Solution Structure of a 2:1 Quindoline–c-MYC G-Quadruplex: Insights into G-Quadruplex-Interactive Small Molecule Drug Design

ARTICLE *in* JOURNAL OF THE AMERICAN CHEMICAL SOCIETY · OCTOBER 2011

Impact Factor: 12.11 · DOI: 10.1021/ja205646q · Source: PubMed

4 AUTHORS, INCLUDING:

Danzhou Yang
The University of Arizona
59 PUBLICATIONS **3,033** CITATIONS

SEE PROFILE

Published in final edited form as:
J Am Chem Soc. 2011 November 9; 133(44): 17673–17680. doi:10.1021/ja205646q.

Solution structure of a 2:1 quindoline-c-MYC G-quadruplex: insights into G-quadruplex-interactive small molecule drug design

Jixun Dai[‡], **Megan Carver**[‡], **Laurence H. Hurley**[‡,§,‖,†], and **Danzhou Yang**[‡,§,‖,†]

[‡]College of Pharmacy, The University of Arizona, 1703 E. Mabel St, Tucson, AZ 85721

[§]BIO5 Institute, The University of Arizona, Tucson, AZ

[‖]Arizona Cancer Center, 1515 N. Campbell Avenue, Tucson, AZ

[†]Department of Chemistry, The University of Arizona, Tucson, AZ

Abstract

Unimolecular parallel-stranded G-quadruplex structures are found to be prevalent in gene promoters. The nuclease hypersensitivity element III_1 (NHE III_1) of the c-MYC promoter can form transcriptionally active and silenced forms and the formation of DNA G-quadruplex structures has been shown to be critical for c-MYC transcriptional silencing. The solution structure of a 2:1 quindoline–G-quadruplex complex has been solved and shows unexpected features, including the drug-induced reorientation of the flanking sequences to form a new binding pocket. While both 3’ and 5’ complexes show overall similar features, there are identifiable differences which emphasize the importance of both stacking and electronic interactions. For the first time we describe the importance of the shape of the ligand as well as the two flanking bases in determining drug binding specificity. These structures provide important insights for the structure-based rational design of drugs that bind to unimolecular parallel G-quadruplexes commonly found in promoter elements.

Keywords

unimolecular parallel G-quadruplex; *c-MYC promoter*; *NMR*; *small molecule quindoline*; *structure-based rational drug design*; *induced-fit binding*

Correspondence to: Danzhou Yang.

Corresponding Author yang@pharmacy.arizona.edu.

Supporting Information Placeholder

Author Contributions

The manuscript was written through contributions of all authors. D.Y. and J.D. designed and interpreted the experiments. J.D. performed NMR experiments and structure calculation. M. C. prepared DNA samples. J.D. and D.Y. prepared figures. D.Y., L.H.H., J.D., and M.C. wrote the manuscript. All authors have given approval to the final version of the manuscript.

Supporting Information Available: Proton chemical shifts for the quindoline-Pu22 complex. 1D ^{1}H NMR titration spectra of the wild-type mycPu22 with quindoline. Imino proton assignments of the 2:1 quindoline:MycG4 complex. Various expanded regions of the 2D-NOESY spectra of the 2:1 quindoline:MycG4 complex in pH7 100 mM K^+ solution and in pH6 10 mM K^+ solution. The hydrogen-deuterium exchange experiments of free MycG4 and 2:1 quindoline:MycG4 complex. The NMR structure of the free MycG4 formed in Pu22 in K^+ solution. The molecular modeling structure of the quindoline:MycG4 complex with the wild-type G23. 1D ^{1}H NMR spectra of 2:1 quindoline complexes of various modified c-MYC promoter sequences. Surface potential visualization of the 5’-face and 3’-face of MycG4 in the 2:1 quindoline:MycG4 complex. This material is available free of charge via the Internet at http://pubs.acs.org.

Introduction

DNA G-quadruplexes are a family of secondary DNA structures that consist of stacked G-tetrads connected by hydrogen bonds that involve both the Hoogsteen and Watson-Crick edges of guanine and stabilized by monovalent cations such as potassium and sodium [1]. Intramolecular G-quadruplexes have been found in a number of G-rich regions with biological significance, such as human telomeres, oncogene promoters, and 5'-UTR regions[1-4]. Parallel-stranded G-quadruplex structures are found to be prevalent in gene promoters including those of c-MYC, VEGF, Hif-1α, and c-KIT genes [5-9]. The NHE III_1 in the c-MYC promoter (mycPu27, Figure 1A) is the best-characterized GC-rich promoter element that has the capacity to form G-quadruplex and i-motif on the purine- and pyrimidine-rich strands, respectively. Over-expression of the c-MYC oncogene is one of the most common genetic aberrations found in a wide variety of solid tumors[10-12] and is mediated by a myriad of different molecular mechanisms[13, 14]. So far, a successful therapeutic approach to directly target modulation of c-MYC gene expression or a downstream target has not reached clinical trials[15], although a number of different approaches have been attempted. Among these more recent attempts is the targeting of c-MYC transcriptional control using small drug-like molecules that bind to the G-quadruplex in the c-MYC promoter[16]. Quarfloxin, which had reached phase 2 clinical trials, is a first-in-class G-quadruplex-interactive drug that disrupts nucleolin–G-quadruplex complexes in the non-template strand of rDNA, causing selective inhibition of RNA polymerase I and displacement of nucleolin into the nucleoplasm [17], where it can bind to the c-MYC G-quadruplex to inhibit gene expression [18]. This, along with other stress-induced p53-dependent pathways, leads to selective apoptosis and cell death in cancer cells[19]. The majority of the G-quadruplex-interactive compounds are planar molecules with multiple fused rings, which have been proposed to intercalate between the end G-tetrad and the flanking residues of the pre-formed "drug binding pockets" [1, 20-25]. Although there is a growing list of molecular structures reported for G-quadruplexes formed in gene promoters, the structures of their drug complexes have been more difficult to obtain. Here we report the solution structure of a novel 2:1 complex between a quindoline and a major G-quadruplex found in the c-MYC promoter that has important implications for the design of selectively G-quadruplex-interactive compounds that bind to this and other parallel unimolecular G-quadruplexes found commonly in promoter elements.

The G-rich strand of the c-MYC NHE III_1 is a 27-nt segment comprised of five consecutive runs of guanines (mycPu27, Figure 1A). Mutational analysis in conjunction with a luciferase reporter system has previously shown that the major G-quadruplex structure responsible for c-MYC transcriptional silencing in K^+ solution appears to involve the four consecutive 3′ runs of gua-nines in mycPu27 (mycPu22, Figure 1A) [5, 26], although this may also be in equilibrium with a second similar G-quadruplex involving an adjacent 5' run of guanines [27]. Compounds that stabilize the c-MYC G-quadruplex can repress c-MYC gene transcription [5, 28]. The major c-MYC promoter G-quadruplex adopts a parallel-stranded folding topology in K^+ solution [26, 29]. We have determined the molecular structure of the major loop isomer in K^+ solution using the sequence Pu22 (Figure 1A) [20], in which G14 and G23 are mutated to thymines to isolate the single major conformation (MycG4, Figure 1B) [26]. Transcriptional factors that bind to either the duplex (e.g., Sp1) or single-stranded (e.g., CNBP, hnRNP K) elements of the NHE III_1 in the c-MYC promoter cause transactivation, while the secondary DNA structures can form from the same element under negative superhelicity to silence transcription [27, 30]. NM23-H2 and nucleolin have been identified as proteins that facilitate the unwinding and folding of the G-quadruplex, respectively [18, 31]. Nucleolin, which binds selectively to the c-MYC G-quadruplex structure, antagonizes the transactivation effect of Sp1 and has been shown by ChIP analysis to bind to

the NHE III_1 in the c-MYC promoter [18], providing persuasive evidence for the presence of this G-quadruplex in a cellular context.

The quindoline compound (Figure 1C) is a derivative of the natural product cryptolepine [32, 33] and has been shown to stabilize the G-quadruplex formed in the c-MYC promoter and thus inhibit the expression of c-MYC in the hepatocellular carcinoma cell line H2p G2 [28]. Here we report that in this quadruplex drug complex, two quindoline molecules bind the MYC G-quadruplex, with one quindoline bound to each of the external tetrads. The drug binding mode to this unimolecular parallel-stranded G-quadruplex is distinct from what has been proposed [20]. Instead of binding within a preformed pocket provided by the flanking bases, both quindoline molecules bind to the MYC quadruplex in an "induced-fit" manner such that, upon drug binding, the conformation of flanking segments is changed dramatically from their free state so that a new drug binding pocket is formed. Together with the –1 or +1 flanking residue, each quindoline molecule forms an additional plane of stacking over a total of three of the four guanines in the external G-tetrads. The +2 and –2 flanking residues then wrap over the newly formed quindoline-base planes at each end, respectively. The structure of this 2:1 quindoline-quadruplex complex provides important insights into the structure-based design of small molecules targeting parallel G-quadruplexes for transcriptional modulation.

Results

Quindoline binds the MycG4 to form a well-defined 2:1 drug-DNA complex

We report here the solution structure of the drug complex of this major c-MYC promoter G-quadruplex with a quindoline compound in physiologically relevant K^+ solution using NOE-restrained molecular dynamics calculation. The free Pu22 DNA (Figure 1A) in pH 7 100 mM K^+ solution forms a single G-quadruplex conformation as indicated by twelve well-resolved imino proton peaks, which come from the twelve tetrad-guanines of three G-tetrad planes of MycG4 (Figure 1B), as shown previously[20]. Upon addition of quindoline compound to the Pu22 solution, the imino protons of the DNA first broadened at lower drug equivalence (0.5-1N) and then became markedly sharper at higher drug equivalence (Figure 2A) indicating a medium exchange-rate of quindoline binding to Pu22 on the NMR time-scale. The observation of a new set of twelve well-resolved imino proton peaks suggested the formation of a well-defined drug-DNA complex. The drug binding stoichiometry was shown to be 2, as the imino region did not change (but became somewhat broader) at the drug equivalence higher than 2N. The Pu22 sequence (Figure 1A) contains two G-to-T mutations at positions 14 and 23 to isolate the major conformation MycG4 (Figure 1B)[20]. To make sure that the same G-quadruplex is formed in the wild-type sequence, we also tested the binding of quindoline to the wild-type mycPu22 sequence by ^{1}H NMR. Our results showed a very similar imino region of the 2:1 wild-type mycPu22-quindoline complex as compared to that of the 2:1 Pu22-quindoline complex (Figure S1), suggesting the same G-quadruplex structure formed in both cases.

Complete proton assignment of the 2:1 quindoline-Pu22 complex, in which one quindoline covers the top G-tetrad and another covers the bottom G-tetrad

We prepared site-specifically labeled DNA samples at each guanine of Pu22 with low-enrichment (6%) incorporation of ^{15}N-labeled-G, and their quindoline complexes at drug equivalence of 2 in pH 7 100 mM K^+ solution. Using ^{15}N-edited experiments as previously reported [22, 34], the imino (one-bond connection to N1) (Figure S2) and base aromatic H8 (two-bond connection to N7) protons of guanines were unambiguously assigned. Based on this assignment, NMR titration data indicated that one quindoline covers the top G-tetrad and another quindoline covers the bottom G-tetrad, as indicated by the chemical shift

changes between the imino protons of the free and bound DNA (Figure 2A). A complete set of 2D-NMR data for the 2:1 quindoline-Pu22 complex was collected in pH 7 100 mM K^+ at temperatures from 10 to 45 °C. Using the assignment of guanine H1s and H8s and sequential assignment strategy [20, 22], we were able to assign all the protons of the 2:1 quindoline-Pu22 complex except for those of the 3'-flanking segment, i.e., T23, A24 and A25, for which more than one set of proton peaks were observed, indicating the presence of multiple conformations (Figure S3). In addition, the aromatic ring protons of quindoline are quite broad and not well resolved (Figure S4A). A set of minor peaks, presumably from the free quindoline, was clearly observed for the quindoline side-chain protons, e.g. QuiMe5' and QuiH4' (Figure 1C), and exchange NOE crosspeaks were observed for those protons (Figure S5). To optimize drug-DNA interactions, we examined the binding of quindoline with Pu22 at various salt concentrations. At lower salt concentration, e.g., 10 mM K^+, the binding of quindoline to Pu22 G-quadruplex was improved, as indicated by the sharper peaks from the complex DNA (Figure 2B). A new complete set of 2D-NMR data was collected for the 2:1 quindoline-Pu22 complex in 10 mM K^+, from 15 to 45 °C, at pH 6 which was used to slow the exchange rates of imino protons. Encouragingly, one set of protons peaks was observed for T23, A24 and A25, indicating a well-defined conformation of the 3'-flanking segment (Figure 3A), in contrast to that at 100 mM K^+ (Figure S3). Moreover, the aromatic ring protons of quindoline became markedly sharper and well resolved (Figure S4B). It is thus indicated that quindoline binds the 3'-end of MycG4 with a higher affinity at the low salt condition, which resulted in a well-resolved NMR spectrum of the 2:1 drug-DNA complex that is sufficient for complete NMR structure determination (Figures 3, S4B, S6). We were able to assign all the protons of the 2:1 quindoline-Pu22 complex in pH 6 10 mM K^+ (Table S1), and the changes in proton chemical shifts between the free and bound DNA at 25 °C are shown in Figure 2C. It is important to note that the resolved portion of the NMR spectra of the 2:1 quindoline-Pu22 complex in pH 7 100 mM K^+ and pH 6 10 mM K^+ (Figure 2A and 2B, Figure 3A and S3) are almost the same, indicating the same complex structure formed at both conditions. The NMR variable temperature study of the 2:1 quindoline-Pu22 complex showed the imino protons of the MycG4 can be clearly detected at 95°C in 10 mM K^+ solution (Figure 4), while the imino protons of the free MycG4 disappeared at 80°C, indicating that binding of quindoline increased the stability of the MycG4 by more than 15°C in its melting temperature. In addition, the NMR hydrogen-deuterium exchange experiments showed that the $t_{1/2}$ of hydrogen-deuterium exchange for the imino protons of the middle G-tetrad is about 12 days for the 2:1 Quindoline:MycG4 complex compared to 2 days for the free MycG4 at 25°C (Figure S7).

The NMR structure of the 2:1 quindoline-Pu22 complex showed a dramatic rearrangement of the flanking segments at both ends induced by quindoline binding

Using our established NMR methods [20, 22], we were able to determine the molecular structure of the quindoline-MycG4 complex using a NOE-restrained distance geometry and molecular dynamics approach (Figure 5A, PDB ID 2L7V, Table 1). A representative model of the complex structure is shown in two different views in Figure 5B. The quindoline-bound MYC G-quadruplex was found to adopt the same folding as that of the free MycG4 formed by Pu22 (Figure 1B) [20], a parallel-stranded G-quadruplex with three double-chain-reversal loops of T, TA, and T that are mostly exposed to solvent. The glycosidic torsion angles of all DNA nucleotides were in the *anti* conformation as indicated by the intraresidue H8-H1′ NOE intensities (Figure 3A). In the complex structure, inter-residue NOE connectivities were clearly observed for the 5'-T(-3) G(-2)A(-1) and 3'-T(+1)A(+2)A(+3) flanking segments (Figure 3A), suggesting a continuous stacked conformation at both ends. However, upon drug binding, a dramatic rearrangement was observed for the flanking segments at both ends. Specifically, with each quindoline molecule, the –1 and +1 flanking

residues, i.e., A6 and T23, respectively, formed a plane capping the 5’ and 3’ external G-tetrads of the MycG4 quadruplex (Figure 6). In addition, the –2 or +2 flanking residues, i.e., G5 and A24, rearranged their conformation to stack over and wrap the newly formed quindoline-DNA plane at each end, respectively (Figure 6). For the 5’-TGA flanking segment, in the free MycG4 structure, A6 stacks over the top G-tetrad while G5 is looped out in the groove [20]. Rather than binding in this preformed pocket by intercalating between the top G-tetrad and A6, the binding of quindoline recruited A6 to form a plane covering the top G-tetrad (Figure 6A). A6 was pushed towards the G20-G7 edge of the top G-tetrad, as defined by the NOE interactions between A6H2/G20H1 (medium), A6H8/G20H1 (weak) (Figure 3B), and A6H2/G20H1’ (medium) (Figure 3A), as well as by the marked down-field shifting of the chemical shifts of the H8 and H1’ protons of A6 as compared to those of A6 in the free MycG4 (Figure 2C). The –2 flanking residue, G5, rearranged its conformation to completely stack over and wrap the quindoline-A6 plane. The well-stacked position of G5 was supported by the significant up-field shifting of G5H8 and G5 sugar protons (Figure 2C); e.g., G5H8 is upfield-shifted for about 1 ppm from its free form, suggesting a strong ring-current effect [35] of the stacking interaction. The profound upfield-shifting of the G5 protons, combined with the clear NOE connectivity between A6H8 and G5 sugar protons (Figure 3A & S6), suggested that the H8 end and the sugar moiety of G5 stacked right above the base ring of A6. Interestingly, the inter-residue NOE connectivity patterns observed for the T4-G5-A6-G7 segment between the sugar protons and (i+1) H8 protons were quite different from those of a regular right-handed-twist DNA backbone, and suggested a somewhat left-handed-twist backbone conformation. For example, an unusual pattern of NOE connectivity was observed between G7H8 and A6 sugar protons, such that the NOE interactions of G7H8 with A6H4’ and A6H1’ were much stronger than that with A6H3’ (Figure 3A & S6), a phenomenon observed for the zigzag sugar conformation of *anti*-cytosines in Z-DNA [36]. Moreover, the NOE between A6H8/G5H3’ was unusually markedly stronger than that between A6H8/G5H1’ (Figure 3A & S6). The conformation of the 5’-terminal T4, as expected, appeared to be quite flexible and not as well defined (Figure 5A). For the 3’-TAA flanking segment, a distinct conformation was also observed in the bound form as compared to that in the free form (Figure 6B). In the free MycG4 structure, A25 folds back to form a potential base pair with T23 to cover the bottom G-tetrad, while A24 stays above the T23:A25 base pair (Figure S8). Instead of intercalating between the bottom G-tetrad and the T23:A25 base pair of the preformed binding pocket, a significant rearrangement of the 3’ flanking segment was observed upon quindoline binding. As indicated by NMR data, similar to the 5’-end, the quindoline molecule recruited the +1 T23 to form a new plane covering the 3’ G-tetrad, with the +2 A24 rearranged to stack over the quindoline-T23 plane (Figure 6B). Clear NOE connectivity patterns of a continuous right-handed DNA were observed for the 3’-flanking segment, i.e., between G22-T23, T23-A24, and A24-A25 steps. T23 was pushed away to the G22-G9 edge of the 3’ G-tetrad, as indicated by weaker NOE interactions between T23H6 and G22 sugar protons (Figure 3A and S6). The down-field shifting of T23H6 (Figure 2C) was likely due to its positioning outside of the G22 and A24 bases. A24 appeared to largely stack over the T23-quindoline plane, as indicated by the NOE connectivity pattern between T23 and A24 (Figure 3A), and the upfield-shifting of most of the A24 protons (Figure 2C). The 3’-terminal A25 was quite flexible in its conformation and appeared to be outside of A24 (Figure 5).

The orientations of the quindoline molecules at both ends were clearly defined by inter-molecular NOE interactions between quindoline and DNA

As shown in the complex structure, the quindoline molecule at the 5’-end covers the G11 and G16 bases of the top tetrad, while the quindoline molecule at the 3’-end covers G13 and G18 of the bottom tetrad (Figure 6A and 6B). This is supported by the NOE interactions between quindoline and MycG4 (Figure 3C), as well as the chemical shift change of tetrad-

guanine imino protons (Figure 2C). The proton resonances of the quindoline molecules are broader as compared to those of DNA (Figure 3A). While the two quindolines bind at different ends of MycG4, proton resonances between the two molecules are not resolved, i.e., one set of peaks is observed for the two quindoline molecules. This may be another reason for the broader peaks of quindoline protons, in addition to the dynamic process of the drug binding. Although inter-molecular NOE interactions between quindoline and DNA are broader and weaker, critical NOE interactions were observed between quindoline and DNA (Figure 3C) which clearly define the orientation of the quindoline molecules at both ends. For example, for the quindoline molecule that binds the 5' G-tetrad, inter-molecular NOE interactions were observed between the protons of quindoline ring A with G7 and G11, e.g., QuiH2/G7H1 and QuiH4/G11H1', and between the protons of quindoline ring D with G16, e.g., QuiH11/G16H1 and QuiH8/G16H1' (Figure 6C). These NOE interactions unequivocally defined the binding position of the quindoline molecule on the top G-tetrad (Figure 6A). Similarly, the binding position of the quindoline molecule on the bottom G-tetrad was clearly defined (Figure 6B). Inter-molecular NOE interactions were observed between the protons of quindoline ring A with G18, e.g., QuiH4/G18H2'&2'' and QuiH2/G18H1, and between the protons of quindoline ring D with G13, e.g., QuiH9/G13H2'&2'', QuiH9/G13H8, and QuiH8/G13H1' (Figure 6D). In addition to NOE interactions between quindoline protons and the stacking G-tetrads, inter-molecular NOE interactions were observed between quindoline protons and flanking residues, for example, QuiH10/A24H2 defines the capping conformation of A24 above the quindoline-T23 plane (Figure 6D), and QuiH11/A6H2 defines the position of A6 which forms a plane with quindoline covering the 5' G-tetrad (Figure 6C).

A potential hydrogen bond appeared to form between T23 and the 3'-quindoline, whereas the –2 residue G5 is more critical for the specific binding of the 5'-quindoline

The quindoline molecule is protonated at the N1 site at pH 7, as suggested by its pKa (~8.3) as reported previously [28]. In the complex structure derived from NOE-restrained molecular dynamics calculation using the N1-protonated quindoline, a potential hydrogen bond appeared to be formed between T23O4 and QuiN1H, with a distance of ~1.9 Å (Figure 6B). However, the QuiN1H could not be detected by NMR due to its fast exchange with water, thus we also performed the NOE-restrained molecular dynamics calculation using the N1-nonprotonated quindoline. The results showed that the intermolecular electrostatic energy between the 3'-quindoline and MycG4 is about –469 (± 8.5) and –426 (± 10.9) kCal/mol for the protonated and nonprotonated forms, respectively. This result indicated that the N1-protonated quindoline is more favored for the quindoline binding at the 3'-end, likely due to the potential H-bond of T23O4:QuiN1H. This H-bond interaction between DNA and quindoline would be strengthened at lower salt concentration, which could explain the improved binding of quindoline at the 3'-end of MycG4. As T23 is a mutation from the wild-type G23, we performed molecular dynamics calculations of the complex structure with the wild-type G23 and found a very similar structure in which a potential H-bond could be formed between G23O6 and QuiN1H (Figure S9). Significantly, while the 5' quindoline binding pocket appeared to be stable at both high and low salt conditions, the mutational analysis of the binding of quindoline with various modified Myc promoter sequences (Figure S10) showed that the –2 residue G5 is more critical for the specific binding of quindoline to Pu22 than the +2 residue A24, and the G-to-T mutation at position 5 markedly weakened the quindoline binding (Figure S10D). The need for the –2 G5 at the 5'-end may be related to the lack of H-bond interaction between the 5' quindoline and the –1 A6, and suggesting a sequence-specific capping interaction. The NOE-restrained molecular dynamics calculation also showed that the intermolecular energy between the 5'-quindoline and MycG4 was about the same for both the protonated and non-protonated drug forms.

Discussion

Solving the structure of a 2:1 quindoline–G-quadruplex promoter complex by NMR allows us to address two important questions: (1) what are the specific molecular recognition determinants between the two quindoline molecules and this c-MYC G-quadruplex? and (2) what insights does this structure provide for future structure-based rational drug design of molecules that interact with unimolecular parallel-stranded G-quadruplexes? The structure of the 2:1 quindoline–Myc G-quadruplex complex is unique in that, at both ends of the unimolecular c-MYC G-quadruplex, a quindoline molecule binds and induces a similarly large reorientation of the flanking sequences to form new drug binding pockets. It is important to note that the same G-quadruplex appears to form in the 2:1 quindoline complex of the wild-type mycPu22 sequence (Figure S1). While the overall structures of both 5’ and 3’ quindoline complexes have largely common features, such as base stacking over two adjacent guanines and recruitment of either the –1 or +1 base (adenine or thymine) that is aligned in the same plane as quindoline, there are also some important differences. The origin of these differences is due to inherent structural features that are associated with the 3’ and 5’ faces as well as the flanking sequences. The binding of quindoline at the 5’-end is more favored at physiologically relevant K^+ concentration, because the 5’-face is more hydrophobic and more accessible for ligand stacking (Figures 5B and S11). Binding of the 5’-end quindoline is more dependent on the stacking interactions that are inherent in the induced binding pocket created by the 5’ tetrad and the two flanking bases at –1 and –2, thus it is K^+ concentration independent. In contrast, the 3’-end is more hydrophilic and less accessible for ligand stacking (Figures 5B and S11). Indeed, the overlay of the 3’-quindoline with the adjacent G-tetrad is clearly less profound as compared to that of the 5’-quindoline (Figure 6A and 6B). Consequently, the increased stability of the 3’-end quindoline complex at lower ionic strength is likely to be related with the specific H-bond interaction between QuiN1H and T23O4 (Figure 6B) and relies less on capping interactions. These types of subtle but important differences related to the effect of ionic strength in ligand recognition between the 3’ and 5’ ends can only be monitored in solution and thus are amenable to NMR studies. In contrast to the G-tetrad interactions, both flanking two-base sequences undergo unexpected large conformational changes to assemble new capping structures, in a manner somewhat analogous to a reorganized ligand-induced fit observed in riboswitches [37]. This is distinct from the previously proposed model wherein the planar G-quadruplex-interactive compounds would intercalate between the external G-tetrad and the flanking residues of the pre-formed “drug binding pockets” [1]. For each case there is proof of a defined overlay of the quindoline molecule with just two of the four guanines from each external tetrad and recruitment of the flanking +1 or –1 base to form a quasi-triad plane that is intercalated between the external 3’ or 5’ tetrad and the +2 or –2 base. We term this induced rearrangement of the flanking sequence with the external tetrad and quindoline an “induced intercalated triad pocket” to describe the coplanar arrangements of the quindoline (which overlay two gua-nines) with the adenine (–1) or thymine (+1) and their combined intercalation between the external tetrad and –2 or +2 base. The 5’-end complex in which an adenine is recruited into the same plane as the quindoline and then capped with a guanine is the most stable and its inherent stability is highly dependent on the capping guanine since deletion or mutation to a thymine dramatically reduces the stability. The 3’-end complex, where thymine is recruited as the in-plane base, is only stable under low ionic strength, emphasizing the importance of ionic, rather than stacking, binding interactions.

We next addressed the question of how this quindoline-induced rearrangement of the flanking sequence is related to previously determined structures of ligand G-quadruplex complexes. Most of the known structures of ligand–G-quadruplex complexes have been determined by X-ray crystallography and are derived from telomeric sequences that form bimolecular and tetramolecular species [38]. The only unimolecular ligand complex with a

promoter G-quadruplex is the NMR-derived structure of a TMPyP4–MYC G-quadruplex complex, which contains a guanine-to-inosine substitution that induces a G-strand discontinuity [39] that did not agree with the DMS footprinting data of the wild-type mycPu27 sequence [5]. In this complex the TMPyP4 stacks over the 5'-end but the orientation of TMPyP4 was not resolved by NMR data and a well-defined binding pocket was not observed. There are two dimeric telomeric G-quadruplexes in which a thymine residue is recruited into a similar in-line triad plane: in one case it is a thymine from the 4T diagonal loop in a dimeric antiparallel structure [40], and in a second case one of the 3' terminal thymines is in the plane of BRACO-19 that stacks between the two bimolecular quadruplexes [41]. However, because of the multimeric nature of both structures, they are less relevant to the unimolecular species described here although the hijacking of a base into the plane of the ligand reinforces the principle described here for both quindoline G-quadruplex complexes

Finally, to address the insights into structure-based design of molecules that interact with unimolecular parallel G-quadruplex structures such as those found in c-MYC, VEGF, Hif-1α, and c-KIT promoters, we first categorize externally bound G-quadruplex ligands into two main types: those typified by telomestatin and TMPyP4, which contain symmetric cyclic fused rings that are quasi or fully 4-fold symmetrical and overlap equally with all four guanines in the external tetrad, and ligands such as acridine-based compounds or Quarfloxin and the quindoline described here, which are inherently asymmetric and contain a smaller stacking moiety. Compounds with symmetric cyclic fused rings, such as telomestatin [42] and TMPyP4 [43], have been considered to be the optimal G-quadruplex-interactive compounds as they provide maximized stacking interactions with the external G-tetrad. However, we have examined the binding of many G-quadruplex-interactive compounds to DNA G-quadruplexes and found that, in general, those compounds with symmetric cyclic fused rings do not appear to bind G-quadruplexes specifically, whereas asymmetric compounds containing a smaller stacking moiety, in particular the crescent shape moiety, are more likely to bind in a defined manner to a specific G-quadruplex. An important implication from our structure is that, unlike the symmetrical cyclic ligand typified by TMPyP4, specific binding and selectivity of the quindoline and other similar crescent-shaped molecules are determined by both the identity of the binding end (3' or 5') and the flanking two bases. It is important to note that the crescent-shaped quindoline provides optimal overlay of two guanines of a G-tetrad. In addition, the electrostatic interaction between the diethylamino group in the side chain of quindoline and the DNA phosphate backbone could help orient and stabilize the quindoline nucleus, which in turn specifically pinpoints the potential location of substituents that can interact in the grooves and with the loops. It is anticipated that small changes in both the shape and electronic structure of the ligand, as well as the identity of the flanking bases making up the intercalated triad pockets, will affect the precise positioning of the ligand relative to the adjacent G-tetrad. The 2:1 quindoline–c-MYC G-quadruplex complex described here provides an important case study for the biologically important parallel intramolecular G-quadruplexes in promoter elements. For this "induced intercalated triad pocket" to form, a suitable single-stranded flanking segment containing at least two bases must exist. Under negative supercoiled conditions there is clear evidence of a single-stranded region, which is quite extensive, at either side of the G-quadruplex in a duplex element [27]. The inherent flexibility of the deoxyribose phosphate backbone of DNA in the single-stranded region flanking the G-quadruplex would likely provide a not too dissimilar environment to that found in the flanking regions here. Therefore it is not unreasonable to expect the same drug-induced binding pocket to form in the context of an associated i-motif in a duplex element. Furthermore, this reorientation occurs at both ends of the c-MYC G-quadruplex and a similar recruited thymine has been shown to be present in crystal structures of two bimolecular structures [40, 41], suggestive of a significant feature of drug complexes with G-quadruplexes. This divides the molecular mechanism for G-quadruplex

recognition by small molecules of this type into two clear sequential steps, involving first the newly recognized induced inter-calated triad pocket recognition and second the groove/loop interactions. The first step provides new opportunities for selective recognition by small molecules that bind to unimolecular parallel structures. It is noted that different paradigms likely will exist for anti-parallel or mixed parallel/antiparallel structures.

Methods

The DNA oligonucleotides were synthesized and purified as described previously [20-22]. 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 0.1-2.5 mM DNA in 25 mM K-phosphate buffer (pH 7.0) and 75 mM KCl, or in the low-salt solution, 10 mM K-phosphate buffer.

NMR experiments were performed on Bruker DRX-600 and 800MHz spectrometers as described previously [20-22]. The 1D GE-JRSE HMQC experiments were used to measure ^{15}N-edited spectra [44]. The jump-return spin-echo was used to suppress the water peak with maximum intensity tuned to 11ppm [45]. Homonuclear 2D NMR experiments, including NOESY, TOCSY, and DQF-COSY, were collected at 10, 15, 20, 25, 30, 35, and 45°C, for the 2:1 quindoline-Pu22 complex samples in water and D2O solution at both pH 7 100 mM K^+ and pH 6 10 mM K^+. The linewidths of the quindoline complex became markedly broader at temperatures lower than 15 °C. Jump-return and watergate were used to suppress the water signal in the spectra. The relaxation delays were set to 2.5 s in 2D experiments. The software Sparky (UCSF) was used in peak assignments and integrations. 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 of ±30%. The methyl base proton Me-H6 distance (2.99 Å) was used as a reference. The distances involving the unresolved protons, e.g., methyl protons, were assigned using pseudo-atom notation in X-PLOR. Proton resonances between the two quindoline molecules are not resolved, therefore the intermolecular quindoline-DNA restraints were used with higher boundaries (±30%).

The structures of complex were calculated using XPLOR [46] and Insight II/Discover (Accelrys, CA). Metric matrix distance geometry (MMDG) and simulated annealing calculations were carried out in X-PLOR to embed and optimize 100 initial structures for the Pu22 sequence, as described previously [21, 22]. The experimentally obtained distance restraints and G-tetrad hydrogen-bonding distance restraints were included during the calculation. A total of 643 distance restraints, of which 184 are from inter-residue NOEs, and 25 are from inter-molecular NOEs between quindoline and MycG4, were incorporated into the NOE-restrained structure calculation. Dihedral angle restraints were used for the glycosidic torsion angle (χ) based on the experimentally determined *anti* conformations. The 20 best molecules were selected based both on their minimal energy terms and number of NOE violations, and were further subjected to 350 ps of NOE distance-restrained molecular dynamics simulation with time steps of 1fs at 300K in explicit solvents in Insight II/Discover using CFF forcefield, as described previously [20]. The quindoline-MycG4 complex was soaked into water solvent using a 10 Å water layer. Two K^+ ions were included between the three G-tetrad planes, and 21 K^+ ions were used to counter the negative charge of DNA backbone. Hydrogen bond restraints were applied to the G-tetrads, using a quadratic energy function with a force constant of 200 kcal/mol/$Å^2$. The 10 lowest energy structures were extracted from MD simulation trajectory and were further energy-minimized.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank Dr. Vijay Gokhale for his help with molecular modeling. This research was supported by the National Institutes of Health (1S10 RR16659 and CA122952 to DY).

References

1. Yang D, Okamoto K. Future Medicinal Chemistry. 2010; 2(4):619–646. [PubMed: 20563318]
2. Punchihewa, C.; Yang, DZ. Therapeutic Targets and Drugs-G-quadruplex inhibitors.. In: Hiyama, K., editor. Telomeres and Telomerase in Cancer. Springer; NJ, USA: 2009. p. 251-280.
3. Qin Y, Hurley LH. Biochimie. 2008; 90(8):1149–1171. [PubMed: 18355457]
4. Kumari S, Bugaut A, Huppert JL, Balasubramanian S. Nature Chemical Biology. 2007; 3(4):218–221.
5. Siddiqui-Jain A, Grand CL, Bearss DJ, Hurley LH. PNAS. 2002; 99(18):11593–11598. [PubMed: 12195017]
6. Sun DY, Guo KX, Rusche JJ, Hurley LH. Nucleic Acids Research. 2005; 33(18):6070–6080. [PubMed: 16239639]
7. De Armond R, Wood S, Sun DY, Hurley LH, Ebbinghaus SW. Biochemistry. 2005; 44(49):16341–16350. [PubMed: 16331995]
8. Rankin S, Reszka AP, Huppert J, Zloh M, Parkinson GN, Todd AK, Ladame S, Balasubramanian S, Neidle S. Journal of the American Chemical Society. 2005; 127(30):10584–10589. [PubMed: 16045346]
9. Fernando H, Reszka AP, Huppert J, Ladame S, Rankin S, Venkitaraman AR, Neidle S, Balasubramanian S. Biochemistry. 2006; 45(25):7854–7860. [PubMed: 16784237]
10. Lutz W, Leon J, Eilers M. Biochimica Et Biophysica Acta-Reviews on Cancer. 2002; 1602(1):61–71.
11. Dang CV. Mol. Cell. Biol. 1999; 19(1):1–11. [PubMed: 9858526]
12. Eilers M, Eisenman RN. Genes & Development. 2008; 22(20):2755–2766. [PubMed: 18923074]
13. Marcu KB, Bossone SA, Patel AJ. Annual Review of Biochemistry. 1992; 61(1):809–858.
14. Spencer CA, Groudine M. Adv Cancer Res. 1991; 56:1–48. [PubMed: 2028839]
15. Prochownik EV, Vogt PK. Genes Cancer. 2010; 1(6):650–659. [PubMed: 21132100]
16. Brooks TA, Hurley LH. Genes Cancer. 2010; 1(6):641–649. [PubMed: 21113409]
17. Drygin D, Siddiqui-Jain A, O'Brien S, Schwaebe M, Lin A, Bliesath J, Ho CB, Proffitt C, Trent K, Whitten JP, Lim JKC, Von Hoff D, Anderes K, Rice WG. Cancer Res. 2009; 69(19):7653–7661. [PubMed: 19738048]
18. Gonzalez V, Guo K, Hurley L, Sun D. J. Biol. Chem. 2009; 284(35):23622–23635. [PubMed: 19581307]
19. Daniely Y, Dimitrova DD, Borowiec JA. Mol Cell Biol. 2002; 22(16):6014–22. [PubMed: 12138209]
20. Ambrus A, Chen D, Dai J, Jones RA, Yang DZ. Biochemistry. 2005; 44(6):2048–58. [PubMed: 15697230]
21. Dai J, Chen D, Jones RA, Hurley LH, Yang DZ. Nucleic Acids Research. 2006; 34(18):5133–5144. [PubMed: 16998187]
22. Dai J, Carver M, Punchihewa C, Jones RA, Yang DZ. Nucleic Acids Research. 2007; 35(15):4927–4940. [PubMed: 17626043]
23. Neidle S, Parkinson G. Nature Reviews Drug Discovery. 2002; 1(5):383–393.
24. Mergny JL, Helene C. Nature Medicine. 1998; 4(12):1366–1367.
25. De Cian A, Lacroix L, Douarre C, Temime-Smaali N, Trentesaux C, Riou JF, Mergny JL. Biochimie. 2007; 90(1):131–155. [PubMed: 17822826]

26. Seenisamy J, Rezler EM, Powell TJ, Tye D, Gokhale V, Joshi CS, Siddiqui-Jain A, Hurley LH. J Am Chem Soc. 2004; 126(28):8702–9. [PubMed: 15250722]
27. Sun D, Hurley LH. J Med Chem. 2009; 52(9):2863–74. [PubMed: 19385599]
28. Ou TM, Lu YJ, Zhang C, Huang ZS, Wang XD, Tan JH, Chen Y, Ma DL, Wong KY, Tang JCO, Chan ASC, Gu LQ. Journal of Medicinal Chemistry. 2007; 50(7):1465–1474. [PubMed: 17346034]
29. Phan AT, Modi YS, Patel DJ. J Am Chem Soc. 2004; 126(28):8710–8716. [PubMed: 15250723]
30. Brooks TA, Hurley LH. Nature Reviews Cancer. 2009; 9(12):849–861.
31. Dexheimer TS, Carey SS, Zuohe S, Gokhale VM, Hu X, Murata LB, Maes EM, Weichsel A, Sun D, Meuillet EJ, Montfort WR, Hurley LH. Mol Cancer Ther. 2009; 8(5):1363–77. [PubMed: 19435876]
32. Guyen B, Schultes CM, Hazel P, Mann J, Neidle S. Organic & Biomolecular Chemistry. 2004; 2(7):981–988. [PubMed: 15034620]
33. Zhou J-L, Lu Y-J, Ou T-M, Zhou J-M, Huang Z-S, Zhu X-F, Du C-J, Bu X-Z, Ma L, Gu L-Q, Li Y-M, Chan AS-C. Journal of Medicinal Chemistry. 2005; 48(23):7315–7321. [PubMed: 16279791]
34. Dai J, Dexheimer TS, Chen D, Carver M, Ambrus A, Jones RA, Yang DZ. J Am Chem Soc. 2006; 128(4):1096–1098. [PubMed: 16433524]
35. Lazzeretti P. Progress in Nuclear Magnetic Resonance Spectroscopy. 2000; 36(1):1–88.
36. Wang AHJ, Quigley GJ, Kolpak FJ, Crawford JL, Vanboom JH, Vandermarel G, Rich A. Nature. 1979; 282(5740):680–686. [PubMed: 514347]
37. Mandal M, Breaker RR. Nat Rev Mol Cell Biol. 2004; 5(6):463.
38. Neidle S. Current Opinion in Structural Biology. 2009; 19(3):239–250. [PubMed: 19487118]
39. Phan AT, Kuryavyi V, Gaw HY, Patel DJ. Nature Chemical Biology. 2005; 1(3):167–173.
40. Haider SM, Parkinson GN, Neidle S. Journal of Molecular Biology. 2003; 326(1):117–125. [PubMed: 12547195]
41. Campbell NH, Parkinson GN, Reszka AP, Neidle S. Journal of the American Chemical Society. 2008; 130(21):6722–6724. [PubMed: 18457389]
42. Shin-ya K, Wierzba K, Matsuo K, Ohtani T, Yamada Y, Furihata K, Hayakawa Y, Seto H. Journal of the American Chemical Society. 2001; 123(6):1262–1263. [PubMed: 11456694]
43. Han FXG, Wheelhouse RT, Hurley LH. Journal of the American Chemical Society. 1999; 121(15): 3561–3570.
44. Szewczak AA, Kellogg GW, Moore PB. Febs Letters. 1993; 327(3):261–264. [PubMed: 7688695]
45. Sklenar V, Bax A. J Magnetic Resonance. 1987; 74(3):469–479.
46. Brünger, AT. X-PLOR Version 3.1: A system for X-ray crystallography and NMR. Yale University Press; New Haven, CT: 1993.

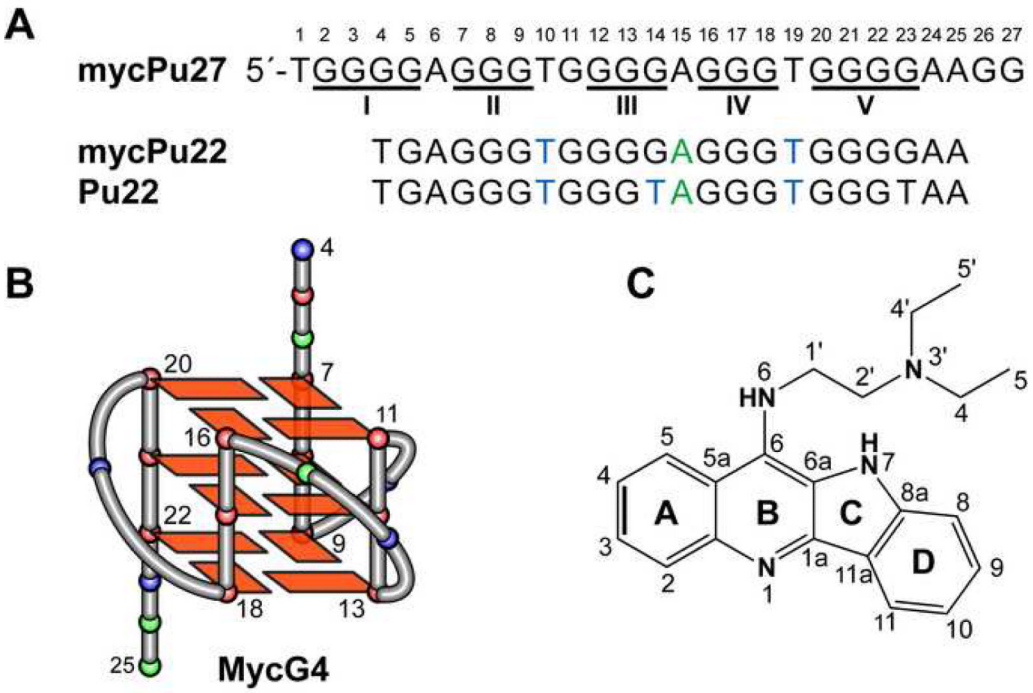

Figure 1.

(A) The promoter sequences of the NHE III_1 element of the c-MYC gene and its modifications. mycPu27 is the wild-type 27-mer G-rich sequence of the c-MYC NHE III_1. mycPu22 is the wild-type 22-mer G-rich sequence of the c-MYC NHE III_1 that forms the major G-quadruplex in physiologically relevant K^+ solution. Pu22 is the modified mycPu22 sequence with G-to-T substitutions at 14 and 23 positions, which adopts the single predominant c-MYC promoter G-quadruplex (MycG4) in K^+ solution and whose structure was determined by NMR (20). (B) The folding topology of MycG4 adopted by Pu22, the major c-MYC promoter G-quadruplex in K^+ solution. Red box = guanine; green ball = adenine, blue ball = thymine. (C) The quindoline molecule with numbering used in this study.

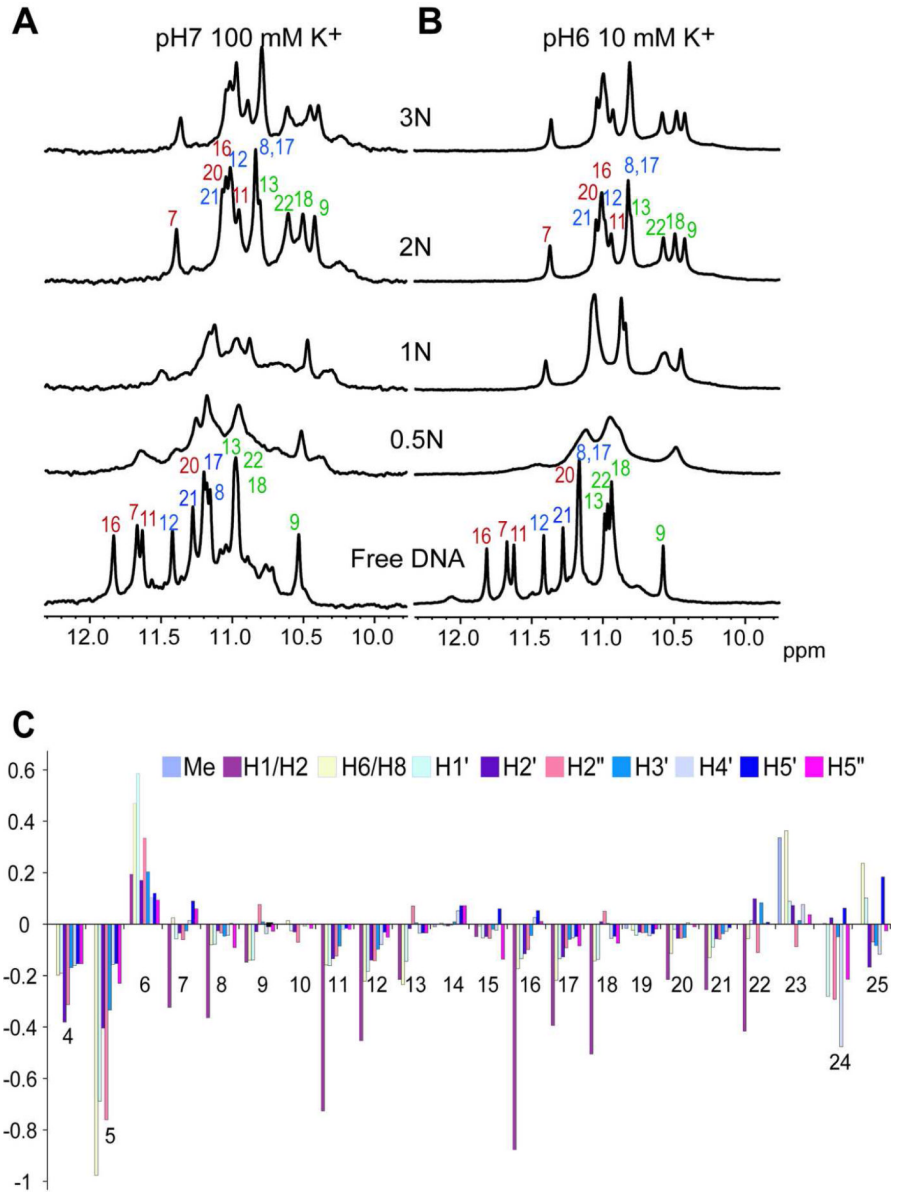

Figure 2.

Imino proton regions of the 1D ^{1}H NMR titration spectra of Pu22 with quindoline in pH7 100 mM K^{+} solution (A) and pH6 10 mM K^{+} solution (B) at 25 °C. The assignments of imino protons of the free Pu22 DNA and 2:1 quindoline:Pu22 complex are shown above the spectra. The imino protons from the 5' G-tetrad are colored in red, the middle G-tetrad in blue, and the 3' G-tetrad in green (see Figure 1B). (C) The chemical shift difference of the Pu22 protons between the free form and the 2:1 drug complex at 25 °C. The residue numbers of Pu22 are based on the numbering shown in Figure 1A.

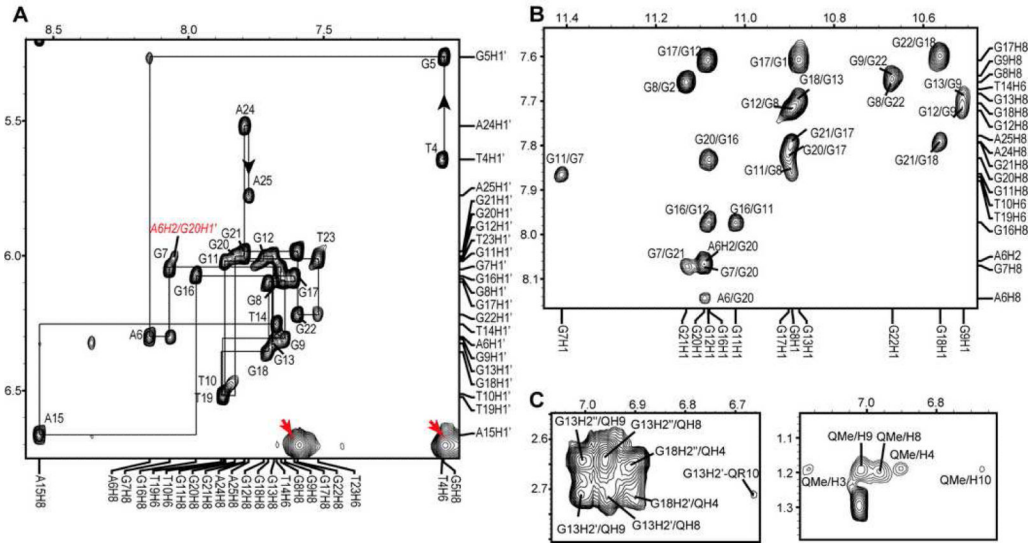

Figure 3.

(A) The expanded H8/H6-H1′ region of the 2D-NOESY spectrum of the 2:1 quindoline:MycG4 complex. The complete sequential assignment pathway is shown. The two NOE peaks from the quindoline molecules are labeled by red arrows, and show broader linewidths as compared to those from DNA. (B) The expanded H1-H8/H2/H6 region, with labeling, of the 2:1 quindoline:MycG4 complex. (C) The expanded regions of the 2D-NOESY spectrum of the 2:1 quindoline:MycG4 complex, showing some examples of the intermolecular NOEs between the quindoline and DNA (left) and between quindoline itself (right). Condition: pH6, 10 mM K^+, 35°C.

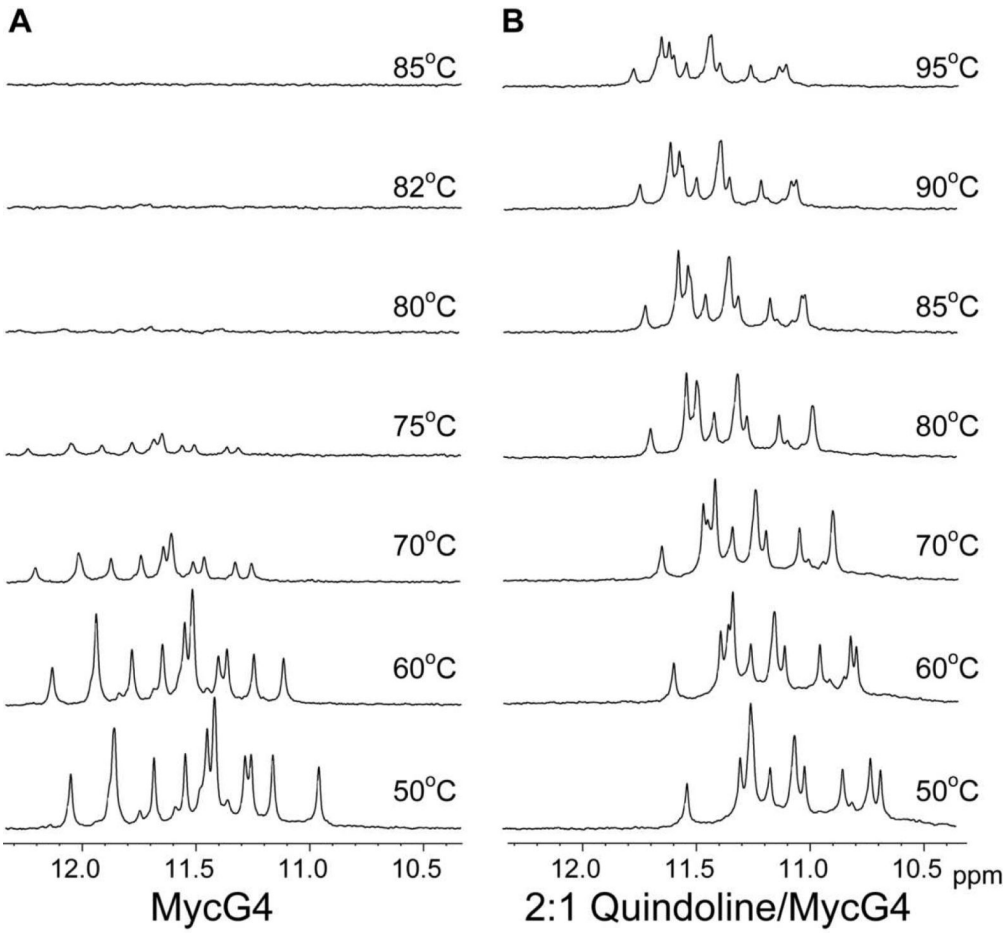

Figure 4.
The imino regions of 1D ^{1}H NMR spectra of MycG4 (A) and the 2:1 quindoline:MycG4 complex at various temperatures in pH6 10 mM K^{+} solution.

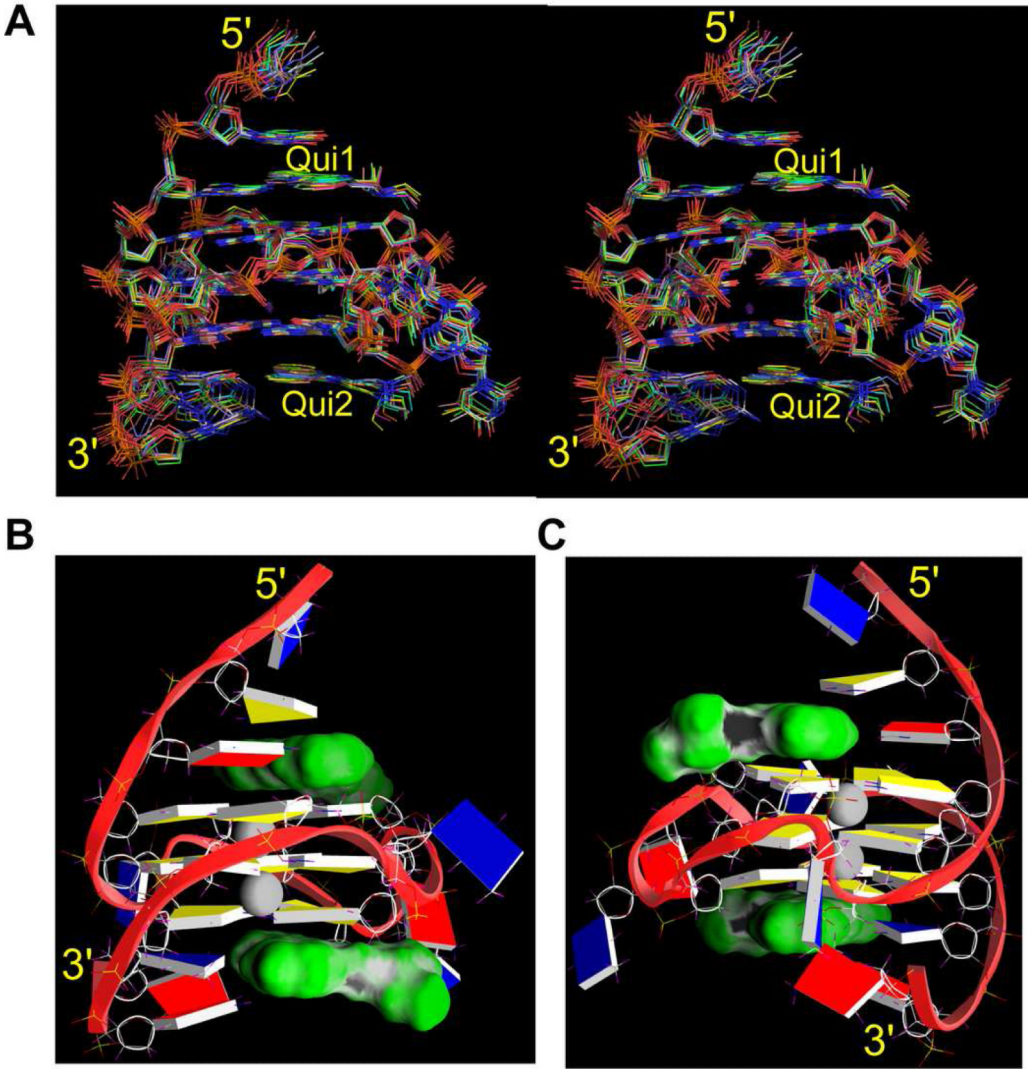

Figure 5.

(A) The ensemble of the superimposed 10 NMR-refined structures of the 2:1 quindoline:MycG4 complex in K^+ solution in stereo view. The two quindoline molecules are labeled. (B) A representative model of the NMR-refined 2:1 quindoline:MycG4 complex structure from two different views, prepared using GRASP (guanine = yellow, adenine = red, thymine = blue). The quindoline molecules are shown in space-filling model in green. The two potassium ions are shown as white balls.

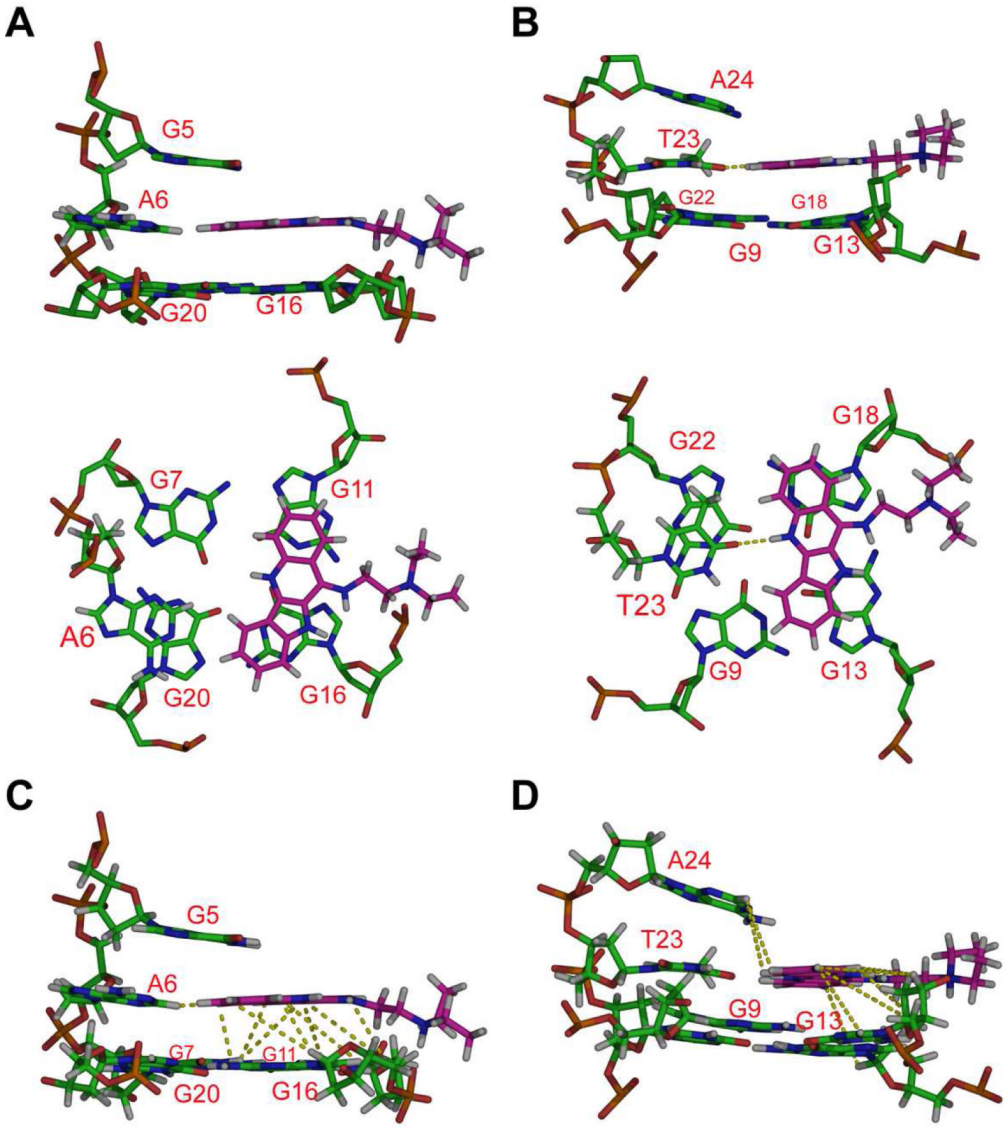

Figure 6.

Two different views of the drug-induced binding pockets at the 5’-end (A) and at the 3’-end (B). Some examples of intermolecular NOE interactions between quindoline and MycG4 DNA are shown as yellow lines for the 5’ drug-binding pocket (C) and the 3’ drug-binding pocket (D)

Table 1

Structural statistics for the 2:1 quindoline-Pu22 complex

Structural statistics	
Distance restraints	643
MycG4	594
Intraresidue	410
Interresidue	184
Intermolecule	25
Hydrogen bonds	24
Deviations from standard geometry	
Bond length (Å)	0.017± 0.007
Bond angle (°)	1.80 ± 0.03
NOE violations	
numbers (>0.2Å)	0.2 ± 0.42
pairwise r.m.s.d. of heavy atoms (Å)	
MycG4	1.05 ± 0.19
All	1.14 ± 0.21