JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

Article

pubs.acs.org/JACS

Solution Structure of an Intramolecular (3 + 1) Human Telomeric G-Quadruplex Bound to a Telomestatin Derivative

Wan Jun Chung,[†,∥] Brahim Heddi,[†,∥] Masayuki Tera,[‡,§] Keisuke Iida,[‡] Kazuo Nagasawa,[‡] and Anh Tuân Phan[†,*]

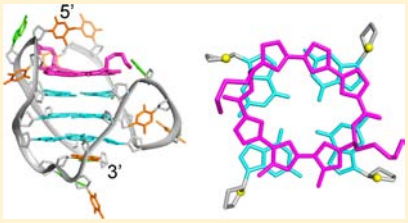
[†]School of Physical and Mathematical Sciences, Nanyang Technological University, Singapore

[‡]Department of Biotechnology and Life Science, Faculty of Technology, Tokyo University of Agriculture and Technology, Japan

[§]Bioorganic Research Institute, Suntory Foundation for Life Science, Japan

Supporting Information

ABSTRACT: Guanine-rich human telomeric DNA can adopt secondary structures known as G-quadruplexes, which can be targeted by small molecules to achieve anticancer effects. So far, the structural information on complexes between human telomeric DNA and ligands is limited to the parallel G-quadruplex conformation, despite the high structural polymorphism of human telomeric G-quadruplexes. No structure has been yet resolved for the complex with telomestatin, one of the most promising G-quadruplex-targeting anticancer drug candidates. Here we present the first high-resolution structure of the complex between an intramolecular (3 + 1) human telomeric G-quadruplex and a telomestatin derivative, the macrocyclic hexaoxazole L2H2-6M(2)OTD. This compound is observed to interact with the G-quadruplex through π-stacking and electrostatic interactions. This structural information provides a platform for the design of topology-specific G-quadruplex-targeting compounds and is valuable for the development of new potent anticancer drugs.



INTRODUCTION

Human telomeric DNA contains numerous tandem repeats of the d(TTAGGG) sequences terminating with a 100−200 nt single-stranded 3′ overhang.[1,2] They are situated at the termini of eukaryotic chromosomes and capped by nucleoprotein complexes.[3] This construct plays an important role in preventing chromosomal degradation and end-to-end fusion.[4,5] In somatic cells, telomeres shorten progressively after each replication, triggering cell senescence and apoptosis when the Hayflick limit is reached.[6,7] Telomerase is an enzyme with reverse transcriptase activity that can elongate the telomeres by adding nucleotides to their 3′ ends.[8] In 80−85% of cancer cells, the up-regulation of telomerase activity interferes with the telomere shortening process, hence, inhibiting cell apoptosis and leading to the development of cancer cells.[9,10]

G-rich human telomeric sequences have the propensity to fold into G-quadruplex structures.[11,12] When G-quadruplex formation at the telomere 3′ overhang is promoted, inhibition of the telomerase activity can be achieved through the disruption of telomerase recognition and interaction capability with the single-stranded DNA 3′ overhang.[13] Moreover, G-quadruplex formation could also lead to the displacement of protein(s) bound at the 3′ overhang,[14,15] inducing cell senescence and apoptosis. Studies have demonstrated that the above inhibitory effect could be promoted through the introduction of small molecules targeting and stabilizing G-quadruplex formation at the telomeric ends.[16,17] Hence, this approach is regarded as a potential anticancer therapeutic strategy.

For the past decades, numerous G-quadruplex-targeting small molecules have been developed. Many of these molecules contain aromatic rings (simple or fused) and cationic side chains, aiming to target the G-tetrad core and anionic phosphate backbone, respectively.[18,19] In order for a small molecule to serve as a potential drug candidate, it has to exhibit high binding affinity and selectivity toward G-quadruplexes. Among the reported G-quadruplex-targeting small molecules, telomestatin (Figure S1A of the Supporting Information, SI), a naturally occurring compound isolated from *Streptomyces anulatus* 3533-SV4,[20] was shown to display highly promising results as an anticancer drug candidate, drawing great interest from researchers.[15,21−25] This compound was shown to be a potent telomerase inhibitor, giving an IC_{50}-TRAP value of 5 nM,[20] and was also shown to target various cancer cell lines with IC_{50} values ranging from 0.5 to 4.0 μM while displaying low cytotoxicity toward normal cell lines.[21,23,26−28] Telomestatin also exhibits selectivity toward G-quadruplex over duplex and single-stranded DNA as demonstrated in competitive FRET experiments.[18,29] However, despite its superior activity and selectivity, there are limitations toward further modification and optimization of telomestatin as it does not possess convenient functional groups. Moreover, the total synthesis of telomestatin is difficult, rendering limited access to this compound especially for an industrial-scale production.[30−32]

Received: June 11, 2013
Published: August 2, 2013

dx.doi.org/10.1021/ja405843r | J. Am. Chem. Soc. 2013, 135, 13495−13501

Telomestatin also has a low water solubility that can hinder its bioavailability, an important factor in drug design.

More recently, a series of macrocyclic molecules (telomestatin analogues) with improved features over telomestatin were developed.[18,33−39] Tera et al.[34] showed that these telomestatin analogues exhibit high selectivity toward G-quadruplex structures and display potent telomerase inhibitory activity.[34] One of the most promising compounds is L2H2-6OTD (also known as HXDL by Rzuczek et al.[33]), a telomestatin derivative containing two alkyl amine side chains and six oxazole rings (Figure S1B of the SI) and possessing a comparable activity as telomestatin (IC_{50}-TRAP = 20 nM, IC_{50} = 7.4 μM).[34] The amine side chains are positively charged at physiological pH,[18] thus improving the solubility of L2H2-6OTD. They are also hypothesized to contribute to the interaction with the phosphate backbone of G-quadruplexes.

Despite the considerable amount of promising data obtained for telomestatin and its analogues, the actual binding modes of these compounds to G-quadruplexes are unknown, with no high-resolution structures reported to date. Currently, the structural knowledge on G-quadruplex-ligand complexes is still limited. The resolved structures of most G-quadruplex-ligand complexes involve the parallel conformation in spite of the broad conformational diversity of G-quadruplexes and a large pool of reported ligands.[40−46] Moreover, only a small number of these resolved structures are intramolecular G-quadruplexes.[40,43,44,47−50] Herein, we present the first NMR solution structure of a complex formed between a (3 + 1) hybrid intramolecular human telomeric G-quadruplex DNA (named *HT*)[51−53] and a telomestatin analogue, L2H2-6M(2)OTD[38] (named **L2H** in this paper) (Figure 1). This compound has two

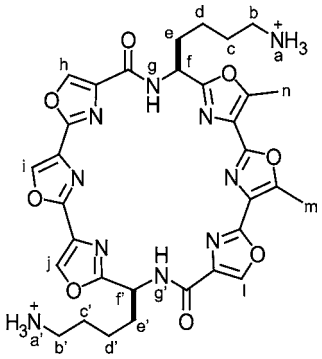


Figure 1. Chemical structure of the macrocyclic hexaoxazole L2H2-6M(2)OTD (**L2H**) with proton naming indicated.

additional methyl groups on two of the six oxazole rings as compared to L2H2-6OTD. Detailed structural information on the complex would be most useful for further modifications and optimizations of these macrocycles into a potential anticancer drug candidate.

RESULTS AND DISCUSSION

Formation of a 1:1 DNA-Drug Complex between *HT* and L2H. The human telomeric *HT* sequence d-(TTGGGTTAGGGTTAGGGTTAGGGA), which forms a well-defined intramolecular (3 + 1) hybrid G-quadruplex,[52] displays twelve well-resolved peaks in the imino proton region of the NMR spectrum (Figure 2). As an increasing amount of **L2H** was titrated into *HT*, we observed the emergence of a new

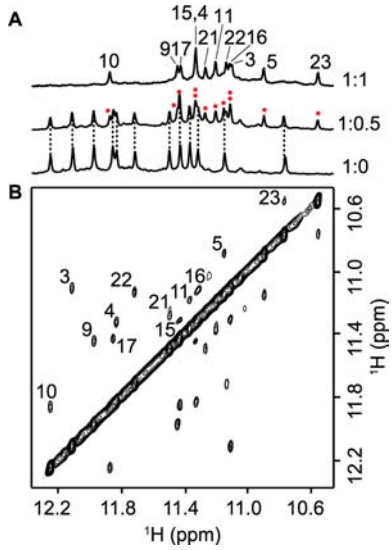


Figure 2. NMR study of the *HT*-**L2H** complex at 37 °C. (A) NMR imino proton spectra showing the titration of *HT* with increasing concentrations of **L2H**. The ratio of DNA to ligand is indicated on the right-hand side of the NMR spectra. Peaks arising from the complex formation at the [ligand]/[DNA] ratio of 0.5 are marked with red asterisk. (B) NOESY spectrum (mixing time, 200 ms) of *HT* in the presence of 50% **L2H**. Exchange cross-peaks between the free and bound *HT* are labeled with residue numbers.

set of twelve well-resolved imino protons for the *HT*-**L2H** complex, while the peaks from the free *HT* gradually decreased after each round of titration. This indicated that the binding of **L2H** to *HT* occurred in a slow exchange regime on the NMR time-scale, allowing both the free and bound states of DNA to be observed. When one equivalent of **L2H** was added to *HT*, only the signals from the *HT*-**L2H** complex could be observed (Figure 2A). As only one set of twelve imino protons was observed, we deduced that only a single dominant conformation of the 1:1 DNA-drug complex was present.

Assignment of Resonances Belonging to the Bound DNA. At the DNA-drug ratio of 1:0.5, exchange cross-peaks between the free and bound states of *HT* could be observed in the 2D NOESY spectrum (Figure 2B). Using these cross-peaks together with the previously reported imino proton assignments of the free *HT*,[52] we unambiguously assigned the 12 imino protons of the bound *HT*. It should be noted that upon ligand binding imino protons from the top G-tetrad underwent the largest chemical-shift changes (Figure 2B and Figure S2 of the SI) and were more protected from solvent exchange (Figure S3 of the SI), consistent with ligand binding at the top G-tetrad layer.

Similarly, some other well-resolved protons of the bound *HT* were unambiguously assigned through exchange cross-peaks with those of the free form. Independently, several aromatic and methyl protons of the bound *HT* were also unambiguously assigned using site-specific deuteration of guanines and adenines[54] (Figures S4 and S5 of the SI) and site-specific low-enrichment (4%) ^{13}C-labeling of thymines[55] (Figures S6 and S7 of the SI). Following the assignment of the aromatic protons, the H8/H6−H1′ NOE sequential connectivity could be traced from T1 to A24 in the NOESY spectrum (Figure S8 of the SI). The remaining protons of the bound *HT* were assigned accordingly using data obtained from through-bond

dx.doi.org/10.1021/ja405843r | J. Am. Chem. Soc. 2013, 135, 13495−13501

(TOCSY, COSY, {^{1}H−^{13}C}-HSQC, and {^{1}H−^{31}P}-HSQC) and through-space (NOESY) correlation experiments.[56]

Assignment of Resonances Belonging to the Bound Ligand. The protons of the ligand **L2H** are named as shown in Figure 1. Aromatic protons (Hh/Hi/Hj/Hl) of **L2H** could be unambiguously distinguished from those of *HT* by comparing {^{1}H−^{13}C}-HSQC spectra of two *HT*-**L2H** complexes, where *HT* was unlabeled (Figure S9A of the SI) or uniformly ^{13}C,^{15}N-labeled (Figure S9B of the SI), respectively. Individual aromatic protons of **L2H** were further identified through intramolecular NOE cross-peaks from the 2D NOESY spectrum (Figure S10B of the SI).

Alkyl side chains of **L2H** were assigned using through-bond correlations in a 2D TOCSY experiment, as they possessed connectivity characteristics in alkyl linkage (Figure S11B of the SI). Furthermore, these resonances were observed in the {^{1}H−^{13}C}-HSQC spectra (Figure S11A of the SI) consistent with the number of carbons on the alkyl side chains. Methyl protons of **L2H** were also observed in {^{1}H−^{13}C}-HSQC (Figure S11A of the SI). Identity of the methyl protons were confirmed through intramolecular NOE cross-peaks from the 2D NOESY spectra.

Alkyl protons (Hf/Hf′) were assigned through the three-bond correlation with amide protons (Hg/Hg′) (Watergate TOCSY, Figure S12 of the SI). Assignments of the individual protons of **L2H** were further affirmed through intramolecular connectivity found in the 2D NOESY spectrum (Figure S10 of the SI).

The Folding Topology of *HT* Remains Unchanged upon Ligand Binding. The folding topology of *HT* in the complex was determined by analyzing NOE cross-peaks between the imino (H1) and aromatic (H8) protons of the guanine bases (Figure 3 and Figure S13A of the SI). These data show that upon binding to **L2H**, *HT* retains its original (3 + 1) hybrid G-quadruplex folding topology, involving three G-tetrad layers, G3·G21·G17·G9, G4·G10·G16·G22, and G5·G11·G15·G23. CD spectra of the *HT*-**L2H** complex (Figure S13C of the SI) display two bands at 260 and 290 nm, characteristic of a (3 + 1) hybrid G-quadruplex.[51−53] However, the ratio between the 260- and 290-nm peaks was observed to increase upon ligand binding.

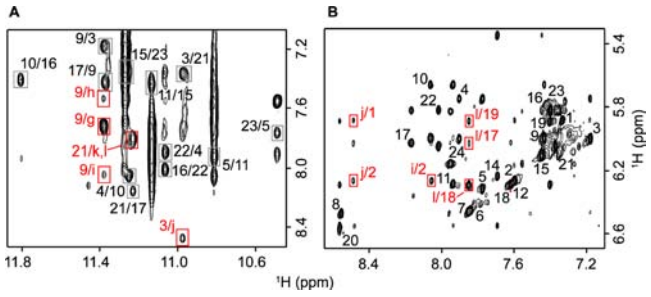


Figure 3. NOESY spectra the *HT*-**L2H** complex. (A) NOESY spectrum (mixing time, 300 ms) showing the NOEs between the imino (H1) protons of *HT* and the aromatic and amide protons of **L2H**. Intermolecular cross-peaks between *HT* and **L2H** are labeled and framed in red with H1 protons of *HT* in the first position and aromatic/amide protons of **L2H** in the second position. Intramolecular NOEs within *HT* are also shown: cross-peaks between guanine imino (H1) protons and aromatic (H8) protons are framed and labeled with the residue number of H1 protons in the first position and that of H8 protons in the second position. (B) NOESY spectrum (mixing time, 500 ms) showing the NOEs between the aromatic protons of **L2H** and the sugar (H1′) protons of *HT*. Intermolecular cross-peaks between *HT* and **L2H** are framed and indicated in red with the aromatic protons of **L2H** in the first position and H1′ protons of *HT* in the second position. Intra-residue H8/H6-H1′ cross-peaks belonging to *HT* are labeled with respective residue numbers.

Solution Structure of the *HT*-L2H Complex. Following the unambiguous assignment of the *HT* and **L2H** resonances, we identified 79 intermolecular NOEs (Figure 3 and Table 1).

Table 1. NMR Restraints and Structure Statistics

A. NMR Restraints		
distance restraints	exchangeable	nonexchangeable
DNA restraints		
intraresidue	2	277
inter-residue	44	151
ligand restraints		
intramolecular	5	21
DNA-ligand restraints		
intermolecular	13	66
other restraints		
hydrogen-bond restraints	51	
torsion angle restraints	12	
repulsive restraints	13	
B. Structure Statistics		
NOE violations		
numbers (>0.2 Å)	0.20 ± 0.63	
deviations from standard geometry		
bond length (Å)	0.005 ± 0.000	
bond angle (°)	0.810 ± 0.017	
impropers (°)	0.933 ± 0.071	
pairwise all heavy atom r.m.s.d. values (Å)		
all heavy atoms	1.990 ± 0.336	
all heavy atoms except T1, T2, T12, T18, T19, A20	1.110 ± 0.159	
G-tetrad core	0.864 ± 0.156	

The solution structure of the *HT*-**L2H** complex was computed using NMR restraints including intermolecular and intramolecular NOEs (Table 1 and Figure S14 of the SI). The lowest-energy refined structures of the complex are presented in Figure 4. The structure of the bottom part of *HT* is almost unchanged from the free to the bound form (Figure S15A of the SI), with the root-mean-square deviation (r.m.s.d.) calculated for residues G4, G5, G10−G16, and G22−A24 being 1.86 Å. Interestingly, **L2H** was observed to replace the Watson−Crick base pair initially formed between T1 and A20 (Figure S15B of the SI), leaving a free hanging 5′ end. The ligand was observed to interact with the top G-tetrad of *HT* through π-stacking (Figure 4C) with the distance between **L2H** and the top tetrad being ∼3.5 Å, similar to the distance between two consecutive G-tetrad layers.[57] The two cationic side chains of **L2H** are directed toward the negatively charged phosphate backbone of *HT* (Figure 4D), while the methyl groups on the oxazoles were positioned above the wide groove between G9 and G17, away from the bases at the 5′ end of *HT* (Figure 4C).

Cationic Side Chains of L2H Interact with the Phosphate Backbone of *HT*. The two side chains of **L2H** are sufficiently close to the phosphate groups of *HT* for significant electrostatic interactions (Figure 4D). To analyze the electrostatic interactions between the flexible cationic side chains of **L2H** and the phosphate groups of *HT*, a 10-ns restrained molecular dynamics simulation of the complex was

dx.doi.org/10.1021/ja405843r | J. Am. Chem. Soc. 2013, 135, 13495−13501

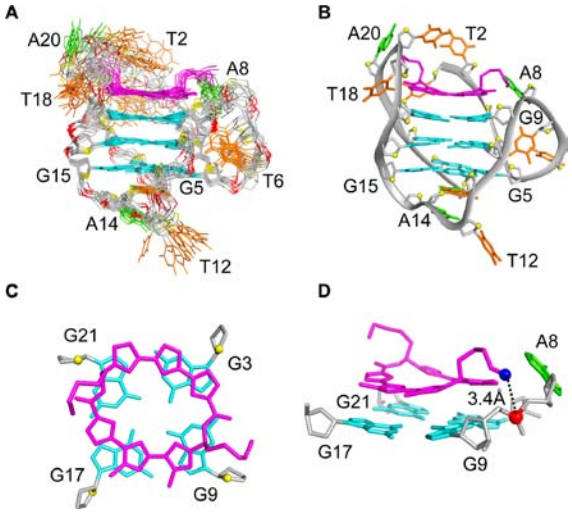


Figure 4. Solution structure of the *HT*-**L2H** complex. (A) Ten superimposed refined structures of the *HT*-**L2H** complex. (B) Ribbon view of a representative structure of the *HT*-**L2H** complex. (C) Top view showing the stacking of **L2H** on the top G-tetrad of *HT*. (D) Side view showing the close proximity between one of the cationic side chain (shown as blue sphere) of **L2H** with the anionic phosphate group (shown as red sphere) of A8. **L2H** is colored in magenta; Guanine residues are colored in cyan; Adenine residues are colored in green, Thymine residues are colored in orange; Backbone and sugars are colored in gray; Phosphorus atoms are colored in red.

conducted in explicit solvent. During the simulation, one of the **L2H**'s cationic amine group spent 90% of the time in close contact (<4.5 Å) with the phosphate group of residue G9, while the other spent 75% and 10% of the time in close contact with the phosphate group of residue A20 and G21, respectively (Figure S16 of the SI). We also observed that the orientation of the side chains changed in accordance with that of the loop bases in order to maintain a close contact with a phosphate group, indicating the role of electrostatic interactions between the cationic side chains of **L2H** and the phosphate backbone of *HT*.

To further investigate the contribution of the cationic amine groups in the interaction, we titrated in *HT* an uncharged compound **L2A**,[38] which is structurally similar to **L2H**, but having the side chains terminating with an amide instead of an amine (Figure S17A,B of the SI). The titration data (Figure S17C,D of the SI) showed that, as compared to **L2H**, a higher concentration of **L2A** was required to achieve the full complex formation. This indicated a lower binding affinity between *HT* and **L2A**, highlighting the importance of cationic groups in the interaction with *HT*.

Ion Coordination between L2H and *HT*. Besides π-stacking interaction, the strong G-quadruplex stabilizing ability of telomestatin and its analogs could also be enhanced by the possible presence of a potassium ion between the macrocycle and the top G-tetrad. In the resolved structure of the *HT*-**L2H** complex, we can observe that there is sufficient space (~3.5 Å) between the top G-tetrad and **L2H** to accommodate a cation or a water molecule, with the macrocycle acting as a pseudo tetrad layer (Figure 5). This ion/molecule can be coordinated to the carbonyl groups of the top G-tetrad as well as the nitrogens on the six oxazole rings of **L2H** (Figure 5).

To investigate on a possible existence of a potassium ion between **L2H** and *HT*, a 10-ns molecular dynamics simulation of the *HT*-**L2H** complex containing a potassium ion manually

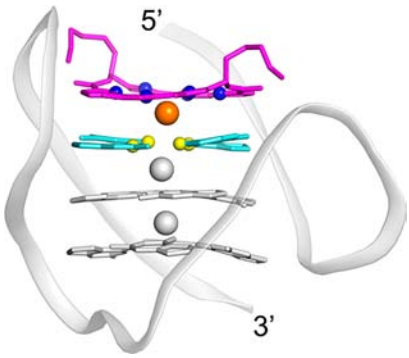


Figure 5. Ribbon representation of the *HT*-**L2H** complex showing a potassium ion between the top G-tetrad and **L2H**. Side view showing the close proximity of the potassium ion to the carbonyl oxygens from the top G-tetrad and the nitrogens from the **L2H** oxazole rings. The top G-tetrad and **L2H** are highlighted: potassium ion is colored in orange, nitrogen atoms are colored in blue, oxygen atoms are colored in yellow, **L2H** is colored in magenta and guanine residues are colored in cyan. The remaining residues of *HT* and potassium ions in the channel of the G-tetrad core are colored in grey.

docked between **L2H** and the top G-tetrad was performed. Throughout the simulation, the potassium ion was observed to remain between **L2H** and the top G-tetrad. This result is consistent with that of mass spectrometry and docking experiments previously reported.[32,58,59]

Interaction of L2H with Different G-quadruplexes. We observed the binding of **L2H** with DNA and RNA G-quadruplexes of different folding topologies (Figures S17–S19 of the SI). From the resolved structure of the *HT*-**L2H** complex, we can deduce that **L2H** interacts with these G-quadruplexes in a similar manner (π-stacking), while the orientation of **L2H** on the G-tetrad depends on the specific features (such as type of loops) of the G-quadruplexes.

For the tetrameric $[d(TAGGGT)]_4$ parallel G-quadruplex, we observed the emergence of a new set of peaks corresponding to a second complex formation when 2 equivalents of **L2H** were added (Figure S19 of the SI). On the basis of our resolved structure of the *HT*-**L2H** complex, we believe that the second ligand molecule interacts with the G-tetrad core through π-stacking via the 3′ bottom G-tetrad. This finding would encourage the G-quadruplex ligand design strategy of Iida et al.[39,60] to improve the binding affinity and selectivity using macrocyclic hexaoxazole dimers. Indeed, the second binding site (likely by stacking at the bottom tetrad) was also observed for the intramolecular G-quadruplex of *HT* when titrated with high concentrations of **L2H** (Figure S17D of the SI). The higher affinity of the ligand to the top tetrad as compared to bottom should reflect the structural difference at the two ends of the G-tetrad core including probably a better capping at the bottom by the T13·A24 base pair.

Planarity of Small Molecules Might Not Be a Crucial Feature for π-Stacking. Most of the small molecules designed to target G-quadruplexes bear a highly planar core, aiming to target the G-tetrad core. However in this instance, we showed that this feature might not be a crucial factor for the design of G-quadruplex-targeting drugs; and if the small molecule possesses sufficient flexibility, it can mold into a specific conformation to maximized interaction with its target. Indeed, despite being highly aromatic, **L2H** was shown by quantum calculations (see Materials and Methods) to be nonplanar and adopt a "roof-like" bent conformation in its free

state (Figure 6A). However upon binding with *HT*, **L2H** flattens out, becoming more planar to maximize the $\pi-\pi$ stacking interaction with the top G-tetrad layer (Figure 6B). The aromatic rings of **L2H** interact with all four guanines on the top G-tetrad. Comparison made with different families of resolved G-quadruplex-ligand complex structures revealed that this interaction differs from that of most other ligands, where single compounds were not observed to interact with all four guanines of the G-tetrad.[40−44,47,48,61−63]

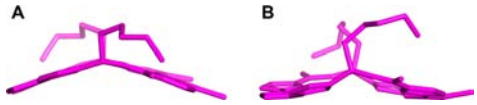


Figure 6. Stick representation of **L2H** in the (A) free and (B) bound state.

Roles of Methyl Groups and Alkyl Side Chains for Enhanced Selectivity toward a Specific G-Quadruplex. While the overlap pattern of **L2H** over a G-tetrad can repeat every 90° rotation between them due to the 4-fold symmetry of the latter, we observed only one specific relative orientation between the ligand and DNA in the *HT*-**L2H** complex. Structural analysis of the complex hinted to us that this specific orientation of the ligand could be adopted to minimize steric clashes between the methyl groups of **L2H** and the bases in the lateral loop, and maximize the electrostatic interaction between the amine groups and the loops. This suggested the role of these groups in the orientation of **L2H** with respect to *HT*.

As one of the methyl groups of **L2H** is positioned directly above the wide groove of *HT* (Figure 7), extending this group to achieve groove interactions might improve the binding affinity and specificity. We can also modify the position and bulkiness of this functional group accordingly to target a specific groove width.

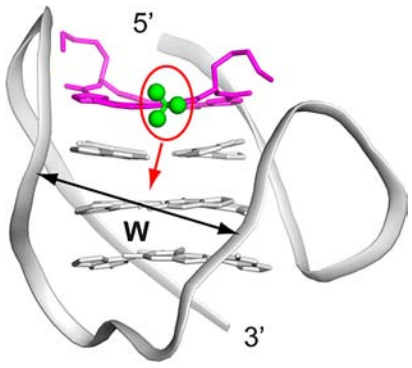


Figure 7. Ribbon representation of the *HT*-**L2H** complex. The protons of a methyl group of **L2H** above the wide groove between G9 and G17 is highlighted as green spheres and framed. The groove, in which the methyl group can be extended into, is highlighted with a red arrow. **HT** is colored in light gray, *L2H* is highlighted in magenta.

On the other hand, the alkyl side chains of **L2H** are relatively flexible with their orientation changes in accordance with that of the DNA loop bases. This suggests that the length and flexibility of the side chains should be taken into consideration for drug optimization. This side chain modification can help to achieve an interaction with a specific type of loop. Furthermore, this side chain modification can be incorporated with the methyl group optimization for a dual targeting effect to enhance drug's selectivity toward a specific G-quadruplex.

CONCLUSIONS

We have shown that the macrocycle L2H2-6M(2)OTD (or **L2H**), a telomestatin derivative, interacts with a human telomeric (3 + 1) hybrid G-quadruplex through π-stacking and electrostatic interactions. From the resolved structure of the complex, we can deduce that planarity is not the utmost important feature in the design of a G-quadruplex-targeting small molecule. More attention should be paid to the flexibility of the molecule. Further enhancement on the binding affinity or selectivity of small molecules toward a specific G-quadruplex scaffold can also be achieved by refining onthe length of their cationic side chain substituents, or by attaching specific linkers to the side chains to target the specific loops/grooves of the desired G-quadruplex. Steric effect can also be incorporated to enhance the selectivity toward a specific G-quadruplex scaffold.

MATERIALS AND METHODS

Sample Preparation. The unlabeled and site-specific labeled DNA oligonucleotides were chemically synthesized on an ABI 394 DNA/RNA synthesizer. Uniformly $^{13}C,^{15}N$ labeling of *HT* was performed by enzymatic synthesis using a protocol adapted from that of Zimmer and Crothers[64] with some modifications. A template hairpin with the sequence d[$TC_3(TAACCC)_3A_3GATCCGA_3GGAT$-CUT], was chemically synthesized with a deoxyribose uracil (dU) preceding the last residue. Enzymatic synthesis was then performed using Klenow fragment (Fermentas) and $^{13}C,^{15}N$-labeled dNTPs (Cambridge Isotopes). dU base was removed by using Uracil DNA Glycosylase (Fermentas) which generated an abasic site. DNA of interest was separated from the template following a heat treatment. The product was then purified using reversed phase HPLC. DNA sample concentrations were determined by measuring the UV absorbance at 260 nm. A 20 mM stock solution of **L2H** was prepared by dissolving the powders in deuterated DMSO (Cambridge Isotope).

Circular Dichroism. Circular dichroism (CD) spectra were recorded at 25 °C using a JASCO-815 spectropolarimeter with a 1-cm path length quartz cuvette containing a solution volume of 500 μL. Spectra (220 to 320 nm) were taken with a scan speed of 200 nm/min. DNA (~5 μM) was dissolved in a buffer containing 70 mM KCl and 20 mM potassium phosphate, pH 7.0. For each measurement, an average of three scans was taken, and the spectral contribution of the buffer was subtracted.

NMR Spectroscopy. Unless otherwise stated, all NMR experiments were performed on 600 and 700 MHz Bruker spectrometers at 25 °C. The concentration of DNA samples was typically 0.1−2 mM. Solution contained 70 mM KCl, 20 mM potassium phosphate (pH 7.0). Spectral analyses were performed using SpinWorks software (http://home.cc.umanitoba.ca/~wolowiec/spinworks/), FELIX (Felix NMR, Inc.), and SPARKY[65] programs.

Quantum Mechanical Computation. Quantum mechanical (QM) computation of the free **L2H** was carried out using the Gaussian 03 software.[66] Geometry optimization at the Hartree−Fock level with a basis valence set 6-31G* was used to determine the structure of **L2H**.

L2H Parametrization. Parameterization of **L2H** was performed using the R.E.D software[67] to determine the partial charges of the individual atoms and using the general AMBER force field (GAFF)[68] to determine the bond, angle and dihedral parameters of the molecule.

NMR-Restrained Structure Calculation. The structure of the *HT*-**L2H** complex was calculated using the XPLOR-NIH[69] and AMBER 9.0 and 10.0[70] programs. NMR-restrained computations were performed as described below. Structures were displayed using PyMOL.[71]

NOE Distance Restraints. Distances between the nonexchangeable protons of the *HT*-**L2H** complex were obtained from NOESY

experiments at various mixing times (100, 300, and 500 ms). The peaks were classified as strong, medium and weak with the distance restraints of 2.9 ± 1.1, 3.75 ± 1.75, and 5.5 ± 1.5 Å, respectively. Resonances from exchangeable protons were classified manually using a NOESY spectrum with a mixing time of 200 ms. The peaks were classified as strong, medium, and weak with distance restraints of 3.0 ± 0.9, 4.0 ± 1.2, and 5.0 ± 1.5 Å, respectively.

Dihedral Restraints. All of the glycosidic bonds of guanines were restrained according to their *anti* and *syn* conformation. The glycosidic χ torsion angle for experimentally determined *syn*-guanine residues were restrained to 60° ± 40°, while that for the *anti*-guanine residues were fixed at 240° ± 40°.

Hydrogen-Bond and Planarity Restraints. Hoogsteen hydrogen bonds between guanines were restrained using H21−N7, N2−N7, H1−O6, and N1−O6 distances, which were set to 1.99 ± 0.1, 2.9 ± 0.1, 1.95 ± 0.1, and 2.95 ± 0.1 Å, respectively. The hydrogen bonds of the T13·A24 base pair was restrained with H3−N1, O2−H61, and O2−N6 distances, which were set to 2.00 ± 0.2, 2.0 ± 0.2, and 3.00 ± 0.2 Å, respectively.

Planarity restraints were used for the G3·G9·G17·G21, G4·G10·G16·G22 and G5·G11·G15·G23 tetrads.

Repulsive Restraints. Repulsive restraints (4.0−20.0 Å) were applied on some pairs of protons that do not exhibit cross-peaks in NOESY.

Distance Geometry Simulated Annealing. An initial extended conformation of *HT* was generated using the XPLOR program and the initial conformation of **L2H** was obtained by quantum mechanical calculation. The extended *HT* and **L2H** were separated by ∼30 Å. The system was then subjected to distance geometry simulated annealing by incorporating the hydrogen-bond, distance, dihedral, planarity and repulsive restraints. Distance restraints involving the top (bases T1, T2, T18, T19, and A20) of *HT* are released during the simulated annealing. 100 structures were generated and subjected to further refinements.

Distance-Restrained Molecular Dynamics Refinement (XPLOR). The 100 structures obtained from the simulated annealing step were refined with a distance-restrained molecular dynamics protocol incorporating all distance restraints. The system was heated from 300 to 1000 K in 5 ps and allowed to equilibrate for 1 ps, during which force constants for the distance restraints were kept at 2 kcal·mol^{-1}·Å^{-2}. The force constants for nonexchangeable proton, exchangeable proton, and repulsive distance restraints were then increased to 64 kcal·mol^{-1}·Å^{-2} in 26 ps. Next, the system was cooled down to 300 K in 14 ps, after which an equilibration was performed for 10 ps. Coordinates of the molecule were saved every 0.5 ps during the last 4.0 ps and averaged. The average structure obtained was then subjected to minimization until the gradient of energy was less than 0.1 kcal·mol^{-1}. Dihedral (50 kcal·mol^{-1}·rad^{-2}) and planarity (1 kcal·mol^{-1}·Å^{-2} for tetrads) restraints were maintained throughout the course of refinement. The ten lowest-energy structures were further refined in AMBER.

Distance-Restrained Molecular Dynamics Refinement in Explicit Solvent (AMBER). The system was neutralized by K^+ cations (including two in the G-tetrad core and one between **L2H** and the top G-tetrad) and solvated with water molecules (TIP3P)[72] in a truncated octahedral box. Hydrogen-bond restraints, interproton distance restraints and repulsive restraints were imposed during the molecular dynamics refinement in the AMBER program. The system was first minimized with harmonic potential position restraints (25 kcal·mol^{-1}·Å^{-2}) on *HT* and **L2H** and the three potassium ions present in the G-tetrad core over 500 steps of steepest descent minimization followed by 500 steps of conjugated gradient minimization. The system was then heated from 100 to 300 K over 10 ps under a constant volume while maintaining 25 kcal·mol^{-1}·Å^{-2} position restraints on aforementioned atoms. The system underwent further steps of minimization and equilibration in which the positional restraints were gradually reduced to 5, 4, 3, 2, 1, and 0.5 kcal·mol^{-1}·Å^{-2}, after which the system was equilibrated without positional restraints for 1 ns at 300 K and 1 atm (in one case, the system was set to run for another 9 ns to investigate on the occupancy and dynamics of the K^+ ion between the DNA and ligand). The system was further minimized in vacuum. Ten lowest-energy structures were chosen.

Data Deposition. The coordinates for the *HT*-**L2H** complex have been deposited in the Protein Data Bank (accession code 2MB3).

■ ASSOCIATED CONTENT

Ⓢ Supporting Information

Additional tables and figures cited in the text. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

phantuan@ntu.edu.sg

Author Contributions

‖These authors contributed equally to this work.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This research was supported by Singapore Ministry of Education Grant ARC33/12 to A.T.P., and a Grant-in-Aid for Scientific Research (B) from JSPS (No. 23310158) to K.N. K.N. is also grateful for financial support from the Mukai Science and Technology Foundation, Tokyo, Japan, and Mochida Memorial Foundation for Medical and Pharmaceutical Research, Tokyo, Japan. M.T. and K.I. are grateful for financial support in the form of JSPS Predoctoral Fellowships for Young Scientists.

■ REFERENCES

(1) Moyzis, R. K.; Buckingham, J. M.; Cram, L. S.; Dani, M.; Deaven, L. L.; Jones, M. D.; Meyne, J.; Ratliff, R. L.; Wu, J. R. *Proc. Natl. Acad. Sci. U.S.A.* **1988**, *85*, 6622−6626.

(2) Wright, W. E.; Tesmer, V. M.; Huffman, K. E.; Levene, S. D.; Shay, J. W. *Genes Dev.* **1997**, *11*, 2801−2809.

(3) de Lange, T. *Genes Dev.* **2005**, *19*, 2100−2110.

(4) de Lange, T. *Oncogene* **2002**, *21*, 532−540.

(5) Lei, M.; Podell, E. R.; Cech, T. R. *Nat. Struct. Mol. Biol.* **2004**, *11*, 1223−1229.

(6) Harley, C. B.; Futcher, A. B.; Greider, C. W. *Nature* **1990**, *345*, 458−460.

(7) Zakian, V. A. *Science* **1995**, *270*, 1601−1607.

(8) Greider, C. W.; Blackburn, E. H. *Cell* **1985**, *43*, 405−413.

(9) Stewart, S. A.; Weinberg, R. A. *Annu. Rev. Cell Dev. Biol.* **2006**, *22*, 531−557.

(10) Kim, N. W.; Piatyszek, M. A.; Prowse, K. R.; Harley, C. B.; West, M. D.; Ho, P. L.; Coviello, G. M.; Wright, W. E.; Weinrich, S. L.; Shay, J. W. *Science* **1994**, *266*, 2011−2015.

(11) Patel, D. J.; Phan, A. T.; Kuryavyi, V. *Nucleic Acids Res.* **2007**, *35*, 7429−7455.

(12) Phan, A. T. *FEBS J.* **2010**, *277*, 1107−1117.

(13) Zahler, A. M.; Williamson, J. R.; Cech, T. R.; Prescott, D. M. *Nature* **1991**, *350*, 718−720.

(14) Brassart, B.; Gomez, D.; De Cian, A.; Paterski, R.; Montagnac, A.; Qui, K. H.; Temime-Smaali, N.; Trentesaux, C.; Mergny, J. L.; Gueritte, F.; Riou, J. F. *Mol. Pharmacol.* **2007**, *72*, 631−640.

(15) Gomez, D.; Wenner, T.; Brassart, B.; Douarre, C.; O'Donohue, M. F.; El Khoury, V.; Shin-Ya, K.; Morjani, H.; Trentesaux, C.; Riou, J. F. *J. Biol. Chem.* **2006**, *281*, 38721−38729.

(16) Balasubramanian, S.; Neidle, S. *Curr. Opin. Chem. Biol.* **2009**, *13*, 345−353.

(17) Neidle, S. *FEBS J.* **2010**, *277*, 1118−1125.

(18) Monchaud, D.; Granzhan, A.; Saettel, N.; Guedin, A.; Mergny, J. L.; Teulade-Fichou, M. P. *J. Nucleic Acids* 2010.

(19) Monchaud, D.; Teulade-Fichou, M. P. *Org. Biomol. Chem.* **2008**, *6*, 627−636.

(20) Shin-ya, K.; Wierzba, K.; Matsuo, K.; Ohtani, T.; Yamada, Y.; Furihata, K.; Hayakawa, Y.; Seto, H. *J. Am. Chem. Soc.* **2001**, *123*, 1262−1263.
(21) Binz, N.; Shalaby, T.; Rivera, P.; Shin-ya, K.; Grotzer, M. A. *Eur. J. Cancer* **2005**, *41*, 2873−2881.
(22) Miyazaki, T.; Pan, Y.; Joshi, K.; Purohit, D.; Hu, B.; Demir, H.; Mazumder, S.; Okabe, S.; Yamori, T.; Viapiano, M.; Shin-ya, K.; Seimiya, H.; Nakano, I. *Clin. Cancer Res.* **2012**, *18*, 1268−1280.
(23) Shammas, M. A.; Shmookler Reis, R. J.; Li, C.; Koley, H.; Hurley, L. H.; Anderson, K. C.; Munshi, N. C. *Clin. Cancer Res.* **2004**, *10*, 770−776.
(24) Temime-Smaali, N.; Guittat, L.; Sidibe, A.; Shin-ya, K.; Trentesaux, C.; Riou, J. F. *PloS one* **2009**, *4*, e6919.
(25) Kim, M. Y.; Vankayalapati, H.; Shin-Ya, K.; Wierzba, K.; Hurley, L. H. *J. Am. Chem. Soc.* **2002**, *124*, 2098−2099.
(26) Liu, W.; Sun, D.; Hurley, L. H. *Nucleos. Nucleot. Nucl. Acids* **2005**, *24*, 1801−1815.
(27) Tahara, H.; Shin-Ya, K.; Seimiya, H.; Yamada, H.; Tsuruo, T.; Ide, T. *Oncogene* **2006**, *25*, 1955−1966.
(28) Tauchi, T.; Shin-Ya, K.; Sashida, G.; Sumi, M.; Nakajima, A.; Shimamoto, T.; Ohyashiki, J. H.; Ohyashiki, K. *Oncogene* **2003**, *22*, 5338−5347.
(29) De Cian, A.; Guittat, L.; Kaiser, M.; Sacca, B.; Amrane, S.; Bourdoncle, A.; Alberti, P.; Teulade-Fichou, M. P.; Lacroix, L.; Mergny, J. L. *Methods* **2007**, *42*, 183−195.
(30) Doi, T.; Shibata, K.; Yoshida, M.; Takagi, M.; Tera, M.; Nagasawa, K.; Shin-ya, K.; Takahashi, T. *Org. Biomol. Chem.* **2011**, *9*, 387−393.
(31) Doi, T.; Yoshida, M.; Shin-ya, K.; Takahashi, T. *Org. Lett.* **2006**, *8*, 4165−4167.
(32) Linder, J.; Garner, T. P.; Williams, H. E.; Searle, M. S.; Moody, C. J. *J. Am. Chem. Soc.* **2011**, *133*, 1044−1051.
(33) Rzuczek, S. G.; Pilch, D. S.; LaVoie, E. J.; Rice, J. E. *Bioorg. Med. Chem. Lett.* **2008**, *18*, 913−917.
(34) Tera, M.; Ishizuka, H.; Takagi, M.; Suganuma, M.; Shin-ya, K.; Nagasawa, K. *Angew. Chem., Int. Ed.* **2008**, *47*, 5557−5560.
(35) Pilch, D. S.; Barbieri, C. M.; Rzuczek, S. G.; Lavoie, E. J.; Rice, J. E. *Biochimie* **2008**, *90*, 1233−1249.
(36) Satyanarayana, M.; Rzuczek, S. G.; Lavoie, E. J.; Pilch, D. S.; Liu, A.; Liu, L. F.; Rice, J. E. *Bioorg. Med. Chem. Lett.* **2008**, *18*, 3802−3804.
(37) Minhas, G. S.; Pilch, D. S.; Kerrigan, J. E.; LaVoie, E. J.; Rice, J. E. *Bioorg. Med. Chem. Lett.* **2006**, *16*, 3891−3895.
(38) Majima, S.; Tera, M.; Iida, K.; Shin-ya, K.; Nagasawa, K. *Heterocycles* **2011**, *82*, 1345−1357.
(39) Iida, K.; Tera, M.; Hirokawa, T.; Shin-Ya, K.; Nagasawa, K. *J. Nucleic Acids* 2010.
(40) Dai, J.; Carver, M.; Hurley, L. H.; Yang, D. *J. Am. Chem. Soc.* **2011**, *133*, 17673−17680.
(41) Gavathiotis, E.; Heald, R. A.; Stevens, M. F.; Searle, M. S. *J. Mol. Biol.* **2003**, *334*, 25−36.
(42) Hounsou, C.; Guittat, L.; Monchaud, D.; Jourdan, M.; Saettel, N.; Mergny, J. L.; Teulade-Fichou, M. P. *ChemMedChem* **2007**, *2*, 655−666.
(43) Nicoludis, J. M.; Miller, S. T.; Jeffrey, P. D.; Barrett, S. P.; Rablen, P. R.; Lawton, T. J.; Yatsunyk, L. A. *J. Am. Chem. Soc.* **2012**, *134*, 20446−20456.
(44) Phan, A. T.; Kuryavyi, V.; Gaw, H. Y.; Patel, D. J. *Nat. Chem. Biol.* **2005**, *1*, 167−173.
(45) Cosconati, S.; Marinelli, L.; Trotta, R.; Virno, A.; De Tito, S.; Romagnoli, R.; Pagano, B.; Limongelli, V.; Giancola, C.; Baraldi, P. G.; Mayol, L.; Novellino, E.; Randazzo, A. *J. Am. Chem. Soc.* **2010**, *132*, 6425−6433.
(46) Campbell, N.; Collie, G. W.; Neidle, S. *Current Protocols in Nucleic Acid Chemistry*; Wiley: New York, 2012; Chapter 17, Unit 17.16.
(47) Bazzicalupi, C.; Ferraroni, M.; Bilia, A. R.; Scheggi, F.; Gratteri, P. *Nucleic Acids Res.* **2013**, *41*, 632−638.
(48) Collie, G. W.; Promontorio, R.; Hampel, S. M.; Micco, M.; Neidle, S.; Parkinson, G. N. *J. Am. Chem. Soc.* **2012**, *134*, 2723−2731.
(49) Micco, M.; Collie, G. W.; Dale, A. G.; Ohnmacht, S. A.; Pazitna, I.; Gunaratnam, M.; Reszka, A. P.; Neidle, S. *J. Med. Chem.* **2013**, *56*, 2959−2974.
(50) Parkinson, G. N.; Cuenca, F.; Neidle, S. *J. Mol. Biol.* **2008**, *381*, 1145−1156.
(51) Xu, Y.; Noguchi, Y.; Sugiyama, H. *Bioorg. Med. Chem.* **2006**, *14*, 5584−5591.
(52) Luu, K. N.; Phan, A. T.; Kuryavyi, V.; Lacroix, L.; Patel, D. J. *J. Am. Chem. Soc.* **2006**, *128*, 9963−9970.
(53) Ambrus, A.; Chen, D.; Dai, J.; Bialis, T.; Jones, R. A.; Yang, D. *Nucleic Acids Res.* **2006**, *34*, 2723−2735.
(54) Huang, X.; Yu, P.; LeProust, E.; Gao, X. *Nucleic Acids Res.* **1997**, *25*, 4758−4763.
(55) Phan, A. T.; Patel, D. J. *J. Am. Chem. Soc.* **2002**, *124*, 1160−1161.
(56) Phan, A. T.; Guéron, M.; Leroy, J. L. *Methods Enzymol.* **2001**, *338*, 341−371.
(57) Lech, C. J.; Heddi, B.; Phan, A. T. *Nucleic Acids Res.* **2013**, *41*, 2034−2046.
(58) Rosu, F.; Gabelica, V.; Smargiasso, N.; Mazzucchelli, G.; Shin-Ya, K.; De Pauw, E. *J. Nucleic Acids*, 2010.
(59) Agrawal, S.; Ojha, R. P.; Maiti, S. *J. Phys. Chem. B* **2008**, *112*, 6828−6836.
(60) Iida, K.; Tera, M.; Hirokawa, T.; Shin-ya, K.; Nagasawa, K. *Chem. Commun.* **2009**, 6481−6483.
(61) Campbell, N. H.; Karim, N. H.; Parkinson, G. N.; Gunaratnam, M.; Petrucci, V.; Todd, A. K.; Vilar, R.; Neidle, S. *J. Med. Chem.* **2012**, *55*, 209−222.
(62) Campbell, N. H.; Parkinson, G. N.; Reszka, A. P.; Neidle, S. *J. Am. Chem. Soc.* **2008**, *130*, 6722−6724.
(63) Clark, G. R.; Pytel, P. D.; Squire, C. J. *Nucleic Acids Res.* **2012**, *40*, 5731−5738.
(64) Zimmer, D. P.; Crothers, D. M. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 3091−3095.
(65) Goddard, T. D.; Kneller, D. G. *SPARKY 3*; University of California: San Francisco.
(66) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, J. A.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Laham, A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; Pople, J. A.; *Gaussian 03*, Revision B.05; Gaussian, Inc.: Pittsburgh, PA, 2003.
(67) Dupradeau, F. Y.; Pigache, A.; Zaffran, T.; Savineau, C.; Lelong, R.; Grivel, N.; Lelong, D.; Rosanski, W.; Cieplak, P. *Phys. Chem. Chem. Phys.* **2010**, *12*, 7821−7839.
(68) Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. *J. Comput. Chem.* **2004**, *25*, 1157−1174.
(69) Schwieters, C. D.; Kuszewski, J. J.; Tjandra, N.; Clore, G. M. *J. Magn. Reson.* **2003**, *160*, 65−73.
(70) Case, D. A.; Cheatham, T. E., 3rd; Darden, T.; Gohlke, H.; Luo, R.; Merz, K. M., Jr.; Onufriev, A.; Simmerling, C.; Wang, B.; Woods, R. J. *J. Comput. Chem.* **2005**, *26*, 1668−1688.
(71) *The PyMOL Molecular Graphics System*, Version 1.5.0.4; Schrödinger, LLC.
(72) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926.