Unique Structural Features of Interconverting Monomeric and Dimeric G-Quadruplexes Adopted by a Sequence from the Intron of the N-myc Gene

ARTICLE *in* JOURNAL OF THE AMERICAN CHEMICAL SOCIETY · MARCH 2012

Impact Factor: 12.11 · DOI: 10.1021/ja208483v · Source: PubMed

3 AUTHORS:

Marko Trajkovski
National Institute of Chemistry
5 PUBLICATIONS **80** CITATIONS
SEE PROFILE

Mateus Webba da Silva
Ulster University
32 PUBLICATIONS **640** CITATIONS
SEE PROFILE

Janez Plavec
National Institute of Chemistry
198 PUBLICATIONS **2,541** CITATIONS
SEE PROFILE

JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY
Article
pubs.acs.org/JACS

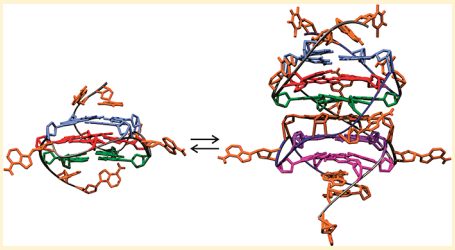
Unique Structural Features of Interconverting Monomeric and Dimeric G-Quadruplexes Adopted by a Sequence from the Intron of the N-myc Gene

Marko Trajkovski,[†] Mateus Webba da Silva,[‡] and Janez Plavec*[,†,§,∥]

[†]Slovenian NMR Center, National Institute of Chemistry, Hajdrihova 19, SI-1000 Ljubljana, Slovenia
[‡]School of Biomedical Sciences, University of Ulster, Cromore Road, Colerine BT52 1SA, U.K.
[§]Faculty of Chemistry and Chemical Technology, University of Ljubljana, Askerceva cesta 5, SI-1000 Ljubljana, Slovenia
[∥]EN→FIST Centre of Excellence, Dunajska 156, SI-1000 Ljubljana, Slovenia

Supporting Information

ABSTRACT: A multidimensional heteronuclear NMR study has demonstrated that a guanine-rich DNA oligonucleotide originating from the N-myc gene folds into G-quadruplex structures in the presence of K^+, NH_4^+, and Na^+ ions. A monomeric G-quadruplex formed in K^+ ion containing solution exhibits three G-quartets and flexible propeller-type loops. The 3D structure with three single nucleotide loops represents a missing element in structures of parallel G-quadruplexes. The structural features together with the high temperature stability are suggestive of the specific biological role of G-quadruplex formation within the intron of the N-myc gene. An increase in K^+ ion and oligonucleotide concentrations resulted in transformation of the monomeric G-quadruplex into a dimeric form. The dimeric G-quadruplex exhibits six stacked G-quartets, parallel strand orientations, and propeller-type loops. A link between the third and the fourth G-quartets consists of two adenine residues that are flipped out to facilitate consecutive stacking of six G-quartets.



INTRODUCTION

DNA containing guanine-rich regions can in the presence of cations, including the biologically most relevant K^+ and Na^+ ions, form four stranded structures called G-quadruplexes. G-quadruplexes consist of stacked G-quartets, each formed by assembly of four Hoogsteen hydrogen-bonded guanines. The stability of G-quadruplexes depends on a number of stacked G-quartets as well as on the length and sequence details of the loops that connect the guanines involved in G-quartets. The orientations of the loops are tightly related to the strand directionality of a G-quadruplex and give rise to high heterogeneity of the G-quadruplex folding topologies.[1−5]

The potential of G-quadruplexes as therapeutic targets has been well established,[6,7] with evidence of their prevalence in functional genomic domains such as telomeres and gene promoters,[8,9] together with altered gene regulation mediated by G-quadruplex stabilization.[10−15] Possible biological roles of G-quadruplexes have been implicated in many vital cell mechanisms.[16−23] Understanding of the roles of G-quadruplexes at the molecular level requires high-resolution structural information. G-quadruplexes with particular biological relevance that have recently been characterized with atomic resolution include c-myc,[24,25] c-kit,[26] c-kit2,[27] bcl-2,[28] and RET[29] from promoters of oncogenes, as well as hTERT[30] and chl1[31] originating from promoter and intronic regions, respectively. Studies on G-rich regions downstream of the transcription start site of genes and within introns of primary mRNA transcripts demonstrated the importance of G-quadruplex formation in DNA transcription and mRNA processing.[32−36]

The Myc family comprises c-myc, N-myc, and L-myc proto-oncogenes, which encode proteins involved in mechanisms regulating cell growth, differentiation, proliferation, and apoptosis.[37] Under normal conditions, the c-myc gene is expressed in all proliferating cells. Aberrance in c-myc regulation accompanies most tumor types. Expression of N-myc gene is restricted to specific tissues during their development. Its deregulation is associated with neuronal tumors.[38−40] The amplification and level of expression of N-myc oncogene have been widely studied due to their connection with the pathophysiology of cancers, most notably neuroblastoma.[41−43] To the best of our knowledge, there is no structural information on G-rich regions within the N-myc gene that could be related to its biological roles.

We initiated an NMR structural study to obtain insights into the structural features of the potentially highly relevant G-rich region located in the first intron of the N-myc gene. Our initial experiments on oligonucleotide $d[TAG_3CG_3AG_3AG_3A_2]$, N-myc, originating from the first intron of the gene, demonstrated formation of G-quadruplex in K^+ ion containing solution. An oligonucleotide, such as N-myc, containing four GGG tracts separated with single nucleotide residues was expected to fold into a monomeric G-quadruplex.[44−47] The 3D structure of the

Received: September 8, 2011
Published: February 3, 2012

dx.doi.org/10.1021/ja208483v | J. Am. Chem. Soc. 2012, 134, 4132−4141

monomeric form of N-myc represents the first example in solution of a G-quadruplex comprising three G-quartets that are bridged with three single nucleotide propeller-type loops. To our surprise, the monomeric form is involved in equilibrium with a dimeric G-quadruplex. The topology of the dimeric G-quadruplex resembles the earlier fold proposed for the c-kit2 dimer.[27] However, the dimeric N-myc exhibits unprecedented single nucleotide residue links that are flipped out from the G-quadruplex core to enable consecutive stacking of six G-quartets. The interconversion between the two forms provides novel insights into G-quadruplex polymorphism by demonstrating the intricate role of loops and their plasticity, which in combination with K^+ ion concentration drive a folding of the G-rich oligonucleotide.

MATERIALS AND METHODS

Sample Preparation. The unlabeled and the residue-specific low-enrichment (6% ^{15}N-labeled) N-myc DNA oligonucleotides were synthesized on an Expedite 8909 synthesizer using standard phosphoramidite chemistry and deprotected with the use of aqueous ammonia. Samples were purified with RP-HPLC and desalted by gel-filtration through a Sephadex G25 column. The oligonucleotide concentrations of NMR samples were between 0.4 and 6.6 mM in the strand. KCl, K-phosphate buffer, $^{15}NH_4Cl$, and NaCl were titrated into samples with the final aliquots not exceeding more than 5% of the sample volume. The concentration of KCl was varied from 0 to 100 mM, while the concentration of K-phosphate buffer (pH 6.5) was between 0 and 20 mM. Samples at different oligonucleotide and K^+ ion concentrations were folded by heating to 95 °C, followed by fast cooling to 0 °C or slow cooling (5 h) to room temperature to avoid kinetically unstable intermediate states. Formation of monomeric and dimeric species was not sensitive to the choice of the annealing protocol. The monomer form could be formed without annealing. Conversion of the dimeric form into the monomeric form was triggered by the change of the oligonucleotide concentration and did not require annealing. On the other hand, formation of the dimeric form as a predominant species in solution required annealing. In samples containing $^{15}NH_4^+$ and Na^+ ions, their concentrations were varied from 0 to 100 and from 0 to 150 mM, respectively. The pH was 6.5 for titrations with NaCl, and it was varied between 4.5 and 7.0 in the presence of $^{15}NH_4Cl$.

NMR Spectroscopy. NMR data were collected on Varian NMR System 600 and 800 MHz spectrometers in the temperature range from 5 to 90 °C. However, most of the spectra included in the manuscript were collected at 25 °C. NOESY spectra were acquired with mixing times of 80, 100, and 200 ms in 90% H_2O, 10% ^2H_2O, and 100% ^2H_2O. DQF-COSY, $^{31}P-^1H$ COSY, and TOCSY spectra with mixing time of 80 ms were acquired in 100% ^2H_2O. Intranucleotide correlations between imino and aromatic protons were acquired on a natural abundance sample with the use of a jump-and-return HMBC (JRHMBC) experiment. Assignment of imino protons was done by ^{15}N-filtered ^{1}H spectra of 6% residue-specific ^{15}N-labeled samples. In the case of the dimeric form, assignment of H8 protons was accomplished by $^{15}N-^1H$ HMQC spectra of 6% residue-specific ^{15}N-labeled samples in combination with other 2D NMR spectra. Thirty different gradient strengths (0.49−29.06 G/cm) were used in diffusion experiments, which were acquired in 100% ^2H_2O. NMR spectra were processed and analyzed using VNMRJ (Varian Inc.) and Sparky (UCSF) software.

UV spectroscopy. Melting experiments were performed on a Cary 100 spectrometer equipped with a temperature control system. A temperature range from 10 to 95 °C was scanned while monitoring absorbance at 295 nm with a 0.1 °C min^{-1} melting/annealing rate. Sample solutions contained 100 mM KCl and 20 mM K-phosphate buffer (pH 6.5). Oligonucleotide concentrations were in the range from 40 to 210 μM. Measurements were made in 0.2, 0.5, and 1 cm path-length cells.

CD Spectroscopy. CD experiments were carried out on an Applied Photophysics Chirascan CD spectrometer at 25 °C over the 200−320 nm wavelength range. Measurements were made in 0.1 and 0.01 cm path-length quartz cells. The oligonucleotide concentrations were in the range from 20 to 370 μM. The samples contained 100 mM KCl and 20 mM K-phosphate buffer (pH 6.5).

Native PAGE. N-myc samples together with a GeneRuler Ultra Low Range DNA ladder with 10−300 base pairs (Fermentas) were loaded and resolved on 20% native PAGE gels, which were supplemented with 20 mM or 100 mM KCl. Stains All (Sigma-Aldrich) was used for staining. The samples that were loaded on gels were from 0.05 to 6.00 mM in oligonucleotide concentration per strand, 20 mM K-phosphate buffer (pH 6.5), with no additional KCl or alternatively with 100 mM KCl. Bromophenol blue and xylene cyanol were used as migration markers in DNA ladder only.

Analytical Ultracentrifugation. Sedimentation velocity experiments were performed as described earlier.[48]

Restraints and Structure Calculations. NOESY spectra recorded at 80, 100, and 200 ms mixing times in 90% H_2O, 10% ^2H_2O, and 100% ^2H_2O were used to obtain distance restraints for exchangeable and nonexchangeable protons, respectively. The volumes of NOE cross-peaks corresponding to H2′−H2″ (1.9 Å) and cytosine H5−H6 (2.5 Å) were used as references. Cross-peaks were classified as strong (1.8−3.6 Å), medium (2.6−5.0 Å), and weak (3.5−6.5 Å). Torsion angle restraints along glycosidic bonds (torsion angle χ) were based on intraresidual NOE correlations of H8 with H1′, H2′, and H2″ protons, which clearly established the *anti* conformation of all guanine residues in monomeric and dimeric forms. For other residues, no torsion angle restraints were used to restrict glycosidic torsion angles. The values of the $^3J_{H1'-H2'}$ and $^3J_{H1'-H2''}$ coupling constants measured in DQF-COSY spectra were consistent with predominance of South-type sugar puckering for all residues. The endocyclic torsion angles ν_1 and ν_2 were restrained in order to limit the available conformational space of pseudorotational parameters in the range from C3′-exo to C1′-exo canonical forms. The observed values of 13 Hz or less for the $^3J_{P-H3'}$ coupling constants were used to restrain torsion angles ε to a range of 225 ± 75° for all residues except the 3′ terminal residues. The chemical shift range of ^{31}P resonances was 2 and 3 ppm for monomeric and dimeric species, respectively. These narrow chemical shift ranges are consistent with standard backbone torsion angles.

Initial starting structures consisted of a single strand in the case of the monomeric form and two single strands in the case of the dimeric form, which were created using X3DNA software[49] and the LEAP module of the AMBER 9 program.[50] They were subjected to extensive conformational search utilizing the restraints listed in Tables 1 and 2. Simulated annealing (SA) calculations were performed with the AMBER 9 program with the parmbsc0 force field[51] using the generalized Born implicit solvation model. Random velocity was used for each of the SA calculations. The SHAKE algorithm for hydrogen atoms was used with a tolerance of 0.00005 Å. The cutoff for nonbonded interactions was 20 Å. After SA calculations, 10 structures with the lowest energy were subjected to a maximum of 100000 steps of steepest descent energy minimization. Hydrogen bond restraints were omitted prior to the final stages of structural refinement. Structural statistics for 10 structures for monomeric and dimeric forms are given in Tables 1 and 2, respectively. UCSF Chimera software was used for visualization and preparation of figures.[52] We used 3DNA 2.0 software to determine helical parameters.[49] Details of the SA protocol are given in the Supporting Information. The coordinates of the monomeric and dimeric G-quadruplexes adopted by N-myc have been deposited in the Protein Data Bank with accession codes 2LEE and 2LED, respectively.

RESULTS

Equilibrium between Two G-quadruplex Forms. The imino region of the ^{1}H NMR spectrum of a desalted sample of N-myc exhibited two broad signals at δ 11.63 and 11.93 ppm, indicating the absence of a well-defined structure. The spectrum changed dramatically upon addition of K^+ ions. At

Table 1. NMR Restraints and Structural Statistics for the N-myc Monomeric G-Quadruplex

NOE-derived distance restraints	non-exchangeable	exchangeable
intranucleotide NOEs	238	0
sequential (i, i + 1)	50	9
long-range (i, >i + 1)	3	32
torsion angle restraints	68	
hydrogen bond restraints[a]	24	
structural statistics		
NOE violations >0.3 Å	0	
deviations from idealized covalent geometry		
bonds (Å)	0.011 ± 0.000	
angles (deg)	2.355 ± 0.021	
pairwise heavy atom rmsd (Å)		
overall	1.430	
without 5′ TA overhang	1.493	
without 3′ AA overhang	1.197	
without 5′ TA and 3′ AA overhangs	1.253	
without C6, A10, and A14	1.038	
only guanine residues	0.488	

[a]Hydrogen bond restraints were omitted prior to the final stages of structural refinement.

Table 2. NMR Restraints and Structural Statistics for the N-myc Dimeric G-Quadruplex

NOE-derived distance restraints	non-exchangeable	exchangeable
intranucleotide NOEs	428	0
sequential (i, i + 1)	84	14
long-range (i, >i + 1)	40	102
torsion angle restraints	136	
hydrogen bond restraints[a]	48	
structural statistics		
NOE violations 0>.3 Å	0	
deviations from idealized covalent geometry		
bonds (Å)	0.011 ± 0.000	
angles (deg)	2.545 ± 0.040	
pairwise heavy atom rmsd (Å)		
overall	1.302	
without 5′ TA overhangs	1.356	
without 3′ AA overhangs	1.056	
without 5′ TA and 3′ AA overhangs	1.100	
without C6, A10, and A14	1.029	
only guanine residues	0.528	

[a]Hydrogen bond restraints were omitted prior to the final stages of structural refinement.

around 20 mM concentration of K^+ ions, twelve narrow peaks were observed in the range from δ 10.82 to 11.96 ppm. The signals corresponded to twelve imino resonances (Figure 1a), which were consistent with the formation of one predominant G-quadruplex structure with three G-quartets. A translational diffusion coefficient of $(1.32 \pm 0.05) \times 10^{-6}$ cm^2 s^{-1} suggested formation of a monomeric G-quadruplex topology. Increase of the total K^+ ion concentration to 120 mM induced transformation of monomeric species into a different G-quadruplex structure with the corresponding set of twelve narrow peaks in the range from δ 10.62 to 11.48 ppm of the ^{1}H NMR spectrum (Figure 1b). The corresponding translational diffusion coefficient of $(1.05 \pm 0.05) \times 10^{-6}$ cm^2 s^{-1} indicated formation of a larger species. Later analysis (*vide infra*) with the use of the HYDROPRO program[53] based on the calculated structure of the dimeric G-quadruplex demonstrated good agreement with the above experimental value. Imino protons of G15 and especially G4, G8, G12, and G16 of the dimeric G-quadruplex adopted by N-myc exhibited a slow rate of exchange in deuterated water, which suggested their involvement in the inner G-quartets (Figure S1 of the Supporting Information). NMR spectra of N-myc at conditions favoring the dimeric form exhibited signals corresponding to dimeric and monomeric forms as well as minor broad signals corresponding to higher order structures, which could not be assigned. The presence of higher order structures in addition to dimeric and monomeric forms was confirmed by PAGE gel analysis (Figure S2). The oligonucleotide concentration dependence of the ratio between monomeric, dimeric, and higher order species observed with PAGE was consistent with NMR data. In order to complement our NMR results on the hydrodynamic properties of different species adopted by N-myc in the presence of K^+ ions, sedimentation velocity ultracentrifugation analysis[54] was performed, which clearly established folding of N-myc into monomeric, dimeric, and higher order structures and the role of K^+ ion concentration on the equilibrium (Figure S3).

The equilibrium between the monomeric and dimeric G-quadruplexes adopted by N-myc was also found to depend on oligonucleotide concentration. At 120 mM K^+ ion concentration the monomeric form predominated over the dimeric form at concentrations of oligonucleotide below 1.3 mM. At 1.3 mM oligonucleotide and 120 mM K^+ ion concentrations, the ratio between the two species was 1:1, whereas above 1.3 mM concentration of oligonucleotide the dimeric form predominated. Additionally, the equilibrium exhibited a pronounced variation with concentration of K^+ ions. The monomeric G-quadruplex predominated at up to 6 mM oligonucleotide concentration as long as the concentration of K^+ ions was kept below 20 mM. Conversion of the dimeric form into the monomeric form was triggered by the change of the oligonucleotide concentration and did not require annealing. On the other hand, formation of the dimeric form as a predominant species in solution required annealing.

Heating of samples to 95 °C followed by different annealing protocols (fast or slow) did not result in a notable difference in ratio between monomeric and dimeric forms. This observation was consistent with the fast rate of folding of both forms. The comparison of resolved ^{1}H NMR signals and their relative intensities corresponding to monomeric and dimeric forms in equilibrium revealed that both forms exhibit very similar properties in the range between 25 and 65 °C. To examine the respective time-dependent stabilities of monomeric and dimeric forms, we compared the NMR spectra of N-myc at conditions with predominating either monomeric or dimeric form before and after the samples were held at 4 °C for ca. 6 months. The lack of spectral changes demonstrated long-term stability for both forms.

UV melting experiments demonstrated a very high temperature stability of N-myc in the presence of K^+ ions. The process of unfolding was not complete even at 95 °C for oligonucleotide concentrations from ca. 40 to 200 μM (Figure S4 of the Supporting Information). The high thermal stability precluded determination of the melting temperature. Melting curves, however, indicated that the melting transition of N-myc was not concentration-dependent, which was consistent with a

dx.doi.org/10.1021/ja208483v | J. Am. Chem. Soc. 2012, 134, 4132–4141

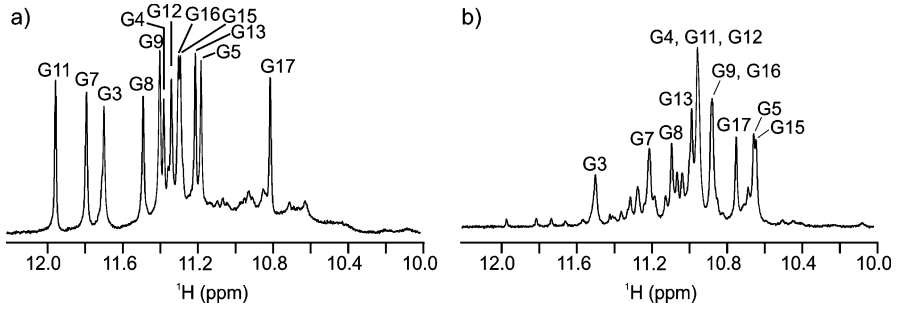


Figure 1. Imino region of the ^{1}H NMR spectra of N-myc at (a) 2.7 mM oligonucleotide concentration per strand, 20 mM K-phosphate buffer and (b) 6.6 mM oligonucleotide concentration per strand, 100 mM KCl, 20 mM K-phosphate buffer. The spectra were recorded at 25 °C and pH 6.5.

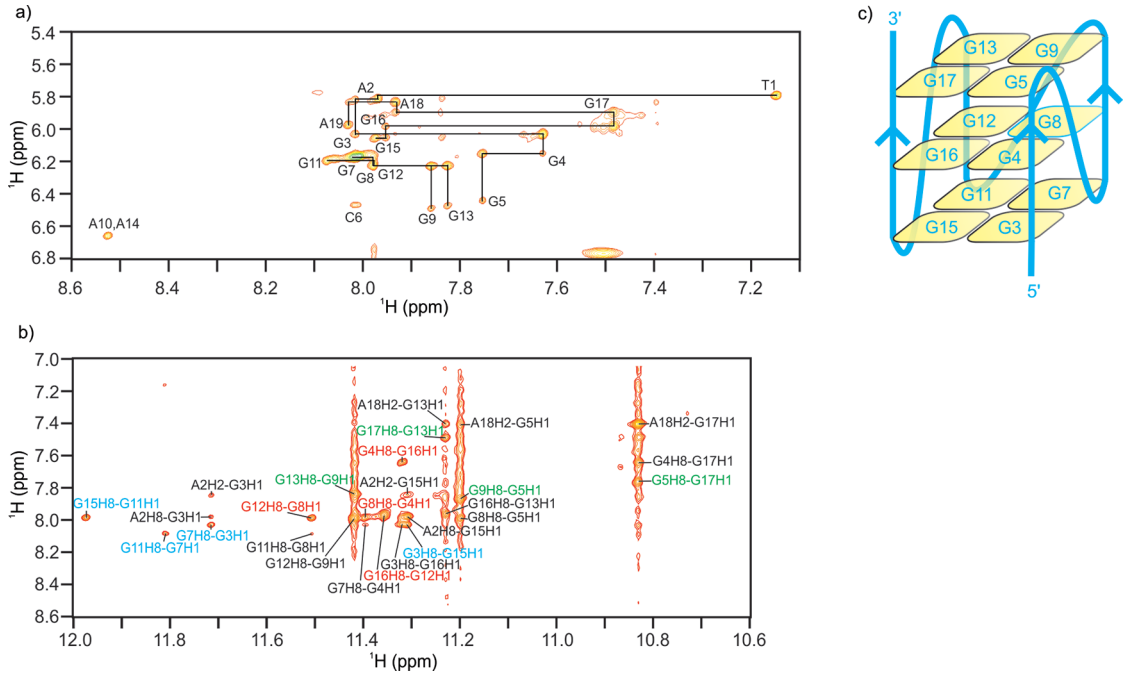


Figure 2. (a) Anomeric–aromatic and (b) aromatic–imino regions of the NOESY spectrum of N-myc with a mixing time of 200 ms (25 °C, pH 6.5). The concentrations of oligonucleotide and K^{+} ions were 2.0 mM and 20 mM, respectively. In panel a, intraresidual H1′-H6/H8 NOE interactions are labeled and sequential correlations are denoted with lines. In panel b, H8–H1 cross-peaks within G3-G7-G11-G15, G4-G8-G12-G16, and G5-G9-G13-G17 quartets are labeled in blue, red, and green, respectively. Cross-peaks corresponding to residues that are not part of the same G-quartets are labeled in black. (c) Topology of the N-myc monomeric G-quadruplex.

monomolecular fold. Furthermore, the melting of the dimeric form was not reflected in temperature-dependent UV absorbance profiles.

The CD spectra of N-myc were consistent with the formation of parallel G-quadruplexes at oligonucleotide concentrations in the range from ca. 20 to 400 μM. CD profiles of N-myc at different oligonucleotide concentrations are shown in Figure S5 of the Supporting Information.

Assignment of NMR Spectra of Monomeric and Dimeric G-Quadruplexes Adopted by N-myc. All twelve imino protons of the monomeric G-quadruplex adopted by N-myc were unambiguously assigned using NMR experiments in 90% H_2O (Figure S6a and b of the Supporting Information). NOESY (Figure 2a), JRHMBC (Figure S7a and b), and other 2D NMR spectra (e.g., TOCSY) were used for assignment of imino, aromatic, sugar, and methyl protons. Intraresidual correlations in the aromatic–anomeric region of the NOESY spectrum were consistent with *anti* glycosidic torsion angle values for all guanine residues. Perusal of NOE correlations between imino and H8 protons for guanines demonstrated formation of three G-quartets featuring the following hydrogen bond directionalities: G3-G7-G11-G15, G4-G8-G12-G16, and G5-G9-G13-G17 (Figure 2b). The four imino protons of the central G4-G8-G12-G16 quartet exhibited the slowest rate of exchange in deuterated water (Figure S6c) and more intense intraresidual NOE correlations with amino protons. Altogether our NMR data were consistent with the folding topology with all four strands in a parallel orientation (Figure 2c). The absence of interresidual correlations in NOESY spectra for residues constituting the three propeller-type loops, C6, A10, and A14, suggested their flexibility. Isochronous ^{1}H and ^{13}C chemical shifts of the signals corresponding to A10 and A14 were consistent with the very similar shielding environments experienced by these two residues, further supporting the model of the monomeric G-quadruplex with three flexible loops.

Assignment of aromatic, sugar, and methyl protons of the dimeric G-quadruplex adopted by N-myc was achieved with the use of NOESY (Figure 3a and b) in combination with TOCSY and JRHMBC (Figure S8 of the Supporting Information)

dx.doi.org/10.1021/ja208483v | J. Am. Chem. Soc. 2012, 134, 4132–4141

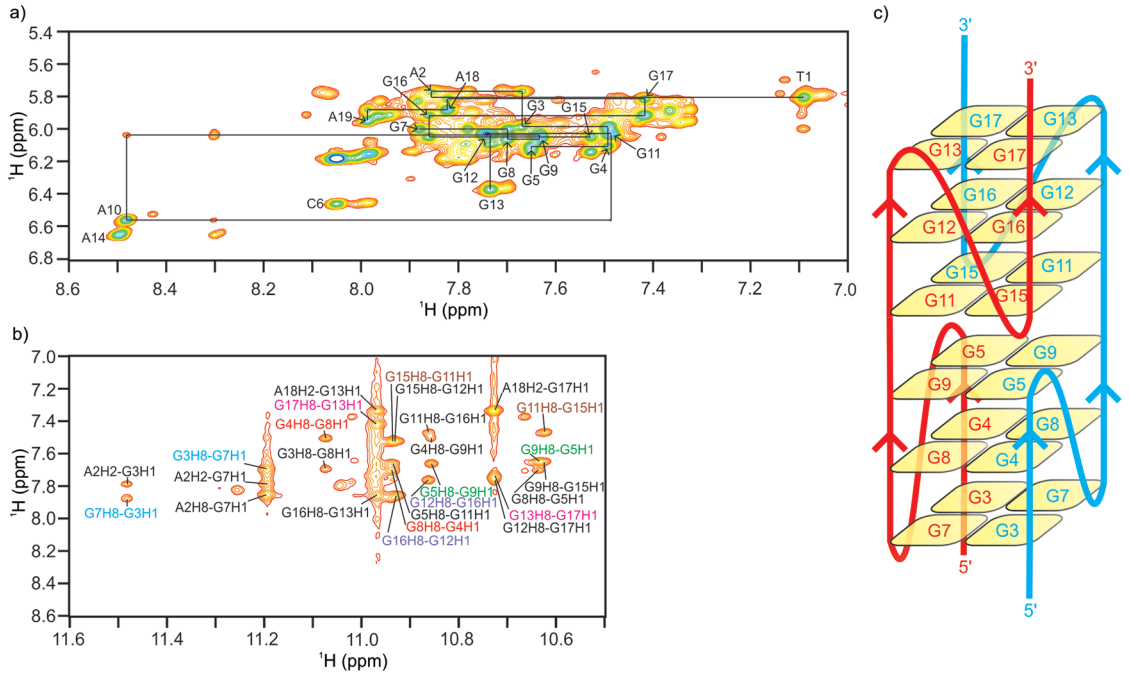


Figure 3. (a) Anomeric–aromatic and (b) aromatic–imino regions of the NOESY spectrum of N-myc with a mixing time of 200 ms (25 °C, pH 6.5). The concentrations of oligonucleotide and K^+ ions were 6.6 mM and 120 mM, respectively. In panel a, intraresidual H1′-H6/H8 NOE interactions are labeled and sequential correlations are denoted with lines. In panel b, H8–H1 cross-peaks within G3-G7-G3-G7, G4-G8-G4-G8, G5-G9-G5-G8, G11-G15-G11-G15, G12-G16-G12-G16, and G13-G17-G13-G17 quartets are labeled in blue, red, green, brown, purple, and magenta, respectively. Cross-peaks corresponding to residues that are not part of the same G-quartets are labeled in black. (c) Topology of dimeric N-myc G-quadruplex.

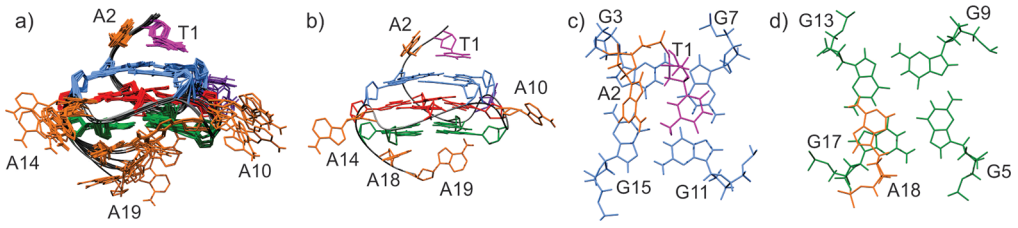


Figure 4. (a) Family of 10 superpositioned refined structures of the N-myc monomeric G-quadruplex. Guanine residues in G3-G7-G11-G15, G4-G8-G12-G16, and G5-G9-G13-G17 quartets are colored in light blue, red, and green, respectively. Adenine, cytosine, and thymine residues are orange, purple, and magenta, respectively. (b) Representative structure of the N-myc monomeric G-quadruplex. (c) Bird's-eye view of stacking of the 5′ TA overhang over the G3-G7-G11-G15 quartet. (d) Bird's-eye view of stacking of A18 over the G5-G9-G13-G17 quartet.

spectra as well as with ^{15}N-filtered ^{1}H (Figure S9) and ^{15}N–^{1}H HMQC (Figure S10) spectra of residue-specific ^{15}N-labeled samples. Intraresidual NOE interactions were consistent with *anti* glycosidic torsion angle values for all guanine residues. NOESY spectra between imino and H8 protons exhibited characteristic correlations within six G-quartets with the following hydrogen bond directionalities: G3-G7-G3-G7, G4-G8-G4-G8, G5-G9-G5-G9, G11-G15-G11-G15, G12-G16-G12-G16, and G13-G17-G13-G17. NOE interactions within each G-quartet together with interactions among different G-quartets, most notably G5-G9-G5-G9 and G11-G15-G11-G15 quartets, confirmed the topology of the dimeric form presented in Figure 3c. In particular, G5 H1′-G15 H8, G5 H8-G15 H8, G5 H2″-G15 H8, G9 H2″-G11 H8, G9 H2′-H11 H8, etc. as well as NOE interactions between overhanging residues and guanine residues constituting the two outer G-quartets were consistent solely with formation of a symmetric dimeric G-quadruplex comprised of six G-quartets with all strands in a parallel orientation. The proposed topology with imino protons of G7 and especially G3 exposed to solvent exchange was consistent with their larger signal half-widths (Figure S1a).

The large number of NOE interactions for monomeric and dimeric G-quadruplexes adopted by N-myc warranted determination of the corresponding high-resolution structures.

Structure of the Monomeric G-Quadruplex Adopted by N-myc. The solution-state structure of the N-myc monomeric G-quadruplex in K^+ solution was calculated using 332 NOE-derived distance restraints together with 68 torsion angle and 24 hydrogen bond restraints (Table 1). A family of 10 structures of the monomeric G-quadruplex adopted by the N-myc presented in Figure 4a exhibited a pairwise heavy atom rmsd of 1.4 Å. The translational diffusion coefficient calculated by the HYDROPRO program[53] using coordinates of the calculated structure of the N-myc monomeric G-quadruplex is 1.25×10^{-6} cm^2 s^{-1}, which is in good agreement with our experimental data. The core of the structure consists of stacked G3-G7-G11-G15, G4-G8-G12-G16, and G5-G9-G13-G17 quartets, which are formed by four parallel GGG tracts characterized by G_{anti}-G_{anti}-G_{anti} steps. The average helical rise

dx.doi.org/10.1021/ja208483v | *J. Am. Chem. Soc.* 2012, 134, 4132–4141

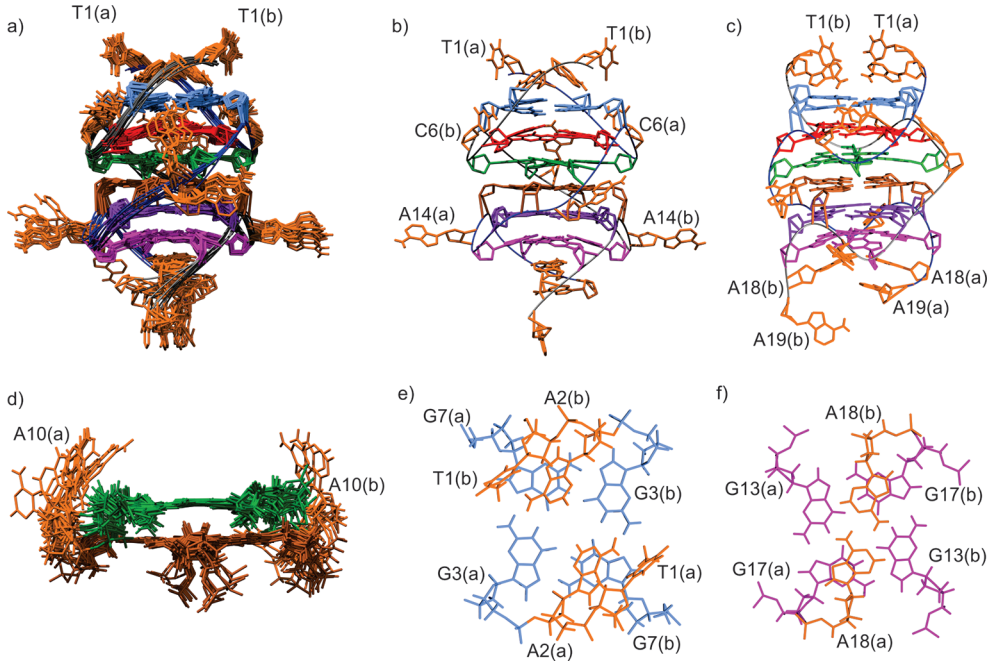


Figure 5. (a) Family of 10 superpositioned refined structures of the dimeric N-myc G-quadruplex. T1 residues of the two strands are designated with letters a and b. Guanines in G3-G7-G3-G7, G4-G8-G4-G8, G5-G9-G5-G9, G11-G15-G11-G15, G12-G16-G12-G16, and G13-G17-G13-G17 quartets are colored in light blue, red, green, brown, purple, and magenta, respectively, while residues from the 5′ and 3′ overhangs as well as from the loops are colored in orange. (b and c) Two different side views of a representative structure. (d) Side view of G5-G9-G5-G9 and G11-G15-G11-G15 quartets (in green and brown, respectively) connected by A10 (in orange). (e) Bird's-eye view of the 5′ TA overhangs' stacking over the outer G-quartet. (f) Bird's-eye view of the stacking of A18 over the outer G13(a)-G17(a)-G13(b)-G17(b).

and twist for consecutive G-quartets are 3.2 Å and 29.5°, respectively. The helical rise in the N-myc monomeric G-quadruplex is close to the typical value for the B-DNA double helix (3.4 Å), whereas the helical twist of the N-myc monomeric G-quadruplex is ca. 7° smaller in comparison to the case of B-DNA and ca. 3° smaller in comparison to the A-DNA double helix. The four groove widths of the N-myc monomeric G-quadruplex are on average 10.7 ± 0.5 Å, which is close to a typical minor groove width in A-DNA (11 Å). Out-of-plane bending toward the 5′ end is observed for all guanine residues, except for G11 and G12. The N-myc monomeric G-quadruplex exhibits three propeller-type single nucleotide loops, each bridging a three G-quartet core. The residues in the loops, C6, A10, and A14, are not well-defined.

Both 5′ overhanging residues, T1 and A2, exhibit well-defined conformations and are stacked on top of the G3-G7-G11-G15 quartet (Figure 4c). T1 is stacked on a nearby G-quartet and positioned close to G7. A2 is positioned on top of the G3 and G15 interface with the purine ring inclined by ca. 45° relative to the plane of the nearby G-quartet. The first overhanging residue at the 3′ end, A18, is well-defined and stacked on G17 (Figure 4d). Stacking of G17 and A18 involves overlap of imidazole moieties. The disposition of A19 is not well-defined.

Dimeric G-Quadruplex Adopted by N-myc Exhibits Unique Structure. The solution-state structure of the dimeric G-quadruplex adopted by N-myc in K^+ solution was calculated with the use of 668 NOE-derived distance, 136 torsion angle, and 48 hydrogen bond restraints. Hydrogen bond restraints were not used in the final stages of structure refinement. A family of 10 refined structures of the dimeric G-quadruplex adopted by N-myc exhibit a pairwise heavy atom rmsd of 1.3 Å (Table 2 and Figure 5). With the use of the HYDROPRO program,[53] a value of 1.06×10^{-6} cm^2 s^{-1} for the translational diffusion coefficient of the calculated structure of the N-myc dimeric G-quadruplex is obtained, which is in good agreement with our experimental data. All guanines exhibit *anti* glycosidic torsion angle values and are involved in G-quartet formation. The six stacked G-quartets, G3-G7-G3-G7, G4-G8-G4-G8, G5-G9-G5-G9, G11-G15-G11-G15, G12-G16-G12-G16, and G13-G17-G13-G17, represent the core of the structure. The first three G-quartets at the 5′ end of the structure are formed by guanines of the first and the second GGG tracts of the oligonucleotide sequence, while the three G-quartets at the 3′ end of the structure are formed by guanines of the third and the fourth GGG tracts. Residues G3, G4, G7, G8, G12, G13, G16, and G17 exhibit out-of-plane bending with O6 oriented toward the 5′ end. As a result, all G-quartets except G5-G9-G5-G9 and G11-G15-G11-G15 exhibit a slight distortion in planarity. C6 and A14 form single nucleotide propeller-type loops, which bridge a three G-quartet core. The conformation of C6 is well-defined, and the proximity of its amino protons to N3, O3′, and O4′ of G3 as well as to O4′ of G4 suggests that they are involved in hydrogen-bonding interactions. Multiple possible acceptors are in accordance with the fast exchanging regime of C6 amino protons, which is consistent with the absence of the corresponding signals in the NMR spectra. A10 in each of the strands connects G9 and G11, which are part of the neighboring G-quartets (Figure 5d). Stacked G5-G9-G5-G9 and G11-G15-G11-G15 quartets represent the part of the structure where nonconsecutive guanines are involved in formation of neighboring G-quartets. Although all six G-quartets are directly stacked on top of each other, the rise and helical twist for consecutive G-quartets vary substantially through the entire structure. The N-myc dimeric G-quadruplex exhibits the largest helical twist for G5-G9-G5-G9 and G11-

dx.doi.org/10.1021/ja208483v | J. Am. Chem. Soc. 2012, 134, 4132–4141

G15-G11-G15 quartets with an average value of 33.4°. Interestingly, the G15 imino proton exhibits a slower rate of exchange with the solvent with respect to G5, G9, and G11. Solvent exposure of G5, G9, G11, and G15 in the calculated structure is consistent with the solvent exchange experiments (Figure S1 of the Supporting Information).

The groove widths exhibited by the N-myc dimeric G-quadruplex are larger between strands belonging to the same oligonucleotide in comparison to those of different molecules for the region comprising G3−G9, whereas the opposite is true for the region comprising G11−G17. Widths of grooves in the regions comprising G3−G9 and G11−G17 are ca. 11.4 Å and 10.5 Å, respectively, between strands of the same molecule, and ca. 10.1 Å and 11.4 Å, respectively, between strands of different molecules. Thus, this set of values stretches over the range typical for the minor groove in A-DNA (11.0 Å) and for antiparallel (12.0 Å) and parallel (10.2 Å) G-quadruplexes.

Both T1 and A2 residues of 5′ overhangs are well-defined (Figure 5e) and positioned above G7 of the complementary strand, leaving the central cation cavity inside the G-quadruplex opened. The 3′ overhang residues A18 are structurally well-defined, with their imidazole moieties involved in stacking interactions with the pyrimidine moieties of G17 (Figure 5f). The pyrimidine moieties of A18 are oriented toward the top of the central cation cavity of the G-quadruplex. Although mutual arrangements of A18 residues from the two strands are not planar, their positioning indicates formation of an A18-A18 N1-amino symmetrical noncanonical base pair. There is, however, no experimental evidence for amino protons of A18 residues involved in hydrogen-bonding interactions. A19 exhibits high conformational flexibility.

Signals in the ^{1}H NMR spectrum corresponding to the imino protons of G3 and G7 are broader than those corresponding to the other guanine residues, thus reflecting higher dynamics of the G3-G7-G3-G7 quartet. Possibly two factors contribute to this. First, the imino protons of these two guanines are more exposed to solvent, as they are part of the outer G-quartet at the 5′ end. Second, the dynamics of the G3-G7-G3-G7 quartet is higher in comparison to the other G-quartets due to interactions with 5′ TA overhangs. Inspection of the family of 10 refined structures of the N-myc dimeric G-quadruplex reveals that G3 and G4 exhibit the largest out of G-quartet plane bending toward 5′ TA overhangs. These observations imply that 5′ TA overhangs indeed induce elevated dynamics of the G3-G7-G3-G7 quartet.

Assessment of the Role of NaCl and $^{15}NH_4Cl$ in Folding of N-myc. We attempted to fold N-myc in the presence of Na^+ and $^{15}NH_4^+$ ions. Folding of N-myc was monitored by recording ^{1}H NMR spectra during titrations with NaCl and $^{15}NH_4Cl$. The imino region of ^{1}H NMR spectra at up to 150 mM Na^+ ion concentration exhibited five major sharp and additional broad minor signals superimposed on a broad hump, all within the chemical shift range characteristic for G-quadruplexes formation. Signals in the aromatic and methyl regions of the ^{1}H NMR spectrum observed before addition of Na^+ ions, which correspond to unfolded oligonucleotide, were preserved at up to 150 mM Na^+ ion concentrations. This indicated that Na^+ ions are much less effective in stabilizing N-myc G-quadruplex formation in comparison to K^+ ions, which induced formation already at concentration below 5 mM. ^{1}H NMR spectra were recorded for N-myc in the presence of Na^+ ions after the sample had been stored overnight at 25 °C followed by 4 °C for 3 months. The spectra showed that with time the initial hump in the imino region gained in intensity, while all the sharp imino signals were broadened almost to the baseline. Consistently, signals in the aromatic and methyl regions were also broadened with time. NMR results were consistent with G-quadruplex formation of N-myc at 150 mM concentration of Na^+ ions being an intermediary step in the process, eventually leading to coexistence of multiple species and aggregation.

Upon addition of $^{15}NH_4Cl$, the imino region of the ^{1}H NMR spectra of N-myc exhibited fifteen major and additional minor signals, all within the chemical shift range characteristic for G-quadruplex formation, thus indicating formation of several G-quadruplex species. By varying oligonucleotide and $^{15}NH_4Cl$ concentrations, pH, and temperature, we were not able to push the equilibrium to a single G-quadruplex species. With the use of ^{15}N-filtered ^{1}H and ^{1}H−^{1}H NOESY spectra, we identified four major signals corresponding to bound $^{15}NH_4^+$ ions with corresponding ^{1}H chemical shifts of δ 6.75, 6.79, 6.86, and 6.93 ppm. $^{15}NH_4^+$ ions bound within the inter G-quartet cavities of a G-quadruplex structure are correlated to eight imino proton signals in 2D ^{1}H−^{1}H NOESY spectra. Interestingly, each of the four major signals corresponding to bound $^{15}NH_4^+$ ions within the N-myc G-quadruplex was correlated only to four imino signals. This implied that they were localized within the G-quadruplex formed by symmetrically arranged N-myc oligonucleotide strands. In NOESY spectra the signals corresponding to bound $^{15}NH_4^+$ ions at δ 6.75 and 6.93 ppm were correlated to the same two imino signals, indicating that they were localized in consecutive inter G-quartet cavities. The same was true for signals corresponding to bound $^{15}NH_4^+$ ions at δ 6.79 and 6.86 ppm. Due to multiple species present and poor dispersion of signals we were not able to complete assignment of NMR signals.

DISCUSSION

The oligonucleotide d[$TAG_3CG_3AG_3AG_3A_2$], N-myc, with four GGG runs each separated from the preceding one by a single nucleotide residue, was expected to adopt a monomeric G-quadruplex structure with all strands in a parallel orientation in the presence of K^+ ions. In fact, our experimental observations based on NMR data confirmed these indications and showed that N-myc folded into a monomeric G-quadruplex. However, the determined structure of the monomeric G-quadruplex adopted by N-myc is novel with regard to the loop lengths and yet represents the simplest among the high-resolution structures of G-quadruplexes containing three G-quartets. Interestingly, residues in all three propeller-type loops are exposed to solvent. On the other hand, 5′ overhang residues T1 and A2 as well as 3′ overhang residue A18 are well-defined. The interactions of overhanging T and A residues with the guanine residues in the neighboring G-quartets are possibly responsible for their out-of-plane bending. The G-quadruplex features slightly wider grooves (10.7 ± 0.5 Å) than typically observed for parallel G-quadruplexes (10.2 Å). The monomeric G-quadruplex adopted by N-myc with three G-quartets and three propeller-type loops is common to c-myc,[11,24,25] c-kit2,[27] hTERT,[30] HIF-1α,[16] RET,[29] VEGF,[22] and KRAS[55] promoters as well as human telomeric repeat.[56] All these structures exhibit, at least to some extent, well-defined loops and at least one loop comprised of more than a single residue. The G-quadruplex adopted by N-myc exhibiting flexible single nucleotide loops was thus a missing element among high-resolution structures consisting of three G-

quartets. Its high temperature stability implies the potential role of G-quadruplex formation in transcription of the N-myc gene. The unstructured loops could serve as a recognizing element for selective binding of proteins regulating transcription.

The dimeric N-myc G-quadruplex exhibits all strands in a parallel orientation, and the core of the structure is comprised of six G-quartets. Overhanging residues T1 and A2 at the 5′ end are well-defined and positioned on top of the nearby G-quartet, leaving the cation cavity along the center of the G-quadruplex core opened. The orientation of the A18 residues at the 3′ end suggests formation of a noncanonical A-A base pair. Stacking of the pyrimidine moieties of A18 on top of the central cation cavity could control the dynamics of cation movement along the G-quadruplex.[57−59] The dimeric G-quadruplex of N-myc with a single adenine residue (A10) flipped out from the core enables direct stacking of the third and the fourth G-quartets in a contiguous stack of six G-quartets. Furthermore, the structure exhibits grooves that vary in their dimensions along the strands and are related to specific arrangements of loops. The first loop comprising a single cytosine residue that bridges the three G-quartet core is structurally well-defined, while the other two loop residues, A10 and A14, are not.

N-myc, with single nucleotide residues connecting the GGG tracts, comprises a sequence of minimal length used as a criterion in algorithms developed to identify primary sequences within the human genome with the potential to adopt G-quadruplex structures.[8] Such bioinformatics analysis combined with rules and examples enabling prediction of G-quadruplex folding topology and stability[60−65] could in the future lead to new ways of identifying potential G-quadruplexes as therapeutic targets for selective gene regulation. Monomeric and dimeric forms of N-myc are interesting examples of G-quadruplex topologies that are not predefined by the length of the loops between guanines constituting the G-quadruplex core. Both forms exhibit flexible loop residues A10 and A14, which is in agreement with the suggestion that G-quadruplexes with loops containing adenine instead of cytosine or thymine residues may permit interconversion between different structural forms.[66] The equilibrium between monomeric and dimeric G-quadruplexes adopted by N-myc is controlled by concentrations of oligonucleotides and K^+ ions. At around 100 mM concentration of K^+ ions coupled with oligonucleotide concentration higher than 1.3 mM, the dimeric G-quadruplex is predominant over the monomeric form. Interestingly, both forms exhibit very similar temperature-dependent profiles and folding rates as well as stability over a longer time period.

Recently it has been shown that depending on the concentration of K^+ ions the human c-kit2 promoter sequence adopts monomeric and dimeric G-quadruplexes.[27] Both forms exhibit topologies similar to those of the monomeric and dimeric G-quadruplexes adopted by N-myc. Interestingly, without considering the 5′ and 3′ overhangs, c-kit2 and N-myc differ only in the central loop, which is comprised of five residues (i.e., CGCGA) in c-kit2 and a single residue (i.e., A) in N-myc. The dimeric G-quadruplex adopted by c-kit2 exhibits an A13-A13 noncanonical base pair, which is intercalated between the third and the fourth G-quartets and thus interrupts their direct stacking. In the case of N-myc, the third and the fourth G-quartets are stacked directly while the two adenine residues (A10) are turned away from the G-quadruplex core. Two propeller-type loops are comprised of single cytosine and adenine residues in the dimeric G-quadruplexes adopted by N-myc and c-kit2. In the case of N-myc, the orientation of the C6 loop residue is well-defined, whereas the A14 loop residue is flexible. In comparison, the C5 loop residue is flexible in the c-kit2 dimeric G-quadruplex, while the A17 loop residue is well-defined and clustered with residues comprising the loop between the third and the fourth G-quartets.

Recent observations indicate involvement of potentially G-quadruplex forming sequences in recombination events.[19,21] For the dimeric G-quadruplex adopted by c-kit2, a strand exchange mechanism involving cleavage in the region of an A-A noncanonical base pair was suggested in the postulated recombination mechanism.[27] It is tempting to speculate that the same part of the structure comprising the third and the fourth G-quartets, with an intercalated A-A noncanonical base pair in c-kit2 in comparison to exposed A residues in N-myc, could represent different branch points enabling specific recombination events.

Depending on concentration of K^+ ions, N-myc adopted two discrete G-quadruplexes. Addition of Na^+ and $^{15}NH_4^+$ ions induced formation of multiple G-quadruplex species of N-myc, with the corresponding equilibria not being substantially dependent on the concentration of cations. The G-quadruplex formed upon initial addition of Na^+ ions was transformed over time into multiple G-quadruplex species followed by their aggregation. Na^+ ions induce formation of intermolecular assemblies through kinetically stabilized G-quadruplex intermediate(s). In comparison, G-quadruplexes formed in the presence of K^+ and $^{15}NH_4^+$ ions were stable over longer time periods. In the presence of $^{15}NH_4^+$ ions, N-myc adopted predominantly intermolecular G-quadruplexes with the highly symmetric cation binding sites. Cation dependence of the folding of N-myc together with the high-resolution structural information presented here provide important new insights that will help in design of ligands targeting specific structural elements in a selective manner as well as provide novel incentives into the role of G-quadruplex forming regions within introns.

■ CONCLUSIONS

We demonstrate that a guanine-rich DNA oligonucleotide $d[TAG_3CG_3AG_3AG_3A_2]$ originating from the first intron of the N-myc gene in the presence of K^+, $^{15}NH_4^+$, and Na^+ ions folds into G-quadruplex structures with the folding highly influenced by the nature of the cations. Furthermore, in K^+ ions containing solution, an unexpected equilibrium between monomeric and dimeric G-quadruplexes was observed and characterized in terms of oligonucleotide and K^+ ion concentrations dependence. Despite different molecularity, the two forms exhibit similar folding rates and temperature-dependent properties altogether demonstrating interesting examples of G-quadruplex polymorphism. The presented work, with the focus on interconversion between monomeric and dimeric forms of N-myc, consists of detailed structural characterization providing new insights into the impact of loops on driving the folding process of G-rich sequences. The monomeric form exhibits three flexible single nucleotide loops and thus represents a missing element in structures of parallel G-quadruplexes comprising three G-quartets. Its structural features together with the high temperature stability suggest that G-quadruplex formation within the intron of the N-myc gene may play a specific biological role. The dimeric G-quadruplex adopted by N-myc comprises six consecutively stacked G-quartets in the core of the structure, a single nucleotide propeller-type loops,

and flipped out adenine residues enabling stacking of the third and the fourth G-quartets.

■ ASSOCIATED CONTENT

Supporting Information

Additional 1D ^{1}H- and ^{15}N-filtered ^{1}H spectra and 2D ^{15}N−^{1}H HMQC and JRHMBC NMR spectra of N-myc monomeric and dimeric G-quadruplexes, nondenaturating PAGE gel analysis, sedimentation velocity distributions, CD spectra, UV melting curves, and complete authorship for ref 50. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

janez.plavec@ki.si

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work was supported by the Slovenian Research Agency (ARRS) and the Ministry of Higher Education, Science and Technology of the Republic of Slovenia (Grants P1-0242 and J1-4020), COST MP0802, and EU FP7 projects with acronyms EAST-NMR (Grant 228461) and Bio-NMR (Grant 261863). We gratefully acknowledge the help of Professor Jonathan B. Chaires and Dr. William L. Dean, James Graham Brown Cancer Center, University of Louisville, Louisville, USA, for performing AUC studies.

■ REFERENCES

(1) *Quadruplex Nucleic Acids*; Neidle, S., Balasubramanian, S., Eds.; The Royal Society of Chemistry: Cambridge, U.K., 2006.
(2) Patel, D. J.; Phan, A. T.; Kuryavyi, V. *Nucleic Acids Res.* **2007**, *35*, 7429−7455.
(3) Neidle, S. *Curr. Opin. Struct. Biol.* **2009**, *19*, 239−250.
(4) Phan, A. T. *FEBS J.* **2010**, *277*, 1107−1117.
(5) Bauer, L.; Tluckova, K.; Tothova, P.; Viglasky, V. *Biochemistry* **2011**, *50*, 7484−7492.
(6) Balasubramanian, S.; Neidle, S. *Curr. Opin. Chem. Biol.* **2009**, *13*, 345−353.
(7) Cosconati, S.; Marinelli, L.; Trotta, R.; Virno, A.; De Tito, S.; Romagnoli, R.; Pagano, B.; Limongelli, V.; Giancola, C.; Baraldi, P. G.; Mayol, L.; Novellino, E.; Randazzo, A. *J. Am. Chem. Soc.* **2010**, *132*, 6425−6433.
(8) Huppert, J. L.; Balasubramanian, S. *Nucleic Acids Res.* **2005**, *33*, 2908−2916.
(9) Todd, A. K.; Johnston, M.; Neidle, S. *Nucleic Acids Res.* **2005**, *33*, 2901−2907.
(10) Siddiqui-Jain, A.; Grand, C. L.; Bearss, D. J.; Hurley, L. H. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 11593−11598.
(11) Phan, A. T.; Kuryavyi, V.; Gaw, H. Y.; Patel, D. J. *Nat. Chem. Biol.* **2005**, *1*, 167−173.
(12) Verma, A.; Yadav, V. K.; Basundra, R.; Kumar, A.; Chowdhury, S. *Nucleic Acids Res.* **2009**, *37*, 4194−4204.
(13) Gunaratnam, M.; Swank, S.; Haider, S. M.; Galesa, K.; Reszka, A. P.; Beltran, M.; Cuenca, F.; Fletcher, J. A.; Neidle, S. *J. Med. Chem.* **2009**, *52*, 3774−3783.
(14) Brooks, T. A.; Kendrick, S.; Hurley, L. *FEBS J.* **2010**, *277*, 3459−3469.
(15) Brosh, R. M.; Wu, Y. L. *FEBS J.* **2010**, *277*, 3470−3488.
(16) De Armond, R.; Wood, S.; Sun, D. Y.; Hurley, L. H.; Ebbinghaus, S. W. *Biochemistry* **2005**, *44*, 16341−16350.
(17) Cogoi, S.; Paramasivam, M.; Spolaore, B.; Xodo, L. E. *Nucleic Acids Res.* **2008**, *36*, 3765−3780.
(18) Palumbo, S. L.; Memmott, R. M.; Uribe, D. J.; Krotova-Khan, Y.; Hurley, L. H.; Ebbinghaus, S. W. *Nucleic Acids Res.* **2008**, *36*, 1755−1769.
(19) Mani, P.; Yadav, V. K.; Das, S. K.; Chowdhury, S. *PLoS One* **2009**, *4*, e4399.
(20) Lipps, H. J.; Rhodes, D. *Trends Cell Biol.* **2009**, *19*, 414−422.
(21) Barros, P.; Boan, F.; Blanco, M. G.; Gomez-Marquez, J. *FEBS J.* **2009**, *276*, 2983−2993.
(22) Sun, D. K.; Guo, K. X.; Shin, Y. J. *Nucleic Acids Res.* **2011**, *39*, 1256−1265.
(23) Phan, A. T.; Kuryavyi, V.; Darnell, J. C.; Serganov, A.; Majumdar, A.; Ilin, S.; Raslin, T.; Polonskaia, A.; Chen, C.; Clain, D.; Darnell, R. B.; Patel, D. J. *Nat. Struct. Mol. Biol.* **2011**, *18*, 796−804.
(24) Phan, A. T.; Modi, Y. S.; Patel, D. J. *J. Am. Chem. Soc.* **2004**, *126*, 8710−8716.
(25) Ambrus, A.; Chen, D.; Dai, J. X.; Jones, R. A.; Yang, D. Z. *Biochemistry* **2005**, *44*, 2048−2058.
(26) Phan, A. T.; Kuryavyi, V.; Burge, S.; Neidle, S.; Patel, D. J. *J. Am. Chem. Soc.* **2007**, *129*, 4386−4392.
(27) Kuryavyi, V.; Phan, A. T.; Patel, D. J. *Nucleic Acids Res.* **2010**, *38*, 6757−6773.
(28) Dai, J. X.; Chen, D.; Jones, R. A.; Hurley, L. H.; Yang, D. Z. *Nucleic Acids Res.* **2006**, *34*, 5133−5144.
(29) Tong, X.; Lan, W.; Zhang, X.; Wu, H.; Liu, M.; Cao, C. *Nucleic Acids Res.* **2011**, *39*, 6753−6763.
(30) Lim, K. W.; Lacroix, L.; Yue, D. J. E.; Lim, J. K. C.; Lim, J. M. W.; Phan, A. T. *J. Am. Chem. Soc.* **2010**, *132*, 12331−12342.
(31) Patel, D. J.; Kuryavyi, V. *Structure* **2010**, *18*, 73−82.
(32) Gomez, D.; Lemarteleur, T.; Lacroix, L.; Mailliet, P.; Mergny, J. L.; Riou, J. F. *Nucleic Acids Res.* **2004**, *32*, 371−379.
(33) Didiot, M. C.; Tian, Z. X.; Schaeffer, C.; Subramanian, M.; Mandel, J. L.; Moine, H. *Nucleic Acids Res.* **2008**, *36*, 4902−4912.
(34) Eddy, J.; Maizels, N. *Mol. Carcinog.* **2009**, *48*, 319−325.
(35) Broxson, C.; Beckett, J.; Tornaletti, S. *Biochemistry* **2011**, *50*, 4162−4172.
(36) Marcel, V.; Tran, P. L. T.; Sagne, C.; Martel-Planche, G.; Vaslin, L.; Teulade-Fichou, M. P.; Hall, J.; Mergny, J. L.; Hainaut, P.; Van Dyck, E. *Carcinogenesis* **2011**, *32*, 271−278.
(37) Meyer, N.; Penn, L. Z. *Nat. Rev. Cancer* **2008**, *8*, 976−990.
(38) Schwab, M. *Cancer Lett.* **2004**, *204*, 179−187.
(39) Brooks, T. A.; Hurley, L. H. *Nat. Rev. Cancer* **2009**, *9*, 849−861.
(40) Bell, E.; Chen, L. D.; Liu, T.; Marshall, G. M.; Lunec, J.; Tweddle, D. A. *Cancer Lett.* **2010**, *293*, 144−157.
(41) Corvi, R.; Savelyeva, L.; Schwab, M. *Cancer Res.* **1995**, *55*, 3471−3474.
(42) Sivak, L. E.; Pont-Kingdon, G.; Le, K.; Mayr, G.; Tai, K. F.; Stevens, B. T.; Carroll, W. L. *Mol. Cell. Biol.* **1999**, *19*, 155−163.
(43) Tang, X. X.; Zhao, H. Q.; Kung, B.; Kim, D. Y.; Hicks, S. L.; Cohn, S. L.; Cheung, N. K.; Seeger, R. C.; Evans, A. E.; Ikegaki, N. *Cancer Res.* **2006**, *66*, 2826−2833.
(44) Hazel, P.; Huppert, J.; Balasubramanian, S.; Neidle, S. *J. Am. Chem. Soc.* **2004**, *126*, 16405−16415.
(45) Rachwal, P. A.; Findlow, I. S.; Werner, J. M.; Brown, T.; Fox, K. R. *Nucleic Acids Res.* **2007**, *35*, 4214−4222.
(46) Bugaut, A.; Balasubramanian, S. *Biochemistry* **2008**, *47*, 689−697.
(47) Guedin, A.; De Cian, A.; Gros, J.; Lacroix, L.; Mergny, J. L. *Biochimie* **2008**, *90*, 686−696.
(48) Garbett, N. C.; Mekmaysy, C. S.; Chaires, J. B. In *G-Quadruplex DNA*; Baumann, P., Ed.; *Methods Mol. Biol.* 2010, *608*, 97−120.
(49) Lu, X. J.; Olson, W. K. *Nucleic Acids Res.* **2003**, *31*, 5108−5121.
(50) Case, D. A. et al. AMBER 9; University of California: San Francisco, 2006.
(51) Perez, A.; Marchan, I.; Svozil, D.; Sponer, J.; Cheatham, T. E.; Laughton, C. A.; Orozco, M. *Biophys. J.* **2007**, *92*, 3817−3829.
(52) Pettersen, E. F.; Goddard, T. D.; Huang, C. C.; Couch, G. S.; Greenblatt, D. M.; Meng, E. C.; Ferrin, T. E. *J. Comput. Chem.* **2004**, *25*, 1605−1612.

(53) Fernandes, M. X.; Ortega, A.; Martinez, M. C. L.; de la Torre, J. G. *Nucleic Acids Res.* **2002**, *30*, 1782−1788.
(54) Petraccone, L.; Spink, C.; Trent, J. O.; Garbett, N. C.; Mekmaysy, C. S.; Giancola, C.; Chaires, J. B. *J. Am. Chem. Soc.* **2011**, *133*, 20951−20961.
(55) Cogoi, S.; Xodo, L. E. *Nucleic Acids Res.* **2006**, *34*, 2536−2549.
(56) Parkinson, G. N.; Lee, M. P. H.; Neidle, S. *Nature* **2002**, *417*, 876−880.
(57) Podbevsek, P.; Hud, N. V.; Plavec, J. *Nucleic Acids Res.* **2007**, *35*, 2554−2563.
(58) Podbevsek, P.; Sket, P.; Plavec, J. *J. Am. Chem. Soc.* **2008**, *130*, 14287−14293.
(59) Sket, P.; Plavec, J. *J. Am. Chem. Soc.* **2010**, *132*, 12724−12732.
(60) Risitano, A.; Fox, K. R. *Nucleic Acids Res.* **2004**, *32*, 2598−2606.
(61) Olsen, C. M.; Lee, H. T.; Marky, L. A. *J. Phys. Chem. B* **2009**, *113*, 2587−2595.
(62) Guedin, A.; Alberti, P.; Mergny, J. L. *Nucleic Acids Res.* **2009**, *37*, 5559−5567.
(63) da Silva, M. W.; Trajkovski, M.; Sannohe, Y.; Hessari, N. M.; Sugiyama, H.; Plavec, J. *Angew. Chem., Int. Ed.* **2009**, *48*, 9167−9170.
(64) Guedin, A.; Gros, J.; Alberti, P.; Mergny, J. L. *Nucleic Acids Res.* **2010**, *38*, 7858−7868.
(65) Cang, X.; Sponer, J.; Cheatham, I. T. E. *J. Am. Chem. Soc.* **2011**, *133*, 14270−14279.
(66) Rachwal, P. A.; Brown, T.; Fox, K. R. *FEBS Lett.* **2007**, *581*, 1657−1660.