Published online 22 September 2006 *Nucleic Acids Research, 2006, Vol. 34, No. 18* ***5133–5144***
doi:10.1093/nar/gkl610

NMR solution structure of the major G-quadruplex structure formed in the human BCL2 promoter region

Jixun Dai[1], Ding Chen[1], Roger A. Jones[2], Laurence H. Hurley[1,3,4] and Danzhou Yang[1,3,4,*]

[1]College of Pharmacy, The University of Arizona, 1703 E. Mabel Street, Tucson, AZ 85721, USA, [2]Department of Chemistry and Chemical Biology, Rutgers University, 610 Taylor Road, Piscataway, NJ 08854, USA, [3]Arizona Cancer Center, 1515 N. Campbell Avenue, Tucson, AZ 85724, USA and [4]BIO5 Institute, The University of Arizona, 1140 E. South Campus Dr, Tucson, AZ 85721, USA

Received March 22, 2006; Revised and Accepted August 3, 2006

Protein Data Bank accession no. PDB ID 2F8U
BioMagResBank accession no. 6975

ABSTRACT

BCL2 protein functions as an inhibitor of cell apoptosis and has been found to be aberrantly expressed in a wide range of human diseases. A highly GC-rich region upstream of the P1 promoter plays an important role in the transcriptional regulation of BCL2. Here we report the NMR solution structure of the major intramolecular G-quadruplex formed on the G-rich strand of this region in K^+ solution. This well-defined mixed parallel/antiparallel-stranded G-quadruplex structure contains three G-tetrads of mixed G-arrangements, which are connected with two lateral loops and one side loop, and four grooves of different widths. The three loops interact with the core G-tetrads in a specific way that defines and stabilizes the overall G-quadruplex structure. The loop conformations are in accord with the experimental mutation and footprinting data. The first 3-nt loop adopts a lateral loop conformation and appears to determine the overall folding of the BCL2 G-quadruplex. The third 1-nt double-chain-reversal loop defines another example of a stable parallel-stranded structural motif using the G_3NG_3 sequence. Significantly, the distinct major BCL2 promoter G-quadruplex structure suggests that it can be specifically involved in gene modulation and can be an attractive target for pathway-specific drug design.

INTRODUCTION

The BCL2 (B-cell CLL/lymphoma 2) gene product is a mitochondrial membrane protein that plays an essential role in cell survival. The BCL2 protein exists in delicate balance with other apoptosis-related proteins and functions as an inhibitor of cell apoptosis (1,2). Deregulation of the *BCL2* gene is associated with aberrant cell growth in many human diseases. For example, BCL2 has been found to be aberrantly overexpressed in a wide range of human tumors, including B-cell and T-cell lymphomas (3,4) and breast (5), prostate (6), cervical (7), colorectal (8), and non-small cell lung carcinomas (9). Moreover, BCL2 overexpression has also been associated with poor prognosis and has been found to interfere with traditional cancer therapeutics (10,11). Inhibition of BCL2 expression by small molecules (12,13), peptidomimetics (14), or antisense oligonucleotides (15,16) has been shown to reduce cellular proliferation and to enhance chemotherapy efficacy. In contrast, aberrantly reduced BCL2 expression is associated with cardiovascular diseases and neurological disorders, such as ischemia/reperfusion injury of cardiac and renal tissues (17,18), multiple sclerosis (19), Alzheimer's and Parkinson's diseases (20), and tissue damage related with stroke (21) and spinal cord injuries (22). Thus, BCL2 has also emerged as an attractive target for neuroprotective and tissue-protective therapies.

There is accumulating *in vitro* evidence of G-quadruplex structures formed in promoter regions of several genes as transcriptional regulators, such as the transcriptional repressors reported in the human MYC (c-MYC) gene (23,24), the chicken β-globulin gene (25) and several muscle-specific genes (26,27), as well as a transcriptional enhancer reported in the human insulin gene (28,29). The human BCL2 gene has two promoters, P1 and P2. The major promoter, P1, located 1386–1423 bp upstream of the translation start site, is a TATA-less, GC-rich promoter containing multiple transcriptional start sites and positioned within a nuclease hypersensitive site (30,31). A highly GC-rich 39-bp region located 58–19 bp upstream of the P1 promoter has been implicated in playing a major role in the regulation of BCL2 transcription (Figure 1A) (32). Multiple transcription factors have been reported to bind to or regulate BCL2 gene expression through this region, including CREB (33), WT1 (34), Sp1 (30), E2F (35), NF-κB (36), and NGF (37). For example, CREB functions to activate BCL2 expression (33), while WT1 functions to repress BCL2 expression (34).

*To whom correspondence should be addressed: Tel: +1 520 626 5969; Fax: +1 520 626 6988; Email: yangd@pharmacy.arizona.edu

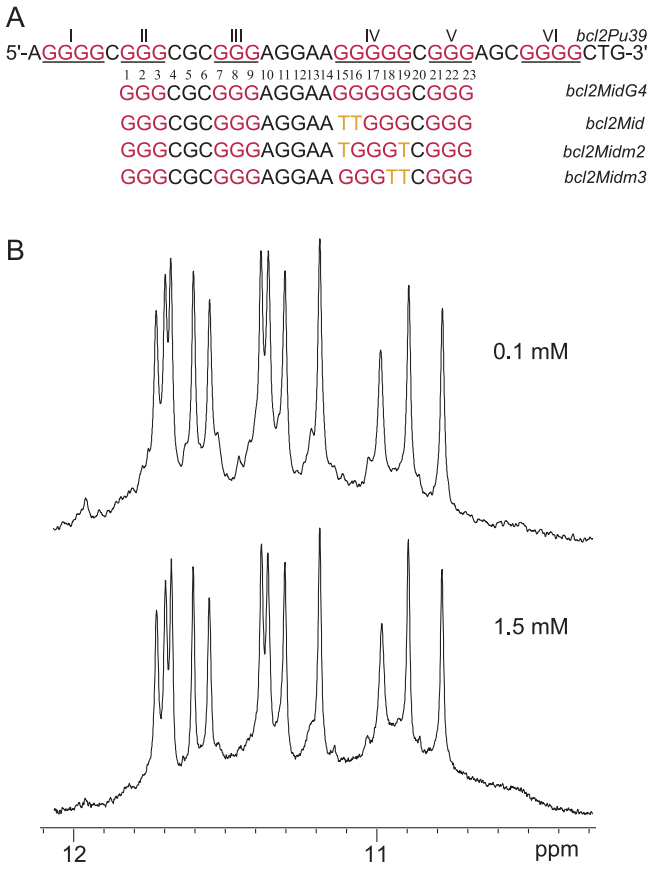

Figure 1. (**A**) The promoter sequence of the BCL2 gene and its modifications. Bcl2Pu39 is the wild-type BCL2 39mer sequence; bcl2MidG4 is the BCL2 23mer sequence containing the middle four consecutive G-runs, which forms the most stable G-quadruplex structure; bcl2Mid is the mutant 23mer with G-to-T mutations at positions 15 and 16, bcl2Midm2 is the mutant 23mer with G-to-T mutations at positions 15 and 19, and bcl2Midm3 is the mutant 23mer with G-to-T mutations at positions 18 and 19. The six G-runs are underlined and numbered using Roman numerals; the BCL2 23mer is numbered using Arabic numerals. (**B**) The imino regions of 1D ^{1}H NMR spectra of bcl2Mid samples at 0.1 mM (upper) and 1.5 mM (lower) strand concentrations. Conditions: 25°C, 20 mM K-phosphate, 40 mM KCl, pH 7.0.

Our previous results have shown that the 39mer guanine-rich strand (bcl2Pu39, Figure 1A) in the BCL2 promoter can form a mixture of three distinct intramolecular G-quadruplexes in K^+-containing solution, and that the G-quadruplex formed on the middle four consecutive runs of guanines is the most stable one (38,39), which adopts a novel folding of mixed parallel/antiparallel-stranded structure (39). In this paper we report the NMR solution structure for the predominant G-quadruplex structure formed in the BCL2 promoter region. The solution structure provides not only the molecular details of this G-quadruplex but also important insights for its loop conformations and interactions with the core tetrad structures.

MATERIALS AND METHODS

Sample preparation

The DNA oligonucleotides were synthesized using β-cyanoethylphosphoramidite solid-phase chemistry on an Expedite™ 8909 Nucleic Acid Synthesis System (Applied Biosystem, Inc.) in DMT-on mode, and were purified using C18 reverse-phase HPLC chromatography, as described previously (39–41). Deprotection was carried out using 80% AcOH for 1 h, followed by ether extraction. DMT-off DNA was further purified by HPLC followed by successive dialysis against 150 mM NaCl and H_2O. 6% 1,2,7-^{15}N, 2-^{13}C-labeled guanine phosphoramidite (42) was used for site-specific labeled DNA synthesis. Samples in D_2O were prepared by repeated lyophilization and final dissolution in 99.96% D_2O. Samples in water were prepared in 10%/90% D_2O/H_2O solution. The final NMR samples contained 0.2–2 mM DNA oligonucleotides in 20 mM K-phosphate buffer (pH 7.0) and 40 mM KCl.

NMR experiments

NMR experiments were performed on a Bruker DRX-600 spectrometer as described before (39–41,43). At the determined melting point (61°C), stoichiometric titration of the melted and the folded strands as a function of total strand concentration from 0.01 to 0.1 mM was performed (43). Assignment of guanine imino and H8 protons were obtained by 1D ^{15}N-filtered experiments using site-specific labeling. Standard 2D NMR experiments, including NOESY, TOCSY, and DQF-COSY, were used at 1, 7, 15, 20, 25, 30 and 35°C to obtain complete proton resonance assignment. The NMR experiments for samples in water solution were performed with Watergate or Jump-and-Return water suppression techniques. Relaxation delays were set to 3 s. The acquisition data points were set to 2048 × (350−512) (complex points). The 45° or 60° shifted sine-squared functions were applied to NOESY and TOCSY spectra. The fifth-order polynomial functions were employed for the baseline corrections. The final spectral sizes are 2048 × 1024. Peak assignments and integrations were achieved using the software Sparky (UCSF). Only relatively isolated peaks were used for NOE-restrained structure calculation. Severely overlapping peaks were discarded. The NOE peaks were integrated using the peak fitting function and volume integration of Sparky. We have manually checked every peak to make sure the fitted lineshape agrees with the experimental data. The average linewidth of NMR peaks is 10–15 Hz, and the digitized resolution of the NOESY spectra is sufficient for the line-fitting and accurate volume integration in Sparky. Distances between non-exchangeable protons were estimated based on the NOE cross-peak volumes at 50–300 ms mixing times, with the upper and lower boundaries assigned to ±20% of the estimated distances. Distances between exchangeable protons were assigned with looser boundaries (±1–1.5 Å). The cytosine base proton H5–H6 distance (2.45 Å) was used as a reference. The distances involving the unresolved protons, e.g. methyl protons, were assigned using pseudo-atom notation to make use of the pseudo-atom correction automatically computed by X-PLOR.

The ^{31}P NMR spectra were collected on a DNA sample at 1.5 mM in D_2O (20 mM potassium-phosphate buffer, 40 mM KCl, pH 7.0) at 25°C and were referenced to an external standard of 85% H_3PO_4, including the 1D proton-decoupled phosphorus spectrum, and 2D heteronuclear ^{31}P–^{1}H Correlation Spectroscopy (COSY) and Heteronuclear Single Quantum

Correlation Spectroscopy (HSQC). A series of ^{31}P–^{1}H HSQC spectra were collected at spectral widths of 10 p.p.m. (^{1}H) × 5 p.p.m. (^{31}P) with 2048 × 128 complex points, using various heteronuclear INEPT transfer delays corresponding to J-couplings of 5, 10, 15, 20 and 25 Hz. The non-selective and H3′-selective ^{31}P–^{1}H COSY experiments (44) were carried out in States-TPPI mode using the same spectral width with 2048 × 128 complex points and 256 scans. Assignments of the individual ^{31}P resonance were accomplished by a combination of 2D ^{1}H/^{1}H NOESY, COSY, TOCSY and heteronuclear ^{31}P–^{1}H COSY.

Distance geometry and simulated annealing (DGSA) calculations

Metric matrix distance geometry (MMDG) calculations were carried out using X-PLOR (45) to embed and optimize 100 initial structures. An arbitrary extended conformation was first generated for the single-stranded bcl2Mid sequence. Substructure embedding was performed to produce a family of 100 DG structures. The embedded DG structures were then subjected to simulated annealing regularization. The experimentally obtained distance restraints and G-tetrad hydrogen-bonding distance restraints were included during the calculation. All distance restraints were specified with the SUM averaging option in X-PLOR (45). After simulated annealing, 97% of the molecules were folded in the correct topology, whereas 3% of the molecules were misfolded, e.g. in left-handed folding topology.

NOE-distance restrained molecular dynamics calculations

All of the 100 molecules obtained from the DGSA calculations were subjected to NOE-restrained Simulated Annealing refinement in XPLOR (45) with a distance-dependent dielectric constant. Atoms participating in hydrogen bonds in the G-tetrad planes were restrained with distances corresponding to ideal hydrogen bond geometry. Each individual hydrogen bond was restrained using two distance restraints (heavy atom–heavy atom and heavy atom–proton). Hydrogen-bonding distance restraints were also applied to A10:T15 bp, with larger distance bounds (±0.4 Å). All experimentally obtained distance restraints were specified ambiguously with the sum-averaging option. The force constants were scaled at 30 and 100 kcal mol^{-1} Å^{-2} for NOE and hydrogen bond distance restraints, respectively. A total of 476 NOE-distance restraints, of which 168 are from inter-residue NOE interactions, were incorporated into the NOE-restrained structure calculation.

Dihedral angle restraints were used to restrict the glycosidic torsion angle (χ) for the experimentally assigned *syn* configuration, i.e. G1, G7, G8, G17 and G21, tetrad-guanines [60(±30)°], and A13 and A14 in the 3′-loop [60(±60)°], as well as for some of the experimentally assigned *anti* configuration bases, i.e. C4, G5 and C6 in the 5′-loop [220(±40)°]. Dihedral angle restraints were also used to restrain the sugar backbone torsion angles β, γ and ε (46). Based on the J-coupling constants of ^{31}P(n)–H5′/H5″(n) and H3′ (n − 1)–^{31}P(n) obtained from ^{31}P–^{1}H COSY and HSQC experiments with various J-couplings, the β angles were restrained to the $\beta-$t conformation at 180(±90)° for all the residues, except for G21 whose β angle was restrained to −60(±20)°. The ε angles were restrained to 80(±20)° for G19, and to 90(±40)° for T15 and C6. Based on the relative intensities of H3′–H5′/H5″ and H4′–H5′/H5″, the γ angles of the majority of residues with resolved H5′/H5″ are in the regular $\gamma+$ conformation (∼60°) or sometimes in the $\gamma-$ conformation (∼−60°), because for each residue, the H3′–H5′ (or H3′–H5″) NOE is clearly stronger than the H3′–H5″ (or H3′–H5′) NOE, except for A14 and T15 which show similar intensities for H3′–H5′ and H3′–H5″, and thus fall in the $\gamma-$t region (46). Thus only the γ angles of A14 and T15 were restrained to 170(±90)°. The force constants of dihedral angle restraints were 10 kcal mol^{-1} rad^{-2} for χ and 5 kcal mol^{-1} rad^{-2} for β, γ and ε.

NOE-restrained simulated annealing refinement calculations were initiated at 300 K. The temperature was gradually increased to 1000 K in 4 ps. The system was equilibrated at 1000 K for 20 ps, and was then slowly cooled to 300 K in 10 ps. The 20 best molecules were selected based both on their minimal energy terms and number of NOE violations, and were further subjected to NOE-restrained molecular dynamics calculations at 300 K for 25 ps. The coordinates saved every 0.1 ps during the last 2.0 ps of NOE-restrained molecular dynamics calculations were averaged, and the resulting averaged-structure was subjected to minimization until the energy gradient of 0.1 kcal·mol^{-1} was achieved. A soft planarity restraint of 1 kcal mol^{-1} Å^{-2} was imposed on the tetrads before the heating process and was removed at the beginning of the equilibration stage. The time steps for all processes of heating, cooling, and equilibration were equal to 0.5 fs. The 10 best molecules were selected based both on their minimal energy terms and number of NOE violations and have been deposited in the Protein Data Bank (accession no. 2F8U).

RESULTS AND DISCUSSION

Basis for selection of the bcl2Mid sequence with its dual G-to-T mutations as the predominant G-quadruplex in the BCL2 promoter

In a recently published study (38), we have demonstrated that the six runs of G-tracts containing three or more contiguous guanines in the wild-type sequence (bcl2Pu39, Figure 1A) can form three overlapping G-quadruplex structures, i.e., 5′G4, midG4, and 3′G4 (see Supplementary Table S1). Of these three overlapping sequences, midG4, or bcl2MidG4 (Figure 1A), is by far the most stable. Since this sequence contains a run of five guanines, it is possible to form three loop isomers in which the three guanines involved in the G-tetrads are either at the 3′ or the 5′ end or in the middle of the five continuous guanines (Supplementary Table S1 and Figure 1A bottom three sequences). The three possible loop isomers can be isolated by dual G-to-T mutations as shown in Figure 1A. The bcl2MidG4 sequence gives rise to an unambiguous DMS cleavage pattern in which there is a clear preference for the loop isomer as formed in bcl2Mid (Figure 1A). The bcl2Mid sequence also forms the most stable G-quadruplex structure, as demonstrated by NMR (Supplementary Figure S1). This sequence gave a well-defined spectrum that has many overlapping ^{1}H NMR resonance signals with the wild-type sequence (Supplementary Figure S1)

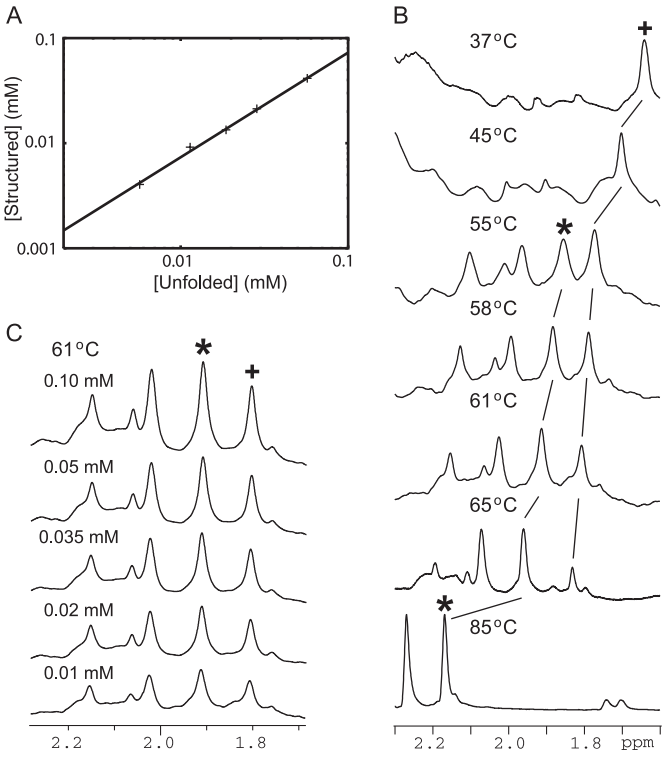

Figure 2. (**A**) Determination of stoichiometry by NMR titration for bclMid in K^+ solution. The slope of the fitted line is 0.99, meaning that the quadruplex structure existing in solution is unimolecular. (**B**) The variable-temperature study of bcl2Mid by NMR. The peak intensities of two resolved peaks at 61°C (the one belonging to the melted form of bcl2Mid is labeled with asterisk and the one belonging to the folded forms is labeled with cross) were used for the calculation. (**C**) The extended region of 1D ^{1}H NMR spectra of bcl2Mid at various concentrations showing the two peaks from the folded and the melted forms. Conditions: 20 mM K-phosphate, 40 mM KCl, pH 7.0.

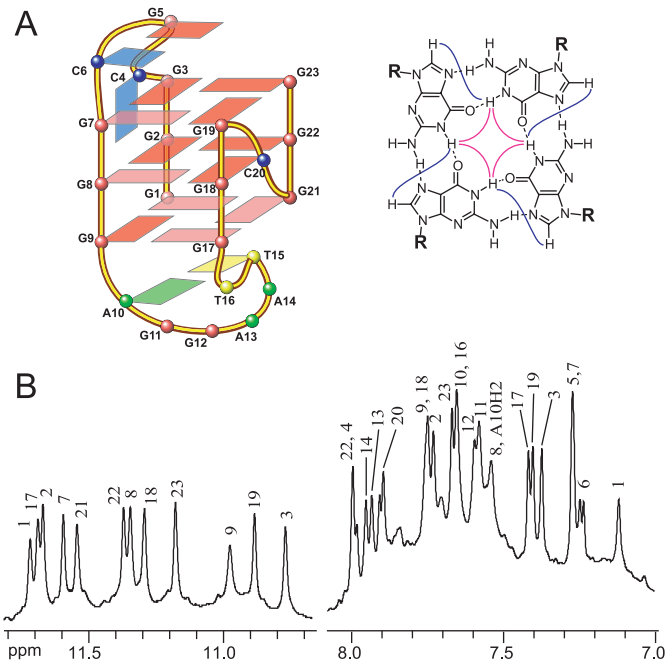

Figure 3. (**A**) (Left) Schematic drawing of the folding topology of bcl2Mid. Red, guanine (*anti*); light red, guanine (*syn*); green, adenine; yellow, thymine; blue, cytosine. (Right) A G-tetrad with H1–H1 and H1–H8 connectivity pattern detectable in NOESY experiments. (**B**) Imino and aromatic regions of the 1D ^{1}H NMR spectrum of bcl2Mid. The imino and aromatic protons are assigned over the resonances. Conditions: 25°C, 20 mM K-phosphate, 40 mM KCl, pH 7.0, 1.5 mM DNA.

(39). In addition this same mutated sequence was previously shown to be the most stable using polymerase stop assay (see figure 6 in Ref. 38). Therefore the bcl2Mid sequence with the dual (G-to-T) mutant at the 5′-end of bcl2MidG4 (positions 15 and 16, Figure 1A) was chosen as the sequence for NMR structure determination. The reason why the other possible contiguous runs of three guanines in Figure 1A are unstable is that these mutational sequences eliminate the G_3NG_3 single-nt double-chain-reversal loop which provides stability to the quadruplex. Thus it is clear that the dual 5′-(G-to-T) mutation selected can give rise to a specific and stable fold containing this G_3NG_3 single-nt double-chain-reversal motif. Importantly, the two mutated guanines (G15 and G16, Figure 1A) give rise to the predominant DMS cleavage pattern found in the wild-type sequence (38). Furthermore, the bcl2Mid sequence gives rise to a CD spectrum very similar to that of the wild-type sequence (39). We are therefore confident that the G-quadruplex from this sequence is not only the predominant one found in a single-stranded BCL2 promoter DNA template, but its dual G-to-T mutations force the sequence into the predominant loop isomer. This is an important point because mutants that change the folding pattern or result in isolation of the least stable species are less likely to be biologically relevant.

bcl2Mid forms a monomeric G-quadruplex structure

The bcl2Mid molecule exhibits a stronger propensity to aggregate than other G-quadruplex forming sequences we have worked with. Minor conformations are also present as indicated by the presence of weak resonances (Figure 1B), whose intensities are <5% of the major species and thus do not interfere with the unambiguous structural analysis of the predominant BCL2 G-quadruplex structure. A sample of bcl2Mid in potassium solution with a concentration over 3 mM, as has been routinely used for other G-quadruplex structures (41,43,47,48), shows a markedly increased background after a week and readily aggregates to a gel form after annealing from 95°C. We therefore used a lower concentration of 1.5 mM of bcl2Mid in potassium for our NMR structure determination (Figure 1B, lower panel). The sample of 1.5 mM bcl2Mid is stable in potassium solution for over 1 month which is sufficient for multiple 2D experiments. The two 1D ^{1}H spectra of the 1.5 mM bcl2Mid NMR sample, freshly made or in NMR solution for 5 weeks, are shown in Supplementary Figure S2A&B. No noticeable differences in 1D NMR spectra were observed for this sample after 5 weeks in solution. However, after 5 months in NMR solution, this sample gave rise to an observable higher background (Supplementary Figure S2C). After over 1 year in NMR solution, a significantly increased level of broad resonance background in 1D NMR spectra is observed for this sample, even though the sharp peaks from the major species are still clearly observable (Supplementary Figure S2D).

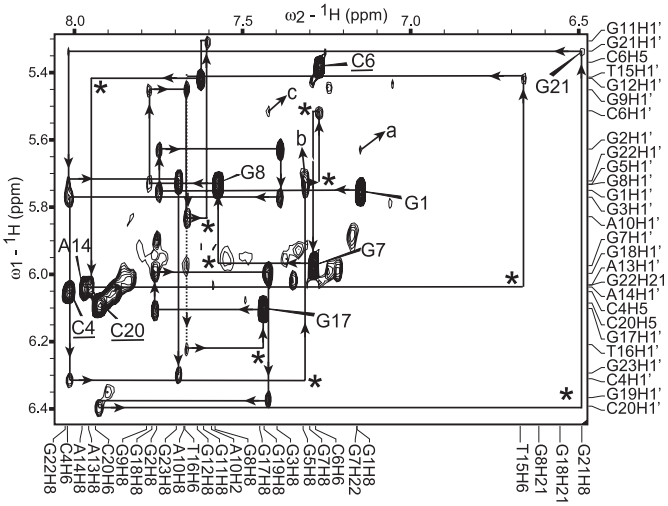

Figure 4. The expanded H8/H6–H1′ region with assignments of the non-exchangeable 2D-NOESY spectrum of bcl2Mid.The sequential assignment pathway is shown. Missing connectivities are labeled with asterisks. The H8–H1′ NOEs of the nucleotides with *syn* configuration are labeled by residue names, while the H5–H6 NOEs of cytosines are also labeled for reference. The H8–H1′ NOE crosspeak of G21 has a large integration value; however, the H8 and H1′ of G21 are much broader than those of other guanines and give rise to a much broader H8–H1′ NOE crosspeak. The characteristic G(*i*)H8/G(*i*+1)H1′ NOEs for the *syn* G(*i*)s are labeled as *a–b* for G1–G2 and G7–G8. The NOE between C6H1′ and G19H8 is labeled as *c*. Conditions: 25°C, 20 mM K-phosphate, 40 mM KCl, pH 7.0, 1.5 mM DNA.

1D proton spectra of bcl2Mid at various temperatures (VTs) indicate that the melting temperature for the quadruplex structure in 60 mM K^+ is around 65°C (Supplementary Figure S3). The melting process of the G-quadruplex structure of bcl2Mid appears to be rather homogeneous throughout the G-tetrad core structure. As described in our previous report (39), bcl2Mid forms a monomeric intramolecular G-quadruplex structure, as demonstrated by the independence of the melting temperature from the concentration, the sharp NMR spectral line widths (Figure 1B), and the EMSA (electrophoretic mobility shift assay) results. The formation of a unimolecular structure in bcl2Mid is unambiguously confirmed by the NMR stoichiometry titration experiment at the melting temperature using DNA concentrations from 0.01 to 0.1 mM (43,49) (Figure 2). The NMR spectra of bcl2Mid at a strand concentration of 1.5 mM (for 2D NMR experiments and structure determination) and 0.1 mM (for 1D experiments and the titration experiment) are the same (Figure 1B, upper and lower panels), indicating that the same unimolecular G-quadruplex structure is formed in all the conditions used for our NMR analysis. In addition, DOSY (Diffusion Ordered Spectroscopy) experiments on bcl2Mid also confirmed that the sharp NMR signals are from a molecular weight equivalent to a monomeric structure. (The DOSY results will be published elsewhere.)

Proton resonance assignments of bcl2Mid

The major G-quadruplex formed on the central BCL2 promoter sequence has been shown to adopt a mixed parallel/antiparallel-stranded folding (Figure 3A) (39). The imino NH1 and base aromatic H8 protons of guanine residues of this bcl2Mid were unambiguously assigned by the site-specific low-concentration (6%) incorporation of 1, 2, 7-^{15}N, 2-^{13}C-labeled guanine nucleoside at each guanine position of the sequence, one base at a time (Figure 3B and Supplementary Figure S4) (39). The base H6 proton resonances of thymines and cytosines were unambiguously assigned by substituting deoxyuridine (dU) for dT/dC one at a time at each thymine/cytosine position of the sequence. Multiple 2D experiments, including 2D-NOESY, TOCSY and COSY, were carried out for bcl2Mid at a strand concentration of 1.5 mM in 20 mM pH 7.0 K-phosphate and 40 mM KCl at various temperatures. Standard DNA sequential assignment procedure was utilized for the assignment of the proton resonances of bcl2Mid (Figure 4). The assignment of the aromatic protons allowed the direct assignment of H1′ and H2′/H2″ resonances, which was then extended to other regions, including those of H3′, H4′ and H5′/H5″. The proton chemical shifts at 25°C are listed in Table 1. All proton resonances have been unambiguously assigned, except some H5′/H5″ protons which cannot be differentiated. The ambiguity of H5′/H5″ protons should not affect the NMR structure calculation, since the NOE intensities associated with H5′ or H5″ do not vary much regardless whether a resonance is assigned as H5′ or H5″. In addition, very few NOE intensities associated with H5′ or H5″ were used (see below). For all the residues with resolved H2′ and H2″ resonance, the H1′–H3′ NOE is weaker than both the H1′–H2′ and H1′–H2″ NOEs, while the H1′–H2″ NOE is stronger than the H1′–H2′ NOE, indicating a C2′-endo sugar pucker conformation.

Phosphorus resonance assignments of bcl2Mid

We have carried out ^{31}P experiments on bcl2Mid, including proton-decoupled 1D phosphorus spectroscopy, as well as 2D heteronuclear ^{31}P–^{1}H COSY and ^{31}P–^{1}H HSQC. The 1D ^{31}P spectrum at 25°C in D_2O is shown in Figure 5A, referenced to the phosphoric acid standard. While the majority of the ^{31}P resonances of bcl2Mid are clustered around -1 ± 0.3 p.p.m., which has been shown to be the chemical shifts for ^{31}P resonances of regular right-handed B-DNAs, there are a number of ^{31}P resonances that are shifted outside this region. Using 2D ^{31}P–^{1}H COSY (Figure 5B) and the proton assignments, we were able to assign all of the ^{31}P resonances of bcl2Mid (Table 1). Every phosphorus resonance was assigned by using the assignments of sugar protons H3′, H4′, and H5′/H5″, and correlating it to its 5′-coupled H3′ proton and 3′-coupled H4′ and H5′/H5″protons. Moreover, the phosphorus assignments confirmed the proton assignments of H3′($n-1$), H4′(n), and H5′/H5″(n) for each step. Very interestingly, the phosphorus resonances that are observed out of the -1 ± 0.3 p.p.m. region are all from the residues in the loop regions of unique conformations, including the downfield-shifted (from −1 p.p.m.) G21P (at C20–G21 step, △P 1.5 p.p.m.), T15 (△P 0.88 p.p.m.), C4 (△P 0.54 p.p.m.), and C6 (△P 0.43 p.p.m.), and the upfield-shifted G19 (△P −0.93 p.p.m.), C20 (△P −1.12 p.p.m.), and T16 (△P −1.61 p.p.m.). In particular, those phosphorus resonances associated with the unique double-chain-reversal single-nt loop (G19, C20, G21, see Figure 3A) and with the unique conformation at T15–T16 (Figure 3A) are the most shifted.

Table 1. Proton chemical shifts for the bcl2Mid at 25°C[a]

	H6/H8	NH21/NH22[b] H2/H5/Me	H1′	H2′,H2″	H3′	H4′	H5′,H5″	NH1/NH3	^{31}P[c]	ΔP[d]
G1[e]	7.15	5.87/9.20	5.75	2.69/3.00	4.86	4.23	3.82	11.74		
G2	7.75	4.96/9.49	5.63	2.37/2.58	4.95	4.28	4.22/4.16	11.68	−1.24	−0.24
G3	7.39	5.48/9.11	5.78	2.43/2.54	4.95	4.38	4.12/4.26	10.79	−0.86	0.14
C4	8.02	6.06	6.32	2.07/2.52	4.74	4.33	4.18/4.07		−0.46	**0.54**
G5	7.31		5.74	1.85/2.13	4.49	3.74	3.84/3.91		−1.31	−0.31
C6	7.27	5.37	5.52	1.54/2.23	4.46	3.77	2.57/3.31		−0.57	**0.43**
G7	7.29	—/6.93	5.98	3.69/3.06	4.78	4.21	4.03/4.32	11.61	−1.29	−0.29
G8	7.57	6.24/9.62	5.74	2.63/2.54	4.92	4.18	4.09/4.23	11.36	−1.26	−0.26
G9	7.78	5.48/9.16	5.46	2.34/2.46	4.89	4.24	4.06	10.98	−1.41	−0.41
A10	7.66	7.58	5.83	1.66/1.94	4.66	4.24	3.98/4.06		−1.23	−0.23
G11	7.61		5.31	2.32/2.20	4.71	3.97	3.80/3.85		−1.31	−0.31
G12	7.62		5.42	2.36/2.25	4.67	3.66	3.54		−1.07	−0.07
A13	7.95		6.04	2.47/2.36	4.72	3.92	3.64/3.73		−0.97	0.03
A14	7.97		6.04	2.11/2.44	4.57	4.12	3.84/3.92		−0.79	0.21
T15	6.66	1.15	5.42	1.43/2.19	4.61	4.13	3.62/4.02	10.29[b]	−0.12	**0.88**
T16	7.66	1.53	6.21	2.02/2.45	4.62	3.74	3.05		−2.61	**−1.61**
G17	7.44	6.19/9.21	6.11	3.37/2.96	4.81	4.46	4.13/4.41	11.69	−1.23	−0.23
G18	7.76	6.35/9.22	6.00	2.41/2.82	4.92	4.44	4.16/4.31	11.30	−1.10	−0.10
G19	7.43	6.12/8.46	6.37	2.64	5.04	4.57	4.29/4.34	10.89	−1.93	**−0.93**
C20	7.93	6.09	6.40	2.26/2.76	4.98	4.77	4.27		−2.12	**−1.12**
G21	6.48	5.71/9.20	5.34	2.85	5.00	4.79	3.91/4.30	11.55	0.52	**1.52**
G22	8.02	5.89/9.38	5.73	2.53/2.61	4.98	4.33	3.98/4.15	11.38	−0.60	0.40
G23	7.69	6.09/8.66	6.30	2.44/2.34	4.58	4.28	4.13/4.18	11.18	−0.78	0.22

[a]The chemical shifts are measured in 20 mM K-phosphate, 40 mM KCl, pH 7.0 referenced to DSS.
[b]These chemical shifts are measured at 1°C.
[c]The ^{31}P chemical shifts are referenced to H_3PO_4.
[d]The difference from −1 p.p.m.
[e]Guanines in *syn* glycosidic conformation are in bold.

The J-coupling constants of ^{31}P(*n*)–H5′/H5″(*n*) and H3′(*n*)–^{31}P(*n* + 1) were obtained from ^{31}P–^{1}H COSY and confirmed by HSQC experiments with various J-couplings. The ^{31}P–H5′ and ^{31}P–H5″ coupling constants can define the β torsion angle quite well (46). Except for G21, all other observed ^{31}P(*n*)–H5′/H5″(*n*) coupling constants are below 10 Hz, indicating the β torsion angles fall in the common *trans*-region, 180(±90)°. G21 gives rise to a ^{31}P–H5′ coupling constant of ∼12 Hz and no observable ^{31}P–H5″ coupling, indicating a β torsion angle in the β− region, −60(±20)°. Most H3′-^{31}P couplings are <10 Hz, indicating that the ε angles fall in the common *trans*-/*gauche*- conformation (170°–300°), at which range the ε angles can not be accurately determined by the H3′(*n*)–^{31}P(*n* + 1) coupling constants due to several possible ε angles for one H3′-^{31}P coupling in this region (46). However, when the H3′(*n*)–^{31}P(*n* + 1) coupling constants are larger than 10 Hz, it is possible to more accurately determine the ε angles. The H3′(*n*)–^{31}P(*n* + 1) coupling constants for G19(−C20P), T15(−T16P) and C6(−G7P) are 15, 12 and 12 Hz, respectively, indicative of ε angles in the *gauche*+/*trans*+ region (centered ∼80–90°C). It is again interesting to note that the residues exhibiting unusual large coupling constants are all associated with the loop regions, especially with the single-nt double-chain-reversal loop. It may be noted that for non-B-DNA structures, the ^{31}P chemical shifts show a very poor correlation with the ε angle values (50).

NOE interactions within G-tetrads regions in the bcl2Mid G-quadruplex

Many inter-residue NOE crosspeaks are observed in 2D-NOESY and are summarized in Figure 6. Only one NOE intensity is associated with H5′ or H5″. An expanded NOESY spectrum of base and sugar H1′ protons is shown in Figure 4. Five guanine residues of the core G-tetrads are in the *syn*-configuration, including G1, G7, G8, G17 and G21 (Figure 3A), as indicated by the very strong H8–H1′ NOE intensities (Figure 4). A characteristic downfield shift is observed for the H2′/H2″ sugar protons of the *syn*-guanines (Table 1). The sequential NOE crosspeak connectivities of the base H8 protons to the 5′-flanking residue sugar H1′/H2′/H2″ protons, typical for right-handed DNA twist, are observed for bcl2Mid (Figure 4). Bcl2Mid adopts a mixed parallel/antiparallel-stranded G-quadruplex structure, with the first, third and fourth G-strands being parallel with each other, and the second G-strand being antiparallel with the rest of the strands (Figure 3A) (39). The top two G-tetrads have the same *anti*/*syn* distribution and are connected by guanines with the same sugar configurations, while the third G-tetrad has a reversed *anti*/*syn* distribution and is connected with the middle G-tetrad by guanines with different sugar configurations (Figure 3A). In accord with the topology, the sequential NOE connectivities are indeed either missing or very weak at the N(*i*)–*syn*G(*i* + 1) steps, i.e. C6–G7, G7–G8, T16–G17 and C20–G21. The characteristic G(*i*)H8/G(*i* + 1)H1′ NOEs are observed when G(*i*) is in the *syn* configuration, such as G1–G2, G7–G8, G8–G9 and G17–G18 (Figure 4). The same configurations of guanines (*syn*–*syn* or *anti*–*anti* steps) connecting the top two G-tetrads are reflected by clear NOE interactions of adjacent guanine base protons, e.g. G2H1/G3H1, G7H1/G8H1, G18H1/G19H1 and G22H1/G23H1 (Supplementary Figure S5). The reversed configurations of guanines connecting the bottom two G-tetrads are clearly reflected by strong inter-tetrad NOE interactions of guanine imino protons, such as G2H1/G9H1,

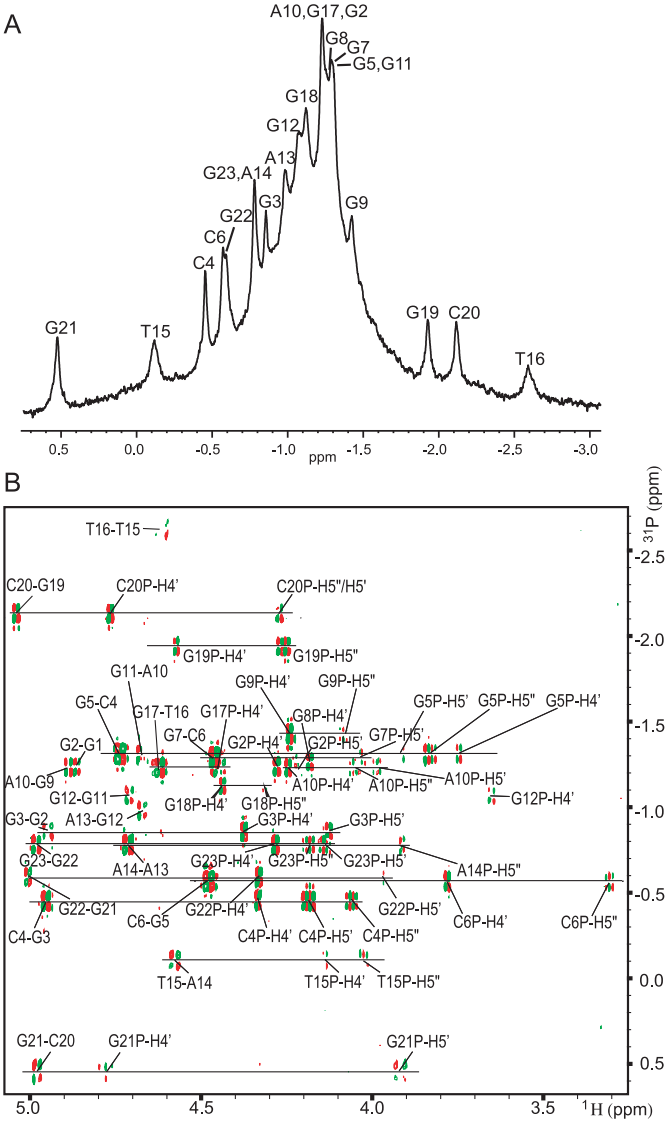

Figure 5. (**A**) 1D proton-decoupled ^{31}P NMR spectrum of bcl2Mid with phosphorus assignments. (**B**) 2D heteronuclear ^{31}P–^{1}H Correlation Spectroscopy (COSY) of bcl2Mid with peak assignments. The NnP-N(n-1)H3′ crosspeaks are labeled as Nn-N(n-1). The spectra were referenced to external H_3PO_4. Conditions: 25°C, 20 mM K-phosphate, 40 mM KCl, pH 7.0, 1.5 mM DNA.

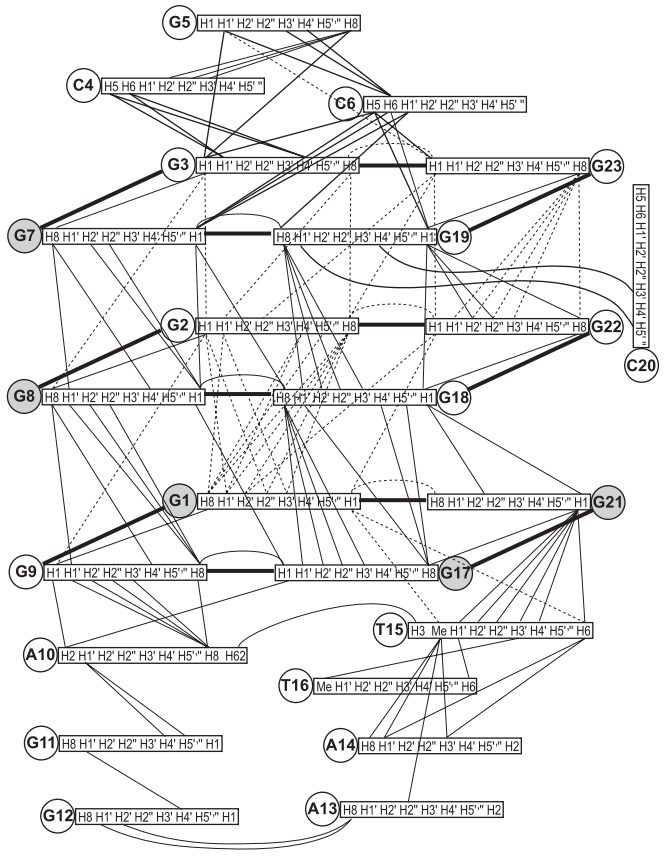

Figure 6. Schematic diagram of inter-residue NOE connectivities of bcl2Mid. The guanines in *syn* configuration are represented using gray circles. The NOE connectivities clearly define the quadruplex conformation and provide distance restraints for structure calculation.

G8H1/G17H1, G18H1/G21H1 and G22H1/G1H1 (Supplementary Figure S5). Furthermore, the right-handedness of the DNA backbone of the G-quadruplex is clearly indicated by inter-tetrad NOE interactions, including G3H1/G8H8, G7H1/G18H8, G19H1/G22H8 and G23H1/G2H8 (Figure 7).

NOE interactions within loop regions in the bcl2Mid G-quadruplex

The sequential NOE connectivities are interrupted at the loop regions, including C4–G5–C6, A10–G11–G12–A13–A14–T15–T16, and G19–C20–G21 (Figure 4), indicating a poorly stacked loop conformation. All residues in the loop regions are in the *anti* conformation except A13 and A14, as indicated by the strong H8–H1′ NOE intensities (Figure 4). A number of interesting NOE interactions are observed for both the lateral loops, whereas the sequential NOE connectivities are almost all missing for the double-chain-reversal single-nucleotide C loop, except a couple of NOEs between the sugar protons of G19 and C20 (Figure 6). For the 5′-CGC lateral loop, sequential connectivities are clearly observable within G5–C6–G7. Interestingly, clear NOEs are observed between G5/C6 and the imino protons of the top tetrad, such as G7H1/C6H6, H5&H1′, G19H1/C6H5&H6, G3H1/G5H8&H1′, and G23H1/G5H1′ (Figure 4, complete summary in Figure 6), indicating that the G5 and C6 residues are stacking over the top G-tetrad. On the other hand, unusual NOEs are observed within the G3–C4–G5 region. Strong NOEs are observed for C4 base protons H5&H6 with G3 sugar protons H1′&H4′, e.g. C4H5/G3H4′ and C4H6/G3H1′ (strong), C4H5/G3H1′ and C4H6/G3H4′ (medium strong), but not with G3H3′ or H2′. No NOEs are observed between C4H6/H5 and G3H8 for base stacking. For the C4–G5 step, clear NOEs are observed for G5H8 with C4H3′ (strong), H2′ (strong) and H2″ (medium strong), but not with C4H1′ or H4′. These NOE data suggest an unusual structure adopted within this region, where the C4 base is not stacked over the top G-tetrad but is positioned into the groove that is close to the H1′/H4′ side of the G3 sugar. This will be discussed in more detail later.

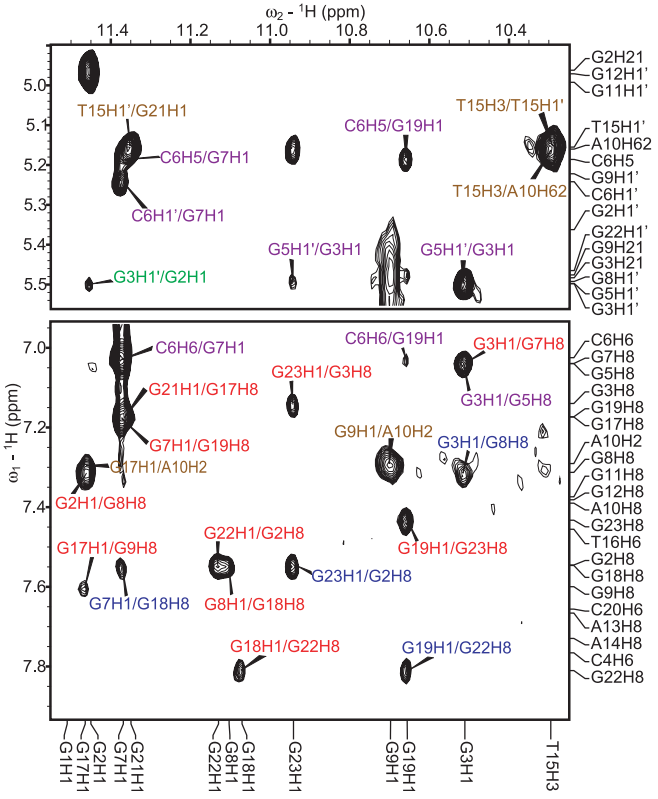

Figure 7. The expanded H1–H8/H6 and H1–H1′ regions with assignments of the exchangeable proton 2D-NOESY spectrum of bcl2Mid. For NOEs involved in G-tetrads: intra-tetrad NOEs are labeled in red, sequential NOEs are labeled in green, and inter-tetrad NOEs are labeled in blue. NOEs involved in the first C4–G5–C6 lateral loop region are labeled in purple, and NOEs involved in the second A10–G11–G12–A13–A14–T15–T16 lateral loop region are labeled in brown. Conditions: 1°C, 20 mM K-phosphate, 40 mM KCl, pH 7.0, 1.5 mM DNA.

For the middle 7-nt lateral loop, A10–G11–G12–A13–A14–T15–T16, A10 is very well stacked with G9 and the bottom G-tetrad, as is evident by the clear NOE interactions, such as A10H2/G9H1&G17H1 (Figures 6 and 7). Very interestingly, clear NOEs are observed between T15, instead of T16, and the bottom G-tetrad, such as T15H6/G1H1&G21H1, and T15H1′/G21H1 (Figures 6 and 7), indicating that T15, but not T16, is stacked right on the bottom G-tetrad. Furthermore, a clear imino proton is observed for T15H3, which exhibits a very strong NOE interaction with A10H62 (Figure 7), indicating that T15 is likely to be involved in a stable H-bonded conformation with A10.

Mutational analysis for loop segments

We have carried out systematic analysis of BCL2 promoter sequences with mutated residues in the loop regions to determine their functional role in BCL2 G-quadruplex formation and stability. We incorporated one mutation at a time for each loop residue, using various substitutions as listed in Table 2, and then collected NMR spectra on each sample. Mutations that do not induce changes in the NMR spectra are marked with a '–' symbol, whereas those that do induce clear spectral changes are marked with a '+' symbol.

Table 2. Effect of single base substitutions on the G-quadruplex formation of the Bcl2Mid sequence

	C	T	U	A	I
C4		+	+		
G5		–		–	
C6			–		
A10	+	+			+
G11		–		–	
G12		–		–	
A13		–			
A14		–			–
T15			–	–[a]	
T16	–		–	–[a]	
C20			–	–	

–: The G-quadruplex formation is not affected; +: the G-quadruplex formation is destabilized.
[a]A stable G-quadruplex with some minor changes in proton chemical shifts.

Table 3. Structural statistics for the Bcl2Mid[a]

NMR distance and dihedral constraints	
Distance restraints	
Total NOE	476
Intraresidue	309
Interresidue	167
Sequential ($\|i-j\|=1$)	107
Non-sequential ($\|i-j\|>1$)	60
Hydrogen bonds	26
Total dihedral angle restraints	38
Structure statistics	
Violations (mean and s.d.)	
Distance constraints (Å)	0.03 ± 0.003
Dihedral angle constraints (°)	0.81 ± 0.63
Max. dihedral angle violation(°)	10.6
Max. distance constraint violation (Å)	0.37
Deviations from idealized geometry	
Bond length (Å)	0.007 ± 0.0001
Bond angle (°)	1.40 ± 0.01
Impropers (°)	1.05 ± 0.01
Average pairwise r.m.s.d. of heavy atoms (Å)	
G-tetrads	1.01 ± 0.12
With C20	1.14 ± 0.14
With C4,G5,C6	1.04 ± 0.12
With A10,T15	1.13 ± 0.14
All residues	2.55 ± 0.65

[a]The ensemble of 10 structures is selected based both on the minimal energy terms and number of NOE violations.

Interestingly, the most sensitive positions are C4 and A10, which cannot tolerate any substitutions.

NOE-restrained structure calculation

NOE data (Figure 6) define the overall topology of this mixed antiparallel/parallel G-quadruplex and were used for NOE-restrained structure calculation of the major BCL2 G-quadruplex formed with bcl2Mid. Solution structures of this G-quadruplex were calculated using a NOE-restrained distance geometry (DGSA) and molecular dynamics (RMD) approach, starting from an arbitrary extended single-stranded DNA model. A total of 476 NOE distance restraints, of which 168 are from inter-residue NOE interactions, were

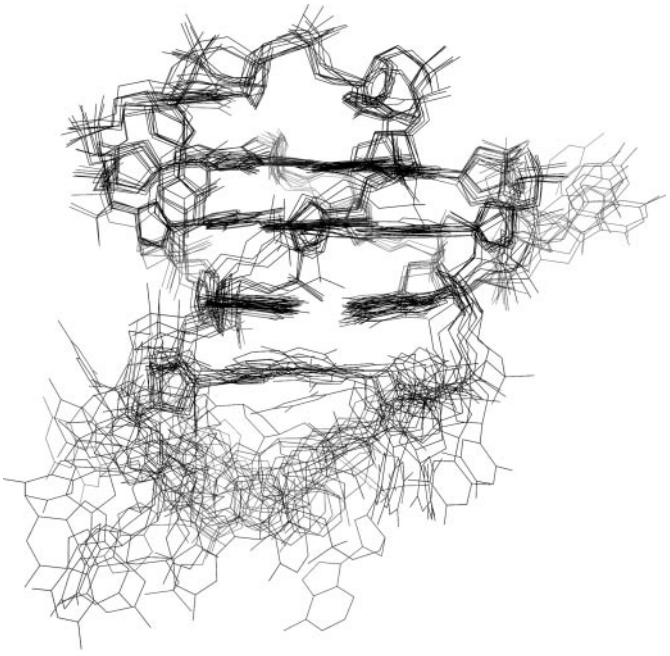

Figure 8. The superimposed 10 lowest energy structures of the bcl2Mid G-quadruplex by NOE-restrained structure refinement.

incorporated into the NOE-restrained structure calculation (Table 3). Dihedral angle restraints were used for the glycosidic torsion angle (χ) for all the *syn* residues, and for the three *anti* residues, i.e. C4, G5 and C6, in the 5′-loop. Dihedral angle restraints were also used to restrain the sugar backbone torsion angles β, γ, and ε (Materials and Methods). The superimposition of the 10 lowest energy structures produced by the refinement is shown in Figure 8 (PDB ID 2F8U). The structure statistics are listed in Table 3. An average of ~23 restraints per residue was used for the solution structure calculation, including experimentally observed H-bond interactions for the G-tetrads, whereas no planarity constraints were used for RMD calculations. Remarkably, out of the 10 lowest energy structures, the RMS deviation for distance violations is only 0.03 Å (Table 3). The bcl2Mid G-quadruplex structure is very well defined, with a RMSD of 1.01 Å for the three G-tetrads. The 5′-CGC lateral loop is also well defined, as the RMSD of the three G-tetrads and the 5′-CGC loop is 1.04 Å. The RMSDs are 1.14 Å and 1.13 Å, respectively, when including the C20 single-nt double-chain-reversal loop or the A10/T15 of the middle 7-nt lateral loop. The RMSD for all residues is 2.55 Å, indicating a much less-defined middle 7-nt loop.

Molecular structure of the major G-quadruplex structure in the BCL2 promoter

A representative model of the bcl2Mid G-quadruplex structure is shown in different views in Figure 9 and Supplementary Figure S6. The G-quadruplex consists of three G-tetrads linked with mixed parallel/antiparallel right-handed G-strands that are connected by the first two lateral loops (CGC and AGGAATT) and a third single-nucleotide (nt) double-chain-reversal side loop (C), as also shown in our previous report (39). The first, third and fourth G-strands of this G-quadruplex are parallel with each other, while the second G-strand is antiparallel with the rest of the G-strands. The bcl2Mid G-quadruplex contains one wide groove (groove I, between the first and second antiparallel G-strands), one narrow groove (groove II, between the second and third antiparallel G-strands), and two intermediate grooves (groove III and IV, between the third, fourth, and first parallel G-strands) (Supplementary Figure S6). All the δ torsion angels fall in

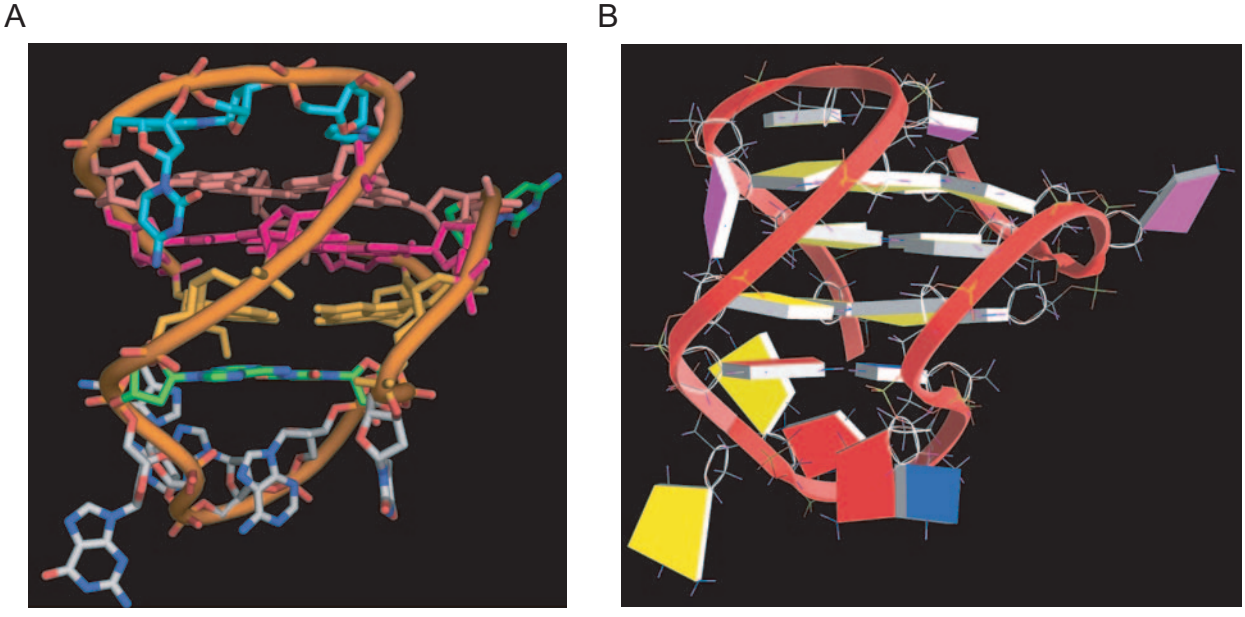

Figure 9. A representative model of the NMR-refined bcl2Mid G-quadruplex structure from two different views. (**A**) is prepared using PyMOL. (**B**) is prepared using GRASP (57) (guanine, yellow; adenine, red; thymine, blue; and cytosine, magenta).

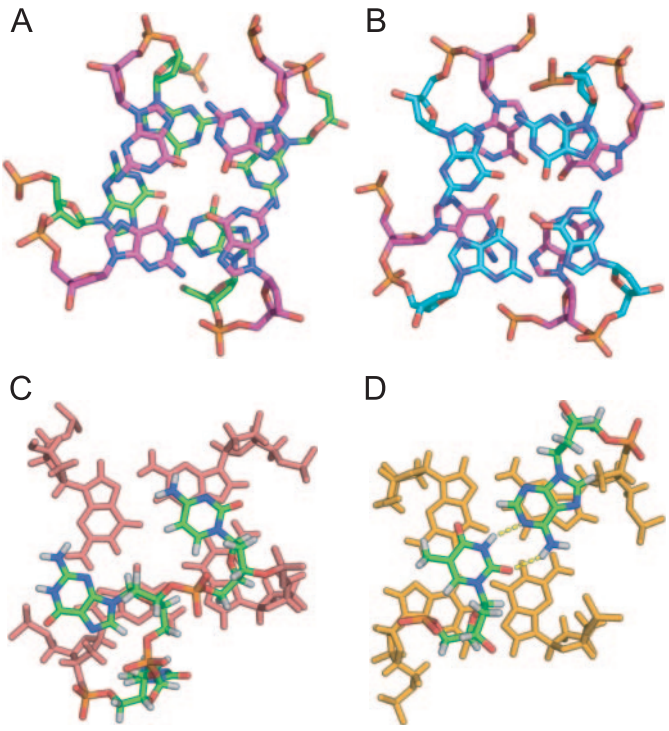

Figure 10. Stacking interactions between (**A**) the middle (magenta) and bottom (green) G-tetrads, which have reversed *anti:syn:anti:anti* and *syn:anti:syn:syn* arrangements, and (**B**) the top (cyan) and middle (magenta) G-tetrads, which have the same *anti:syn:anti:anti* arrangements; and stacking interactions between (**C**) the first lateral loop C4–G5–C6 (cyan) and the top G-tetrad (pink), and (**D**) the bottom G-tetrad (orange) and the A10:T15 bp (green) from the second lateral loop, A10–G11–G12–A13–A14–T15–T16. Figures are prepared using PyMOL.

the range of 110–150°C, consistent with the C2′−sugar pucker conformations indicated by the NMR data.

The three G-tetrads are well defined, with the top two G-tetrads having the same arrangement of guanine configuration (*anti:syn:anti:anti*) and the bottom G-tetrad having the reversed arrangement (*syn:anti:syn:syn*). Extensive stacking between the guanine five-membered rings is observed for adjacent guanines of the bottom two G-tetrads with the alternate guanine glycosidic configurations (Figure 10A), while only partial stacking is observed for adjacent guanines of the top two G-tetrads with the same guanine glycosidic configurations (Figure 10B).

Loop conformation and functional role of loop residues

The NMR solution structures indicate that the three loop regions interact with the core G-tetrads in a specific way that defines and stabilizes the unique BCL2 G-quadruplex structure. The first C4–G5–C6 linker that connects the first and second antiparallel G-strands forms a unique lateral loop conformation which is very well defined (Figures 9A and 10C). G5 and C6 stack right on the top G-tetrad, where G5 stacks very well with G3 and C6 is positioned above the center of G7 and G19 (Figure 10C). Interestingly, the C4 residue adopts a unique conformation in which its base is positioned in the wide groove I and is perpendicular to the G-tetrad planes (Figure 9A), covering the outside of the G3 residue of the top tetrad. This unique conformation of C4 appears to be rather favored, as over 90% of the 100 structures from our DGSA calculation are in this conformation. The conformation of the C4–G5–C6 loop is well defined as indicated by the small root mean square deviation of this loop region (Figure 8 and Table 3) (see above). The unique conformation of C4 explains the unusual NOE interactions observed between C4 and G3/G5 as discussed before, e.g. strong NOEs between the C4 H5/H6 and G3 H1′/H4′, and the lack of NOE between the G5H8 and the C4 H1′ (Figure 6).

The second 7-nt A10–G11–G12–A13–A14–T15–T16 linker that connects the second and third antiparallel G-strands also forms a lateral loop conformation in the NMR structure (Figure 9). A10 stacks very well with G9 and the bottom tetrad, while T15, which is the second residue from the other end of the loop, stacks well with the bottom G-tetrad. Potential reversed Watson–Crick hydrogen bonds could be formed between A10 and T15 (Figure 10D), as the imino proton H3 of T15 can be detected at 1°C (Figure 7). T16, which is sequentially adjacent to G17, however, is looped out from the bottom G-tetrad (Figure 9). The remaining four residues of this long lateral loop are not very well defined (Figure 8) and are mostly exposed to the solvent (Figure 9).

Significantly, bcl2Mid contains a G_3NG_3 sequence motif, which adopts a single-nucleotide (C20) double-chain-reversal loop conformation connecting two parallel G-strands (Figure 3A), representing another example of a very stable parallel-stranded structural motif, in addition to the one first observed in the MYC promoter sequence (41,51,52). The conformation of this single-nucleotide double-chain-reversal loop (C20) is very similar to those of the MYC G-quadruplex (41). The right-handed DNA backbone twist brings the ends of the two adjacent parallel G-strands very close to each other spatially so that the single-nucleotide double-chain-reversal loop conformation is rather favored (Figure 9B and Supplementary Figure S6).

The loop conformations are defined by the experimental NOEs and agree well with the experimental data (Figure 6 and Table 2). For the first 3-nt CGC lateral loop, C4, which is positioned in groove I, cannot even tolerate a substitution of uridine, which is the closest nucleotide to cytosine, indicating that the amino group of cytosine is clearly more favored for the groove positioning, as reported previously for DNA groove-binding ligands (53,54). On the other hand, G5 and C6, which appear to be able to tolerate various substitutions, can presumably still retain the stacking interactions with different residues. For the second 7-nt AGGAATT lateral loop, in accord with the NMR structure in which A10 forms a well-defined H-bonded conformation that stacks extensively with G9 of the bottom tetrad, the A10 residue cannot tolerate any substitution in our mutational analysis. Interestingly, T15 and T16, which are the two mutated residues, display much higher tolerance and can be substituted with uridines and adenosines, indicating a greater conformational flexibility of these two positions. The rest of the loop residues display much higher flexibility and can tolerate most substitutions. Indeed, G11 and G12, which adopt an extended loop conformation, show enhanced cleavage in the DMS footprinting study (38). Finally, the C20 single-nucleotide double-chain-reversal loop appears to be able to tolerate various substitutions.

The BCL2 G-quadruplexes and their relationship to transcriptional control and other G-quadruplex-containing promoter elements

The G-quadruplexes in the BCL2 promoter element represent a more complex architecture than those found in other promoters, such as the c-Myc NHE III_1. In this contribution we have characterized the major loop isomer from the bcl2MidG4 sequence that overlaps with two other G-quadruplexes (5′G4 and 3′G4). In principle there is a total of 15 different possible loop and structural isomers in the BCL2 promoter element. Since this element (bcl2Pu39, Figure 1A) also overlaps with a binding site for WT1 transcriptional factor, which is a suppressor protein, it may be that formation of one or more of the G-quadruplexes results in transcriptional activation. In addition to the possible transcriptional activation role of the G-quadruplexes in this region, it is also possible that the emergence or elimination of individual G-quadruplexes in this region may differentially regulate gene expression by selectively interacting with transcriptional factors that either activate or suppress BCL2 gene expression.

A common feature of the most stable BCL-2 loop isomer studied and the c-Myc promoter G-quadruplex (41,51,52), and likely the VEGF and HIF-1α promoter G-quadruplexes (55,56), is the single-nt G_3NG_3 double-chain-reversal loop. Only this isomer of the bcl2MidG4Pu23 retains this feature that is essential for stability of these G-quadruplexes. The single less stable G_3–TTA–G_3 double-chain-reversal loop found in the human telomeric G-quadruplex (43) may be important for the inherent more facile reversibility of the G-quadruplex to alternative folding structures.

Last, the diversity in sequences and resulting folding patterns found within the BCL2 promoter gives rise to differential drug binding (38), which may mimic protein recognition differences. From a drug design perspective, this diversity in drug binding provides opportunities for selectivity. Furthermore, the major G-quadruplexes formed in BCL2 and MYC promoter sequences are different (41), with distinct folding patterns, G-tetrad conformations, loop conformations, and groove conformations, suggesting that such regions can be differentially targeted by G-quadruplex-interactive agents. The structural diversity of the G-quadruplexes formed in oncogene promoter regions makes such regions attractive targets for pathway-specific drug design.

SUPPLEMENTARY DATA

Supplementary Data is available at NAR Online.

ACKNOWLEDGEMENTS

This research was supported by the National Institutes of Health (1K01CA83886 and 1S10 RR16659) and Arizona Biomedical Research Commission (05-002A). We thank Dr Megan Carver for technical help and proofreading the manuscript. Funding to pay the Open Access publication charges for this article has been waivered by Oxford University Press.

Conflict of interest statement. None declared.

REFERENCES

1. Adams,J.M. and Cory,S. (1998) The Bcl-2 protein family: arbiters of cell survival. *Science*, **281**, 1322–1326.
2. Chao,D.T. and Korsmeyer,S.J. (1998) BCL-2 family: regulators of cell death. *Annu. Rev. Immunol.*, **16**, 395–419.
3. Cleary,M.L., Smith,S.D. and Sklar,J. (1986) Cloning and structural analysis of cDNAs for bcl-2 and a hybrid bcl-2/immunoglobulin transcript resulting from the t(14;18) translocation. *Cell*, **47**, 19–28.
4. Akagi,T., Kondo,E. and Yoshino,T. (1994) Expression of Bcl-2 protein and Bcl-2 mRNA in normal and neoplastic lymphoid tissues. *Leuk. Lymphoma*, **13**, 81–87.
5. Joensuu,H., Pylkkanen,L. and Toikkanen,S. (1994) Bcl-2 protein expression and long-term survival in breast cancer. *Am. J. Pathol.*, **145**, 1191–1198.
6. McDonnell,T.J., Troncoso,P., Brisbay,S.M., Logothetis,C., Chung,L.W., Hsieh,J.T., Tu,S.M. and Campbell,M.L. (1992) Expression of the protooncogene bcl-2 in the prostate and its association with emergence of androgen-independent prostate cancer. *Cancer Res.*, **52**, 6940–6944.
7. Tjalma,W., De Cuyper,E., Weyler,J., Van Marck,E., De Pooter,C., Albertyn,G. and van Dam,P. (1998) Expression of bcl-2 in invasive and *in situ* carcinoma of the uterine cervix. *Am. J. Obstet. Gynecol.*, **178**, 113–117.
8. Baretton,G.B., Diebold,J., Christoforis,G., Vogt,M., Muller,C., Dopfer,K., Schneiderbanger,K., Schmidt,M. and Lohrs,U. (1996) Apoptosis and immunohistochemical bcl-2 expression in colorectal adenomas and carcinomas. Aspects of carcinogenesis and prognostic significance. *Cancer*, **77**, 255–264.
9. Pezzella,F., Turley,H., Kuzu,I., Tungekar,M.F., Dunnill,M.S., Pierce,C.B., Harris,A., Gatter,K.C. and Mason,D.Y. (1993) Bcl-2 protein in non-small-cell lung carcinoma. [see comment]. *N. Engl. J. Med.*, **329**, 690–694.
10. Desoize,B. (1994) Anticancer drug resistance and inhibition of apoptosis. *Anticancer Res.*, **14**, 2291–2294.
11. Reed,J.C., Kitada,S., Takayama,S. and Miyashita,T. (1994) Regulation of chemoresistance by the bcl-2 oncoprotein in non-Hodgkin's lymphoma and lymphocytic leukemia cell lines. *Ann. Oncol.*, **5** (Suppl. 1), 61–65.
12. Oltersdorf,T., Elmore,S.W., Shoemaker,A.R., Armstrong,R.C., Augeri,D.J., Belli,B.A., Bruncko,M., Deckwerth,T.L., Dinges,J., Hajduk,P.J. *et al.* (2005) An inhibitor of Bcl-2 family proteins induces regression of solid tumours. *Nature*, **435**, 677–681.
13. Enyedy,I.J., Ling,Y., Nacro,K., Tomita,Y., Wu,X., Cao,Y., Guo,R., Li,B., Zhu,X., Huang,Y. *et al.* (2001) Discovery of small-molecule inhibitors of Bcl-2 through structure-based computer screening. *J. Med. Chem.*, **44**, 4313–4324.
14. Tzung,S.P., Kim,K.M., Basanez,G., Giedt,C.D., Simon,J., Zimmerberg,J., Zhang,K.Y. and Hockenbery,D.M. (2001) Antimycin A mimics a cell-death-inducing Bcl-2 homology domain 3. *Nature Cell. Biol.*, **3**, 183–191.
15. Marshall,J., Chen,H., Yang,D., Figueira,M., Bouker,K.B., Ling,Y., Lippman,M., Frankel,S.R. and Hayes,D.F. (2004) A phase I trial of a Bcl-2 antisense (G3139) and weekly docetaxel in patients with advanced breast cancer and other solid tumors. *Ann. Oncol.*, **15**, 1274–1283.
16. Klasa,R.J., Gillum,A.M., Klem,R.E. and Frankel,S.R. (2002) Oblimersen Bcl-2 antisense: facilitating apoptosis in anticancer treatment. *Antisense Nucleic Acid Drug Dev.*, **12**, 193–213.
17. Yang,J., Marden,J.J., Fan,C.G., Sanlioglu,S., Weiss,R.M., Ritchie,T.C., Davisson,R.L. and Engelhardt,J.F. (2003) Genetic redox preconditioning differentially modulates AP-1 and NF-κB responses following cardiac ischemia/reperfusion injury and protects against necrosis and apoptosis. *Mol. Ther.*, **7**, 341–353.
18. Gobe,G., Zhang,X.J., Willgoss,D.A., Schoch,E., Hogg,N.A. and Endre,Z.H. (2000) Relationship between expression of Bcl-2 genes and growth factors in ischemic acute renal failure in the rat. *J. Am. Soc. Nephrol.*, **11**, 454–467.
19. Tews,D.S. (2002) Apoptosis and muscle fibre loss in neuromuscular disorders. *Neuromusc. Disord.*, **12**, 613–622.
20. Hetts,S.W. (1998) To die or not to die—an overview of apoptosis and its role in disease. *J. Am. Med. Assoc.*, **279**, 300–307.
21. Yenari,M.A., Zhao,H., Giffard,R.G., Sobel,R.A., Sapolsky,R.M. and Steinberg,G.K. (2003) Gene therapy and hypothermia for stroke treatment. *Ann. N. Y. Acad. Sci.*, **993**, 54–68.

22. Seki,T., Hida,K., Tada,M. and Iwasaki,Y. (2003) Role of the bcl-2 gene after contusive spinal cord injury in mice. *Neurosurgery*, **53**, 192–198.
23. Siddiqui-Jain,A., Grand,C.L., Bearss,D.J. and Hurley,L.H. (2002) Direct evidence for a G-quadruplex in a promoter region and its targeting with a small molecule to repress c-MYC transcription. *Proc. Natl Acad. Sci. USA*, **99**, 11593–11598.
24. Simonsson,T., Pecinka,P. and Kubista,M. (1998) DNA tetraplex formation in the control region of c-myc. *Nucleic Acids Res.*, **26**, 1167–1172.
25. Howell,R.M., Woodford,K.J., Weitzmann,M.N. and Usdin,K. (1996) The chicken beta-globin gene promoter forms a novel 'cinched' tetrahelical structure. *J. Biol. Chem.*, **271**, 5208–5214.
26. Etzioni,S., Yafe,A., Khateb,S., Weisman-Shomer,P., Bengal,E. and Fry,M. (2005) Homodimeric MyoD preferentially binds tetraplex structures of regulatory sequences of muscle-specific genes. *J. Biol. Chem.*, **280**, 26805–26812.
27. Yafe,A., Etzioni,S., Weisman-Shomer,P. and Fry,M. (2005) Formation and properties of hairpin and tetraplex structures of guanine-rich regulatory sequences of muscle-specific genes. *Nucleic Acids Res.*, **33**, 2887–2900.
28. Lew,A., Rutter,W.J. and Kennedy,G.C. (2000) Unusual DNA structure of the diabetes susceptibility locus IDDM2 and its effect on transcription by the insulin promoter factor Pur-1/MAZ. *Proc. Natl Acad. Sci. USA*, **97**, 12508–12512.
29. Hammondkosack,M.C.U., Dobrinski,B., Lurz,R., Docherty,K. and Kilpatrick,M.W. (1992) The human insulin gene linked polymorphic region exhibits an altered DNA-structure. *Nucleic Acids Res.*, **20**, 231–236.
30. Seto,M., Jaeger,U., Hockett,R.D., Graninger,W., Bennett,S., Goldman,P. and Korsmeyer,S.J. (1988) Alternative promoters and exons, somatic mutation and deregulation of the Bcl-2-Ig fusion gene in lymphoma. *EMBO J.*, **7**, 123–131.
31. Tsujimoto,Y. and Croce,C.M. (1986) Analysis of the structure, transcripts, and protein products of bcl-2, the gene involved in human follicular lymphoma. *Proc. Natl Acad. Sci. USA*, **83**, 5214–5218.
32. Young,R.L. and Korsmeyer,S.J. (1993) A negative regulatory element in the bcl-2 5′-untranslated region inhibits expression from an upstream promoter. *Mol. Cell. Biol.*, **13**, 3686–3697.
33. Ji,L., Mochon,E., Arcinas,M. and Boxer,L.M. (1996) CREB proteins function as positive regulators of the translocated bcl-2 allele in t(14;18) lymphomas. *J. Biol. Chem.*, **271**, 22687–22691.
34. Heckman,C., Mochon,E., Arcinas,M. and Boxer,L.M. (1997) The WT1 protein is a negative regulator of the normal bcl-2 allele in t(14;18) lymphomas. *J. Biol. Chem.*, **272**, 19609–19614.
35. Gomez-Manzano,C., Mitlianga,P., Fueyo,J., Lee,H.Y., Hu,M., Spurgers,K.B., Glass,T.L., Koul,D., Liu,T.J., McDonnell,T.J. *et al.* (2001) Transfer of E2F-1 to human glioma cells results in transcriptional up-regulation of Bcl-2. *Cancer Res.*, **61**, 6693–6697.
36. Heckman,C.A., Mehew,J.W. and Boxer,L.M. (2002) NF-κB activates Bcl-2 expression in t(14;18) lymphoma cells. *Oncogene*, **21**, 3898–3908.
37. Liu,Y.Z., Boxer,L.M. and Latchman,D.S. (1999) Activation of the Bcl-2 promoter by nerve growth factor is mediated by the p42/p44 MAPK cascade. *Nucleic Acids Res.*, **27**, 2086–2090.
38. Dexheimer,T.S., Sun,D. and Hurley,L.H. (2006) Deconvoluting the structural and drug-recognition complexity of the G-quadruplex-forming region upstream of the bcl-2 P1 promoter. *J. Am. Chem. Soc.*, **128**, 5404–5415.
39. Dai,J., Dexheimer,T.S., Chen,D., Carver,M., Ambrus,A., Jones,R.A. and Yang,D. (2006) An intramolecular G-quadruplex structure with mixed parallel/antiparallel G-strands formed in the human Bcl-2 promoter region in solution. *J. Am. Chem. Soc.*, **128**, 1096–1098.
40. Dai,J.X., Punchihewa,C., Mistry,P., Ooi,A.T. and Yang,D.Z. (2004) Novel DNA bis-intercalation by MLN944, a potent clinical bisphenazine anticancer drug. *J. Biol. Chem.*, **279**, 46096–46103.
41. Ambrus,A., Chen,D., Dai,J., Jones,R.A. and Yang,D. (2005) Solution structure of the biologically relevant G-quadruplex element in the human c-MYC promoter. Implications for G-quadruplex stabilization. *Biochemistry*, **44**, 2048–2058.
42. Zhao,H., Pagano,A.R., Wang,W.M., Shallop,A., Gaffney,B.L. and Jones,R.A. (1997) Use of a C-13 atom to differentiate two N-15-labeled nucleosides. Syntheses of [(NH2)-N-15]-adenosine, [1,NH2-N-15(2)]- and [2-C-13-1,NH2-N-15(2)]-guanosine, and [1,7,NH2-N-15(3)]- and [2-C-13-1,7,NH2-N-15(3)]-2′-deoxyguanosine. *J. Org. Chem.*, **62**, 7832–7835.
43. Ambrus,A., Chen,D., Dai,J., Bialis,T., Jones,R.A. and Yang,D. (2006) Human telomeric sequence forms a hybrid-type intramolecular G-quadruplex structure with mixed parallel/antiparallel strands in potassium solution. *Nucleic Acids Res.*, **34**, 2723–2735.
44. Sklenar,V. and Bax,A. (1987) Measurement of H-1-P-31 NMR coupling-constants in double-stranded DNA fragments. *J. Am. Chem. Soc.*, **109**, 7525–7526.
45. Brünger,A.T. (1993) *X-PLOR Version 3.1: A System for X-ray Crystallography and NMR. Book X-PLOR Version 3.1: A System for X-ray Crystallography and NMR*. Yale University Press, New Haven, CT.
46. Wijmenga,S.S. and van Buuren,B.N.M. (1998) The use of NMR methods for conformational studies of nucleic acids. *Prog. Nucl. Magn. Reson. Spectrosc.*, **32**, 287–387.
47. Wang,Y. and Patel,D.J. (1993) Solution structure of the human telomeric repeat D[Ag(3)(T(2)Ag(3))3] G-tetraplex. *Structure*, **1**, 263–282.
48. Zhang,N., Phan,A.T. and Patel,D.J. (2005) (3+1) assembly of three human telomeric repeats into an asymmetric dimeric G-quadruplex. *J. Am. Chem. Soc.*, **127**, 17277–17285.
49. Phan,A.T., Gueron,M. and Leroy,J.L. (2001) *Investigation of Unusual DNA Motifs. Nuclear Magnetic Resonance of Biologica Macromolecules*. Pt A. Academic Press Inc., San DiegoVol. 338, pp. 341–371.
50. Canalia,M. and Leroy,J.L. (2005) Structure, internal motions and association-dissociation kinetics of the i-motif dimer of d(5mCCTCACTCC). *Nucleic Acids Res.*, **33**, 5471–5481.
51. Seenisamy,J., Rezler,E.M., Powell,T.J., Tye,D., Gokhale,V., Joshi,C.S., Siddiqui-Jain,A. and Hurley,L.H. (2004) The dynamic character of the G-quadruplex element in the c-MYC promoter and modification by TMPyP4. *J. Am. Chem. Soc.*, **126**, 8702–8709.
52. Phan,A.T., Modi,Y.S. and Patel,D.J. (2004) Propeller-type parallel-stranded G-quadruplexes in the human c-myc promoter. *J. Am. Chem. Soc.*, **126**, 8710–8716.
53. Bailly,C. and Chaires,J.B. (1998) Sequence-specific DNA minor groove binders. Design and synthesis of netropsin and distamycin analogues. *Bioconjug. Chem.*, **9**, 513–538.
54. Dervan,P.B. (2001) Molecular recognition of DNA by small molecules. *Bioorg. Med. Chem.*, **9**, 2215–2235.
55. Sun,D., Guo,K., Rusche,J.J. and Hurley,L.H. (2005) Facilitation of a structural transition in the polypurine/polypyrimidine tract within the proximal promoter region of the human VEGF gene by the presence of potassium and G-quadruplex-interactive agents. *Nucleic Acids Res.*, **33**, 6070–6080.
56. De Armond,R., Wood,S., Sun,D.Y., Hurley,L.H. and Ebbinghaus,S.W. (2005) Evidence for the presence of a guanine quadruplex forming region within a polypurine tract of the hypoxia inducible factor 1 alpha promoter. *Biochemistry*, **44**, 16341–16350.
57. Nicholls,A., Sharp,K.A. and Honig,B. (1991) Protein folding and association—insights from the interfacial and thermodynamic properties of hydrocarbons. *Proteins*, **11**, 281–296.