



**FACULTY
OF MATHEMATICS
AND PHYSICS**
Charles University

BACHELOR THESIS

Anton Bokach

**Concentration-dependent topology of
telomere DNA quadruplexes**

Department of Low Temperature Physics

Supervisor of the bachelor thesis: Mgr. Václav Římal, Ph.D.

Study programme: Physics

Study branch: General Physics

Prague 2020

I declare that I carried out this bachelor thesis independently, and only with the cited sources, literature and other professional sources. It has not been used to obtain another or the same degree.

I understand that my work relates to the rights and obligations under the Act No. 121/2000 Sb., the Copyright Act, as amended, in particular the fact that the Charles University has the right to conclude a license agreement on the use of this work as a school work pursuant to Section 60 subsection 1 of the Copyright Act.

In date

Author's signature

“La chance ne sourit qu’aux esprits bien préparés.” Louis Pasteur

Title: Concentration-dependent topology of telomere DNA quadruplexes

Author: Anton Bokach

Department: Department of Low Temperature Physics

Supervisor: Mgr. Václav Římal, Ph.D., Department of Low Temperature Physics

Abstract: Telomeres contain repetitive DNA sequences and play a key role in cellular ageing and cancer. Human telomeric DNA contains guanine-rich sequences, such as $(GGGTTA)_n$, that are capable of four-stranded quadruplex structures assembling. Quadruplexes can adopt various topologies depending on the ionic environment, concentration, thermal treatment and other factors. Those topologies include parallel, anti-parallel and hybrid (3+1) types.

We explored the effect of dilution and annealing of two human telomere sequences (21 and 22 nucleotide long) in a potassium-containing solution by nuclear magnetic resonance (NMR) and Raman spectroscopies. We found that multimolecular complexes with parallel orientation of the DNA strands, which are formed at high DNA concentration (≈ 10 mM), persist even after strong dilution. They are disrupted only by thermal annealing. Furthermore, annealing led to the creation of a unimolecular G-quadruplex that appeared to be in (3 + 1) conformation. The (3 + 1) conformation of G-quadruplex is the most populated monomeric quadruplex structure present in the solution, regardless of oligonucleotide concentration.

We examined the differences between quadruplex structures in crowded and diluted solutions, as well as, before and after annealing. The presented thesis brings new insights about dependence of the quadruplex type on ionic composition of the solvent and oligonucleotide concentration.

Keywords: DNA, Quadruplex, NMR spectroscopy, Raman spectroscopy

Contents

Introduction	3
1 Nucleic acids and nucleotides	4
1.1 Nucleotides and nucleic acids	4
1.2 DNA and its structure	5
1.3 G-quadruplex	7
1.4 Telomere structures	8
2 Experimental Methods	11
2.1 NMR spectroscopy	11
2.1.1 NMR essence and resonance frequency	11
2.1.2 Chemical shifts and nuclear shielding	12
2.1.3 Bloch Equations	12
2.1.4 Free Precession	13
2.1.5 Radiofrequency pulses and spin echo	14
2.1.6 Water signal suppression	15
2.1.7 Dipole-dipole interaction	16
2.2 Raman spectroscopy	17
2.2.1 Introduction to the general idea	17
2.2.2 Raman Effect	18
3 Materials and Experiment description	19
3.1 Samples	19
3.2 NMR experiment	20
3.2.1 Pulse experiments	21
3.3 Determination of concentrations	21
3.4 Raman experiment	22
4 Results	24
4.1 NMR experiment results	24
4.1.1 Tel22 nucleotide sequence	24
4.1.2 Tel21 nucleotide sequence	32
4.2 Raman experiment results	40
5 Discussion	44
5.1 Tel22 NMR experiment	44
5.2 Tel21 NMR experiment	45
5.3 Tel22 Raman experiment	46
5.4 G4 topologies observed	46
Conclusion	47
Bibliography	48
List of Figures	51

List of Tables	54
List of Abbreviations	55

Introduction

Oligonucleotides are frequently used for live organisms' studies, as they represent a short DNA or RNA sequence and have certain similar qualities and features of naturally-long DNAs. One of the notable types of nucleotide sequences in DNA molecules are telomeres, that are proven to serve multiple purposes. Along with protectional purpose, they play a key role in cellular ageing, cancer or hereditary diseases. Every cell replication costs a telomere 50-200 bases. After reaching critical length, DNA experiences programmed cell death. However, telomeres can be elongated by the enzyme telomerase. Telomerase activity was not observed in healthy somatic cells, but it exhibits activity in a vast majority of cancer cells, making them, essentially, immortal. Telomeric DNA, rich in guanine, forms quadruplex structures of various topologies depending on the ionic environment, concentration, thermal treatment and other factors. Inhibition of telomerase activity may be succeeded by intramolecular G-quadruplexes assembling, which is a reason for searching for ligand capable of intramolecular telomeric G-quadruplex stabilization.

Studies of oligonucleotides folding topologies has become an appealing direction of a research in the last few decades. However, due to G-quadruplex topology dependency on an impressive set of parameters, many telomeric DNA structures are still far from being solved.

There are multiple spectroscopy technics, including NMR, Raman, or CD spectroscopy that contribute to different oligonucleotides folding topologies solutions. According to the previously conducted experiments, all the spectroscopy techniques mentioned above were proven of the G-quadruplex structure elucidation. The main goal of this thesis is to bring new insights regarding the dependence of the G-quadruplex structure type on the ionic composition of the solvent and oligonucleotide concentration with a special interest in the differences caused by annealing. The nuclear magnetic resonance (NMR) spectroscopy and Raman spectroscopy will be used in this thesis in order to attempt answering those questions.

1. Nucleic acids and nucleotides

1.1 Nucleotides and nucleic acids

Nucleotides play key roles in metabolism, one of which is being an energy currency in the metabolic transactions of a cell. Furthermore, nucleotides are the building blocks of the nucleic acids, DNA and RNA, that are widely known as storage units of genetic information. It is known that every species has to transmit this genetic information further in order for the following generations to persist and survive. The nucleotide sequences of a cell's nucleic acids dictate the topological features of each protein and other cellular components. More precisely, every protein's amino acid sequence, likewise every RNA nucleotide sequence, is rooted and coded in the cell's DNA nucleotide sequence. A segment of a DNA molecule containing instructions for functional biological product synthesis is called a gene. Cells usually contain thousands of genes. Along with its frequent subsequent usage to produce other proteins, RNA is assigned to execute numerous functions. Some RNAs are involved in protein synthesis, others are involved in post-transcriptional modification, etc. For instance, mRNA carries genetic information from the cell's nucleus to the cytoplasm and transcribes the genetic code from DNA into a form that can be read and used for protein production, whereas, the main goal of transfer RNA is to bring amino acids to ribosomes during the translation process.

A nucleotide is a molecule composed of a nucleoside with one or several phosphate groups attached. A nucleoside is a molecule that consists of aldopentose sugar bonded to a nitrogen atom of a nitrogenous base (a heterocyclic purine or a pyrimidine base) through the sugar's anomeric carbon. The mentioned sugars can be either ribose or deoxyribose (see Figure 1.1).

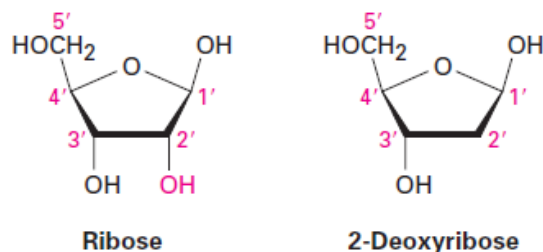


Figure 1.1: Major pyrimidine and purine bases of nucleic acids (adapted from [1])

We define nucleic acids as biopolymers consisting of nucleotides joined together by the formation of covalent phosphodiester bonds between the 5'-phosphate group attached to the sugar of one nucleotide unit and a 3'-hydroxyl group on the sugar of the next nucleotide unit in order to form a long nucleotide chain. Since phosphodiester linkages have the same orientation within the linear nucleotide chain, they result in the polarity of a strand and the distinctive 5' and 3' ends of the strand.

RNA aldopentose sugar is ribose, DNA assigned sugar is 2-deoxyribose. DNA contains four different nitrogen-containing nucleobases, two substituted purines (adenine and guanine) and two substituted pyrimidines (cytosine and thymine).

Adenine, guanine, and cytosine also occur in RNA, but thymine is replaced in RNA by a closely related pyrimidine base called uracil. Nevertheless, cases of thymine occurrence in RNA, together with uracil occurrence in DNA, are present [1]. Major pyrimidine and purine bases of nucleic acids are shown in Figure 1.2.

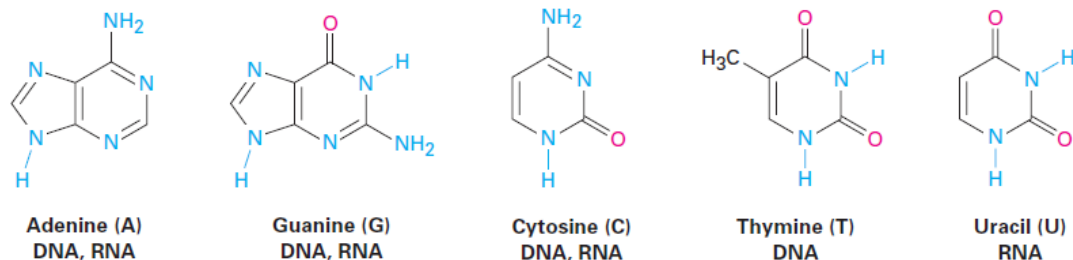


Figure 1.2: Major pyrimidine and purine bases of nucleic acids (adapted from [1])

1.2 DNA and its structure

The first successful DNA isolation was executed by Friedrich Miescher in 1868. An important contribution towards DNA structure understanding came from Erwin Chargaff, who has formulated four quantitative relationships, based on his experiments and observations, later named Chargaff rules:

1. The base composition of DNA generally varies from one species to another.
2. DNA specimens isolated from different tissues of the same species have the same base composition.
3. The base composition of DNA in a given specie does not change with an organism's age, nutritional state, or changing environment.
4. In all cellular DNAs, regardless of the species, the number of adenosine residues is equal to the number of thymidine residues (that is, $A = T$), and the number of guanosine residues is equal to the number of cytidine residues ($G = C$). From these relationships it follows that the sum of the purine residues equals the sum of the pyrimidine residues; that is, $A+G = T+C$. [2]

In 1953 James Watson and Francis Crick proposed their three-dimensional DNA structure model, which was composed of two helical complementary polynucleotide chains curling around the same axis, thereby assembling a right-handed double helix [3]. The two polynucleotide chains run in opposite directions and are kept together via hydrogen bonds formed between 2 pairs of bases, precisely Adenine with Thymin and Cytosine with Guanin. Each A-T pair forms two hydrogen bonds, while each G-C pair forms three hydrogen bonds. That is the essence of the mentioned complementary rule. The DNA double helix molecule structurally has internal and external parts. The internal part is represented by tightly located hydrophobic planar ring structures of pyrimidine and purine bases,

that are perpendicular to the common axis, whereas the external part consists of alternating deoxyribose and phosphate groups forming a hydrophilic backbone.

With 10 base pairs per one turn, the double helix is 20 Å wide. Base pairing in accordance to the complementary rule results in the formation of so-called grooves, with the major groove being 12 Å wide and the minor groove being 6 Å wide. Along with the hydrogen bonding between two pairs of bases, there is, however, another weak interaction, that plays the key role in double helix's stability establishment – base-stacking interaction. A DNA molecule may occur in distinctive three-dimensional forms due to its outstanding flexibility. The first factor that maintains the DNA flexibility is thermal fluctuation, which in particular cases may provide the extension, bending or even melting of a polynucleotide chain. The second factor that can be pointed out is the rotation possibility around several bonds in a sugar–phosphate backbone. Despite structural variations, key properties of DNA (according to Watson and Crick) are not affected. Atypical DNA structures can assemble, if instead of two DNA strands, three or four strands are present. Watson–Crick base pair nucleotides are capable of establishing further hydrogen bonds, notably with major groove's functional groups. Nevertheless, Watson–Crick base pairing is not the only pairing type possible. So-called Hoogsteen positions, namely N-7 and the C-6 amino group of purines, in case of A-T pair, are in charge of triplex DNA hydrogen-bonding formation. This pairing is referred to as Hoogsteen pairing. In Figure 1.3 the essential differences of Watson–Crick and Hoogsteen pairing are highlighted.

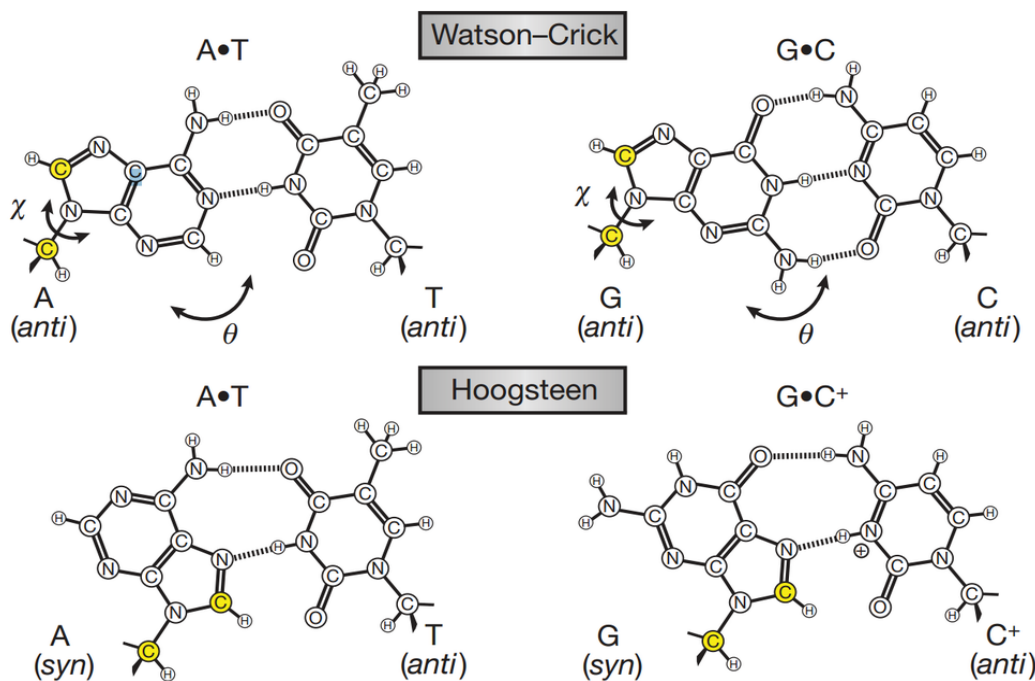


Figure 1.3: Watson–Crick and Hoogsteen A-T and G-C base pair structures. The Hoogsteen geometry can be achieved by the purine rotation around the glycosidic bond (χ) and base-flipping (θ), affecting simultaneously C-8 and C-1' (coloured yellow) (adapted from [4])

1.3 G-quadruplex

If a large number of guanosine residues are present in four DNA polynucleotide chains, then they might pair in order to construct a quadruplex, also known as the G-tetraplex. The quadruplex is a considerably stable structure even under various physical impacts, also with the possibility of polynucleotide chains being in a variety of orientations. When exposed to an ionic environment, composed of K^+ and Na^+ G-rich DNA sequences are likely to establish quadruplex constructions in vitro, being assembled from several planar G-G-G-G structures called tetrads, pictured in Figure 1.4 [5]. In figure 1.4, on the left side, we may also see a base-pairing pattern in the guanosine quadruplex structure and on the right side we see possible options of nucleotide chains' orientations in a G-quadruplex.

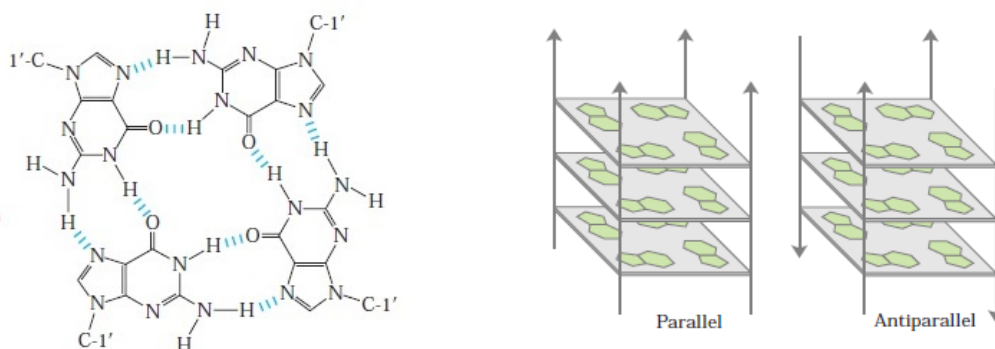


Figure 1.4: Left - Guanosine quadruplex base-pairing pattern; Right - Guanosine quadruplex's possible nucleotide chains' orientations (adapted from [2])

A large amount of DNA is required for encoding genetic information for a single-celled bacterium or multi-cellular organism to be produced. As we know, double helix DNA molecules in eukaryotic cells are arranged into chromosomes, which are capable of being replicated and later being equally divided between daughter cells during the cell cycle. The DNA molecule replication occurs during the cell cycle's interphase. There are two notable types of nucleotide sequences in DNA molecules that are designed for replication process efficiency. The first type that serves as a DNA replication start is called replication origin. The second type, telomeres, is in charge of marking chromosome ends. Telomeres are composed of repeated DNA sequences, mainly $(GGGTTA)_n$ for humans, and are believed to have a protective purpose along with playing an important role in cellular ageing and cancer. Human telomeric DNA has a typical length of 5-8 thousand base pairs with a 3' single-stranded overhang of 100-200 nucleotides. With every replication 50-200 telomere bases are lost. If telomeric DNA reaches critical length, the cell later experiences apoptosis or programmed cell death [6]. However, there exists an enzyme called telomerase that is capable of telomere elongation. This enzyme, nonetheless, is deactivated in normal human somatic cells, while exposing activity in 80-85 % of cancer cells [7]. Inhibition of telomerase activity may be succeeded by intramolecular G-quadruplexes assembling, by dint of telomeric nucleotide sequences with guanine excess [8, 9, 10]. Thus, seeking for a stabilizing ligand of the previously described construction is a captivating approach for anticancer drug evolution.

1.4 Telomere structures

In a Na^+ solution, a four-repeat human telomeric nucleotide sequence $\text{d}[\text{AGGG}(\text{TTAGGG})_3]$ assembles an unimolecular G-quadruplex, where tetrad surrounding guanines are syn-syn-anti-anti and loops are consecutively edgewise-diagonal-edgewise¹. Furthermore, every G-tract neighbours both parallel and antiparallel strands[12]. The described G-quadruplex topological conformation is called antiparallel basket conformation, also illustrated in figure 1.5 (a). Nevertheless, there is a second antiparallel G-quadruplex conformation called antiparallel chair conformation.

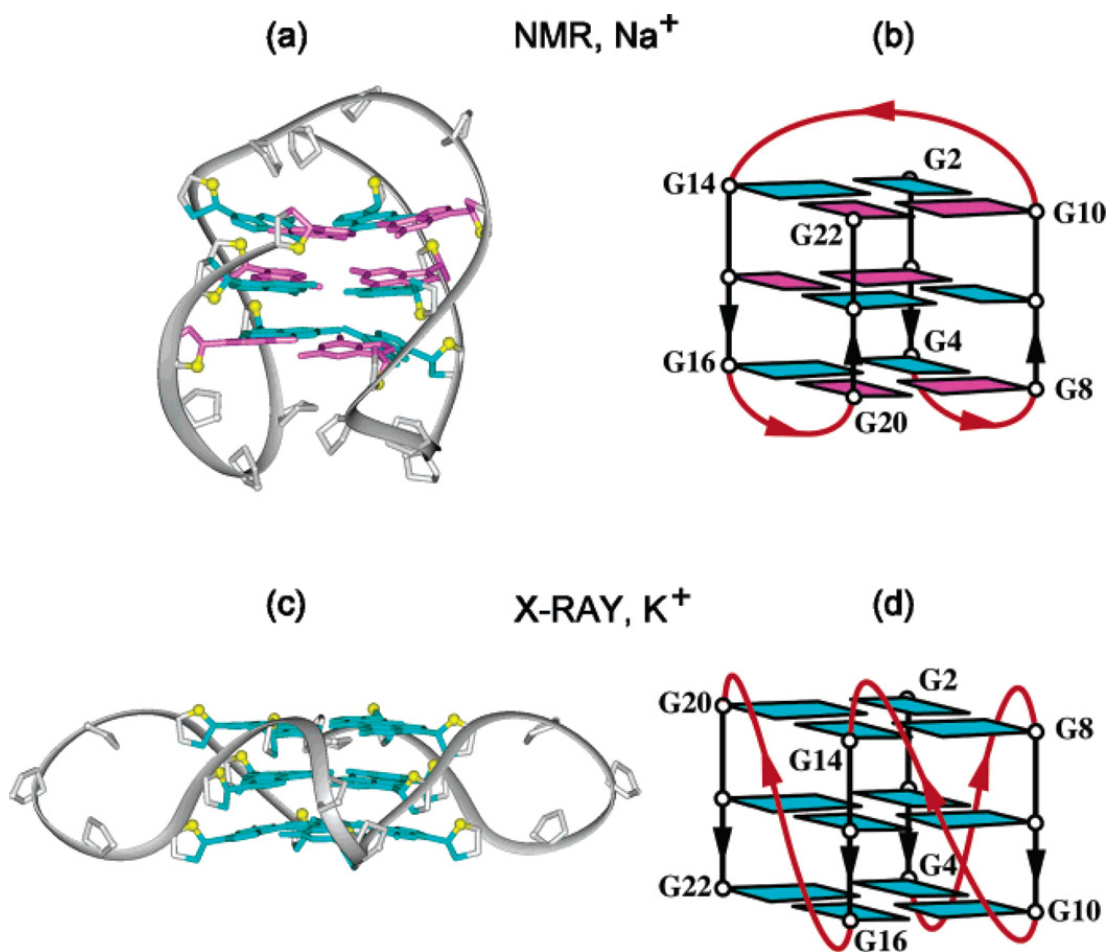


Figure 1.5: Structure of intramolecular G-quadruplexes formed by the human telomeric sequence $\text{d}[\text{AGGG}(\text{TTAGGG})_3]$: (a, b) in Na^+ solution; (c, d) in a K^+ -containing crystal. Loops are colored red; anti and syn guanines are colored cyan and magenta, respectively. O4' atoms are colored yellow. (adapted from [13])

A completely different intramolecular G-quadruplex assembles in a K^+ -rich ion environment in the crystal structure of the identical nucleotide sequence presented in previous paragraph (four-repeat human telomeric nucleotide sequence $\text{d}[\text{AGGG}(\text{TTAGGG})_3]$). In the observed G-quadruplex, firstly, all strands are parallel to each other, secondly, guanines are in anti conformation and thirdly,

¹For a detailed description of different G-quadruplexes, their structural entities, properties and other features see e.g. [11].

double-chain reversal loops are detected [14].

However, the formerly described intramolecular G-quadruplex structure in a K^+ -containing crystal of a $d[AGGG(TTAGGG)_3]$ human telomeric nucleotide sequence, illustrated in figure 1.5 (b), in a K^+ -rich ion solution occurs not to be the most probable form of the G-quadruplex. In a K^+ -rich ion solution, the most probable G-quadruplex fold of a $d[AGGG(TTAGGG)_3]$ nucleotide sequence is a (3 + 1) G-quadruplex fold. Therefore, G-tetrads are either anti-syn-syn-syn or syn-anti-anti-anti [13]. It follows from the above that multiple types of the G-quadruplex folds can be present in a K^+ solution simultaneously, as conformational heterogeneity is a prevailing phenomenon, we can only predict in terms of probability, which conformation in a particular case is preferable. As an example of conformational heterogeneity – a simultaneous coexistence of the parallel and antiparallel quadruplex assemblies in a K^+ solution [15].

Flanking sequences, on 5' and 3' ends of a quadruplexe-assembling sequence play key roles in the stabilization of a particular G-quadruplex structure fold, especially the 3'-flanking sequence. This fact can be used purposely, for instance, for the creation of a telomeric sequence, which would assemble a single stable unimolecular G-quadruplex structure in a K^+ environment [6]. Some of the mentioned intramolecular G-quadruplex structures are pictured and elucidated in Figure 1.6.

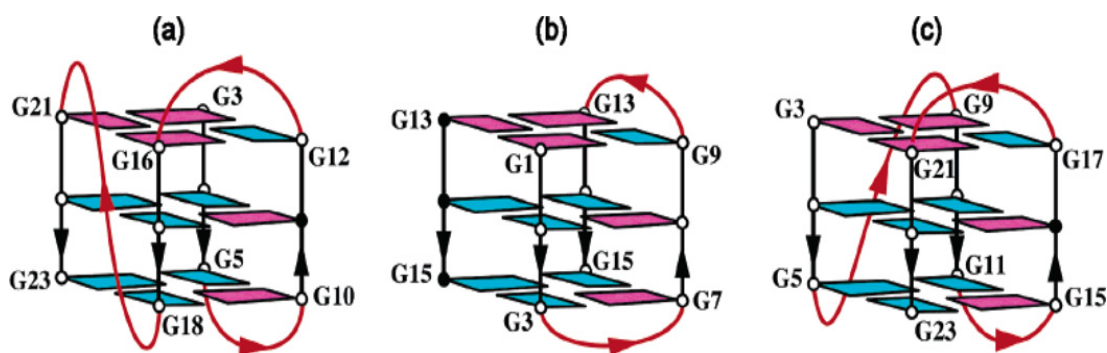


Figure 1.6: Schematic structures of the (3 + 1) core-containing G-quadruplexes formed by (a) the four-repeat *Tetrahymena* telomeric $d[(TTGGGG)_4]$ sequence in a Na^+ solution, (b) the three-repeat human telomeric $d[GGG(TTAGGG)_2T]$ sequence in a Na^+ solution, and (c) the four-repeat human telomeric $d[TAGGG(TTAGGG)_3]$ sequence in a K^+ solution (adapted from [13]).

Some of the conformations may be predominant, as in the case of the $A_3(G_3TTA)_3G_3AA$ nucleotide sequence in a K^+ environment, regardless of Na^+ presence or absence. For this sequence it was observed that the addition of K^+ to the pre-assembled basket-type G-quadruplex structure of the $AG_3(G_3TTA)_3$ sequence (in the presence of 100 mM Na^+) converts the G-quadruplex fold to the K^+ -form.

Another factor that has been proven to have an impact on human telomere structures in a K^+ -rich ion environment is DNA concentration [16], along with stabilizing cation type and concentration, molecular crowding or dehydrating conditions and different protocols of sample preparation. Long telomeric DNA behaves in the same way, as the $GGG_3(TTAGGG)_3$ nucleotide sequence that forms a G-quadruplex of a antiparallel basket conformation in Na^+ or K^+ en-

vironment. The Na^+ -stabilized $\text{G}_3(\text{TTAG}_3)_3$ nucleotide sequence remains in its antiparallel assemble despite high Na^+ concentration, annealing or thermal destabilization. However, transitions of both $\text{G}_3(\text{TTAG}_3)_3$ and long telomeric DNA from antiparallel G-quadruplex conformation to a parallel G-quadruplex arrangement, through $(3 + 1)$ conformation, has been proven to be induced in a K^+ -rich solution by the $\text{G}_3(\text{TTAG}_3)_3$ concentration increase [17]. In other words, human telomeric DNA in K^+ can adopt the antiparallel G-quadruplex conformation, the $(3 + 1)$ conformation, or the parallel quadruplex conformation depending on the DNA concentration. In addition to concentration dependence, the quadruplex structure was proven to depend on the sample preparation method. Influencing factors, that have not yet been pointed out, include annealing, the presence or absence of thermal destabilization, storage conditions and the time elapsed between the sample preparation and management. On the other hand, DNA concentration differing by orders of magnitude for various experimental methods can be responsible for contradictory structures observed under otherwise comparable conditions [16].

2. Experimental Methods

2.1 NMR spectroscopy

2.1.1 NMR essence and resonance frequency

Magnetic moments of a nuclei tend to be highly sensitive towards neighbouring environment, however, their interaction with that environment is remarkably weak. Usually, nuclei in their ground, unexcited states are used in NMR experiments. The nuclear magnetic moment, $\boldsymbol{\mu}$, is closely connected to the magnetic nuclei angular momentum, \boldsymbol{I} , also known as spin, through the following relation,

$$\boldsymbol{\mu} = \gamma \boldsymbol{I} \quad (2.1)$$

where constant γ is the gyromagnetic ratio of the certain nucleus (also dependent on the nuclide state). The magnitude and orientation of both $\boldsymbol{\mu}$ and \boldsymbol{I} , since both of those quantities are vector quantities, are quantized [18].

Different isotopes of the same chemical element possess different nuclear angular momenta I . The spin value of a nuclei, that possess an even number of nucleons, is equal to $k/2$, whereas spin value of nuclei possessing an even number of nucleons is equal to k , where k is integer. If an isotope has an even number of protons and neutrons, the protons and neutrons end up pairing with antiparallel spins, which results in a total zero spin angular momentum I .

Without magnetic field presence, every I -spin nucleus direction, of a total $2I+1$ orientations, has the same energy value. After magnetic field \boldsymbol{B} application, the mentioned energy level degeneration is no longer present. The magnetic field separates energy levels assigned to different nuclear spin projections. As for nuclei, that posses a $1/2$ value of spin, only 2 orientations can be measured. That is a case of a hydrogen atom ^1H . A static magnetic field in NMR technique causes equidistant separation of nuclear energy levels in respect to corresponding magnetic quantum number, whereas a radiofrequency electromagnetic field of an appropriate frequency causes energy state flips. The energy, E , of nuclear magnetic moment, $\boldsymbol{\mu}$, in magnetic field, \boldsymbol{B} , can be described as[18]

$$E = -\boldsymbol{\mu} \cdot \boldsymbol{B} \quad (2.2)$$

For NMR, only transitions between adjacent states are permitted, which is the essence of the selection rule. Energy required for transition between permitted states, also known as energy resonance condition, can be expressed as

$$\Delta E = h\nu = \hbar\gamma B \quad (2.3)$$

where ν is the electromagnetic radiation frequency. The mentioned radiation frequency, ν , also known as Larmor frequency can be described as

$$\nu = \frac{\gamma B}{2\pi} \quad (2.4)$$

2.1.2 Chemical shifts and nuclear shielding

The gyromagnetic ratio, γ , of a nucleus is a principally important unit for NMR spectroscopy, along with the magnetic field strength it is exposed to. One of the core concepts of the NMR technique is founded on the effect, called chemical shift. Chemical shift, in turn, is rooted in the difference of a local field B that is experienced by a nucleus in an atom and external field B_0 , that would impact the bare nucleus of an atom without the presence of related electrons. In the case of a ^1H proton, different NMR frequencies will correspond to each inequivalent ^1H proton in a molecule, due to different electron distribution around each molecular atom. By cause of this shielding, the experimenter is capable of distinguishing between signals coming from different ^1H protons. The resulting magnetic field B felt by the nucleus (surrounded by corresponding electrons) is usually smaller than the external magnetic field B_0 , as the external magnetic field B_0 is diminished by the magnetic field B' , created by the nucleus-surrounding electrons. Therefore, the mentioned relation, in case of an isotropic liquid as a medium, can be written as follows

$$B = B_0 - B' = B_0(1 - \sigma) \quad (2.5)$$

where σ denotes the so-called shielding constant. The resonance frequency of a nucleus can be calculated as

$$\nu = \frac{\gamma B_0(1 - \sigma)}{2\pi} \quad (2.6)$$

However, the usage of the shielding constant σ is not considered practically comfortable, thus the chemical shift δ is more commonly used, which is defined as

$$\delta = 10^6 \cdot \frac{(\nu - \nu_{ref})}{\nu_{ref}} \quad (2.7)$$

where ν_{ref} is a resonance frequency of a reference nucleus. Chemical shift δ values are usually provided in parts per million, or ppm. If a liquid is being explored, chemical shifts in a ^1H spectrum will be found in the interval of 0-15 ppm in respect to reference compound tetramethylsilane, also known as TMS. Approximate chemical shifts of hydrogen nuclei in respect to a different atomic surrounding are presented in figure 2.1.

2.1.3 Bloch Equations

When placed in a magnetic field, a collection of magnetic nuclei spread themselves unevenly amongst the $2I+1$ available energy levels according to Boltzmann distribution [18]. This results to a macroscopic nuclear magnetization \mathbf{M} . Phenomenological equations that describe the time development of magnetic moment vector \mathbf{M} in a magnetic field in terms of classic (non-quantum) physics are called Bloch Equations. In equilibrium state \mathbf{M} has the same direction as a vector of an external static magnetic field \mathbf{B}_0 . According to classical theory, the behaviour of a magnetic moment \mathbf{M} in a magnetic field \mathbf{B} is described by the equation of motion [20]:

$$\frac{d\mathbf{M}}{dt} = \gamma[\mathbf{M} \times \mathbf{B}] \quad (2.8)$$

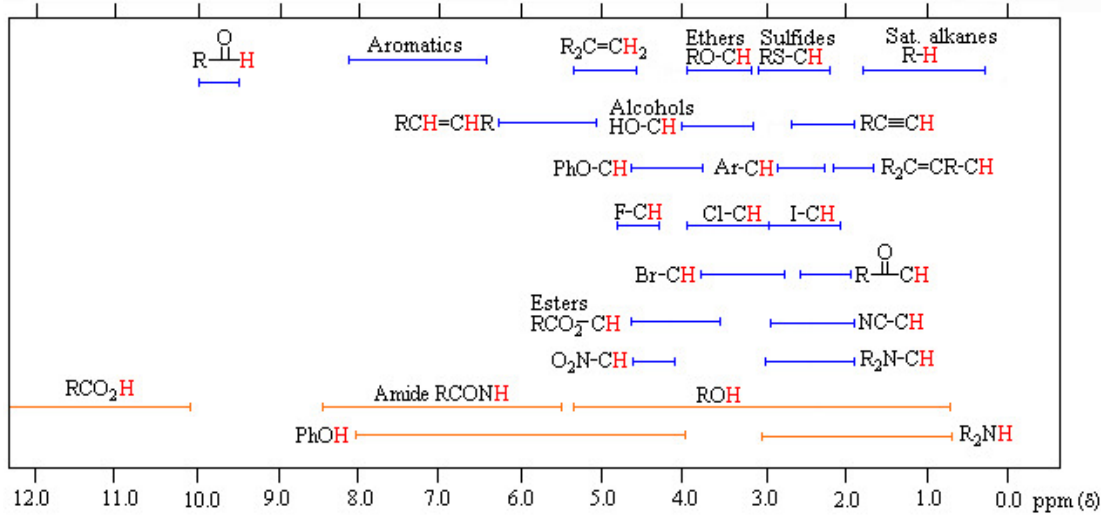


Figure 2.1: Chemical shifts of hydrogen nuclei in respect to a different atomic surrounding (adapted from [19]).

In the NMR experiment, field \mathbf{B} is obtained as a sum of static magnetic field \mathbf{B}_0 and the rotating field \mathbf{B}_1 . The direction of static magnetic field \mathbf{B}_0 vector, pointing along the z axis, is called longitudinal, whereas the perpendicular plane in respect to the z axis is called the transverse plane. Bloch equations, that include relaxation effects, in their complete form look as follows:

$$\frac{dM_x}{dt} = \gamma[\mathbf{M} \times \mathbf{B}]_x - \frac{M_x}{T_2} \quad (2.9)$$

$$\frac{dM_y}{dt} = \gamma[\mathbf{M} \times \mathbf{B}]_y - \frac{M_y}{T_2} \quad (2.10)$$

$$\frac{dM_z}{dt} = \gamma[\mathbf{M} \times \mathbf{B}]_z - \frac{M_z - M_0}{T_1} \quad (2.11)$$

where T_1 is longitudinal relaxation time, T_2 is transverse relaxation time (also known as spin-spin relaxation), M_0 is equilibrium magnetization, M_x , M_y , M_z are total magnetization vector \mathbf{M} components. Bloch equations describe magnetic nuclei moment interaction with external field only. They do not contain any component that would be responsible for inter-nuclei interactions description, except for relaxation. The relaxation time T_2 is characteristic for the transverse magnetization in the x - y plane while the variation of the longitudinal magnetization along the z -axis is a function of T_1 [20].

2.1.4 Free Precession

Assuming $\mathbf{M}(0)=(M_0,0,0)$, $\mathbf{B}(0)=(0,0,B_0)$ and $T_1, T_2 \rightarrow \infty$ equations (2.9)-(2.11) will take following simplified form

$$\frac{dM_x}{dt} = \gamma B_0 M_y \quad (2.12)$$

$$\frac{dM_y}{dt} = -\gamma B_0 M_x \quad (2.13)$$

$$\frac{dM_z}{dt} = 0 \quad (2.14)$$

From the last set of equations we can obtain a subsequent, general, non-trivial solution

$$M_x = C \cos(\gamma B_0 t) \quad (2.15)$$

$$M_y = -C \sin(\gamma B_0 t) \quad (2.16)$$

$$M_z = \text{const} \quad (2.17)$$

where C is a constant. Using complex calculus, we may rewrite obtained M_x and M_y components of magnetic moment vector \mathbf{M} , applying identity $M_{\perp} = M_x + iM_y$, as a transversal component

$$M_{\perp}(t) = C e^{-i\omega t} \quad (2.18)$$

where equation $\omega = 2\pi\nu$ and equation (2.4) for the case of external magnetic field \mathbf{B}_0 were used. From equation (2.18) we may conclude, that for non-zero gyromagnetic ratio nuclei, in the ideal case of no relaxation present, the precession of magnetic moment vector \mathbf{M} would be happening with angular frequency ω . Successive Fourier transformation applied on an obtained signal from the experiment provides a frequency spectrum. Those frequencies represent resonance frequencies of a nuclei present in different chemical environments.

2.1.5 Radiofrequency pulses and spin echo

The main principle behind the modern NMR spectroscopy is the usage of short pulse sequences, generated by a radiofrequency field, which results in the direction diversion of a total magnetization vector \mathbf{M} away from the z axis, followed by a measurement of voltage time dependency on the detection coil, which is caused by Larmor precession of total nuclear magnetization vector \mathbf{M} . Assume that the sample of interest is placed into homogeneous static magnetic field \mathbf{B}_0 , directed along the z axis. In addition to homogeneous magnetic field \mathbf{B}_0 , there is a weak radiofrequency magnetic field $\mathbf{B}_1(t)$, whose vector is rotating in a plane orthogonal to z axis (xy plane) at a ω_{rf} frequency, neighbouring Larmor frequency ω_0 . The summary field cants from a z axis direction and exhibits a rotating motion around the z axis at a ω_{rf} frequency. In a rotating frame (coordinate system rotating around z axis in step at ω_{rf} frequency), field $\mathbf{B}_1(t)$ is constant, therefore magnetic moment vector \mathbf{M} (assuming radiofrequency pulse duration $t_p \ll T_1, T_2$) exhibits precession around \mathbf{B}_1 vector. The angle through which the magnetization vector \mathbf{B}_1 turns (away from z axis) is called a flip angle β [18]

$$\beta = \gamma B_1 t_p \quad (2.19)$$

Radiofrequency magnetic field $\mathbf{B}_1(t)$ usually operates in short pulses. The most frequently used pulses have $\pi/2$ or π flip angles. The first one would make magnetic moment vector \mathbf{M} perpendicular to the z axis, while π -pulse would

invert the magnetization vector direction, placing it along the negative z axis. The record of a free precession signal, presented earlier in this chapter, after a $\pi/2$ -pulse application, is the essence of one of the main pulse NMR spectroscopy techniques. Other techniques contain more complicated pulse sequences, however they all assume non-zero transversal magnetization vector component in order for free precession signal to be detected and recorded.

Certainly, radiofrequency pulse may effect only those nuclei, whose resonance frequency is equal to pulse frequency or is in its frequency neighbourhood. The longer the pulse, the lesser the range of resonance frequencies being affected, therefore, pulse construction for certain frequency intervals is possible. The selection of pulse duration should be correlated with the intensity of radiofrequency magnetic field B_1 for the sake of flip angle β value conservation.

The π -pulse is used for the creation of the so-called spin echo. The diagram with the π -pulse sequence for the spin echo experiment is available in the following figure 2.2.

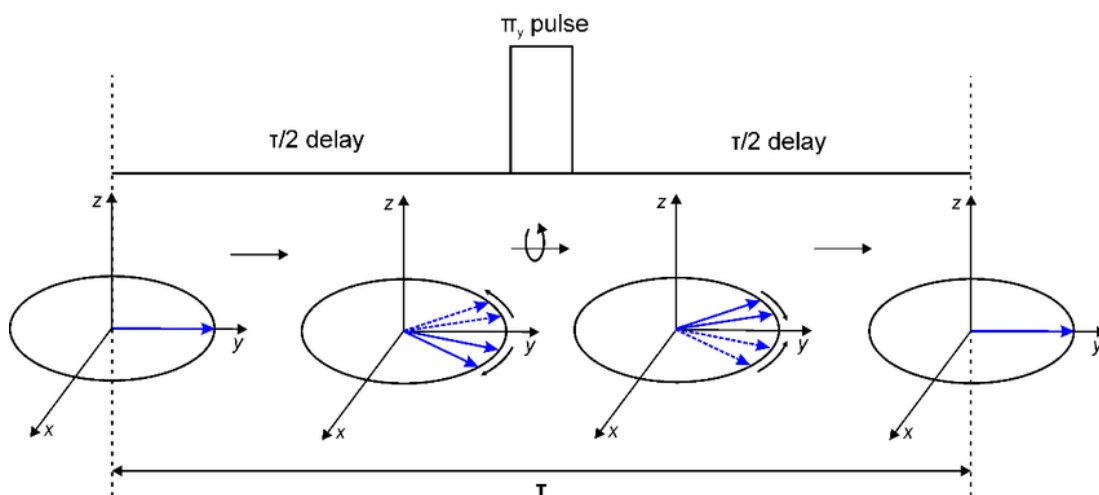


Figure 2.2: Operation of the spin echo experiment (adapted from [21]).

As shown in the previous figure, right after the elapse of $\tau/2$ following the π -pulse, most of the nuclear spins will find themselves in the phase coherence state, suitable for the signal measurement, due to its high intensity. Subsequently, the signal will start fading away due to transversal magnetization vector component M_{\perp} decay. The entire described process of the signal's growth, climax and fading, as a whole, is called spin echo. The Fourier transformation application, as with the free precession experiment, will give us a frequency spectrum; however, the transversal relaxation process occurs during the delays $\tau/2$ that modifies the spectrum.

2.1.6 Water signal suppression

Water signal suppression is a mandatory step, when ^1H NMR spectrum of a compound in aqueous solution is being recorded. There were several techniques invented for water signal suppression, and certainly, each of them has its advantages, disadvantages and range of applicability. However, all mentioned techniques have a similar feature – all of them also suppress the solution-exchangeable protons'

signal (in case of nucleic acids we refer to amino and imino protons of the bases). If due to the proton exchange, some of sample's protons become incorporated into water molecules during the water suppression part of a pulse sequence, the signal corresponding to them will also be suppressed, therefore it won't be observed in a spectrum. The described effect will be certainly followed by respective peaks' intensity decrease, representing smaller quantity of the atoms.

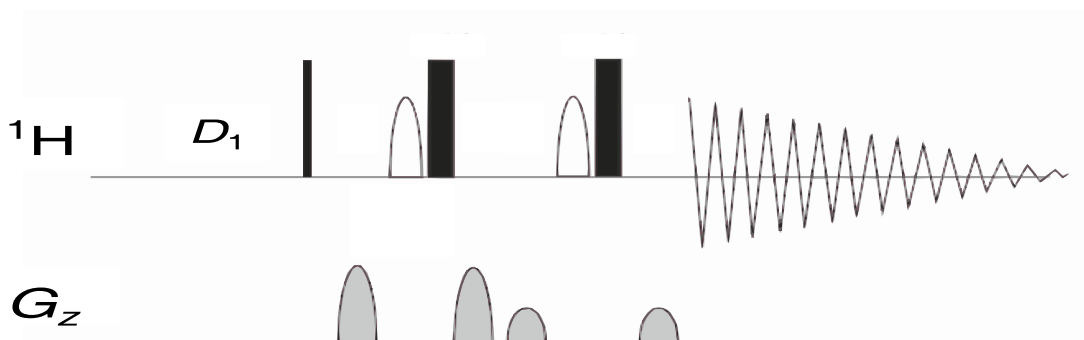


Figure 2.3: Zgesgp pulse sequence (adapted from [22])

One of the widely used gradient-based solvent suppression pulse sequence called zgesgp, uses hard 180° , selective 180° pulses and gradient pulses on H_2O , in case of water as a solvent[23]. The mentioned pulse sequence illustration is available in figure 2.3. Spin echo is created for all ${}^1\text{H}$ resonances except for those, at the water resonance frequency, which are de-phased by the magnetic field gradients.

The adequate adjustment of selective pulse parameters, like duration, shape, range or intensity will result in water spectrum line suppression in an accurate range. The water spectrum signal can be decreased to that extent, so it becomes comparable to the scale of the other spectral lines, so even oligonucleotide protons signal from chemical shift range, neighbouring the water chemical shift interval, become distinguishable.

2.1.7 Dipole-dipole interaction

The nuclear spin, along with magnetic field \mathbf{B}_0 's impact and local magnetic fields induced by electrons, is also affected by the magnetic field, created by other nuclei. Two nuclear magnetic dipolar moments interaction throughout space is called direct dipole–dipole interaction. Dipole–dipole interaction energy is inversely proportional to r^3 [18]. Due to chaotic molecular motion, this interaction averages out to zero, therefore, the direct dipole–dipole interaction is not observable in a liquid spectra, however it does impact relaxation process.

On the other hand, indirect dipole–dipole interaction (i.e. J-coupling) engages nuclear spins through electrons involved in chemical bonds, which make an indirect dipole–dipole interaction extremely dependent on the bond's chemical nature, as well as on the nature of the observed nuclei. The J-coupling is notably helpful for the structure determination of liquid samples using NMR spectroscopy [18]. The indirect dipole–dipole interaction is in charge of spectral lines splitting between magnetically non-equivalent nuclei. Due to the nuclear spin projection

onto the direction of the magnetic field, the magnetic field intensity around the active nuclei tends to shift, which, for every nuclei with 1/2 spin in range of the possible indirect dipole–dipole interaction, results in the spectral line splitting into two lines (i.e. a doublet). If a greater quantity of active nuclei are in range of a possible indirect dipole–dipole interaction, then more complicated multiplets can be experimentally observed in the spectrum.

If the impact of the indirect dipole–dipole interaction is not of interest for the experiment, or even causes spectrum elucidation problems, then its effect can be suppressed with the decoupling technique, which can be homonuclear or heteronuclear [20]. The essence of the decoupling technique is based on sample irradiation at a certain frequency or frequency range, which, to a certain extent, results in the cancelling of the coupling effect.

2.2 Raman spectroscopy

2.2.1 Introduction to the general idea

Raman spectroscopy is a type of vibrational spectroscopy, which is based molecular vibrational states' studies for obtaining information about the structure of the explored molecule, along with its dynamics and corresponding environment. Samples may be examined either in bulk or in microscopic amounts over a wide range of temperatures and physical states (e.g., gases, liquids, latexes, powders, films, fibers, or as a surface or embedded layer) [24]. Phenomenologically, Raman spectroscopy can be described as a two-photon inelastic light-scattering event. First, an incoming high energy photon (of energy greater than the molecular vibration energy) partially transfers its energy to a molecule in a form of molecular vibration, whereas the rest of the energy is being re-emitted in a form of a reduced frequency photon. In general, a molecular energy transfer may happen through photon absorption/emission. The mentioned process of energy increase/decrease may be expressed as follows:

$$h\nu_s = h\nu_0 \pm (E_2 - E_1) \quad (2.20)$$

where ν_s is the frequency of the scattered photon, ν_0 is the frequency of the incident photon, E_2 and E_1 account for energy level values, describing energy increase/decrease.

Approximately, we may consider total molecular energy as consisting of three components

$$E_{total} = E_{rot} + E_{vib} + E_{el} \quad (2.21)$$

where E_{rot} is rotational energy, E_{vib} is vibrational energy and E_{el} is the electronic energy of a molecule. In Raman spectroscopy, the point of our interest is E_{vib} . Vibrational energy levels of a molecule are reliant on the corresponding atom masses, the bonds' strength and geometric setup, therefore they are exceptional for each molecule. As a result, every Raman vibrational band can be uniquely described by its intensity, frequency, and band shape.

2.2.2 Raman Effect

Usually, in a Raman experiment, a sample is being exposed to monochromatic radiation produced by a laser. Scattered light happens to have a dominant Rayleigh scattered light constituent, which has incident beam frequency ν_0 , and a minor Raman scattered light constituent, which happens to have frequencies $\nu_0 \pm \nu_m$ (where ν_m is a vibrational frequency of the molecule) and approximately a 10^{-5} intensity of the incident beam. The $\nu_0 - \nu_m$ and $\nu_0 + \nu_m$ lines are called the Stokes and anti-Stokes lines, respectively [25]. Certainly, the frequency of interest in a Raman experiment is ν_m .

An electromagnetic wave interacting with an explored molecule, causes perturbation and dislocation of the molecule's charged particles and further induction of dipole moment μ , which can be described as [24]

$$\mu = \alpha E \tag{2.22}$$

where E is incident electric field, α is polarizability (generally a tensor quantity). The induced dipole moment occurs as a result of molecular polarizability, where polarizability is the deformability of the electron cloud about the molecule by an external electric field [24]. A classic description of the light scattering process can be done by using the oscillating dipoles model, affected by the electromagnetic field incident radiation and further electromagnetic radiation production (the magnetic constituent of an electromagnetic wave is ignored).

There is no energy loss in respect to elastically scattered Rayleigh light, on the contrary to Raman scattered light, where energy loss relative to the exciting energy occurs in specific vibrational sample coordinates. Change in molecular polarizability caused by molecular vibration will result in the amplitude modulation of the dipole moment oscillation and thus, Raman Stokes or anti-Stokes emission, and therefore, the possibility of Raman bands observation [24].

3. Materials and Experiment description

3.1 Samples

At our disposal, we had two human telomere nucleotide sequences purchased from ATDBio Ltd, Southampton. Regarding the first, consisting of 21 nucleotides - GGGTTAGGGTTAGGGTTAGGG - we will refer to it as Tel21. The second sequence, consisting of 22 nucleotides - AGGGTTAGGGTTAGGGTTAGGG - will be referred to as Tel22. Before the experiment, the following steps were executed by the bachelor's thesis advisor:

1. Dissolution of lyophilised samples to the total volume of $500\mu\text{L}$ in deionized water.
2. Desalination and concentration with an Amicon 3k filter for accidental dirt and counter-ion removal.
3. Dissolution to a volume of $1200\mu\text{L}$ in a 50 mM Na^+ -PBS (PH 7.0) + 150 mM NaCl + 150 mM KCl + 10% D_2O + DSS ($\approx 0,01\text{ mM}$) buffer, where PBS is Phosphate-buffered saline and DSS is 4,4-dimethyl-4-silapentane-1-sulfonic acid.
4. Annealing to the temperature of $97\text{ }^\circ\text{C}$ for 20 minutes, followed by slow cooling in order to break possible non-quadruplex structures and create a unitary intramolecular antiparallel Na^+ -quadruplex.
5. Condensing with Amicon 3k to approximately $40\mu\text{L}$.
6. Transfer to a K^+ -PBS buffer (same as the previously mentioned Na^+ - PBS, but with the addition of 150 mM of KCl).
7. Reduction of total volume, using Amicon 3k filter, to $33\mu\text{L}$ for Tel21 and $45\mu\text{L}$ for Tel22 (accurately measured by weighing).
8. Incubation for 22 hours at $38\text{ }^\circ\text{C}$.

Losses could have occurred during the condensing process while using the Amicon 3k filter. However, assuming zero losses, Tel 21 concentration is $c_{\text{Tel}21} = 12.0\text{ mmol} \cdot \text{dm}^{-3}$ and Tel 22 concentration is $c_{\text{Tel}22} = 12.4\text{ mmol} \cdot \text{dm}^{-3}$. For the experiments, the same sample volume $V_{\text{Tel}21, \text{Tel}22} = 30\mu\text{L}$ was used.

PBS is mainly in charge of maintaining a constant pH. NaCl has been added in order to increase the total sodium cation concentration, whereas 10% volume fraction of D_2O has been added in order to maintain the magnetic field homogeneity and stability during NMR experiments. The purpose of DSS addition is to serve as an NMR standard.

As a buffer, for all the samples we used the one mentioned in point 3 of the previous list

Samples with variable concentrations were prepared by adding buffer to the Tel21 and Tel22 initial samples. Each sample, intentionally, was meant to have approximately half of the concentration of the previous sample (for both Tel21 and Tel22 sequence). As for the NMR experiment, each sample was then transported to a test tube, which would be later used for the measurements. For the Raman spectroscopy experiment, which was carried out for the Tel22 nucleotide sequence only, we would always pipette $10\mu\text{L}$ out of each Tel22 sample into an individual tube. The planned concentrations of the concentrated Tel 22 and Tel21 samples are presented in two following tables¹(Table 3.1, Table 3.2) as c_{calc} .

sample	n[nmol]	V[μL]	c_{calc} [mM]	c_{exp} [mM]
1	372	350	1,06	0,571
2	361	640	0,56	0,381
3	175	655	0,26	0,095

Table 3.1: Concentration table for Tel22 NMR experiment

sample	n[nmol]	V[μL]	c_{calc} [mM]	c_{exp} [mM]
1	360	300	1,20	0,482
2	343	630	0,54	0,397
3	169	655	0,26	0,072

Table 3.2: Concentration table for Tel21 NMR experiment

3.2 NMR experiment

For the NMR experiment, Bruker Avance 500 III HD spectrometer was used, operating at a magnetic field of 11,7 T with a corresponding 500 MHz hydrogen ^1H resonance frequency and a 202 MHz phosphorus ^{31}P resonance frequency. Spectra were recorded using a dual-channel(^1H , broadband) probehead, suitable for liquid samples placed in a 5 mm diameter test tube. A magnetic field homogeneity, with deviations by an order of 10^{-9} magnetic field intensity, can be achieved.

There were two types of test tubes used for the NMR experiment. First, a Norell S500 for higher overall volume ($\approx 600\mu\text{L}$) measurements and second, a Shigemi BMS3 for lower overall volume ($\approx 250\mu\text{L}$) measurements. Both of the test tubes had 5 mm diameters. The sample temperature was monitored by TopSpin 3.5 software of the spectrometer. The flow of nitrogen gas through the probe was set to 600 l/h. All the samples for both nucleotide sequences were measured at the temperature of $T = 25^\circ\text{C}$.

For concentrated samples of Tel22 with concentrations $c_{calc} = 0,56$ mM and $c_{calc} = 0,26$ mM and concentrated samples of Tel21 with concentrations $c_{calc} = 0,54$ mM and $c_{calc} = 0,26$ mM we also conducted an annealing with Techne DB200/2 Dri-block heater, which lasted 22 minutes at the temperature of $T = 99^\circ\text{C}$ followed by a self-cooling process until the temperature of the sample within the heater would reach room temperature. For the Tel22 sample with $c_{calc} =$

¹Experimental sample concentration values c_{exp} were determined with the aid of the NMR ^{31}P zg-pulse experiment and their review will be provided in the next chapter of this thesis.

0,56 mM, annealing was executed twice. We executed adjustments of the radiofrequency circuit and shimming coils before every experiment, as well as we secured same physical conditions for each experiment. The chemical shifts of ^1H NMR spectra were calibrated in respect to the DSS standard at 0 ppm. The obtained data from the experiments were processed with TopSpin software, including Fourier transformation with an exponential apodization of 2 Hz line-broadening factor, phase adjustment, baseline correction.

3.2.1 Pulse experiments

The zg pulse sequence with long relaxation delays ($D1 = 60$ s) between pulses was used for oligonucleotide concentration determination from ^{31}P NMR spectra. The acquisition parameters of the used zg-pulse sequence is available in Tables 3.3 and 3.4.

The zgesgp pulse sequence was used for the oligonucleotides' ^1H NMR spectra acquiring. The acquisition parameters of the used zgesqp pulse sequence is available in Tables 3.3 and 3.4. In every ^1H experiment, we would run at least 3 dummy scans before acquisition.

Tel22						
^1H zgesgp-pulse sequence						
sample	annealing	c_{calc} [mM]	TD	NS	AQ[s]	D1[s]
1	-	1,06	32768	256	1,31	1,0
2	before	0,56	32768	1024	1,31	1,0
2	after	0,56	32768	2000	1,31	1,0
2	after	0,56	32768	16384	1,31	1,0
3	before	0,26	32768	16384	1,31	1,0
3	after	0,26	32768	16384	1,31	1,0
^{31}P zg-pulse sequence						
sample	annealing	c_{calc} [mM]	TD	NS	AQ[s]	D1[s]
1	-	1,06	16384	700	2,03	60
2	before	-	-	-	-	-
2	after	-	-	-	-	-
2	after	0,56	16384	2512	2,03	60
3	before	0,26	16384	400	2,03	60
3	after	0,26	16384	400	2,03	60

Table 3.3: Acquisition parameters for ^1H , ^{31}P Tel22 NMR experiments, where c_{calc} is expected concentrations of the concentrated samples, TD is size of FID (number of data points), NS is number of scans, AQ is acquisition time, D1 is relaxation delay. The sample number correspond to those in Table 3.1. As it was mentioned in section 3.1, annealing for the second sample was executed twice, one after another.

3.3 Determination of concentrations

Experimental concentrations of Tel21 and Tel22 were determined with the help of ^{31}P NMR zg-pulse spectra using TopSpin 3.5 software. Since molar concentra-

Tel21						
¹ H zgesgp-pulse sequence						
sample	annealing	c_{calc} [mM]	TD	NS	AQ[s]	D1[s]
1	-	1,20	32768	1024	1,31	1,0
2	before	0,54	32768	1024	1,31	1,0
2	after	0,54	32768	1024	1,31	1,0
3	before	0,26	32768	16384	1,31	1,0
3	after	0,26	32768	6144	1,31	1,0
³¹ P zg-pulse sequence						
sample	annealing	c_{calc} [mM]	TD	NS	AQ[s]	D1[s]
1	-	1,20	16384	800	2,03	60
2	before	-	-	-	-	-
2	after	0,54	16384	1120	2,03	60
3	before	0,26	16384	4480	2,03	60
3	after	0,26	16384	2600	2,03	60

Table 3.4: Acquisition parameters for ¹H, ³¹P Tel21 NMR experiments, where c_{calc} is expected concentrations of the concentrated samples, TD is size of FID (number of data points), NS is number of scans, AQ is acquisition time, D1 is relaxation delay. The sample number correspond to those in Table 3.2.

tions are directly proportional to a peak's area under the curve, by integrating peaks corresponding to the oligonucleotides and the buffer of a known concentration(which in our case was a phosphorus-containing 50 mM buffer), oligonucleotide concentration may be determined. The mentioned concentrations were calculated with approximately 20% accuracy. The computation accuracy was also based on the experiments' results(repeating the ³¹P experiments before and after annealing of the same sample).

3.4 Raman experiment

The Raman experiment was carried out for the Tel22 nucleotide sequence with the usage of samples (1,2 and 3) from table 3.1. before annealing and the samples 2 and 3 after annealing. Each particular concentration sample was provided in the same amount - 10 μ L. The volume of used samples was determined based on pipetting by a Gilson Pipetman pipette, as each particular sample was prepared simultaneously with samples used in the NMR experiment.

The 532.2 nm excitation Nd:YAG laser with frequency doubling (Verdi V2, Coherent) produced a power of approximately 0.5 W at the sample. Scattered light collected in a right-angle geometry was analysed by a Spex 270 M spectrograph (Jobin Yvon, single grating with 1800 grooves per millimetre) with a CCD detector cooled by liquid nitrogen. An edge filter in front of the spectrograph was used to suppress the elastic scattering. Raman spectra were recorded in the region of Stokes shifts between 300 cm^{-1} and 2000 cm^{-1} with the resolution of 1 cm^{-1} . The total acquisition time for every spectrum was 1000 s. A cylindrical quartz microcuvette with 10 μ l of the solution was placed into a thermo-stabilised chamber. After each measurement, the spectrum of a neon lamp was recorded for precise

spectral calibration. The background of Raman spectra was corrected by subtracting the optimal fifth-degree polynomial function, the properly scaled Raman spectra of the buffer, pure water, and the quartz cuvette wall. For subsequent data processing and spectra plotting, Gnuplot environment was used.

4. Results

4.1 NMR experiment results

In the obtained spectra from ^1H NMR experiments we are primarily interested in certain chemical shift intervals. The [6,5:12,5]ppm interval is of our principal interest. The interval [12:15]ppm corresponds to exchangeable imino hydrogens (H1), involved in regular Watson-Crick A-T and G-C base pairing, whereas imino hydrogens, related to guanines in G-tetrad construction, resonate at frequencies accountable for a [10,5:12,5]ppm interval, which expresses their engagement in a Hoogsteen H-bonded network. Imino hydrogens signal correlated with a [10,5:12,5]ppm interval is detached from the imino hydrogens' signal coming from other DNA arrangements and the number of imino resonances correspond to the number of G-tetrad-associated guanines present in the system [26].

The interval [6:10]ppm corresponds to aromatic hydrogens directly attached to an arene ring and to hydrogens of cytosine amino group, while the sugar protons' signal is exhibited in a [1:6,5]ppm interval. The chemical shift chart can be found in figure 2.1.

As an example of the oligonucleotide concentration computation using a ^{31}P NMR zg pulse spectrum, the corresponding spectrum of the Tel22 sample for $c_{calc} = 1,06$ mM (table 3.1) is presented in Figure 4.1 and the corresponding Tel21 sample for $c_{calc} = 1,20$ mM (table 3.2) is presented in Figure 4.8. The rest of the oligonucleotide concentration computations were done in identical manner. The calculatively determined (expected) c_{calc} and experimentally obtained concentrations c_{calc} resp. c_{exp} are presented in table 3.1 and table 3.2 in the previous chapter.

4.1.1 Tel22 nucleotide sequence

In figure 4.2 we observe the comparison of Tel22 ^1H NMR spectra in respect to c_{exp} , while figure 4.3 represents figure 4.2 restricted to [10:12,5] ppm interval. In figure 4.4 the comparison of Tel22 ^1H NMR spectra of the sample with $c_{exp} = 0,381$ mM before and after annealing is presented, while figure 4.5 represents figure 4.4 restricted to [10:12,5] ppm interval. A similar comparison is presented in figure 4.6, but for $c_{exp} = 0,095$ mM, whereas figure 4.7 represents figure 4.6 restricted to [10:12,5] ppm interval.

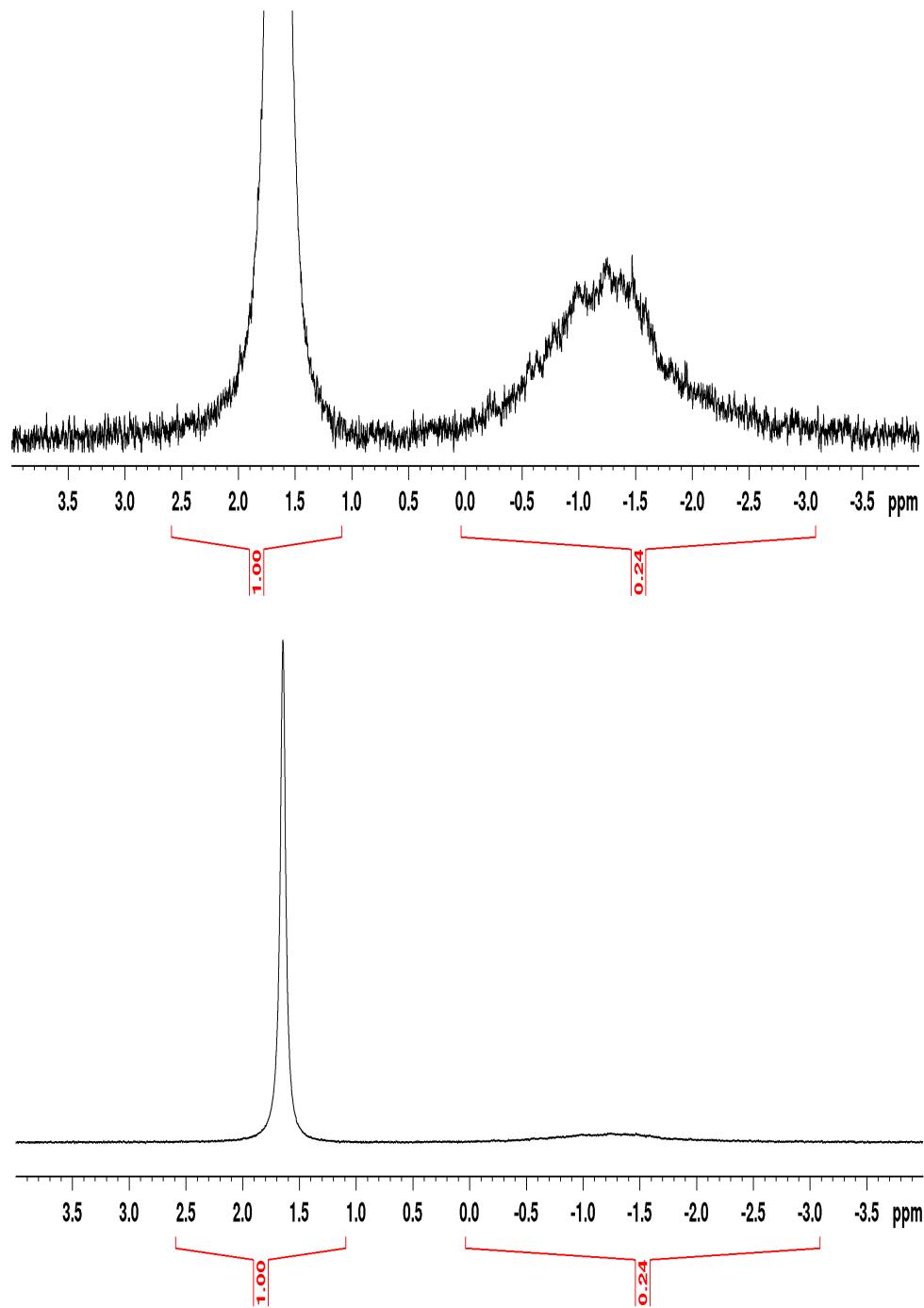


Figure 4.1: An example of experimental determination of Tel22 oligonucleotide concentration $c_{exp} = 0,571$ mM with the help of ^{31}P spectrum. The left peak corresponds to a buffer signal, the right peak corresponds to a oligonucleotide signal(containing 21 phosphorus nuclei).

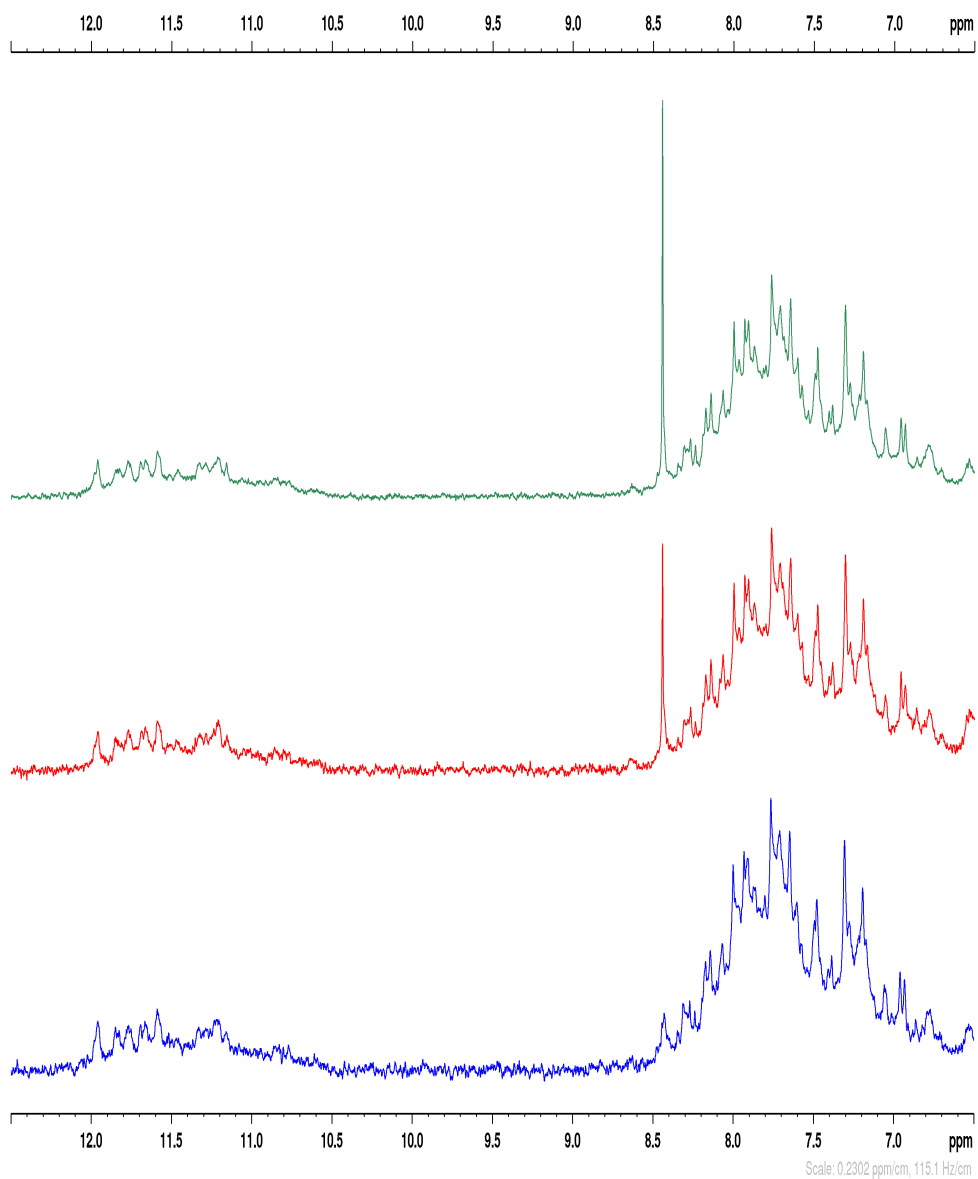


Figure 4.2: A comparison of Tel22 ^1H NMR spectra in respect to c_{exp} . **Blue** line corresponds to $c_{exp} = 0,571$ mM, **red** line corresponds to $c_{exp} = 0,381$ mM, **green** line corresponds to $c_{exp} = 0,095$ mM. Spectral noise levels are different due to different NS values, processing parameters are identical. Spectra were scaled to maximum intensity.

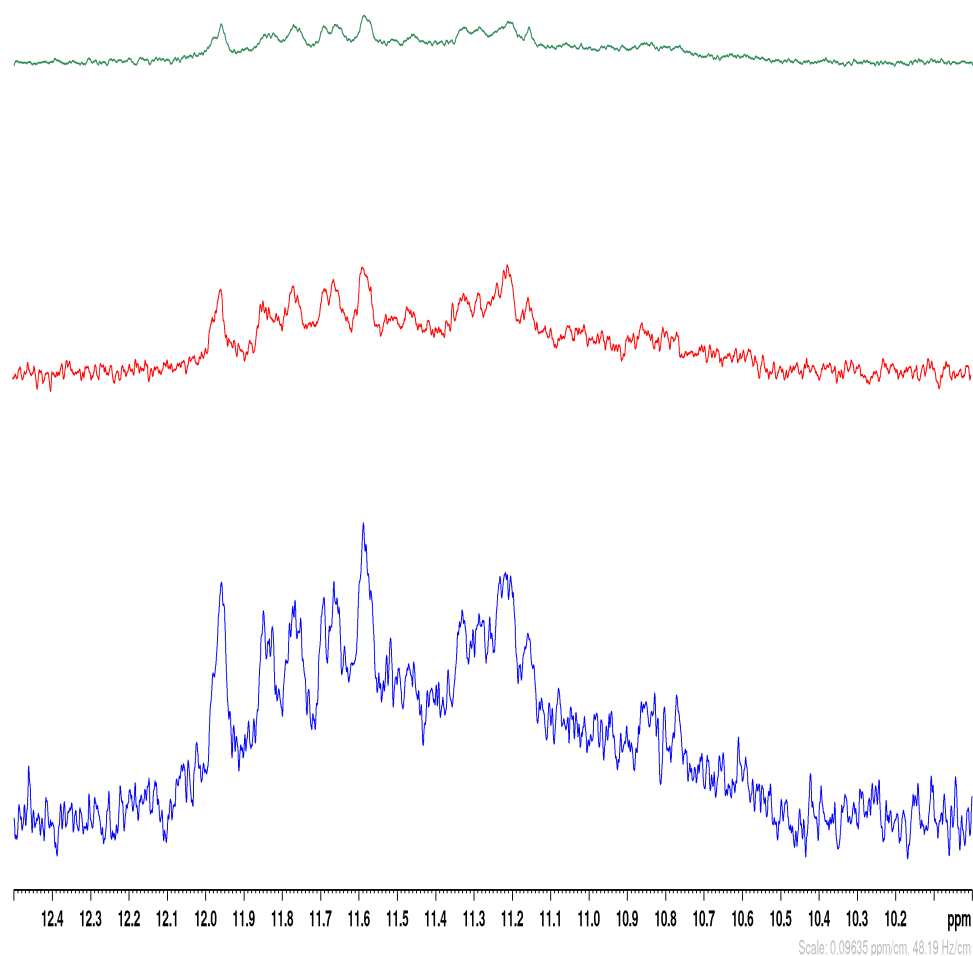


Figure 4.3: A comparison of Tel22 ¹H NMR spectra in respect to c_{exp} restricted to [10:12,5] ppm interval. **Blue** line corresponds to $c_{exp} = 0,571$ mM, **red** line corresponds to $c_{exp} = 0,381$ mM, **green** line corresponds to $c_{exp} = 0,095$ mM. Spectral noise levels are different due to different NS values, processing parameters are identical. Spectra were intensity calibrated in respect to DSS peak at 0 ppm.

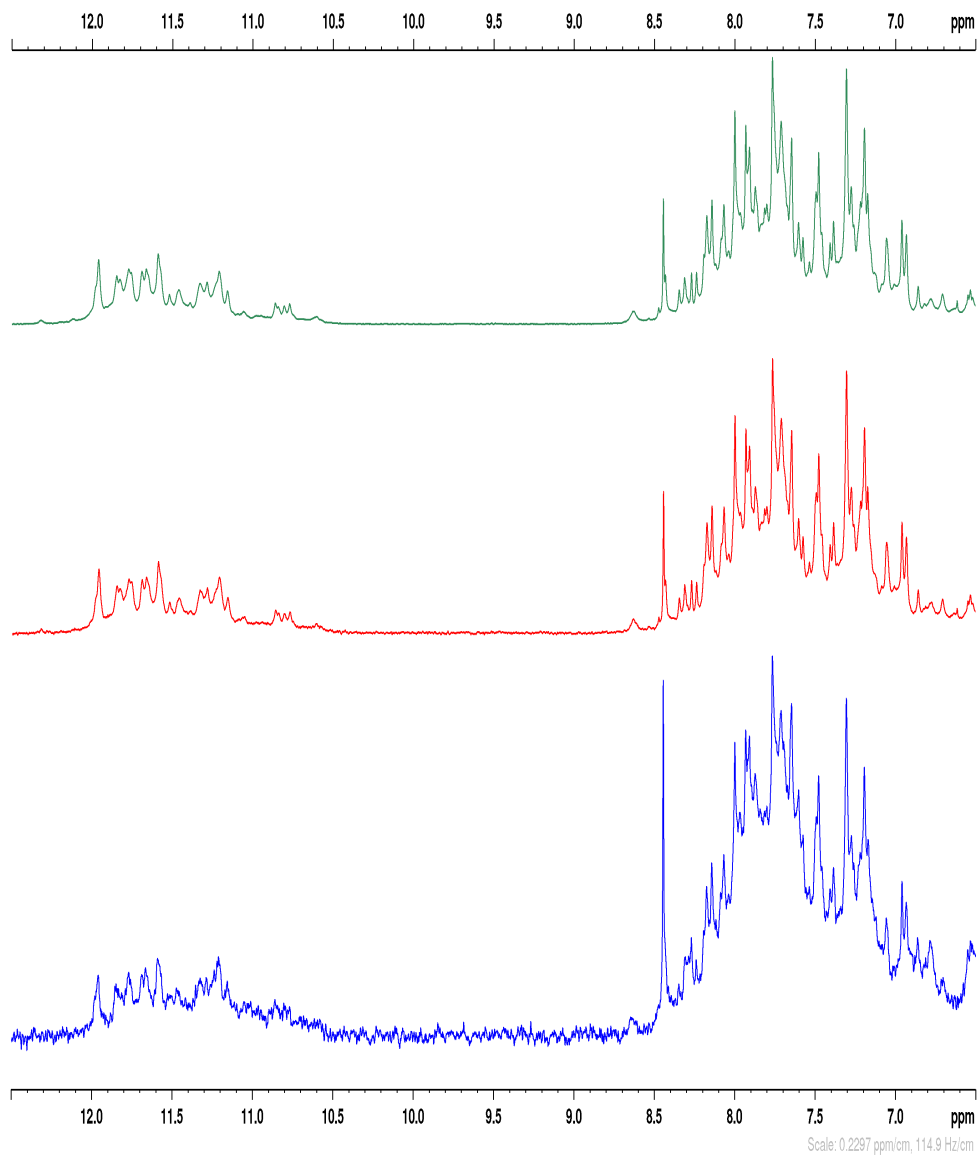


Figure 4.4: Tel22 ^1H NMR spectra for $c_{exp} = 0$, 381 mM before and after annealing. Blue line corresponds to the sample before annealing, red line corresponds to the sample after the first annealing, green line corresponds to the sample after the second annealing. Spectral noise levels are different due to different NS values, processing parameters are identical. Spectra were scaled to maximum intensity.

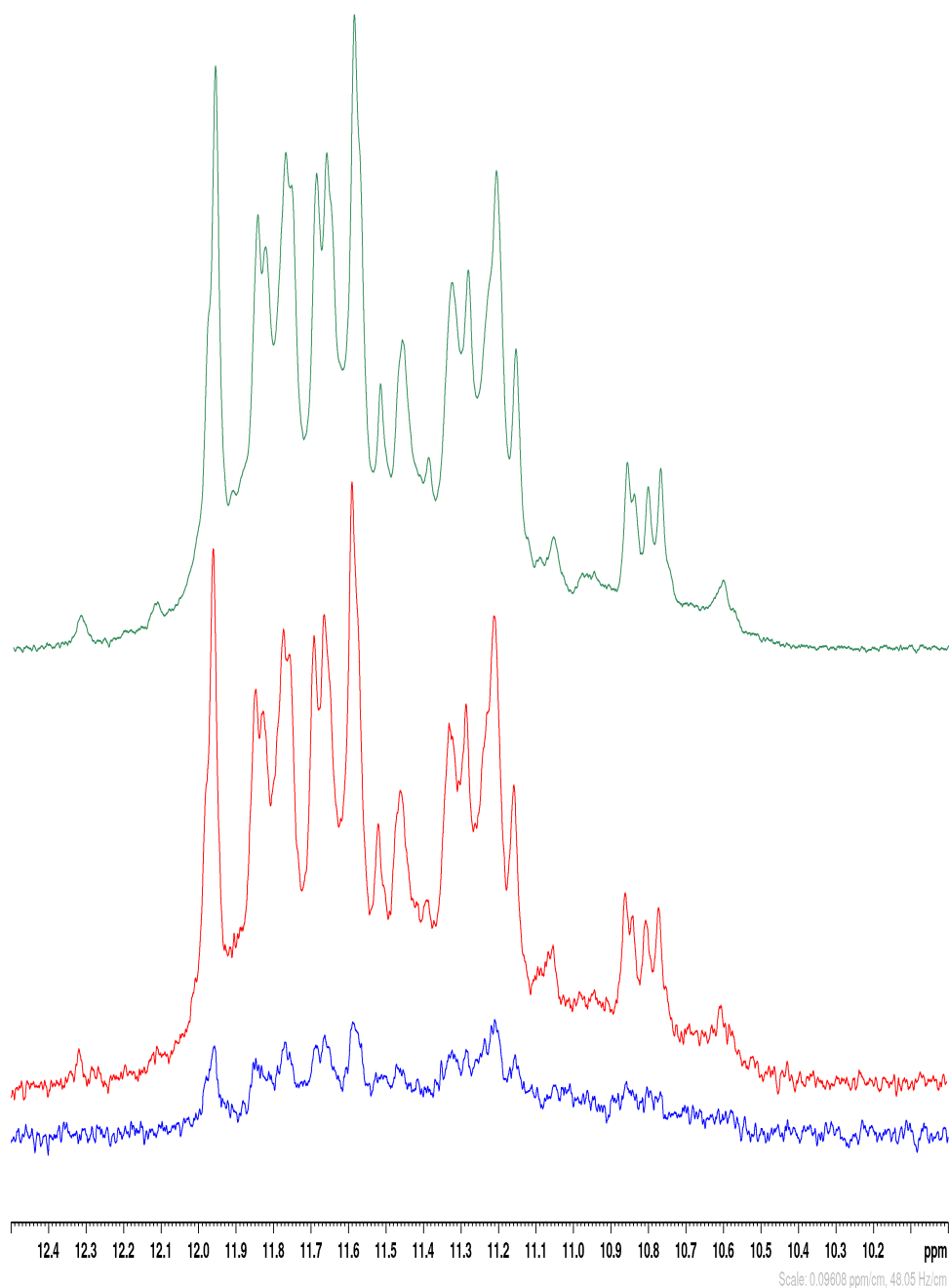


Figure 4.5: Tel21 ¹H NMR spectra for $c_{exp} = 0$, 381 mM before and after annealing restricted to [10:12,5] ppm interval. **Blue** line corresponds to the sample before annealing, **red** line corresponds to the sample after annealing. Spectral noise levels are different due to different NS values, processing parameters are identical. Spectra were intensity calibrated in respect to DSS peak at 0 ppm.

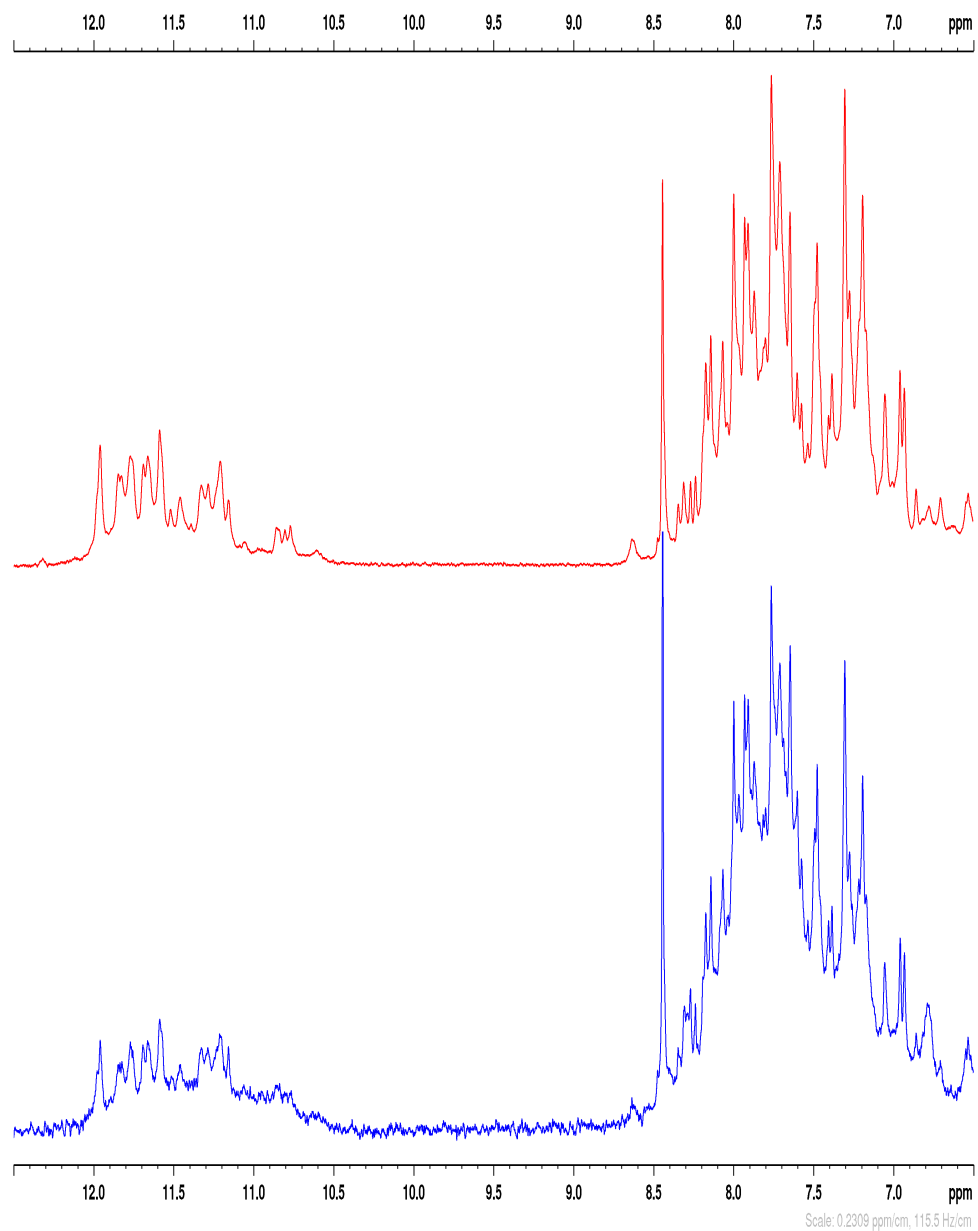


Figure 4.6: Tel22 ^1H NMR spectra for $c_{exp} = 0,095$ mM before and after annealing. **Blue** line corresponds to the sample before annealing, **red** line corresponds to the sample after the annealing. Spectra were scaled to maximum intensity, possess the same acquisition and data processing parameters.

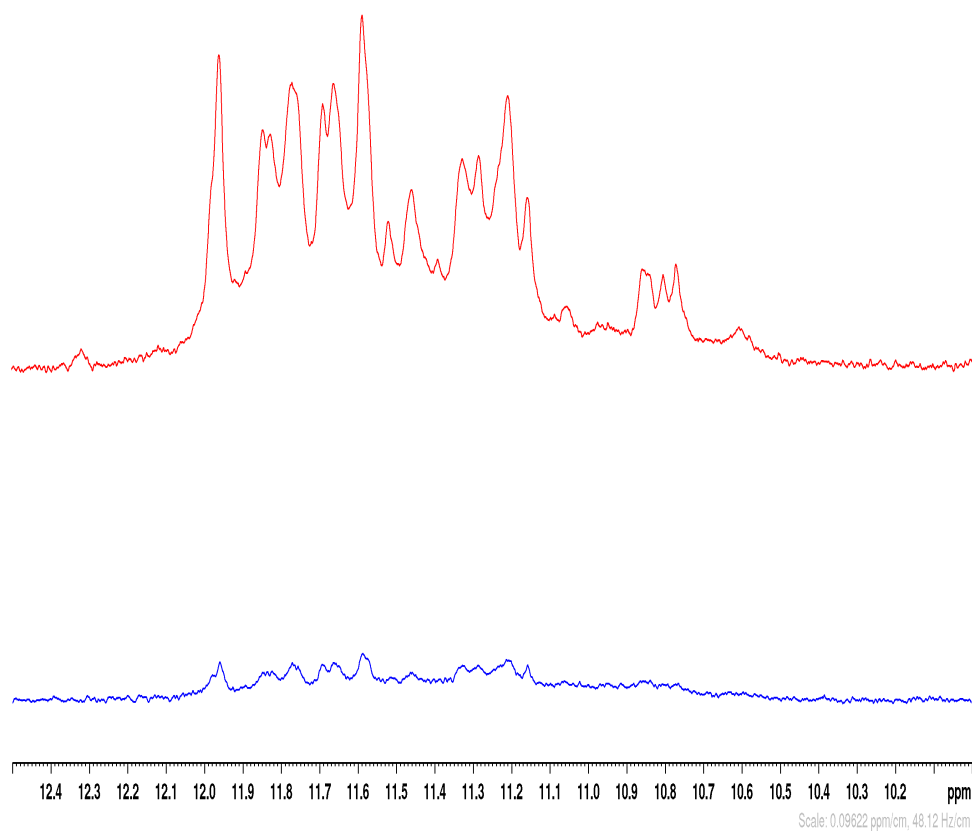


Figure 4.7: Tel21 ^1H NMR spectra for $c_{exp} = 0,095$ mM before and after annealing restricted to [10:12,5] ppm interval. **Blue** line corresponds to the sample before annealing, **red** line corresponds to the sample after annealing. Spectral noise levels are different due to different NS values, processing parameters are identical. Spectra were intensity calibrated in respect to DSS peak at 0 ppm.

4.1.2 Tel21 nucleotide sequence

In figure 4.9 we observe a comparison of Tel21 ^1H NMR spectra in respect to c_{exp} , while figure 4.10 represents figure 4.9 restricted to [10:12,5] ppm interval. In figure 4.11 the comparison of Tel21 ^1H NMR spectra of the sample with $c_{exp} = 0,397$ mM before and after annealing is presented, while figure 4.12 represents figure 4.11 restricted to [10:12,5] ppm interval. The similar comparison is presented in figure 4.13, but for $c_{exp} = 0,072$ mM, whereas figure 4.14 represents figure 4.3 restricted to [10:12,5] ppm interval.

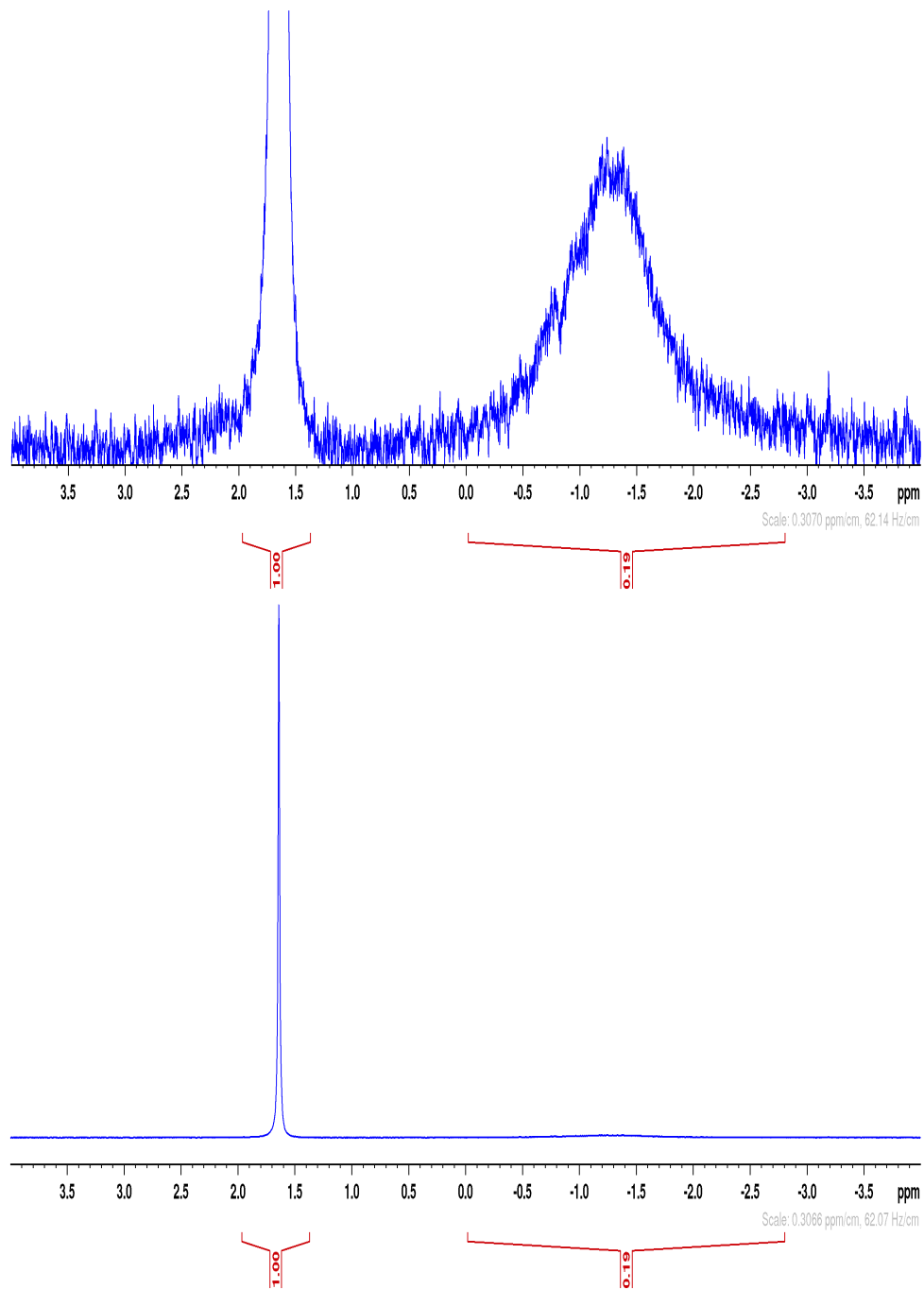


Figure 4.8: An example of experimental determination of Tel21 oligonucleotide concentration $c_{exp} = 0,482$ mM with the help of ^{31}P spectrum. The left peak corresponds to buffer signal, the right peak corresponds to an oligonucleotide signal(containing 20 phosphorus nuclei).

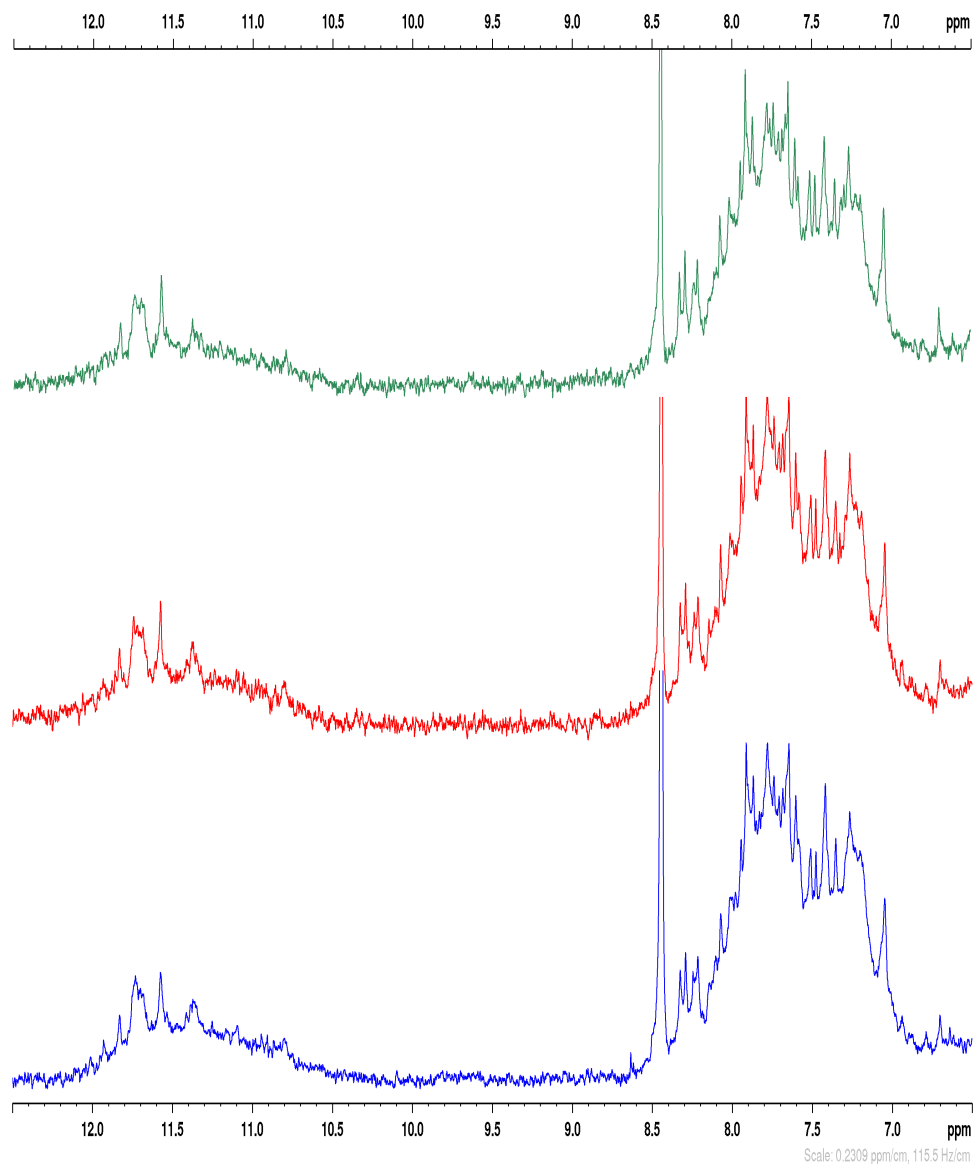


Figure 4.9: A comparison of Tel21 ^1H NMR spectra in respect to c_{exp} . **Blue** line corresponds to $c_{exp} = 0,482$ mM, **red** line corresponds to $c_{exp} = 0,397$ mM, **green** line corresponds to $c_{exp} = 0,072$ mM. Spectral noise levels are different due to different NS values, processing parameters are identical. Spectra were scaled to maximum intensity.

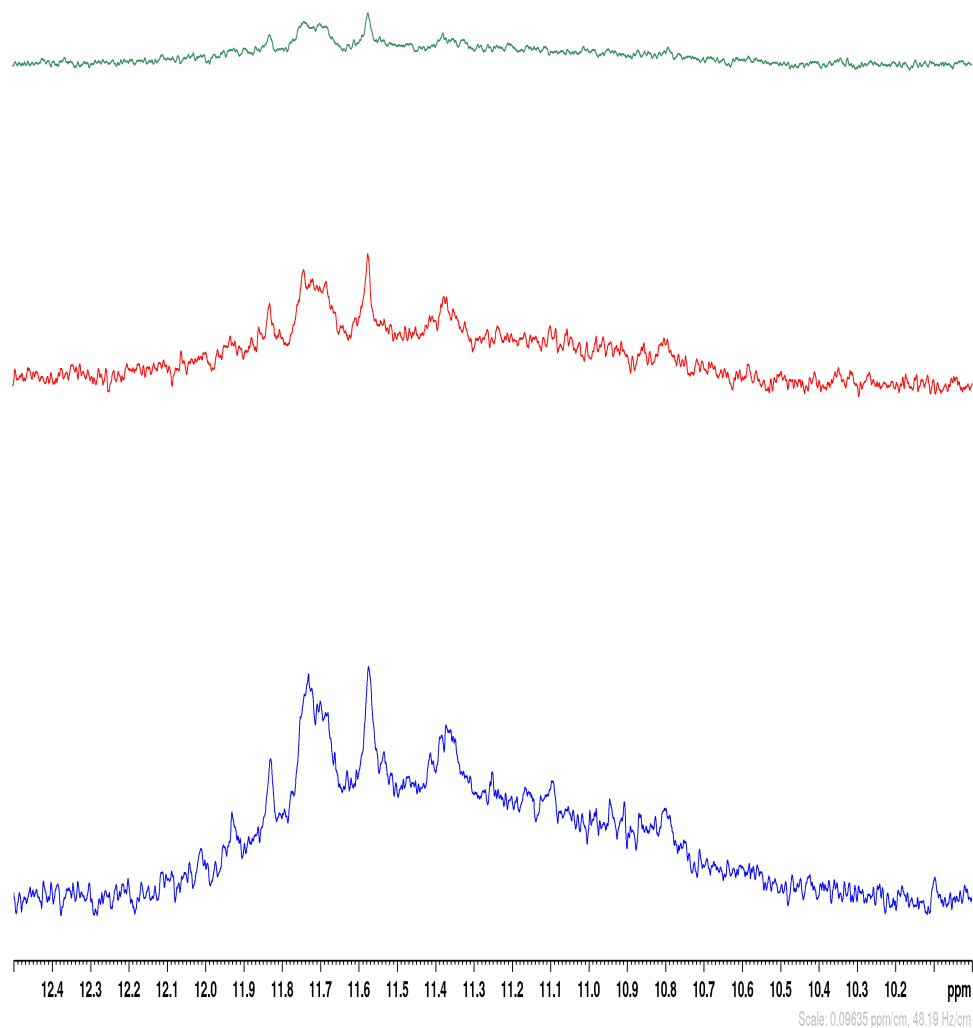


Figure 4.10: A comparison of Tel21 ¹H NMR spectra in respect to c_{exp} restricted to [10:12,5] ppm interval. **Blue** line corresponds to $c_{exp} = 0,482$ mM, **red** line corresponds to $c_{exp} = 0,397$ mM, **green** line corresponds to $c_{exp} = 0,072$ mM. Spectral noise levels are different due to different NS values, processing parameters are identical. Spectra were intensity calibrated in respect to DSS peak at 0 ppm.

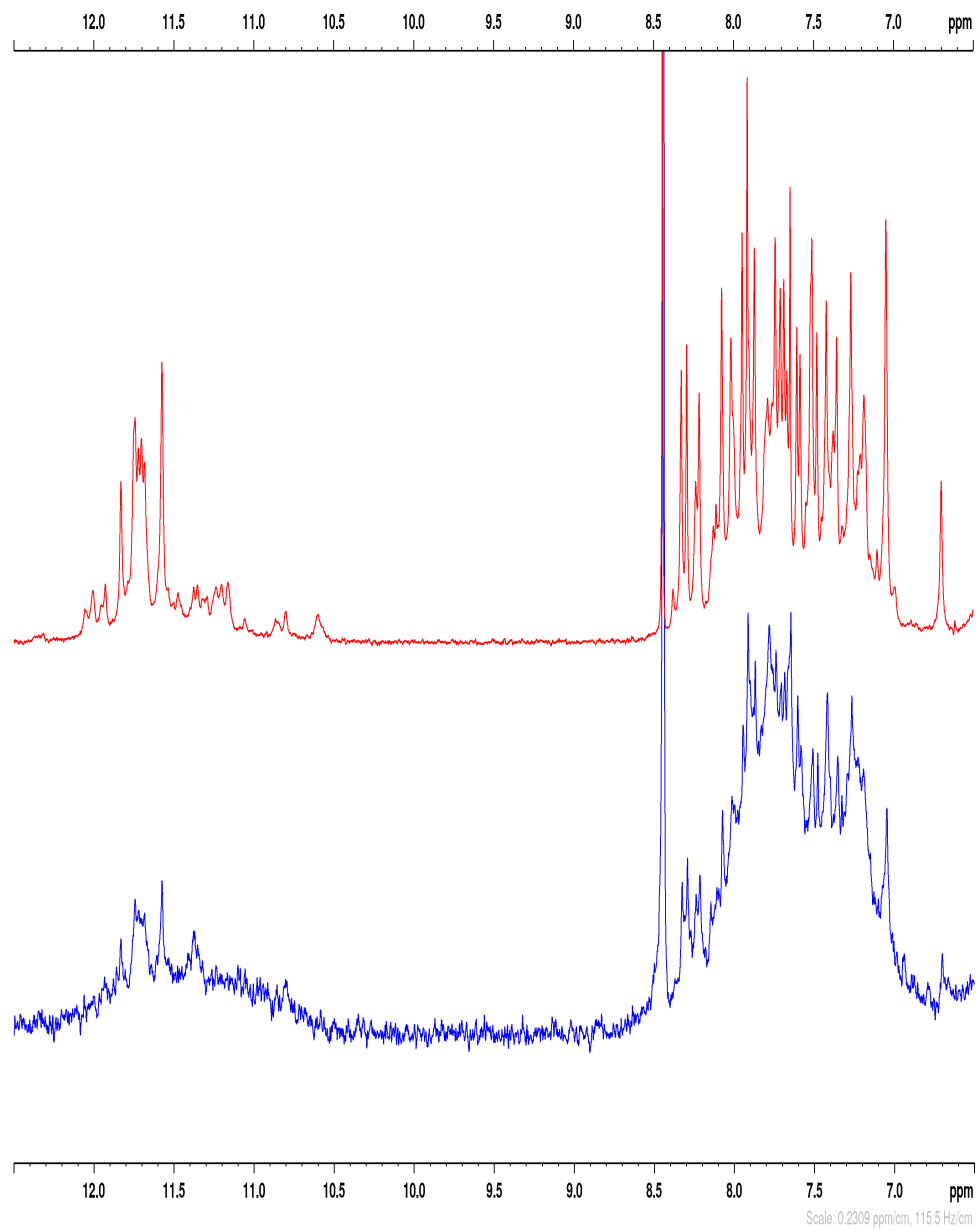


Figure 4.11: Tel21 ^1H NMR spectra for $c_{exp} = 0, 397$ mM before and after annealing. **Blue** line corresponds to the sample before annealing, **red** line corresponds to the sample after annealing. Spectral noise levels are different due to different NS values, processing parameters are identical. Spectra were scaled to maximum intensity.

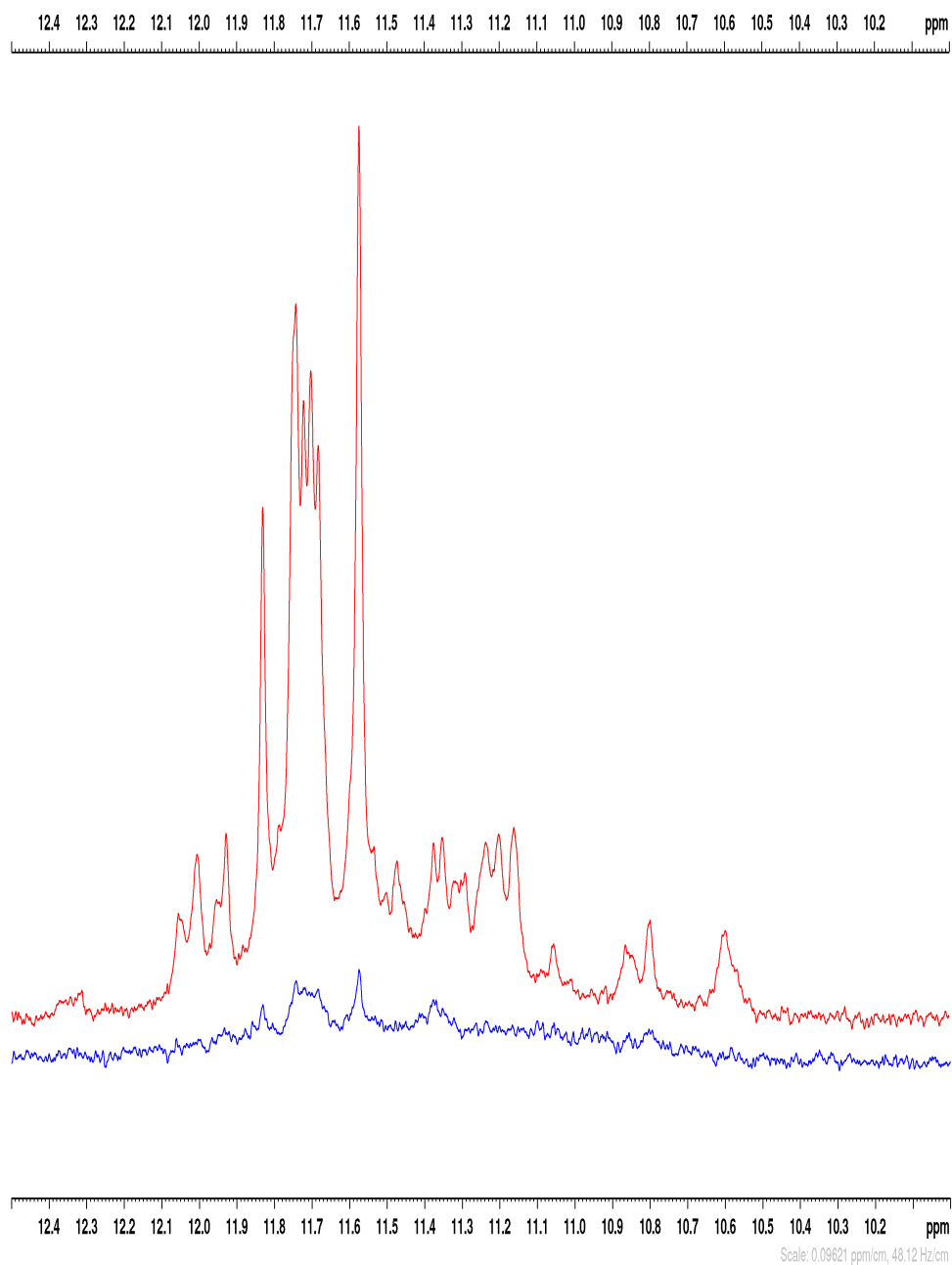


Figure 4.12: Tel21 ^1H NMR spectra for $c_{exp} = 0,397$ mM before and after annealing restricted to $[10:12,5]$ ppm interval. **Blue** line corresponds to the sample before annealing, **red** line corresponds to the sample after annealing. Spectral noise levels are different due to different NS values, processing parameters are identical. Spectra were intensity calibrated in respect to DSS peak at 0 ppm.

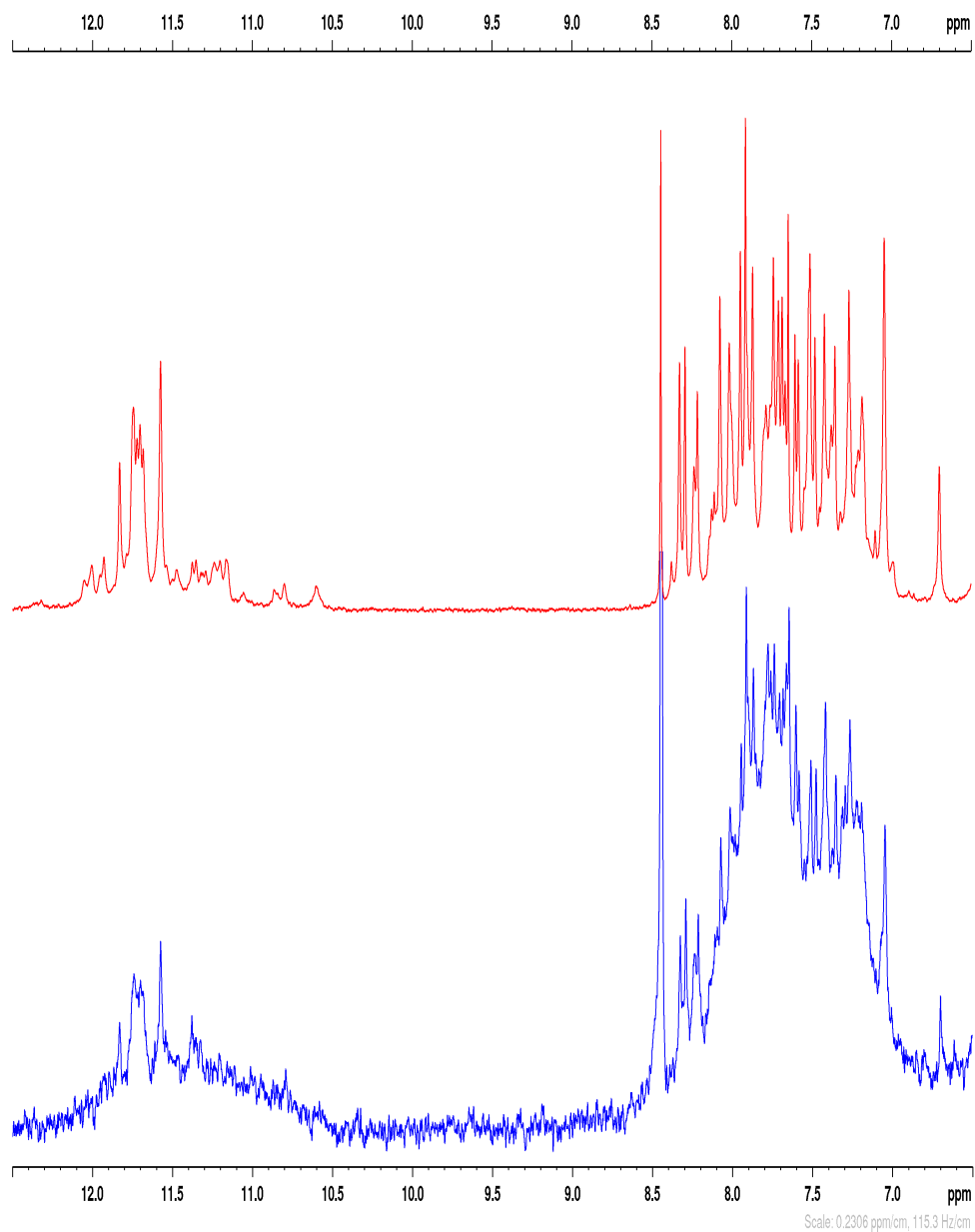


Figure 4.13: Tel21 ^1H NMR spectra for $c_{exp} = 0,072$ mM before and after annealing. **Blue** line corresponds to the sample before annealing, **red** line corresponds to the sample after annealing. Spectral noise levels are different due to different NS values, processing parameters are identical. Spectra were scaled to maximum intensity.

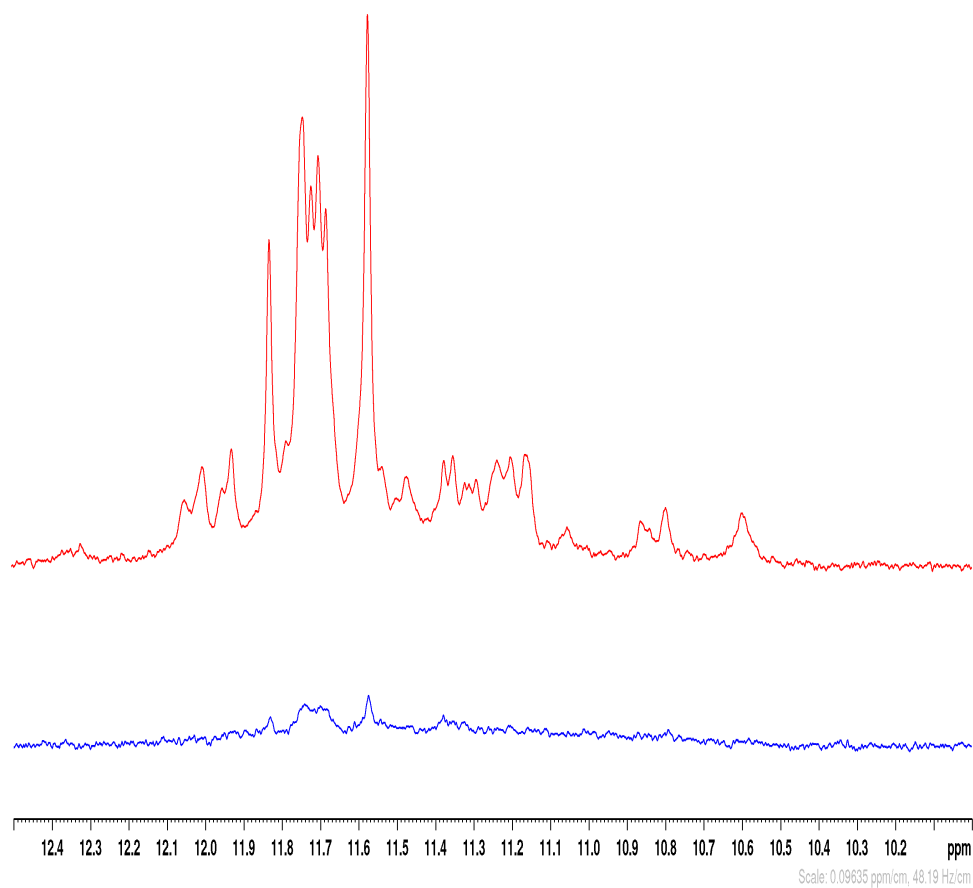


Figure 4.14: Tel21 ^1H NMR spectra for $c_{exp} = 0,072$ mM before and after annealing restricted to [10:12,5] ppm interval. **Blue** line corresponds to the sample before annealing, **red** line corresponds to the sample after annealing. Spectral noise levels are different due to different NS values, processing parameters are identical. Spectra were intensity calibrated in respect to DSS peak at 0 ppm.

4.2 Raman experiment results

The Raman experiment was carried out for the Tel22 nucleotide sequence only with the usage of sample numbers 1,2,3 from table 3.1. In the same table, experimentally obtained Tel22 sample concentrations c_{exp} , related to the Raman experiment, are presented under c_{exp} .

In figure 4.15 we may observe a comparison of Tel22 Raman spectra in respect to experimentally determined c_{exp} . In figure 4.16 the comparison of Tel22 Raman spectra of the $c_{exp} = 0,381$ mM before and after annealing is presented. A similar comparison is presented in figure 4.17, but for $c_{exp} = 0,095$ mM.

In the obtained spectra from Raman experiments we are primarily interested in certain Raman markers, that play key role in specific G-quadruplex structure determination. For instance, general G-quadruplex Raman marker, if present, is situated at approximately 1715 cm^{-1} reflecting the dG-O6 Hoogsteen hydrogen bonding within G-tetrads. The Raman shift at approximately 1580 cm^{-1} indicates strong guanine band linked to interbase H-bonding at dG-N2H express maintenance of G-quadruplex structure in Na^+ or K^+ buffers by formerly unfolded oligonucleotides¹. Distribution of glycosidic torsion angles of guanosines(C2'-endo/syn-G at approximately 1326 cm^{-1} and C2'-endo/anti-G at approximately 1338 cm^{-1}) appear to be the key Raman markers that are used to differentiate G-quadruplex structures [27]. Changes in guanosine glycosidic torsion angles cause band changes at approximately 1338 cm^{-1} . On the other hand, for differentiation of G-quadruplex structures we could use other distinctive pair of Raman markers, that can be found approximately at 671 cm^{-1} (C2'-endo/syn-dG) and approximately at 686 cm^{-1} (C2'-endo/anti-dG) playing the same role as the markers 1326 cm^{-1} and 1338 cm^{-1} , however using 1326 cm^{-1} , 1338 cm^{-1} markers we consider to be more simple.

¹Table of Raman markers can be found e.g. here [27]

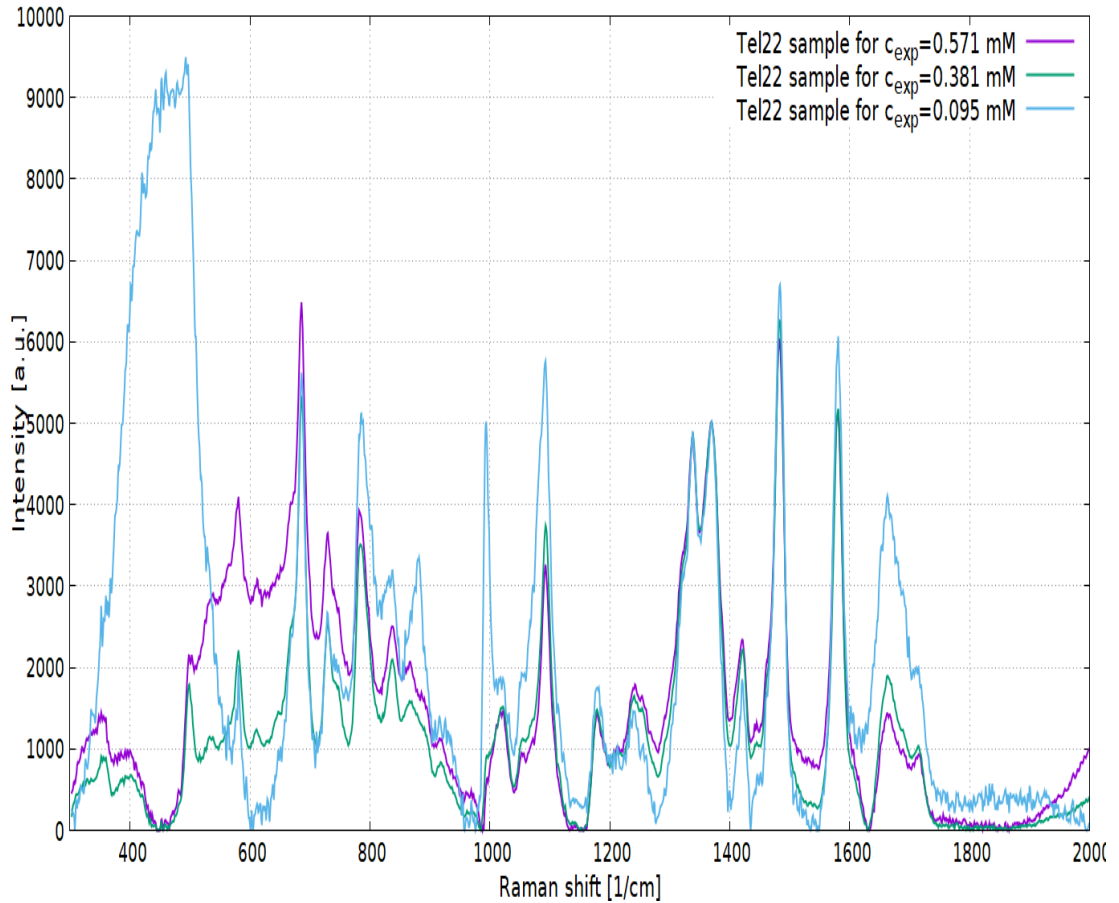


Figure 4.15: Tel22 Raman spectra in respect to c_{exp} . Purple line corresponds to $c_{exp} = 0,571$ mM, green line corresponds to $c_{exp} = 0,381$ mM, Blue line corresponds to $c_{exp} = 0,095$ mM. Spectra were processed in Gnuplot.

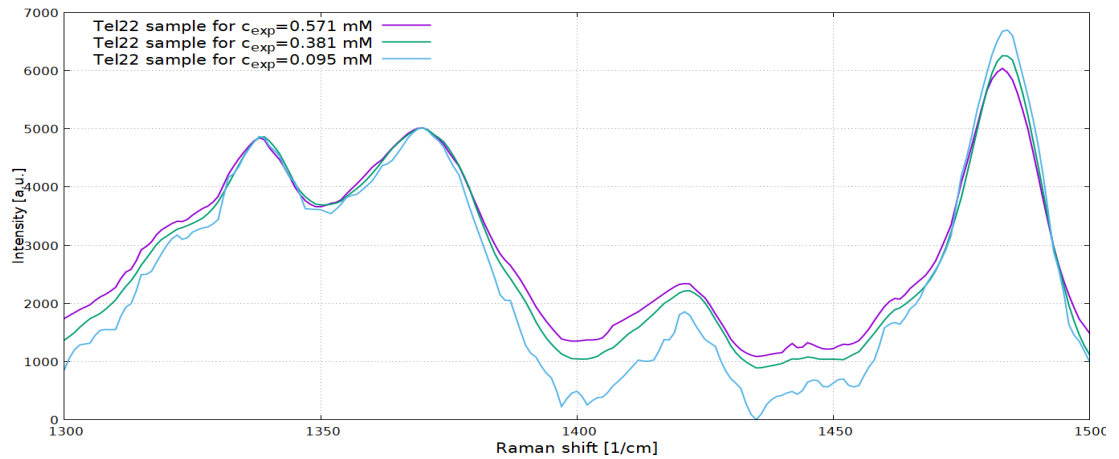


Figure 4.16: Tel22 Raman spectra in respect to c_{exp} restricted to $[1300:1500]$ cm^{-1} . Purple line corresponds to $c_{exp} = 0,571$ mM, green line corresponds to $c_{exp} = 0,381$ mM, Blue line corresponds to $c_{exp} = 0,095$ mM. Spectra were processed in Gnuplot.

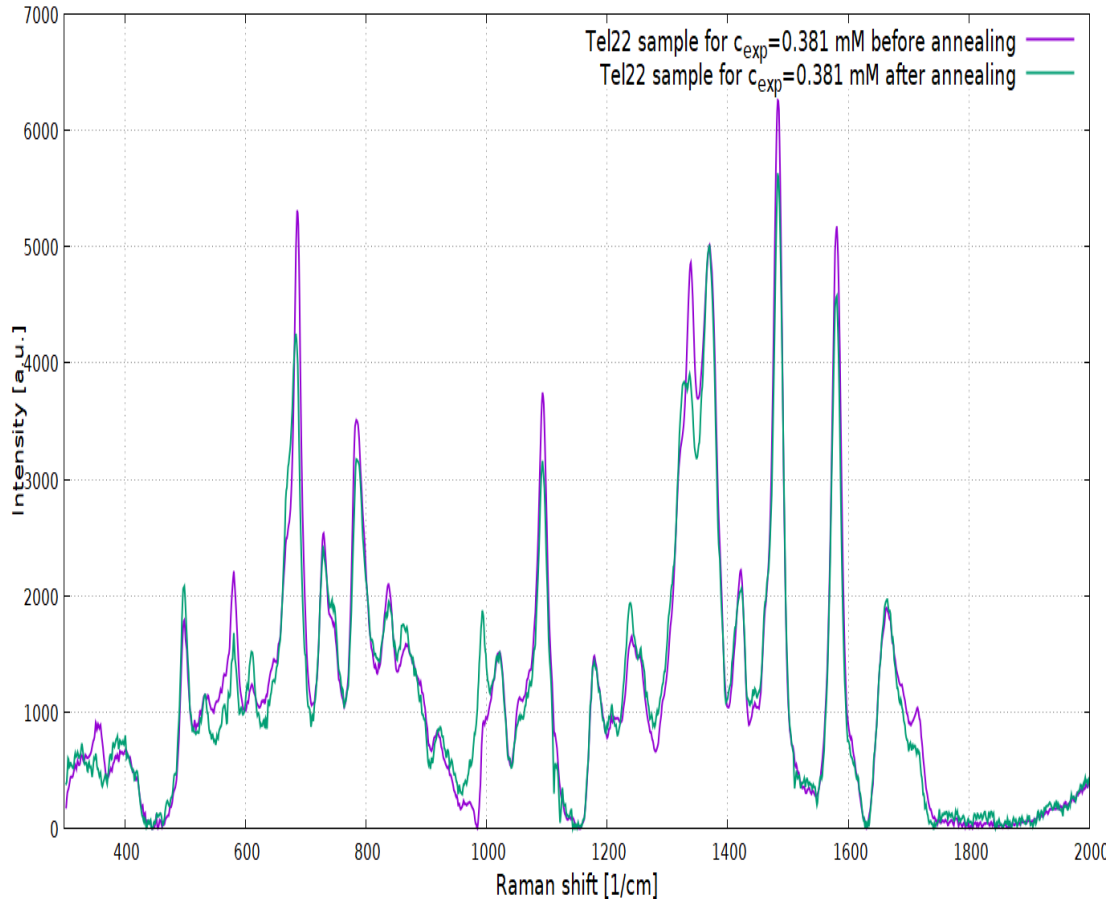


Figure 4.17: Tel22 Raman spectra for $c_{exp} = 0,381$ mM before and after annealing. Purple line corresponds to the sample before annealing, green line corresponds to the sample after annealing. Spectra were processed in Gnuplot.

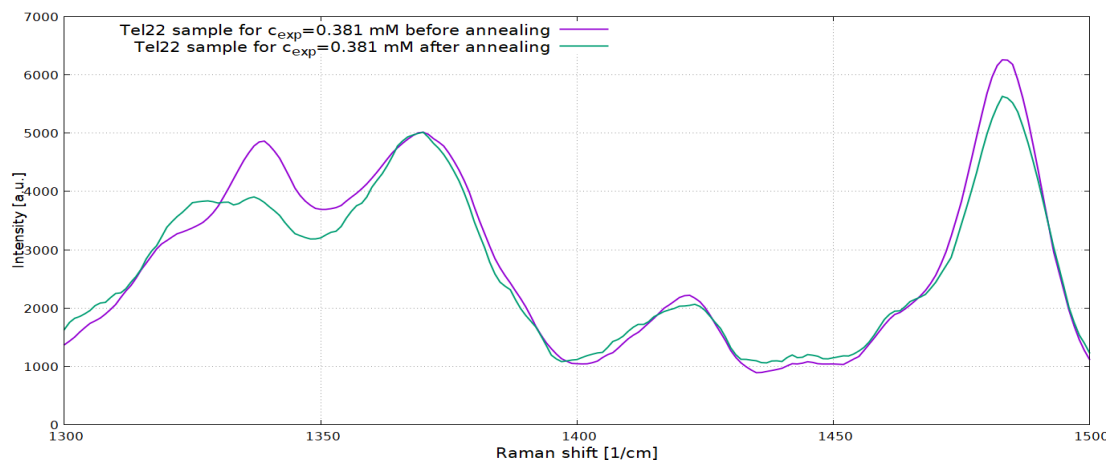


Figure 4.18: Tel22 Raman spectra for $c_{exp} = 0,381$ mM before and after annealing restricted to $[1300:1500]$ cm^{-1} . Purple line corresponds to the sample before annealing, green line corresponds to the sample after annealing. Spectra were processed in Gnuplot.

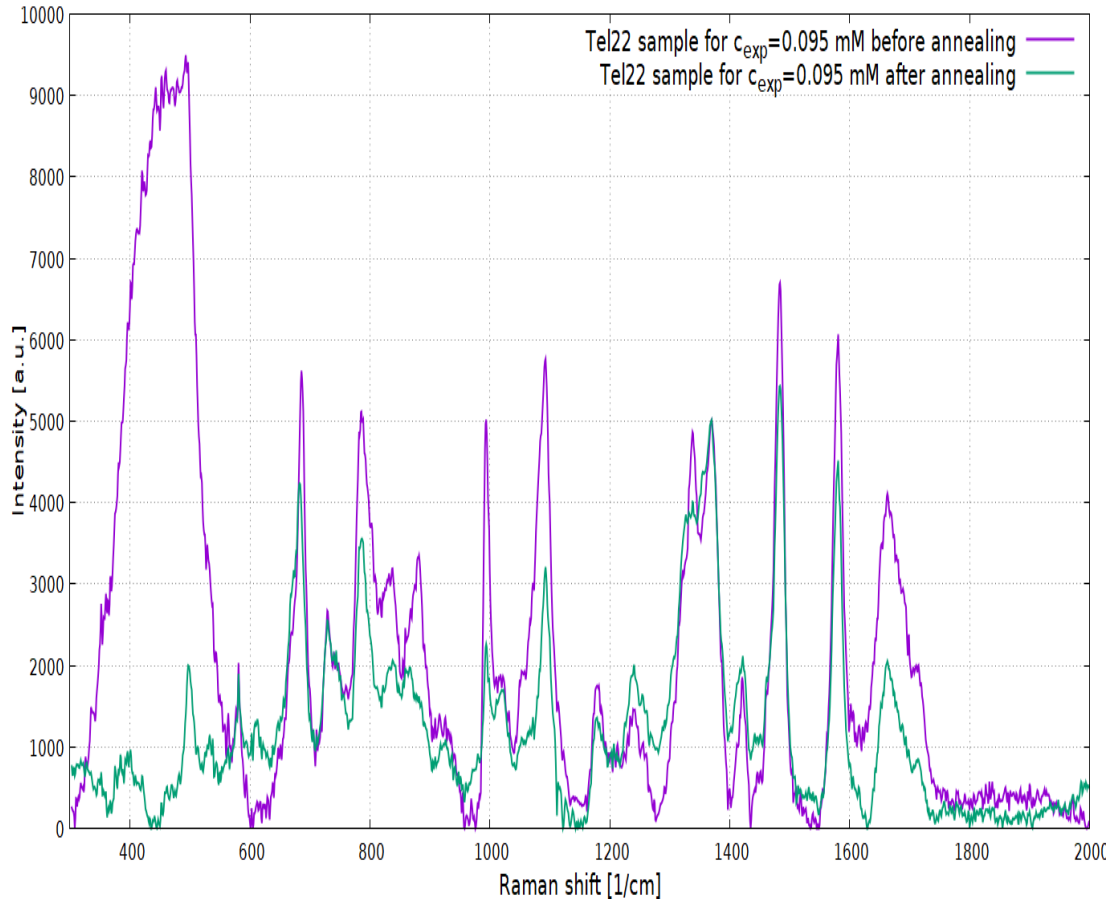


Figure 4.19: Tel22 Raman spectra for $c_{exp} = 0,095$ mM before and after annealing. Purple line corresponds to the sample before annealing, green line corresponds to the sample after annealing. Spectra were processed in Gnuplot.

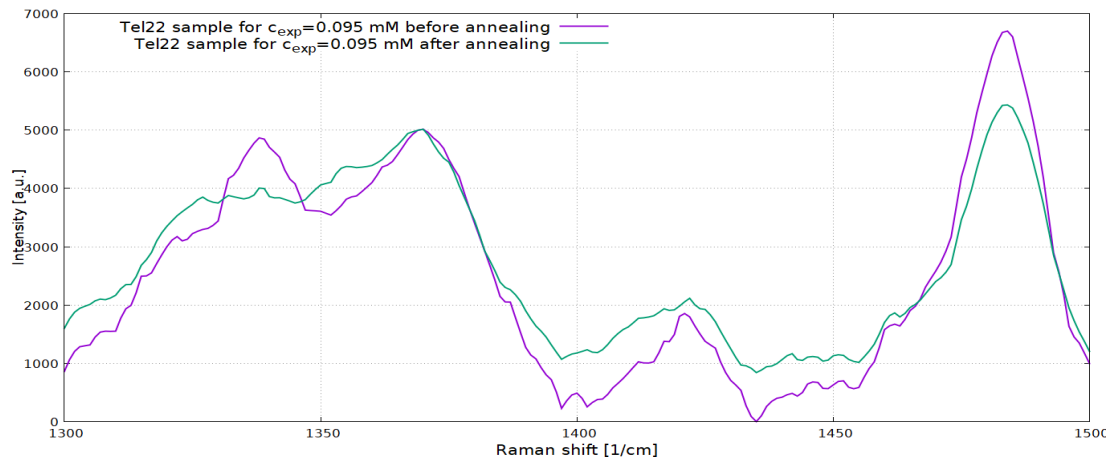


Figure 4.20: Tel22 Raman spectra for $c_{exp} = 0,095$ mM before and after annealing restricted to $[1300:1500]$ cm^{-1} . Purple line corresponds to the sample before annealing, green line corresponds to the sample after annealing. Spectra were processed in Gnuplot.

5. Discussion

The processes mentioned in the introduction, conformational changes of human telomeric quadruplexes, were observed with the aid of Raman spectroscopy [27, 28]. We were expected to observe same phenomena also with the usage of NMR spectroscopy.

First of all, as we see from tables 3.1 and 3.2, experimentally determined oligonucleotide concentrations c_{exp} of different samples do not match those previously calculated ones c_{calc} , being 2-3 times less than expected in cases of both oligonucleotide sequences. That happened, mainly, due to the transportational losses of the sample. It could not happen due to annealing, as it was determined that concentration values obtained from ^{31}P NMR spectra before and after annealing can differ only within a 20% range.

5.1 Tel22 NMR experiment

There are 4 imino hydrogens in one tetrad, all of different chemical environments. In a Tel22 nucleotide sequence, there are 12 guanines that can be involved in the intramolecular G-quadruplex assembling. Those 12 guanines can form 3 G-tetrad of a G-quadruplex's core, consequently we may expect 12 peaks in a [10:12,5]ppm interval to be observed in a ^1H NMR spectrum. However, in figures 4.2-4.7 a greater quantity of peaks(>12) is observed, which means that a Tel22 nucleotide sequence, in a K^+ -rich chemical environment, forms a mixture of multiple stable G-quadruplex conformations, regardless of the sample concentration, annealing or its absence. Indeed, analysing figures 4.2-4.7, we can conclude that no dramatic structural change in spectra is observable, neither due to oligonucleotide concentration change(decrease), nor due to annealing. However, there is a significant growth of the signal's intensity for samples after annealing(figure 4.4-4.7) observable. Also, as we see from figures 4.4 and 4.5, the spectrum corresponding to the sample after the first annealing(figures 4.4 and 4.5 middle spectrum) looks identical to the spectrum corresponding to the sample after the second annealing(figures 4.4 and 4.5 middle spectrum). The peak at 8.5 ppm should correspond to an impurity in the sample that probably entered the sample during the experiment.

We may compare obtained ^1H NMR spectra with the one in figure 5.1[6] now. The obtained spectra, in the [10:12,5] ppm interval do look very similar to the one in figure 5.1. In figures 4.2 and 4.3, we see a higher noise level in the spectrum for a more concentrated sample ($c_{exp} = 0,571\text{mM}$) being present due to a lower NS value.

In figures 4.6 and 4.7, where both spectra have the same NS value and identical data processing parameters, we also see a higher noise level(wider baseline) in the bottom spectrum than in the top one(for both intervals of interest

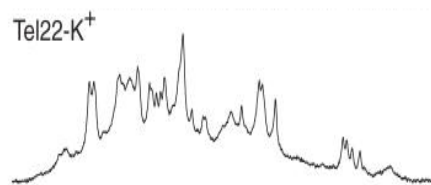


Figure 5.1: Imino regions of 1D ^1H NMR spectrum of Tel22 in a K^+ solution at $\text{K} = 298,15\text{ K}$ (adapted from [6])

- [10:12,5] ppm and [6,5:9] ppm). According to the provided facts, we may conclude that a wider baseline is present in spectra before annealing (bottom spectra in figures 4.3-4.7) due to a multimolecularity of the G-quadruplex structures, since the shape and quantity of peaks is given by the predominant unimolecular G-quadruplex structures. In other words, multimolecular G-quadruplex structures are reflected in spectra as a wider baseline.

It is known, that transversal relaxation time T_2 is connected to the loss of phase coherence among nuclei in a sample. The shorter the T_2 values is, the broader the spectral peak is. Greater molecules possess shorter transverse relaxation times, consequently, NMR signal obtained from bigger molecules (of higher molecular mass) decays more swiftly, resulting in line broadening in a NMR spectrum. Consequently, annealing results in multimolecular G-quadruplex structures' disintegration and creation of unimolecular G-quadruplex assemblies. Generally speaking, it is the reason why the top spectrum in figures 4.6 and 4.7, even though acquisition and data processing parameters are identical for both experiments, has sharper peaks and a lower baseline comparing to the bottom spectrum. As for [6,5:9] ppm interval corresponding to aromatic hydrogens, process of the base line reduction is observable within the peak group. In figures 4.4 and 4.6 we can only see narrowing of certain peaks within [6,5:9] ppm interval of the spectrum.

5.2 Tel21 NMR experiment

In a Tel21 sequence, as in a Tel22 nucleotide sequence, there are 12 guanines that can be involved in the intramolecular G-quadruplex assembling. As in case of Tel21, we may expect 12 peaks in a [10:12,5] ppm interval to be observed in a ^1H NMR spectrum. However, in figures 4.9-4.14 a greater quantity of peaks (>12) is observed, which means that a Tel21 nucleotide sequence, in a K^+ -rich chemical environment, forms a mixture of multiple stable G-quadruplex conformations, regardless of the sample concentration, annealing or its absence. In figure 4.9 we clearly see that concentration changes do not affect the spectrum shape in the entire [6,5:12,5] ppm interval. In figure 4.10 we see that a higher concentration sample contributes to higher intensity spectrum, but with no structural changes observed between bottom and middle spectra, while top spectrum can not be qualitatively evaluated due to the weak signal. In figures 4.11 and 4.12 we see an impressive intensity growth for the sample after annealing in [10:12,5] ppm interval, however no shape changes can be observed, as spectra in figure 4.12 closely resemble each other, based on the amount of signal obtained in bottom spectrum for more concentrated sample. The spectral peaks also became more narrow, comparing to the bottom spectrum. The same phenomenon, observed in figures 4.11 and 4.12, is now observed in figures 4.13 and 4.14. As for [10:12,5] ppm interval, similarly as in case of Tel22 sequence, we can state, that no observable shape change of spectra in figures 4.9-4.14 was observed neither due to oligonucleotide concentration change (decrease), nor due to annealing. Only change in intensity of a signal is observable. The peak at 8.5 ppm should correspond to an impurity in the sample that probably entered the sample during the experiment, as in case of the Tel22 sequence experiment.

In figures 4.9 and 4.10, we see a higher noise level in the spectrum with

$c_{exp} = 0,482$ mM and $c_{exp} = 0,397$ mM than with $c_{exp} = 0,072$ mM being present due to a lower NS value.

In figures 4.11 and 4.12, where both spectra have the same NS value and identical data processing parameters, we also see a higher noise level (wider baseline) in the bottom spectrum than in the top one (for both intervals of interest - [10:12,5] ppm and [6,5:9] ppm). As in case of Tel22 nucleotide sequence, we may conclude that a wider baseline is present in spectra before annealing (bottom spectra in figures 4.11-4.14) due to a multimolecularity of the G-quadruplex structures, since the shape and quantity of peaks is given by the predominant unimolecular G-quadruplex structures. All the conclusions connected to multimolecular G-quadruplex structures' disintegration, creation of unimolecular G-quadruplex assemblies and their effects on the spectra are identical, as for Tel22 sequence.

5.3 Tel22 Raman experiment

Spectra were calibrated to the same intensity value at 1370 cm^{-1} .¹ Analysing figures 4.15 and 4.16, we come to the conclusion that no changes in characteristic bands are observable. Therefore, no G-quadruplex conformational change is observed in respect to different concentrations of the sample.

In figures 4.17 and 4.18, with the help of specific Raman markers 1326 cm^{-1} and 1338 cm^{-1} , we witness G-quadruplex conformational change of a $c_{exp} = 0,381\text{ mM}$ concentration sample before annealing from parallel to hybrid (3 + 1) conformation after annealing, reflected in characteristic peak shape change [27]. Similar characteristic peak shape change is observable in figures 4.19 and 4.20, corresponding to $c_{exp} = 0,095$ mM.

5.4 G4 topologies observed

Due to the presented facts, we come to the conclusion that parallel conformation G-quadruplexes are not observable with the help of an NMR spectrometer, therefore, according to the experiment results of two previous paragraphs, they should be multimolecular. On the other hand, unimolecular 3 + 1 hybrid conformation of G-quadruplexes are present (observable by NMR spectroscopy, as well as Raman spectroscopy).

The variability of stacked G-quadruplex multimers was previously proposed [29, 30]. A more precise structure determination of the G-quadruplexes should be done with 2D NMR spectroscopy methods, although there is an application limit of NMR in respect to molecular mass. Some additional CD spectroscopy measurements can be executed as well.

¹In figures 4.15 and 4.19, we see abnormal intensity growth in the case of the sample with $c_{exp} = 0,095\text{ mM}$ concentration in approximately [200:600] cm^{-1} interval due to incorrect buffer signal subtraction.

Conclusion

In this bachelor thesis, we have conducted NMR and Raman spectroscopy experiments with two human telomere nucleotide sequences. The first one, Tel21, consisted of 21 nucleotides - $G_3(G_3TTA)_3$. The second sequence, Tel22, consisted of 22 nucleotides - $AG_3(G_3TTA)_3$. Our goal was to explore the dependence of the G-quadruplex type fold on the ionic composition (K^+), as well as exploring structural differences of a quadruplex before and after annealing.

We found that :

- Experimentally determined oligonucleotide concentrations of different samples did not match those previously calculated ones. It did not happen due to annealing, as it was determined that concentration values obtained from ^{31}P NMR spectra before and after annealing can differ only within a 20% range. Therefore, transportational losses of the sample could be considered as the main reason.
- Spectra of Tel22 and Tel21 samples before annealing in respect to different concentrations did not differ from each other shapewise, but only by intensity. More concentrated samples exhibited higher intensity of a signal.
- Spectra of Tel22 and Tel21 samples after annealing illustrated remarkable growth of the signal's intensity for every particular concentration. Most of the spectra peaks narrowed. The significant baseline reduction was observable.
- 1H NMR spectra after one or multiple annealings do not differ.
- Our spectra, obtained for Tel22 samples do resemble one gained earlier [6].
- Tel22 nucleotide sequence, as well as Tel21 nucleotide sequence in a K^+ -rich chemical environment, forms a mixture of multiple stable monomeric G-quadruplex conformations, that are the same regardless of the sample concentration, annealing or its absence.
- The wider baseline was present in spectra before annealing due to a multimolecularity of the G-quadruplex structures. Therefore, parallel conformation G-quadruplexes are not observable with the help of an NMR spectrometer. An annealing resulted in disintegration of multimolecular G-quadruplex structures and creation of unimolecular G-quadruplex structures.
- No G-quadruplex conformational changes could be observed for Tel22 samples in respect to different concentrations before annealing with the aid of Raman spectroscopy.
- Conformational change of G-quadruplex from parallel to hybrid (3 + 1) conformation after annealing could be observed with the help of Raman spectroscopy.

Bibliography

- [1] Mcmurry J. *Organic Chemistry*. 7th edition. Thomson Higher Education, 10 Davis Drive, Belmont, CA 94002-3098, USA, 2008.
- [2] Nelson D. and Cox M. *Principles of Biochemistry*. 5th edition. W.H.Freeman and Company Ltd., 41 Madison Avenue, New York, 10010, 2008.
- [3] J.D. Watson and F.H.C. Crick. Molecular structure of nucleic acids. *Nature*, 171:737–738, 1953.
- [4] E. N. Nikolova, E. Kim, A. A. Wise, P. J. O'Brien, and I. Andricioaei. Transient Hoogsteen base pairs in canonical duplex DNA. *Nature*, 470:498–502, 2011.
- [5] D. J. Patel, S. Bouaziz, A. Kettani, and Y. Wang. Structures of guanine-rich and cytosine-rich quadruplexes formed in vitro by telomeric, centromeric, triplet repeat disease DNA sequences. *Oxford Handbook of Nucleic Acid Structures*, pages 389–453, 1999.
- [6] A. Ambrus, C. Ding, D. Jixun, T. Bialis, R.A. Jones, and Y. Danzhou. Human telomeric sequence forms a hybrid-type intramolecular G-quadruplex structure with mixed parallel/antiparallel strands in potassium solution. *Nucleic Acids Research*, 34:2723–2735, 2006.
- [7] 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*, 266:2011–2015, 1994.
- [8] A.M. Zahler, J.R. Williamson, T.R. Cech, and D.M. Prescott. Inhibition of telomerase by G-quartet DNA structures. *Nature*, 350:718–720, 1991.
- [9] M. Lei, A.J. Zaug, E.R. Podell, and T.R. Cech. Switching human telomerase on and off with hPOT1 protein in vitro. *The journal of biological chemistry*, 102:10864–10869, 2005.
- [10] L. Oganessian, I.K. Moon, T.M. Bryan, and M.B. Jarstfer. Extension of G-quadruplex DNA by ciliate telomerase. *The EMBO journal*, 25:1148–1159, 2006.
- [11] N.G. Dolinnaya, A.M. Oglobina, and M.G. Yakubovskaya. Structure, Properties, and Biological Relevance of the DNA and RNA G-Quadruplexes: Overview 50 Years after Their Discovery. *Biochemistry(Moscow)*, 81:1602–1649, 2016.
- [12] Y. Wang and D.J. Patel. Solution structure of the human telomeric repeat d[AG₃(T₂AG₃)₃] G-tetraplex. *Structure*, 1:263–282, 1993.
- [13] K.N. Luu, A.T. Phan, V. Kuryavyi, L. Lacroix, and D.J. Patel. Structure of the Human Telomere in K⁺ Solution: An Intramolecular (3+1) G-quadruplex Scaffold. *Journal of American chemical society*, 128:9963–9970, 2006.

- [14] G.N. Parkinson, M.P. Lee, and S. Neidle. Crystal structure of parallel quadruplexes from human telomeric DNA. *Nature*, 417:876–880, 2002.
- [15] L. Ying, J.J. Green, H. Li, D. Klenerman, and S. Balasubramanian. Studies on the structure and dynamics of the human telomeric G quadruplex by single-molecule fluorescence resonance energy transfer. *Proceedings of the National Academy of Sciences of the United States of America*, 25:14629–14634, 2003.
- [16] J. Palacký, M. Vorlicková, I. Kejnovská, and P. Mojzes. Polymorphism of human telomeric quadruplex structure controlled by DNA concentration: a Raman study. *Nucleic Acids Research*, 41:1005–1016, 2013.
- [17] D. Renčiuk, I. Kejnovská, P. Školáková, K. Bednářová, J. Motlová, and M. Vorlíčková. Arrangements of human telomere DNA quadruplex in physiologically relevant K^+ solutions. *Nucleic Acids Research*, 37(19):6625–6634, 2009.
- [18] Hore P.J. *Nuclear Magnetic Resonance*. 1st edition. Oxford University Press Inc., 1995.
- [19] William Reusch. Department of Chemistry, Michigan State University. Nuclear Magnetic Resonance Spectroscopy. <https://www2.chemistry.msu.edu/faculty/reusch/VirtTxtJml/Spectrpy/nmr/nmr1.htm>, 2013. Accessed: 2019-12-25.
- [20] Gunther H. *NMR Spectroscopy. Basic Principles, Concepts and Applications in Chemistry*. 3rd edition. Wiley-VCH Verlag GmbH and Co., 2013.
- [21] Rudraksha Dutta Majumdar. *A nuclear magnetic resonance spectroscopic investigation of the molecular structure and aggregation behavior of asphaltene*. PhD thesis, University of Lethbridge, 12 2015.
- [22] Teodor Parella. *Pulse Program Catalogue: I. 1D & 2D NMR Experiments*. Bruker BioSpin GmbH, 2010.
- [23] Tsang-Lin Hwang and A.J. Shaka. Water Suppression That Works. Excitation Sculpting Using Arbitrary Waveforms and Pulsed Field Gradients. *Journal of Magnetic Resonance, Series A*, 112(2):275–279, feb 1995.
- [24] Larkin P. *Infrared and Raman Spectroscopy. Principles and Spectral Interpretation*. 1st edition. Elsevier, 2011.
- [25] J. Ferraro, K. Nakamoto, and C. Brown. *Introductory Raman Spectroscopy*. 2nd edition. Academic Press, 2003.
- [26] R.I. Mathad and D. Yang. G-Quadruplex Structures and G-Quadruplex-Interactive Compounds. *Telomeres and Telomerase. Methods in Molecular Biology*, 735, 2011.
- [27] J. Palacký, P. Mojzeš, I. Kejnovská, and M. Vorlíčková. Does Raman spectroscopy recognize different G-quadruplex arrangements? *Wiley. Journal of Raman Spectroscopy*, pages 1–12, 2019.

- [28] J. Palacký, M. Vorlíčková, I. Kejnovská, and P. Mojzeš. Polymorphism of human telomeric quadruplex structure controlled by the DNA concentration: a Raman study. *Nucleic Acids Research*, 41(2):1005–1016, 2013.
- [29] Rashid M. Abu-Ghazalah, Steve Rutledge, Lewis W. Y. Lau, David N. Dubins, Robert B. Macgregor, and Amr S. Helmy. Concentration-Dependent Structural Transitions of Human Telomeric DNA Sequences. *Biochemistry*, 51(37):7357–7366, sep 2012.
- [30] Brahim Heddi and Anh Tuan Phan. Structure of Human Telomeric DNA in Crowded Solution. *Journal of the American Chemical Society*, 133(25):9824–9833, 6 2011.

List of Figures

1.1	Major pyrimidine and purine bases of nucleic acids (adapted from [1])	4
1.2	Major pyrimidine and purine bases of nucleic acids (adapted from [1])	5
1.3	Watson–Crick and Hoogsteen A-T and G-C base pair structures. The Hoogsteen geometry can be achieved by the purine rotation around the glycosidic bond (χ) and base-flipping (θ), affecting simultaneously C-8 and C-1'(coloured yellow) (adapted from [4]) . .	6
1.4	Left - Guanosine quadruplex base-pairing pattern; Right - Guanosine quadruplex's possible nucleotide chains' orientations (adapted from [2])	7
1.5	Structure of intramolecular G-quadruplexes formed by the human telomeric sequence d[AGGG(TTAGGG) ₃]: (a, b) in Na ⁺ solution; (c, d) in a K ⁺ -containing crystal. Loops are colored red; anti and syn guanines are colored cyan and magenta, respectively. O4' atoms are colored yellow. (adapted from [13])	8
1.6	Schematic structures of the (3 + 1) core-containing G-quadruplexes formed by (a) the four-repeat Tetrahymena telomeric d[(TTGGGG) ₄] sequence in a Na ⁺ solution, (b) the three-repeat human telomeric d[GGG(TTAGGG) ₂ T] sequence in a Na ⁺ solution, and (c) the four-repeat human telomeric d[TAGGG(TTAGGG) ₃] sequence in a K ⁺ solution (adapted from [13]).	9
2.1	Chemical shifts of hydrogen nuclei in respect to a different atomic surrounding (adapted from [19]).	13
2.2	Operation of the spin echo experiment (adapted from [21]).	15
2.3	Zgesgp pulse sequence (adapted from [22])	16
4.1	An example of experimental determination of Tel22 oligonucleotide concentration $c_{exp} = 0,571$ mM with the help of ³¹ P spectrum. The left peak corresponds to a buffer signal, the right peak corresponds to a oligonucleotide signal(containing 21 phosphorus nuclei). . . .	25
4.2	A comparison of Tel22 ¹ H NMR spectra in respect to c_{exp} . Blue line corresponds to $c_{exp} = 0,571$ mM, red line corresponds to $c_{exp} = 0,381$ mM, green line corresponds to $c_{exp} = 0,095$ mM. Spectral noise levels are different due to different NS values, processing parameters are identical. Spectra were scaled to maximum intensity.	26
4.3	A comparison of Tel22 ¹ H NMR spectra in respect to c_{exp} restricted to [10:12,5] ppm interval. Blue line corresponds to $c_{exp} = 0,571$ mM, red line corresponds to $c_{exp} = 0,381$ mM, green line corresponds to $c_{exp} = 0,095$ mM. Spectral noise levels are different due to different NS values, processing parameters are identical. Spectra were intensity calibrated in respect to DSS peak at 0 ppm.	27

4.4	Tel22 ^1H NMR spectra for $c_{exp} = 0,381$ mM before and after annealing. Blue line corresponds to the sample before annealing, red line corresponds to the sample after the first annealing, green line corresponds to the sample after the second annealing. Spectral noise levels are different due to different NS values, processing parameters are identical. Spectra were scaled to maximum intensity.	28
4.5	Tel21 ^1H NMR spectra for $c_{exp} = 0,381$ mM before and after annealing restricted to [10:12,5] ppm interval. Blue line corresponds to the sample before annealing, red line corresponds to the sample after annealing. Spectral noise levels are different due to different NS values, processing parameters are identical. Spectra were intensity calibrated in respect to DSS peak at 0 ppm.	29
4.6	Tel22 ^1H NMR spectra for $c_{exp} = 0,095$ mM before and after annealing. Blue line corresponds to the sample before annealing, red line corresponds to the sample after the annealing. Spectra were scaled to maximum intensity, possess the same acquisition and data processing parameters.	30
4.7	Tel21 ^1H NMR spectra for $c_{exp} = 0,095$ mM before and after annealing restricted to [10:12,5] ppm interval. Blue line corresponds to the sample before annealing, red line corresponds to the sample after annealing. Spectral noise levels are different due to different NS values, processing parameters are identical. Spectra were intensity calibrated in respect to DSS peak at 0 ppm.	31
4.8	An example of experimental determination of Tel21 oligonucleotide concentration $c_{exp} = 0,482$ mM with the help of ^{31}P spectrum. The left peak corresponds to buffer signal, the right peak corresponds to an oligonucleotide signal(containing 20 phosphorus nuclei). . .	33
4.9	A comparison of Tel21 ^1H NMR spectra in respect to c_{exp} . Blue line corresponds to $c_{exp} = 0,482$ mM, red line corresponds to $c_{exp} = 0,397$ mM, green line corresponds to $c_{exp} = 0,072$ mM. Spectral noise levels are different due to different NS values, processing parameters are identical. Spectra were scaled to maximum intensity.	34
4.10	A comparison of Tel21 ^1H NMR spectra in respect to c_{exp} restricted to [10:12,5] ppm interval. Blue line corresponds to $c_{exp} = 0,482$ mM, red line corresponds to $c_{exp} = 0,397$ mM, green line corresponds to $c_{exp} = 0,072$ mM. Spectral noise levels are different due to different NS values, processing parameters are identical. Spectra were intensity calibrated in respect to DSS peak at 0 ppm.	35
4.11	Tel21 ^1H NMR spectra for $c_{exp} = 0,397$ mM before and after annealing. Blue line corresponds to the sample before annealing, red line corresponds to the sample after annealing. Spectral noise levels are different due to different NS values, processing parameters are identical. Spectra were scaled to maximum intensity.	36

4.12	Tel21 ^1H NMR spectra for $c_{exp} = 0,397$ mM before and after annealing restricted to [10:12,5] ppm interval. Blue line corresponds to the sample before annealing, red line corresponds to the sample after annealing. Spectral noise levels are different due to different NS values, processing parameters are identical. Spectra were intensity calibrated in respect to DSS peak at 0 ppm.	37
4.13	Tel21 ^1H NMR spectra for $c_{exp} = 0,072$ mM before and after annealing. Blue line corresponds to the sample before annealing, red line corresponds to the sample after annealing. Spectral noise levels are different due to different NS values, processing parameters are identical. Spectra were scaled to maximum intensity.	38
4.14	Tel21 ^1H NMR spectra for $c_{exp} = 0,072$ mM before and after annealing restricted to [10:12,5] ppm interval. Blue line corresponds to the sample before annealing, red line corresponds to the sample after annealing. Spectral noise levels are different due to different NS values, processing parameters are identical. Spectra were intensity calibrated in respect to DSS peak at 0 ppm.	39
4.15	Tel22 Raman spectra in respect to c_{exp} . Purple line corresponds to $c_{exp} = 0,571$ mM, green line corresponds to $c_{exp} = 0,381$ mM, Blue line corresponds to $c_{exp} = 0,095$ mM. Spectra were processed in Gnuplot.	41
4.16	Tel22 Raman spectra in respect to c_{exp} restricted to [1300:1500] cm^{-1} . Purple line corresponds to $c_{exp} = 0,571$ mM, green line corresponds to $c_{exp} = 0,381$ mM, Blue line corresponds to $c_{exp} = 0,095$ mM. Spectra were processed in Gnuplot.	41
4.17	Tel22 Raman spectra for $c_{exp} = 0,381$ mM before and after annealing. Purple line corresponds to the sample before annealing, green line corresponds to the sample after annealing. Spectra were processed in Gnuplot.	42
4.18	Tel22 Raman spectra for $c_{exp} = 0,381$ mM before and after annealing restricted to [1300:1500] cm^{-1} . Purple line corresponds to the sample before annealing, green line corresponds to the sample after annealing. Spectra were processed in Gnuplot.	42
4.19	Tel22 Raman spectra for $c_{exp} = 0,095$ mM before and after annealing. Purple line corresponds to the sample before annealing, green line corresponds to the sample after annealing. Spectra were processed in Gnuplot.	43
4.20	Tel22 Raman spectra for $c_{exp} = 0,095$ mM before and after annealing restricted to [1300:1500] cm^{-1} . Purple line corresponds to the sample before annealing, green line corresponds to the sample after annealing. Spectra were processed in Gnuplot.	43
5.1	Imino regions of 1D ^1H NMR spectrum of Tel22 in a K^+ solution at $K = 298,15$ K (adapted from [6])	44

List of Tables

3.1	Concentration table for Tel22 NMR experiment	20
3.2	Concentration table for Tel21 NMR experiment	20
3.3	Acquisition parameters for ^1H , ^{31}P Tel22 NMR experiments, where c_{calc} is expected concentrations of the concentrated samples, TD is size of FID (number of data points), NS is number of scans, AQ is acquisition time, D1 is relaxation delay. The sample number correspond to those in Table 3.1. As it was mentioned in section 3.1, annealing for the second sample was executed twice, one after another.	21
3.4	Acquisition parameters for ^1H , ^{31}P Tel21 NMR experiments, where c_{calc} is expected concentrations of the concentrated samples, TD is size of FID (number of data points), NS is number of scans, AQ is acquisition time, D1 is relaxation delay. The sample number correspond to those in Table 3.2.	22

List of Abbreviations

DNA deoxyribonucleic acid

RNA deoxyribonucleic acid

A adenine

C cytosine

G guanine

T thymine

NMR nuclear magnetic resonance

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

ppm parts per million