## Rational Optimization of Solvent Conditions for G-Quadruplex Oligonucleotides

- CD spectroscopy as a tool to characterize G-Quadruplex sequences as potential drug targets
- Orthogonal approach used to drive optimization of Mass Spectrometry experiments
- Discrimination of different G-Quadruplex topologies and assessment of buffer-dependent stabilities

As the regulatory role of G-quadruplex structures in crucial biological processes emerges, CD spectroscopy can be expected to gain increased importance in their characterization as drug targets for anti-viral and anticancer therapies. In this study, CD spectroscopy was used as an orthogonal technique to optimize experimental conditions for Mass Spectrometry experiments, which can be used for high-throughput screening in drug development with regards to structural analysis and binding interactions.

*Data courtesy of Sara N. Richter, Ph.D., Microbiology and Clinical Microbiology, Department of Molecular Medicine, University of Padua [1].*

### Introduction

Nucleic acid sequences rich in guanine can fold into G-quadruplexes (G4s), intramolecular conformations that are characterized by stacked tetrads of these nucleotide bases. As G4-forming sequences participate in the genomic regulation in humans and viruses, they are promising targets for anti-viral drugs and new cancer therapies.

For example, ligands binding to G4s in virus pathogens might enable new therapeutic approaches by blocking the viral DNA replication.

In humans, G4-stabilizing molecules might act on the G4-forming sequences of telomeres and thus inhibit proliferation of tumour cells while leaving normal cells unaffected. Clinical applications appear likely as such activities have already been demonstrated for potential drug candidates in both in vitro and in vivo tumour models. Moreover, the capability of G4-ligand interactions to down-regulate

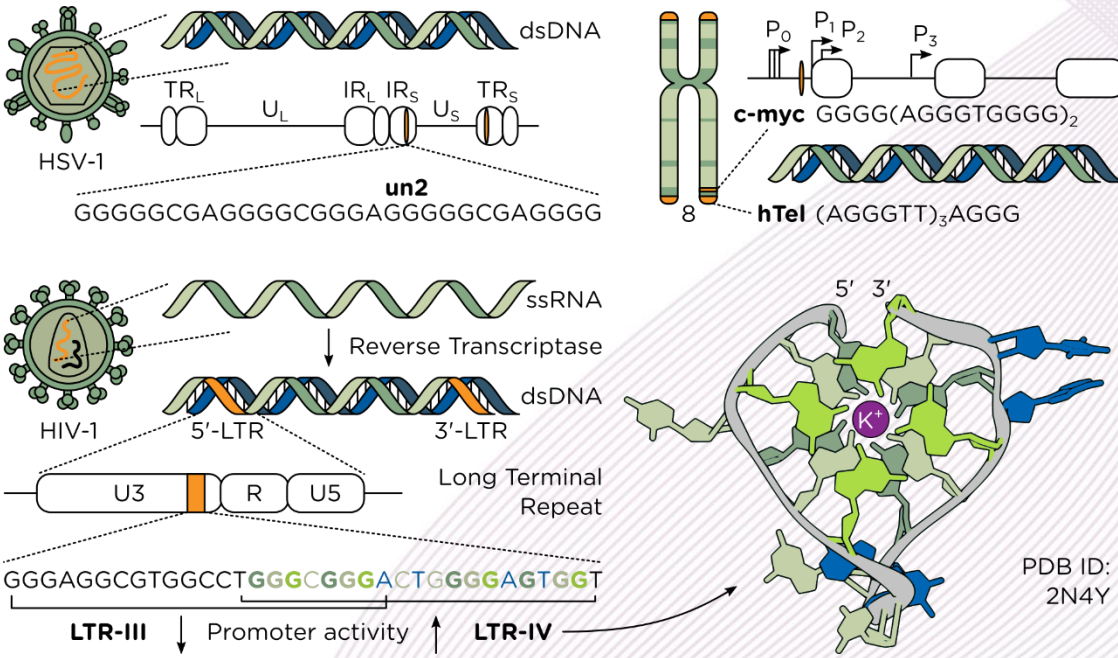



**Figure 1:** Locations and sequences of the G-quadruplexes used in this study.

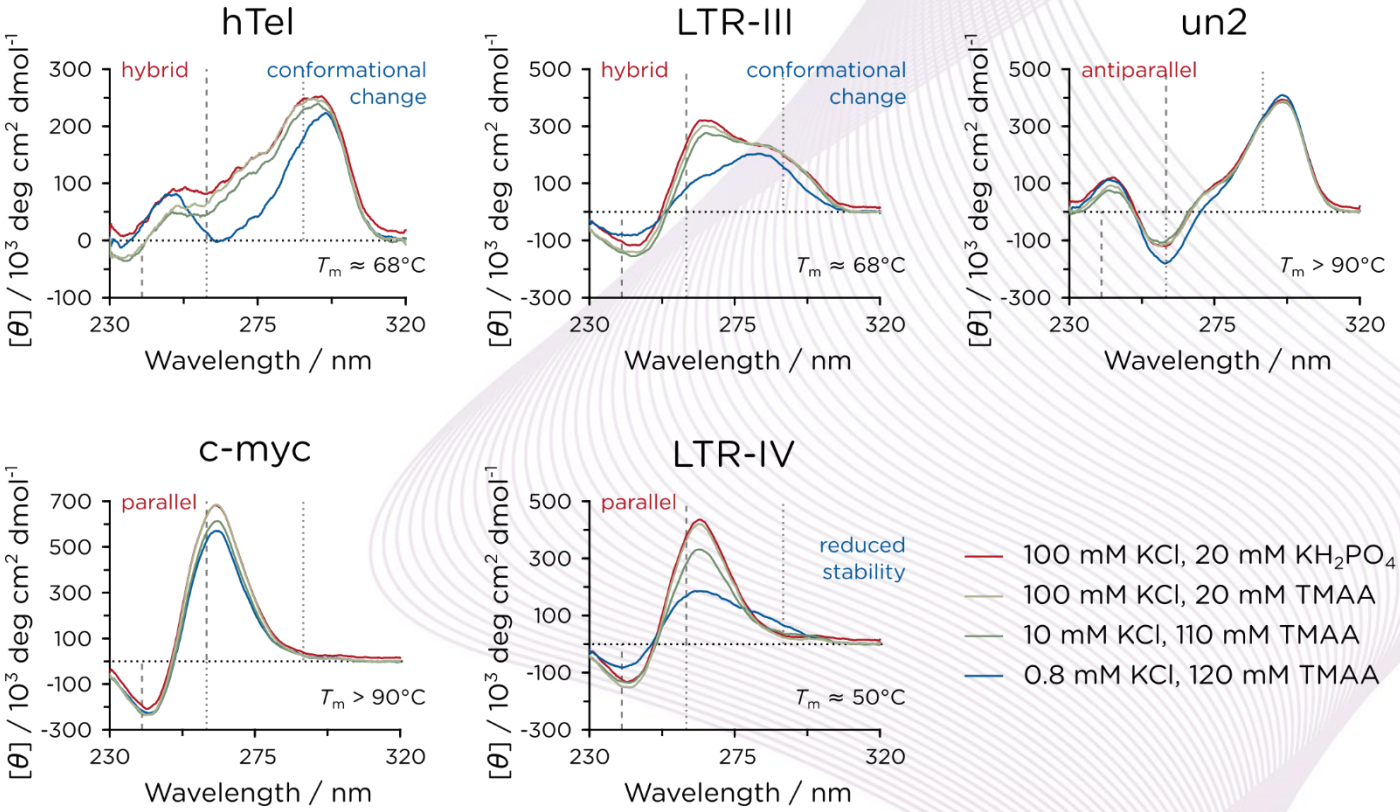



**Figure 2:** Comparison of CD spectra of five G4 forming oligonucleotides at physiological potassium concentration (red) and in TMAA buffer with lowered potassium concentrations.

oncogenes by acting on their corresponding promoters has been shown [2].

As the importance of regulatory functions exerted by G4s unfolds, efforts are increasing to develop new G4 therapeutics with better recognition and improved binding.

## Case Study

This study involved five G4s: Un2 [3], LTR-III [4] and LTR-IV [5], c-myc [6] and hTel [7] (Figure 1).

Un2 is located in the double-stranded DNA genome of herpes simplex virus 1 (HSV-1) which is known for causing latent, recurring infections. The HSV-1 genome consists of a short and a long region with unique sequences, US and UL, flanked by terminal and inverted repeats, TRL/IRL and IRS/TRS. While other putative G4s are known to exist in almost all of these regions, un2 is located in the inverted repeat flanking the short unique sequence.

LTR-III and LTR-IV are located in the genome of the human immunodeficiency virus 1 (HIV-1) which is well known for causing the acquired immune deficiency syndrome (AIDS). After transcription of the viral single-stranded RNA genome into double-stranded DNA, long terminal repeats (LTR) enable its insertion into the genome of the human host. The U3 region of the LTR promoter contains three overlapping sequences that can form mutually exclusive G4s, LTR-II, -III and -IV, and formation of LTR-III and LTR-IV is known to silence and enhance promoter activity, respectively.

Finally, c-myc and hTel form in the human chromosome, with c-myc located in the promoter of the c-MYC oncogene on chromosome 8 and hTel present in chromosome telomeres.

## Results

Native G4 conformations. As a reference for subsequent measurements, the native conformations of the G4s in a phosphate buffer at physiological $K^+$ concentration (**Figure 2**, red spectra) were characterized.

In this buffer (100 mM KCl, 20 mM $KH_2PO_4$), CD spectra of LTR-IV and c-myc were characteristic for a parallel G4 topology with a negative peak at ~240 nm and a positive peak at ~260 nm, whereas those of hTel and LTR-III indicated a hybrid topology with positive signals at both 260 nm and 290 nm. The CD spectrum of un2 showed a profile typical for antiparallel G4 topology with a negative peak at ~260 nm and a positive peak at ~290 nm.

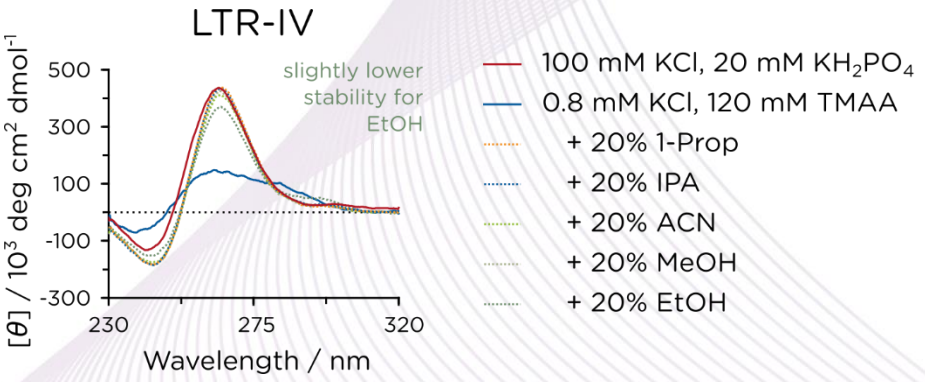



**Figure 3:** Comparison of G4-folding in a range of different organic solvents for LTR-IV (left) and known NMR structure of LTR-IV together with its base squence (right).

G4 stability in TMAA at reduced K+. The physiological $K^+$ concentration required for G4s to fold into their native conformation (100 to 150 mM) strongly affects G4 detection and data quality in electrospray ionization mass spectrometry (ESI-MS).

One strategy to circumvent this problem is to lower the $K^+$ concentration to a level compatible with ESI-MS (<1 mM) while providing physiological ionic strength with a volatile, bulky buffer such as trimethylammonium acetate (TMAA). Such buffers do not fit into the G4 cavity and, thus, do not impede coordination of $K^+$ by the G4 at low $K^+$ concentration.

The ability of TMAA to allow the stabilization of G4s even at low $K^+$ concentrations is illustrated in **Figure 2**. Replacing the buffer with TMAA while maintaining the total ionic strength resulted in CD spectra virtually identical (gray spectra) to the reference spectra (red spectra) even at a $K^+$ concentration lowered by an order of magnitude. Substantial structural changes occurred only at a $K^+$ concentration below 1 mM for the G4s of hTel and LTR-III (blue spectra). LTR-IV retained its parallel conformation but showed a decrease in molar ellipticity, indicating a lowered stability at strongly reduced $K^+$

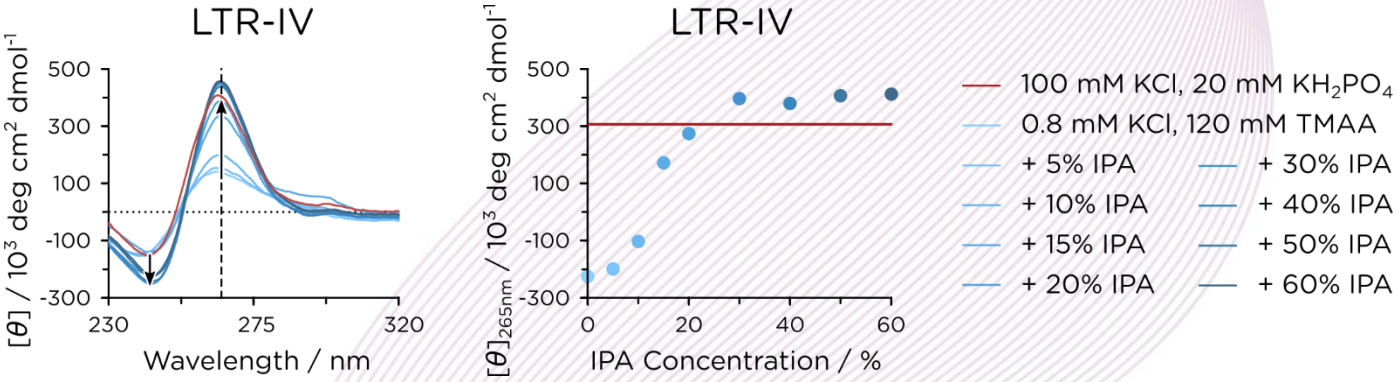



**Figure 4:** Concentration dependence of G4 stabilization by IPA for LTR-IV.

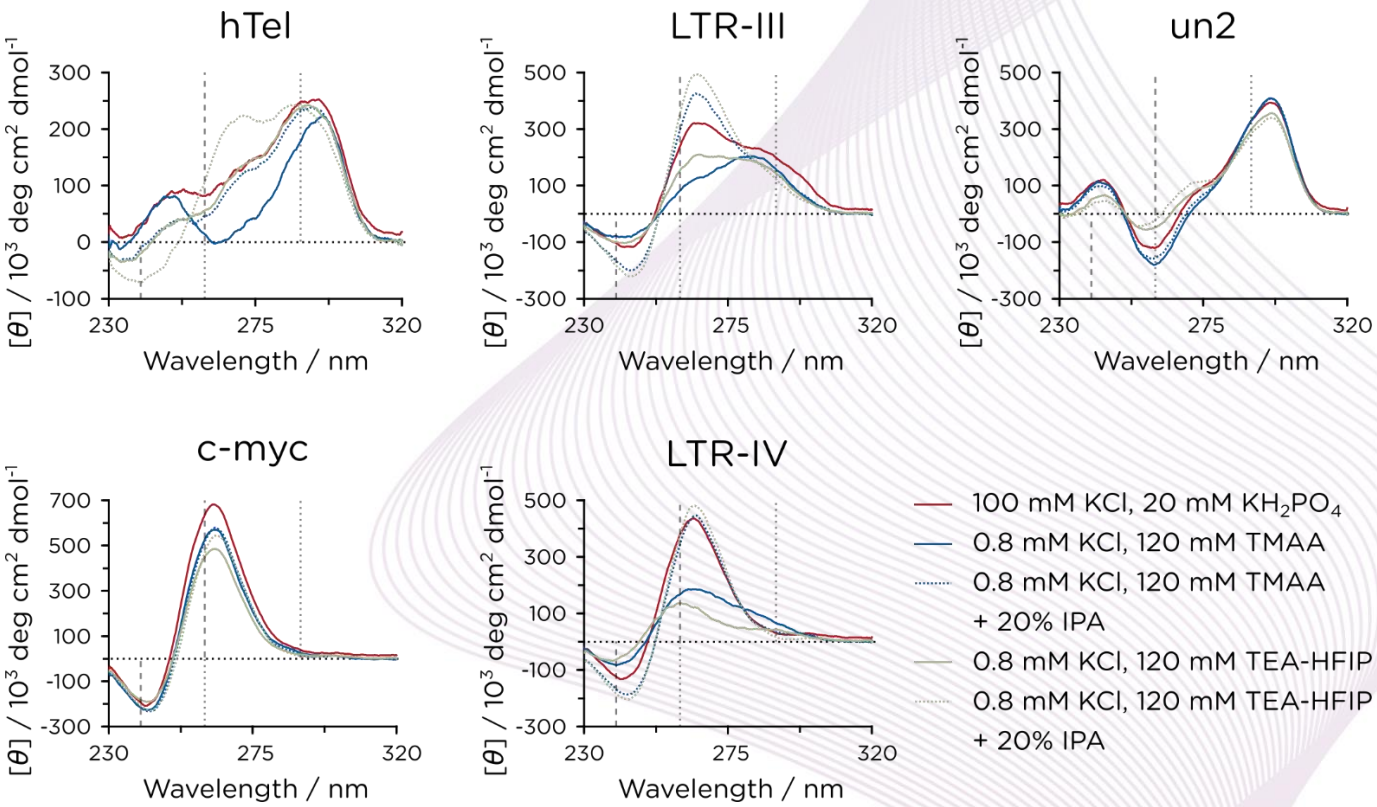



**Figure 3:** Comparison of CD spectra of five G4 forming oligonucleotides at physiological potassium concentration (red) and in different buffers optimized for Mass Spectrometry experiments.

concentration. As indicated by its rather low melting temperature of ~50°C, the G4 of LTR-IV is known to be less stable than the other G4s. In contrast, the native conformation was maintained even at 0.8 mM KCl for the G4s of c-myc and LTR-IV, which are the most stable G4s examined herein as their melting temperatures are known to be higher than 90°C.

These results illustrate that the approach using TMAA to improve resolution and detection of G4s in ESI-MS is well suited for stable G4s such as c-myc and un2 but less optimal for G4s of lower stability. Therefore, new solvent conditions were searched for in the following.

Improvement of G4-folding by organic solvents. It is known that primary alcohols promote folding of G4s due to molecular crowding and release of water upon G4 assembly. Therefore, the effect of several organic solvents on G4 stability and conformation of the low-stability G4 of LTR-IV was investigated (**Figure 3**). Basically, adding 20% of any of the tested organic solvents to 120 mM TMAA, 0.8 mM $K^+$ promoted the physiological G4 conformation as corresponding CD spectra were virtually identical to that in 100 mM KCl, 20 mM $KH_2PO_4$, with ethanol being slightly less effective.

Concentration dependence of G4 stabilization by IPA. Although all tested organic solvents performed equally well in stabilizing the G4 of LTR-IV, IPA was chosen for subsequent optimization because it is less toxic than methanol or acetonitrile. To find out the minimum concentration of IPA required to stabilize the native G4 conformation of LTR-IV, its CD at a range of IPA concentrations was recorded (**Figure 4**). The magnitude of the CD spectrum increased with IPA concentration and reached a plateau just beyond ~20% IPA.

Therefore, this concentration was used in the following process for buffer optimization.

Effect on stability by TMAA vs TEA-HFIP. Buffers other than TMAA are known to also improve detection of nucleic acids in ESI-MS. Therefore, a comparison of the ability to stabilize G4s between TMAA and triethylamine/ hexafluoroisopropanol (TEA-HFIP) was conducted, both in the absence and presence of 20% IPA (**Figure 3**).

For most of the G4s investigated, the two different buffers performed similar in stabilizing the native conformation. Only for hTel a marked difference manifested, with the CD signal increasing at 265 nm and decreasing at 250 nm in TEA-HFIP, 20% IPA. As the shoulder in the CD spectrum at 250 nm is known to correlate with a basket-type G4 with only two G-tetrad layers ('Form 3', [8]), this conformation appeared to disappear at these buffer conditions.

## Conclusions

By using the Chirascan V100 system for a rational optimization approach, optimal buffer conditions were found for the stabilization of several G4s, allowing for an increased signal intensity in the analysis of G4s by ESI-MS without affecting physiological folding.

The established conditions are expected to benefit high-throughput screening of G4 ligands in the development of potential drug candidates against viral infections for new anti-cancer therapies.

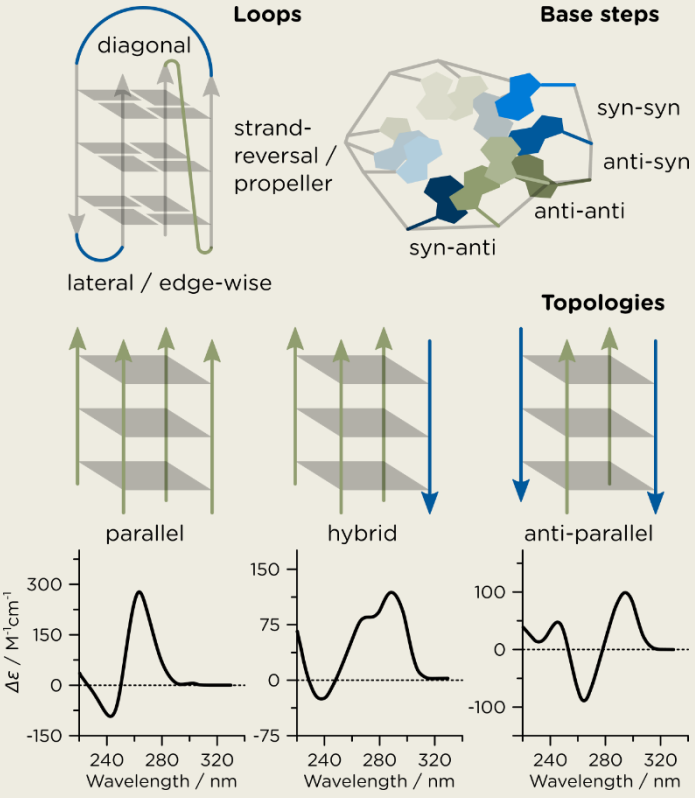
## G-Quadruplex Secondary Structure

CD spectroscopy is a powerful tool to characterize G4s [9-11]. The basic unit of G4s is a tetrad of co-planar guanine bases, with their carbonyl oxygen atoms directed towards the centre where monovalent cations are coordinated. These can be e.g. $Na^+$ ions in plane or between planes of two successive tetrads, or $K^+$ ions that are always located equidistant between tetrad planes.

The different folds that a stack of tetrads can adopt can be rationalized in terms of several properties (see figure). These include different types of the loops connecting the strands, and base steps, i.e. different orientation patterns of successive glycosidic bonds [12]. The different possible G4 topologies allow only for certain combinations of these properties, e.g. parallel G4s always have propeller loops and anti-anti base steps only.

Other parameters determining the overall G4 fold are glycosidic bond angles, quartet stacking geometry and groove dimensions [13]. Moreover, the number of stacked tetrads can vary and so can the molecularity of G4s if they are formed by intermolecular association of strands [14].

However, CD spectral profiles typical for the different G4 topologies (see figure, [15]) arise mainly from the stacking arrangements of the guanine base steps. In fact, the amounts of base steps can be estimated based on G4 CD spectra in a manner similar to secondary structure analysis for proteins [15].



## Experimental

CD spectra shown were obtained with a Chirascan V100 using a quartz glass cuvette with a pathlength of 5 mm. Acquisition settings: 1 nm bandwidth, 0.2 nm step size, 0.3 s time-per-point.

For further experimental details refer to the main reference [1].

## References

[1] M. Scalabrin, M. Palumbo and S. N. Richter, 2017, *Anal. Chem.*, 89:17, 8632–8637.

[2] E. Ruggiero and S. N. Richter, 2018, *Nucleic Acids Res.*, 46, 3270–3283.

[3] S. Artusi, M. Nadai, R. Perrone, M. A. Biasolo, G. Palù, L. Flamand, A. Calistri and S. N. Richter, 2015, *Antiviral Res.*, 118, 123–131.

[4] R. Perrone, M. Nadai, I. Frasson, J. A. Poe, E. Butovskaya, T. E. Smithgall, M. Palumbo, G. Palù and S. N. Richter, 2013, *J. Med. Chem.*, 56:16, 6521–6530.

[5] B. De Nicola, C. J. Lech, B. Heddi, S. Regmi, I. Frasson, R. Perrone, S. N. Richter and A. T. Phan, 2016, *Nucleic Acids Res.*, 44:13, pp. 6442–6451.

[6] A. Siddiqui-Jain, C. L. Grand, D. J. Bearss and L. H. Hurley, 2002, *Proc. Natl. Acad. Sci. USA*, 99:18, 11593–11598.

[7] Y. Wang and D. J. Patel, 1993, *Structure*, 1:4, 263–282.

[8] W. K. Lim, S. Amrane, S. Bouaziz, X. Weixin, M. Yuguang, D. J. Patel, K. N. Luu and A. T. Phan, 2009, *J. Am. Chem. Soc.*, 131:12, 4301–4309.

[9] M. Vorlíčková, I. Kejnovská, K. Bednářová, D. Renčiuk and J. Kypr, 2012, *Chirality*, 24:9, 691–698.

[10] M. Vorlíčková, I. Kejnovská, J. Sagi, D. Renčiuk, K. Bednářová, J. Motlová and J. Kypr, 2012, *Methods*, 57:1, 64–75.

[11] J. Kypr, I. Kejnovská, D. Renčiuk and M. Vorlíčková, 2009, *Nucleic Acids Res.*, 37:6, 1713–1725.

[12] X. Cang, J. Šponer and T. E. Cheatham, III, 2011., *Nucelic Acids Res.*, 39:10, 4499–4512

[13] A. I. Karsisiotis, C. O'Kane and M. W. da Silva, 2013, *Methods*, 64:1, 28–35.

[14] S. Burge, N. G. Parkinson, P. Hazel, A. K. Todd and S. Neidle, 2006, *Nucleic Acids Res.*, 34:19, 5402–5415.

[15] R. del Villar-Guerra, J. O. Trent and B. J. Chaires, 2018, *Angew. Chem. Int. Ed. Engl.*, 57:24, 7171–7175.