HHS Public Access

Author manuscript

J Phys Chem B. Author manuscript; available in PMC 2018 November 22.

Published in final edited form as:

J Phys Chem B. 2017 November 22; 121(46): 10484–10497. doi:10.1021/acs.jpcb.7b09406.

Resolving the Ligand Binding Specificity in c-MYC G-quadruplex DNA: Absolute Binding Free Energy Calculations and SPR Experiment

Nanjie Deng[1,*], **Lauren Wickstrom**[2], **Piotr Cieplak**[3], **Clement Lin**[4], and **Danzhou Yang**[4,*]

[1]Department of Chemistry and Physical Sciences, Pace University, 1 Pace Plaza, New York, NY 10038

[2]Borough of Manhattan Community College, the City University of New York, Department of Science, New York, NY 10007

[3]Sanford Burnham Prebys Medical Discovery Institute, La Jolla, CA 92037

[4]Department of Medicinal Chemistry and Molecular Pharmacology, College of Pharmacy, Purdue University, West Lafayette, IN 47907

Abstract

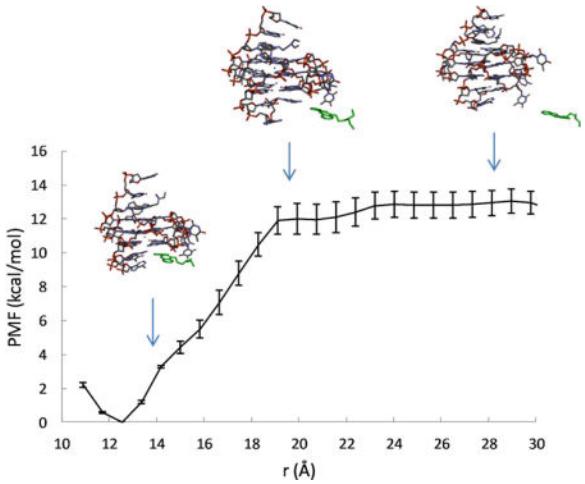
We report the absolute binding free energy calculation and surface plasmon resonance (SPR) experiment for ligand binding with the cMYC G-quadruplex DNA. The unimolecular parallel DNA G-quadruplex formed in the nuclease hypersensitivity element III1 of the c-MYC gene promoter regulates the c-MYC transcription and is recognized as an emerging drug target for cancer therapy. Quindoline derivatives have been shown to stabilize the G-quadruplex and inhibit the c-MYC expression in cancer cells. NMR revealed two binding sites located at the 5′ and 3′ termini of the G-quadruplex. Questions about which site is more favored and the basis for the ligand-induced binding site formation remain unresolved. Here, we employ two absolute binding free energy methods, the double decoupling and the potential of mean force methods, to dissect the ligand binding specificity in the c-MYC G-quadruplex. The calculated absolute binding free energies are in general agreement with the SPR result and suggest that the quindoline has a slight preference for the 5′ site. The flanking residues around the two sites undergo significant reorganization as the ligand unbinds, which provides evidence for ligand-induced binding pocket formation. The results help interpret experimental data and inform rational design of small molecules targeting the c-MYC G-quadruplex.

Graphical abstract

naniie.deng@gmail.com, telephone: (212) 346-1592; yangdz@purdue.edu, telephone: (765) 494-814.

SUPPORTING INFORMATION AVAILABLE

Contains the details of the methodology, and two tables of HSA calculation on the energetics of cavity water molecules in the two unliganded binding cavities.



INTRODUCTION

In recent years, there has been a growing interest in targeting the DNA G-quadruplex secondary structure formed in the guanine-rich regions in human telomeres and gene promoters for developing novel anti-cancer therapy[1–5]. A G-quadruplex is composed of stacked planar structures (G-tetrads) formed by four hydrogen bonded guanines, with additional stabilizing forces from K^+ or Na^+ ions located in the central ion channel of the G-tetrad planes[1]. DNA G-quadruplexes formed in the proximal location of promoters in a number of human proto-oncogenes have been found to regulate gene transcription[1,6,7], such as c-MYC[8,9], VEGF[10,11], HIF-1α[12], BCL-2[13–15], KRAS[16], Rb[17], c-KIT[18,19], RET[20], hTERT[21], and PDGFR-β[22,23]. c-MYC, one of the most commonly deregulated genes in human cancers, has a G-quadruplex-forming motif in the promoter Nuclease Hypersensitive Element (NHE) III1 which regulates 80–95% of its total transcription[8,9]. The G-quadruplex formed in the c-MYC promoter region NHE III_1 is the first and most extensively studied gene promoter G-quadruplex[9]; the major G-quadruplex formed in the c-MYC promoter in K^+ solution is a parallel-stranded structure[24]. Parallel-stranded G-quadruplex structures are found to be common in the human promoter sequences[1,6]. The G-quadruplex formed in the NHE III_1 of the c-MYC gene promoter has been recognized as a promising target for small molecule agents to modulate the transcription of c-MYC[25]. Quindoline derivatives (Fig. 1A) have been shown to stabilize the c-MYC G-quadruplex and down-regulate the expression of c-MYC in cancer cells[26,27]. The molecular structures of the c-MYC G-quadruplex and its drug complexes have been determined by solution NMR[24,28,29]. New G-quadruplex-interactive ligands with micromolar activities[30–33] have been identified in recent studies using the NMR structures of apo and ligand-bound c-MYC G-quadruplex molecules.

Despite the progress in the discovery of new G-quadruplex-interactive ligands, important questions remain regarding the ligand-binding specificity in the c-MYC G-quadruplex. Initially, the binding mode of the quindoline with the c-MYC G-quadruplex was thought to involve a single quindoline molecule end-stacking onto the surface of the 3′-end G-tetrad[26]. Subsequently, the NMR structure of the c-MYC G-quadruplex-quindoline complex revealed that each of the 5′- and the 3′-end binds a quindoline molecule[29], and that the two ligand

specific binding pockets were only formed upon binding of ligands (Fig. 1B). From the structure, it is unclear which of these two sites interacts more favorably with the ligand and why. In addition, the physical reason why the two binding pockets observed in NMR[29] are not preformed in the free G-quadruplex[24] is also not well-understood. A better understanding of binding site specificity and the mechanism for the ligand-induced binding pocket formation will assist in the interpretation of experimental data and inform on rational design of more potent c-MYC G-quadruplex-interactive small molecules.

Experimentally, equilibrium binding constants can be determined by methods such as surface plasmon resonance (SPR) and isothermal titration calorimetry (ITC). However, these methods cannot directly measure ligands binding affinities at different binding sites. Although site-specific binding parameters may be extracted by various fitting methods, the reliability of such procedures is low for different binding affinities at different binding sites unless these binding affinities are significantly different. More importantly, even when the individual binding constants can be extracted from model fitting, they cannot be assigned to the specific binding sites by these methods alone. While NMR spectroscopy can provide important structural information on binding-site conformation[29,31], it does not directly provide accurate measurement of binding affinities for individual binding sites, nor does it provide quantitative mechanistic information for binding specificity and affinity.

To gain atomistic insights into the binding site specificity in the c-MYC G-quadruplex, here we employ two absolute binding free energy methods[34–37], the potential of mean force method (PMF)[38–41] and double decoupling method (DDM)[42–44], on the quindoline compound used in the NMR structural study[26,29] to compute the intrinsic binding affinities for each binding site in the c-MYC G-quadruplex. In recent years, statistical mechanics based absolute binding free energy methods such as DDM[42–44] have increasingly been applied to compute protein-small molecule binding free energy[45–50], and we and others have shown that such methods can generate more accurate results compared with the more rapid docking and MM-PB(GB)/SA in estimating binding affinity for a number of protein-ligand systems[51–53]. However, the DDM method has not yet been applied to small molecule-DNA (or RNA) complexes. A DDM calculation of ligand-DNA binding will involve decoupling the ligand from the ionic solution containing the flexible and highly charged DNA. Such calculations can be computationally challenging for charged ligands because of the large magnitude of their electrostatic decoupling free energies[38,40,54] Here, we apply DDM for the first time to compute the absolute binding free energy for a ligand-nucleic acid system, taking into account the electrostatic corrections for the effects of the finite-size periodic simulation system on the electrostatic free energy.[54] We also employ the PMF method to study the binding thermodynamics of the quindoline-c-MYC G-quadruplex complex, using a new formula for the absolute binding free energy (see Methods). The PMF approach has a distinct advantage in computing the absolute binding free energy of charged ligand in that it uses a physical pathway to move the ligand out of the binding site, rather than alchemically decouples the charged ligand from the ionic solution.[38,47,55,56] Using two different absolute binding free energy methods with the same energy function can cross validate the internal consistency of the computational results. We have also experimentally determined the apparent binding constant of quindoline binding to the c-MYC G-quadruplex using SPR, by assuming that the two binding sites have the same binding affinity.

We find that the calculated absolute binding free energies of the quindoline molecule for the two binding sites are in general agreement with the SPR-derived overall binding free energy to within ~2 kcal/mol. The calculated equilibrium binding constant is also compatible with the IC_{50} of inhibition by this compound. Both the PMF and DDM calculations suggest that the quindoline compound has a slight preference for the 5′-end pocket of the G-quadruplex; this is consistent with the NMR spectral of the flanking residues forming the two binding sites. The DDM calculations indicate that the difference between the affinities of the 5′ and 3′ sites is mainly attributable to the favorable nonpolar interaction between the planar quindoline core and the 5′ binding cavity. In contrast, the electrostatic interactions are found to actually favor the 3′ binding, which is consistent with the NMR data showing that the 3′-end binding is ionic concentration-dependent[29]. The simulations also reveal significant conformational rearrangements in the 5′-end and 3′-end flanking residues as the ligands unbind, which eventually causes the instability of the binding pockets in the apo G-quadruplex. This result supports the notion that the binding pockets in the c-MYC G-quadruplex are induced by ligand binding[29]. To understand the instability of the binding pocket in the apo G-quadruplex, we computed the energetics of the water molecules occupying the unliganded 5′- and 3′ cavities using the Hydration Site Analysis (HSA)[57,58] and found that unfavorable hydration and poor solute-water interactions contribute to the instability of these pockets and promote ligand binding. These insights into the binding site specificity may help inform the discovery and optimization of small molecules targeting the c-MYC G-quadruplex.

MATERIALS AND METHODS

Determination of Equilibrium Dissociation Constant K_d from SPR

The SPR measurements were performed with a four-channel BIAcore 2000 optical biosensor system (GE Healthcare). The oligonucleotide 5′-Biotin-d[$T_5TGAG_3TG_3TAG_3TG_3TA_2$] (Myc 14/23) was purchased from Sigma-Aldrich (Woodlands, TX). 5′-Biotin labeled DNA was immobilized onto streptavidin coated sensor chips (SA chip from BIAcore, GE Healthcare) by injection of DNA stock solution at 10 μg/ml with a flow rate of 10 μl/min for 30 minutes. One flow cell was used to immobilize the DNA oligomer sample, while an additional cell was left blank as a control. The total response unit (RU) for Myc 14/23 was 2193. The SPR experiments were performed at 25°C in filtered (0.22 μm) and degassed 10 mM HEPES buffer (pH 7.4) containing 100 mM KCl, 3 mM EDTA and 0.005% surfactant Polysorbate 20. Quindoline drug solutions (0.1, 0.2, 0.5, 1, 5, and 10 μM) were prepared by serial dilutions from stock solution and injected from 7 mm plastic vials with pierceable plastic crimp caps. Solutions of various drug concentrations were injected through the flow cells until a constant steady state response was obtained. Drug solution flow was then replaced by buffer flow resulting in dissociation of the complex. In general, a series of different drug concentrations were injected onto the chip with a flow rate of 30 μl/min for a period of 5 minutes followed by a dissociation period of 7 minutes. After each cycle, the chip surface was regenerated with 30 seconds injection of pH 2.5 glycine solution and multiple 1 minute buffer injections. The reference response from the blank cell was subtracted from the response in each cell containing DNA to give a signal (RU) that was directly proportional to the amount of bound compound. A set of sensorgrams

at different concentrations for binding of quindoline to DNA was obtained. The equilibrium constant was obtained from the fitting plots of RU versus free drug concentration. The equilibrium dissociation constant (K_d) was obtained using the 1:1 Langmuir binding model. The association (k_{on}) and dissociation (k_{off}) rate constants were fitted simultaneously using the following equation:

$$dR/dt = k_{on}C(R_{max} - R) - k_{off}R$$

where R represents the response unit, C is the concentration of quindoline. The experiments were done in triplicate.

Molecular dynamics simulation setup

The starting structure for the molecular dynamics simulations is the solution NMR structure of the 2:1 quindoline-c-MYC G-quadruplex (PDB entry 2L7V, Fig. 1B)[29]. The G-rich sequence (Myc 14/23) used in the NMR structure determination is the modified wild-type sequence with two G-to-T substitutions at the 14 and 23 positions (Fig. 1C). This sequence has been shown to adopt the single predominant c-MYC promoter G-quadruplex in K^+ solution and whose apo- and ligand bound structures were determined by NMR[24,29]. The 5′-flanking segment of the Myc 14/23 sequence is the same as that in the wild-type sequence. The T23 of the 3′-flanking segment TAA is a mutation from the wild-type G23 of the corresponding 3′-flanking GAA segment. As shown previously[24], both the mutant TAA and wild-type GAA 3′-flanking segments can form similar stable fold-back conformations. In this study, all the MD simulations were started from the first structure of the NMR ensemble, since the difference among the 10 structures in the NMR ensemble is small: the RMS deviations between structure 1 and all the rest of the structures in the NMR ensemble are below 1 Å, comparing with the average RMS fluctuation of ~ 2 Å observed in the MD simulation of the G-quadruplex. Before running the free energy simulations, we determine the pKa of the N1 atom of the quindoline to be ≈ 10.2 (Fig. 1A) using the QM Jaguar pKa program[59]. Therefore, at physiological pH, the N1 site is treated as protonated in the simulations.

In this work, the MD binding free energy simulations using the PMF and DDM approaches were performed using the GROMACS 4.6.4[60]. The AMBER parmbsc0 force field[61] is used to model the c-MYC G-quadruplex DNA in aqueous solutions. We have also performed MD simulations on the quindoline-cMYC G-quadruplex complex using the AMBER parmbsc1 force field[62]. Starting from the NMR structure of the complex, the ligand binding pockets are severely distorted in these simulations using the parmbsc1 force field. By contrast, the two ligand binding pockets remain stable in the MD simulations run with the parmbsc0 force field. Therefore, all the free energy simulations were carried out using the parmbsc0 force field. The quindoline ligand is modeled by the Amber GAFF parameters set[63] and the AM1-BCC charge model[64]. A truncated octahedral box containing TIP3P water [65] olecules previously equilibrated at 300 K and 1 atm pressure was used to solvate the quindoline-DNA complex. The dimension of the solvent box is set up to ensure that the distance between solute atoms from nearest walls of the box is at least 10 Å. K^+ ions are added to the solvent box to maintain charge neutrality. The Lennard-Jones parameters developed by Joung and

Cheatham are used to model the K^+ metal ion[66]. A recent report shows that with the use of the Åqvist parameter[67] for K^+ the K+ ions in the central ion channel of the G-tetrads can escape from the G-quadruplex, while such channel ion escape was not observed with the use of the Joung and Cheatham parameter[68]. The electrostatic interactions were computed using the particle-mesh Ewald (PME) method [69] with a real space cutoff of 10 Å and a grid spacing of 1.0 Å. MD simulations were performed in the NPT ensemble with a time step of 2 fs.

Using the PMF approach to calculate absolute binding free energy

The PMF method[38,39,70] uses a physical pathway connecting the bound and unbound states to compute the absolute binding free energy. We have derived a new formulation for the absolute binding free energy expression in the PMF approach: see Eq. (1), whose derivation is given in the Supporting Information. The computation of the $\Delta G^{\circ}_{\mathrm{bind}}$ involves the following steps: (1) apply harmonic restraint on the three Euler angles (Fig. S1) of the ligand in the bound state to restrain ligand orientation; (2) apply harmonic restraint on the polar and azimuth angles (Fig. S1) in the spherical coordinates to restrain the ligand center along a fixed axis when it binds/unbinds; (3) reversibly remove the ligand from the binding pocket along the chosen axis until it reaches the bulk region; (4) compute the free energy change when the restraints on the ligand center and ligand orientation are removed and the ligand is allowed to occupy the standard volume $1/C^{\circ}$ and rotate freely in the bulk region. The corresponding expression for the absolute binding free energy is

$$\Delta G^{\circ}_{bind} = -\Delta G^{bound}_{restr} - w(r^*) - k_B T \ln \frac{\int_{bound} e^{-w(r)/k_B T} dr}{(2\pi k_B T/k_r)^{\frac{1}{2}}} + \Delta G^{bulk}_{restr} \quad (1)$$

Here $-\Delta G^{bound}_{restr}$ is the free energy of restraining the ligand orientation in the bound state; $w(r)$ is the reversible work (potential of mean force or PMF) of pulling the ligand center away from the binding site to the current location r. r^* is an arbitrary location of the ligand center in the bulk region. The computed value of ΔG°_{bind} does not depends on the choice of the bulk location r^*[40]. $\int_{bound} e^{-w(r)/k_B T} dr$ is a 1D line integral which spans the entire region corresponding to the bound state. k_r is the force constant of the harmonic restraint which constrains the ligand to be in the proximity of the bulk location r^*. ΔG^{bulk}_{restr} is the free energy of applying the positional and orientational restraints on the ligand in the bulk solvent. The complete derivation and details of the PMF calculation are given in the SI.

Using the DDM to dissect absolute binding free energy into physically meaningful terms

In order to understand the molecular driving force underlying the binding site selectivity, we also performed DDM calculation[42–44,46], which can decompose the absolute binding free energy ΔG°_{bind} into contributions from electrostatic and nonpolar interactions (Supporting Information). The calculation involves two legs of alchemical decoupling simulations, in which a spatially restrained ligand is decoupled from the receptor binding pocket or from the bulk solution by gradually turning off the nonbond interactions (first Coulomb interaction

and then the Lennard-Jones interaction) between the ligand and its environment. The absolute binding free energy $\Delta G^{\circ}_{\mathrm{bind}}$ can be written as

$$\Delta G^{\circ}_{\mathrm{bind}} = -\Delta G^{coupled}_{restr} + \Delta G_{Coulomb} + \Delta G_{LJ} + \Delta G^{finite\_size}_{elec\_corr} + \Delta G^{gas}_{restr} \quad (2)$$

where $\Delta G_{Coulomb}$ and ΔG_{LJ} can be interpreted as the contributions of the effective electrostatic interactions and the nonpolar interactions to the total binding free energy, respectively[46,71]. The term $\Delta G^{finite\_size}_{elec\_corr}$ is the correction that accounts for the effects of finite-size, periodic simulation system on the calculated electrostatic free energy of decoupling a charged ligand. Here, $\Delta G^{finite\_size}_{elec\_corr}$ is estimated using a scheme developed by Rocklin and coworkers.[54] The remaining terms $\Delta G^{coupled}_{restr}$ and ΔG^{gas}_{restr} are the free energy of spatially restraining the ligand to remain in the binding site for a fully coupled ligand and that for a fully decoupled ligand, respectively (see Supporting Information for more details of the DDM and its binding free energy decomposition scheme).

Hydration Sites Analysis

To understand the origin of the instability of the binding pocket in the apo G-quadruplex, we performed Hydration Sites Analysis (HSA)[58,72] to compute the energetics of the water molecules occupying the unliganded 5′- and 3′ cavities. The HSA calculations involved three steps: (1) molecular dynamics simulations to generate water molecule distributions in the binding cavity, (2) cluster analysis of these water molecule distributions to isolate high density regions of water molecules (i.e. hydration sites) and (3) energy calculations of the hydration sites using inhomogeneous solvation theory. Additional structural properties such as enclosure of the hydration site and hydrogen bonds to water molecules in the hydration site allow for a more complete understanding of the solvation properties of these two cavities. In this work, we performed HSA analysis on a subset of 10,000 frames from a 100 ns production trajectory. High density spherical regions (hydration sites) of 1Å radius were identified using a clustering procedure[58] on the water molecules that were found within 3 Å of the heavy atoms of the quindoline rings in an aligned holo conformation. The resulting hydration sites were each populated by retrieving all water molecules, which had oxygen atoms within 1.0 Å from the corresponding hydration site center. The hydration sites were then enumerated according to their occupancies, with the highest populated site given the index 0.

The energy of each hydration site (E_{total}) is calculated as a function of the pairwise solute-water energies (E_{sw}) and water-water energies (E_{ww})[58],

$$E_{tot} = \frac{1}{2}(E_{sw} + E_{ww})$$

where a factor of one-half takes into account the assignment of pairwise interactions to each molecule in the pair.

Hydrogen bonds were identified based on the following geometric criteria: the donor–acceptor heavy atom distance is less than or equal to 3.5 Å and the hydrogen–donor–acceptor angle is less than or equal to 30°[72]. The total number of hydrogen (HB_{tot}) is equal to sum of the solute-water (HBsw) and water-water hydrogen bonds (HBww).

Fractional enclosure (f_{enc}) was calculated based on the following function:[72]

$$f_{enc}=1-\frac{N_{nbr}}{N_{nbr-bulk}}$$

where N_{nbr} is mean number of water molecules found in the first hydration shell of a hydration site or first shell neighbors and $N_{nbr\text{-}bulk}$ is mean number of first-shell neighbors of a TIP3P water molecule in neat water simulation as defined in this previous work. This quantity indicates the degree to which the water in a hydration site is blocked from contact with other water molecules.

More details of the MD simulations used to perform the HSA are given in the Supporting Information.

RESULTS

Binding affinity determined by SPR measurements indicates that quindoline binds the c-MYC G-quadruplex tightly with a slow k_{on}

SPR is used to experimentally determine the equilibrium binding constant K_a of quindoline to Myc 14/23 d[$TGAG_3TG_3TAG_3TG_3TA_2$][24,29] from the direct measurements of complex association (k_{on}) and dissociation (k_{off}) rates in real time. Biotinylated oligonucleotide was immobilized on a streptavidin-coated sensor chip. Increasing concentrations of quindoline (0.1, 0.2, 0.5, 1, 5, and 10 μM) were injected and the interaction of the quindoline with the oligonucleotides was measured. Binding between quindoline and MycG4 was observed at quindoline concentration as low as 0.1 μM (Fig. 2). The intensity of the observed response increased with increasing concentrations of quindoline in a dose dependent manner. By fitting the sensorgrams using a 1:1 Langmuir model, the association constant K_a for Myc 14/23 was calculated as $3.27\times10^{6}M^{-1}$. Directly measured k_{on} and k_{off} are 2.57×10^{4} ± 140 M^{-1} s^{-1} and 7.87×10^{-3} ± 2.91×10^{-5} s^{-1}, respectively (Table 1). It should be noted that the on-rate k_{on} is much slower than the diffusion-limited on-rate for bimolecular binding, which is on the order of ~10^{9} M^{-1} s^{-1}. The slow binding of quindoline to the c-MYC G-quadruplex is likely to reflect the ligand-induced conformational rearrangement near the binding sites, as shown by computation simulations described later.

Absolute binding free energies for the two binding sites from PMF approach suggest that the quindoline binds the 5′ end with a slightly higher affinity

Using the NMR structure of the 2:1 quindoline-c-MYC G-quadruplex[29] (Fig. 1B) as the starting point for MD simulation, we computed the absolute binding free energies $\Delta G^{\circ}_{\mathrm{bind}}$ for quindoline molecules bound at the 5′ and 3′ sites: see Table 2.

The $\Delta G^{\circ}_{\mathrm{bind}}$ calculated from the PMF for the 5′ and 3′ binding sites are −8.8 (±0.6) kcal/mol and −7.7 (±0.7) kcal/mol respectively, in good agreement with the effective ΔG°_{bind} of −8.94 kcal/mol determined from SPR. The agreement with the experiment supports the free energy model used in this study and allows us to further analyze the binding site specificity in the c-MYC G-quadruplex.

The calculated binding free energies shown in Table 2 suggest that the quindoline compound has a slight preference for the 5′ pocket. Later on we will compare this result with previous NMR results and with the DDM calculations. To gain insights into the nature of the quindoline binding at the two sites, we examine the different free energy contributions in the PMF calculation of ΔG°_{bind} (Table 3). These include: ΔG^{bulk}_{restr}, the free energy associated with restraining the position and orientation of the ligand to a bulk location r^*; $-w(r^*)$, the reversible work of physically moving the orientationally restrained ligand from the bulk into the center of the binding site; and $-\Delta G^{bound}_{restr}$ the free energy of releasing all the restraints in the bound state. Table 3 shows that all these three terms make important contributions to the final ΔG°_{bind}; the highly attractive PMF term $-w(r^*)$ is largely offset by ΔG^{bulk}_{restr}, the free energy cost of spatially and orientationally restraining the unbound ligand to the bulk location.

PMF results show that both binding pockets at the 5′- and 3′-ends undergo significant rearrangement as the ligand unbinds

To investigate the effect of ligand binding on the conformation and dynamics of the G-quadruplex, we examine the intermediate structures of the ligand-DNA complex observed as the ligand is pulled out of the binding pocket, together with the calculated free energy profile $w(r)$ as a function of the ligandbinding site distance r (see Fig. 3 and Fig. 4). The $w(r)$ curves measure the energy expended to pull the ligand out of the binding pocket. For both binding sites, the free energy minimum in $w(r)$ coincides with the ligand-binding site distance found in the NMR structure of the quindoline-DNA complex[29]. From the free energy minimum, the $w(r)$ increases steadily as the ligand leaves the binding pocket; the curve levels off after a characteristic distance (≈ 19 Å for the 3′ pocket, and ≈ 20 Å for the 5′ pocket), where the quindoline fused ring completely loses its stacking interaction with the terminal G-tetrads and enters the bulk solution.

An important observation from the structures along the ligand unbinding pathway is that both binding pockets undergo significant rearrangement as the ligand unbinds (Fig. 5). For the 5′ binding pocket, in the ligand bound structure (Fig. 5A), the flanking bases of G5, A6 and the 5′-end G-tetrad together form a concave binding pocket, which enables the base ring of G5 to stack over and wrap around the plane of the quindoline fused ring; as the ligand unbinds from the 5′ pocket (r > 20 Å), the 5′ flanking residues undergo large fluctuations and adopt a variety of structures. Fig. 5B and Fig. 5C show two representative arrangements in which the drug-induced binding pocket at the 5′-end collapses after unbinding of the ligand. In Fig. 5B, the G5 base ring stacks over the A6; in Fig. 5C, the base rings of both G5 and A6 stack on the 5′ end G-tetrad plane. The collapse of the binding cavity caused by ligand unbinding observed here from MD simulation is consistent with the NMR structures

of the free c-MYC G-quadruplex and its quindoline complexes: in the complex structures the 5′ binding pocket only exists in the ligand bound DNA G-quadruplex due to the ligand-induced orientation of the flanking bases of G5 and A6[24,29]. Significant conformational change of the flanking residues due to ligand unbinding is also seen in the simulated structures of the 3′ pocket: as seen from the Fig. 5D, in the ligand-bound state, the A25 base ring is at an angle of 60° relative to the T23 base; as the ligand unbinds, the base ring of A25 becomes coplanar with respect to that of T23: see Fig. 5E and 5F. Comparing the NMR structures of the apo and ligand-bound G-quadruplex[24,29], we find similar conformational rearrangements at the 3′-end. Therefore, it appears that the MD simulations used to compute PMF capture the key conformational reorganization in the two binding pockets. It should be noted that the free energy change associated with ligand induced binding pocket reorganization is an important component of the absolute binding free energy. This free energy of reorganization is implicitly contained in the PMF function $w(r)$. To understand how $w(r)$ implicitly contains the information of the receptor binding site reorganization, we consider for example what happens to the $w(r)$ during the collapse of the 5′ pocket upon ligand unbinding: when the ligand is about to leave the binding pocket, suppose that the pocket residues are rigid, then the ligand will still experience the remaining attractive forces which will cause the $w(r)$ to continue to rise; in reality however, as the ligand is about to leave the pocket the pocket will undergo conformational change that will "push" the ligand out such that the pocket will reach the collapsed state as soon as the ligand is completely out. This tendency of the ligand being "pushed out" by the pocket when the ligand is about to leave the pocket helps cancel out some of the remaining attractive forces from the pocket and effectively lowers the PMF $w(r)$. It should be noted that while this free energy of the pocket reorganization is accounted for here by the PMF function $w(r)$ generated by MD, such receptor reorganization effects are difficult to be treated properly using the more rapid methods such as docking or MM-PB(GB)/SA[53].

PMF results reveal differences in the 5′ and 3′ complexes

Comparisons of the PMF curves shown in Figures 3 and Fig. 4 reveal that, while the increase in $w(r)$ for the 5′ complex is smooth, the corresponding $w(r)$ for the 3′ complex suggests a two-step process: starting from the free energy minimum at r = 12.6 Å, w(r) rises relatively sharply within a short range from r = 12.6 Å to 14.2 Å, which corresponds to the breaking of the intermolecular hydrogen bond between the protonated N1 of the quindoline and O4 atom of the base of the flanking T23 at the 3′-end (Fig. 4; see also Fig. S2 in the Supporting Information); thereafter $w(r)$ rises somewhat less rapidly, which reflects the gradual loss of the quindoline ring stacking with the 3′-end G-tetrad plane with increasing r. The direct contribution of the intermolecular H-bond between the quindoline-N1 and T23-O4 to the binding free energy of the 3′ complex is therefore estimated to be ~ −4 kcal/mol. By comparison, no intermolecular H-bond is formed in the 5′ complex, and the rise in the $w(r)$ is attributable to the loss of the hydrophobic enclosure of the quindoline ring in the 5′ pocket formed by the 5′-end G-tetrad plane and the flanking G5 base (Fig. 3 and Fig. 5A). While the intermolecular hydrogen bond formation in the 3′ complex was suggested by NMR structural study[29], it was not directly observed in NMR experiments; the result from the PMF calculations confirms the differences between the 5′ and 3′ complexes and the intermolecular hydrogen bond formation between quindoline and the 3′ end DNA.

Decomposition of absolute binding free energy into physically meaningful contributions by DDM

To understand the physical reasons for the difference in the binding affinities of quindoline at the two binding sites (Table 2), we decompose the absolute binding free energy into physically meaningful components using the DDM method[71]. In this work, we used two different DDM setups involving different thermodynamic states for the fully decoupled ligand to compute ΔG°_{bind}. We first discuss the results obtained using the DDM setup in which all the intramolecular nonbonded interactions are turned off in the decoupled ligand (i.e. as the ligand is decoupled intermolecularly from the environment, its intramolecular noncovalent interaction is also turned off). The calculated ΔG°_{bind} for the 5′-end and 3′- end complexes are −11.2 (±1.3) kcal/mol and −9.6 (±0.6) kcal/mol, respectively (Table 2). This result is in general agreement with the experimental ΔG°_{bind} of −8.94 kcal/mol to within ~ 2 kcal/mol. Consistent with the binding free energies estimated from PMF, the $\Delta G^{\circ}_{\mathrm{bind}}$ calculated using this DDM setup also suggests that the quindoline binding at the 5′ site is somewhat stronger than that at the 3′ site.

Next, we examine the contributions to the ΔG°_{bind} computed by DDM. In DDM, the $\Delta G^{\circ}_{\mathrm{bind}}$ is obtained as the sum of several terms including the effective nonpolar free energy contribution δG_{LJ} and the effective electrostatic free energy contribution $\delta G_{Coulomb}$ (see Methods). Table 4 shows the free energy decompositions obtained using the DDM setup in which the nonbonded intramolecular interactions in the decoupled ligand are turned off. The difference in the electrostatic free energy contributions $\Delta G_{Coulomb}$ actually favors the binding at the 3′-end by about −1.7 kcal/mol, which likely reflects the QuiN1H-T23O4 intermolecular hydrogen bond present at the 3′ site (Fig. 1B, Ref[29]). However, the more favorable $\Delta G_{Coulomb}$ is outweighed by the larger difference in the nonpolar free energy contribution δG_{LJ} which favors the binding at the 5′ site by about −3.3 kcal/mol; the net result is that the 5′ complex is more stable than the 3′ complex by about −1.6 kcal/mol. Table 4 also shows that the corrections for the effect of the finite-size periodic simulation system on the electrostatic free energy are on the order of ~0.8 kcal/mol.

We have also computed binding free energies using a DDM setup in which the ligand intramolecular nonbonded interaction is left unchanged when the ligand is decoupled intermolecularly from its environment (Table 2 and Table 5). Here, the estimated ΔG°_{bind} for the binding at the 5′-end and 3′-end are −12.4 (±1.7) kcal/mol and −7.5 (±0.3) kcal/mol, respectively, which deviate from the experimental ΔG°_{bind} = −8.94 kcal/mol by 1.5–3.5 kcal/mol. This DDM result also shows that the quindoline binding at the 5′-end is stronger than that at the 3′-end. Qualitatively similar results on the relative importance of the electrostatic contribution $\Delta G_{Coulomb}$ and nonpolar contribution ΔG_{LJ} are also obtained (Table 5).

The slight preference of quindoline for the 5′ site is likely related to its greater hydrophobic enclosure

As seen in Table 4 and Table 5, the difference in the binding free energy ΔG°_{bind} for the 5′ and 3′ sites is dominated by δG_{LJ}. This term reflects the effective nonpolar interaction

between the ligand and the DNA binding site, which includes contributions from several thermodynamic processes, including the desolvation of the ligand from the bulk solvent, the displacement of the binding cavity water molecules by the ligand, and the direct intermolecular Lennard-Jones interactions. To understand qualitatively why the effective nonpolar ligand-DNA interaction is stronger at the 5′ site, we examine the difference in the geometries of the two binding pockets (Fig. 6). At the 5′ end the flanking residues G5 and A6 and the 5′-end G-tetrad arrange themselves to form the cavity which wraps around the quindoline ring, which results in an intercalation type of binding. Binding interfaces displaying such hydrophobic enclosure characters are known to contribute additional affinity to binding[73], because (a) the enclosed ligand interact with receptor atoms on both sides of the binding cavity and (b) the water molecules occupying the unliganded cavity are thermodynamically less stable since they cannot form extensive hydrogen bonds with the neighboring water molecules or solute[57]. Importantly, comparing with the 5′ cavity, the extent of hydrophobic enclosure in the 3′ cavity is more limited (Fig. 6). The magnitude of the effect of hydrophobic enclosure to binding free energy can be estimated using a simple surface area model: from Fig. 6, the 5′ pocket has a larger buried van der Waals surface area than that at the 3′ site by $\Delta A \approx 26$ Å^2; Using a surface tension coefficient $\gamma = 0.072$ kcal mol^{-1} Å^2 in the surface area model for binding free energy, $\Delta G = -\gamma \Delta A$[74,75], the hydrophobic enclosure would favor the binding at the 5′ site by −1.8 kcal/mol. This value is understandably smaller than the difference of −3.3 kcal/mol in the effective nonpolar contribution ΔG_{LJ} estimated from DDM, since ΔG_{LJ} also includes the contribution from direct intermolecular Lennard-Jones interactions, which also favors the binding at the 5′ site. The above analysis suggests that residue G5 is crucial for ligand binding at the 5′ site (Fig. 6). Importantly, the critical role played by the G5 in the specific binding of quindoline was borne out by the experimental mutational analysis, which shows that the G-to-T mutation at position 5 markedly weakened the quindoline binding[29]. In contrast, the quindoline binding at the 3′-end relies less on capping interactions as shown by mutational analysis[29].

The results shown in Table 4 and Table 5 indicate that the errors in the $\Delta \mathrm{G}^{\circ}_{bind}$ calculated using the DDM setup without turning off intramolecular nonbonded interactions in the decoupled ligand are larger than those obtained with the intramolecular nonbonded interactions turned off in the decoupled ligand. The larger absolute error in the $\Delta \mathrm{G}^{\circ}_{bind}$ is likely related to conformational sampling problems with the decoupled ligand in which strong, unscreened intramolecular electrostatic forces lead to kinetically trapped states. Note that the decoupled states of the gas-phase ligand appear in both $Z_{R_{wat}\cdots L_{gas}}$ and $Z_{L_{gas}}$ in Eq. (9.2) and Eq. (9.4) (Supporting Information); in the limit of infinite sampling time, the two setups should yield the identical result. However, in practice, where the simulation time is finite, we found that the DDM setup with the nonbonded intramolecular interactions turned off in the decoupled ligand yielded more accurate results. By comparing the conformations of the fully decoupled ligands sampled by the decoupling simulations, we found that retaining the intramolecular interactions in the decoupled ligand resulted in sampling problems as the ligand can be kinetically trapped in more compact conformations because of the unscreened intramolecular electrostatic interactions. One representative structure of such kinetically trapped conformations sampled during the decoupling simulations is shown Fig.

7A. While such compact conformations are over-represented in the simulations that decouple the ligand from the bulk solvent, they are underrepresented in the simulations that decouple the ligand from the DNA binding site. This reflects conformational sampling problems caused by strong ligand intramolecular interactions. In the DDM setup with the intramolecular nonbonded interactions turned off, such conformational sampling problems with the decoupled ligand states are avoided (Fig. 7B). As a result, for the quindoline derivative which carries a flexible and highly polar side chain, turning off the intramolecular interactions in the decoupled ligand leads to better ΔG°_{bind} estimates from DDM.

Energetically unfavorable water molecules may contribute to the instability of the binding pockets in the apo G-quadruplex

To help understand the reason why the binding pockets are unstable in the free G-quadruplex, we analyzed the energetics of the water molecules that would occupy the unliganded binding cavities. Using the HSA method[58,72], we have identified thermodynamically favorable and unfavorable water molecules located inside the both binding pockets based on an energetic cutoff of −9.53 kcal/mol, the energy of a TIP3P[65] bulk water molecule (see Methods, Fig. 8). The 5′ cavity contains two energetically unfavorable hydration sites in the back of the cavity while the 3′ cavity contains three energetically unfavorable hydration sites in the shallow groove of this binding cavity (Table S1 and S2, Supporting Information). In these cases, the energetically unfavorable water molecules which are located in partially enclosed regions of the binding site are unable to simultaneously maintain favorable energetic contacts with both the solute surface and neighboring water molecules. In addition, many energetically favorable water molecules in both cavities are only slightly more favorable than bulk water and make poor interactions with the solute surface as reflected in small number of solute-water hydrogen bonds and overall high solute-water energies (Tables S1 and S2). The energetics of the cavity water molecules, therefore, provide a physical explanation for the role of water displacement in binding and the reorganization upon unbinding because the weak solute-water interactions can be readily compensated either by the strong direct ligand-cavity interactions or the favorable intramolecular collapse of the binding pockets in the absence of the ligand, which is observed in both the NMR study and our computation simulations.

DISCUSSIONS

Most of the previous computational studies of small molecule binding with nucleic acids have used the end-point method MM-PB(GB)/SA (molecular mechanics-Poisson-Boltzmann/generalized Born/surface area)[31,76–78]. While these studies have provided insights into the energetics of ligand-DNA binding, the MM-PB(GB)/SA method has limited accuracy due to its intrinsic approximation of energy contributions from receptor conformational change and solvent effect[79,80]. In this study we show that using statistical mechanics-based methods DDM and PMF, the absolute binding free energy of a potent ligand of the c-MYC G-quadruplex DNA can be computed reasonably accurately to within ~ 2 kcal/mol from the experiment (Table 2). The accuracy in the calculated absolute binding free energy obtained here is similar to that reported recently by Moraca et al. who applied metadynamics to study the binding of alkaloid berberine to a human telomeric G-

quadruplex[81]. The results from that study and from the present work demonstrate the potential utility of statistical mechanics based absolute binding free energy methods in characterizing the thermodynamics of ligand-DNA binding. Comparing with MM-PB(GB)/SA, the improvement from these molecular dynamics based methods can be attributed to two key factors: first, the effects of binding pocket reorganization induced by the ligand binding/unbinding, such as those shown in Fig. 5, can be captured by MD-based free energy methods; second, the effect of structured water molecules on binding can be accounted for naturally by these explicit solvent based methods. We are testing whether the promising results reported here can be extended broadly to more diverse ligand-DNA complexes.

According to both PMF and DDM calculations, the 5′ site binds the quindoline with slightly higher affinities compared with the 3′ site (Table 2). We compare this computational result with NMR experiments on the 2:1 quindoline-G-quadruplex[29]. In pH 7 and 100 mM K^+ solution, the NMR spectrum for the 3′-flanking segment T23, A24 and A25 are not well resolved compared with the 5′ binding site, indicating that the 3′-end are more flexible and may exist in multiple conformations. This suggests that under the physiological conditions the quindoline-G-quadruplex complex at the 5′-end is more stable than the 3′-end complex, which is consistent with the computational result that quindoline shows a slight preference for binding at the 5′-end. Furthermore, the DDM calculations provide a plausible explanation for this binding site specificity. As seen from Table 4 and Table 5, the electrostatic free energy contribution actually favors the binding at the 3′-end, which likely reflects the intermolecular hydrogen bond formed at the 3′ site. However, this is outweighed by the nonpolar free energy contribution which much more significantly favors the binding at the 5′ site, which is related to a greater hydrophobic enclosure at the 5′ end. Consequently, the 5′ complex is slightly more stable than the 3′ complex. A very recent report on another quindoline derivative 4bm with an aromatic ring-containing side-chain showed an even stronger binding preference at the 5′ end of the c-MYC G-quadruplex[82]. This result could be nicely explained by our computational study: first, the aromatic ring-containing side chain of the 4bm is likely to favor the nonpolar interactions at the 5′-end even more than the quindoline molecule studied in this work does: compare Fig. 9 with Fig. 1A; second, because the 4bm features a CH_3 group attached to the positively charged nitrogen atom in the quindoline ring, it is unable to form intermolecular hydrogen bond with the O4 atom of the T23 at the 3′-end, which would further reduce the binding at the 3′ end.

In this work the PMF calculations suggest that the unbinding of the ligand causes major conformational changes in the flanking residues at both the 5′- and 3′-ends, resulting in the collapse of the binding cavities in the free c-MYC G-quadruplex (Fig. 5). This observation from the simulation supports the notion that the binding pockets in the G-quadruplex are not preformed but induced by the ligand, a result consistent with the solution NMR structures of the apo and holo G-quadruplex[29]. The observation of the ligand-induced conformational rearrangements of the flanking residues forming the binding pockets is also supported by the binding kinetics measurements: the on rate of association k_{on} observed in SPR is several orders of magnitude slower than the diffusion-limiting values (Table 1). This could be caused by the kinetic barrier encountered by the incoming ligand to induce the formation of the binding pocket.

The quindoline-induced conformational rearrangement in the flanking segments observed in both the computer simulations and previous NMR structures provide important considerations for the biologically important parallel G-quadruplexes in gene promoters. As shown in the complex structures by NMR[29] and by computational simulations here, for the induced drug binding pocket to form, an appropriate single-stranded flanking segment containing at least two bases must exist. There is clear evidence of extensive single-stranded regions at either side of the G-quadruplex in a duplex element under negative supercoiled conditions[83]. The inherent conformational flexibility of the DNA backbone in the single-stranded region flanking the G-quadruplex in gene promoters would likely provide an environment for the observed drug binding pockets to form.

Intriguingly, the calculated binding affinity of 0.39 μM (converted from the PMF-calculated binding free energy of −8.8 kcal/mol, see Table 2) is in close agreement with the IC_{50} of 0.3 μM of the inhibition of the expression of c-MYC in the cancer cell line by the same ligand[26]. While IC_{50} is not a direct indicator of binding affinity, the two are closely related[84]. Since the conformational reorganization of the flanking segments is crucial to the binding pocket formation, the fact that our calculated binding affinity correlates well with the experimental IC_{50} of inhibition in the cell line assay provides an additional validation for the computational results, and it supports the view that the observed conformational dynamics likely plays a similar role in the flanking regions around the G-quadruplex formed in the c-MYC promoter. Further evidence that the binding pockets formed by the flanking segments of the c-MYC G-quadruplex are relevant drug targets comes from recent studies[31,33] of in silico virtual screening targeting the NMR structures of the c-MYC G-quadruplex[24,29]. In these studies candidate ligands are docked into the binding sites formed by the flanking residues of the c-MYC G-quadruplex molecule from the NMR structures of the drug complexes. These modeling work have either led to the discovery of G-quadruplex-selective binders with micromolar activities[31], or helped rationalize the binding mechanism of new active compounds[33].

The results of this work have implications for structure-based rational design targeting the c-MYC G-quadruplex. First, our results show that it is possible to obtain accurate absolute binding free energies for binding of charged ligand with DNA using the statistical mechanics-based methods DDM and PMF. These more detailed methods, although slower than docking and MM-PB(GB)/SA, can provide greater accuracy and could be employed in the post-docking refinement stage to more reliably identify true binders to help prioritize synthetic chemistry in drug design. Second, the results suggest that absolute binding free energy calculations can shed lights on binding site selectivity, which is difficult to obtain by experimental methods such as ITC and SPR. Such calculations can be useful for discovering ligands that selectively bind G-quadruplex relative to duplex DNA[31]. Third, the finding that the 5′ end flanking segment of the c-Myc G-quadruplex can rearrange itself to form more hydrophobically enclosed binding pocket suggests that the 5′ site could be the primary focus of in silico virtual screening of small molecule libraries. Furthermore, the observation that ligand binding induces conformational rearrangement in the flanking residues of the G-quadruplex suggests that an in silico virtual screening should optimally target the holo structures of the G-quadruplex generated from either NMR or MD simulation, instead of targeting the apo structure of the c-MYC G-quadruplex[30,31].

CONCLUSIONS

We report absolute binding free energy calculations on the quindoline-c-MYC G-quadruplex complex using two statistical mechanics-based free energy simulation methods DDM and PMF. The study shows that small molecule binding affinities at two different binding sites within one DNA G-quadruplex structure can be computed in good agreement with the experiment. While equilibrium binding constants can be determined by experimental methods such as ITC and SPR, these methods can not directly measure the binding affinities at different binding sites. Our study also provides a plausible explanation for the binding site selectivity and the ligand-induced conformational reorganization in the c-MYC G-quadruplex. These insights could be important for structure-based rational design targeting the c-MYC G-quadruplex. The methods described in this study can be applied to other DNA or RNA systems to aid in the interpretation of experiments and inform rational drug discovery targeting nucleic acids.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This research was supported by grants from the National Institutes of Health (CA177585 (DY)). We gratefully acknowledge Steven Ramsey and Tom Kurtzman for providing helpful discussions.

References

1. Yang D, Okamoto K. Structural Insights into G-Quadruplexes: Towards New Anticancer Drugs. Future Med Chem. 2010; 2(4):619–646. [PubMed: 20563318]
2. Balasubramanian S, Hurley LH, Neidle S. Targeting G-Quadruplexes in Gene Promoters: A Novel Anticancer Strategy? Nat Rev Drug Discov. 2011; 10(4):261–275. [PubMed: 21455236]
3. Neidle S. Quadruplex Nucleic Acids as Novel Therapeutic Targets. J Med Chem. 2016; 59(13): 5987–6011. [PubMed: 26840940]
4. Fadrná E, Špačková N, Štefl R, Koča J, Cheatham TE, Šponer J. Molecular Dynamics Simulations of Guanine Quadruplex Loops: Advances and Force Field Limitations. Biophys J. 2004; 87(1):227–242. [PubMed: 15240460]
5. Šponer J, Cang X, Cheatham TE. Molecular Dynamics Simulations of G-DNA and Perspectives on the Simulation of Nucleic Acid Structures. Methods. 2012; 57(1):25–39. [PubMed: 22525788]
6. Qin Y, Hurley LH. Structures, Folding Patterns, and Functions of Intramolecular DNA G-Quadruplexes Found in Eukaryotic Promoter Regions. Biochimie. 2008; 90(8):1149–1171. [PubMed: 18355457]
7. Brooks TA, Hurley LH. The Role of Supercoiling in Transcriptional Control of MYC and Its Importance in Molecular Therapeutics. Nat Rev Cancer. 2009; 9(12):949.
8. Simonsson T. DNA Tetraplex Formation in the Control Region of C-Myc. Nucleic Acids Res. 1998; 26(5):1167–1172. [PubMed: 9469822]
9. Siddiqui-Jain A, Grand CL, Bearss DJ, Hurley LH. 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. 2002; 99(18):11593–11598. [PubMed: 12195017]
10. Sun D. 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. 2005; 33(18):6070–6080. [PubMed: 16239639]

11. Agrawal P, Hatzakis E, Guo K, Carver M, Yang D. Solution Structure of the Major G-Quadruplex Formed in the Human VEGF Promoter in K+: Insights into Loop Interactions of the Parallel G-Quadruplexes. Nucleic Acids Res. 2013; 41(22):10584–10592. [PubMed: 24005038]
12. De Armond R, Wood S, Sun D, Hurley LH, Ebbinghaus SW. Evidence for the Presence of a Guanine Quadruplex Forming Region within a Polypurine Tract of the Hypoxia Inducible Factor 1α Promoter†. Biochemistry (Mosc). 2005; 44(49):16341–16350.
13. Dai J, Dexheimer TS, Chen D, Carver M, Ambrus A, Jones RA, Yang D. 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. 2006; 128(4):1096–1098. [PubMed: 16433524]
14. Dexheimer TS, Sun D, Hurley LH. Deconvoluting the Structural and Drug-Recognition Complexity of the G-Quadruplex-Forming Region Upstream of the Bcl-2 P1 Promoter. J Am Chem Soc. 2006; 128(16):5404–5415. [PubMed: 16620112]
15. Agrawal P, Lin C, Mathad RI, Carver M, Yang D. The Major G-Quadruplex Formed in the Human BCL-2 Proximal Promoter Adopts a Parallel Structure with a 13-Nt Loop in K^+ Solution. J Am Chem Soc. 2014; 136(5):1750–1753. [PubMed: 24450880]
16. Cogoi S, Xodo LE. G-Quadruplex Formation within the Promoter of the KRAS Proto-Oncogene and Its Effect on Transcription. Nucleic Acids Res. 2006; 34(9):2536–2549. [PubMed: 16687659]
17. Xu Y, Sugiyama H. Formation of the G-Quadruplex and i-Motif Structures in Retinoblastoma Susceptibility Genes (Rb). Nucleic Acids Res. 2006; 34(3):949–954. [PubMed: 16464825]
18. Rankin S, Reszka AP, Huppert J, Zloh M, Parkinson GN, Todd AK, Ladame S, Balasubramanian S, Neidle S. Putative DNA Quadruplex Formation within the Human C-Kit Oncogene. J Am Chem Soc. 2005; 127(30):10584–10589. [PubMed: 16045346]
19. Fernando H, Reszka AP, Huppert J, Ladame S, Rankin S, Venkitaraman AR, Neidle S, Balasubramanian S. A Conserved Quadruplex Motif Located in a Transcription Activation Site of the Human C-Kit Oncogene. Biochemistry (Mosc). 2006; 45(25):7854–7860.
20. Guo K, Pourpak A, Beetz-Rogers K, Gokhale V, Sun D, Hurley LH. Formation of Pseudosymmetrical G-Quadruplex and i-Motif Structures in the Proximal Promoter Region of the RET Oncogene. J Am Chem Soc. 2007; 129(33):10220–10228. [PubMed: 17672459]
21. Palumbo SL, Ebbinghaus SW, Hurley LH. Formation of a Unique End-to-End Stacked Pair of G-Quadruplexes in the HTERT Core Promoter with Implications for Inhibition of Telomerase by G-Quadruplex-Interactive Ligands. J Am Chem Soc. 2009; 131(31):10878–10891. [PubMed: 19601575]
22. Qin Y, Fortin JS, Tye D, Gleason-Guzman M, Brooks TA, Hurley LH. Molecular Cloning of the Human Platelet-Derived Growth Factor Receptor Beta (PDGFR-Beta) Promoter and Drug Targeting of the G-Quadruplex-Forming Region to Repress PDGFR-Beta Expression. Biochemistry (Mosc). 2010; 49(19):4208–4219.
23. Brown RV, Wang T, Chappeta VR, Wu G, Onel B, Chawla R, Quijada H, Camp SM, Chiang ET, Lassiter QR, et al. The Consequences of Overlapping G-Quadruplexes and i-Motifs in the Platelet-Derived Growth Factor Receptor β Core Promoter Nuclease Hypersensitive Element Can Explain the Unexpected Effects of Mutations and Provide Opportunities for Selective Targeting of Both Structures by Small Molecules To Downregulate Gene Expression. J Am Chem Soc. 2017; 139(22):7456–7475. [PubMed: 28471683]
24. Ambrus A, Chen D, Dai J, Jones RA, Yang D. Solution Structure of the Biologically Relevant G-Quadruplex Element in the Human c-MYC Promoter. Implications for G-Quadruplex Stabilization†. Biochemistry (Mosc). 2005; 44(6):2048–2058.
25. Chen B-J, Wu Y-L, Tanaka Y, Zhang W. Small Molecules Targeting C-Myc Oncogene: Promising Anti-Cancer Therapeutics. Int J Biol Sci. 2014; 10(10):1084–1096. [PubMed: 25332683]
26. Ou T-M, Lu Y-J, Zhang C, Huang Z-S, Wang X-D, Tan J-H, Chen Y, Ma D-L, Wong K-Y, Tang JC-O, et al. Stabilization of G-Quadruplex DNA and Down-Regulation of Oncogene *c – Myc* by Quindoline Derivatives. J Med Chem. 2007; 50(7):1465–1474. [PubMed: 17346034]
27. Ou T-M, Lin J, Lu Y-J, Hou J-Q, Tan J-H, Chen S-H, Li Z, Li Y-P, Li D, Gu L-Q, et al. Inhibition of Cell Proliferation by Quindoline Derivative (SYUIQ-05) through Its Preferential Interaction with *c – Myc* Promoter G-Quadruplex. J Med Chem. 2011; 54(16):5671–5679. [PubMed: 21774525]

28. Phan AT, Kuryavyi V, Gaw HY, Patel DJ. Small-Molecule Interaction with a Five-Guanine-Tract G-Quadruplex Structure from the Human MYC Promoter. Nat Chem Biol. 2005; 1(3):167–173. [PubMed: 16408022]
29. Dai J, Carver M, Hurley LH, Yang D. Solution Structure of a 2:1 Quindoline–c-MYC G-Quadruplex: Insights into G-Quadruplex-Interactive Small Molecule Drug Design. J Am Chem Soc. 2011; 133(44):17673–17680. [PubMed: 21967482]
30. Kang H-J, Park H-J. In Silico Identification of Novel Ligands for G-Quadruplex in the c-MYC Promoter. J Comput Aided Mol Des. 2015; 29(4):339–348. [PubMed: 25527072]
31. Hou J-Q, Chen S-B, Zan L-P, Ou T-M, Tan J-H, Luyt LG, Huang Z-S. Identification of a Selective G-Quadruplex DNA Binder Using a Multistep Virtual Screening Approach. Chem Commun. 2015; 51(1):198–201.
32. Rocca R, Costa G, Artese A, Parrotta L, Ortuso F, Maccioni E, Pinato O, Greco ML, Sissi C, Alcaro S, et al. Hit Identification of a Novel Dual Binder for *h-Telo/c-Myc* G-Quadruplex by a Combination of Pharmacophore Structure-Based Virtual Screening and Docking Refinement. ChemMedChem. 2016; 11(16):1721–1733. [PubMed: 27008476]
33. Zeng DY, Kuang GT, Wang SK, Peng W, Lin SL, Zhang Q, Su XX, Hu MH, Wang H, Tan JH, et al. Discovery of Novel 11-Triazole Substituted Benzofuro[3,2-*b*]Quinolone Derivatives as *c-Myc* G-Quadruplex Specific Stabilizers via Click Chemistry. J Med Chem. 2017
34. Deng Y, Roux B. Computations of Standard Binding Free Energies with Molecular Dynamics Simulations. J Phys Chem B. 2009; 113(8):2234–2246. [PubMed: 19146384]
35. Chodera JD, Mobley DL, Shirts MR, Dixon RW, Branson K, Pande VS. Alchemical Free Energy Methods for Drug Discovery: Progress and Challenges. Curr Opin Struct Biol. 2011; 21(2):150–160. [PubMed: 21349700]
36. Gallicchio E, Levy RM. Advances in All Atom Sampling Methods for Modeling Protein–ligand Binding Affinities. Curr Opin Struct Biol. 2011; 21(2):161–166. [PubMed: 21339062]
37. Mobley DL, Gilson MK. Predicting Binding Free Energies: Frontiers and Benchmarks. Annu Rev Biophys. 2017; 46(1):531–558. [PubMed: 28399632]
38. Woo H-J, Roux B. Calculation of Absolute Protein-Ligand Binding Free Energy from Computer Simulations. Proc Natl Acad Sci. 2005; 102(19):6825–6830. [PubMed: 15867154]
39. Lee MS, Olson MA. Calculation of Absolute Protein-Ligand Binding Affinity Using Path and Endpoint Approaches. Biophys J. 2006; 90(3):864–877. [PubMed: 16284269]
40. Gumbart JC, Roux B, Chipot C. Standard Binding Free Energies from Computer Simulations: What Is the Best Strategy? J Chem Theory Comput. 2013; 9(1):794–802. [PubMed: 23794960]
41. Velez-Vega C, Gilson MK. Overcoming Dissipation in the Calculation of Standard Binding Free Energies by Ligand Extraction. J Comput Chem. 2013:n/a–n/a.
42. Gilson M, Given J, Bush B, Mccammon J. The Statistical-Thermodynamic Basis for Computation of Binding Affinities: A Critical Review. Biophys J. 1997; 72(3):1047–1069. [PubMed: 9138555]
43. Boresch S, Tettinger F, Leitgeb M, Karplus M. Absolute Binding Free Energies: A Quantitative Approach for Their Calculation. J Phys Chem B. 2003; 107(35):9535–9551.
44. Deng Y, Roux B. Calculation of Standard Binding Free Energies: Aromatic Molecules in the T4 Lysozyme L99A Mutant. J Chem Theory Comput. 2006; 2(5):1255–1273. [PubMed: 26626834]
45. Mobley DL, Graves AP, Chodera JD, McReynolds AC, Shoichet BK, Dill KA. Predicting Absolute Ligand Binding Free Energies to a Simple Model Site. J Mol Biol. 2007; 371(4):1118–1134. [PubMed: 17599350]
46. Deng N, Zhang P, Cieplak P, Lai L. Elucidating the Energetics of Entropically Driven Protein–Ligand Association: Calculations of Absolute Binding Free Energy and Entropy. J Phys Chem B. 2011; 115(41):11902–11910. [PubMed: 21899337]
47. Rocklin GJ, Boyce SE, Fischer M, Fish I, Mobley DL, Shoichet BK, Dill KA. Blind Prediction of Charged Ligand Binding Affinities in a Model Binding Site. J Mol Biol. 2013; 425(22):4569–4583. [PubMed: 23896298]
48. Ge X, Roux B. Calculation of the Standard Binding Free Energy of Sparsomycin to the Ribosomal Peptidyl-Transferase P-Site Using Molecular Dynamics Simulations with Restraining Potentials. J Mol Recognit. 2010; 23(2):128–141. [PubMed: 20151411]

49. Lawrenz M, Baron R, McCammon JA. Independent-Trajectories Thermodynamic-Integration Free-Energy Changes for Biomolecular Systems: Determinants of H5N1 Avian Influenza Virus Neuraminidase Inhibition by Peramivir. J Chem Theory Comput. 2009; 5(4):1106–1116. [PubMed: 19461872]
50. Aldeghi M, Heifetz A, Bodkin MJ, Knapp S, Biggin PC. Predictions of Ligand Selectivity from Absolute Binding Free Energy Calculations. J Am Chem Soc. 2017; 139(2):946–957. [PubMed: 28009512]
51. Deng N, Forli S, He P, Perryman A, Wickstrom L, Vijayan SKV, Tiefenbrunn T, Stout CD, Gallicchio E, Olson AJ, et al. Distinguishing Binders from False Positives by Free Energy Calculations: Fragment Screening Against the Flap Site of HIV Protease. J Phys Chem B. 2014; 119(3):976–988. [PubMed: 25189630]
52. Slaughter A, Jurado KA, Deng N, Feng L, Kessl JJ, Shkriabai N, Larue RC, Fadel HJ, Patel PA, Jena N, et al. The Mechanism of H171T Resistance Reveals the Importance of Ndelta-Protonated His171 for the Binding of Allosteric Inhibitor BI-D to HIV-1 Integrase. Retrovirology. 2014; 11(1):100. [PubMed: 25421939]
53. Deng N, Flynn WF, Xia J, Vijayan RSK, Zhang B, He P, Mentes A, Gallicchio E, Levy RM. Large Scale Free Energy Calculations for Blind Predictions of Protein–ligand Binding: The D3R Grand Challenge 2015. J Comput Aided Mol Des. 2016; 30(9):743–751. [PubMed: 27562018]
54. Rocklin GJ, Mobley DL, Dill KA, Hünenberger PH. Calculating the Binding Free Energies of Charged Species Based on Explicit-Solvent Simulations Employing Lattice-Sum Methods: An Accurate Correction Scheme for Electrostatic Finite-Size Effects. J Chem Phys. 2013; 139(18): 184103. [PubMed: 24320250]
55. Wilhelm M, Mukherjee A, Bouvier B, Zakrzewska K, Hynes JT, Lavery R. Multistep Drug Intercalation: Molecular Dynamics and Free Energy Studies of the Binding of Daunomycin to DNA. J Am Chem Soc. 2012; 134(20):8588–8596. [PubMed: 22548344]
56. Zhou J-K, Yang D-Y, Sheu S-Y. The Molecular Mechanism of Ligand Unbinding from the Human Telomeric G-Quadruplex by Steered Molecular Dynamics and Umbrella Sampling Simulations. Phys Chem Chem Phys PCCP. 2015; 17(19):12857–12869. [PubMed: 25908641]
57. Young T, Abel R, Kim B, Berne BJ, Friesner RA. Motifs for Molecular Recognition Exploiting Hydrophobic Enclosure in Protein-Ligand Binding. Proc Natl Acad Sci. 2007; 104(3):808–813. [PubMed: 17204562]
58. Nguyen CN, Cruz A, Gilson MK, Kurtzman T. Thermodynamics of Water in an Enzyme Active Site: Grid-Based Hydration Analysis of Coagulation Factor Xa. J Chem Theory Comput. 2014; 10(7):2769–2780. [PubMed: 25018673]
59. Bochevarov AD, Harder E, Hughes TF, Greenwood JR, Braden DA, Philipp DM, Rinaldo D, Halls MD, Zhang J, Friesner RA. Jaguar: A High-Performance Quantum Chemistry Software Program with Strengths in Life and Materials Sciences. Int J Quantum Chem. 2013; 113(18):2110–2142.
60. Pronk S, Pall S, Schulz R, Larsson P, Bjelkmar P, Apostolov R, Shirts MR, Smith JC, Kasson PM, van der Spoel D, et al. GROMACS 4.5: A High-Throughput and Highly Parallel Open Source Molecular Simulation Toolkit. Bioinformatics. 2013; 29(7):845–854. [PubMed: 23407358]
61. Pérez A, Marchán I, Svozil D, Sponer J, Cheatham TE, Laughton CA, Orozco M. Refinement of the AMBER Force Field for Nucleic Acids: Improving the Description of α/γ Conformers. Biophys J. 2007; 92(11):3817–3829. [PubMed: 17351000]
62. Ivani I, Dans PD, Noy A, Pérez A, Faustino I, Hospital A, Walther J, Andrio P, Goñi R, Balaceanu A, et al. Parmbsc1: A Refined Force Field for DNA Simulations. Nat Methods. 2015
63. Wang J, Wolf RM, Caldwell JW, Kollman PA, Case DA. Development and Testing of a General Amber Force Field. J Comput Chem. 2004; 25(9):1157–1174. [PubMed: 15116359]
64. Jakalian A, Bush BL, Jack DB, Bayly CI. Fast, Efficient Generation of High-Quality Atomic Charges. AM1-BCC Model: I. Method. J Comput Chem. 2000; 21(2):132–146.
65. Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML. Comparison of Simple Potential Functions for Simulating Liquid Water. J Chem Phys. 1983; 79(2):926.
66. Joung IS, Cheatham TE. Determination of Alkali and Halide Monovalent Ion Parameters for Use in Explicitly Solvated Biomolecular Simulations. J Phys Chem B. 2008; 112(30):9020–9041. [PubMed: 18593145]

67. Àqvist J. Ion-Water Interaction Potentials Derived from Free Energy Perturbation Simulations. J Phys Chem. 1990; 94(21):8021–8024.
68. Rebič M, Laaksonen A, Šponer J, Uličný J, Mocci F. Molecular Dynamics Simulation Study of Parallel Telomeric DNA Quadruplexes at Different Ionic Strengths: Evaluation of Water and Ion Models. J Phys Chem B. 2016; 120(30):7380–7391. [PubMed: 27379924]
69. Essmann U, Perera L, Berkowitz ML, Darden T, Lee H, Pedersen LG. A Smooth Particle Mesh Ewald Method. J Chem Phys. 1995; 103(19):8577.
70. Yin J, Henriksen NM, Slochower DR, Gilson MK. The SAMPL5 Host-guest Challenge: Computing Binding Free Energies and Enthalpies from Explicit Solvent Simulations by the Attach-Pull-Release (APR) Method. J Comput Aided Mol Des. 2017; 31(1):133–145. [PubMed: 27638809]
71. Lin Y-L, Meng Y, Huang L, Roux B. Computational Study of Gleevec and G6G Reveals Molecular Determinants of Kinase Inhibitor Selectivity. J Am Chem Soc. 2014; 136(42):14753–14762. [PubMed: 25243930]
72. Haider K, Wickstrom L, Ramsey S, Gilson MK, Kurtzman T. Enthalpic Breakdown of Water Structure on Protein Active-Site Surfaces. J Phys Chem B. 2016; 120(34):8743–8756. [PubMed: 27169482]
73. Friesner RA, Murphy RB, Repasky MP, Frye LL, Greenwood JR, Halgren TA, Sanschagrin PC, Mainz DT. Extra Precision Glide: Docking and Scoring Incorporating a Model of Hydrophobic Enclosure for Protein–Ligand Complexes. J Med Chem. 2006; 49(21):6177–6196. [PubMed: 17034125]
74. Gallicchio E, Kubo MM, Levy RM. Enthalpy-Entropy and Cavity Decomposition of Alkane Hydration Free Energies: Numerical Results and Implications for Theories of Hydrophobic Solvation. J Phys Chem B. 2000; 104(26):6271–6285.
75. Levy RM, Zhang LY, Gallicchio E, Felts AK. On the Nonpolar Hydration Free Energy of Proteins: Surface Area and Continuum Solvent Models for the Solute-Solvent Interaction Energy. J Am Chem Soc. 2003; 125(31):9523–9530. [PubMed: 12889983]
76. Shaikh SA, Ahmed SR, Jayaram B. A Molecular Thermodynamic View of DNA–drug Interactions: A Case Study of 25 Minor-Groove Binders. Arch Biochem Biophys. 2004; 429(1):81–99. [PubMed: 15288812]
77. Treesuwan W, Wittayanarakul K, Anthony NG, Huchet G, Alniss H, Hannongbua S, Khalaf AI, Suckling CJ, Parkinson JA, Mackay SP. A Detailed Binding Free Energy Study of 2 : 1 Ligand–DNA Complex Formation by Experiment and Simulation. Phys Chem Chem Phys. 2009; 11(45): 10682. [PubMed: 20145812]
78. Mulholland K, Wu C. Binding of Telomestatin to a Telomeric G-Quadruplex DNA Probed by All-Atom Molecular Dynamics Simulations with Explicit Solvent. J Chem Inf Model. 2016; 56(10): 2093–2102. [PubMed: 27632666]
79. Genheden S, Ryde U. The MM/PBSA and MM/GBSA Methods to Estimate Ligand-Binding Affinities. Expert Opin Drug Discov. 2015; 10(5):449–461. [PubMed: 25835573]
80. Islam B, Stadlbauer P, Neidle S, Haider S, Sponer J. Can We Execute Reliable MM-PBSA Free Energy Computations of Relative Stabilities of Different Guanine Quadruplex Folds? J Phys Chem B. 2016; 120(11):2899–2912. [PubMed: 26918369]
81. Moraca F, Amato J, Ortuso F, Artese A, Pagano B, Novellino E, Alcaro S, Parrinello M, Limongelli V. Ligand Binding to Telomeric G-Quadruplex DNA Investigated by Funnel-Metadynamics Simulations. Proc Natl Acad Sci. 2017; 114(11):E2136–E2145. [PubMed: 28232513]
82. Funke A, Dickerhoff J, Weisz K. Towards the Development of Structure-Selective G-Quadruplex-Binding Indolo[3,2-*b*]Quinolines. Chem – Eur J. 2016; 22(9):3170–3181. [PubMed: 26808655]
83. Sun D, Hurley LH. The Importance of Negative Superhelicity in Inducing the Formation of G-Quadruplex and i-Motif Structures in the c-Myc Promoter: Implications for Drug Targeting and Control of Gene Expression. J Med Chem. 2009; 52(9):2863–2874. [PubMed: 19385599]
84. Felsenstein KM, Saunders LB, Simmons JK, Leon E, Calabrese DR, Zhang S, Michalowski A, Gareiss P, Mock BA, Schneekloth JS. Small Molecule Microarrays Enable the Identification of a

Selective, Quadruplex-Binding Inhibitor of MYC Expression. ACS Chem Biol. 2016; 11(1):139–148. [PubMed: 26462961]

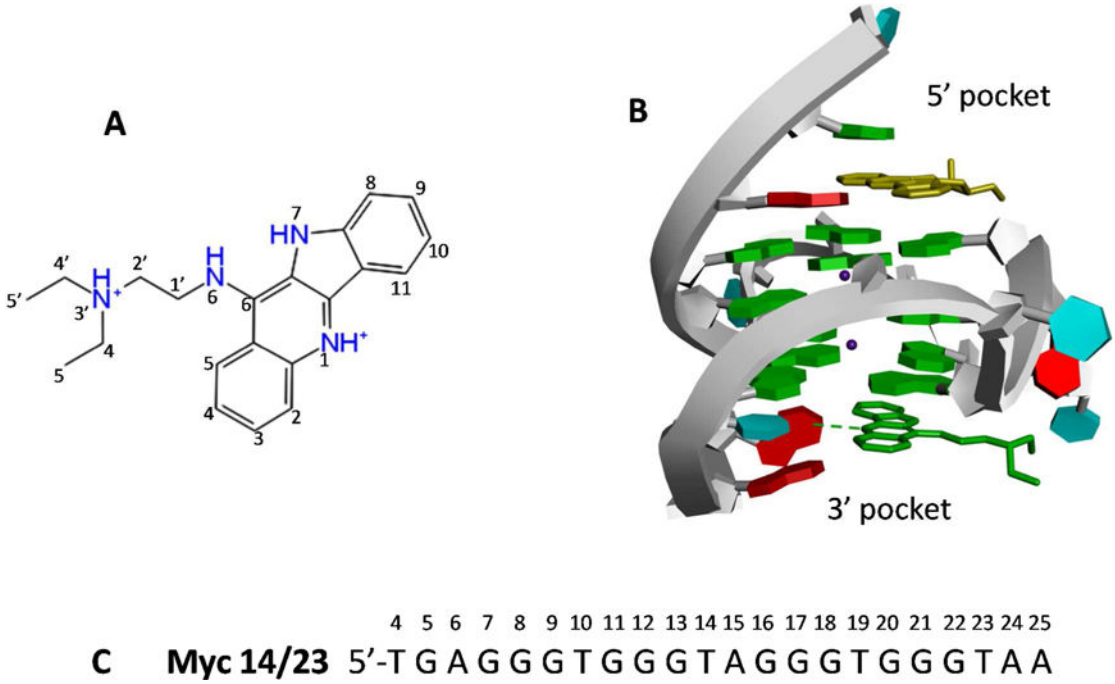


Figure 1.

(A). The 2D structure of the quindoline derivative. (B)The NMR structure of the 2:1 quindoline-G-quadruplex complex. The two quindoline molecules at the 5′-end and 3′ end are shown in yellow and green sticks, respectively. At the 3′-end the intermolecular hydrogen bond between the T23O4 of the DNA and N1 of the quindoline is shown as green dashed line. The two potassium ions are shown as purple dots. (C). The sequence of c-MYC G-quadruplex (Myc 14/23).

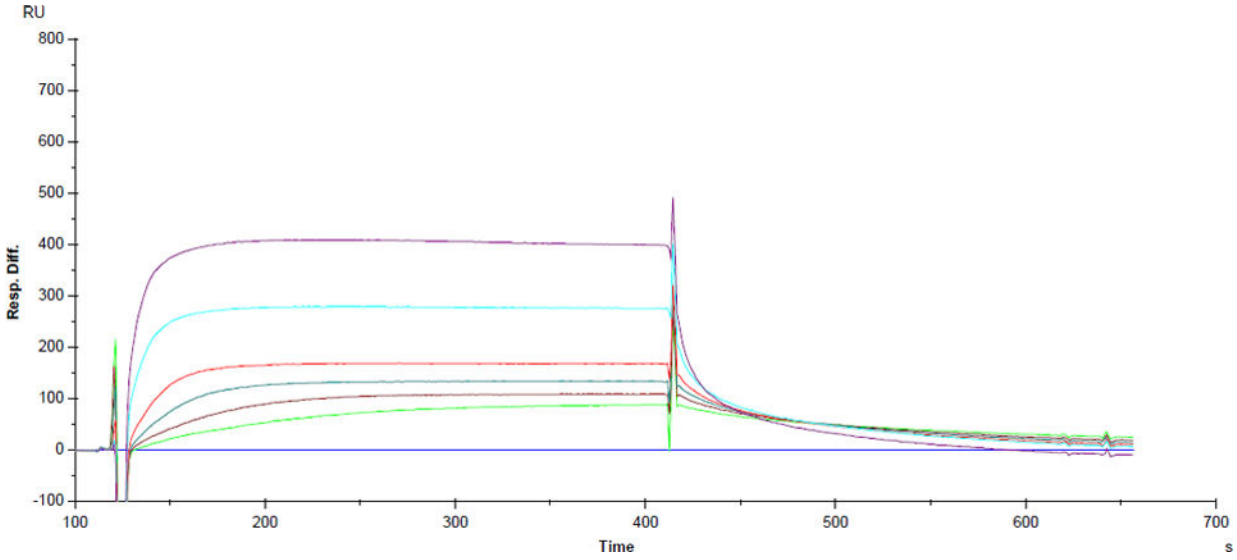


Figure 2.

SPR sensorgram for binding of increasing concentrations of quindoline to Myc 14/23 d[$TGAG_3TG_3TAG_3TG_3TA_2$] in 10mM Hepes, 100mM KCl, 3mM EDTA, 0.005% Polysorbate 20, pH 7.4. The concentrations of quindoline are 0.1, 0.2, 0.5, 1, 5, and 10 μM.

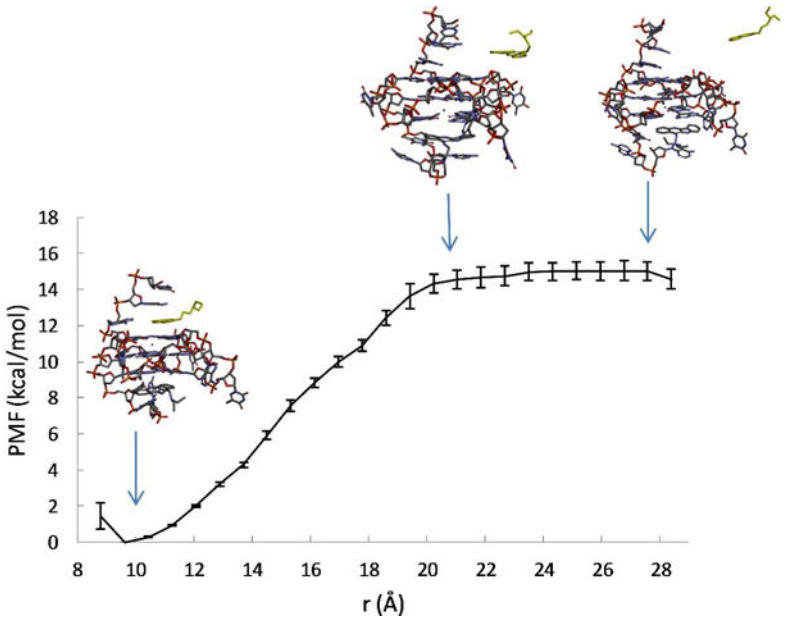


Figure 3.

Calculated free energy profile $w(r)$ for the 5′ quindoline of as a function of the distance r between the ligand atom A and the DNA atom a. The representative conformations of the G-quadruplex-quindoline complex observed at different r are also shown, with the quindoline molecule represented by yellow sticks.

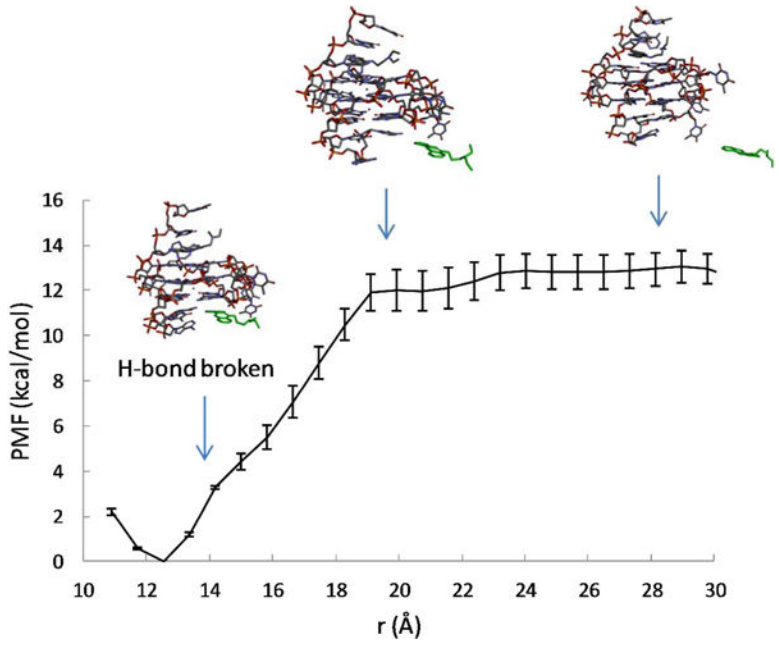


Figure 4.

Calculated free energy profile $w(r)$ for the 3′ quindoline. The representative conformations of the G-quadruplex-quindoline complex observed at the different r are also shown, with the quindoline molecule represented as green sticks.

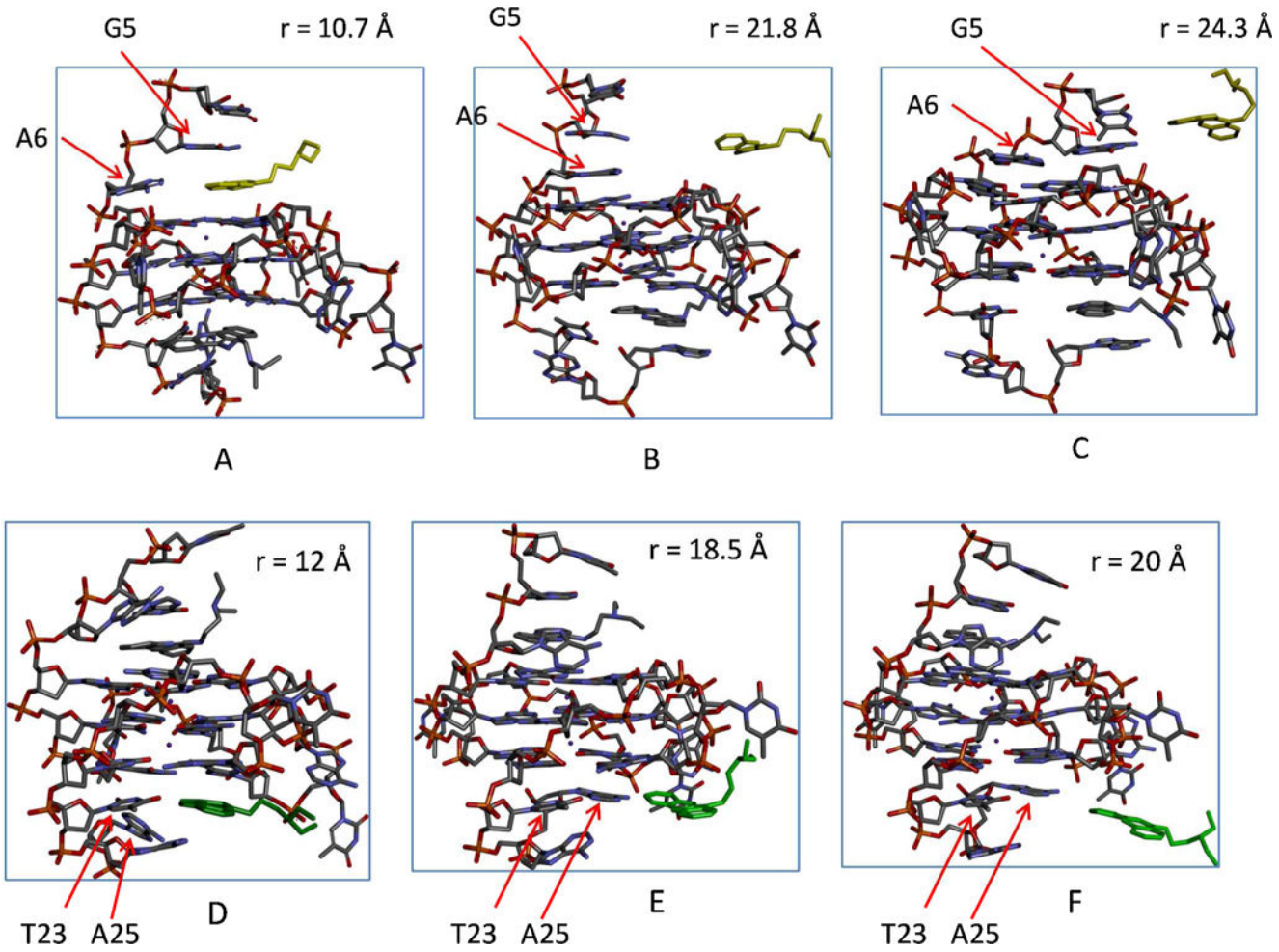


Figure 5.

(A), (B) and (C): conformational changes near the 5′-end pocket induced by ligand unbinding with increasing ligand-DNA distance. (D), (E), and (F): conformational changes near the 3′-end site induced by the ligand unbinding at different ligand-DNA distances.

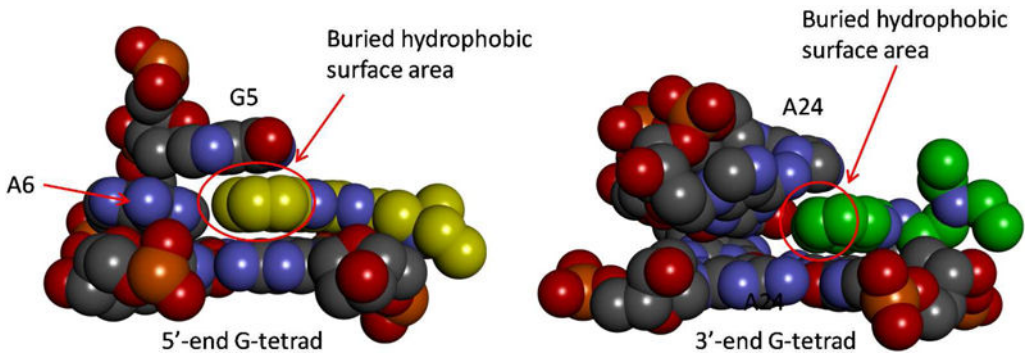


Figure 6.

Detailed views of the 5′-end (left) and 3′-end (right) binding pockets. The carbon atoms in the two quindolines are displayed as yellow (left) and green (right) spheres. The red circles indicate the buried hydrophobic surface areas in each of the binding sites.

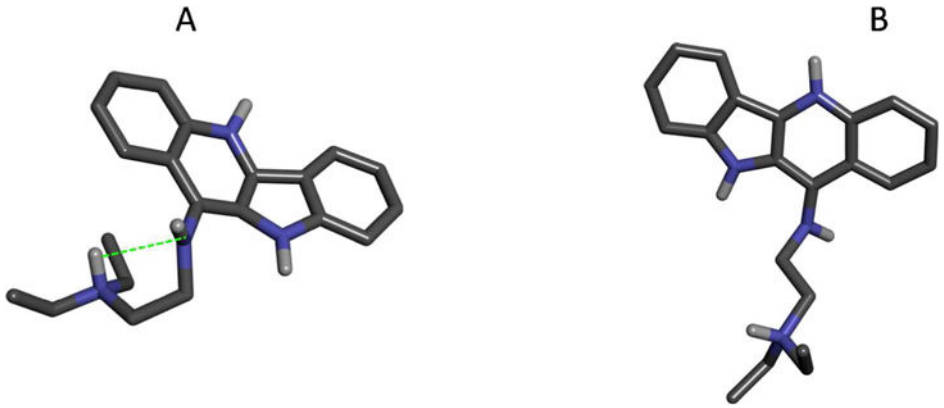


Figure 7.

(A) A representative structure of a fully decoupled ligand sampled without switching off the ligand intramolecular interaction. The green dotted line indicates an electrostatic intramolecular interaction in the quindoline side chain. (B) A structure of the fully decoupled ligand sampled with the intramolecular nonbonded interactions turned off.

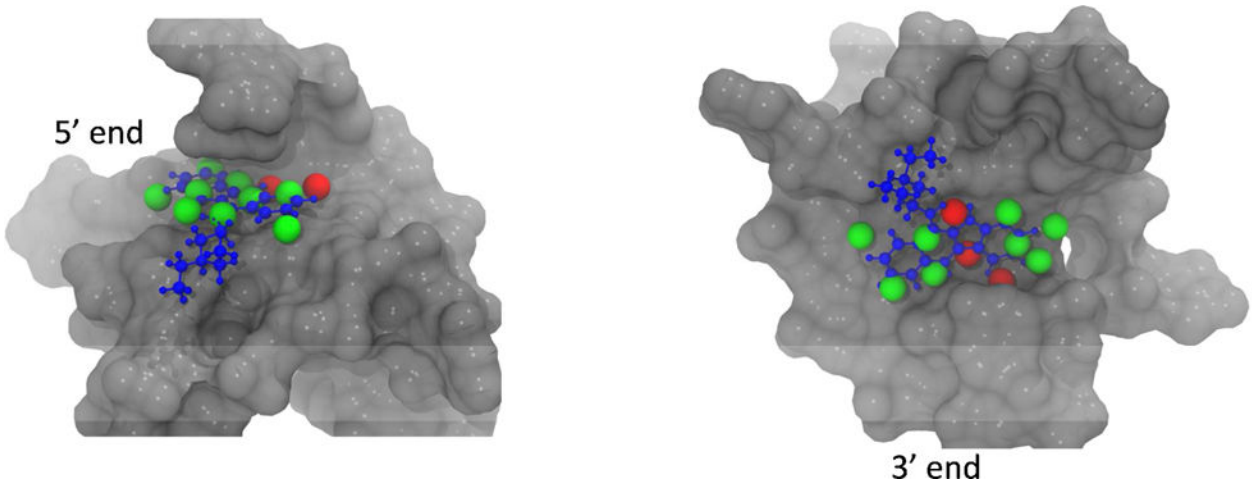


Figure 8.

Hydration sites overlapping with the ligand (blue) in the 5′ end and 3′ end binding cavities. Hydrations sites with energies greater than −9.53 kcal/mol are shown in red and hydration sites with energies less than −9.53 kcal/mol are shown in green. See also Table S1 and S2.

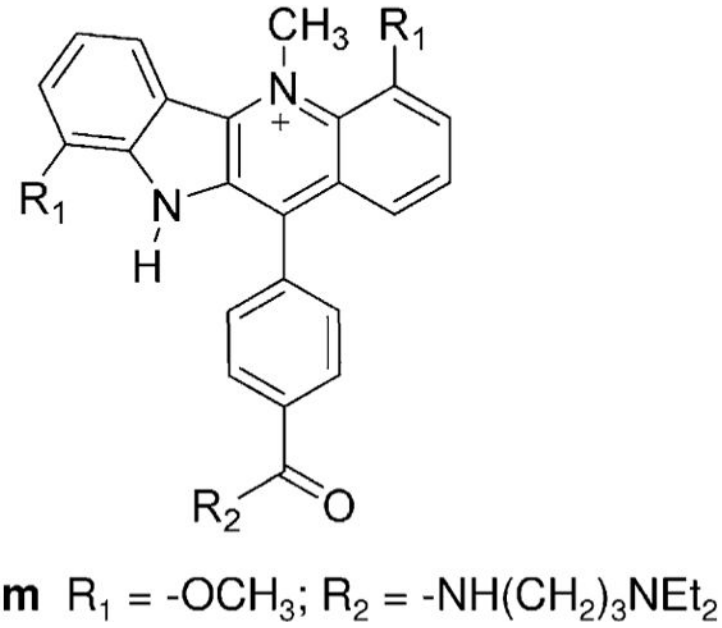


Figure 9.
The 2D structure of the quindoline derivative 4bm studied in the Reference[81].

Table 1

Binding kinetic and thermodynamic parameters from SPR.

k_{on} (M^{-1} s^{-1})	k_{off} (s^{-1})	K_a (M^{-1})	$\Delta G^{\circ}_{bind}(kcal/mol)^a$
$2.57\times10^4 \pm 140$	$7.87\times10^{-3} \pm 2.91\times10^{-5}$	$3.27\times10^6 \pm 3.0\times10^4$	−8.94 ± 0.005

[a] $\Delta G^{\circ}_{bind} = -k_B T \times \ln K_a$.

Table 2

Absolute binding free energies of the quindoline at the two binding sites from computation and experiment. (in kcal/mol.)

Binding Site	ΔG°_{bind} **PMF approach**	ΔG°_{bind} **DDM, no-ligand-intramolecular**[a]	ΔG°_{bind} **DDM, with-ligand-intramolecular**[b]	ΔG°_{bind} **Experiment**[c]
5′-end	−8.8 (±0.6)	−11.2 (±1.3)	−12.4 (±1.7)	−8.94
3′-end	−7.7 (±0.7)	−9.6 (±0.6)	−7.5 (±0.3)	

[a] The intramolecular nonbonded interactions in the decoupled ligand are turned off.

[b] The intramolecular interactions in the decoupled state of the ligand are left unchanged.

[c] Derived from SPR.

Table 3

Contributions to the absolute binding free energy ΔG°_{bind} calculated using the PMF approach.[a] (in kcal/mol.)

Binding free energy components	5′ site	3′ site
$-w(r^*)$	−15.1 (±0.5)	−12.8 (±0.8)
$-k_B T \ln \frac{\int_{bound} e^{-w(r)/k_B T} dr}{(2\pi k_B T/k_r)^{\frac{1}{2}}}$	−0.1 (±0.05)	−0.0 (±0.0)
ΔG^{bulk}_{restr}	10.1 (±0.0)	9.4 (±0.0)
$-\Delta G^{bound}_{restr}$	−3.7 (±0.06)	−4.3 (±0.05)
ΔG°_{bind}	−8.8 (±0.6)	−7.7 (±0.7)

[a] See Methods for the definition of the different binding free energy components.

Table 4

Contributions to the binding free energy from DDM with the nonbonded intramolecular interactions in the decoupled ligand turned off.[a] (in kcal/mol.)

Binding Site	$\Delta G_{Coulomb}$	ΔG_{LJ}	ΔG^{gas}_{restr}	$\Delta G^{finite\_size}_{elec\_corr}$	$-\Delta G^{coupled}_{restr}$	ΔG°_{bind}
5′-end	1.2 (±0.8)	−14.7 (±1.4)	3.2 (±0)	−0.79 (±0.1)	−0.07 (±0.01)	−11.2 (±1.3)
3′-end	−0.5 (±1.7)	−11.4 (±1.1)	3.2 (±0)	−0.81(±0.03)	−0.12 (±0.05)	−9.6 (±0.6)

[a] See Methods for the definition of different binding free energy terms.

Table 5

Contributions to the binding free energy from the DDM setup in which the nonbonded intramolecular interactions in the decoupled ligand are not scaled. (in kcal/mol.)

Binding Site	$\Delta G_{Coulomb}$	ΔG_{LJ}	ΔG^{gas}_{restr}	$\Delta G^{finite\_size}_{elec\_corr}$	$-\Delta G^{coupled}_{restr}$	ΔG°_{bind}
5′-end	−1.6 (±0.4)	−13.5 (±1.4)	3.2 (±0.0)	−0.36 (±0.07)	−0.09 (±0.03)	−12.4 (±1.7)
3′-end	−1.0 (±0.5)	−8.7 (±0.7)	3.2 (±0.0)	−0.89 (±0.04)	−0.12 (±0.04)	−7.5 (±0.3)