Novel G-quadruplex stabilizing agents: in-silico approach and dynamics

Rajiv Kumar Kar, Priyanka Suryadevarab, Jagannath Janaa, Anirban Bhunia* and Subhrangsu Chatterjee*

*aBiomolecular NMR and Drug Design Laboratory, Department of Biophysics, Bose Institute, P-1/12 CIT Scheme VII M, Kolkata- 700054, India; bDepartment of Pharmacoinformatics, NIPER, Hajipur, EPIP Campus, Bihar 844101, India

Communicated by Ramaswamy H. Sarma

(Received 27 March 2012; final version received 10 October 2012)

The stabilization of overhang G-rich repetitive DNA units at the 3′-end of telomeres, which are well known to form functionally important G-quadruplex structures, is a current goal in designing novel anticancer drugs. In the present study, we have undertaken an in silico approach by molecular docking using a small molecule library to find potential G-quadruplex stabilizing agents. Two molecules, A, [N′1-imino(2-pyridyl)methyl-3,4,5-trimethoxybenzene-1-carboxyhydrazide] and B, [3-[4-(5-[2cyanoethyl](methyl)amino)benzylidene-amino)propyl]imino)propanenitrile], that had good docking scores have been investigated for interaction with G-quadruplexes in a Molecular Dynamics simulation study. Fluorescence spectroscopy of G-quadruplexes bound to the screened molecules A and B was used to experimentally validate the theoretical results. The binding of ligands A and B to G-quadruplexes resulted in blue shifts of 10–18 nm, respectively, in the fluorescence emission spectra of the G-quadruplexes, demonstrating that both molecules bind to the G-face of the quadruplex. The same experiment was performed for the complexation of these small molecules with a G-rich DNA duplex, 5′-GCGCATGCTACGC-3′ 3′-CGCGTACGAGCGC-5′. Interestingly, no blue shift was observed in the fluorescence emission spectra of the DNA duplex in the presence of these small molecules. Thus, these findings indicated that these ligands very selectively bind to G-quadruplexes instead of the duplex DNA. In addition, a one-dimensional water ligand observed via a gradient spectroscopy Nuclear Magnetic Resonance (NMR) experiment showed that both molecules bound to the 23-mer G-quadruplex DNA. The molecular properties of the ligand–quadruplex complex have been analyzed with the help of the Adaptive Poisson-Boltzmann Solver, revealing that electrostatics govern the binding of the small molecules to G-quadruplexes. Both molecules were investigated in detail using solvation free energy calculations and Absorption, Distribution, Metabolism, Elimination and Toxicity (ADMET) predictions, which provide insight into lead optimization for designing G-quadruplex stabilizing agents; therefore, these molecules have potential as new therapeutic agents.

Keywords: G-quadruplex; docking; MD simulation; APBS; WaterLOGSY NMR; FEP

Introduction

The 3′-end region of chromosomes consists of almost 200 nucleotides composed of tandem repeats of hexanucleotide (TTAGGG), and ends with an overhanging single strand. Telomeres are the main location in which these crucial functional secondary structures of DNA form (Cech, 2004). Telomeres have a vital role in life because of the various functions of telomere-associated proteins, their attachment to the nuclear matrix, and their higher order chromatin structures (Campbell, 2012; Collado, Blasco, & Serrano, 2007). Conservation of telomeric length is an important biological condition for cell growth (Engelhardt & Finke, 2001; Lange, 2009). Telomerase is the enzyme that functionalizes the addition of hexanucleotide repeats of TTAGGG to the 3′-end of a telomere and maintains the telomere during normal cell division (mitosis) (Eisenstein, 2011; Flores et al., 2005). Unusual overexpression of telomerase causes massive extension of telomeric ends and anomalous cell proliferation, which causes cancer. Many studies report that 80–85% of cancerous cells have overexpression of telomerase (Cong, Wright, & Shay, 2002; Lin et al., 2008; Rodriguez-Brenes, & Peskin, 2009). Researchers have hypothesized that the inhibition of telomerase can be an effective approach to return cancer cells to natural apoptosis (Finkel, Serrano, & Blasco, 2007; Sun & Hurley, 2001; Zaug, Podell, & Cech, 2005). The finding that cells switch to alternative mechanisms to maintain telomeric extension is valuable for adopting different strategies for the down regulation of genes (Hu et al., 2012; Ward & Autexier, 2005).
The overhanging G-rich repetitive DNA units at the 3'-end of telomeres can form various tertiary structures, including G-quadruplexes, where the guanine bases stack over each other and are stabilized by the cyclic Hoogsteen type of hydrogen bonding (Burge, Parkinson, Hazel, Todd, & Neidle, 2006; Dai, Carver, Punchihewa, Jones, & Yang, 2007; Dai, Punchihewa et al., 2007; Luu, Phan, Kuryavyi, Lacroix, & Patel, 2008). These unusual but functionally important DNA structures are also present at other genomic positions, such as promoter oncogenic regions, including c-kit (Fernando et al., 2006), c-MYC (Jain, Grand, Bearss, & Hurley, 2002), bel-2 (Dexheimer, Sun, & Hurley, 2006), VEGF (Sun, Guo, Rusche & Hurley, 2005), and others. There has been immense research on the biological and physical properties of these noncanonical DNA structures; this research indicates that G-quadruplexes might be a target for designing new anticancer therapeutics (Balasubramanian & Neidle, 2009; Neidle, 2009; Neidle & Parkinson, 2008; Neidle & Read, 2001). The most promising aspect of the G-quadruplex structure is its topology, which cannot be recognized by the single-stranded RNA component of the telomerase enzyme (Rodriguez et al., 2008). In contrast, the G-quadruplex itself can be a signal of DNA damage and can, therefore, induce apoptosis (Rodriguez et al., 2008). Currently, researchers are more focused on developing agents that can stabilize the secondary G-rich structures of DNA rather than developing agents to inhibit telomerase activity (Dailey et al., 2009; Monchau & Teulade-Fichou, 2008).

A number of quadruplex-stabilizing ligands have been characterized both by laboratory synthesis and assays as well as by in silico approaches (Martins et al., 2007; Read et al., 2001). Only a few of the G-quadruplex-stabilizing agents have been successful in various clinical trials (Drygin, Jain, & O’Brien, 2009; Simonsson, Pecinka, & Kubista, 1998). The popular stabilizing agents that have already been tested and studied are mainly derivatives of acridine (Burger et al., 2005), bisquinoline, macrocycle, porphyrin, (Hu et al., 1999) and anthracene (Granotier et al., 2005; Kim, Vankayalapati, Shin-ya, Wierzba, & Hurley, 2002; Tauchi et al., 2006). In addition to those types of molecules, the list of other organic molecules that have been synthesized and tested to design novel therapeutic agents is endless (Du et al., 2010; Li et al., 2012; Peduto et al., 2011). Many popular computational methods, such as virtual screening, pharmacophore mapping, 3D-QSAR, and CoMFA-COMSIA, have also been used to determine which structural features can be changed for successful future anticancer drug discovery (Chen et al., 2011; Ma, Chan, Lee, Kwan, & Leugn, 2011; Zambre, Murumkar, Giridhar, & Yadav, 2010).

Determining which organic chemical class should be investigated experimentally is a crucial procedure in the initial stage of drug development. The synthesis, derivatization with modification in functional groups, and addition–deletion of excluded volumes for achieving a molecule that is truly specific for the target are indeed tedious, requiring a huge investment of manpower and finances. Therefore, despite much investment in research, no drug has been developed as a potential stabilizing agent of G-quadruplexes. Alternatively, structure-based drug design is an efficient method to screen several potential chemical classes of compounds; after lead optimization, the resulting chemical moiety may prove worthy to the pharmaceutical community (Holt, Buscaglia, Trent, & Chaires 2011; Neidle & Thurston, 2005). Because of the advancement of computing power, in silico drug design has proven to be advantageous in the rational design of novel therapeutic agents; this method is currently in high demand. In the present study, we have adopted the most consistent and efficient methodology of molecular docking to screen a drug library for potential G-quadruplex-stabilizing agents. We have attempted to validate the methodology at regular intervals and to reduce the selection criteria for molecules to produce the most accurate positive result. The docking result was further strengthened theoretically with another promising method, Molecular Dynamics (MD) simulation. In addition, the results of the theoretical studies were validated using low-resolution and high-resolution spectroscopy. Fluorescence and Nuclear Magnetic Resonance (NMR) spectroscopies were employed to show that the hit molecules bind selectively to the G-quadruplex structure. The fundamental nature and various other characteristics of the hypothesized molecules, such as electrostatic properties, solvation free energy, and pharmacokinetics (Absorption, Distribution, Metabolism, Elimination and Toxicity [ADMET]), have been analyzed; this information could be used by various research communities to test and develop these G-quadruplex stabilizing agents.

We search for a promising algorithm that can be adapted successfully to satisfy our goal of finding a G-quadruplex stabilizing agent via in silico docking/virtual screening. All of the available programs have cutting-edge results against protein targets, yet their efficacies against nucleic acids are still debated. Anthony and coworkers have performed docking simulation experiments where the small molecules are targeted to the minor groove of DNA (Anthony et al., 2005). The simulation showed an excellent agreement with experimental results from X-ray crystallography, NMR, and gel-based footprinting. In another type of assessment, it was found that protein-based docking programs such as Glide and GOLD are efficient enough to reproduce the binding modes of the 60 tested RNA complexes (Li et al., 2010). A similar type of validation was also reported recently by a molecular modeling group; they proved that the performance of Glide and GOLD with the simulated structure is very good and produces accurate results in accordance with the
structures elucidated by spectroscopic techniques (Srivastava, Chourasia, Kumar, & Sastry, 2011). Moreover, a comparative analysis of different DNA binding drugs for disease targets such as leishmania has also been reported where Glide is used for docking studies with nucleic acid targets (Chauhan, Vidyarthi, & Poddar, 2012). Considering these reports, we have adopted the efficient Glide docking algorithm to screen a drug library for potential G-quadruplex stabilizing agents.

**Materials and methods**

**Chemicals and reagents**

The DNA duplex and G-quadruplex sequences were purchased from Kric, India. The small molecules, **A** and **B**, were purchased from Fisher Scientific Inc (Maybridge, Cornwall, UK). Deuterium oxide was obtained from Cambridge Isotope Laboratories Inc.

**G-quadruplex and molecular data-set**

For this investigation, the NMR structure of human telomeric G-quadruplex T(AGGG(TTAGGG))₃ (PDB accession code: 2ld8) (Heddi & Phan, 2011) was chosen. The potassium (K⁺) ion was described as an important factor contributing to the structural topology of G-quadruplexes (Lim et al., 2009; Renčiuk et al., 2009). Two K⁺ ions were manually placed between three quartets using the edit option in Maestro. Potential binding sites for G-quadruplex were identified using SiteMap v2.5; the calculations begin with an initial search stage that finds one or more regions called sites, which are suitable for binding to ligand molecules. The molecular data-set containing the small molecules used for virtual screening was obtained from the Maybridge database (HitFinder) of 14,400 molecules. This database was chosen because it has selection of molecules using a clustering algorithm that employs standard Daylight Fingerprints with the Tanimoto (Butina, 1999) similarity index. Clustering with .71 similarities can then be used in NMR for screening against the target (Baurin et al., 2004). Moreover, the entire molecular data-set is in the permissible range of drug-likeness, i.e. the Lipinski rule of five (Lipinski, Lombardo, Dominy, & Feeney, 2001): ClogP ≤ 5, H-bond acceptors ≤ 10, H-bond donors ≤ 5, and mol. wt. ≤ 500.

**Docking Studies**

All of the molecules were docked using Glide v5.7. The conformational flexibility of the ligand molecule is determined by an extensive conformational search augmented by a heuristic screen that eliminates unsuitable conformations with long-range internal H-bonds. For each core conformation of a ligand, an exhaustive search for possible locations and orientations is performed over the surface of the G-quadruplex. The refined ligand undergoes energy minimization based on precomputed OPLS-AA (Jorgensen, Maxwell, & Tirado-Rives, 1996) van der Waals and electrostatic grids to reduce the large energy and gradient terms resulting from inter-atomic contacts that are too close. A grid with dimensions of 40.10 Å × 38.63 Å × 30.56 Å over the G-quadruplex was used to define the binding region and to allow ligands to dock at distances ≤ 20 Å. The defined grid covers the entire topology of the G-quadruplex so that the ligand position and orientation relative to the receptor can be sampled sufficiently; the conformation of the receptor was also fixed during the docking. The standard precision (SP) mode of the software was used for the initial screening of the molecules. Selected top scoring molecules binding to the desired core of the G-quadruplex were selected for the extra precision (XP) mode of docking (Eldridge, Murray, Auton, Paolini, & Mee, 1997). To validate the docking results, we used the docking score of Quarflaxin (CX-3543) (Simonsson et al., 1998), a lead molecule designed and synthesized by Cylene Pharmaceuticals, as a positive control for the docking procedure. Quarflaxin, a pentacyclic system with amino side chains and high selectivity for G-quadruplexes, is already in a phase II clinical trial. Other classes of compounds with reported interaction with G-quadruplexes were also used to perform the control test of docking studies (Zhen, Xi, Chattopadhyaya, & Hecht, 2011). The Glide score is based on the ChemScore scoring function coupled with the ligand–receptor molecular mechanics interaction energy and ligand strain energy; thus, provides a docked pose. Based on the binding site of these drugs and their docking score, we then validated our docking study of small molecules from the Maybridge database. As part of the docking result analysis, we also used a 1000 ligand-like decoy set from Schrödinger (www.schroderger.com/productpage/14/5/74/) to evaluate the results using a receiver operating characteristic curve (ROC) curve approach (Triballeau, Acher, Brabet, Pin, & Bertrand, 2005). The characteristic area under the curve obtained from ROC can be helpful in predicting whether the obtained hits are true positives.

**Molecular dynamics (MD) simulation**

Dynamic simulation of the drug-bound and drug-free states of the G-quadruplex was performed using Desmond Molecular Dynamics System v2.2 with OPLS-AA force fields (Jorgensen et al., 1996). The OPLS force field generates accurate charges and force field parameters for ligands. This force field was adopted with an objective to keep the same set of protonation states and charge parameters for ligand pose refinement and conformational sampling during docking; thus, the same parameter set will also be used in the simulation studies. The solute system containing the drug-bound and drug-
free states of the G-quadruplex was solvated using the TIP3P water model (Jorgensen, Chandrasekhar, Madura, Impey, & Klein, 1983), with orthorhombic (α=β=γ=90°) boundary conditions extending a distance of 10 Å from any solute atom in all directions. The charge of the total system was neutralized by adding the appropriate number of Na⁺ counter ions. The number of explicit water molecules ranged from 3891 to 4008. After the set up of the system, an initial minimization was performed with a convergence threshold of 1.0 kcal.mol⁻¹Å⁻¹ to allow all of the atoms to adjust to the system environment. An MD simulation was performed for each system using an isothermal–isobaric ensemble (NPT) with a relaxation process. The relaxation process consisted of minimization with and without solute restraints and simulation in an canonical ensemble using a Berendsen thermostat with a temperature of 10 K and restraints on solute heavy atoms for 12 ps. In contrast, simulation in the NPT ensemble was performed using a Berendsen thermostat and a Berendsen barostat with a temperature of 10 K and a pressure of 1 atm for 12 ps. In the next stage of relaxation, the system was simulated for 24 ps with Berendsen NPT with a temperature of 300 K and a pressure of 1 atm with and without restraints on solute heavy atoms. The particle mesh Ewald (Darden, York, & Pedersen, 1993) method was used with a tolerance of 1e⁻⁰⁹ as a part of the Coulomb method. A cutoff of 14 Å was used for calculating the nonbonded solvent–solvent and solute–solvent interactions. The pressure and temperature of the system was maintained at 1.013 atm (isotropic) and 300 K, respectively. The SHAKE algorithm (Kutler, Gunsteren, & Hennenberger, 2001) was employed with an integration time step of 2 fs. A recording interval of 1.2 and 4.8 ps for the energy and trajectory frames, respectively, was set. The MD simulation for the entire system was continued for up to 5 ns, and the trajectory analysis was performed using the Simulation Event Analysis module of Desmond (Scheme 2).

Adaptive Poisson-Boltzmann Solver (APBS) iso-surface calculations

The Adaptive Poisson-Boltzmann Solver (APBS) was used to analyze the electrostatic characteristics of the macromolecular complex. The standard all-atom charges generated by the OPLS force field were used for this purpose. The Poisson-Boltzmann equation (PBE), the most widespread model for evaluating electrostatic properties, was adopted from the reference of Baker and coworkers (Baker, Sept, Joseph, Holst, & Camnon, 2001). The essence of the model is that it relates the electrostatic potential to the dielectric properties and the distribution of the atomic partial charges of the biomolecular system. The various computations required to evaluate the electrostatic properties were performed using a plugin option of PyMOL v1.2r3pre, and the resulting electrostatic contour visualizations were collected at the appropriate positive and negative iso-surface values, optimized for visualization.

Free energy perturbation (FEP) and concept of free energy

The main objective of performing the MD simulations for an ensemble system is to explore the macroscopic properties of that system based on its microscopic properties. The macroscopic properties include various thermodynamic terms, such as temperature, pressure, number of particles in the system, and the free energy of binding of a particular compound; in contrast, the atomic positions (r) and the momenta (p) are the microscopic properties of a system, which are defined as the coordinates of the multi-dimensional space called the phase space. The concept of free energy calculations was developed in the early nineties (Kirkwood, 1935; Zwanzig, 1954) but was not popular because of the complexity. The first successful treatment of free energy calculations was for the proton transfer in the lysozyme (Warshel & Levitt, 1976). These calculations gained popularity with significant works by several scientists (Kollman, 1993; McCarrick & Kollman, 1999) who integrated thermodynamic parameters with the free energy calculations. This free energy calculation is mainly based on Zwanzig’s equation, which gives the free energy difference between the two states A and B.

\[
\Delta G = G_B - G_A = -\beta^{-1}\ln[\exp(-\beta V)]_A
\]  

(1)

where, \(\beta=1/kT\) (k is Boltzmann constant and \(T\) is temperature), \([\ldots]_A\) denotes an MD or Monte Carlo generated ensemble average of state A, and \(\Delta V = V_A - V_B\), \(V_A\) and \(V_B\) are the potential energies of A and B states. For absolute solvation free energy expressions, A can be defined as the solute molecule in the gas phase and B as the solute molecule in the solution phase.

The free energy calculations were performed in Desmond. From Figure 1, the solvation free energy for

![Figure 1](https://example.com/figure1.png)

Figure 1. Schematic representation of the free energy change in solute–solvent interaction in different phases. Reactant is referred to ligand molecule and since there is no differentiation in the formed product except the molecular conformation, nothing is referred to the products here.
ligand molecules can be calculated by annihilating the reactant (ligand) in solvent and in vacuum.

\[ \Delta G_{\text{solvation}} = \Delta G_{\text{gasphase}} - \Delta G_{\text{solvent}} \]  

(2)

The main advantage of Desmond is that it requires only one step to calculate the solvation free energy of the solute; the reactant in vacuum is recovered from the solvent phase by rejecting the interactions of reactant with any other atoms in system (Shivakumar et al., 2010).

The solvation free energy calculation plays a major role in predicting the binding affinity of the compound and is a powerful tool in the lead optimization step for drug discovery. Implementing a FEP is a multi-step method, which involves the use of a group of intermediate potential energy functions defined between the paths from state A to state B (Brandsdal et al., 2003). In these functions, \( \lambda \) is the coupling factor whose value ranges from 0 to 1, and \( m = (1 \text{ to } n) \) is the number of points; each point possesses an individual potential energy function corresponding to the particular \( \lambda \) value.

\[ V_m = (1 - \lambda_m)V_A + \lambda_mV_B \]  

(3)

A new process of multiple step free energy calculations was proposed, which is called \( \lambda \)-dynamics.” In this process, the \( \lambda \) value is treated as a dynamic variable whose value ranges from 1 to \( n \) (Kong & Brooks, 1996).

The main concept involved in this method is that different ligands compete for the same receptor in the body; this competition can be modeled by the simultaneous searching of the various ligands with a varying value of \( \lambda \) (Elber & Karplus, 1990). The use of various intermediate states increases the accuracy of the free energy calculations. A more accurate method for the calculation of free energies was proposed by Bennett (1976) and referred to as the Bennett Acceptance Ratio. The Free energy perturbation (FEP) calculations were performed in Desmond using this method in which a number of simulations were run in parallel for a single FEP calculation.

For the solvation free energy calculation in Desmond, the total free energy by the FEP workflow was employed: the small molecules were annihilated separately in the solvent system with the TIP3P water models, the simulation was run for up to 2 ns, and \( \lambda \) was given a value of 11. A correction value for the energy, which represents the misplaced long-range dispersion energy because of the cutoff of the van der Waal potential energy, was also obtained using FEP calculations.

**Fluorescence studies**

Fluorescence emission spectroscopy was used to analyze the binding behavior of both molecules (A and B) with a G-quadruplex, 5′-TAGGG(TTAGGG)3-3′. A G-rich DNA duplex was used as a control in this study with the following sequence: 5′-GCGCATGCTACGCG-3′ and 3′-CGCGTACGATGCGC-5′.

![Figure 2. Binding sites of G-quadruplex showing the hydrophilic loop regions and the central hydrophobic core formed by the co-facial guanine residues.](image)
Emission spectra were recorded at room temperature using a Hitachi spectrometer (F-700 FL spectrophotometer) with 1 cm path-length quartz cuvette. Each emission spectrum was recorded from 300 to 500 nm with an excitation wavelength of 260 nm. The excitation and emission slits were set to 5 nm each. The emission spectra were recorded by titrating the small molecules with concentrations of up to 50 μM against 10 μM of the G-quadruplex and the duplex DNA. All fluorescence experiments were performed in 10 mM phosphate buffer, 10 mM EDTA, and 100 mM KCl (pH 7.0). Because of solubility problems with molecule A, a small amount of NaOH was added to dissolve it in at the same buffer.

WaterLOGSY NMR experiments
All spectra were recorded at 298 K with a Bruker AVANCE III 500 MHz NMR spectrometer equipped with a 5 mm smart probe. For each sample, a reference spectrum and a 1D WaterLOGSY spectrum were recorded in 10 mM phosphate buffer, 10 mM EDTA, and 100 mM KCl (pH 7.0) containing 10% D2O. The first water-selective 180° pulse was 25 ms long. A weak rectangular PFG was applied for the entire mixing time (1.5 s). A short gradient recovery time of 1 ms was introduced at the end of the mixing time before the detection pulse. The two water-selective 180° square pulses of the double spin-echo sequence were 2.6 ms long. The gradient recovery time was 0.2 s. The data were collected with a sweep width of 10 ppm, an acquisition time of 1.5 s and a relaxation delay of 2.0 s. Prior to Fourier transformation, the data were multiplied with an exponential function with a line broadening of 2.0 Hz. All chemical shifts were referenced using (2,2-Dimethyl-2-silapentane-5-sulfonate sodium salt) as an internal standard.

Results and discussion
Docking Studies
Two types of binding sites were found for the G-quadruplex from SiteMap analysis (Figure 2). By structure, these two binding sites can be classified as: (a) a loop region, which formed small cavities and (b) a co-facial/ end stacking region at the top of the macromolecule (Scheme 1). The loop regions are highly rich in electron density and are favorable for binding with small molecular fragments. These loop regions provide good binding affinity for hydrogen bond acceptors. The other binding site, i.e. the co-facial end stacking region containing the stacked guanine residues (G-face) (forming the first quartet), is enriched with π-electron clouds from all four directions (Scheme 1). This site is important for the design of novel G-quadruplex stabilizing agents because it is the most suitable region for binding of small molecules, peptides or DNA–RNA aptamers (Murat, Singh, & Defrancq, 2011), which can add favorable entropy to the system, resulting in its stabilization. The region was explored as the most suitable region when we docked the previously reported stabilizing agents with the G-quadruplex (Figure 3); hence, we excluded the investigation with the loop binders. This initial docking study not only enables us to explore the favorable binding site for interaction but also provides a means to validate our docking studies (Scheme 1). We screened a set of 14,400 molecules that were available in the Maybridge database using the SP mode of Glide.
The docked poses obtained for the molecules from this study were superimposed with the docked poses obtained for the test set of molecules. The selection criteria for the molecules to be processed in the XP mode of docking were that the Glide SP score was comparable to that of the test set with a rmsd fit of up to 2 Å for their binding pose. A total of 42 molecules of comparable Glide score were selected; these molecules had a docked pose in the co-facial region only and did not have any unusual binding poses in the loop region. Although the basic docking algorithm utilized by Glide is the same for both SP and XP modes, the alternative sampling of ligand conformation with ligand–receptor complementarity and scoring function is more extensive in the XP mode. Hence, with the XP mode of docking, the false positives can be excluded. The top Glide score obtained from the XP mode of docking was –5.65 for Quarflloxin (the docked pose is represented in Figure 3).

Two molecules, A and B, showed satisfactory results in agreement with the score and conformation of Quarflloxin. The two molecules, N’1-imino(2-pyridyl)methyl-3,4,5-trimethoxybenzene-1-carbohydrazide (molecule A) and (3-[4-([3-([4-([2-cyanoethyl)(methyl)amino]benzyli- dene] amino)propyl]imino)methyl]methyl(anilino) propanenitrile) (molecule B) are shown with partial atomic charges in Figure 4. The competence of the docking result was also assessed with a set of ligand decoys containing 1000 molecules. These decoys are nonspecific binders and can bind to any part of the receptor macromolecule. This data-set is known to consist of inactive molecules; they will be treated as false positive even if they bind to the receptor site. Thus, the docking efficiency and accuracy of the 42 selected molecules can be compared with respect to their binding pose using a receiver operating characteristic curve (ROC) (Figure 5). The ROC value for the plot, “sensitivity” vs. “1-specificity”, was .865; interestingly, the area under the fitted curve was found to be .850. This result dem-

Figure 3. Alignment of docked pose of test molecules over G-quadruplex. All molecules are shown in tube representation highlighted with different colors codes. Docked pose of Quarflloxacin is shown in top view and side view. The purple balls indicate the K+ in the G-core of quadruplex.
onstrates that the docking was accurate. The high ROC value (close to 1) indicates that the 42 selected mole-
cules in our study selectively bind to the G-core of the
quadruplexes (Table S1).

Structural stability of drug-bound G-quadruplexes
The high docking score for two molecules, A and B, suggested that they may have a high affinity for G-quad-

MD simulation was performed with both ligand mole-
cules bound to the G-quadruplex. The simulation fetch
average conformations of the drug-like molecule bound
to the G-quadruplex with very few fluctuations in rmsd

Figure 4. Molecule A is N’1-imino (2-pyridyl)methyl-3,4,5-
trimethoxybenzene-1-carboxydrazide. Molecule B is (3-[[4-([3-
(4-[{(2-cyanoethyl)(methyl)amino]benzylidene}amino)propyl]
imino]methyl)(methyl)anilino]propanenitrile) with partial
atomic charges as derived from OPLS force field. The
arrow indicates the H-bond donor (A) and acceptor (B) atoms.

Figure 5. ROC plot as obtained from docking with decoy set
and 42 selected molecules.

Figure 6. RMSD plot for G-quadruplex at its free and bound
states with ligand A and B. (A) a plot of RMSD of nucleobases vs. time, (B) a plot of RMSD of sugar-phosphate
backbone vs. time, (C) a plot of RMSD of all atoms vs. time.
A free-state G-quadruplex was also simulated with the same time scale and used as a control and reference for the simulation results of the drug-bound state. Figure 6 (A–C) shows the rmsd plot for all of the models compared with the initial structure and considering various atoms under investigation. In the ligand-bound state, the fluctuations, based on the rmsd, of the backbone are higher than those for the nucleotide bases. The greater variation in the rmsd of the free-state G-quadruplex than that of the ligand-bound states indicates the absence of steric strain by ligands. Ensemble structures of various MD frames are aligned and shown in Figure 7. The rmsd deviation of molecules A and B bound to G-quadruplexes appears to be almost comparable for the bases and for the backbone; this finding demonstrates that both ligands behave as stabilizing agents with prominent contacts and are potent in controlling the fluctuations of G-quadruplex structures (Table S2). An initial jump in the rmsd scale during first few ps of simulation compared with that of the starting frame corresponds to the relaxation of the model. The stability of the trajectories of the drug-bound states can be observed from 1 to 5 ns, with minimal fluctuations ranging from 0.5 to 1.5 Å. The rmsd of both drug-bound states are limited to this deviation and indicate that the average structures of the bound states are approximately convergent; thus, the MD simulation was terminated after 5 ns. In contrast, the rmsd plot for the free-state G-quadruplex does not appear to have stabilized in 5 ns (Figure 6). Although achieving stability in the free-state model was not the subject of investigation, it was used as a control for the study of A and B in their bound states. A plot of the all-atom rmsd vs. time also explains the similar behaviors (Figure 6(C)). The standard deviations for the all atoms rmsd of A and B were ~0.487 and ~0.427 Å, respectively, whereas that of the free-state G-quadruplex was ~0.697 Å.

**Coulombic and van der Waal interaction energies**

The van der Waal and Coulomb energies are two important contributors to the stabilization of any biomolecule. Table 1 describes the van der Waal and Coulomb interaction energies (kcal.mol$^{-1}$) of G-quadruplex with self-atoms, solvent atoms (water models), and solute atoms of ligands and ions. The Coulomb energy is the major contributor for the stabilization of the drug-bound state of G-quadruplex (Table 1). The Coulomb energy for the model favors binding to the ligand compared with the free-state of the model. The strain of ligand molecules makes the aromatic nuclei of guanine bases stack over each other, which can be judged by the average torsion angle deviations (Table S2). Furthermore, the protonation by the ligand compensates for the electronegativity of O6 by restricting its fluctuations. When shared, the electron orbitals of O6 might become limited to a certain number of degrees of freedom, which ultimately limits the deviation in geometry of the guanine core. This type of effect was shown by comparing the rmsd of O6 specifically over time in all three models, as shown in Figure 8. The central N atom of molecule A is able to

![Figure 7](image_url). Superimposed frames for ligand–quadruplex complex at different time scale. The frames are collected at an interval of 500 ps during the MD simulation. (A) G-quadruplex bound to molecule A, (B) G-quadruplex bound to molecule B.
Scheme 2. Schematic representation of work flow.

Table 1. Describing nonbonded interaction energy$^a$ of G-quadruplex, with and without the ligand molecules.

| Interacting atoms | Coulomb energy | vdw energy | Total energy |
|-------------------|----------------|------------|-------------|
|                   | (kcal. mol$^{-1}$) | (kcal. mol$^{-1}$) | (kcal. mol$^{-1}$) |
| **A$^b$** | | | |
| Self | $-2961.63$ (15.40) | $-138.16$ (12.09) | $-3099.79$ |
| Solute | $-39.03$ (5.16) | $-33.21$ (2.92) | $-72.24$ |
| Solvent | $-2848.97$ (104.66) | $-86.9$ (20.04) | $-2935.87$ |
| **B$^c$** | | | |
| Self | $-2911.46$ (19.89) | $-144.32$ (12.91) | $-3055.78$ |
| Solute | $-35.59$ (6.22) | $-52.6$ (4.32) | $-88.19$ |
| Solvent | $-2736.65$ (141.41) | $-78.48$ (20.66) | $-2815.13$ |
| Free-state | | | |
| Self | $-2845.06$ (29.52) | $-151.96$ (0.97) | $-2985.58$ |
| Solute | | | $-151.96$ |
| Solvent | $-2900.23$ (111.37) | $-117.14$ (21.92) | $-3017.37$ |

$^a$Units of the energy computed are in kcal. mol$^{-1}$.

$^b$Molecule A bound to G-quadruplex.

$^c$Molecule B bound to G-quadruplex.
molecule corresponding energy contributions shown in Table 1. In information). The same difference can be observed in the atoms act as H-bond acceptors (Figure S1 in supporting bond donor, whereas for molecule \( B \) makes a very strong H-bond with O6 and acts as a H-

Figure 8. A plot RMSD of hydrophobic G-core Guanine O6 atoms vs. time for all three models.

make a very strong H-bond with O6 and acts as a H-bond donor, whereas for molecule \( B \), the terminal N atoms act as H-bond acceptors (Figure S1 in supporting information). The same difference can be observed in the corresponding energy contributions shown in Table 1. In molecule \( A \), there are three methoxy groups and one of these methyl groups interacts with the \( \pi \) cloud of the guanine ring. This type of \( \sigma-\pi \) contact is very favorable and is another reason for stable binding of molecule A–G-quadruplex. A diagrammatic explanation is given in the supplementary data (Figure S2 in the supporting information). The Coulomb and van der Waal contributions to the stacking energy are almost comparable for molecules \( A \) and \( B \). As in the free state, there is no ligand interaction with the G-quadruplex; the contribution from Coulomb interaction is almost zero. Therefore, the total energy comes from the van der Waal interactions between two K\(^+\) ions and the G-core. The majority of the effects of the continuum system of the TIP3P water model are from the Coulomb interactions rather than the van der Waal interactions. This interaction is more pronounced in the absence of a ligand. A similar type of explanation is given using a Poisson equation describing the multi-pole representations of the electronic distribution in Figure 9.

**Hydrogen bonding analysis**

Hoogsteen type H-bonding is the key factor behind the topology of the G-quadruplex. The average inter-atomic distances between N1H-O6 and N2H-N7 are shown in Table 2. The hydrophobic core of the first quartet is the most suitable region for the binding of stabilizing agents. Searching for structural deviations in the inter-atomic distances of polar atoms can provide valuable information regarding the interaction of DNA with small molecules. One terminal of molecule \( B \) was buried sideways between the first and second quartet, binding it to the model with a strong H-bond, and the other terminal was above the structure (Figure S3 in the supporting information). The Hoogsteen type H-bonding plays a vital role in the stabilization of guanine structures. The planarity and uniformity of the H-bonds involved in Hoogsteen G-pairing is well controlled by the position of the O6 atom. A plot of rmsd (O6) vs. time for the drug-bound states of the G-quadruplex showed that the displacement of the O6 atom from the plane of the Hoogsteen base pairing is negligible during the 5ns of simulation (Figure 8). For the ligand–receptor complex, we found that the Hoogsteen H-bonding in the G-pairing and H-bonding between the ligand and receptor play a key role in stabilizing the complex; \( \pi-\pi \) stacking interactions also enable the correct positioning of the ligands over the G-quadruplex for the entire time scale. As indicated in Figure 4, the atoms of ligands that are emphasized with the arrow H-bond to the macromolecule. The two hydrogen atoms that have a central protonated N atom (H-bond donor) in molecule \( A \) make hydrogen bonds with the O6 of guanines (G5 and G23) (Figure S1 in the supporting information). In contrast, one terminal N atom, a H-bond acceptor, from a cyanide (–CN) group of molecule \( B \) makes a H-bond with the amine group (–N(2)H) of a guanine residue (G4) in the second quartet (Figure S1 in the supporting information) (below the facial quartet). The average H-bond distances computed for molecules \( A \) and \( B \) were ~2.54 and ~2.45 Å, respectively, which emphasizes that H-bonds are the main contributor to the binding of both ligands to the G-quadruplex. The strength of the H-bonding for the H-bond acceptor in molecule \( B \) is much greater than that of the H-bond donor of molecule \( A \). This finding can be seen in a histogram plot (Figure 10), which counts the frequency of H-bond distances for both molecules. The H-bonding pattern provides a steric restriction on the rotational freedom of the C–C bond in the ligands. This restriction in the rotational degrees of freedom reinforces the stacking of the aromatic moiety of the ligand over the G-core. The H-bonding arrangement of molecule \( B \) is biased toward the side of the molecule in which the geometry is most affected by the conformational constraints of the ligand–DNA complex. This bias comprises the flexibility to its other half; the dynamicity is shown in a plot of distance vs. time (Supplementary data Figure S3). Examining the overall structure indicates that a small molecule should be rationally designed with a N atom that could be utilized as a donor (molecule \( A \)) or acceptor (molecule \( B \)) to form a contact with the G-core; doing so can strengthen strong binding in the system.

**Electrostatic Iso-surface by APBS**

A standard practice in molecular biophysics is the evaluation of the electrostatic properties of a ligand–macromolecule complex. The APBS, developed by the group of Nathan Baker, is a routine using the PBE to illustrate
the electrostatic attributes of a biomolecular complex; thus, the molecular electrostatic surface can be understood and used to develop a good stabilizing agent for G-quadruplexes. Figure 9(A) shows the electrostatic iso-surface contours of the G-quadruplex model with positive and negative iso-surfaces at ±4 kT/e. If we carefully analyze the G-quadruplex model, the macromolecule can be assumed to be a system with highly anionic characteristics. For the ease of comparison, a top view of the G-quadruplex model is shown with the phosphorous atoms in sphere representation (Figure 9(B)), indicating the relative position of the phosphate groups, sugars, and bases in the quadruplex topology. In the G-quadruplex structure, the electronegative charge is diffused in two ways: (1) the negative charges of the backbone phosphates surround the nucleobases and sugar moieties and (2) the lone pair electrons of guanine O6 atoms are directed toward the center, increasing the electron density in the G-core. Because of the dense cloud caused by negative charge diffusion, the iso-surface contours of the

Figure 9. Electrostatic iso-surface contours calculated using APBS. The positive and negative iso-surface was calculated at ±4 kT/e. (A) electrostatic iso-surfaces of free-state G-Quadruplex, (B) top view of G-quadruplex model 5'-TAGGG (TTAGGG)3, showing the phosphorous atoms in sphere representation and relative positional arrangement of sugars and bases in the topology, (C) and (D) electrostatic iso-surfaces of ligand-bound state G-quadruplex with ligands, A and B.
G-quadruplex form a large dome shape. The π-electrons of the purine rings also add to the overall electronegativity of the system. When we compared the electrostatic properties of the G-quadruplex system bound to the drug, as shown in Figure 9(C) and (D), we found a reduced overall charge in the system in the hydrophobic region. Because of this reduction, a G-quadruplex can only bind to (and thus recognize) a molecule if it has sufficient positive electrostatic fields. In molecule A, the methyl protons also interact with the aromatic system of the preceding guanine base, which tightens the binding of molecule A with its host. The electrostatic iso-surfaces of the G-quadruplex bound to both molecules A and B were obtained at ± 4 kT/e. The electrostatic iso-surfaces of both ligand systems are shown in Figure 11, where the iso-surfaces were achieved at ± 2 kT/e, to illustrate the broad range of electropositive values. The partial atomic charges of individual molecules were assigned according to the OPLS force-field and are shown in Figure 4.

Solvation free energy

The solvation energies of molecules A and B were computed in pure solvent. The annihilation of the compounds resulted in a total solvent energy of $-48.7 \pm 0.3$ kcal mol$^{-1}$ and $-74.5 \pm 0.6$ kcal mol$^{-1}$ for molecules A and B, respectively. These values are the lowest terms of the 11 λ windows that were separately annihilated to find the various possible low energy conformations in which the ligands were expected to bind G-quadruplexes. Because the ligands contained no intra-perturbed groups for the solvation free energy calculation, the interaction of both of the small molecules with the solvent must be favorable. In molecule A, the central charged amine group is bound to an aromatic system (pyridyl group) from one side, which hinders the solvation shell of the amino group. Similarly for molecule B, the two charged central amino groups are affected by a benzylidene group that utilizes the lone pair availability of the N atom and thus increases the solvent accessibility. This effect can be understood in terms of the solvation free energy difference between molecules A and B. Graphical plots

Table 2. Average distances$^a$ between various heavy atoms in G-quadruplex.

| Free-state DNA | N1H1-O6 | N2H1-N7 | O6-O6a | O6-O6b | O6-O6c |
|----------------|---------|---------|--------|--------|--------|
| Q-1            | 1.94    | 2.72    | 3.42   | 4.82   | 3.73   |
| Q-2            | 1.94    | 2.47    | 4.01   | 5.68   | 3.75   |
| Q-3            | 1.98    | 2.72    | 3.83   | 5.3    | 3.78   |
| A$^b$          |         |         |        |        |        |
| Q-1            | 1.92    | 2.61    | 3.52   | 5.14   | 3.60   |
| Q-2            | 1.93    | 2.85    | 3.93   | 5.5    | 3.71   |
| Q-3            | 1.99    | 2.73    | 3.83   | 5.37   | 3.82   |
| B$^c$          |         |         |        |        |        |
| Q-1            | 1.87    | 2.34    | 3.63   | 5.00   | 4.15   |
| Q-2            | 2.04    | 2.62    | 4.23   | 5.71   | 3.95   |
| Q-3            | 2.69    | 3.16    | 3.76   | 6.54   | 3.75   |

Notes: O6-O6a – average distance between O6 of nearest residue of same quartet.O6-O6b – average distance between O6 of diagonal guanine residues.O6-O6c – average distance between O6 of stacked guanine residue of two quartets.

$^a$Units of distances calculated are in angstrom (Å).

$^b$Molecule A bound to G-Quadruplex.

$^c$Molecule B bound to G-Quadruplex.
of the energy variations, solvation free energy change $\Delta G$ (kcal.mol$^{-1}$) in the 2-ns time scale for the molecules and their sliding free energy change $\Delta G$ (kcal.mol$^{-1}$) computed for a sliding time of 1.5 ns, are shown in Figure 12. The initial fluctuations visible for molecule B are attributed to the conformational flexibility of the molecule, which later becomes almost constant. The overall solvent accessibility of the molecules can be easily seen in these plots, which show a good correlation between the solvent model and the structural features of the respective molecules. To the best of our knowledge, there is no report of a study emphasizing the calculation of the standard molecular solvation free energy of G-quadruplex stabilizing agents. Without this experimental value, it is very difficult to validate our theoretical data.

When analyzing the plots for both ligands for the solvation free energy with the same time scale, molecule B shows a relatively higher free energy. This finding
indicates that molecule B is more solvated than molecule A. These data are based on the larger solvent accessible surface area (SASA) of B and attributed to the higher degree of freedom in the state bound to the G-quadruplex. The solvation free energy calculation is also accompanied by a correction factor, which is typically known as a long-range dispersion energy correction. The correction values were −3.0 and −4.2 for molecules A and B, respectively. The details of the solvation free energy calculations are given in Table 3. Few of the other relevant properties computed from QikProp can be correlated with the discussion of solvation free energy in Figure 12. (Upper panel) Plot of solvation free energy (G) for A (left) and B (Right) as calculated from FEP with respect to time. (Lower panel) Plot of sliding energy (G) for A (left) and B (right) with respect to time. Error bar indicated the upper and lower range as conferred among the 11 λ windows used for annihilation.

### Table 3. Results of solvation free energy\(^a\) computed for molecules.

| Parameters                              | Molecule A | Molecule B |
|-----------------------------------------|------------|------------|
| ΔG solvent transfer                     | −48.7 ± 0.3| −74.5 ± 0.6|
| ΔG solvent solvation                    | −50.6 ± 0.3| −76.4 ± 0.6|
| ΔG solvent long range dispersion correction | −3.0       | −4.2       |

\(^a\)Units of solvation free energy computed is kcal. mol\(^−1\).

### Table 4. Computed SASA and related properties of both molecules.

| Properties\(^a\)           | Molecule A | Molecule B | Range\(^b\)      |
|----------------------------|------------|------------|------------------|
| SASA                       | 636.79     | 868.86     | 300.0–1000.0     |
| FOSA                       | 268.93     | 435.33     | 0.0–750.0        |
| FISA                       | 137.22     | 155.00     | 7.0–330          |
| PISA                       | 230.63     | 278.51     | 0.0–450.0        |
| Solvent accessible volume  | 1077.40    | 1520.01    | 500.0–2000.0     |
| Computed dipole moment     | 2.08       | 8.90       | 1.0–12.5         |

\(^a\)SASA – Total SASA in Å\(^2\). FOSA – Hydrophobic component of SASA; saturated C and attached H. FISA – Hydrophilic component of SASA: N, O, and H on hetero atoms. PISA – Carbon and attached H component of SASA. Units for solvent accessible volume are in Å\(^3\).

\(^b\)Range indicates the values desired for drug-like compound.
Table 5. In silico predicted ADMET properties of both molecules.

| Properties\(^a\) | Molecule A | Molecule B | Range\(^b\) |
|------------------|------------|------------|-------------|
| QPlogHERG (predicted IC\(_{50}\)) | \(-5.90\) | \(-7.13\) | Concern below \(-5\) |
| QPPCaco | 494.98 | 335.76 | <25 poor, >500 great |
| QPlogBB | \(-1.35\) | \(-2.31\) | \(-3.0\) to \(1.2\) |
| QPPMDCK | 231.33 | 152.07 | <25 poor, >500 great |
| QPlogKp | \(-2.47\) | \(-1.86\) | \(-8.0\) to \(-1.0\) |
| QPlogKhsa | \(-0.31\) | 0.21 | \(-1.5\) to \(1.5\) |
| Percentage of human oral absorption | 86.33 | 100 |  |
| Formal charge | 1.0 | 2.0 |  |

\(^a\)QPlogHERG – Predicted IC\(_{50}\) value for HERG K\(^+\) channels. QPPCaco – Predicted apparent Caco-2 cell (model for gut-blood barrier) permeability in nm/s. QPlogBB – Predicted brain/blood partition coefficient. QPPMDCK – Predicted apparent MDCK cell (model for blood-brain barrier) permeability in nm/s. QPlogKp – Predicted skin permeability. QPlogKhsa – Prediction of binding to human serum albumin.

\(^b\)Range indicates the values desired for drug-like compound.

Figure 13. Fluorescence emission spectra of G-quadruplex (represented here as GQ) in presence of molecule A (A) and molecule B (B), determination of G-quadruplex binding affinity of molecule A (C) and molecule B (D) following the changes in emission maxima (\(\lambda_{\text{max}}\)) of G-quadruplex in the context of ligand concentrations. Fluorescence emission spectra of GC-rich duplex DNA (represented here as DD) in presence of molecule A (E) and molecule B (F).
Interestingly, all of the SASA parameters fall in the permissible range for a drug-like compound, indicating that these molecules are strong candidates for further pharmaceutical research.

**ADMET analysis**

Analyzing pharmacokinetic properties helps us to determine the pharmaceutical information and many significant physical characteristics of both molecules, A and B. A total of 44 peer pharmacological and pharmacokinetic properties have been measured. The molecular data-set we used for the screening were in accordance with the basic criteria of Lipinski’s rule of five for drug-likeness. A similar result was obtained using QikProp analysis (mol. wts.: 330.34 and 414.55 Dalton, H-bond donors: 3 and 0, H-bond acceptors: 7 and 8, and Octanol-water partition coefficients [log P]: 1.9 and 4.5, respectively, for A and B), which shows that the in silico predicted results are accurate for other descriptors. The other relevant properties are described in Table 5. Both compounds are Central Nervous System (CNS) -inactive; therefore, it can be hypothesized that the molecules have no central nervous system activity. All of the other predicted properties described in Table 4 appear to be within the permissible range.

### Binding of the G-quadruplex with small molecules measured using fluorescence spectroscopy

To validate the theoretical study, we used the intrinsic fluorescence of the G-quadruplex to monitor its binding with molecules A and B. A significant hypsochromic shift or blue shift in the fluorescence emission provides strong evidence for whether a drug/ligand molecule binds to the G-quadruplex. The fluorescence experiment was performed by titrating molecules A and B (0–50 μM) with the G-quadruplex, 5′-TAGGG (TTAGGG)₃ (10 μM). The characteristic emission of the spectrum is shown in Figure 13. A total of 10–18 nm hypsochromic shifts in the G-quadruplex emission spectra were observed when either molecule A or B was bound to the G-quadruplex (Figure 13(A) and (C)). The changes in the fluorescence emission maxima of the G-quadruplex as a function of the concentration of the ligands, molecules A and B, yields equilibrium dissociation constants ($K_d$) of 137 ± 1.1 and 31 ± 1.2 μM, respectively (Figure 13(B) and (D)). Molecule B is a better binder to the G-quadruplex than molecule A. A control experiment was performed for a GC-rich DNA duplex, 5′-GCGCATGCTACGCG-3′-3′-CGCAGTGATGCGC-5′ in the presence of molecules A and B. Interestingly, no blue shifts of the emission spectra of the DNA duplex occurred with the
addition of the ligands (Figures 13(E) and (F)), but a significant increase in the intensity in the fluorescence spectra upon the addition of the ligands was observed. To determine the cause of the unexpected increase in fluorescence intensity, we measured the fluorescence spectra of only the ligands, A and B, without any interference of the host, the G-quadruplex (Figure S4 in the supporting information). We observed an increase in fluorescence intensity in the emission spectra at ~340 nm for both molecules when they were excited at 260 nm (Figure S4 in the supporting information). This finding explains why the G-quadruplex and DNA duplex showed increases in fluorescence intensity upon the addition of ligands A and B. However, experimental evidence such as blue shift of G-quadruplex upon addition to the molecules A and B clearly showed that molecules A and B bind much more selectively to the G-quadruplex structures than to the DNA duplex.

**Water ligand observed via gradient spectroscopy (WaterLOGSY) NMR experiments**

Water ligand observed via gradient spectroscopy (WaterLOGSY) NMR experiments allow the detection of binding of small molecules to protein targets with dissociation constants in the low micromolar to millimolar range (Dalvit et al., 2000; Dalvit, Fogliatto, Stewart, Veronesi, & Stockman, 2001). Compared with a saturation transfer difference NMR experiment (Bhunia, Saravanan, Mohanram, Mangoni, & Bhattacharjya, 2011; Mayer & Meyer, 1999), which is also an important technique for measuring the binding of small molecules to receptor proteins, WaterLOGSY is more effective for studying the interactions of nucleic acids and small molecules (Bhunia, Bhattacharjya, & Chatterjee, 2012; Lepre, Moore, & Peng, 2004). The WaterLOGSY signal is produced by the selective transfer of bulk water magnetization to the ligand in solution via the macromolecule–ligand complex. In the resulting spectra, the ligands that bind generally have an opposite sign to those of compounds that do not bind.

Figure 14 shows the one-dimensional WaterLOGSY NMR spectra of the small molecules, A and B, in the presence of the G-quadruplex. Interestingly, both molecules bind to the G-quadruplex, as indicated by the positive signals. The aromatic ring protons and the methyl protons of molecule A show positive signals in the presence of the G-quadruplex. The simulation data suggests that the three methoxy groups attached to the benzene moiety are in close proximity to the G5 and G23 groups of the G-quadruplex. The aromatic ring protons of benzene and the pyridyl groups of molecule A severely overlap at 7.35 and 7.52 ppm (Figure 14B). However, the MD simulation data indicate that the aromatic rings, benzene and pyridyl, are stacked to G11, G5, and G17, respectively. The putative binding site and the MD simulation data correlate well with the WaterLOGSY data (Figure 14C). In contrast, the intensity of the peaks of molecule B is significantly stronger than that of molecule A, when binding to the G-quadruplex. Interestingly, all of the protons, except for those in two ethylene groups of molecule B, show positive WaterLOGSY signals. This result demonstrates that the ethylene groups (3.12 and 3.82 ppm) do not interact with the G-quadruplex as indicated by the negative signals (Figure 14E). This finding is not surprising because the MD simulation data clearly show that one terminal of molecule B is quite flexible and that the ethylene groups point toward the solution (Figure 14F). Both N–Me groups (3.67 and 3.38 ppm) and both aromatic ring protons (6.90 and 7.83 ppm) and the flexible linkers in the center show strong WaterLOGSY signals. In the simulated model, G17 and G23 are in close proximity to one terminal of molecule B, whereas G5, G10, and G23 are very close to the other terminal of the G-quadruplex (Figure 14F). The flexible linker, attached to both aromatic ring protons of molecule B, stacked to the G-faces, such as G11, G17, and G5 of G-quadruplex. As a control experiment, we measured the WaterLOGSY signals of molecules A and B in the absence and presence of GC-rich duplex DNA (Figure S5 in the supporting information). None of the molecules bound to duplex DNA structures. The WaterLOGSY NMR experiments in conjunction with the MD simulation characterize the binding epitope of both molecules, which bind to the G-quadruplex at an atomic resolution.

**A final hypothesis: requirements for the rational design of G-quadruplex stabilizing agents:**

From previous literature and this study, we elaborate some of the key features, which can be implemented for the future rational design of a G-quadruplex stabilizing agent. In designing a suitable stabilizing agent for guanine-rich secondary structures, the primary concern is to understand the structural criteria needed to complement the G-quadruplex. Some of the criteria for a rational design of novel G-quadruplex stabilizing agents are as follows:

1. A molecule with enhanced aromaticity is essential to effectively interact with G-quadruplexes. π–π stacking interactions are key requirements for any molecule to be attached to the facial guanine quartets.
2. H-bond acceptability is the main parameter for a ligand to bridge with the loop surrounding the G-core. In molecule B, nitrogen can act as a Lewis base and donate its lone pair of electrons to form a hydrogen bond, which helps place molecule B...
above the quadruplex structure. This property is important because it establishes a criterion to develop stabilizing agents irrespective of the facial topology of G-quadruplex.

(3) In contrast, the donation of a H-bond to the quartets by a ligand molecule is a prerequisite for the facial G-quadruplex stabilization. Ligand molecules with this characteristic will enable the utilization of the lone pairs available from the O6 of guanines. This effect is more pronounced when the H-bond donor moiety is found in the central region of the ligand molecule, as in molecule A. Apart from this, the protonation to this moiety provides a positive potential directed toward the central hydrophobic region.

(4) The methyl-π type interaction also helps to stabilize the ligand and G-quartet assembly. Ligands with a methyl moiety localized on the top of the central part of the G-quadruplex try to stabilize the host–guest assembly with methyl-π type dipole-induced dipole interactions, as is the case for molecule A. This type of nonpolar interaction from the mixing of orbitals of the methyl-proton with the π electron in the purine ring of guanine is of higher order. This interaction can provide a very stable and strong binding affinity with a minimal margin of excluded volumes to any organic core molecule.

(5) The radius of gyration of the molecule is another important criterion. The average distance between the guanine residues of a quartet from all types of G-quadruplex structures is within a range of 15–17Å. If the radius of gyration of the molecule falls in this range, it can successfully find a conformational orientation over the hydrophobic core of the G-quadruplex. Additionally, the middle portion of the molecule should have rotational degrees of rotation with respect to the C–C bond to be able to adjust to the planarity of the guanine face. The radius of gyration of a molecule is related to the dimension of the central core and not of the functional groups attached to the core.

(6) Electrostatic interaction has a more significant role than any other physical descriptor. In our study, we found the G-quadruplex structure to be highly anionic with abundant electronic density surrounding the structure from all directions. The incoming ligand must have sufficient cationic characteristics. In other words, the incoming ligand should have adequate electron-withdrawing effects to stay and bind over the planarity of the quadruplex.

Conclusion

In this study, we revealed that the two molecules, A and B, are pharmacologically potent in arresting G-quadruplex structures in DNA. The data from MD simulations indicate that these screened analogs are capable of binding the quadruplex within a reasonable time. Fluorescence experiments were performed by titrating molecules A and B (0–50 μM) with the quartet DNA of sequence, 5′-TAGGG (TTAGGG)3 (10 μM). A large blue shift was found in the emission spectra of the G-quadruplex (the fluorescence emission of the G-quartet is at 337 nm) with stepwise increase in the ligand molecule concentration over the concentration of the G-quadruplex. Molecules A and B give emission blue shifts of 10–18 nm, respectively, demonstrating that both molecules bind to the G-face of the quadruplex. A one-dimensional WaterLOGSY experiment further proved at atomic resolution that both molecules, A and B, bind to the G-quadruplex structure; the structural information of the bound complex is well correlated with the MDs simulation. Our experimental investigation is underway with these molecules. The framework of these two molecules can be the platform to design novel anticancerous therapeutics in the near future.

Supplementary material

The supplementary material for this paper is available online at http://dx.doi.10.1080/07391102.2012.742246.

Acknowledgments

Anirban Bhunia and Subhrangsu Chatterjee would like to thank the Director of the Bose Institute for providing the initial research support. Subhrangsu Chatterjee would like to thank DST, the Government of India for the Ramanujan Fellowship. The Central Instrument Facility (CIF) of Bose Institute is greatly acknowledged for providing the Fluorescence instrument facility. AB and SC thank Dr. Helena Kovacs (Bruker Biospin AG, Switzerland) for the WaterLOGSY pulse program development and useful technical discussion. We also thank Mr. Sudipto K. Goswami (Bruker India Scientific Pvt. Ltd.) for the fruitful discussion.

References

Anthony, N. G., Huchet, G., Johnston, B. F., Parkinson, J. A., Suckling, C. J., Waigh, R. D., & Mackay, S. P. (2005). In silico footprinting of ligands binding to the minor groove of DNA. Journal of Chemical Information and Modelling, 45, 1896–1907.

Baker, N. A., Sept, D., Joseph, S., Holst, M. J., & Cammon, M. C. (2001). Electrostatics of nanosystems: Application to microtubules and the ribosome. Proceedings of the National Academy of Sciences USA, 98, 10037–10041.

Balasubramanian, S., & Neidle, S. (2009). G-quadruplex nucleic acids as therapeutic targets. Current Opinion in Chemical Biology, 13, 345–353.
Heddi, B., & Phan, A. T. (2011). Structure of human telomeric DNA in crowded solution. *Journal of American Chemical Society*, 133, 9824–9833.

Holt, P. A., Buscaglia, R., Trent, J. O., & Chaires, J. B. (2011). A discovery funnel for nucleic acid binding drug candidates. *Drug Development Research*, 72, 178–186.

Hu, J., Hwang, S. S., Lisesa, M., Gan, B., Sahin, E., Jaskelioff, M., … DePinho, R. A. (2012). Antitelomerase therapy provokes ALT and mitochondrial adaptive mechanisms in cancer. *Cell*, 148, 651–663.

Hu, J., Hwang, S. S., Lisesa, M., Gan, B., Sahin, E., Jaskelioff, M., … Hoff, D. V. (1999). Effects of cationic porphyrins as G-quadruplex interactive agents in human tumor cells. *Cancer Research*, 59, 639–644.

Jain, A. S., Grand, C. L., Bearss, D. J., & Hurley, L. H. (2002). Direct evidence for a G-quadruplex in a promoter region and its targeting with a small molecule to repress c-MYC transcription. *Proceedings of the National Academy of Sciences USA*, 99, 11593–11598.

Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W., & Klein, M. L. (1983). Comparison of simple potential functions for simulating liquid water. *Journal of Chemical Physics*, 79, 926–935.

Jorgensen, W. L., Maxwell, D. S., & Tirado-Rives, J. (1996). Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids. *Journal of American Chemical Society*, 118, 11225–11236.

Kim, M. Y., Vankayalapati, H., Shin-ya, K., Wierzba, K., & Hurley, L. H. (2002). Telomestatin, a potent telomerase inhibitor that interacts quite specifically with the human telomeric intramolecular G-quadruplex. *Journal of American Chemical Society*, 124, 2088–2089.

Kirkwood, J. G. (1935). Statistical mechanics of fluid mixtures. *Journal of Chemical Physics*, 3, 300–313.

Kollman, P. (1993). Free energy calculations: Applications to chemical and biochemical phenomena. *Chemical Review*, 93, 2395–2417.

Kong, X. J., & Brooks, C. L. (1996). Lambda-dynamics: A new approach to free energy calculations. *Journal of Chemical Physics*, 105, 2414–2423.

Krutler, V., Gunsteren, W. F. V., & Hennenberger, P. H. (2001). A fast SHAKE algorithm to solve distance constraint equations for small molecules in molecular dynamics simulations. *Journal of Computational Chemistry*, 22, 501–508.

Lange, T. (2009). How telomeres solve the end-protection problem. *Science*, 326, 948–952.

Lepre, C. A., Moore, J. M., & Peng, J. W. (2004). Theory and applications of NMR-based screening in pharmaceutical research. *Chemical Reviews*, 104, 3641–3676.

Li, Y., Shen, J., Sun, X., Li, W., Liu, G., & Tang, Y. (2010). Accuracy assessment of protein-based docking programs against RNA targets. *Journal of Chemical Information and Modelling*, 50, 1134–1146.

Li, Z., Tan, J. H., He, J. H., Long, Y., Ou, T. M., Li, D., … Huang, Z. S. (2012). Disubstitutedquinazoline derivatives as a new type of highly selective ligands for telomeric G-quadruplex DNA. *European Journal of Medicinal Chemistry*, 47, 299–311.

Lim, K. M., Amnane, S., Bouaziz, S., Xu, W., Yu, M., Patel, D. J., … Phan, A. T. (2009). Structure of the human telomere in K+ solution: A stable basket-type G-quadruplex with only two G-tetrad solution. *Journal of American Chemical Society*, 131, 4301–4309.

Lin, J., Jin, R., Zhang, B., Chen, H., Bai, Y. X., Yang, X., … Huang, J. J. (2008). Nucleolar localization of TERT is unrelated to telomerase function in human cells. *Journal of Cell Science*, 121, 2169–2176.

Lipinski, C. A., Lombardo, F., Dominy, B. W., & Feeney, P. J. (2001). Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Advanced Drug Delivery Reviews*, 46, 3–26.

Luu, K. N., Phan, A. T., Kuryavyi, V., Lacroix, L., & Patel, D. J. (2008). Structure of the human telomere in K+ solution: An intramolecular (3+1) G-quadruplex scaffold. *Journal of American Chemical Society*, 129, 9963–9970.

Ma, D. L., Chan, D. S. H., Lee, P., Kwan, M. H. T., & Leugn, C. H. (2011). Molecular modeling of drugDNA interactions: Virtual screening to structure-based design. *Biochimie*, 93, 1252–1266.

Martins, C., Gunaratnam, M., Stuurt, J., Makwana, V., Greciano, O., Reszka, A. P., … Neidle, S. (2007). Structure-based design of benzylamino-acridine compounds as G-quadruplex DNA telomere targeting agents. *Bioorganic & Medicinal Chemistry Letters*, 17, 2293–2298.

Mayer, M., & Meyer, B. (1999). Characterization of ligand binding by saturation transfer difference NMR spectroscopy. *Angewandte Chemie*, 38, 1784–1788.

McCarrick, M. A., & Kollman, P. A. (1999). Predicting relative binding affinities of non-peptide HIV protease inhibitors with free energy perturbation calculations. *Journal of Computer-Aided Molecular Design*, 13, 109–121.

Monchaud, D., & Teulade-Fichou, M. P. (2008). A hitchiker’s guide to G-quadruplex ligands. *Organic and Biomolecular Chemistry*, 6, 627–636.

Murat, P., Singh, Y., & Defrançq, E. (2011). Methods for investigating G-quadruplex DNA/ligand interactions. *Chemical Society Reviews*, 40, 5293–5307.

Neidle, S., & Read, M. A. (2001). G-quadruplexes as therapeutic targets. *Biopolymers*, 56, 195–208.

Neidle, S. (2009). The structures of quadruplex nucleic acids and their drug complexes. *Current Opinion in Structural Biology*, 19, 239–250.

Neidle, S., & Parkinson, G. N. (2008). Quadruplex DNA crystal structures and drug design. *Biochimie*, 90, 1184–1196.

Neidle, S., & Thurston, D. E. (2005). Chemical approaches to the discovery and development of cancer therapies. *Nature Reviews*, 5, 285–296.

Pederu, A., Pagano, B., Petronzi, C., Massa, A., Esposito, V., Virgilio, A., … Filosa, R. (2011). Design, synthesis, biophysical and biological studies of trisubstitutednaphthalimidines as G-quadruplex ligands. *Bioorganic & Medicinal Chemistry*, 19, 6419–6429.

QikProp. (2011). version 3.4. Schrödinger. New York, NY: LLC.

PyMOL Molecular Graphics System, Version 1.2r3pre, Schrödinger, LLC.

Read, M., Harrison, R. J., Romagnoli, B., Tanius, F. A., Gowen, S. H., Reszka, A. P., … Neidle, S. (2001). Structure-based design of selective and potent G quadruplex-mediated telomerase inhibitors. *Proceedings of the National Academy of Sciences USA*, 98, 4844–4849.

Renčiuk, D., Kejnovska, I., Skolakova, P., Bednarova, K., Motlova, J., & Vorlickova, M. (2009). Arrangements of human telomere DNA quadruplex in physiologically relevant K+ solutions. *Nucleic Acid Research*, 37, 6625–6634.
Rodriguez, R., Muller, S., Yeoman, J. A., Trentesaux, C., Riou, J. F., & Balasubramanian, S. (2008). A novel small molecule that alters shelterin integrity and triggers a DNA-damage response at telomeres. *Journal of American Chemical Society, 130*, 15758–15759.

Rodriguez-Brenes, I. A., & Peskin, C. S. (2009). Quantitative theory of telomere length regulation and cellular senescence. *Proceedings of the National Academy of Sciences USA, 107*, 5387–5392.

Shivakumar, D., Williams, J., Wu, Y., Damm, W., Shelley, J., & Sherman, W. (2010). Prediction of absolute solvation free energies using molecular dynamics free energy perturbation and the OPLS force field. *Journal of Chemical Theory and Computation, 6*, 1509–1519.

Simonsson, T., Pecinka, P., & Kubista, M. (1998). DNA tetraplex formation in the control region of c-myc. *Nucleic Acid Research, 26*, 1167–1172.

SiteMap. (2011). *version 2.5*, Schrödinger. New York, NY: LLC.

Srivastava, H. K., Chourasia, M., Kumar, D., & Sastry, G. N. (2011). Comparison of computational methods to model DNA minor groove binders. *Journal of Chemical Information and Modelling, 51*, 558–571.

Sun, D., Guo, K., Rusche, J. J., & Hurley, L. H. (2005). Facilitation of a structural transition in the polypurine/polypyrimidine tract within the proximal promoter region of the human VEGF gene by the presence of potassium and G-quadruplex-interactive agents. *Nucleic Acid Research, 33*, 6070–6080.

Sun, D., & Hurley, L. H. (2001). Targeting telomeres and telomerase. *Methods in Enzymology, 340*, 573–577.

Tauchi, T., Shin-ya, K., Sashida, G., Sumi, M., Okabe, S., Ohyashiki, J. H., & Ohyashiki, K. (2006). Telomerase inhibition with a novel G-quadruplex-interactive agent, telomestatin: in vitro and in vivo studies in acute leukemia. *Oncogene, 25*, 5719–5725.

Triballeau, N., Acher, I., Pin, J. P., & Bertrand, H. O. (2005). Virtual screening workflow development guided by the "Receiver Operating Characteristic" curve approach. Application to high-throughput docking on metabotropic glutamate receptor subtype 4. *Journal of Medicinal Chemistry, 48*, 2534–2547.

Ward, R. J., & Auterxier, C. (2005). Pharmacological telomerase inhibition can sensitize drug-resistant and drug-sensitive cells to chemotherapeutic treatment. *Molecular Pharmacology, 68*, 779–786.

Warshel, A. P., & Levitt, M. (1976). Theoretical studies of enzymic reactions: Dielectric, electrostatic and steric stabilization of the carbonium ion in the reaction of lysozyme. *Journal of Molecular Biology, 103*, 227–249.

Zambre, V. P., Murumkar, P. R., Giridhar, R., & Yadav, M. R. (2010). Development of highly predictive 3D-QSAR CoMFA models for anthraquinone and acridone derivatives as telomerase inhibitors targeting G-quadruplex DNA telomere. *Journal of Molecular Graphics and Modelling, 29*, 229–239.

Zaug, A. J., Podell, E. R., & Cech, T. R. (2005). Human POT1 disrupts telomeric G-quadruplexes allowing telomerase extension in vitro. *Proceedings of the National Academy of Sciences USA, 102*, 10864–10869.

Zhen, L. H., Xi, Z., Chattopadhyaya, J., & Hecht, S. M. (2011). G-quadruplex and its ligands in anticancer therapy. *Medicinal chemistry of nucleic acids, 206*–257.

Zwanzig, R. W. (1954). High-temperature equation of state by a perturbation method. I Nonpolar gases. *Journal of Chemical Physics, 22*, 1420–1426.