Structural Characterization of a Guanine−Quadruplex Ligand Complex[†]

Martin A. Read and Stephen Neidle*

The CRC Biomolecular Structure Unit, Chester Beatty Laboratories, The Institute of Cancer Research, Fulham Road, London SW3 6JB, UK

Received July 7, 2000; Revised Manuscript Received August 25, 2000

ABSTRACT: The inhibition of telomerase by molecules such as disubstituted amidoanthraquinones is believed to be due to their stabilization of guanine−quadruplex complexes. The characterization is reported of a complex with the intermolecular parallel quadruplex formed from the sequence TGGGGT and a 1,4-bis-piperidino amidoanthraquinone. Crystals obtained did not give single-crystal diffraction; the fiber-like pattern has been interpreted in terms of a repeating unit with four guanine-quartets and two stacked/intercalated ligand molecules. The two categories of possible structures for the complex consistent with this interpretation have been examined by molecular dynamics simulations, with fully solvated environments and 1000 ps simulation times. The two central guanine-quartets in the intercalation model rapidly became highly distorted, whereas the two types of models with ligand stacked externally on the ends of the quadruplex remained very stable. It was concluded that the externally bound ligand complexes best represent the structure of this quadruplex complex, in agreement with earlier NMR results on related systems.

The ends of chromosomes, termed telomeres, comprise specialized guanine-rich DNA sequences with tandem repeats of simple motifs (*1*), such as $(TTAGGG)_n$ in *Homo sapiens* and $(TTGGGG)_n$ in the ciliate *Tetrahymena*. Telomeric DNA serves to protect chromosome ends from damage and recombination. It is double-stranded for most of its length, but the extreme 3′ end is single-stranded. Under conditions of high sodium or potassium ions, telomeric DNA can associate or fold in a variety of intermolecular or intramolecular ways (*2*) to form four-stranded quadruplex structures with characteristic Hoogsteen hydrogen-bonded quartets of guanine bases (Figure 1, panel a). There have been two crystal structures of *Oxytrichia* telomeric DNA sequences published, of d(TGGGGT) and d(GGGGTTTTGGGG) (*3*, *4*). The former structure is formed by four strands arranged in an all-parallel fashion with all the glycosidic bonds in an anti conformation and a stack of four guanine-quartets. This arrangement has been confirmed by NMR studies in solution (*5*).

The enzyme telomerase is a specialized DNA polymerase (*6*) (reverse transcriptase) which synthesizes further telomere repeats on to the 3′ ends of existing telomeric DNA, via an endogenous RNA template, using the terminal telomeric repeat as a primer. Telomerase is activated in the majority of human tumor cells yet is not present, or is inactivated, in somatic cells (*7*). This finding is in accord with the observations of progressive decrease in telomere length upon successive rounds of replication in normal cells as compared with shortened but stable telomeres in tumor cells. This suggests that the ability of tumor cells to maintain their telomeres is an essential element of the immortalization process (*8*).

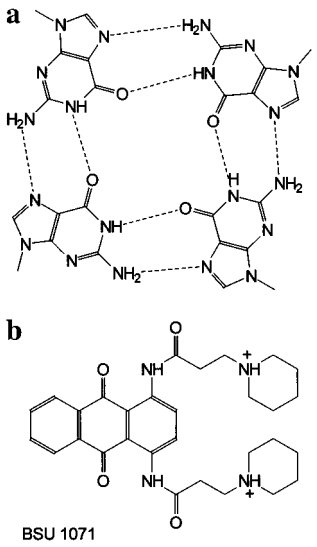

FIGURE 1: (a) Arrangement of guanine bases in the guanine quartet. Dashed lines indicate hydrogen bonds. (b) Structure of the anthraquinone BSU1071.

Telomerase is thus potentially a highly selective anticancer target (*9*). Its inhibition has been demonstrated to lead to a loss of cellular proliferation ability and cell death in tumor cells (*10*, *11*). A distinct strategy has emerged for the design of inhibitors, based on the stabilization of higher-order folding of telomeric DNA repeats (*12*). Such folding into quadruplexes has been shown to effectively inhibit telomeric DNA from acting as a primer for telomerase and thereby

[†] We are grateful to the Cancer Research Campaign for support (Program Grant SP1384 to S.N.) and the Institute of Cancer Research for a research studentship to M.A.R.

* To whom correspondence should be addressed. Chester Beatty Laboratories, The Institute of Cancer Research, 237 Fulham Road, London SW3 6JB, UK. Telephone: 0044 0207 970 6043; Fax: 0044 0207 352 8039; E-mail: s.neidle@icr.ac.uk.

10.1021/bi001584k CCC: $19.00 © 2000 American Chemical Society
Published on Web 10/10/2000

inhibiting the synthesis of further telomeric repeats. It has been shown that a number of amido-anthraquinone derivatives disubstituted at the 1,4-, 1,5-, 1,8-, 2,6-, or 2,7- positions with amidoalkylamino substituents, are effective telomerase inhibitors, with $^{tel}IC_{50}$ values in the 1−10 μM range (*13a*−*d*) and that inhibition depends on the formation of a guanine−quadruplex complex with the ligand (*13a*). Other compounds possessing planar aromatic chromophores have also been found to inhibit telomerase, notably, tetra-(*N*-methylpyridyl) porphyrins (*14*), 2,7-disubstituted fluorenones (*15*), and acridines (*16*), as well as the perylene derivative PIPER (*17*). Stabilization of complexes involving tetra-(*N*-methyl-pyridyl) porphyrin and other porphyrins (*18*), an amidoanthraquinone, and a perylene diimide with guanine quadruplex structures have been demonstrated by UV/visible and NMR spectroscopy. These data have been interpreted in terms of models with the ligands binding external to the stacks of guanine-quartets within a quadruplex, rather than being intercalated between stacks (*14b*, *17a*). A detailed structural model is only available (*17a*) for the 2:1 PIPER-d[TTAGGG]$_4$ complex (from an NMR study); in part this may be due to the mobile nature of the quadruplex−ligand complexes examined to date.

We have previously shown (*16b*) that structural information derived from molecular modeling studies can aid the rational design of new and more effective telomerase inhibitors. The initial objective of the present studies was to obtain experimental structural data from co-crystallization of a DNA quadruplex sequence with the 1,4-disubstituted amidoanthraquinone compound BSU1071 (Figure 1, panel b), which is an active telomerase inhibitor at the low micromolar level. This initial goal was not realized since the crystals that we have obtained did not give single-crystal diffraction. However, the fiber pattern produced, in conjunction with molecular dynamics simulations, have enabled us to define plausible structures for the complex, which are also consistent with the features suggested in several other studies on quadruplex−ligand complexes. We have also taken into account the stoichiometry of this ligand-quadruplex complex, as determined by UV titration as 2:1 in ligand/quadruplex (Dosanjh, Ph.D. Thesis, University of London, 2000, to be published). Thus, all models examined have been consistent with this stoichiometry.

We have used the well-characterized parallel quadruplex (*4*, *5*) formed from four strands of the sequence d(TGGGGT) as a template for binding this ligand. In principle, there are two distinct categories of intercalative-type binding sites for a ligand such as BSU1071: (i) being intercalated within the set of four guanine quartets and (ii) being stacked externally to the quartets. Several possible models are shown schematically in Figure 2. This study has focused on models A and C.

MATERIALS AND METHODS

Crystallization Trials. The DNA sequence d(TGGGGT) was purchased from the Oswell DNA service in a HPLC-purified form. The synthesis of the ligand BSU1071 was as previously reported (*13b*,*c*). Various crystallization conditions were initially tried, and after several systematic variations of methyl-2,4-pentane-dione (MPD), spermine, and salt

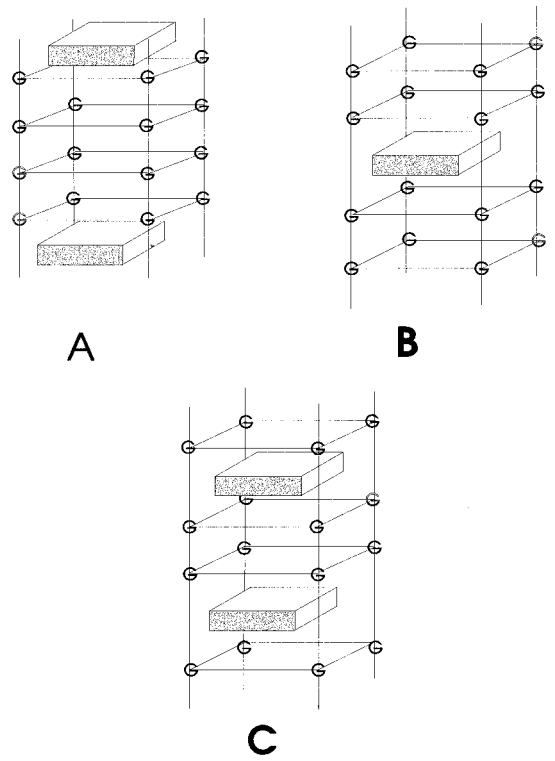

FIGURE 2: Schematic of three possible arrangements for a complex between the [d(GGGG)]$_4$ quadruplex and a ligand such as BSU1071. (A) External binding; (B) a single symmetric intercalation site; (C) two intercalation sites.

concentrations, crystals were obtained at 13 °C using the hanging-drop method and with the following conditions: 0.5 mM oligonucleotide d(TGGGGT); 0.5 mM BSU1071; 5 mM spermine tetrahydrochloride; 50 mM $MgCl_2$; 5% MPD. All oligonucleotide solutions were prepared in 40 mM sodium cacodylate buffer at pH 6.5. Small crystals were first observed after 4−6 weeks, starting as needlelike rods with reservoir concentrations in the range of 50−60% MPD. After 2−3 months, the crystals obtained with BSU1071 were of good morphology and highly colored even when removed from the crystal solution, indicating the presence of ligand in the crystal lattice.

Diffraction Experiments. Crystals were examined at liquid nitrogen temperatures on an RAXIS-II image plate detector (Molecular Structure Corporation), equipped with a rotating anode X-ray source and Yale mirror optics. The crystals chosen were approximately 150 × 200 × 300 μm in dimensions, and all had a deep red coloration and rod-like shape. Single-crystal diffraction was not observed, despite the apparent good quality of the crystals, as evidenced by their morphology.

Molecular Modeling. Molecular modeling and simulation studies on quadruplex−BSU1071 complexes with both external and internal intercalation sites were carried out. The initial coordinates for the model were taken from the crystal structure of the d(TGGGGT) quadruplex (Nucleic Acid Database entry no. UD0009), selecting one of the four molecules in the crystallographic asymmetric unit. The crystal structure (*4*) shows that the terminal thymine residues of each d(TGGGGT) strand project away from the helical axis of the molecule and do not participate in the quadruplex motif, but rather they are more fundamentally involved in

three-dimensional crystal lattice interactions. Accordingly, both the 3′ and the 5′ thymine residues were omitted from all the modeling and simulations, leaving a pure d(GGGG) quadruplex, since the primary aim was to investigate the dynamic properties of the ligand−quadruplex interactions.

A quadruplex structure with two internal intercalation sites (Figure 2, panel C) was constructed. Initially, a pair of G-quartets at each end of the $[d(GGGG)]_4$ quadruplex were interactively moved apart to a separation of 6.8 Å from their original 3.4 Å using molecular modeling tools within the InsightII package (Molecular Simulations Inc, San Diego, CA, 1999) and the XLEAP module within the AMBER v5.0 package (*19*). The Cornell et al force field (*20*) was used for the DNA. This initial construction resulted in poor geometry at several points in the phosphate backbone of each strand, which was relieved by the selective minimization of the affected areas while restraining the individual G-quartets to their crystallographic geometries.

A molecular model of the anthraquinone drug BSU1071 was constructed and minimized using the InsightII package. Partial charges were calculated semiempirically using the AM1 formalism in the MOPAC program. Force-field parameters were obtained from the cff force field within the Insight II package. The ligands were docked into the intercalation site with the AFFINITY docking program, part of the InsightII suite. This program incorporates the use of manual and automatic docking procedures in combination with molecular mechanics within the DISCOVER module of InsightII. The docking method employed enables nonbonded van der Waals and electrostatic interactions to be simultaneously monitored during the docking, and several possible conformations for the ligand were evaluated interactively. "Flexible" ligand docking was then used to define the lowest energy position of each drug using a Monte Carlo-based automated docking protocol. This uses a random iterative algorithm to sample changes in torsion angles and atomic positions while simultaneously recalculating internal and interaction energies. If the energy becomes lower than the initial value, then the new position and conformation is accepted by the algorithm; however, if the energy is higher the change is rejected, and the possess is repeated with a new random change in torsion angles. The automated docking procedure then selected the best structures and subjected the totality of the binding site, including the closest phosphate atoms and sugars, in a predefined subset, to 500 steps of molecular mechanics minimization. During the docking, the four torsion angles between the methylene carbon atoms of the side chains on the ligand were allowed to vary. By allowing specific parts of the binding site and ligand to move during the protocol, the steric clashes produced with normal fixed docking methods may be alleviated, resulting in a greater number of sampled conformations and a closer approach to the global minimum.

The most stable ligand orientations for the internal intercalation complex are shown in Figure 3. The side chains reside in the grooves of the quadruplex, with the chromophore of the ligands positioned directly over the center of the adjacent G-quartet. This model of the complex was then subjected to molecular dynamics simulation using the AMBER 5.0 suite of molecular modeling programs. Parameters for the atom types in the BSU1071 molecule were taken from the AMBER parameter set if available there or were extrapolated from values in the cff force field.

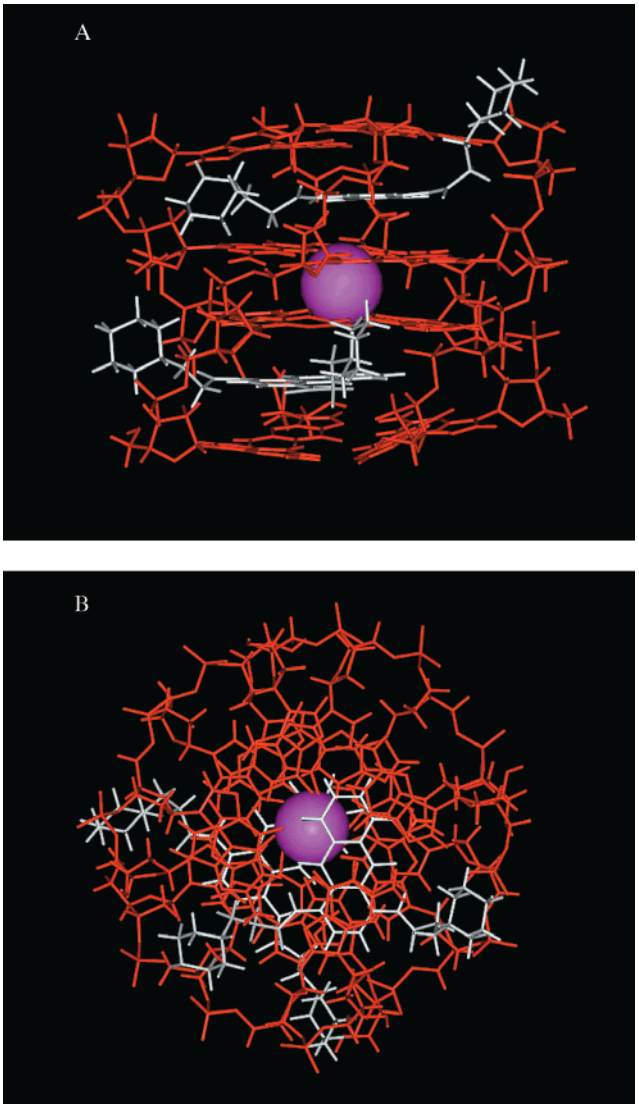

FIGURE 3: Two views of the starting conformations of the internal intercalation model, prior to dynamics simulation. Sodium ions are shown in magenta and the BSU1071 molecules are colored white.

The $[d(GGGG)]_4$−BSU1071 complex was initially solvated in a periodic TIP3 water box, of approximate dimensions 45 × 45 × 45 Å and which extended at least 10 Å from any solute atom. An internal sodium ion was positioned in the central channel between the middle two G-quartets using crystallographically determined distances from the d(TGGGGT) crystal structure. The sodium ion was positioned to allow coordination with the four polarized carbonyl oxygen atoms on two adjacent G-quartets, completing the octahedral coordination sphere. The entire system then required a further 12 sodium ions and 5 chlorine ions for complete neutralization, as each drug contains two cationic charges due to the protonation of the amine side chains. The sodium counterions were initially placed into the most negative locations using the Coulombic potential terms with the LEAP module of the AMBER 5.0 package.

Calculations were carried out using the SANDER module of AMBER 5.0 with the SHAKE algorithm enabled for the hydrogen atoms and a 2 fs time step. Simulations were performed using the Berendsen temperature-coupling algo-

rithm with a time constants of 0.2 ps. The nonbonded pairs list was updated every 20 steps. Each complex was subjected to 1000 steps of conjugate gradient minimization followed by 25 ps of equilibration dynamics at 300 K with full constraints on the DNA of the complex and an 11 Å nonbonded Lennard−Jones and electrostatic cutoff. After this initial equilibration, the Particle Mesh Ewald (P. M. E.) summation term was then activated for all subsequent simulations, for the inclusion of long-range electrostatic interactions into the calculation without truncation. The P. M. E. charge grid spacing was approximately 1.0 Å, and the charge grid was chosen to be products of the powers of 2, 3, and 5 to ensure efficiency of the fast Fourier transform calculation. The complex was then gradually allowed to relax as restraints were sequentially lowered over seven successive rounds of conjugate gradient minimizations. For dynamics runs after minimizations, the initial velocities were assigned using the standard Maxwellian distributions. The final production run of 1000 ps was then carried out on a SGI multiprocessor Origin 200 computer. For comparative purposes, the native G-quadruplex system $[d(GGGG)]_4$ with no bound ligands was also simulated with the same protocol and under identical periodic P. M. E. conditions. Coordinates were written to the output files every 1 ps for the analysis of the trajectory. The trajectory analysis was carried out using the CARNAL and RDPARM modules available in the AMBER 5.0 suite.

Further modeling studies were undertaken to evaluate the possible conformations of an externally stacking $[d(GGGG)]_4$−BSU1071 system (Figure 2, panel A), with again two ligand molecules bound to the quadruplex. Starting coordinates were again taken from the d(TGGGGT) crystal structure. The previously minimized models of the BSU1071 molecule were manually docked, in a stereochemically reasonable orientation, adjacent to each of the terminal G-quartets on the $[d(GGGG)]_4$ quadruplex, resulting in a complex with two bound ligand molecules. The AFFINITY module of InsightII was then used in the same way as described above for the internal intercalation complex to find the most energetically favorable positions for the ligands. Several low-energy ones were found. However, most were equivalent due to the symmetrical nature of the quadruplex and the ligand binding sites. These structures reduced to two distinct groups. The two models in each group with the lowest energy were selected for the next stage of the simulation. (Both models represented many equivalent structures with very similar calculated interaction energies.) Figure 4, panel a shows the first model, where three of the four G-quadruplex grooves are occupied by the two ligands' cationic side chains, one groove containing two side chains in the same groove; this is designated as model A. Conversely, model B (Figure 4, panel b) represents the group where each side chain, on the two separate ligands, resides in different grooves of the quadruplex. The two systems were then treated to molecular dynamics simulations using the AMBER 5.0 package exactly as described above, so a direct comparison of the trajectories could be made. The sole necessary exception was that these models required two additional sodium ions positioned inside the G-quadruplex to replace the drug molecules from the internal intercalation model. The complexes were solvated in a periodic TIP3 water box, of approximate dimensions

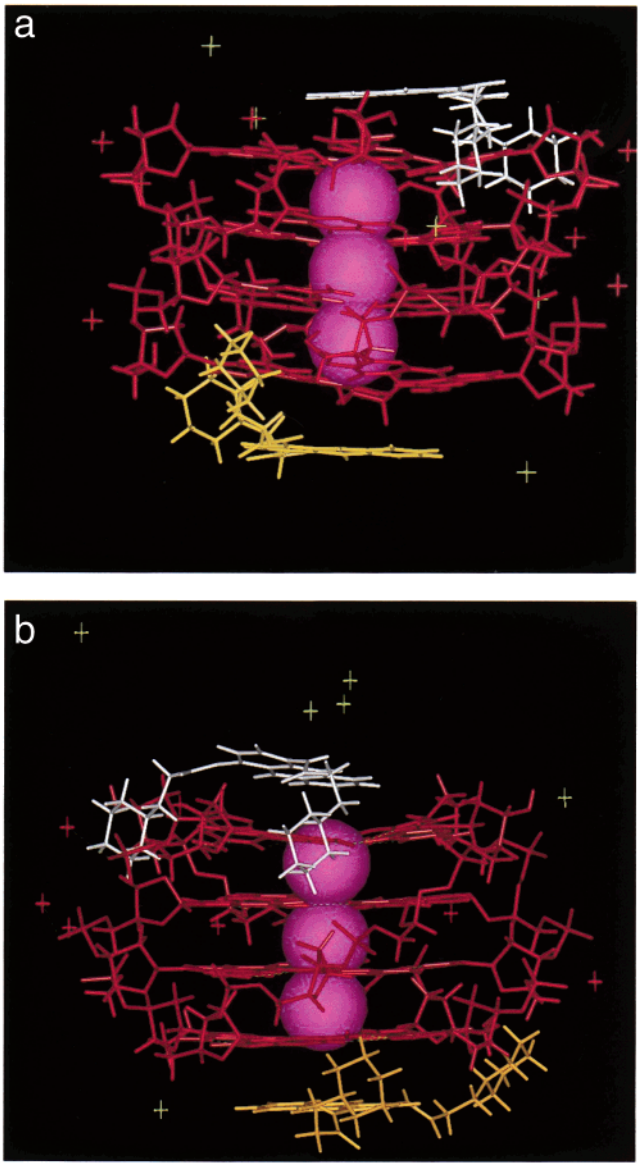

FIGURE 4: Two initial externally bound ligand complexes, prior to dynamics simulations. (a) Has one groove of the quadruplex occupied by side-chains from two ligand molecules; (b) has the two side-chains residing in two separate grooves.

45 × 45 × 45 Å, and which extended at least 10 Å from any solute atom. Internal sodium ions were positioned in the central channel using crystallographically determined distances from the d(TGGGGT) structure.

A parallel simulation of the native $[d(GGGG)]_4$ quadruplex was undertaken, using times and conditions identical to the above, with treatment of bound sodium ions identical to those in the externally bound ligand simulations.

RESULTS

The Diffraction Study. Single crystals, with color indicating bound ligand molecule gave a diffraction pattern on an image plate frame (Figure 5) typical of diffraction from a paracrystalline fiber (e.g. of a polynucleotide), rather than of an ordered single crystal. The pattern has some overall analogies to those from paracrystalline fibers of B-DNA (*21*), as well as those of polyG, and strongly suggests that the bases are all parallel to each other and at right-angles to a helix axis.

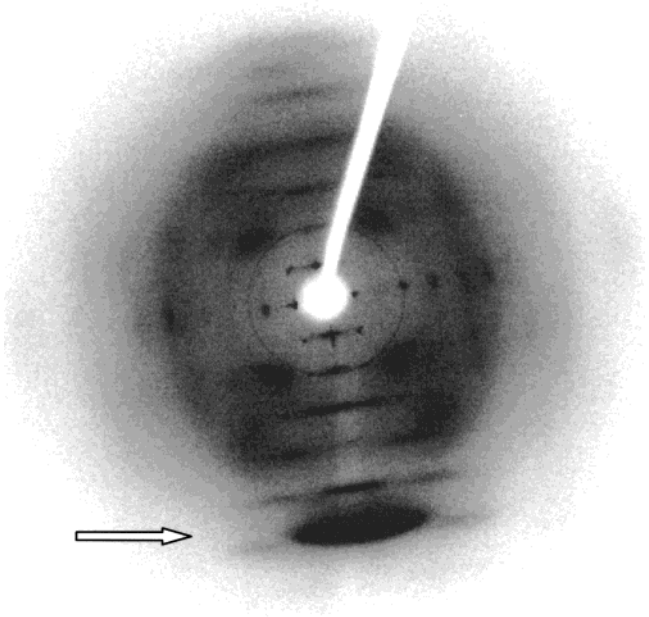

FIGURE 5: Diffraction pattern from a crystal of the BSU1071 complex with d(TGGGGT). The 3.4-Å layer line is indicated.

It corresponds to a symmetric, disordered structure, albeit with a helical arrangement being evident. Although it has not been possible to index individual diffraction maxima, the pattern of layer lines can be readily interpreted. The strong intensity at 3.4 Å is indicative of a structure with pronounced stacking. This occurs at the sixth layer line and corresponds to six stacked planar groups and a pitch of ca. 20 Å. Such a repeating unit of six planes is thus only consistent with models containing four guanine quartets plus two molecules of BSU1071. This is in turn consistent with each BSU1071 molecule being either externally stacked or internally intercalated into the stacks, i.e., to the models shown in Figure 2, panels A and C (although at this stage we cannot discount an asymmetric model with one ligand molecule intercalated as in Figure 2, panel B, and a second externally stacked). The stoichiometry of the complex deduced from the diffraction pattern is also consistent with the conclusions based on UV titration data (see above). It definitively rules out a model with the anthraquinone rings bound in and orientated along the grooves of the quadruplex.

Simulation Studies on the Intercalation Complex Model. Two views of the final snapshot for the trajectory of the internal intercalation model are shown in Figure 6. There has been considerable distortion of the geometry of the two terminal G-quartets during the simulation, and there has been large-scale deterioration of the initial internal intercalation model. The plot of rms deviations (Figure 7, panel a) over the total 1000 ps of the molecular dynamics production run shows a sudden change at the beginning of the trajectory, to an rmsd of 2.0 Å. This occurs after the gradual relaxation of restraints on the system (it is notable from Figure 7, panel b, that the system as a whole is energetically stable during the production run, although the rms fluctuations in total and potential energy, of 295 and 192 kcal/mol, respectively, are much greater than those for the stacked models; see below). The majority of the distortions occur to the two external

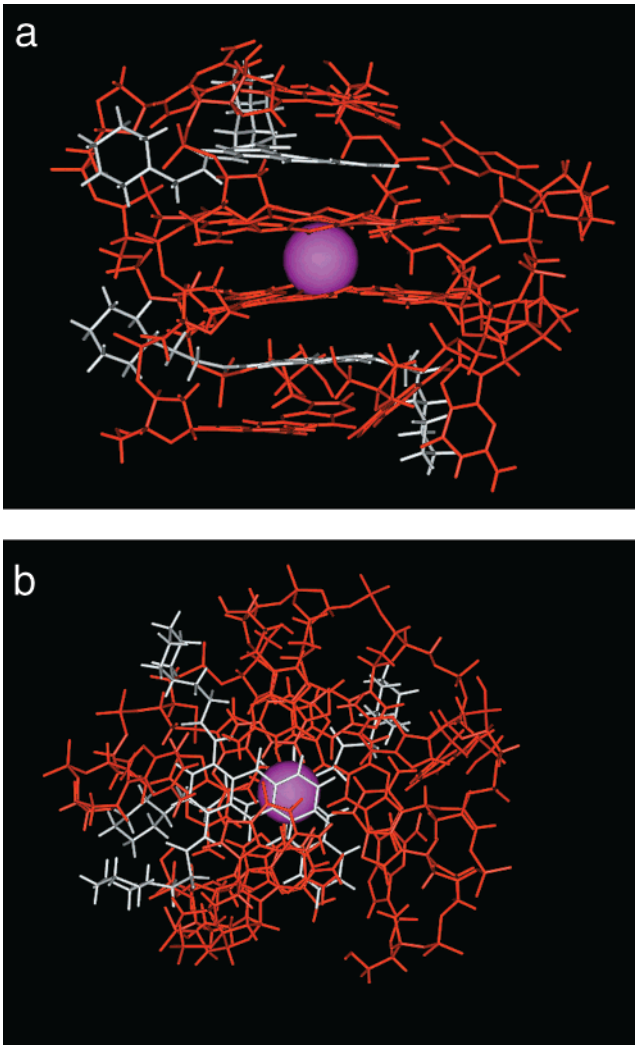

FIGURE 6: Two views of the final trajectory from the simulation of the intercalation model of Figure 2, panel A. Note the distortions to the outer G-quartets, although the central ones remain relatively intact.

Table 1: Average Hydrogen Bond Distances for the G-quartets in the Internal Intercalation Model (Figure 7) over the Last 600 ps of Molecular Dynamics[a]

G-quartet number	ave distance for N7...N2 bond (Å)	deviation from ave	ave distance for O6...N1 bond (Å)	deviation from ave
first	4.895	0.951	4.505	0.683
second	2.889	0.131	2.968	0.172
third	2.922	0.110	2.923	0.149
fourth	4.153	0.590	3.683	0.617

[a] The deviations from the averaged distances are also given. Values for the terminal G-quartets, adjacent to the bound BSU1071 molecules, are listed in the first and fourth rows

G-quartets, with their array of cyclic hydrogen bonds being disrupted by influence of the bound BSU1071 molecules. This is detailed in Table 1, where the average donor−acceptor distances for Hoogsteen hydrogen bonds in the two terminal G-quartets are increased, by between 0.7 and 1.9 Å as compared to the values in the internal G-quartets. The extent of the disruption is clearly seen in the plot of a single N2...N7 hydrogen bond distance for an external G-quartet over the entire trajectory (Figure 7, panel c). The terminal guanine bases rapidly move out from their G-quartet posi-

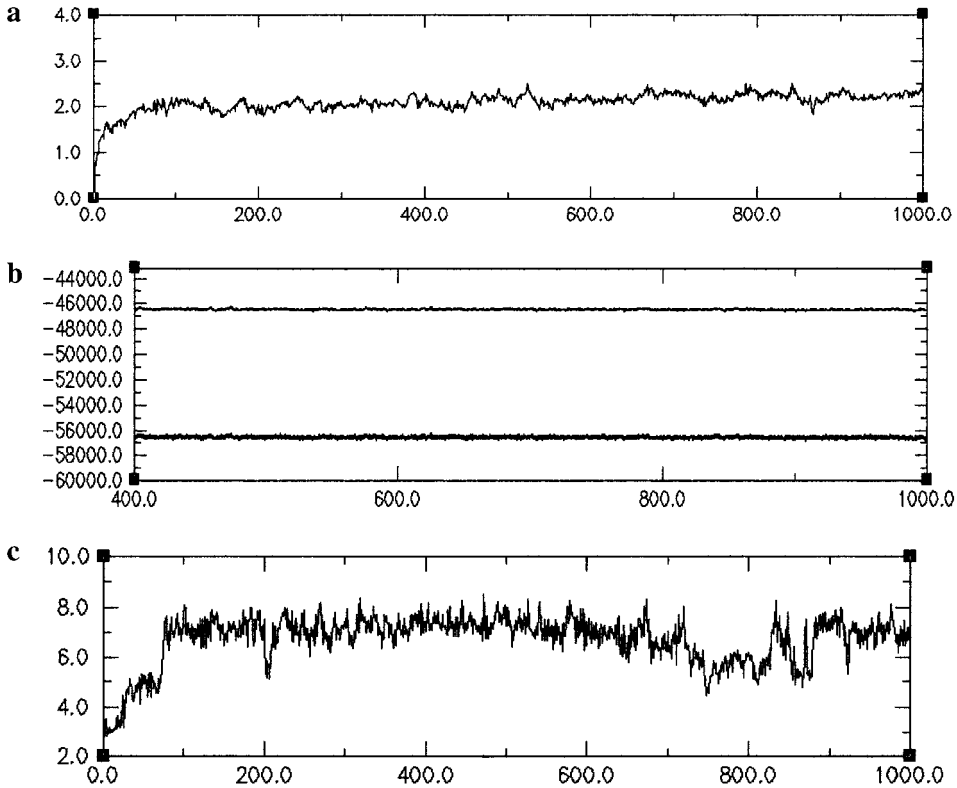

FIGURE 7: Plots for the trajectory of the intercalation model. (a) Showing rms deviations from the averaged structure, in Å. The time-scale is in ps; (b) plots of potential and total energies, in kcal/mol during the simulation; (c) variation in N2...N7 Hoogsteen hydrogen bond distance for one of the external G-quartets during the simulation.

tions after the relaxation of restraints, and there is a concomitant increase in N2...N7 (and O6...N1) distances, indicative of an almost complete deterioration of the hydrogen bond and G-quartet architecture.

Simulations of the Externally Bound Ligand Complexes Models A and B and the Native G-Quadruplex Systems. The structures of the two starting quadruplex−ligand complexes (Figure 4) show a distinction between the two conformations due to the ability of the drugs to anchor themselves in distinct ways into the grooves of the quadruplex. However, when the molecular dynamics trajectory for each of the simulations was examined, it was evident that the flexibility of the side chains allows the entire BSU1071 molecules to rotate through many different positions, so that many of the snapshots were closely similar. The plots of the total and potential energies over the final 600 ps of each trajectory verify that the systems rapidly remained at equilibrium (Figure 8). This was consistent with the temperatures and pressure trajectories of the systems, which also remain virtually constant throughout the molecular dynamics production runs. The rms fluctuations in total and potential energy for models A and B are, respectively, 84 and 103 kcal/mol, and 82 and 101 kcal/mol.

Several key features are apparent after the 1000 ps production run for all these three simulations. Principally, the overall G-quadruplex architectures for all three of these theoretical structures remain intact (Figure 9) and are in excellent overall agreement with the crystal structure geometry (*4*) used for the starting structures. This was reflected in the low rms differences between the starting structures and those formed during each trajectory (Figure 10, panels a, c, and e). Other characteristics of the $[d(GGGG)]_4$ structure, such as all-anti glycosidic bond conformation, helical twists, interstrand distances, hydrogen bond distances, and planarity of the G-quartets were maintained throughout the simulations, highlighting the stability of these systems (see below). The rmsd plots for models A and B (Figure 10, panel a and c) have slightly higher values than the plot for the native quadruplex model (Figure 10, panel e), indicative of the movement and flexibility by the BSU1071 molecules in these complexes. This is highlighted by the rmsd values for the BSU1071 molecules alone in both model A and model B, which show large movements over the course of the simulations (Figure 10, panels b and d). Visual examination of the side chains of the BSU1071 molecule in several snapshots during the trajectories shows that the six-membered piperidine rings form transient weak hydrophobic interactions with the sugars of the terminal G-quartets, as the ligands undergo transitions from one conformation to another. A schematic representation of the interconversion of the two models A and B is shown in Figure 11, panel a. Figure 11, panel b, shows the superposition of orientations for one BSU1071 molecule through the complete molecular dynamics simulation of model A, where the movement of the anthraquinone chromophore and marked flexibility of the side are apparent. The anthraquinone chromophores at each end of the G-quadruplex structure for both model A and model B remain at the distance of 3.4 Å from the terminal G-quartets, with little variation during the simulations.

Analysis of the hydrogen bond distances for all the Hoogsteen N1...O6 and N2...N7 hydrogen bonds between the guanine bases in the four quartets of each quadruplex model shows that all hydrogen bonds remain intact, including those in the terminal G-quartets (Figure 12, panels a−c;

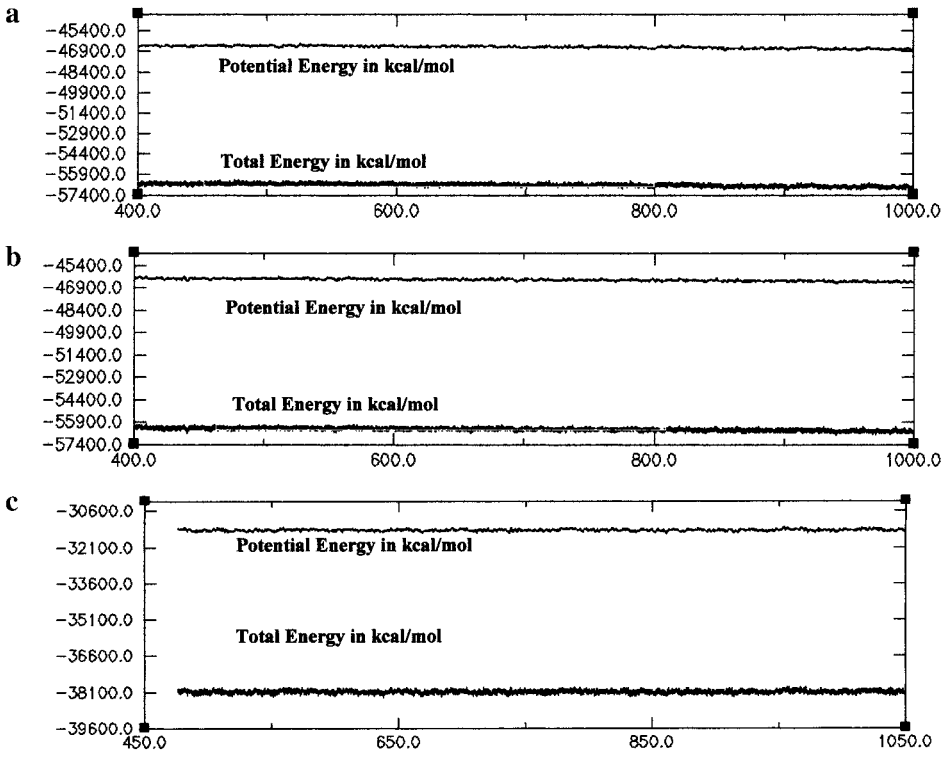

FIGURE 8: Plots of potential and total energies, in kcal/mole, for (a) the external ligand complex, model A; (b) the external ligand complex, model B; (c) the native G-quadruplex.

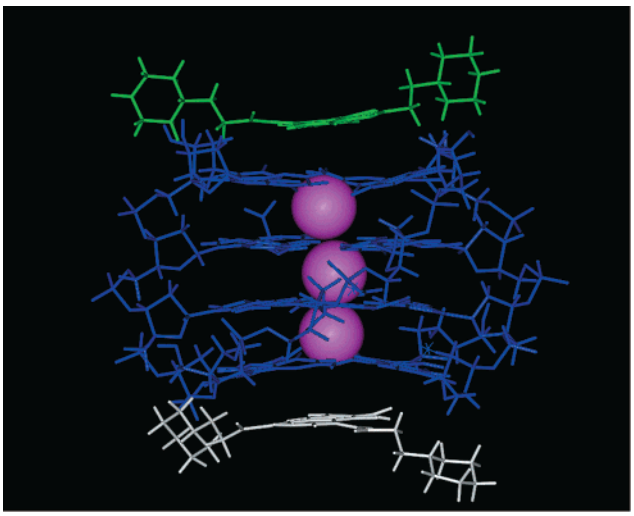

FIGURE 9: Plots of the structure in the final snapshot of the simulation for the external ligand complex, model A. The simulated G-quadruplex structures for model B of the external ligand complex and the native G-quadruplex are closely analogous, and are therefore not shown here.

Tables 2−4). There is no significant difference in hydrogen-bond distances between those in the inner and terminal G-quartets.

Table 2: Average Hydrogen Bond Distances for the G-quartets in the External-Binding Model A, over the Last 600 ps of Molecular Dynamics[a]

G-quartet number	ave distance for N7...N2 bond (Å)	deviation from ave	ave distance for O6...N1 bond (Å)	deviation from ave
first	2.953	0.106	3.038	0.154
second	2.990	0.138	3.274	0.247
third	3.000	0.139	3.285	0.220
fourth	2.936	0.108	3.044	0.162

[a] The external G-quartets, adjacent to the ligands, are listed in the first and fourth rows.

Table 3: Average Hydrogen Bond Distances for the G-quartets in External-Binding Model B over the Last 600 ps of the Molecular Dynamics Simulation[a]

G-quartet number	ave distance for N7...N2 bond (Å)	deviation from ave	ave distance for O6...N1 bond (Å)	deviation from ave
first	2.939	0.103	3.043	0.160
second	3.000	0.144	3.280	0.234
third	3.002	0.146	3.302	0.228
fourth	2.940	0.108	3.059	0.158

[a] The external G-quartets, adjacent to the ligands, are listed in the first and fourth rows.

DISCUSSION

The Nature of the Crystalline Complex. The inability of the crystalline complex between d(TGGGGT) and BSU1071 to give discrete single-crystal diffraction is revealing of inherent disorder. Although we are unable to precisely define the detailed nature of this disorder, the diffraction pattern actually obtained enables us to unequivocally define the nature of the repeating unit and thus to rule out some categories of models. That with a single intercalation site embedded within the quadruplex, as suggested elsewhere (*22*), i.e., between G-quartets 2 and 3 (Figure 2, panel B), is ruled out by the presence of six layer lines in the diffraction pattern, as well as by the observed stoichiometry. Asymmetric variants on this arrangement, with a stacked ligand at one terminus, are in principle possible. However, we are unconvinced of their feasibility on account of the behavior found here for the quadruplex when an intercalation site between G-quartets 1 and 2 is introduced (Figure 2, panel

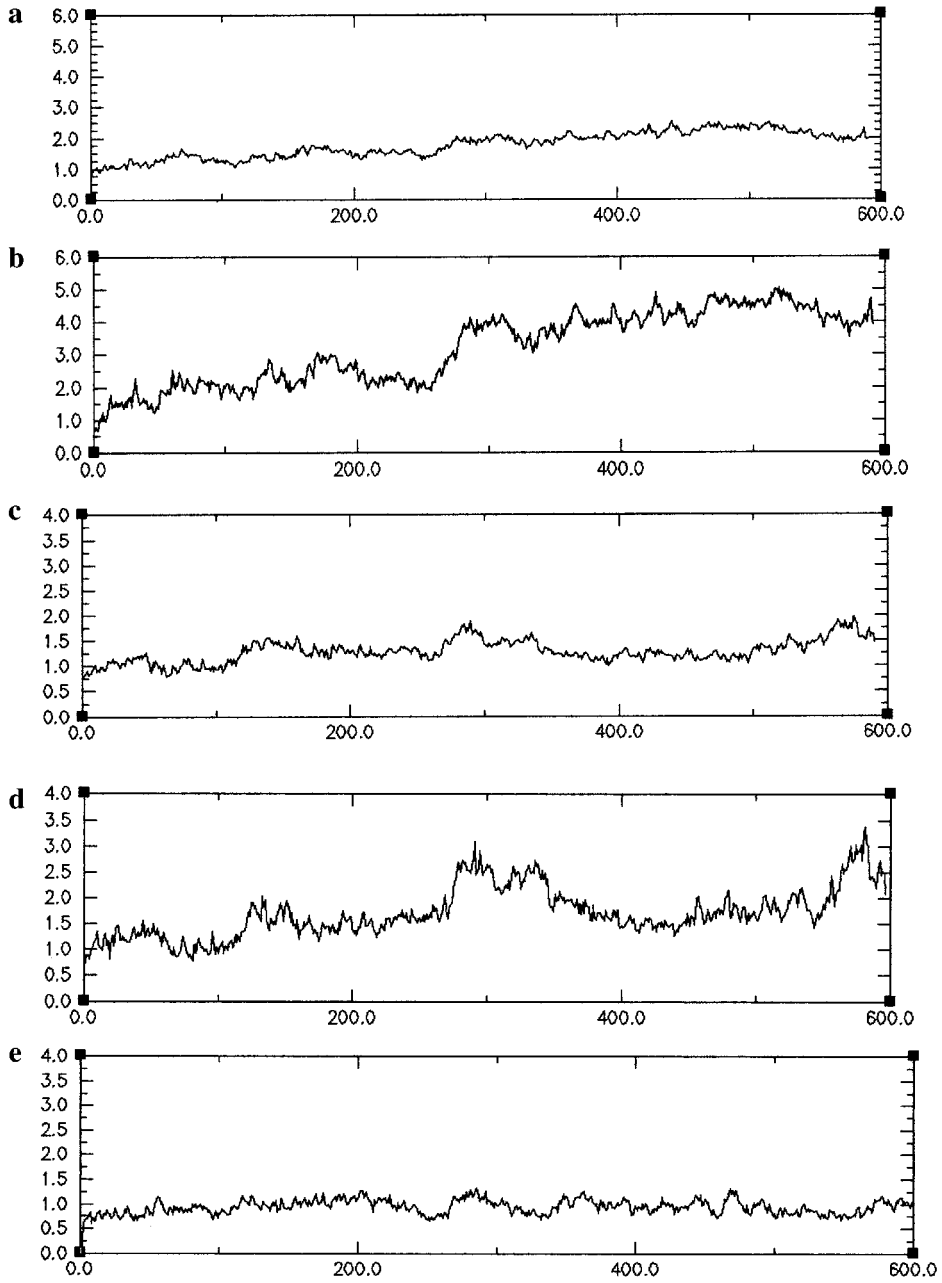

FIGURE 10: Plots of rms deviations from the averaged structures, in Å, during the trajectories. The time-scale is in ps. (a) The external ligand complex, model A; (b) the ligand alone in model A; (c) the external ligand complex, model B; (d) the ligand alone in model B; (e) the native G-quadruplex.

C). The marked distortions in geometry that result from the instability of the intercalation sites (Figures 6 and 7, panel c; Table 1) would also occur with a centrally positioned intercalation site. Any quadruplex structure with such highly distorted bases (we can no longer consider the structure in Figure 6 to contain four G-quartets) would not produce the regular diffraction pattern seen in Figure 2 and can be ruled out. We conclude that the diffraction pattern together with the behavior during the dynamics simulations is consistent only with the externally bound models for the $[d(GGGG)]_4$–BSU1071 complex. This is also in accord with the conclusions of external stacking in the NMR structure determined for the PIPER ligand complexed with a closely similar intermolecular quadruplex (*17*a). External binding is also most consistent with the relatively rapid kinetics observed for this system (*17b*), contrasting with the very slow kinetics of quadruplex formation and dissociation (*2*, *23*).

Table 4: Average Hydrogen Bond Distances for the G-quartets in the G-quadruplex-only Model over the Last 600 ps of Molecular Dynamics Simulation

G-quartet number	ave distance for N7...N2 bond (Å)	deviation from ave	ave distance for O6...N1 bond (in Å)	deviation from ave
first	2.949	0.114	3.107	0.188
second	2.995	0.140	3.296	0.216
third	3.000	0.142	3.299	0.231
fourth	2.953	0.112	3.093	0.179

The disorder found in the crystals is likely to be a direct result of the conformational flexibility exhibited by the disubstituted anthraquinone molecule BSU1071, with the existence of many energetically similar and quasi-equivalent ligand orientations in the externally bound models. Thus the two original starting models A and B readily interconvert,

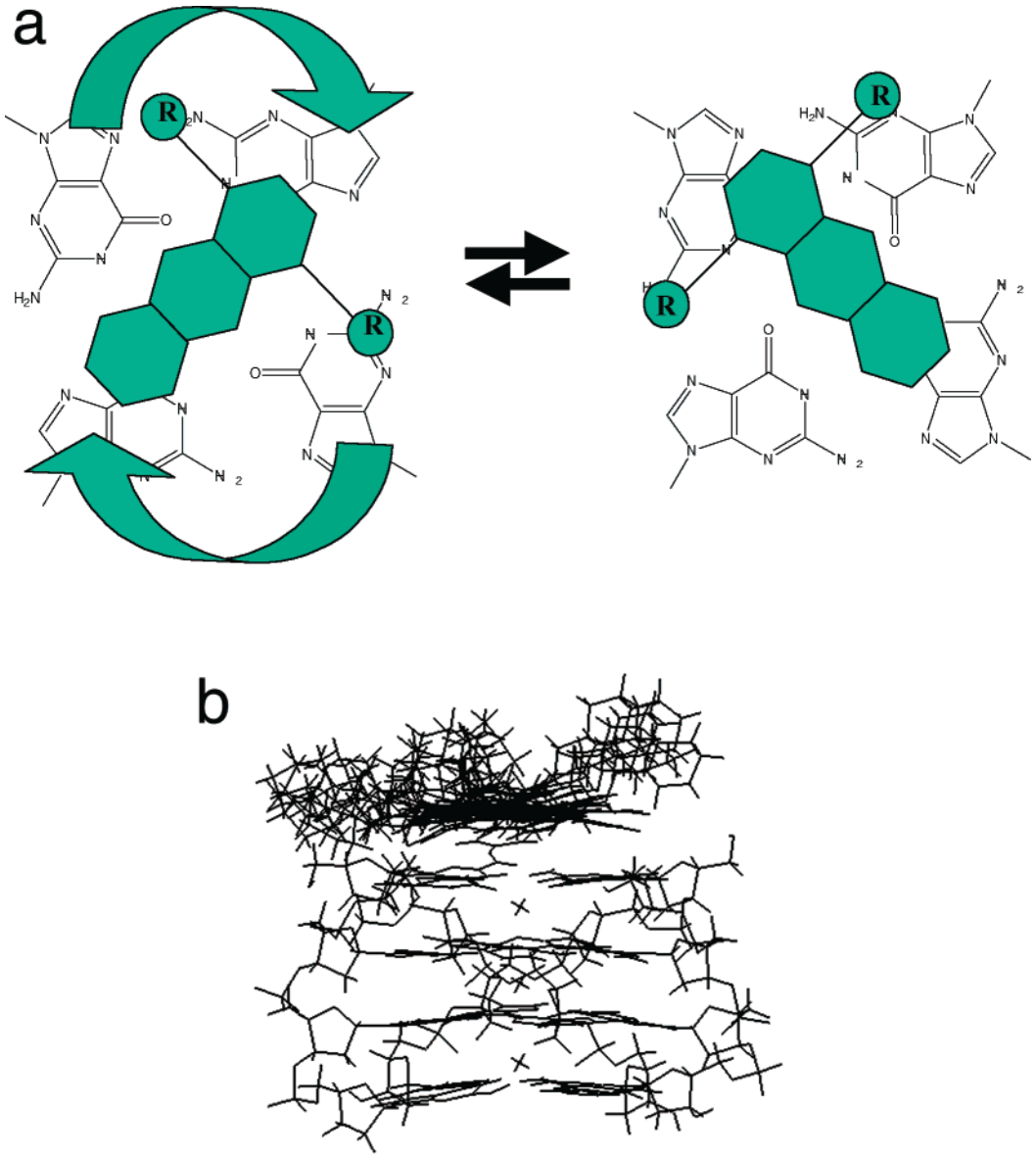

FIGURE 11: (a) Schematic representation of the motion of ligands during the 1000 ps simulations of external complexes models A and B. (b) Plot of the superposition of orientations of BSU1071, taken from snapshots during the 1000 ps simulation of model A.

initiated by a slight rotational change in the orientation of an anthraquinone moiety with respect to the adjacent G-quartet and a subsequent rotation of the torsion angles around methylene linkers in the side chains. The inherent flexibility of the side chains provides a series of low energy barriers between each conformation, which are easily surmounted in the simulations. Thus the structures in the crystal may represent a series of quasi-equivalent conformers and groove populations for the ligands.

In the intercalation model (Figure 2, panel C), the collapse of the G-quartets adjacent to the intercalated anthraquinone groups is because there is only a small amount of guanine−anthraquinone stacking (Figure 6, panel b). This results in a highly strained system, weakening the cyclic array of Hoogsteen hydrogen bonds that hold the quadruplex together. Furthermore, the position of the drug blocks the favorable interactions of any sodium or potassium ions located in the central channel required to counteract the negative potential created by the polarized O6 carbonyl oxygen atoms on the guanine base edges. Inevitably, the distortion escalates with increased simulation times, culminating in the observed disruption of the entire model. Although it is feasible that a longer array of stacked G-quartets maybe able to accommodate such an intercalation site, it is difficult to envisage with the short $[d(GGGG)]_4$ quadruplex.

The alternative externally stacking ligand models have several characteristics that make their structures more energetically stable. The terminal G-quartets lack the complete π-stacking interactions associated with the internal G-quartets as they only interact with one other G-quartet, providing the opportunity to enhance stabilization of the structure by effectively capping and securing the ends of the quadruplex with the ligands. They also represent an easily accessible binding site allowing the ligands to stabilize and interact with the large π-plane, without disruption to the phosphate backbones, and having significantly less kinetic barriers to overcome. In addition, with the molecules sitting externally the crucial internal monovalent cations remain intact in the central channel.

Behavior of the Quadruplexes during the Simulations. The hydrogen bond distances in the external BSU1071 complexes and in the native structure are in accord with their crystal-

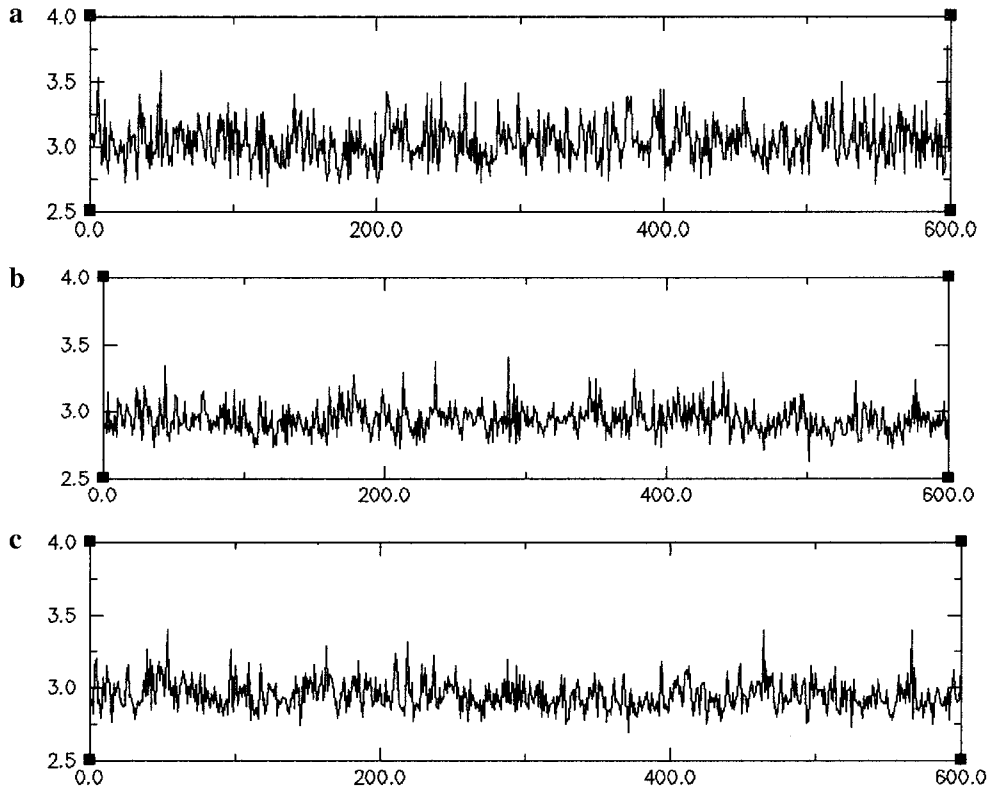

FIGURE 12: Variation in N2...N7 Hoogsteen hydrogen bond distance, in Å, for one of the external G-quartets during the simulations. (a) For model A; (b) for model B; (c) for the native quadruplex.

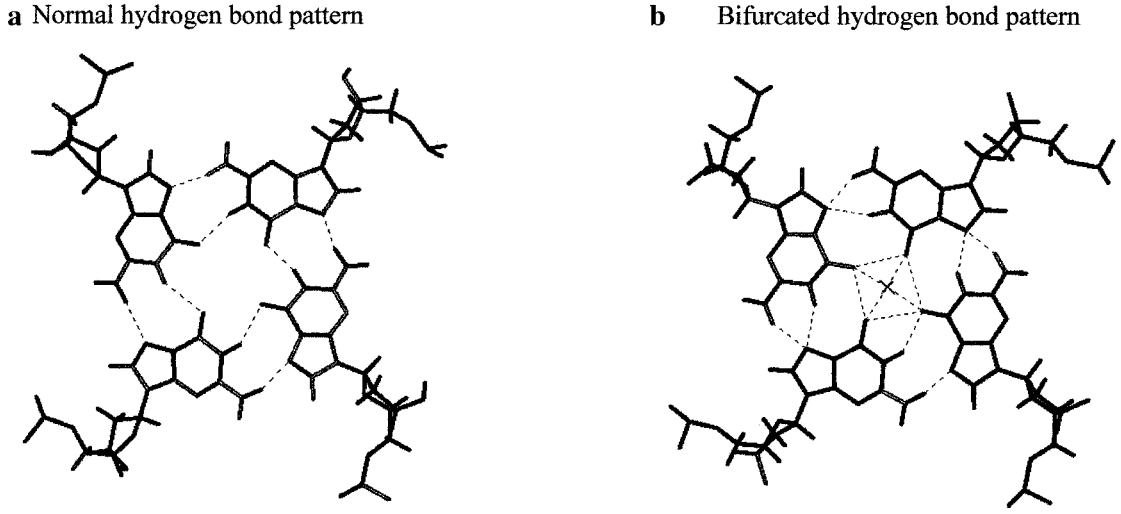

FIGURE 13: Plots of the averaged structure of (a) one of the external G-quartets and (b) one of the internal G-quartets, in the native quadruplex. Models A and B have closely similar arrangements.

lographic values (*4*), although the pair-additive nature of the force field results in the formation of bifurcated hydrogen bonding in the internal G-quartets, independent of the presence of ligands (Figure 13). This feature has not been observed in the crystal structures of quadruplexes. Although it may be attributed, in part, to the dynamic motion of the molecule, it has been suggested as being more likely due to the neglect of polarization effects in the force field, between the central cations and their coordination to the guanine bases (*24*). The effect has been observed in a recent 3.4 ns simulation (*24*) of $[d(TGGGGT)]_4$, as have other features observed in the present simulations.

The position of the BSU1071 ligand in the external models, lying directly over the central cation channel of the quadruplex, signifies that the strongest $\pi-\pi$ stacking interactions exists between the six-membered quinone ring of the ligand and the center of the G-quartet, containing the polarized guanine O6 oxygen atoms. This suggest that the N1...O6 hydrogen bonds on the external G-quartets would be more influenced by the stabilizing effects of the ligand as compared to the relatively peripheral N2...N7 hydrogen bonds. This property is consistent with the trend in distances (Tables 2−4), where the average values for the N1...O6 distances in the external G-quartets for both the externally bound ligand−DNA complex models, are slightly (and consistently) less than those from the corresponding native G-quadruplex simulation. The deviations from the averaged distances are also greater in the native G-quadruplex model,

indicating an extra degree of stability in the N1...O6 hydrogen bonds when adjacent to the ligand.

The strategy for G-quadruplex-mediated inhibition of the telomerase enzyme by small molecules such as BSU1071 requires that formation of a stable G-quadruplex complex occurs rapidly and without the necessity for energetically costly conformational change in the quadruplex. The finding that external stacking is the most plausible arrangement for a ligand complex, at least for the intermolecular quadruplex formed by d(GGGG), is entirely in accord with this requirement and will aid the design of future ligands with enhanced selectivity to quadruplexes.

ACKNOWLEDGMENT

We thank Tony Reszka for synthesis of compound BSU1071 and Struther Arnott and Alex Wood for useful discussions.

REFERENCES

1. Blackburn, E. H. (1991) *Nature 350*, 569−573.
2. Williamson, J. R. (1994) *Annu. Rev. Biophys. Biomol. Struct. 23*, 703−730.
3. Kang, C. H., Zhang, X., Ratliff, R., Moyzis, R., and Rich, A. (1992) *Nature 356*, 126−131.
4. (a) Laughlan, G., Murchie, A. I. H., Norman, D. G., Moore, M. H., Moody, P. C. E., Lilley, D. M. J., and Luisi, B. (1994) *Science 265*, 520−524; (b) Phillips, K., Dauter, Z., Murchie, A. I. H., Lilley, D. M. J., and Luisi, B. (1997) *J. Mol. Biol. 273*, 171−182.
5. Aboul-ela, F., Murchie, A. I. H., Norman, D. G., and Lilley, D. M. J. (1992) *Nature 360*, 280−282.
6. O'Reilly, M., Teichmann, S. A., and Rhodes, D. (1999) *Curr. Opin. Struct. Biol. 9*, 58−65.
7. (a) Kim, N. W., Piatyszek, M. A., Prowse, K. R., Harley, C. B., West, M. D., Ho, P. L. C., Coviello, G. M., Wright, W. E., Weinrich, R., and Shay, J. W. (1994) *Science 266*, 2011−2015; (b) Meyerson, M., Counter, C. M., Eaton, E. N., Ellison, L. W., Steiner, P., Caddle, S. D., Ziaugra, L., Beijersbergen, R. L., Davidoff, M. J., Lie, Q., Bacchetti, S., Haber, D. A., and Weinberg, R. A. (1997) *Cell 90*, 785−795.
8. Hahn, W. C., Counter, C. M., Lundberg, A. S., Beijersbergen, R. L., Brooks, M. W., and Weinberg, R. A. (1999) *Nature 400*, 464−468.
9. Neidle, S., and Kelland, L. R. (1999) *Anti-Cancer Drug Des. 14*, 341−347.
10. (a) Hahn, W. C., Stewart, S. A., Brooks, M. W., York, S. G., Eaton, E., Kurachi, A., Beijersbergen, R. L., Knoll, J. H. M., Meyerson, M., and Weinberg, R. A. (1999) *Nature Med. 5*, 1164−1170; (b) Zhang, X., Mar, V., Zhou, W., Harrington, L., and Robinson, M. O. (1999) *Genes Dev. 13*, 2388−2399.
11. Herbert, B. S., Pitts, A. E., Baker, S. I., Hamilton, S. E., Wright, W. E., Shay, J. W., and Corey, D. R. (1999) *Proc. Natl. Acad. Sci. U.S.A. 96*, 14276−14281.
12. (a) Mergny, J.-L., and Hélène, C. (1998) *Nature Med. 4*, 1366−1367; Mergny, J.-L., Mailliet, P., Lavelle, F., Riou, J.-F., Laoui, A., and Hélène, C. (1999) *Anti-Cancer Drug Des. 14*, 327−339; (b) Han, H., and Hurley, L. H. (2000) *Trends Pharm. Sci. 21*, 136−142.
13. (a) Sun, D. B., Thompson, B., Cathers, B. E., Salazar, M., Kerwin, S. M., Trent, J. O., Jenkins, T. C., Neidle, S., and Hurley, L. H. (1997) *J. Med. Chem. 40*, 2113-2116. (b) Perry, P. J., Gowan, S. M., Reszka, A. P., Polucci, P., Jenkins, T. C., Kelland, L. R., and Neidle, S. (1998) *J. Med. Chem. 41*, 3253−3260. (c) Perry, P. J., Reszka, A. P., Wood, A. A., Read, M. A., Gowan, S. M., Dosanjh, H. S., Jenkins, T. C., Kelland, L. R., and Neidle, S. (1998) *J. Med. Chem. 41*, 4873−4884; (d) Neidle, S., Harrison, R. J., Reszka, A. P., and Read, M. A. (2000) *Pharmacol. Ther. 85*, 133−139.
14. (a) Wheelhouse, R. T., Sun, D., Han, H., Han, F. X., and Hurley, L. H. (1998) *J. Am. Chem. Soc. 120*, 3261−3262. (b) Han, F. X., Wheelhouse, R. T., and Hurley, L. H. (1999) *J. Am. Chem. Soc. 121*, 3561−3570. (c) Izbicka, E., Wheelhouse, R. T., Raymond, E., Davidson, K. L., Lawrence, R. A., Sun, D., Windle, B. E., Hurley, L. H., and von Hoff, D. D. (1999) *Cancer Res. 59*, 639−644.
15. Perry, P. J., Read, M. A., Davies, R. T., Gowan, S. M., Reszka, A. P., Wood, A. A., Kelland, L. R., and Neidle, S. (1999) *J. Med. Chem. 42*, 2679−2684.
16. (a) Harrison, R. J., Gowan, S. M., Kelland L. R., and Neidle, S. (1999) *Bioorg. Med. Chem. Lett. 9*, 2463−2468. (b) Read, M. A., Wood, A. A., Harrison, R. J., Gowan, S. M., Kelland, L. R., Dosanjh, H. S., and Neidle, S. (1999) *J. Med. Chem. 42*, 4538−4546.
17. (a) Federoff, O. Y., Salazar, M., Han, H., Chemeris, V. V., Kerwin, S. M., and Hurley, L. H. (1998) *Biochemistry 37*, 12367-12374. (b) Han, H., Cliff, C. L., and Hurley, L. H. (1999) *Biochemistry 38*, 6981−6986.
18. (a) Arthanari, H., Basu, S., Kawano, T. L., and Bolton, P. H. (1998) *Nucleic Acids Res. 26*, 3724−3728; (b) Anantha, N. V., Azam, M., and Sheardy, R. D. (1998) *Biochemistry 37*, 2709−2714; (c) Arthanari, H., and Bolton, P. H. (1999) *Anti-Cancer Drug Des. 14*, 317−326.
19. Case, D. A., Pearlman, D. A., Caldwell, J. C., Cheatham, T. E., Ross, W. S., Simmerling, C., Darden, T., Merz, K. M., Stanton, R. V., Cheng, A., Vincent, J. J., Cromley, M., Ferguson, D. M., Radmer, R., Seibel, G. L., Singh, U. C., Weiner, P., and Kollman, P. A. (1997) *AMBER5*, University of California, San Francisco.
20. Cornell, W. D., Cieplak, P., Bayly, C. I., Gould, I. R., Merz, K., Ferguson, D. M., Spellmeyer, D. C., Fox, T., Cardwell, J. W., and Kollman, P. A. (1995) *J. Am. Chem. Soc. 117*, 5179−5197.
21. Arnott, S. (1999) in *Oxford Handbook of Nucleic Acid Structure* (Neidle, S., Ed.) Oxford University Press.
22. Haq, I., Trent, J. O., Chowdhry, B., and Jenkins, T. C. (1999) *J. Am. Chem. Soc. 121*, 1768−1779.
23. Wyatt, J. R., Davis, P. W., and Freier, S. M. (1996) *Biochemistry 35*, 8002−8008.
24. Špačková, N., Berger, I., and Šponer, J. (1999) *J. Am. Chem. Soc. 121*, 5519−5534.

BI001584K