Analyst

Estimation of G-Quartet-Forming Guanines in Parallel-Type G-Quadruplexes by Optical Spectroscopy Measurements of Their Single-Nucleobase Substitution Sequences

Journal:	*Analyst*
Manuscript ID	AN-COM-06-2018-001131.R1
Article Type:	Communication
Date Submitted by the Author:	13-Jul-2018

Complete List of Authors:	Maruyama, Ryo; The University of Tokyo, Department of Life Sciences, Graduate School of Arts and Sciences Makino, Kurumi; University of Tokyo, Department of Life Sciences, Graduate School of Arts and Sciences Yoshitomi, Toru; The university of Tokyo, Department of Life Sciences, Graduate School of Arts and Sciences Yui, Hiroharu; Tokyo University of Science, Department of Chemistry, Faculty of Science Furusho, Hitoshi; Nissan Chemical Industries, Ltd., Yoshimoto, Keitaro; The University of Tokyo, Department of Life Sciences, Graduate School of Arts and Sciences; JST, PRESTO

Analyst

COMMUNICATION

Estimation of G-Quartet-Forming Guanines in Parallel-Type G-Quadruplexes by Optical Spectroscopy Measurements of Their Single-Nucleobase Substitution Sequences

Received 00th January 20xx,
Accepted 00th January 20xx

DOI: 10.1039/x0xx00000x

www.rsc.org/

Ryo Maruyama[a], Kurumi Makino[a], Toru Yoshitomi[a], Hiroharu Yui[b], Hitoshi Furusho[c], and Keitaro Yoshimoto[a, d*]

Since much attention has been paid to in vivo biological functions of G-quadruplexes, structural analyses of G-quadruplexes are essential for understanding their functional mechanisms. Here, we established a simple optical-spectroscopy-based method for the estimation of G-quartet-forming guanines in parallel-type G-quadruplexes using measurements of circular dichroism and the thermal melting temperature.

Guanine (G)-rich single-stranded DNA sequences can fold into four-stranded structures called G-quadruplexes, where planer G-quartets are stabilized by Hoogsteen hydrogen bonds between the four Gs and monovalent cations such as potassium ions.[1–3] Because there is growing evidence for the role of these structures in biological processes, such as gene regulation and telomere maintenance,[4–9] an understanding of their structure and function is very important for the development of new drugs. There are several experimental methods for estimating G-quadruplex formation, which provide structural information at different conformational levels.[10] Fluorescence spectroscopy using G-quadruplex-binding fluorescent dyes has been used to confirm G-quadruplex folding. Thioflavin T (ThT) is a fluorescent dye that selectively binds to quadruplex forming DNAs and RNAs with noncovalent bonds, accompanied by fluorescence.[11] This method has significant advantages: the measurements are low cost and do not require complex processing, although conformational information beyond the folding structure cannot be obtained. Analysis using circular dichroism (CD) spectra provides topological information about G-quadruplexes.[12,13] In general, G-quadruplexes are classified into three topological structures, parallel, anti-parallel, and hybrid.[14] CD spectroscopy provides this topological information regarding G-quadruplex structures and allows us to distinguish parallel structures from anti-parallel structures. Thus, analytical methods based on optical spectroscopy provide conformational information of G-quadruplexes with simple techniques and equipment.

To obtain further structural information about G-quadruplexes in addition to the folding and topology, X-ray crystallography and NMR spectroscopy have been used.[15–21] X-ray crystallography can be used to determine the higher order structure, although the growth of single crystals often requires a long period, and there are many cases where the crystallization is unsuccessful because of the large number of experimental conditions requiring optimization. NMR spectroscopy requires significantly less sample preparation time than crystallography. Although NMR spectroscopy is a powerful method for the estimation of the higher order structure, expertise is required for the use of special pulse sequences. In addition, trace nuclear Overhauser enhancement with correlation analyses are needed.

Thus, it would be desirable to develop an analytical method that reliably yields information concerning the higher order structure using simple techniques and easy operation. Analytical methods based on electrophoresis can estimate the G-quartet-forming Gs without expertise. For example, dimethyl sulfate (DMS) footprinting is used to study G-quartet-forming Gs in G-quadruplexes. DMS methylates the N7 position of G, which leads to facile depurination. Although the addition of piperidine then leads to cleavage at the abasic site, little or no cleavage at the Gs is observed when the N7 is protected from methylation by hydrogen bonding arising from G-quadruplex formation.[22] These cleaved fragments can be visualized using gel electrophoresis, resulting in the estimation of G-quartet-forming Gs. However, since generally oligonucleotide samples are prepared using radioisotope labeling in DMS footprinting, special equipment and facilities are required. No method provides the conformational information concerning the G-

[a] Department of Life Sciences, Graduate School of Arts and Sciences, The University of Tokyo, 3-8-1 Komaba, Meguro, Tokyo 153-8902, Japan.
[b] Department of Chemistry, Faculty of Science, Tokyo University of Science, 1-3 Kagurazaka, Shinjuku, Tokyo 162-8601, Japan.
[c] Chemical General Division, Nissan Chemical Industries, Ltd., 2-10-2 Tsuboi-nishi, Funabashi, Chiba 274-8507, Japan.
[d] JST, PRESTO, The University of Tokyo, Komaba 3-8-1, Meguro, Tokyo 153-8902, Japan.
Electronic Supplementary Information (ESI) available: [details of any supplementary information available should be included here]. See DOI: 10.1039/x0xx00000x

quartet-forming Gs in G-quadruplexes without complicated operation and radioisotope labeling.

In this paper, we propose optical spectroscopy-based analytical method that estimates G-quartet-forming Gs in parallel-type G-quadruplexes. The method consists of the measurement of the CD spectra and the thermal melting temperature (T_m) of their single-nucleobase substitution sequences. Without radioisotope labeling of nucleotide, the proposed method can estimate G-quartet-forming Gs using simple techniques and easy operation.

In the case of parallel-type G-quadruplexes, the information concerning G-quartet-forming Gs can be linked to the estimation of their higher order structure. Therefore, in the present study, we investigated the G-quartet-forming Gs in VEGF_Pu22T12T13 (VPT),[21] T95-2T,[22] and VEFG-Pu22 (VP),[21] which form three types of three-tetrad parallel-type G-quadruplexes. The higher order structures of VPT and T95-2T have been estimated in previous studies using NMR spectroscopy, and the G-quartet-forming Gs of VP have been determined using DMS footprinting.[23,24] The substitution of G-quartet-forming Gs with adenine (A) is expected to change the structural properties, including topologies and thermal stability, of the G-quadruplexes.[25] The CD spectra and thermal stabilities of their single-nucleobase substituted sequences were measured to assess the substitution effects on the G-quartet-forming Gs.

As shown in Fig. 1A, VPT is a base substitution sequence of VP,[21] where G12 and G13 of VP are replaced with thymine (T). VPT forms a three-tetrad parallel-type G-quadruplex consisting of four G-runs of three and four consecutive Gs, which are underlined and denoted the I–IV G-runs in Fig. 1A and colored blue in Fig. 1B. Fig. 1C shows the CD spectra of VPT and its single-nucleobase G-to-A substitution sequence (Table S1). The CD spectrum of VPT shows a positive Cotton effect at 260 nm and a negative cotton effect at 240 nm, supporting a parallel-type G-quadruplex, as reported previously.[26] Interestingly, the CD spectra of the G-to-A substitution sequences of the G-quartet-forming Gs from G3 to G5 in run I, G7 to G9 in run II, G14 to G16 in run III, and G18 to G20 in run IV showed remarkable changes compared to that of VPT. Especially, the CD spectrum of the G-to-A substitution sequences in G19 in run IV shows a positive Cotton effect at 290 nm and a negative cotton effect at 265 nm, supporting an anti-parallel-type G-quadruplex. [25]

These spectral changes indicate the changes in G-quadruplex topologies on the substitution of the G-quartet-forming Gs with A. On the other hand, the CD spectra of the G-to-A substitution sequences in G2 in run I and G21 in run IV, which are not G-quartet-forming Gs, showed almost identical spectra to that of VPT, indicating no changes to the G-quadruplex topology. These results indicate that the G-quartet-forming Gs of VPT, which have been previously determined by NMR,[21] can be estimated using CD spectra measurements of single-nucleobase A substitution sequences alone.

The investigation of the T95-2T sequence was carried out. T95-2T consists of the high-definition structure of a simple monomeric G-quadruplex, which could serve as a reference for propeller-type G-quadruplex structures in solution,[22] as shown

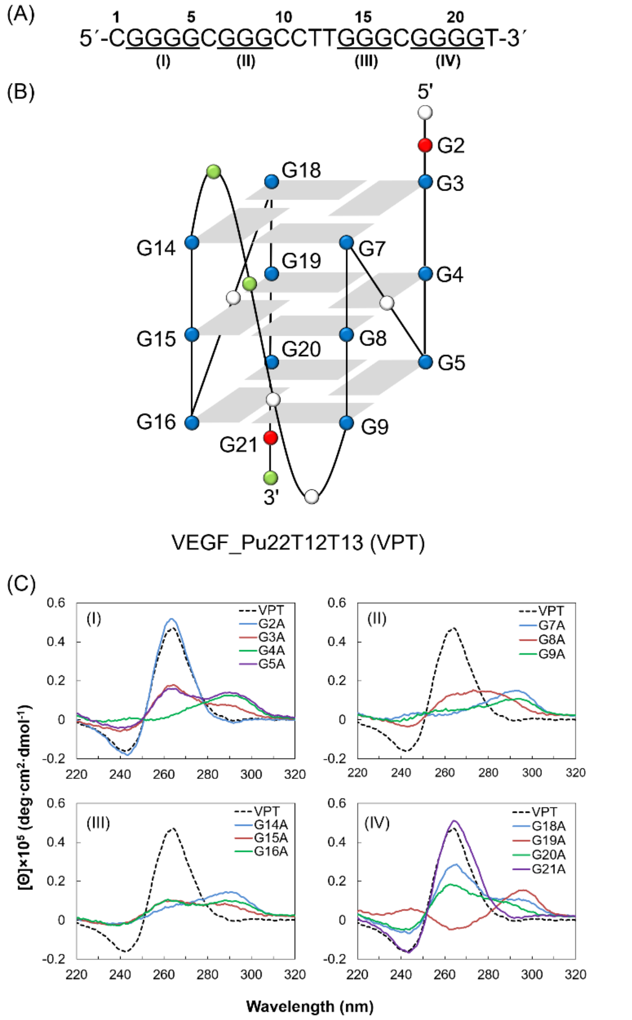


Fig. 1 (A) The sequence of VEGF_Pu22T12T13. VEGF_Pu22T12T13 is a 22-mer G-rich sequence. The four G-runs are underlined and numbered (I–IV). The number above the sequence indicates the base number. (B) Three-dimensional structure of the parallel-type VEGF_Pu22T12T13 G-quadruplex, which is drawn based on PDB ID 2M27 (colors: G-quartet-forming G, blue; non-G-quartet-forming G, red; cytosine, white; T, green). (C) CD spectra of the VEGF_Pu22T12T13T sequence and its single-nucleobase G-to-A substitution sequences. Four-µM ssDNA samples were prepared in 10 mM potassium phosphate buffer containing 40 mM KCl (pH 7.2).

in Fig. S1A and S1B, and was drawn based on PDB ID 2LK7. In this sequence, all the Gs in the T95-2T sequence are G-quartet-forming Gs (Fig. S1A, shown underlined). The CD spectra of T95-2T shown in Fig. S2 support the formation of a parallel-type G-quadruplex under the experimental conditions. Unlike the VPT sequence, almost no significant changes in the CD spectra of its single-nucleobase G-to-A substitution sequences were observed, indicating no topological change in the G-quadruplex after substitution (Fig. S2). Interestingly, the CD spectrum of the G-to-A substitution sequences in G11 in run III indicated an anti-parallel-type G-quadruplex topology. In addition to the CD spectra, the thermal melting curves show a melting transition at 295 nm, which provides evidence concerning the thermodynamic stability of the G-quadruplexes.[27] To obtain further structural information concerning T95-2T, the T_m of T95-2T and its single-nucleobase A substitution sequences were measured from their melting curves. Significant decreases in T_m were observed in all G-quartet-forming Gs (Fig. S3). In contrast, the single-nucleobase T-to-A substitution of T95-2T resulted in almost no change in T_m (Fig. S3). This result indicates that single-

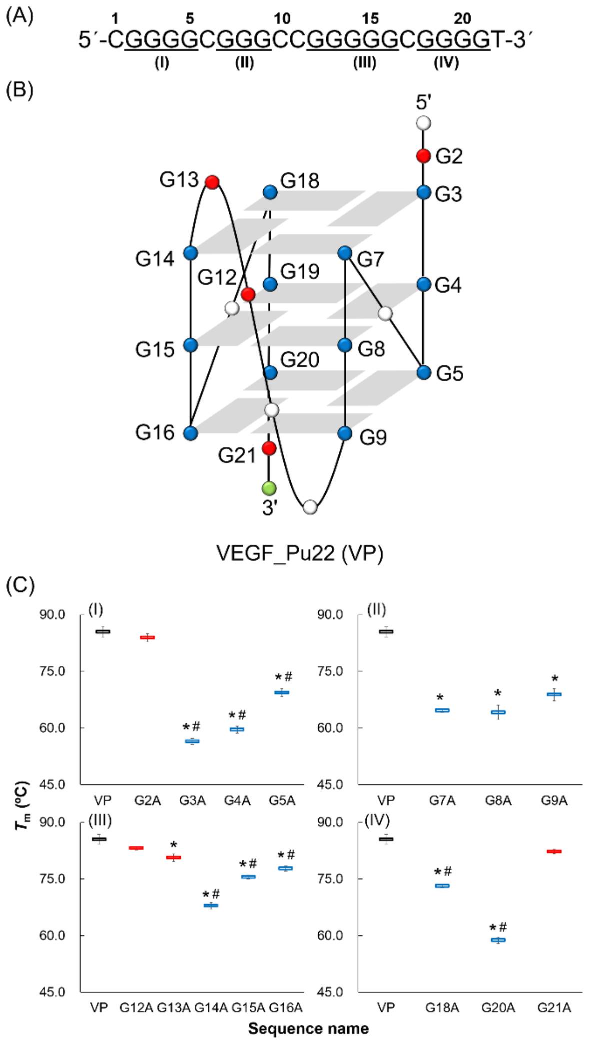


Fig. 2 (A) The VEGF_Pu22 sequence. VEGF_Pu22 is a 22-mer G-rich sequence. The four G-runs are underlined and numbered (I–IV). The numbers above the sequence indicate the base number. (B) Three-dimensional structure of parallel-type G-quadruplex in VEGF_Pu22 (color scheme: G-quartet forming G, blue; G-quartet no-forming G, red; cytosine, white; T, green). (C) T_m values of VEGF_Pu22 sequence and its single-nucleobase G-to-A substitution sequences were measured by UV melting curves at 295 nm. T_m values of VEGF_Pu22 without substitution are colored in black, where T_m is 85.5 ± 1.3 °C. The T_m values of the G-to-A substitution sequences in G-quartet-forming Gs and non-G-quartet-forming Gs are blue and red, respectively. Values are means ± standard error (SE); $n \geq 3$, *Significant difference from VP (Tukey-Kramer, P < 0.01), #Significant difference from non-G-quartet-forming Gs (Tukey-Kramer, P < 0.01). Four-μM ssDNA samples were prepared in 25 mM potassium phosphate buffer containing 70 mM KCl (pH 7.0).

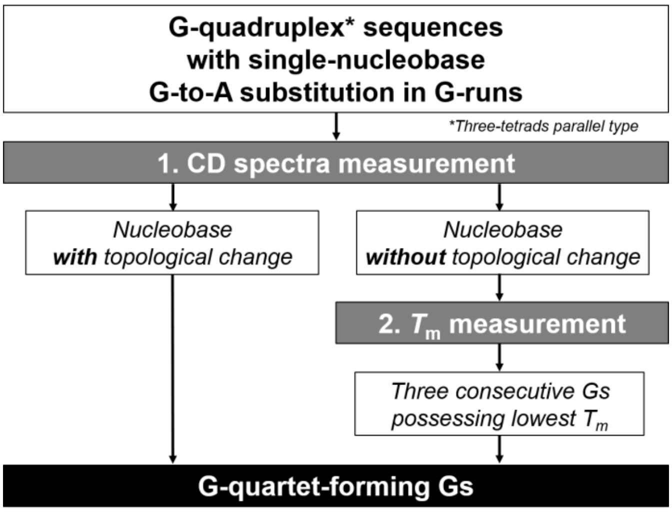
nucleobase substitution of G-quartet-forming Gs by A induced a significant destabilization of the G-quartet structure, resulting in a reduction in the thermodynamic stability of the G-quadruplex. These results suggest that G-quartet-forming Gs can be estimated using the changes in T_m, even when there are no topological changes in the G-quadruplex on single-nucleobase G-to-A substitution.

As shown in Fig. 2A, the VP sequence is contained in the promoter region of VEGF and has been reported to form three-tetrad parallel-type G-quadruplexes,[21] whose G-quartet-forming Gs have been estimated by DMS footprinting previously.[21] The three-dimensional structure of the G-quadruplex is shown schematically in Fig. 2B.

The VP sequence has four G-runs of three, four, and five consecutive Gs, which are underlined and denoted I–IV in Fig. 2A. As shown in Fig. S4, the CD spectrum of VP showed typical parallel-type Cotton effect with negative and positive peaks at approximately 240 and 260 nm, respectively. Almost no topological changes were observed in the single-nucleobase G-to-A substitution sequences, except for G19 in run IV (Fig. S4). In the previous study, G19 was estimated to be a G-quartet-forming G by footprinting analysis, which is also supported by the CD spectral changes measured here. T_m measurements were carried out to investigate other G-quartet-forming Gs. As shown in Fig. 2C, the single-nucleobase G-to-A substitution sequence of the G-quartet-forming Gs estimated by footprinting analysis has a lower T_m than that of the non-G-quartet-forming Gs. Compared to the T_m values of VP, those of single-nucleobase G-to-A substitution sequences, except for G19, were decreased significantly for the G-quartet-forming Gs: from G3 to G5 in run I, G7 to G9 in run II, G14 to G16 in run III, and G18 and G20 in run IV. Among them, the degree of decrease in the T_m of the sequence substituted at G-run III is lower than those substituted at G-runs I, II, and IV. In G-run III, because there are five Gs and they can form three-tetrad G-quartets in several patterns, the degree of reduction in the thermal stability of substitution sequences at G-run III is lower than those of the substitution sequences in the G-runs in other regions. Therefore, the results obtained from CD and T_m measurements strongly support the estimation of G-quartet-forming Gs by footprinting analysis.

We succeeded in acquiring correlation relationships for G-quartet-forming Gs from CD spectra and T_m measurements of single-nucleobase G-to-A substitution. The data collection used in the present study and the analysis process for the estimation the G-quartet-forming Gs of three-tetrad parallel-type G-quadruplexes are shown in Scheme 1.

First, as shown in step 1 in Scheme 1, CD measurements of single-nucleobase G-to-A substitution sequences of G-quadruplex-forming sequences are carried out to investigate the topologies of substitution sequences. Then, destabilization of G-quadruplex by single-nucleobase G-to-A substitution induces topological changes. The G-quartet-forming Gs are estimated from these topological changes. In step 2, T_m measurement of the single-nucleobase G-to-A substitution sequences of the G-quadruplex-forming sequence is carried out to evaluate the thermal stability. Finally, because single-nucleobase G-to-A substitution of the G-quartet-forming Gs reduces the thermal stability, the G-quartet forming Gs are estimated by selection of the lowest three T_m values.



Scheme 1 Scheme showing the steps for the estimation of G-quartet-forming Gs in G-quadruplexes.

Conclusions

In conclusion, three parallel-type G-quadruplexes with different lengths of G-runs were analyzed, and the changes in their CD spectra and T_m of their single-nucleobase substitutions were investigated in detail, resulting in a promising new estimation method for G-quartet-forming Gs. The substitution of G-quartet-forming Gs with A induced the destabilization of the higher order structure of the G-quadruplex and changed the CD spectrum and thermal melting temperature. As mentioned above, changes in the CD spectra of VPT-G19A and T95-2T-G11A are interesting points, indicating that the conformational change between parallel-type quadruplexes and anti-parallel-type quadruplexes can be regulated by single nucleobase substitution. Our proposed method for the estimation of G-quartet-forming Gs does not require specialized equipment or analysis, such as NMR and X-ray crystallography, and easily identified, using optical spectroscopy, the G-quartet-forming Gs without complicated analysis or radioisotope labeling. We believe that the proposed method is promising for the structural analyses of G-quadruplex at the laboratory level.

Acknowledgments

This work was partially supported by the JST-PRESTO [Grant JPMJPR16FB] and Nissan Chemical Industries, Ltd.

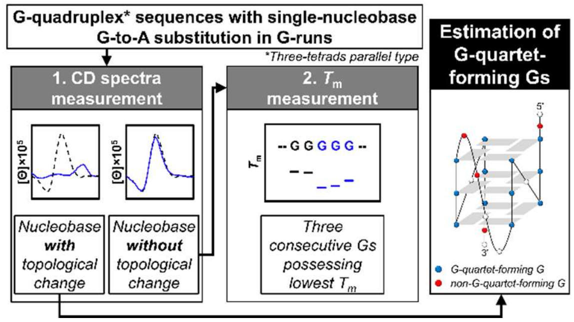
Conflicts of interest

There are no conflicts to declare.

Notes and references

1 D. Sen and W. Gilbert, *Nature*, 1988, **334**, 364–366.
2 J. T. Davis, *Angew. Chemie Int. Ed.*, 2004, **43**, 668–698.
3 X. Yang, D. Liu, P. Lu, Y. Zhang and C. Yu, *Analyst*, 2010, **135**, 2074–2078.
4 Y. Xu, *Chem. Soc. Rev.*, 2011, **40**, 2719–2740.
5 H. J. Lipps and D. Rhodes, *Trends Cell Biol.*, 2009, 19, 414–422.
6 R. K. Moyzis, J. M. Buckingham, L. S. Cram, M. Dani, L. L. Deaven, M. D. Jones, J. Meyne, R. L. Ratliff and J. R. Wu, *Proc. Natl. Acad. Sci. U. S. A.*, 1988, **85**, 6622–6.
7 K. Guo, A. Pourpak, K. Beetz-Rogers, V. Gokhale, D. Sun and L. H. Hurley, *J. Am. Chem. Soc.*, 2007, **129**, 10220–10228.
8 A. T. Phan, V. Kuryavyi, S. Burge, S. Neidle and D. J. Patel, *J. Am. Chem. Soc.*, 2007, **129**, 4386–4392.
9 S. Cogoi and L. E. Xodo, *Nucleic Acids Res.*, 2006, **34**, 2536–2549.
10 J. L. Huppert, *Chem. Soc. Rev.*, 2008, **37**, 1375–1384.
11 J. Mohanty, N. Barooah, V. Dhamodharan, S. Harikrishna, P. I. Pradeepkumar and A. C. Bhasikuttan, *J. Am. Chem. Soc.*, 2013, **135**, 367–376.
12 C. C. Hardin, T. Watson, M. Corregan and C. Bailey, *Biochemistry*, 1992, **31**, 833–41.
13 S. Paramasivan, I. Rujan and P. H. Bolton, *Methods*, 2007, **43**, 324–331.
14 J. L. Huppert, *Chem. Soc. Rev.*, 2008, **37**, 1375–1384.
15 C. Kang, X. Zhang, R. Ratliff, R. Moyzis and A. Rich, *Nature*, 1992, **356**, 126–131.
16 G. Laughlan, a I. Murchie, D. G. Norman, M. H. Moore, P. C. Moody, D. M. Lilley and B. Luisi, *Science*, 1994, **265**, 520–4.
17 S. L. Forman, J. C. Fettinger, S. Pieraccini, G. Gottarelli and J. T. Davis, *J. Am. Chem. Soc.*, 2000, **122**, 4060–4067.
18 J. T. Davis, *Angew. Chemie Int. Ed.*, 2004, **43**, 668–698.
19 M. Marušič, P. Šket, L. Bauer, V. Viglasky and J. Plavec, *Nucleic Acids Res.*, 2012, **40**, 6946–6956.
20 D. Wei, A. K. Todd, M. Zloh, M. Gunaratnam, G. N. Parkinson and S. Neidle, *J. Am. Chem. Soc.*, 2013, **135**, 19319–19329.
21 M. Adrian, D. J. Ang, C. J. Lech, B. Heddi, A. Nicolas and A. T. Phan, *J. Am. Chem. Soc.*, 2014, **136**, 6297–6305.
22 A. Siddiqui-Jain, C. L. Grand, D. J. Bearss and L. H. Hurley, *Proc. Natl. Acad. Sci.*, 2002, **99**, 11593–11598.
23 P. Agrawal, E. Hatzakis, K. Guo, M. Carver and D. Yang, *Nucleic Acids Res.*, 2013, **41**, 10584–10592.
24 N. Q. Do and A. T. Phan, *Chem. - A Eur. J.*, 2012, **18**, 14752–14759.
25 M. Tomaško, M. Vorlíčková and J. Sagi, *Biochimie*, 2009, **91**, 171–179.
26 R. Giraldo, M. Suzuki, L. Chapman and D. Rhodes, *Proc. Natl. Acad. Sci. U. S. A.*, 1994, **91**, 7658–7662.
27 J. L. Mergny, A. T. Phan and L. Lacroix, *FEBS Lett.*, 1998, **435**, 74–78.

Table of contents



Here, we established new estimation method for G-quartet-forming guanines in parallel-type G-quadruplexes using measurements of circular dichroism and thermal stability.