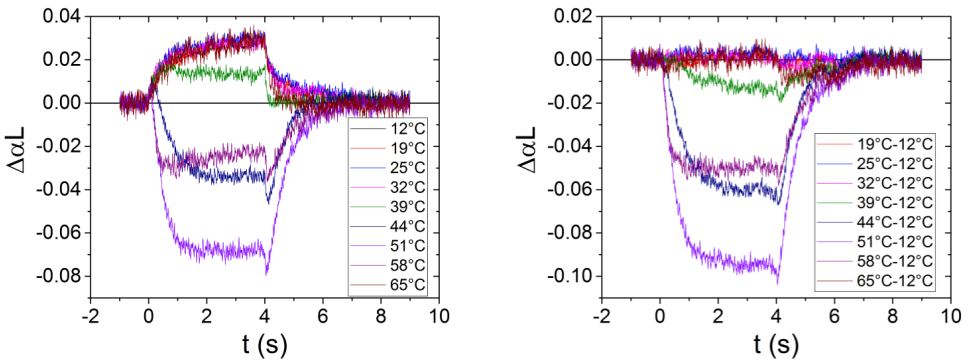


Figure 6.19: Left: Dynamic absorption changes of c-MYC($Na^+$) at different $T_{in}$. It is the same series as for which CD dynamics were shown in Fig. 6.14. Right: Same series but curves were subtracted by the dynamics at 12°C.

# Chapter 7

# Conclusion & Perspectives

My thesis has focused on the study of the folding mechanisms of several biologically-relevant guanine-quadruplexes (G4). To this end, we have used circular dichroism (CD) spectroscopy, both in static and time-resolved experiments, in order to monitor the temperature-induced conformational changes of DNA. G4 folding dynamics have proven to be too slow to be captured entirely by our nanosecond T-jump laser setup. Therefore, we have extended the time window of our setup from a few milliseconds to seconds by using a continuous laser heating to enlarge our temporal window. In the meantime, we had to improve the sensitivity of our detection for measuring the small CD signals of G4. To this end, we have implemented a CD detection using the combination of a PEM and a lock-in amplifier to reach a sensitivity of $10^{-5}$ (see Chapter 3).

Prior to time-resolved measurements, we have performed static synchrotron radiation CD (SRCD) on a series of short G4-forming sequences exhibiting different topologies. Despite the fact that SVD analysis of the SRCD spectra measured at different temperatures cannot clearly distinguish a possible presence of several G4 conformations, they have revealed dominant contributions of the intrastrand interactions in the spectral region below 260 nm, while at longer wavelengths signals mainly reflect the interactions of guanines in the G4 core. Consequently, we have performed time-resolved thermal denaturation and renaturation measurements for four selected G4-forming sequences (Tel21, Tel22, TBA G4 and c-MYC) and one G-triplex structure (TBA G3). These studies presented in Chapter 5 and 6 have revealed biphasic kinetics with time constants of a few 100 ms and a few seconds. Denaturation and renaturation display different kinetics that rule out a simple two-state folding mechanism and suggests instead distinct folding and unfolding pathways. In stark contrast with unfolding, no activation barriers have been found for G4 folding at low temperatures, suggesting the absence of traps close to the native-folded ensemble. The study

of the effect of cation concentration $[Na^+]$ on the denaturation and renaturation kinetics of Tel21 has also emphasized a significant role in the thermodynamical as well as the dynamical properties of G4. Regardless of the data treatment, all these results point towards complex multi-pathway G4 mechanisms that may involve deep traps in the region of the unfolded ensemble. Interestingly, all studied antiparallel G4 (Tel21, Tel22, TBA G4) as well as the all-parallel one (c-MYC) exhibit similar behavior indicating similar complex mechanisms. Future works will focus on hybrid-type G4 topologies, in order to make more generalized statements. Finally, the comparison between the dynamics of TBA G4 and TBA G3 has revealed significantly slower folding rates with respect to those of G4, making the formation of a triplex intermediate in the course of G4 folding unlikely.

In conclusion, this work has allowed us to shed new light on common features of G4 dynamics, whatever their specific conformation. However, many studies remain to be performed to deepen our understanding of their folding mechanisms, either by studying other G4-forming sequences or by exploring the use of different perturbations to induce folding or unfolding. Since heating and cooling dynamics are in the time order of G4 dynamics, further investigations are necessary. First, we are working on a model to extract pure G4 dynamics from our raw data in order to obtain quantitative results. Second, the use of $CaF_2$ glass windows instead of quartz shows a four-fold faster temperature change in our simulations. With that, we would have on one hand faster temperature changes and thus a smaller influence on the G4 dynamics, on the other hand we would have a model that extracts G4 dynamics from two different heating/cooling profiles. This will allow us to accurately quantify G4 dynamics.

A promising approach for a different perturbation to induce unfolding is the use of azobenzene derivative photoswitches for photo-induced denaturation/renaturation of G4 [203]. The use of distinct perturbations allows to probe distinct region of the folding landscapes providing additional information on the folding mechanism of DNA. In addition, an on-going collaboration with the beamline DISCO at SOLEIL is focusing on the use of these perturbations (T-jump, photoswitch) in combination with the time-resolved SRCD measurements in order to obtain the temporal evolution of the entire G4 CD spectra down to 180 nm. Probing the VUV region may reveal interesting information on the fate of the intrastrand interaction especially in the G4 loops in the course of folding.

Last but no least, I have focused a lot on the development of new techniques during my thesis. Therefore, the last two months of my stay at the LOB will be mostly dedicated to test further the new ideas mentioned in Sec. 3.6.3 for the imple-

mentation of the two-pulse TRCD setup in order to have a sensitive method whose time-resolution will be only limited by the time delay between the two pulses and their temporal widths. With these sets of complementary experiments in combination with the development of theoretical approaches for simulating G4 folding, no doubt that the fascinating world of G-quadruplexes will give other parts of their secrets.

# Supplementary information

## SI: CD spectrometer for T-Jump measurements

### Novelties

Many modifications have been performed to improve different aspects of the present CD spectrometer during my thesis.

Long measurement times and UV light sometimes result in sample degradation. Consequently, the internal shutter of monochromator closes now whenever no CD spectra are recorded. This occurs during heating/cooling steps for example. Furthermore, initially one could only record spectra starting from lower to higher wavelengths. Now, both is possible. It is, though, recommended to choose decreasing wavelengths, since with lower wavelengths phototoxicity increases. The waiting time for lock-in equilibrium emergence was reduced without any observed change in CD spectra. Besides, it is highly recommended to choose a step size of 1 nm. Larger steps seem to decrease spectrum quality, smaller steps simply increase measurement time. As mentioned before, the voltage applied to PMT is varied. A slight change in the program that just takes the last applied voltage value as the initial one to adjust PMT voltage allowed also to decrease measurement time. It takes now 40 s per nm.

A complete temperature-control of CD spectra has been added. One can now choose 1 to 5 different wavelength regions and 1 to 3 different temperature regions. Step sizes (nm and temperature) can be varied independently for each region. A region of 1 wavelength is suitable for melting curves (initial wavelength = final wavelength). Additionally, 5 different peaks can be monitored. One can furthermore choose to heat or cool or to do both during one measurement. The option of heating with going back to initial temperature is also possible to verify if nothing was damaged during melting. Since melting curves of whole spectra and also TRCD measurements can take several hours, an option to shut down the lamp at the end has also been added. This increases the life time of the lamp. However, it is not possible to shut down all instruments automatically at the end. Future work could include that.

There is only one challenge remaining. A random communication error between PEM and computer occurs sometimes, especially for long measurements. In that case, all data were lost before. I could not resolve that problem either. However, the program now saves data after each wavelength step (temperature-independent mode), so that at least this data is conserved.

## Alignment of Two-pulse TRCD setup

A brief section is dedicated to alignment, due to its importance for the two-pulse setup. Many procedures of how to perform it exist and were also undertaken. We will discuss here the one we used.

Procedure (see Fig. 3.21(b)):

- Mount linear polarizer (P)
- All intensity goes to PMT1
- Mount birefringent plate (BP)
- Turn BP to 45°
- Both detectors show same intensity
- Turn BP back to 0°
- All intensity goes to PMT1
- Mount 1st QWP
- Turn 1st QWP to 45°
- Adjust it to circular mode until both detectors show same intensity
- Turn QWP back to 0°
- All intensity goes to PMT1
- Mount 2nd QWP
- Turn 2nd QWP to 45°
- Adjust it to circular mode until both detectors show same intensity
- Turn BP to 45°
- No intensity change at detectors should happen
- Alignment completed

# SI: Steady-state CD spectra

| G4 | c (mmol/L) | Cation | c (mmol/L) | L ($\mu m$) |
|---|---|---|---|---|
| First run: | | | | |
| Tel21 | 0.9 | $Na^+$ | 150 | 33.5 |
| 2GKU | 0.9 | $K^+$ | 150 | 37.5 |
| 22CTA | 0.8 | $K^+$ | 150 | 37.5 |
| TBA G4 | 0.4 | $K^+$ | 150 | 37.5 |
| c-MYC | 0.8 | $Na^+$ | 150 | 37.5 |
| Second run: | | | | |
| Tel21 | 0.8 | $Na^+$ | 25 | 33.5 |
| Tel21 | 0.9 | $K^+$ ($Na^+$) | 100 (25) | 37.5 |
| TBA G3 | 0.4 | $K^+$ | 150 | 200 |
| Third run: | | | | |
| Tel21 | 0.9 | $Li^+$ | 150 | 33 |
| Tel22 | 0.9 | $Na^+$ | 150 | 33 |
| Tel22 | 0.9 | $K^+$ | 100 | 33 |
| 26CEB | 0.9 | $K^+$ | 100 | 33 |

Table 7.1: Investigated structures and conditions for steady state CD melting curves at SOLEIL. Concentration is c and L is pathlength through cuvette.

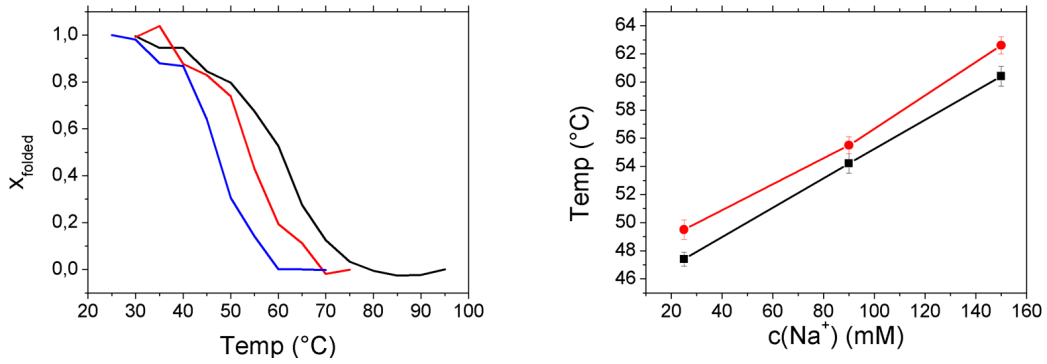


Figure 7.1: Left: Melting curves of different sodium concentrations at 295 nm (raw data): 150 mM black, 90 mM red, 25 mM blue. Right: melting points of different ion concentrations at 293 nm for the raw data (black dots) and the first singular value (red dots). Tel21 concentration was roughly 1 mM for all three curves (see Table 7.1).

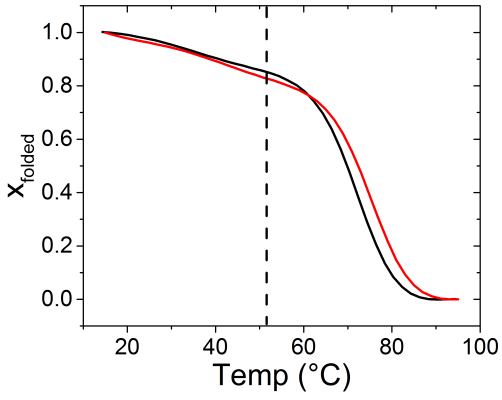


Figure 7.2: Absorption melting curves of Tel21 in 500mM $Na^+$ (black) and 900mM $Na^+$ (red).

# SI: Tel21 with $Li^+$

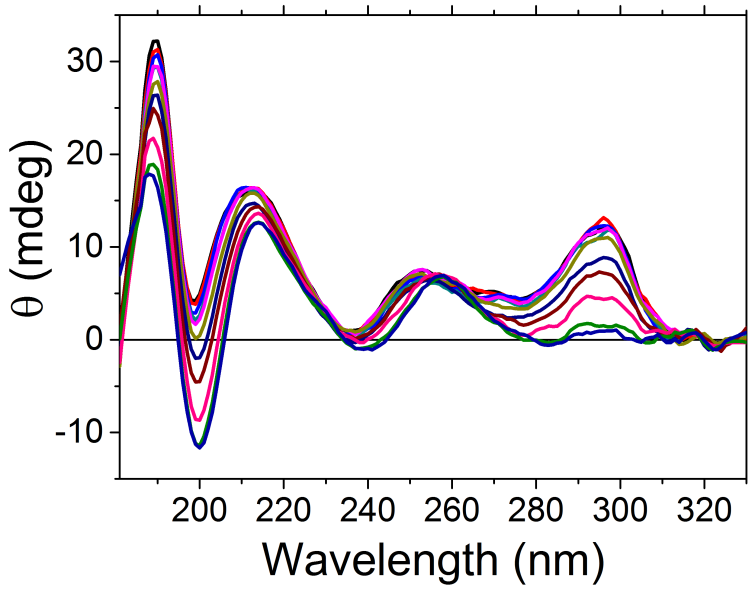


Figure 7.3: CD spectra of Tel21 in 150 mM $Li^+$. Spectral range is 181-330 nm. CD spectra were recorded from 5°C to 55°C every 5°. Black curve: 5°C, dark blue curve: 55°C.

Lithium is know to not stabilize G4 structure [40, 204]. Even though Tel21 does not form a G4 with $Li^+$, it still exhibits some CD signals (see Fig. 7.3). Interestingly, there is a peak at 295 nm that vanishes during heating. This peak was also observed by Marchand and Gabelica, in which they stated that sequences starting with a $5'-G$ tract can form secondary structures without cations. Those structures could exhibit a hairpin-like shape [203]. It could also be that there is a base-stacking between $5'-G$ and $3'-G$. A comparison of Tel22 and Tel21-T with $Li^+$ could be performed in order to verify this.

We also performed a series of dynamical studies on Tel21 with $Li^+$. The results are illustrated in Fig. 7.4. At first glance it looks similar to the curves of Tel21 with $Na^+$, except that denaturation and renaturation are clearly faster for Tel21 with $Li^+$. This is not surprising, since it is significantly less stable with $Li^+$. Dynamics exhibits almost a monoexponential behavior. Only for major amplitude changes a biexponential function yields a slightly better fit. This means that dynamics are less complex than in the case of Tel21 with $Na^+$ at 150 mM. There is again an acceleration of denaturation and a slowing down of renaturation with temperature.

Fig. 7.5 shows the Arrhenius plot of the extracted folding and unfolding rate con-

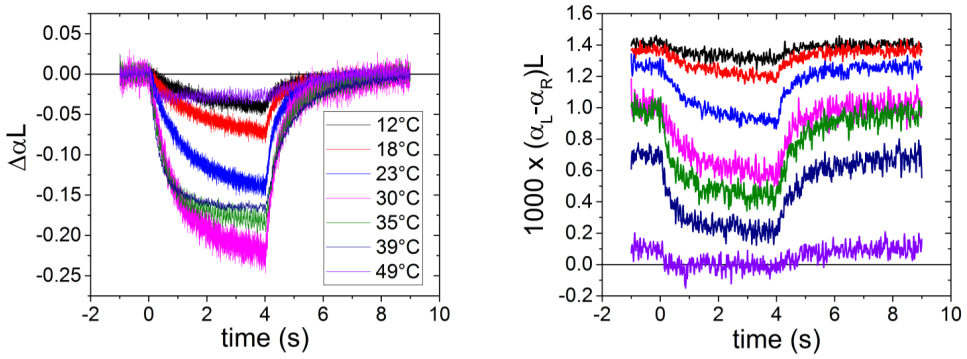


Figure 7.4: Left: Selected absorption curves at different temperatures. Left: selected CD curves. Curves were smoothed with the Savitzky-Golay method.(Abs (only 35, 39, 49°C) = 40pts, CD = 60pts). Curves were recorded at 293 nm. Color codes are equal for both graphs. Heating started at time 0 and heating laser was stopped at t = 4 s. $\Delta T = 10°C$.

stants. The folding rate constants show again Arrhenius-like behavior. Unfolding rate constants are temperature-independent. This figure could suggests, that unfolding rate constants are different for renaturation and denaturation. In order to verify this, we performed supplementary series, that confirmed that unfolding rate constants are temperature-independent.

Those results are interesting, since they show that there is no energy barrier between unfolded and folded species and that there are almost no energy traps. This further confirms the conclusions of Chapter 5 that $Na^+$ stabilize energy traps and indicates that the slow dynamics are probably due to the DNA strand itself.

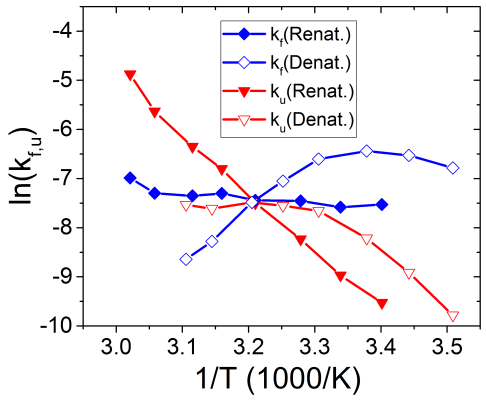


Figure 7.5: Arrhenius plot of $ln(k(T))$ against $1/T$. Values extracted from folding are depicted with filled dots, values from unfolding are shown with empty dots. Values of $k^{fold}$ are blue, values of $k^{unf}$ are red.

# SI: Dynamics of Tel21

### SI - Fitting function

The derivation of the fit function used for the curves obtained by ns T-jump in Sec. 5.4 is described here.

First, we define fraction folded and unfolded as $N_f$ and $N_u$, respectively. Its sum stays constant: $N = N_f + N_u$. Absorption coefficient of folded and unfolded species is defined as $\alpha_f$ and $\alpha_u$, respectively. After the induced T-jump, both fractions change: $N_f \rightarrow N_f + \Delta N_f$ and $N_u \rightarrow N_u + \Delta N_u$. The unfolding absorption coefficient changes equally: $\alpha_u \rightarrow \alpha_u + \Delta\alpha_u$. In that approach $\alpha_f$ is supposed to be temperature independent. From that the absolute absorption change is described as

$$\Delta\alpha = N_u \Delta\alpha_u + \Delta N_u(\alpha_u - \alpha_f) \tag{7.1}$$

Here, some assumptions were made. The term $\Delta N_u \Delta\alpha_u$ is neglected, which is true as long as $\Delta N_u \ll N_u$. In our experiments, $\Delta T$ was only around a few degrees and therefore this approximation is valid. Furthermore, we suppose that the temperature dependence of $\Delta\alpha_u(t)$ and thus also its time dependence occurs from fast dynamics (possibly destacking of Guanines). The observed quasi instantaneous increase of absorbance, followed by a fast decrease is thought to be roughly as fast as the temperature change. Therefore $\Delta\alpha_u(t)$ writes

$$\Delta\alpha_u(t) = \Delta\alpha_{u0} \cdot e^{-t/\Gamma} \tag{7.2}$$

where $\Delta\alpha_{u0}$ reflects the quasi instantaneous absorbance change and $\Gamma$ is the cooling time. The second time-dependent term in Eq. 7.1 that is $\Delta N_u(t)$. It describes mixed dynamics of unfolding/folding and cooling. Without cooling, the fraction of unfolded species is expressed as follows:

$$\frac{dN_u}{dt} = -\gamma(N_u - N_T) \tag{7.3}$$

where $N_T$ is the maximum unfolded fraction induced by a T-jump. Cooling is described as

$$N_T^* = N_{u0} + (N_T - N_{u0})e^{-\Gamma t} \tag{7.4}$$

where $N_T^* = N_T(t)$. Eq. 7.3 and Eq. 7.4 together yield

$$\frac{dN_u(t)}{dt} = -\gamma\Big(N_u(t) - N_{u0} - (N_T - N_{u0}) \cdot e^{-\Gamma t}\Big) \tag{7.5}$$

This leads to

$$\boxed{\Delta N_u(t) = \frac{\gamma}{\Gamma - \gamma} \cdot (N_T - N_{u0}) \cdot \left(e^{-\gamma t} - e^{-\Gamma t}\right)} \tag{7.6}$$

By inserting Eq. 7.2 and Eq. 7.6 in Eq. 7.1 we obtain

$$\Delta\alpha = N_u \Delta\alpha_{u0} \cdot e^{-\Gamma t} + \frac{\gamma}{\Gamma - \gamma}(N_T - N_{u0}) \cdot (\alpha_u - \alpha_f) \cdot \left(e^{-\Gamma t} - e^{-\gamma t}\right) \tag{7.7}$$

Thus, our fitting function becomes

$$\boxed{y = y_0 \cdot e^{-\Gamma t} + K \cdot \left(e^{-\Gamma t} - e^{-\gamma t}\right)} \tag{7.8}$$

with $K = \frac{\gamma}{\Gamma - \gamma}(N_T - N_{u0}) \cdot (\alpha_u - \alpha_f)$.

# Bibliography


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

[2] I. Bessi, H. R. A. Jonker, C. Richter, and H. Schwalbe. Involvement of long-lived intermediate states in the complex folding pathway of the human telomeric G-Quadruplex. *Angew. Chem. Int. Ed.*, 54:8444–8448, 2015.

[3] K. W. Lim, P. Alberti, A. Guédin, L. Lacroix, J.-F. Riou, N. J. Royle, J.-L. Mergny, and A. T. Phan. Sequence variant (CTAGGG)n in the human telomere favors a G-quadruplex structure containing a G·C·G·C tetrad. *Nucleic Acids Res.*, 37:6239–6248, 2009.

[4] A. T. Phan, V. Kuryavyi, K. N. Luu, and J. P. Dinshaw. Structure of two intramolecular G-quadruplexes formed by natural human telomere sequences in $K^+$ solution. *Nucleic Acids Res.*, 35:6517–6525, 2007.

[5] J. Dai, M. Carber, C. Punchihewa, R. Jones, and D. Yang. Structure of the hybrid-2 type intramolecular human telomeric G-quadruplex in $K^+$ solution/ insight into structure polymorphism of the human telomeric sequence. *Nucleic Acids Res.*, 35:4927–4940, 2007.

[6] R. F. Macaya, P. Schultze, F. W. Smith, J. A. Roe, and J. Feigon. Thrombin-binding DNA aptamer forms a unimolecular quadruplex structure in solution. *Proc. Natl. Acdad. Sci.*, 90:3745–3749, 1993.

[7] V. Limongelli, S. De Tito, L. Cerofolini, M. Fragai, B. Pagano, R. Trotta, S. Cosconati, L. Marinelli, E. Novellino, I. Bertini, A. Randazzo, C. Luchinat, and M. Parrinello. The G-Triplex DNA. *Angew. Chem. Int. Ed.*, 52:2269–2273, 2013.

[8] A. T. Phan, Y. S. Modi, and D. J. Patel. Propeller-type parallel-stranded G-Quadruplexes in the human c-myc promoter. *J. Am. Chem. Soc.*, 126:8710–8718, 2004.

[9] A. T. Phan, V. Kuryavyi, H. Y. Gaw, and D. J. Patel. Small-molecule interaction with a five-guanine-tract G-quadruplex structure from the human c-MYC promoter. *Nat. Chem. Biol.*, 1:167–173, 2005.

[10] M. Gellert, M. N. Lipsett, and D. R. Davies. Helix formation by guanylic acid. *Proc. Natl. Acad. Sci.*, 48:2013–2018, 1962.

[11] J. R. Williamson, M. K. Raghuraman, and T.R. Cech. Monovalent cation-induced structure of telomeric DNA: the G-quartet model. *Cell*, 59:971–980, 1989.

[12] J. D. Watson and F. H. Crick. A Structure for Deoxyribose Nucleic Acid. *Nature*, 171:737–738, 1953.

[13] G. Biffi, D. Tannahill, J. McCafferty, and S. Balasubramanian. Quantitative visualization of DNA G-quadruplex structures in human cells. *Nat. Chem.*, 5: 182–186, 2013.

[14] D. Rhodes and H. J. Lipps. Survey and symmary. G-quadruplexes and their regulatory roles in biology. *Nucleic Acids Res.*, 43:8627–8637, 2015.

[15] J. W. Szostak. DNA ends: Just the beginning (nobel lecture). *Angew. Chem. Int. Ed.*, 49:7386–7404, 2010.

[16] R. K. Moyzis, J. M. Buckingham, L. S. Cram, M. Dani, L. L. Deaven, M. D. Jones, J. Meyne, R. L. Ratliff, and J. R. Wu. A highly conserved repetitive DNA sequence, (TTAGGG)n, present at the telomeres of human chromosomes. *Proc. Natl. Acad. Sci.*, 85:6622–6626, 1988.

[17] R. C. Allshire, J. R. Gosden, S. H. Cross, G. Cranston, D. Rout, N. Sugawara, J. W. Szostak, P. A. Fantes, and N. D. Hastie. Telomeric repeat from T. thermophila cross hybridizes with human telomeres. *Nature*, 332:656–659, 1988.

[18] A. G. Bodnar, M. Ouellette, M. Frolkis, S. E. Holt, C. P. Chiu, G. B. Morin, C. B. Harley, J. W. Shay, S. Lichtsteiner, and W. E. Wright. Extension of life-span by introduction of telomerase into normal human cells. *Science*, 279: 349–352, 1998.

[19] N. W. Kim, M. A. Piatyszek, K. R. Prowse, C. B. Harley, M. D. West, P. L. Ho, G. M. Coviello, W. E. Wright, S. L. Weinrich, and J. W. Shay. Specific association of human telomerase activity with immortal cells and cancer. *Science*, 266:2011–2016, 1994.

[20] J. Šponer, G. Bussi, P. Stadlbauer, P. Kührová, P. Banáš, B. Islam, S. Haider, S. Neidle, and M. Otyepka. Folding of guanine quadruplex molecules - funnel-like mechanism or kinetic partitioning? an overview from MD simulation studies. *Biochim. Biophys. Acta*, 1831:1246–1263, 2017.

[21] A. Marchand and V. Gabelica. Folding and misfolding pathways of G-quadruplex DNA. *Nucleic Acids Res.*, 44:10999–11012, 2016.

[22] V. Gabelica. A pilgrim's guide to G-quadruplex nucleic acid folding. *Biochimie*, 105:1–3, 2014.

[23] R. D. Gray, J. O. Trent, and J. B. Chaires. Folding and unfolding pathways of the human telomeric G-quadruplex. *J. Mol. Biol.*, 426:1629–1650, 2014.

[24] S. L. Noer, S. Preus, D. Gudnason, M. Aznauryan, J.-L. Mergny, and V. Birkedal. Folding dynamics and conformational heterogeneity of human telomeric G-quadruplex structures in $Na^+$ solutions by single molecule FRET microscopy. *Nucleic Acids Res.*, 44:464–471, 2016.

[25] C.-K. Chan, J. Hoftrichter, and W. A. Eaton. Optical triggers in protein folding. *Science*, 274:628–629, 1996.

[26] M. Volk. Fast initiation of peptide and protein folding processes. *Eur. J. Org. Chem.*, 14:2605–2621, 2001.

[27] J. Kubelka. Time-resolved methods in biophysics. 9. laser temperature-jump methods for investigating biomolecular dynamics. *Photochem. Photobiol. Sci.*, 8:499–512, 2009.

[28] M.-T. Khuk, L. Menconca, S. Sharma, M. Volk, X. Solinas, and F. Hache. Measurement of circular dichroism dynamics in a nanosecond temperature-jump experiment. *Rev. Sci. Instrum.*, 82:054302, 2011.

[29] G. D. Fasman. *Circular Dichroism and the Conformational Analysis of Biomolecules*. Springer, 1996.

[30] N. Maizels and L. Gray. The G4 genome. *PLoS Genet*, 9, 2013.

[31] N. Maizels. G4-associated human diseases. *EMBO Rep.*, 16:910–922, 2015.

[32] R. Hoffmann, Y. Moshkin, S. Mouton, N. Grzeschik, R. Kalicharan, J. Kuipers, A. Wolters, K. Nishida, A. Romashchenko, J. Postberg, H. Lipps, E. Berezikov, O. C. M. Sibon, B. N. G. Giepmans, and P. M. Lansdorp. Guanine quadruplex structures localize to heterochromatin. *Nucleic Acids Res.*, 44:153–163, 2016.

[33] K. W. Lim, S. Amrane, S. Bouaziz, W. Xu, Y. Mu, D. J. Patel, K. N. Luu, and A. T. Phan. Structure of the human telomere in $K^+$ solution: A stable basket-type G-Quadruplex with only two G-tetrad layers. *J. Am. Chem. Soc.*, 131:4301–4309, 2009.

[34] B. Heddi, N. Martín-Pintado, Z. Serimbetov, T. M. A. Kari, and A. T. Phan. G-quadruplexes with (4n - 1) guanines in the G-tetrad core: formation of a G-triad water complex and implication for small-molecule binding. *Nucleic Acids Res.*, 44:910–916, 2016.

[35] L. A. Yatsunyk, O. Mendoza, and J.-L. Mergny. Nano-oddities: unusual nucleic acid assemblies for DNA-based nanostructures and nanodevices. *Acc. Chem. Res.*, 47:1836–1844, 2014.

[36] S. Neidle and S. Balasubramanian. *Quadruplex Nucleic Acids*. RSC, 2006.

[37] J. L. Huppert and S. Balasubramanian. Prevalence of quadruplexes in the human genome. *Nucleic Acids Res.*, 33:2908–2916, 2005.

[38] D. E. Gilbert and J. Feigon. Multistranded DNA structures. *Curr. Opin. Struct. Biol.*, 9:305–314, 1999.

[39] C. J. Lech, B. Heddi, and A. T. Phan. Guanine base stacking in G-quadruplex nucleic acids. *Nucleic Acids Res.*, 41:2034–2046, 2013.

[40] J. T. Davis. G-quartets 40 years later: From 5-GMP to molecular biology and supramolecular chemistry. *Angew. Chem. Int. Ed.*, 43:668–698, 2004.

[41] T. Fujii, P. Podbevšek, J. Plavec, and N. Sugimoto. Effects of metal ions and cosolutes on G-quadruplex topology. *J. Inorg. Biochem.*, 166:190–198, 2017.

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

[43] G. Parkinson, M. Lee, and S. Neidle. Crystal structure of parallel quadruplexes from human telomeric DNA. *Nature*, 417:876–880, 2002.

[44] K. N. Luu, A. T. Phan, V. Kuryavyi, L. Lacroix, and D. J. Patecl. Structure of the human telomere in $K^+$ solution: an intramolecular (3 + 1) G-quadruplex scaffold. *J. Am. Chem. Soc.*, 128:9963–9970, 2006.

[45] K. W. Lim, V. Chinn Min Ng, B. Martín-Pintado, N. Heddi, and A. T. Phan. Structure of the human telomere in $Na^+$ solution: an antiparallel (2+2) G-quadruplex scaffold reveals additional diversity. *Nucleic Acids Res.*, 41:10556–10562, 2013.

[46] W. J. Chung, B. Heddi, E. Schmitt, K. W. Lim, Y. Mechulam, and A. T. Phan. Structure of a left-handed DNA G-quadruplex. *Proc. Natl. Acad. Sci.*, 112:2729–2733, 2016.

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

[48] N. Q. Do, K. W. Lim, M. H. Teo, B. Heddi, and A. T. Phan. Stacking of G-quadruplexes: NMR structure of a G-rich oligonucleotide with potential anti-HIV and anticancer activity. *Nucleic Acids Res.*, 39:9448–9457, 2011.

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

[50] A. Guédin, A. De Cian, J. Gros, L. Lacroix, and J. L. Mergny. Sequence effects in single-base loops for quadruplexes. *Biochimie*, 90:686–696, 2008.

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

[52] J. L. Huppert. Four-stranded nucleic acids: structure, function and targeting of G-quadruplexes. *Chem. Soc. Rev.*, 37:1375–1384, 2008.

[53] A. N. Lane, J. B. Chaires, R. D. Gray, and J. O. Trent. Stability and kinetics of G-quadruplex structures. *Nucleic Acids Res.*, 36:5482–5515, 2008.

[54] O. Mendoza, J. Elezgaray, and Mergny J.-L. Kinetics of quadruplex to duplex conversion. *Biochimie*, 118:225–233, 2015.

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

[56] C. K. Kwow, M. E. Sherlock, and P. C. Bevilacqua. Effect of loop sequence and loop length on the intrinsic fluorescence of G-quadruplexes. *Biochemistry*, 52:3019–3021, 2013.

[57] X. Cang, J. Šponer, and T. E. Cheatham III. Explaining the varied glycosidic conformational, G-tract length and sequence preferences for anti-parallel g-quadruplexes. *Nucleic Acids Res.*, 39:4499–4512, 2011.

[58] R. Otero, M. Schöck, L. M. Molina, E. Laegsfaard, I. Stensgaard, B. Hammer, and F. Besenbacker. Guanine quartet networks stabilized by cooperative hydrogen bonds. *Angew. Chem. Int. Ed.*, 44:2270–2275, 2005.

[59] E. A. Venczel and D. Sen. Parallel and antiparallel G-DNA structures from a complex telomeric sequence. *Biochemistry*, 32:6220–6228, 1993.

[60] A. Wlodarczyk, P. Grzybowski, A. Patkowski, and A. Dobek. Effect of ions on the polymorphism, effective charge, and stability of human telomeric DNA. photon correlation spectroscopy and circular dichroism studies. *J. Phys. Chem. B*, 109:3594, 2005.

[61] E. Largy, J.-L. Mergny, and V. Gabelica. chapter 7, pages 203–258. Springer International Publishing, 2016.

[62] J. Guab, J. Leszczynski, and M. Bansal. A new insight into the structure and stability of Hoogsteen hydrogen-bonded G-tetrad: an ab initio SCF study. *Chem. Phys. Lett.*, 311:209–214, 1999.

[63] D. Bhattacharyya, G. M. Arachchilage, and S. Basu. Metal cations in G-Quadruplex folding and stability. *Front. Chem.*, 4:38, 2016.

[64] K. Phillips, Z. Dauter, A. I. Murchie, D. M. Lilley, and B. Luisi. The crystal structure of a parallel-stranded guanine tetraplex at 0.95 A resolution. *J. Mol. Biol.*, 17:171–182, 1997.

[65] R. D. Gray and J. B. Chaires. Linkage of cation binding and folding in human telomeric quadruplex DNA. *Biophys. Chem.*, 159:205–209, 2011.

[66] B. G. Kim, H. M. Evans, D. N. Dubins, and T. V. Chalikian. Effects of salt on the stability of a G-Quadruplex from the human c-MYC promoter. *Biochemistry*, 54:3420–3430, 2015.

[67] J. Dai, M. Carver, and D. Yang. Polymorphism of human telomeric quadruplex structures. *Biochimie*, 90:1172–1183, 2008.

[68] N. V. Hud, F. W. Smith, Anet F. A. L., and J. Feigon. The selectivity for $K^+$ versus $Na^+$ in DNA quadruplexes is dominated by relative free energies of hydration: a thermodynamic analysis by 1H NMR. *Biochemistry*, 35:15383, 1996.

[69] J. Gu and J. Leszczynski. A remarkable alteration in the bonding pattern: An HF and DFT study of the interaction between the metal cations and the Hoogsteen hydrogen bonded G-Tetrad. *J. Phys. Chem. A*, 104:6308–6313, 2000.

[70] J. R. Williamson. G-Quartet structures in telomeric DNA. *Annu. Rev. Biophys. Biomol. Struct.*, 23:703, 1994.

[71] W. S. Ross and C. C. Hardin. Ion-induced stabilization of the G-DNA quadruplex: Free energy perturbation studies. *J. Am. Chem. Soc.*, 116:6070, 1994.

[72] B. Sacca, L. Lacroix, and J. L. Mergny. Kinetics of tetramolecular quadruplexes. *Nucleic Acids Res.*, 33:1182, 2005.

[73] R. D. Gray, L. Petraccone, J. O. Trent, and J. B. Chaires. Characterization of a $K^+$-induced conformational switch in a human telomeric DNA oligonucleotide using 2-aminopurine fluorescence. *Biochemistry*, 49:179–194, 2010.

[74] P. Balagurumoorthy and S. K. Brahmachari. Structure and stability of human telomeric sequence. *J. Biol. Chem.*, 269:21858, 1994.

[75] P. Şket and J. Plavec. Tetramolecular DNA quadruplexes in sodium: Insides into structural diversity and cation movement. *J. Am. Chem. Soc.*, 132:12724–12732, 2010.

[76] A. Delville, C. Detellier, and P. Laszlo. Determination of the correlation time for a slowly reorienting spin-3/2 nucleus: binding of $Na^+$ with the 5'-GMP supramolecular assembly. *J. Magn. Reson.*, 34:301–315, 1979.

[77] M. M. Cai, X. D. Shi, D. Sidorov, D. Fabris, Y. F. Lam, and J. T. Davis. Cation-directed self-assembly of lipophilic nucleosides: the cation's central role in the structure and dynamics of a hydrogen-bonded assembly. *Tetrahedron*, 58:661, 2002.

[78] J. S. Lee. The stability of polypurine tetraplexes in the presence of mono- and divalent cations. *Nucleic Acids Res.*, 18:6057–6060, 1990.

[79] N. Nagesh and D. Chatterji. Ammonium ion at low concentration stabilizes the G-quadruplex formation by telomeric sequence. *J. Biochem. Biophys. Meth.*, 30:1–8, 1995.

[80] S. Basu, A. Szewczak, M. Cocco, and S. A. Strobel. Direct detection of monovalent metal ion binding to a DNA G-quartet by 205T NMR. *J. Am. Chem. Soc.*, 122:3240, 2000.

[81] A. Wong and G. Wu. Selective Binding of Monovalent Cations to the Stacking G-Quartet Structure Formed by Guanosine 5-Monophosphate: A Solid-State NMR Study. *J. Am. Chem. Soc.*, 125:13895, 2003.

[82] F. M. Chen. $Sr^{2+}$ facilitates intermolecularG-quadruplex formation of telomeric sequences. *Biochemistry*, 31:3769–2776, 1992.

[83] D. Miyoshi, A. Nakao, T. Toda, and N. Sugimoto. Effect of divalent cations on antiparallel G-quartet structure of d(G4T4G4). *FEBS Lett.*, 496:128–133, 2001.

[84] B. Heddi and A. T. Phan. Structure of human telomeric DNA in crowded solution. *JACS*, 133:9824–9833, 2011.

[85] D. Miyoshi and N. Sugimoto. Molecular crowding effects on structure and stability of DNA. *Biochimie*, 90:1040–1051, 2008.

[86] D. Miyoshi, A. Nakao, and N. Sugimoto. Molecular crowding regulates the structural switch of the DNA G-quadruplex. *Biochemistry*, 41:15017–15024, 2002.

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

[88] V. S. Chambers, G. Marsico, J. M. Boutell, M. Di Antonio, G. P. Smith, and S. Balasubramanian. High-throughput sequencing of DNA G-quadruplex structures in the human genome. *Nat. Biotechnol.*, 33:877–881, 2015.

[89] A. Karsisiotis, C. Okane, and M. Webba da Silva. DNA quadruplex folding formalism–a tutorial on quadruplex topologies. *Methods*, 64:28–35, 2013.

[90] S. A. Ohnmacht and S. Neidle. Small-molecule quadruplex-targeted drug discovery. *Bioorg. Med. Chem. Lett.*, 24:2602–2612, 2014.

[91] L. Crabbe, R. E. Verdun, C. I. Haggblom, and J. Karlseder. Defective telomere lagging strand synthesis in cells lacking WRN helicase activity. *Science*, 306: 1951–1953, 2004.

[92] B. De Nicola, C. J. Lech, B. Heddi, S. Regmi, I. Frasson, R. Perrone, and A. T. Richter, S. N.and Phan. Structure and possible function of a G-quadruplex in the long terminal repeat of the proviral HIV-1 genome. *Nucleic Acids Res.*, 44: 6442–6451, 2016.

[93] J. L. Huppert and S. Balasubramanian. G-quadruplexes in promoters throughout the human genome. *Nucleic Acids Res.*, 35:406, 2007.

[94] D. Sun, B. Thompson, B. E. Cathers, M. Salazar, S. M. Kerwin, J. O. Trent, T. C. Jenkins, S. Neidle, and L. H. Hurley. Inhibition of human telomerase by a G-Quadruplex-Interactive Compound. *J. Med. Chem.*, 40:2113, 1997.

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

[96] D. Drygin, A. Siddiqui-Jain, S. O'Brien, M. Schwaebe, A. Lin, J. Bliesath, C. B. Ho, C. Proffitt, K. Trent, J. P. Whitten, J. K. Lim, D. Von Hoff, K. Anderes, and W. G. Rice. Anticancer activity of CX-3543: a direct inhibitor of rRNA biogenesis. *Cancer Res.*, 69:7653, 2009.

[97] M. Micco, G. W. Collie, A. G. Dale, S. A. Ohnmacht, I. Pazitna, M. Gunaratnam, A. P. Reszka, and S. Neidle. Structure-based design and evaluation of naphthalene diimide G-Quadruplex ligands as telomere targeting agents in pancreatic cancer cells. *J. Med. Chem.*, 56:2959, 2013.

[98] Y. Ma, T.-M. Ou, J.-H. Tan, J.-Q. Hou, S.-L. Huang, L.-Q. Gu, and Z.-S. Huang. Quinolino-benzo-[5, 6]-dihydroisoquindolium compounds derived from berberine: A new class of highly selective ligands for G-quadruplex DNA in c-myc oncogene. *Eur. J. Med. Chem.*, 46:1906–1913, 2011.

[99] K. Shin-ya, K. Wierzba, K. Matsuo, T. Ohtani, Y. Yamada, K. Furihata, Y. Hayakawa, and H. Seto. Telomestatin, a novel telomerase inhibitor from Streptomyces anulatus. *J. Am. Chem. Soc.*, 123:1262–1263, 2001.

[100] T. Doi, K. Shibata, M. Yoshida, M. Takagi, M. Tera, K. Nagasawa, K. Shin-ya, and T. Takahashi. (S)-Stereoisomer of telomestatin as a potent G-quadruplex binder and telomerase inhibitor. *Org. Biomol. Chem.*, 9:387–393, 2011.

[101] D. Monchaud and M.-P. Fichou-Teulade. A hitchhiker's guide to G-quadruplex ligands. *Org. Biomol. Chem.*, 6:627–636, 2008.

[102] S. M. Haider, I. Autiero, and S. Neidle. Surface area accessibility and the preferred topology of telomeric DNA quadruplex-ligand complexes. *Biochimie*, 93:1275–1279, 2011.

[103] M.-Y. Kim, H. Vankayalapati, K. Shin-ya, K. Wierzba, and L. H. Hurley. Telomestatin, a potent telomerase inhibitor that interacts quite specifically with the human telomeric intramolecular G-Quadruplex. *J. Am. Chem. Soc.*, 124: 2098–2099, 2002.

[104] S. L. Forman, J. C. Fettinger, S. Pieraccini, G. Gottarelli, and J. T. Davis. Toward artificial ion channels: A lipophilic G-Quadruplex. *J. Am. Chem. Soc.*, 122:4060–4067, 2000.

[105] X. Shi, K. M. Mullaugh, J. C. Fettinger, Y. Jiang, S. A. Hofstadler, and J. T. Davis. Lipophilic G-Quadruplexes are self-assembled ion pair receptors, and the bound anion modulates the kinetic stability of these complexes. *J. Am. Chem. Soc.*, 125:10830–10831, 2003.

[106] T. Kato. Self-assembly of phase-segregated liquid crystal structures. *Science*, 295:2414–2418, 2002.

[107] Q. W. Xu, H. Deng, and W. H. Braunlin. Selective localization and rotational immobilization of univalent cations on quadruplex DNA. *Biochemistry*, 32: 13130–13137, 1993.

[108] X. X. Zhang, R. M. Izatt, J. S. Bradshaw, and K. E. Krakowiak. Approaches to improvement of metal ion selectivity by cryptands. *Coord. Chem. Rev.*, 174: 179–189, 1998.

[109] R. M. Izatt, K. Pawlak, J. S. Bradshaw, and R. L. Bruening. Thermodynamic and kinetic data for macrocycle interaction with cations, anions and neutral molecules. *Chem. Rev.*, 95:2529–2586, 1995.

[110] T. C. Marsh, J. Vesenca, and E. Henderson. A new DNA nanostructure, the G-wire, imaged by scanning probe microscopy. *Nucleic Acids Res.*, 23:696–700, 1995.

[111] G. I. Livshits, A. Stern, D. Rotem, N. Borovok, G. Eidelshtein, A. Migliore, E. Penzo, S. J. Wind, R. Di Felice, S. S. Skourtis, J. C. Cuevas, L. Gure-

vich, A. B. Kotlyar, and D. Porath. Long-range charge transport in single G-quadruplex DNA molecules. *Nat. Nanotechnol.*, 9:1040–1046, 2014.

[112] V. A. Ngo, R. Di Felice, and S. Haas. Is the G-Quadruplex an effective nanoconductor for ions? *J. Phys. Chem. B*, 118:864–872, 2014.

[113] P. Alberti and Mergny J.-L. DNA duplex–quadruplex exchange as the basis for a nanomolecular machine. *Proc. Natl. Acad. Sci.*, 100:1569–1573, 2003.

[114] J. Zhou, S. Amrane, D. N. Korkut, A. Bourdoncle, H.-Z. He, D.-L. Ma, and J.-L. Mergny. Combination of i-motif and G-Quadruplex structures within the same strand: Formation and application. *Angew. Chem. Int. Ed.*, 125:7896–7900, 2013.

[115] C. R. Cantor and P. R. Schimmel. *Biophysical Chemistry: Part III: Techniques for the Study of Biological Structure and Function*. Oxford, 1981.

[116] J. L. Mergny and L. Lacroix. Analysis of thermal melting curves. *Oligonucleotides*, 13:515–537, 2003.

[117] R. D. Gray, R. Buscaglia, and J. B. Chaires. Dynamics of mismatched base pairs in DNA. *J. Am. Chem. Soc.*, 134:16834–16844, 2012.

[118] M. Yilmaz, C. Ozic, and I. Gok. *Principles of Nucleic Acid Separation by Agarose Gel Electrophoresis - Pirinciples and Basics.* InTech, 2012.

[119] H.-Q. Yu, D. Miyoshi, and N. Sugimoto. Characterization of structure and stability of long telomeric DNA G-quadruplexes. *J. Am. Chem. Soc.*, 128: 15461–15468, 2006.

[120] B. Rupp. *Biomolecular Crystallography: Principles, Practice, and Application to Structural Biology*. Garland Science, 2009.

[121] A. I. Karsisiotis, N. M. Hessari, E. Novellino, G. P. Spada, A. Randazzo, and M. Webba da Silva. Topological characterization of nucleic acid G-Quadruplexes by UV absorption and circular dichroism. *Angew. Chem. Int. Ed.*, 50:10645–10648, 2011.

[122] R. D. Gray and J. B. Chaires. Kinetics and mechanism of $K^+$- and $Na^+$-induced folding of models of human telomeric DNA into G-quadruplex structures. *Nucleic Acids Res.*, 36:4191–4203, 2008.

[123] A. Y. Q. Zhang and S. Balasubramanian. The kinetics and folding pathways of intramolecular G-Quadruplex nucleic acids. *J. Am. Chem. Soc.*, 134:19297–19308, 2012.

[124] J. You, H. Li, X.-M. Lu, W. Li, P.-Y. Wang, S.-X. Dou, and X.-G. Xi. Effects of monovalent cations on folding kinetics of G-quadruplexes. *Bioscience Rep.*, 37:20170771, 2017.

[125] Y. Zhao, Z.-Y. Kan, Z.-X. Zeng, Y.-H. Hao, H. Chen, and T. Zheng. Determining the folding and unfolding rate constant of nucleic acids by biosensor. application to telomere G-Quadruplex. *J. Am. Chem. Soc.*, 126:13255–13264, 2004.

[126] J. J. Green, S. Ladame, L. Ying, D. Klenermann, and S. Balasubramanian. Investigating a quadruplex-ligand interaction by unfolding kinetics. *J. Am. Chem. Soc.*, 128:9809–9812, 2006.

[127] T. Mashimo and H. Sugiyama. Folding pathways of human telomeric hybrid G-quadruplex structure. *Nucleic Acids Symp Ser*, 51:239–240, 2007.

[128] Y. Li, C. Liu, X. Feng, Y. Xu, and B. F. Liu. Ultrafast microfluidic mixer for tracking the early folding kinetics of human telomere G-quadruplex. *Ana. Chem.*, 86:4333–4339, 2014.

[129] L. Ying, J. J. Green, H. Li, D. Klenerman, and S. Balasubramanian. Studies on the stucture and dynamics of the human telomeric G quadruplex by single-molecule fluorescence resonance energy transfer. *Proc. Natl. Acad. Sci.*, 100: 14629–14634, 2003.

[130] J. Y. Lee, B. Okumus, D. S. Kim, and T. Ha. Extreme conformational diversity in human telomeric DNA. *Proc. Natl. Acad. Sci.*, 102:18938–18943, 2005.

[131] M. Aznauryan, S. Søndergaard, S. L. Noer, B. Schiøtt, and V. Birkedal. A direct view of the complex multi-pathway folding of telomeric G-quadruplexes. *Nucleic Acids Res.*, 44:11024–11032, 2016.

[132] K. Okamoto, Y. Sannohe, T. Mashimo, H. Sugiyama, and M. Terazima. G-quadruplex structures of human telomere DNA examined by single molecule FRET and BrG-substitution. *Bioorg. Med. Chem.*, 16:6873–6879, 2008.

[133] X. Long and M. D. Stone. Kinetic partitioning modulates human telomere DNA G-Quadruplex structural polymorphism. *PLOS*, 8:e83420, 2013.

[134] D. Koirala, T. Mashimo, Y. Sannohe, Z. Yu, H. Mao, and H. Sugiyama. Intramolecular folding in three tandem guanine repeats of human telomeric DNA. *Chem. Commun.*, 48:2006–2008, 2012.

[135] Z. Yu, D. Koirala, Y. Cui, L. F. Easterling, Y. Zhao, and H. Mao. Click chemistry assisted single-molecule fingerprinting reveals a 3D biomolecular folding funnel. *J. Am. Chem. Soc.*, 134:12338–12341, 2012.

[136] W. Li, X.-H. Hou, P.-Y. Wang, X.-G. Xi, and M. Li. Direct measurement of sequential folding pathway and energy landscape of human telomeric G-quadruplex structures. *J. Am. Chem. S.*, 135:6423–6426, 2013.

[137] H. You, X. Zeng, Y. Xu, C. J. Lim, A. K. Efremov, A. T. Phan, and J. Yan. Dynamics and stability of polymorphic human telomeric G-quadruplex under tension. *Nucleic Acids Res.*, 42:8789–8795, 2014.

[138] D. Koirala, C. Ghimire, C. Bohrer, Y. Sannohe, Z. Yu, H. Sugiyama, and H. Mao. Long-loop G-quadruplexes are misfolded population minorities with fast transition kinetics in human telomeric sequences. *J. Am. Chem. Soc.*, 136: 2235–2241, 2013.

[139] Z. Yu, V. Gaerig, Y. Cui, H. Kang, V. Gokhale, Y. Zhao, L. H. Hurley, and H. Mao. Tertiary DNA structure in the single-stranded hTERT promoter fragment unfolds and refolds by parallel pathways via cooperative or sequential events. *J. Am. Chem. Soc.*, 134:5157–5184, 2012.

[140] P. Stadlbauer, M. Krepl, T. E. Cheatham III, J. Koća, and J. Šponer. Structural dynamics of possible late-stage intermediates in folding of quadruplex DNA studied by molecular simulations. *Nucleic Acids Res.*, 41:7128–7143, 2013.

[141] P. Stadlbauer, L. Trantiírek, T. E. Cheatham III, J. Koća, and J. Šponer. Triplex intermediates in folding of human telomeric quadruplexes probed by microsecond-scale molecular dynamics simulations. *Biochimie*, 105:22–35, 2014.

[142] P. Stadlbauer, P. Kürová, Banášn P., J. Koča, G. Bussi, L. Trantírek, M. Otyepka, and J. Šponer. Hairpins participating in folding of human telomeric sequence quadruplexes studied by standard and T-REMD simulations. *Nucleic Acids Res.*, 43:9626–9644, 2015.

[143] Gajarský M., M. L. Živković, P. Stadlbauer, B. Pagano, R. Fiala, J. Amato, L. Tomáška, J. Šponer, J. Plavec, and Trantírek L. Structure of a stable G-Hairpin. *J. Am. Chem. Soc.*, 139:3591–3594, 2017.

[144] L. Cerofolini, J. Amato, A. Giachetti, V. Limongelli, E. Novellino, M. Parrinello, M. Fragai, A. Randazoo, and C. Luchinat. G-triplex structure and formation propensity. *Nucleic Acids Res.*, 42:13393–13404, 2014.

[145] P. Stadlbauer, L. Mazzanti, T. Cragnolini, D. J. Wales, P. Derreumaux, S. Pasquali, and J. Šponer. Coarse-grained simulations complemented by atomistic molecular dynamics provide new insights into folding and unfolding of human telomeric G-Quadruplexes. *J. Chem. Theory Comput.*, 12:6077–6097, 2016.

[146] H. Ma, C. Wan, and A. H. Zewail. Ultrafast T-jump in water: studies of conformation and reaction dynamics at the thermal limit. *J. Am. Chem. Soc.*, 128:6338–6340, 2006.

[147] N. Berova, K. Nakanishi, and R. W. Woody. *Circular dichroism: principles and applications, ed.* Wiley-VCH, 2000.

[148] N. Berova, L. Di Bari, and G. Pescitelli. Application of electronic circular dichroism in configurational and conformational analysis of organic compounds. *Chem. Soc. Rev.*, 36:914, 2007.

[149] G. Gottarelli, S. Lena, S. Masiero, S. Pieraccini, and G. P. Spada. The use of circular dichroism spectroscopy for studying the chiral molecular sefl-assembly: An overview. *Chirality*, 20:471, 2008.

[150] S. Masiero, R. Trotta, S. Pieraccini, S. De Tito, R. Perone, A. Randazzo, and G. P. Spada. A non-empirical chromophoric interpretation of CD spectra of DNA G-quadruplex structures. *Org. Biomol. Chem.*, 8:2683–2692, 2010.

[151] J. Meyer-Ilse, D. Akimov, and B. Dietzek. Recent advances in ultrafast time-resolved chirality measurements: perspective and outlook. *Laser Photonics Rev.*, 7:495–505, 2013.

[152] J. W. Lewis, R. A. Goldbeck, X. Kliger, D. S. anx Xie, R. C. Dunn, and J. D. Simon. Time-resolved circular dichroism spectroscopy: Experiment, theory, and applications to biological systems. *J. Phys. Chem.*, 96:5243–5254, 1992.

[153] A. Randazzo, G. P. Spada, and M. W. da Silva. pages 67–86. Top Curr Chem, 2013.

[154] H. De Voe. Optical properties of molecular aggregates. I. classical model of electronic absorption and refraction. *J. Chem. Phys.*, 41:393–400, 1964.

[155] H. De Voe. Optical properties of molecular aggregates. II. classical theory of the refraction, absorption and optical activity of solutions and crystals. *J. Chem. Phys.*, 43:3199–3208, 1965.

[156] L. B. Clark. Electronic spectra of crystalline guanosine: transition moment directions of the guanine chromophore. *J. Am. Chem. Soc.*, 116:5265–52700, 1994.

[157] M. P. Füelscher, L. Serano-Andrés, and B. O. Roos. A theoretical study of the electronic spectra of adenine and guanine. *J. Am. Chem. Soc.*, 119:6168–6176, 1997.

[158] G. Gottarelli, P. Palmieri, and G. P. Spada. The exciton optical activity of the four-stranded helix of poly(G). *Gazz. Chim. Ital.*, 120:101–107, 1990.

[159] D. M. Gray, J. D. Wen, C. W. Gray, R. Repges, C. Repges, G. Raabe, and J. Fleischhauer. Measured and calculated CD spectra of G-quartets stacked with the same or opposite polarities. *Chirality*, 20:431–440, 2008.

[160] A. Fersht. *Structure and mechanism in protein science*. W. H. Freeman and Company, 1998.

[161] L. Mendonca. *Dynamique Conformationelle des Protéines étudiée par Dichroïsme Circulaire Résolu en Temps*. PhD thesis, Ecole Polytechnique, 2013.

[162] J. E. Bertie and Z. Lan. Infrared intensities of liquids XX: The intensity of the OH stretching band of liquid water revisited, and the best current value of the optical constants of H2O(l) at 25 deg between 15,000 and 1 cm-1. *Appl. Spectrosc.*, 8:1047–1057, 1996.

[163] B. Nölting. *Protein folding kinetics. Biophyscial methods*. Springer, 2006.

[164] H. T. Le, R. Buscaglia, W. L. Dean, J. B. Chaires, and J. O. Trent. pages 179–210. Top Curr Chem, 2013.

[165] S. Lindsay. *Introduction to Nanoscience*. OUP Oxford, 2009.

[166] J. Kestin, M. Sokolov, and W. A. Wakeham. Viscosity of liquid water in the range of -8 deg to 150 deg. *J. Phys. Chem. Ref. Data*, 7:941–948, 1978.

[167] D. R. Lide. *Index of refraction of water/ Handbook of Chemistry and Physics*. CRC Press LLC, 2001-2002.

[168] J. Meyer-Ilse, D. Akimov, and B. Dietzek. Ultrafast circular dichroism study of the ring opening of 7-Dehydrocholesterol. *J. Phys. Chem. Lett.*, 3:182–185, 2012.

[169] L. Mendonca and F. Hache. Nanosecond T-jump experiment in poly(glutamic acid): a circular dichroism study. *Int. J. Mol. Sci.*, 13:2239–2248, 2012.

[170] A. Trifonov, I. Buchvarov, A. Lohr, F. Würthner, and T. Fiebig. Broadband femtosecond circular dichroism spectrometer with white-light polarization control. *J. Phys. Chem. Lett.*, 81:043104–1–043104–8, 2010.

[171] A. Steinbacher, H. Hildenbrand, S. Schott, J. Buback, M. Schmid, P. Nürnberger, and T. Brixner. Generating laser-pulse enantiomers. *Opt. Express*, 25: 21735–21752, 2017.

[172] F. C. A. Pockels. Ueber den Einfluss elastischer Deformationen, speciell einseitigen Druckes, auf das optische Verhalten krystallinischer Körper. *Ann. Phys. (Berlin)*, 273:144–172, 1889.

[173] R. C. Jones. A new calculus for the treatment of optical systems I. description and discussion of the calculus. *J. Opt. Soc. Am.*, 31:488–493, 1941.

[174] S. Akiyama, S. Takahashi, K. Ishimori, and I. Morishima. Stepwise formation of a-helices during cytochrome c folding. *Nat. Struct. Biol.*, 7:514–520, 2000.

[175] R. Improta. Quantum mechanical calculations unveil the structure and properties of the absorbing and emitting excited electronic states of guanine quadruplexes. *Chem. Eur. J.*, 20:8106–8115, 2014.

[176] D. Loco, S. Jurinovich, L. D. Di Bari, and B. Mennucci. A fast but accurate excitoni simulation of the electronic circular dichroism of nucleic acids: how can it be achieved? *Phys. Chem. Chem. Phys.*, 18:886, 2016.

[177] H. Gattuso, A. Spinello, A. Terenzi, X. Assfeld, G. Barone, and A. Monari. Circular dichroism of DNA G-Quadruplexes: Combining modeling and spectroscopy to unravel complex structures. *J. Phys. Chem. B*, 22:17011–17019, 2016.

[178] A. I. S. Holm, B. Kohler, S. V. Hoffmann, and S. B. Nielsen. Synchrotron radiation circular dichroism of various G-quadruplex structures. *Biopolymers*, 93:429–433, 2010.

[179] A. I. S. Holm, B. Kohler, S. V. Hoffmann, and S. B. Nielsen. Vacuum-ultraviolet circular dichroism spectroscopy of DNA: a valuable tool to elucidate topology and electronic coupling in DNA. *Biopolymers*, 12:9581–9596, 2010.

[180] L. C. Bock, L. C. Griffin, J. A. Latham, E. H. Vermaas, and J. J. Toole. Selection of single-stranded DNA molecules that bind and inhibit human thrombin. *Nature*, 355:564–566, 1992.

[181] A. Aviñó, C. Fàbrega, M. Tintoré, and R. Eritja. Thrombin binding aptamer, more than a simple aptamer: chemically modified derivatives and biomedical applications. *Curr. Pharm. Des.*, 12:1873–4286, 2012.

[182] T. Mashimo, Y. Sannohe, H. Yagi, and H. Sugiyama. Folding pathways of hybrid-1 and hybrid-2 G-quadruplex structures. *Nucleic Acids Symp. Ser.*, 52: 409–410, 2008.

[183] C. Krejtschi and K. Hauser. Stability and folding dynamics of polyglutamic acid. *Eur. Biophys. J.*, pages 673–685, 2011.

[184] R. W. Henry and J. Hoftrichter. Singular value decomposition: application to analysis of experimental data. pages 129–191. Academic Press, New York, 1992.

[185] R. D. Gray and J. B. Chaires. Analysis of multidimensional G-Quadruplex melting curves. *Curr. Protoc. Nucleic Acid Chem.*, page 17.4, 2011.

[186] C. A. Sprecher and W. C. Johnson Jr. Circular dichroism of the nucleic acids monomer. *Biopolymers*, 16:2243–2264, 1977.

[187] M. Bončina, G. Vesnaver, J. B. Chaires, and J. Lah. Unraveling the thermodynamics of the folding and interconverstion of human telomere G-Quadruplex. *c*, 55:10340–10344, 2016.

[188] A. Ansari, s; V. Kuznetsov, and Y. Shen. Configurational diffusion down a folding funnel describes the dynamics of DNA hairpins. *Proc. Natl. Acad. Sci.*, 98:7771–7776, 2001.

[189] P. Hamm, J. Helbing, and J. Bredenbeck. Stretched versus compressed exponential kinetics in alpha-helix folding. *Chem. Phys.*, 323:54–65, 2006.

[190] G. Portella and M. Orozco. Multiple routes to characterize the folding of a small DNA hairpin. *Angew. Chem. Int. Ed.*, 49:7673–7676, 2010.

[191] M. I. Wallace, L. Ying, S. Balasubramanian, and D. Klenerman. Non-Arrhenius kinetics for the loop closure of a DNA hairpin. *Proc. Natl. Acad. Sci.*, 98:5584–5589, 2001.

[192] M. Volk, L. Milanesi, J. P. Waltho, C. A. Hunter, and G. S. Beddard. The roughness of the protein energy landscape results in anomalous diffusion of the polypeptide backbone. *Phys. Chem. Chem. Phys.*, 17:762–782, 2015.

[193] C. H. Mak. Unraveling base stacking driving forces in DNA. *J. Phys. Chem. B*, 120:6010–6020, 2016.

[194] Z. J. Zhang, J. X. Dai, E. Veliath, R. A. Jones, and D. Z. Yang. Structure of a two-g-tetrad intramolecular G-quadruplex formed by a variant human telomeric sequence in $K^+$ solution: insight into the interconversion of human telomeric G-quadruplex structure. *Nucleic Acids Res.*, 38:1009–1021, 2010.

[195] C. Pérez-Arnáiz, N. Busto, J. M. Leal, and B.s García. New microsecond intramolecular reactions of human telomeric DNA in solution. *RSC Adv.*, 6: 39204–39208, 2016.

[196] X. Zeng, L. Zhang, X. Xiao, Y. Jiang, Y. Guo, X. Yu, X. Pu, and M. Li. Unfolding mechanism of thrombin-binding aptamer revealed by molecular dynamics simulation and Markov state model. *Sci. Rep.*, 6:24065, 2016.

[197] H. You, J. Wu, F. Shao, and J. Yan. Stability and kinetics of c-MYC promoter G-Quadruplexes studied by single-molecule manipulation. *J. Am. Chem. Soc.*, 137:2424–2427, 2015.

[198] A. Siddiqui-Jain, C. L. Grand, D. J. Bearss, and I. Hurley. Direct evidence for a G-quadruplex in a promoter region and its targeting with a small molecule to repress c-MYC transcription. *Proc. Natl. Acad. Sci.*, 99:11593, 2002.

[199] P. S. Shirude, L. Ying, and S. Balasubramanian. Single molecule conformational analysis of the biologically relevant DNA G-quadruplex in the promoter of the proto-oncogene c-MYC. *Chem. Commun.*, pages 2007–2009, 2008.

[200] A. J. Stevens, H. L. Kennedy, and M. A. Kennedy. Fluorescence methods for probing G-Quadruplex structure in single- and double-stranded DNA. *Biochemistry*, 55:3714–3725, 2016.

[201] D. Koirala, T. Mashimo, Y. Sannohe, Z. Yu, H. Mao, and H. Sugiyama. Intramolecular folding in three tandem guanine repeats of human telomeric DNA. *Chem. Commun.*, 48:2006–2008, 2011.

[202] X. Long, J. W. Parks, C. R. Bagshaw, and C. J. Sto. Mechanical unfolding of human telomere G-quadruplex DNA probed by integrated fluorescence and magnetic tweezers spectroscopy. *Nucleic Acids. Res.*, 41:2746, 2013.

[203] S. v Ceru, P. Šket, I. Prsilan, J. Lah, and J. Plavec. A new pathway of DNA G-Quadruplex formation. *Angew. Chem. Int. Ed.*, 126:4981–4984, 2014.

[204] N. T. Dao, R. Haselsberger, M.-E. Michel-Beyerle, and A. T. Phan. Following G-quadruplex formation by its intrinsic fluorescence. *FEBS Lett.*, 585:3969–3977, 2011.

**Titre :** Dynamiques conformationnelles de l'ADN G-quadruplex sondées par dichroïsme circulaire résolu en temps

**Mots clés :** G-quadruplexes, dichroïsme circulaire résolu en temps, T-jump, dynamique de repliement

**Résumé** : Les quadruplexes de guanines (G4) sont des structures d'ADN non-canoniques qui résultent de l'empilement hydrophobe de tétrades de guanines, stabilisé par des cations métalliques (tels que $Na^+$ et $K^+$). Il existe aujourd'hui un nombre croissant de preuves expérimentales qui attestent de l'implication des G4 dans d'importantes fonctions cellulaires corrélées à leur mécanisme de repliement/déplement. Toutefois, très peu d'études ont abordé les aspects dynamiques de leur formation. Aussi, nous avons entrepris l'étude de plusieurs G4 mono-moléculaires à l'aide d'une nouvelle extension d'expériences de saut de température, capable de mesurer la dynamique de dénaturation thermique et de renaturation consécutive de l'ADN, sur une fenêtre temporelle allant de quelques millisecondes aux secondes. Les changements conformationnels ont été sondés par dichroïsme circulaire (CD) résolu en temps, connu pour être très sensible à l'arrangement des guanines dans les G4.
Au préalable des études résolues en temps, en collaboration avec DISCO/SOLEIL, nous avons mesuré les spectres CD statiques de différentes séquences G4 présentant des topologies distinctes, comme celles des télomères humains, de l'aptamère de la thrombine ou des promoteurs de c-MYC. Nous avons observé des cinétiques de dénaturation et renaturation biphasiques avec des constantes de temps de quelques centaines de millisecondes et quelques secondes. Ces cinétiques dépendent fortement de l'amplitude du saut de température et de la concentration de cations métalliques. L'ensemble de ces observations suggère l'existence de plusieurs voies de repliement/déplement des G4 sur des surfaces de potentiel très rugueuses.

**Title :** Conformational dynamics of G-quadruplex DNA probed by time-resolved circular dichroism

**Keywords :** G-quadruplexes, time-resolved circular dichroism, T-jump, folding dynamics

**Abstract**: Guanine-quadruplexes (G4) are non-canonical DNA structures that result from the hydrophobic stacking of guanine quartets stabilized by metal cations (typically $Na^+$ and $K^+$). There is now an increasing body of experimental evidence of their occurrence in important cell functions correlated to their folding/unfolding mechanisms. However, only few studies have addressed the dynamical aspect of their formation. In this context, we have undertaken the study of several intramolecular G4 with a novel extension of temperature-jump experiments capable to measure the thermal denaturation and the consecutive renaturation of DNA over a time window spanning a few ten milliseconds to seconds. Conformational changes have been monitored by time-resolved circular dichroism (CD), which is known to be sensitive to the chiral arrangement of guanines in the G4 scaffolds.
Prior to time-resolved measurements, within the frame of a collaboration with DISCO/SOLEIL, we have performed static synchrotron radiation CD measurements on several short G4-forming sequences, such as human telomere, thrombin-binding aptamer and c-MYC promoter sequences, displaying distinct topologies. Denaturation and renaturation kinetics are found to exhibit biphasic decays with time constants of a few hundred milliseconds and a few seconds, respectively. Those kinetics depend strongly on the amplitude of the temperature jump and the concentration of cations. Taken together these observations suggest the existence of multiple folding pathways on extremely rugged landscapes.