Raman Study on Polymorphism of Human Telomeric Quadruplexes

Jan Palacký[1], Michaela Vorlíčková[2,3*], Iva Kejnovská[2,3], Peter Mojzeš[1*]

[1]Charles University in Prague, Faculty of Mathematics and Physics, Institute of Physics, Ke Karlovu 5, CZ-121 16 Prague 2, Czech Republic

[2]Institute of Biophysics, Academy of Sciences of the Czech Republic, Královopolská 135, CZ-612 65 Brno, Czech Republic

[3]CEITEC – Central European Institute of Technology, Masaryk University, Kamenice 5, CZ-625 00 Brno, Czech Republic

[*]*Corresponding authors:* mojzes@karlov.mff.cuni.cz, mifi@ibp.cz

Supplementary data

Detailed interpretation of Raman markers diagnostic of quadruplex structures

Structural interpretation of the conventional non-resonant Raman spectra of $G_3(TTAG_3)_3$ quadruplexes is based on the ultimate Raman studies of interquadruplex switching between antiparallel and parallel forms of *Oxytricha* telomeric sequence $(T_4G_4)_4$ controlled by the type, concentration, and molar ratio of Na^+ and K^+ cations (34, 35). Raman markers established there, refined in subsequent Raman studies (36-39) and catalogued (32, 39) are listed in Supplementary Tab. S1. Since human telomeric sequences incorporate also dA residues, proper assignment of the $G_3(TTAG_3)_3$ spectra has required analysis of dA contribution missing in the previous Raman quadruplex studies. Raman spectra of homooligonucleotides $(dG)_{15}$, $(dT)_{15}$ and $(dA)_{15}$ in Na^+- and K^+-PBS have been acquired (Supplementary Fig. S1) and used for spectral interpretation (Supplementary Fig. S2).

Quadruplex Raman markers can be assorted according to structural information that they provide. There are **(i)** markers signaling formation of G-tetrads and constitution of the quadruplex, **(ii)** markers sensitive to conformation of dG residues distinguishing between *C2'-endo/anti* and *C2'-endo/syn* conformers, **(iii)** the bands sensitive to the arrangement of dT-loops, and **(iv)** markers of overall geometry and regularity of the sugar-phosphate backbone.

(i) Markers sensitive to formation of G-tetrads

Raman signatures of G-tetrad formation concern structurally informative dG bands located at ~1483±2, 1581±1 and 1719±3 cm^{-1}. The bands are sensitive to the strength of hydrogen bonding at dG-N7, dG-N2H and dG-O6, respectively (see G-tetrad scheme in Fig. 3), distinguishing thus stronger hydrogen bonding of dG within G-tetrads from weaker bonding to surrounding H_2O.

Major contribution to the intensity of the most intense Raman band of $G_3(TTAG_3)_3$ at ~1483±2 cm^{-1} comes from vibrations of dG residues, although there are some minor contributions from dA and dT bands at ~1486 and 1483 cm^{-1}, respectively (Supplementary Fig. S2). The band is sensitive to Hoogsteen hydrogen bonding between dG-N7 serving as an acceptor and the donor dG-N2H group (32). In single-strands and B-DNA duplexes, the band typically occurs at ~1486 – 1490 cm^{-1} due to weak hydrogen bonding with H_2O, and downshifts for ~ 6 – 9 cm^{-1} upon formation of Hoogsteen hydrogen bonds within G-tetrads (34, 39). According to the present study on $G_3(TTAG_3)_3$, exact position of ~1483±2 cm^{-1} band is sensitive to interquadruplex transition induced by K^+ and DNA concentration, as well as to the type of stabilizing cation. The band located at 1484 cm^{-1} in 100 mM K^+ slightly downshifts (~1 cm^{-1}) and decreases in intensity (~10% hypochromism) as the concentration of K^+ increases to 500 mM (Fig. 5). Similar changes can be observed on the antiparallel-to-parallel interquadruplex transition induced by high DNA concentration (Fig. 1 and Supplementary Fig. S3). On the other hand, the same relative downshift and decrease in intensity at ~1483 cm^{-1} can bee seen in Raman spectrum of $(dG)_{15}$ in 200 mM K^+ in comparison to 200 mM Na^+ solution (Supplementary Fig. S1), albeit $(dG)_{15}$ evidently forms an intermolecular parallel four-stranded quadruplex (36) with both cations. Additional downshift of the ~1483±2 cm^{-1} band may thus indicate minor changes in

the microenvironment of the five-membered imidazole ring of dG due to dG-syn/dG-anti glycosyl reorientation, as well as cation-dependent differences in hydration of the grooves or in the stacking of G-tetrads. However, as demonstrated by Raman melting profile of K^+-stabilized $G_3(TTAG_3)_3$ quadruplex (Fig. 9 and Supplementary Fig. S6), thermal dependency of the position of 1484 cm^{-1} band exhibits rather complicated behavior impeding simple and straightforward interpretation of the downshift.

Quadruplex band at ~1581±1 cm^{-1} is sensitive to hydrogen bonding at exocyclic dG-N2H donor site. The band exhibits downshift to ~1574 cm^{-1} on the quadruplex melting (Fig. 9 and Supplementary Fig. S5). For $G_3(TTAG_3)_3$ adopting a single-strand non-quadruplex structure (e.g. in the absence of stabilizing cations), the band is located at 1578 cm^{-1} (Fig. 7). According to our study using solvent- and background-corrected, normalized Raman spectra with precisely calibrated wavenumber scales, the position and intensity of the ~1581±1 cm^{-1} is virtually insensitive to the type of stabilizing cation, as well as to interquadruplex transition, since no significant spectral changes are observed on the K^+/Na^+ replacement (Fig. 4), and on the antiparallel-to-parallel transition induced by increased DNA or K^+ concentration (Figs. 1 and 5). Relative insensitivity of the ~1581±1 cm^{-1} band to interquadruplex switching contrasts with high sensitivity of the nearby located ~1483±2 cm^{-1} band.

Information on the formation of G-tetrads and on the strength of hydrogen bonds can be drawn also from a weak but clearly resolvable shoulder at ~1719±3 cm^{-1} that is sensitive to participation of the dG-O6 in the interbase hydrogen bonding with dG-N1H (39). In non-quadruplex structures the band is located at ~1688±3 cm^{-1} (Fig. 7). Intensity of the band seems to be insensitive to the interquadruplex transition (Fig. 9), and can thus serve as an indicator of general quadruplex structure.

(ii) Markers sensitive to *C2'-endo/anti* and *C2'-endo/syn* conformation of dG

Raman bands at 686±2 and 1338±2 cm^{-1} were ascribed to *C2'-endo/anti* conformers of dG (32, 34-36, 39). On the other hand, shoulder at 671±2 and the band at 1326±2 cm^{-1} indicate *C2'-endo/syn*-dG. Structural transition of antiparallel-to-parallel $G_3(TTAG_3)_3$ comprises increase in the population of *C2'-endo/anti-dG* at the expense of *C2'-endo/syn-dG*, as can be illustrated by relative intensity increase of the bands at 686 and 1338 cm^{-1}, accompanied by intensity decrease of the shoulder at 671 cm^{-1} and the 1326 cm^{-1} band (Fig. 1). Although there is no doubt that *syn/anti*-dG Raman markers qualitatively indicate antiparallel-to-parallel transition of K^+-stabilized $G_3(TTAG_3)_3$, identification of the initial and final structures and accurate quantification of the number of dG residues concerned in the transition simply from Raman spectra are neither straightforward nor simple. As *C2'-endo/syn* and *C2'-endo/anti* populations of dG residues in antiparallel $G_3(TTAG_3)_3$ are equal whereas parallel structure consists exclusively of *C2'-endo/anti*-dG conformers, one could expect twofold intensity increase at 686 and 1338 cm^{-1} and disappearance of the bands at 671 and 1326 cm^{-1} upon *complete* transition from fully antiparallel to fully parallel folding. On the other hand, antiparallel and (3+1) quadruplex structures, differing in a single dG residue switched from *C2'-endo/syn* to *C2'-endo/anti*, are expected to be distinguishable due to 6:6 and 5:7 intensity ratio of the corresponding Raman markers, respectively. Unfortunately, because of the complicated overlap with neighboring Raman bands (Supplementary Fig. S1 and S2) and resulting ambiguity in the absolute non-

Raman background, reliable decomposition of the *syn*-dG/*anti*-dG markers into separate bands and quantification of their *absolute* intensities is affected by a substantial uncertainty. Although *relative* intensity changes can be evaluated rather accurately from differential Raman spectra, the baseline for quantification of the number of dG residues involved in a transition is rather unsure. Consequently, Raman-based conclusions on the K^+-structures of $G_3(TTAG_3)$, especially differentiation between antiparallel and (3+1) K^+-forms, have to be supported by a careful and critical comparison with Raman spectra of the Na^+-stabilized, undoubtedly antiparallel $G_3(TTAG_3)_3$, taking into account particularities of vibrational spectroscopy (i.e. possible non-uniformity of Raman cross-sections of the respective *syn*-dG/*anti*-dG vibrational modes).

(iii) Markers sensitive to arrangement of dT-loops

The band at 611±2 cm^{-1} was initially assigned as a marker of *C2'-endo/syn*-dG (34), since its intensity strongly decays on the antiparallel-to-parallel quadruplex transition of $(T_4G_4)_4$ (35). Later, on the basis of protium-deuterium exchange (36), the band was reassigned to dT conformer preferred within lateral dT_4-loops of the antiparallel foldback of *Oxytricha* $(T_4G_4)_4$ (36, 39). No suitable precursor of the band can be found in Raman spectra of $(dA)_{15}$, $(dG)_{15}$ and $(dT)_{15}$ (Supplementary Figs. S1 and S2). Furthermore, the 611±2 cm^{-1} band was reported not to appear in non-telomeric sequences $d(TG)_8$ and $dT_6(TG)_8$ containing dT_n tracts, but organized in other structures than dT-loops (38, 39). The band is missing also in other dT-containing structures such as A- and B-DNA duplexes (36, 39). On the other hand, it appears at 615 cm^{-1} in the Raman spectra of intramolecular triple helices consisting of purine and pyrimidine tracts interconnected by dT_4 loops (46). It is thus speculated (36, 39) that the 611±2 cm^{-1} band is a marker of an unusual *C2'-endo*/*syn*-dT conformer, occurrence of which was suggested to be restricted exclusively to lateral (edgewise) or diagonal loops of foldback quadruplexes or in other dT-loop-containing DNA structures. Disappearance of the 611±2 cm^{-1} band indicates rearrangement of the loops to double-chain-reversal orientation typical for intramolecullar parallel quadruplex. If the band assignment (32, 36, 39) is valid, then, according to our results on $G_3(TTAG_3)_3$, it could serve as a more universal Raman marker distinguishing between side/diagonal and double-chain-reversal loop arrangement also for dT-loops incorporating dA residue.

Since the 611±2 cm^{-1} band is located in the low-wavenumber region containing few relatively separate bands with simple non-Raman baseline and negligible overlaps, it is suitable for quantitative analysis. The low-wavenumber region below 650 cm^{-1} was only partly exploited in the previous Raman studies (32, 34-36, 39). In addition to the 611±2 cm^{-1} marker, the region contains four relatively separate bands at 499±2, 534±1, 564±1 and 581±1 cm^{-1}. Two of them (499±2 and 581±1 cm^{-1}) exhibit intensity variations on the interquadruplex transition induced by increase of K^+ and DNA concentration (Figs. 1, 3, 4 and 5). Although the spectral changes of the relatively weak low-wavenumber bands may seem to be tiny, they are highly reproducible and constitute characteristic pattern easily recognizable in differential Raman spectra whenever the well-established Raman markers 671/686 cm^{-1} and 1326/1338 cm^{-1} indicate *syn/anti* transition of dG residues.

For the Na^+-stabilized antiparallel $G_3(TTAG_3)_3$, the bands at 581±1 and 611±2 cm^{-1} are of comparable heights. Similar intensity ratio can be seen in the Raman spectra K^+-stabilized $G_3(TTAG_3)_3$ under the

conditions favoring antiparallel quadruplex structure (low-K^+, low-DNA). However, detailed comparison of the putative Na^+- and K^+-antiparallel folds reveals slight differences, e.g. downshift (1 cm^{-1}), narrowing and intensity decrease (~15%) of the 611±2 cm^{-1} band in the K^+-form. On the interquadruplex transition induced by high K^+ and DNA concentration, intensity of the 611±2 cm^{-1} band is reduced, whereas intensity of the 581 cm^{-1} band increases.

The 499 ± 2 cm^{-1} band was formerly assigned to the scissoring mode of the phosphodioxy (PO_2^-) group (47). Its shape and intensity can be satisfactorily explained as a simple linear combination of the band precursors present in the spectra of the corresponding homooligonucleotides (Supplementary Fig. S2) with the main contribution coming from dT residues. The band is slightly more intense at lower K^+ concentrations or in the presence of Na^+ (regardless of the Na^+ concentration), probably reflecting higher flexibility of the sugar-phosphate backbone of d(TTA)-loops in the antiparallel $G_3(TTAG_3)_3$ backfold. Unlike the 499 cm^{-1} band, the principal contribution to the band at 581 cm^{-1} comes from dG residues (Supplementary Figs. S1 and S2). Its intensity increases on the interquadruplex transition induced by high concentration of K^+ or DNA, and follows increase of 686 and 1338 cm^{-1} markers of *C2'-endo/anti*-dG. Albeit weak, the change evidently indicates quadruplex switching to parallel structure. Slight but apparent intensity decrease of the 499 cm^{-1} band accompanied by comparable intensity increase of the 581 cm^{-1} band can be suggested as a new signature of the antiparallel-to-parallel quadruplex transition.

The band at ~1370 cm^{-1} is dominated by the most prominent Raman band of dT (at 1376 cm^{-1} in single-strand $(dT)_{15}$, Supplementary Fig. S1), however with some contributions from dG and dA providing their own precursors at 1363 and 1380 cm^{-1}, respectively (Supplementary Fig. S1). Intensity of the dT band near ~1375 - 1380 cm^{-1} in B-DNA duplexes was previously reported to increase with increasing hydrophobicity of the environment of dT methyl groups (51). Consequently, an intensity decrease at 1370 cm^{-1} observed in $G_3(TTAG_3)_3$ on the increase of K^+ may be interpreted as easier accessibility of the solvent molecules to dT methyl groups in double-chain-reversal d(TTA)-loops of (3+1) or parallel quadruplexes than in the lateral/diagonal loops of the antiparallel fold. Although real molecular mechanism explaining intensity change of the dT band may be different, we suggest that it reflects rearrangement of d(TTA)-loops and can serve as Raman marker of the loop arrangement.

(iv) Markers sensitive to overall geometry and regularity of the sugar-phosphate backbone

Raman spectra of Na^+- and K^+-stabilized $G_3(TTAG_3)_3$ quadruplexes display DNA phosphate markers at 837±1 and 1093±1 cm^{-1} (Fig. 1). Positions of both markers indicate that phospodiester (O-P-O) and phosphodioxy (PO_2^-) groups have the local geometries similar to canonical B-DNA (32, 34). Whereas position and shape of the 837±1 cm^{-1} band seem to be insensitive to interquadruplex transition, band broadening, intensity decrease and finally disappearance is observed on the quadruplex melting at increased temperatures (Fig. 9). On the other hand, intensity of the 1093±1 cm^{-1} band associated with PO_2^- symmetric stretching mode remains largely unchanged on the quadruplex melting, and can serve as intensity standard for Raman melting studies (45).

Further Raman marker characteristic for regular B-DNA-family backbone (32, 34), but diagnostic for antiparallel-to-parallel transition of $G_3(TTAG_3)_3$, is a broad band with a maximum at 788±2 cm^{-1} (Figs. 1, 4 and 5). The band consists of spectral contributions from dG, dT as well as dA residues (51) (Supplementary Figs. S1 and S2). Previously it was attributed to symmetrical stretching mode of the phosphodiester (O-P-O) group (48-50). The band gradually downshifts from ~788 to 786 cm^{-1} (with an intensity increase at the short-wavenumber side) on the antiparallel-to-parallel transition induced by high concentration of DNA or K^+ (Fig. 1, 4 and 5). It is likely that the sharp differential feature located at ~781 cm^{-1} in all differential spectra between parallel and antiparallel $G_3(TTAG_3)_3$ forms is primarily due to rearrangement of the quadruplex sugar-phosphate backbone involving mainly dG (and to a lesser extent dT) residues since the same feature can be found in differential Raman spectra describing antiparallel-to-parallel transition of *Oxytricha* quadruplex $(T_4G_4)_4$ (35) containing no dA residues.

Reliability of spectral differences revealed in Raman spectra

Although spectral changes of several Raman bands of moderate and weak intensities may seem to be too small to be identified reliably, their statistical significance is assured by the mode of data acquisition and treatment used in the present work. Instead of visual inspection and identification of the differences in few intentionally selected Raman spectra obtained under particular physico-chemical conditions, multivariate SVD methods (44) are applied to extensive Raman datasets to take into consideration also inherent spectral variabilities caused by uncontrollable variations of experimental conditions (day-to-day fluctuations, different signal-to-noise ratio, uncertainties in the background- and solvent-correction, different batches used for sample preparation). Representative Raman spectra represent average spectra extracted from several measurements realized on different days under the identical experimental conditions, often by using oligonucleotides from different syntheses. By means of advanced multivariate SVD methods (44), inherent variabilities within datasets consisting of several Raman spectra belonging to the same particular structure or state, were evaluated for their statistical significance. Spectral differences between Raman spectra of two different structures were considered as significant merely in the case when corresponding spectral variability between two datasets was significantly greater (at least 5x) than the spectral variability within a dataset describing identical structure. Instead of comparing two intentionally selected spectra taken under different physico-chemical conditions, spectral differences were recognized by the multivariate SVD methods in the spectral series (consisting of dozens of spectra) obtained with gradually varied intensive variable (concentration, ionic strength, temperature, time). Continuous dependence of the spectral difference as a function of an intensive variable was regarded as a further sign of reliability and evidence of real existence of structural differences behind. For a presentation in a more common mode, Raman spectra corresponding to the edges of the interval of intensive variable under the study have been extracted and their difference used to better highlight spectral differences.

Table S1. Raman markers diagnostic of quadruplex structures

Marker position [cm^{-1}]	Assignment	Diagnostic of telomeric fold or structure
499 ± 2	sc PO_2^-, dT, dG	stronger in ap-Q
581 ± 1	dG	stronger in p-Q
611 ± 2	dT_4 and d(TTA) loops	lateral or diagonal loops in ap-Q or hairpin
671 ± 2	C2'-*endo/syn*-dG	stronger shoulder in ap-Q
686 ± 2	C2'-*endo/anti*-dG	stronger in p-Q
781 ± 1	O-P-O, dG, dT, dA	stronger in p-Q
837 ± 1	O-P-O	marker of regular B-type bk
1093 ± 1	PO_2^-	intensity standard
1322	C2'-*endo/syn*-dG	non-Q
1326 ± 2	C2'-*endo/syn*-dG	stronger in ap-Q
1338 ± 2	C2'-*endo/anti*-dG	stronger in p-Q
1371 ± 1	dT	stronger in ap-Q
1483 ± 2	dG-N7 strong Hoogsteen H-bond	ap-Q or p-Q; downshift and hypochromism in p-Q
~1486	dG-N7 weak H-bond to H_2O	non-Q
~1578	dG-N2H H-bond to H_2O	non-Q
1581 ± 1	dG-N2H interbase H-bond	ap-Q or p-Q
1688 ± 2	dG-O6 H-bond to H_2O	non-Q
1719 ± 3	dG-O6 interbase H-bond	ap-Q or p-Q

Data compiled from ref. (32, 34-36, 39, 46-51) and this work.

Abbreviations: ap-Q, antiparallel quadruplex; p-Q, parallel quadruplex; non-Q, non-quadruplex; bk, backbone; sc, scissoring vibration

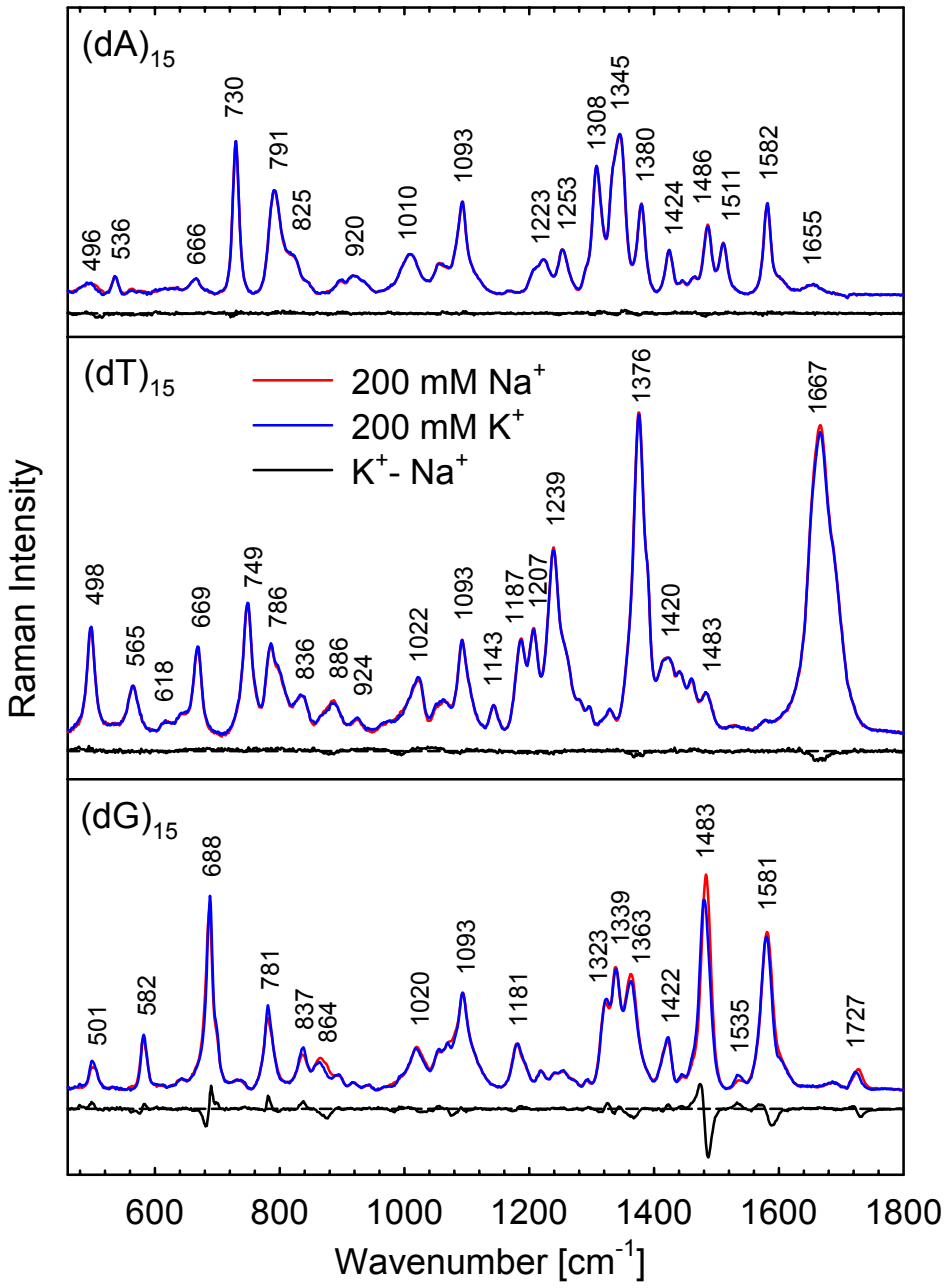

Figure S1. Normalized solvent- and background-corrected Raman spectra of $(dG)_{15}$, $(dT)_{15}$ and $(dA)_{15}$ in 30 mM PBS (pH 6.8) containing 200 mM of Na^+ or K^+ in total. Oligonucleotide concentration was 1.3 mM (in the strand, 20 mM in nucleoside) and the spectra were taken at 5°C. Difference spectra between K^+ and Na^+-solutions are shown to highlight cation-specific spectral variance. Under the conditions used, $(dA)_{15}$ and $(dT)_{15}$ adopt single-stranded B-DNA helices of similar structures in both alkali solutions, whereas $(dG)_{15}$ evidently forms four-stranded intermolecular parallel quadruplex exhibiting structural differences between Na^+- and K^+-stabilized forms.

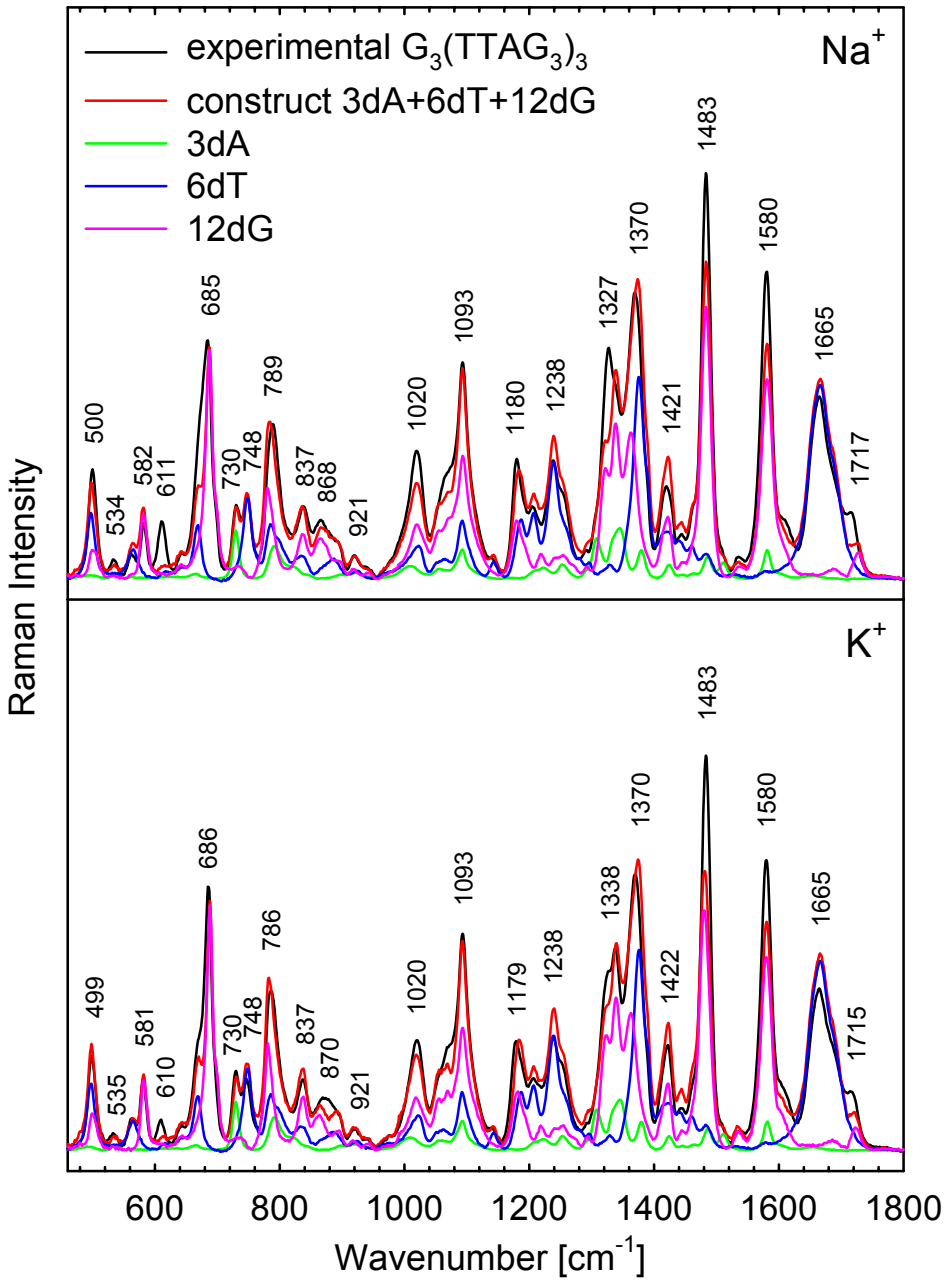

Figure S2. Comparison of experimental Raman spectra of $G_3(TTAG_3)_3$ in 500 mM Na^+ (top) or K^+ (bottom) and the synthetic spectra constructed as linear combinations of the Raman spectra of $(dG)_{15}$, $(dT)_{15}$ and $(dA)_{15}$ (in 200 mM Na^+ or K^+; Fig. S1). Proportional contributions from constituent nucleotides are shown to depict origin of individual Raman bands. Positions of the bands relate to the experimental spectrum.

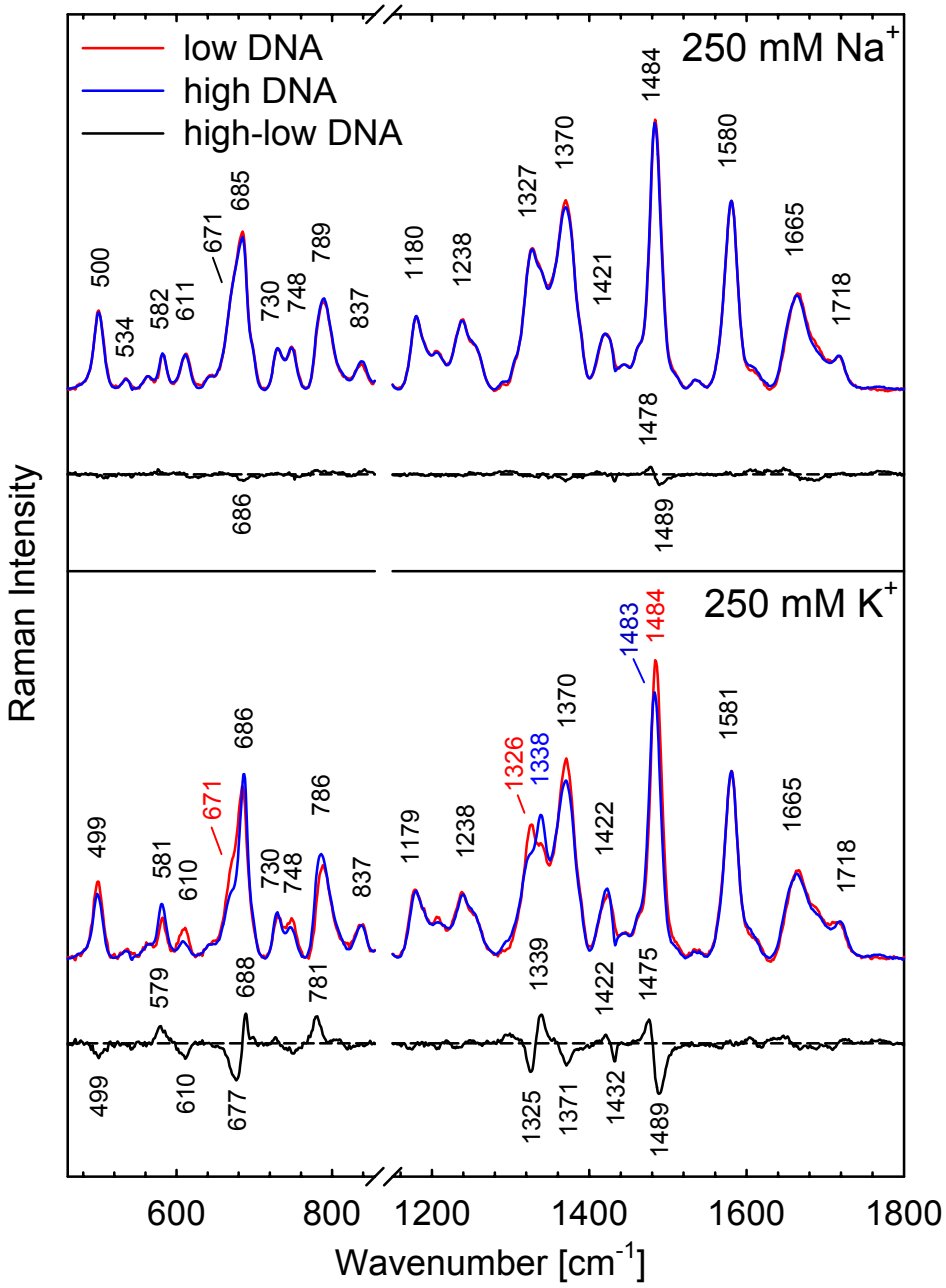

Figure S3. Cation-specific effect of high DNA concentration on the quadruplex structure. Normalized Raman spectra of $G_3(TTAG_3)_3$ at low (~8 mM) and high (~240 mM) DNA concentration in the presence of 250 mM of Na^+ (top) and K^+ (bottom) are compared in detail.

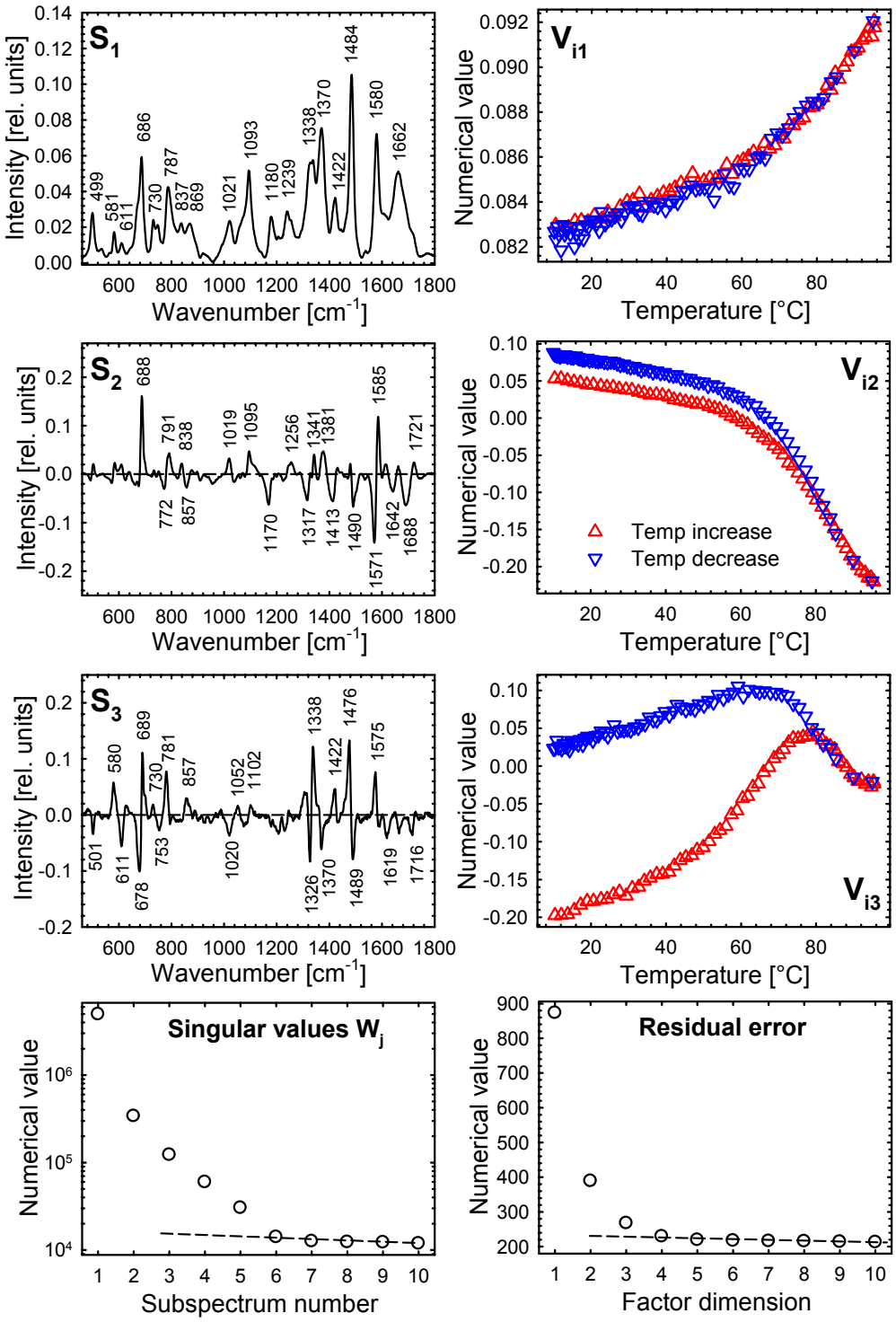

Figure S4. Factor analysis of the spectral changes observed in the course of annealing. K^+-stabilized (230 mM K^+) antiparallel quadruplex $G_3(TTAG_3)_3$ (330 mM in nucleosides) was heated continuously from 10 to 95°C and then cooled down. Rate of temperature increase/decrease was ca 1.5°C/min. Only the most important orthogonal spectral components describing differences related preferentially to overall denaturation of the quadruplex structures (S_2, V_{i2}) and to interquadruplex transition (S_3, V_{i3}) are shown for clarity.

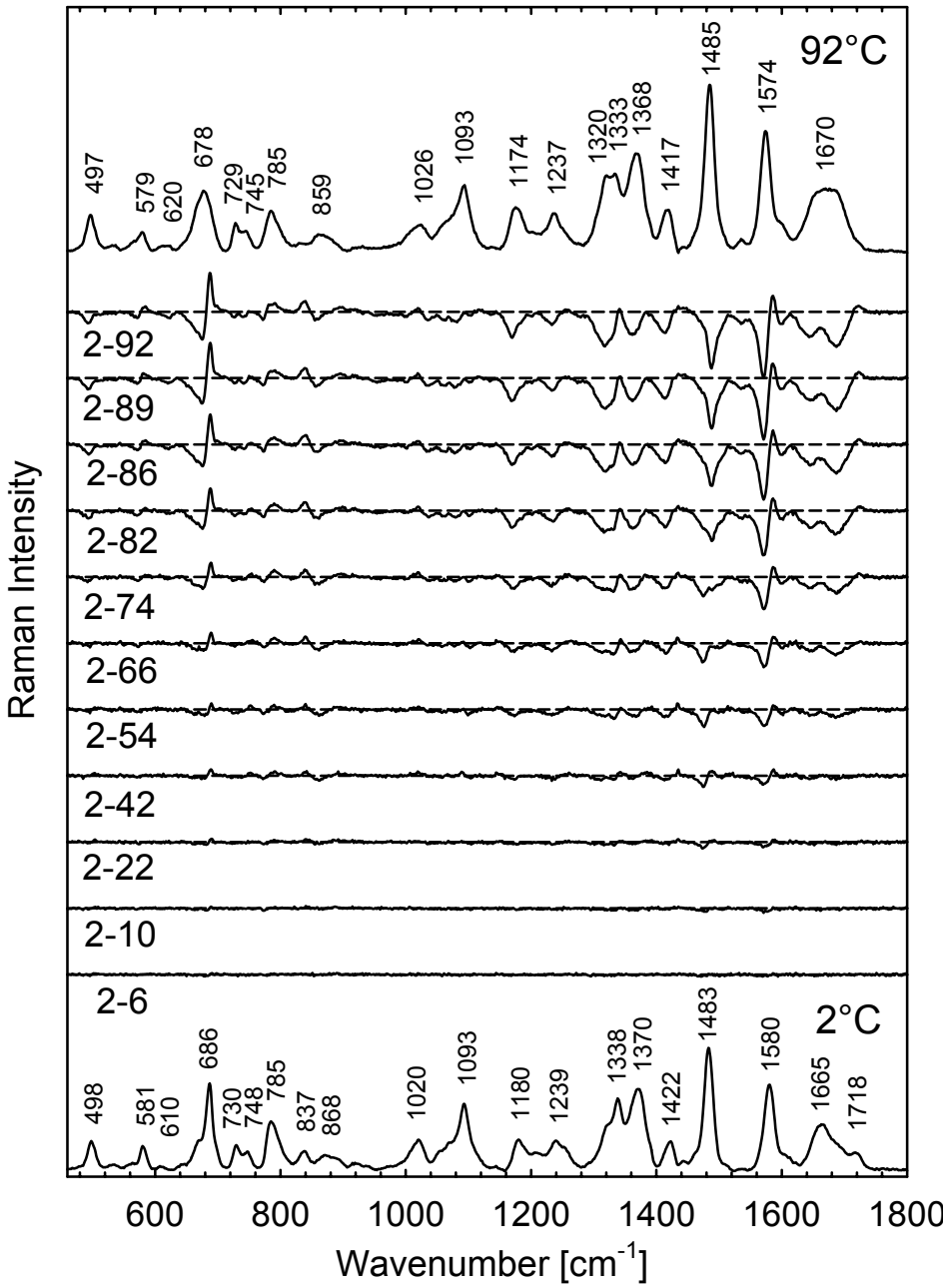

Figure S5. Spectral changes induced by thermal melting of preformed (by slow annealing) parallel K^+-stabilized quadruplex $G_3(TTAG_3)_3$. DNA concentration of 250 mM (in nucleoside), in 250 mM K^+ (30 mM PBS, pH 6.8). Intermediate traces show the differences between the spectrum of fully ordered quadruplex (2°C) and selected spectra taken at the temperatures indicated. Factor analysis of the full dataset comprising 46 temperature-equilibrated Raman spectra taken on the temperature increase and decrease is shown in Fig. S6.

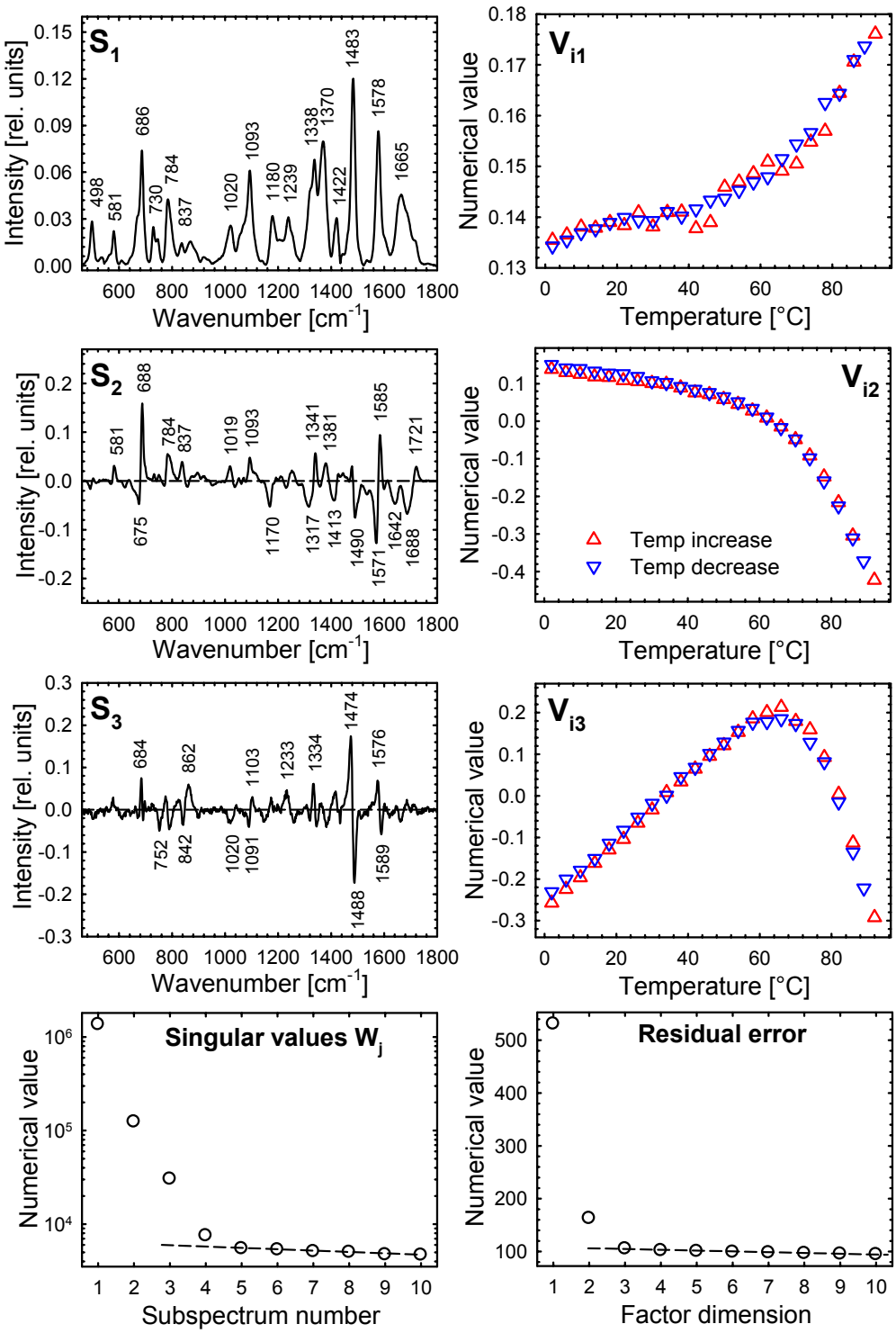

Figure S6. Factor analysis of the spectral changes induced by thermal melting of preformed (by slow annealing) K^+-stabilized parallel quadruplex $G_3(TTAG_3)_3$. DNA concentration of 250 mM (in nucleoside), in 250 mM K^+ (30 mM PBS, pH 6.8), temperature range 2 – 92°C. Analysis comprises dataset of 46 temperature-equilibrated Raman spectra taken on the temperature increase as well as decrease. Selected spectra are shown in Fig. S5 to better depict temperature-induced spectral changes. The third orthogonal spectral component S_3 resolved by SVD is needed to describe the specific temperature dependence of the 1483 cm^{-1} band which differs from thermal behavior of the other bands.