PCCP

Accepted Manuscript

ROYAL SOCIETY OF CHEMISTRY

www.rsc.org/pccp

Cite this: DOI: 10.1039/c0xx00000x

www.rsc.org/xxxxxx

ARTICLE TYPE

Molecular Mechanism of Ligand Unbinding from Human Telomeric G-quadruplex by Steered Molecular Dynamics and Umbrella Sampling Simulations

Jia-Kai Zhou[a], Dah-Yen Yang[b,†,*] and Sheh-Yi Sheu[a,c,†,*]

Received (in XXX, XXX) Xth XXXXXXXXX 20XX, Accepted Xth XXXXXXXXX 20XX
DOI: 10.1039/b000000x

G-quadruplexes are attractive drug targets in cancer therapy. Understanding the mechanisms of the binding/unbinding processes involving biomolecules and molecular recognition is essential in designing new drugs of the G-quadruplexes. We performed steered molecular dynamics and umbrella sampling simulations to investigate the molecular mechanism and kinetics of ligands unbinding processes of the basket, propeller and hybrid G-quadruplex structures. Our studies of the ligand charge effect showed that Coulomb interaction plays a significant role in stabilizing the G-quadruplex structure in the unbinding process. The free energy profiles were carried out and the free energy changes associated with the unbinding process were computed quantitatively, whereas these results could help to identify accessible binding sites and transients interactions. The dynamics of the hydration shell water molecules around the G-quadruplex exhibits an abnormal Brownian motion, and the thickness and free energy of the hydration shell were estimated. A two-step relaxation scheme was theoretically developed to describe the kinetic reaction of BMVC and G-quadruplex interactions. Our computed results fall in a reasonable range of experimental data. The present investigation could be helpful in the structure-based drug design.

Cite this: DOI: 10.1039/c0xx00000x

www.rsc.org/xxxxxx

ARTICLE TYPE

A Introduction

The genomic regions, such as telomeres and oncogene promoters, that have tandem repeats of the guanine (G)-rich sequences can form G-quadruplex structures[1-5] that are attractive drug targets in cancer therapy.[6,7] Human telomere is consisted of the repetitive sequence d(TTAGGG).[8] Telomeres regulate many relevant biological function in cell replication processes.[9] In somatic cells, telomeres shorten after each round of cell division in the absence of DNA polymerase. Once the telomeres reach a critically short length, the cells enter a senescent state and die.[10,11] In cancer cells, telomere length is maintained by the action of the enzyme telomerase, which catalyzes the elongation of telomere repeats at the ends of chromosomes. Telomerase is activated in approximately 85% of tumor cells, and it is not expressed in most somatic cells.[9,12] This finding has been extended to a wide range of human cancers, suggesting that inhibition of human telomerase may arrest cell proliferation, leading to cell death in tumors.[13,14] It has become a strategy to design ligands that can stabilize G-quadruplex structures for anticancer therapy.[15-25]

G-quadruplex contain a structural motif, the G-quartet, which is formed by four G bases that are held in a plane by Hoogsteen hydrogen bonds.[26] Many compounds with planar aromatic chromophores and cationic groups have been used to study the binding affinity and activity of the G-quadruplex.[6,16,27-29] The binding modes of ligand to the G-quadruplex depend on variants such as the topology of G-quadruplexes, the nature of the ligands, the ionic concentrations, and the sample preparation. Known structures of the ligand-bound G-quadruplex complexes have shown that the ligands are bound by end-stacking G-quartets[30-34] or by groove binding.[35] There have been an increasing number of thermodynamic and kinetic experiments that were performed to unravel the stability and specificity of small ligands and the G-quadruplexes.[36-42] In addition, a single-molecule-based method was applied to study the G-quadruplex/pyridostatin interactions.[43] Recently, metadynamics simulations have been used to study the mechanistic insights of ligand binding to the G-quadruplex.[44,45]

In structure-based drug design, the ability of a ligand to stabilize the G-quadruplex depends on not

only the binding affinity of the ligand but also the dynamic process. Thus, the proposition of new drugs for the G-quadruplexes is essential to understanding the specific interactions between ligand and receptor and to underlying the mechanisms of the binding/unbinding processes involving biomolecules and molecular recognition. Nevertheless, the molecular mechanisms behind the binding of drugs to the G-quadruplexes in atomic scale are less understood.[44,45] Due to that the binding/unbinding interactions between a receptor and a ligand have high degrees of freedom including the reaction coordinate, the molecular recognition and the solvated water molecule, which contributes to the hydration shell free energy,[46] it is a challenging task to unravel this problem. The surface water structure, namely the hydration shell, of the biomolecules plays a crucial role in the stabilization of biomolecular structure[47,48] and involves in whether a ligand can pass through the interface region to reach a binding site or reaction center.[49] The hydration shell free energy is also important thermodynamic property that might be estimated by experiments in biological systems.[50] G-quadruplex structures are heterogeneous, and many aspects of the water molecules surrounding the G-quadruplex remain unclear. Thus, a deeper understanding of the dynamics of the ligand penetrating through the surface hydration shell is necessary in the ligand binding/unbinding process.

The ligand 3,6-bis-(1-methyl-4-vinylpyridinium) carbazole diiodide (BMVC) (Fig. 1A) was applied to the human telomeric DNA and the results showed that BMVC could stabilize the G-quadruplex structure by increasing T_m of 13 °C, have a high binding affinity of ~10^9 M^{-1} *in vitro* and the IC_{50} value of ~ 0.05 μM for telomerase.[51-53] The experiments suggested that the binding of BMVC to the G-quadruplex follows a complex process with at least two binding modes.[52,54] Assays of the binding strength of BMVC to the G-quadruplex using single-molecule tethered particle motion method suggested that BMVC is a better end-stacking ligand.[55] In addition, the molecular docking and the binding affinity calculations revealed that BMVC prefers to be at the end of the G-quartet.[56] Based on the results, BMVC is a candidate in the present study for further studying the molecular mechanism and kinetics of the unbinding process in the various G-quadruplex structures.

Even though there have been extensive studies for the association/dissociation of the ligands and proteins by molecular dynamics (MD) and steered molecular dynamics (SMD) simulations,[58-68] the mechanism of ligand intercalation in the G-quadruplex, a critical step in the formation of the ligand/G-quadruplex complex, is still elusive. In the present study, the following questions are addressed. How does the ligand target with the G-quadruplex and stack with the G-quartet? What are the mechanical

properties and the relevant free energy in the ligand binding/unbinding processes? Because it is difficult to produce unfavorable states using the conventional MD simulation, we herein performed computer-based simulations by using SMD[58-69] and umbrella sampling[70,71] to investigate these problems. The SMD is able to pull a molecule out of its binding pocket by applying a guiding potential along a reaction coordinate and is used to generate the initial coordinate for the umbrella sampling computations.[69,72,73] Upon the umbrella sampling to enhance extensive sampling along the reaction coordinate, the potential of mean force (PMF)[67,74] can be constructed.

Because there are multiple pathways of ligands unbinding from the G-quadruplex, the most-stable ligand-bound state and unbinding pathway have to be identified in the reaction coordinate. We here classified two appropriate unbinding pathway categories: direct pathway, in which the ligand escapes from the end-stacked binding site to the bulk state, and indirect pathway, which passes by surface sliding. Charged and neutral carbazole derivatives were used to study the charge effect on the unbinding mechanism. An analysis of the ligand charge effect revealed that Coulomb interaction exhibits a remarkable influence on the structural distortion of the G-quadruplex during the unbinding process. The Coulomb interaction shows significant contribution in the stabilization of the G-quadruplex/ligand complex. Accordingly, we proposed, in energetics, a two-step reaction scheme to describe the kinetic process of ligand binding to a G-quadruplex. First of all, a ligand arrives at an accessible groove region, followed by shuffling on the groove surface, when the ligand is able to adjust its conformation in a manner that leads to a favourable groove surface state. Then, once the thermal motion overcomes the surface free energy, the ligand is therefore able to slide through beneath the loop region and stack onto the G-quartet, entering the bound state. In this scenario, there exists a free energy barrier between the bound state and the groove surface state. In addition, we explored the dynamics and thermodynamic properties of the hydration waters nearby the G-quadruplex surface. Notably, our computed results fall in a reasonable range compared with seminal experimental data. The simulations can provide detailed insights into the mechanical properties and structural changes in the G-quadruplex/BMVC complex. The results could be useful to complement experimental data obtained via atomic force microscopy (AFM).[75,76]

B Computational Methods

All simulations were performed using the CHARMM program[77] and the CHARMM27 force field.[78,79] The optimized structure and partial charges of the ligands (BMVC and 3,6-bis(1-methyl-4-vinyltoluene) carbazole diiodide (BMVC0)) (Fig. 1A) were obtained using the Gaussian09 package with Hartree-Fock theory and a 6-31G* basis set.[80] BMVC0 is an analogue of BMVC in which the N atom in the pyridinium group is replaced with a C atom. Force constants for variant parameters of ligands were taken from similar atoms in the CHARMM27 force field. The ligand structural properties and coordinates are summarized in the Supporting Information. The structure of the human telomeric DNA (PDB 143D)[57] contains 22 nucleotides with the sequence d(AGGGTTAGGGTTAGGGTTAGGG), which has a basket structure consisting of two lateral loops and one diagonal loop (see Fig. 1B). Three Na^+ ions were placed in the center of the G-quartet planes. A docking method was used to identify the ligand binding sites; more detailed procedures have been described previously.[56] Each G-quadruplex/ligand complex structure was subjected to energy minimization using the steep descent (SD) and adopted basis Newton-Raphson (ABNR) algorithms with 100 kcal/(mol $Å^2$) harmonic constraints for the heavy atoms of the G-quadruplex and the ligand, and then was solvated with TIP3P[81] water molecules in a 72.0×72.0×72.0 $Å^3$ cubic box. Sixteen Na^+ ions were added to neutralize the system. The complex was subjected to heating to a temperature of 300 K, followed by 100 ps of equilibration. Subsequently, an additional 1 ns of MD simulation was performed without the constraints. The particle mesh Ewald algorithm[82] was used to treat the long-range electrostatic interactions, and the SHAKE algorithm[83] was applied to constrain the bonds involving hydrogen atoms. The non-bonding interaction was truncated at 12 Å. The solvated complex system was further performed using the MD simulation production. According to the trajectory analysis, the mean square displacement and the dipole-dipole correction function of the hydrated shell water molecules near the G-quadruplex complex surface were measured. The simulation of the diffusional association method[84] was used to calculate the association rate between the G-quadruplex and the ligand.

Steered Molecular Dynamics (SMD). The ligand-bound G-quadruplex complex was chosen for further SMD and umbrella sampling simulations. After energy minimization using SD and ABNR algorithms, each complex system was solvated with TIP3P water cubic box and counter ions were added to neutralize the system. Then, the system was equilibrated for 1ns MD simulation in the NVT and then NPT ensemble. The SMD simulation was performed with small constraint force of 0.5

kcal/(mol Å^2) on the heavy atoms of G-quadruplex and three centered Na^+ ions to retain the structure, and was used to extract the initial coordinates for umbrella sampling simulations along the unbinding pathway. A pulling force with a virtual spring constant of 7.2 kcal/(mol Å^2) was imposed on the ligand center of mass and the speed was 0.01 Å/ps based on considerations of sampling accuracy and computational efficiency. The reaction coordinate, r, is defined as the distance between the ligand center of mass and the O5' atom of the T5 base on the G-quadruplex when pulling the ligand out of the G-quadruplex (i.e., r = 0 in the bound state) (Fig. 1B). For each pathway, a 3 ns SMD simulation was performed to pull the ligand completely out of the G-quadruplex into bulk solvent. Ten repeats of the SMD runs were performed.

Umbrella Sampling. This method adopts a biased potential function imposing on the structure generated by the SMD along the reaction coordinate, according to the following equation:

$$u'(\mathrm{r}) = u(\mathrm{r}) + w(\mathrm{r}) \quad (1)$$

where $u(\mathrm{r})$ is the potential function and $w(\mathrm{r})$ is a weighting function with a quadratic form,

$$w(\mathrm{r}) = k(\mathrm{r} - r_0)^2 / 2 \quad (2)$$

where k is the harmonic force constant and r_0 is the harmonic potential center. The ligand position probability distributions along the reaction coordinate were extracted from the simulations. The size of umbrella window was 1 Å and there were 35 windows along the reaction coordinate. For each umbrella window, the complex structure was subjected to a 500 ps equilibration, followed by a 3 ns production simulation under a harmonic force constant of 5 kcal/(mol Å^2). Convergence of the calculations was examined by extending the simulation times. The weighted histogram analysis method (WHAM)[85] was used to combine all windows to compute the PMF.

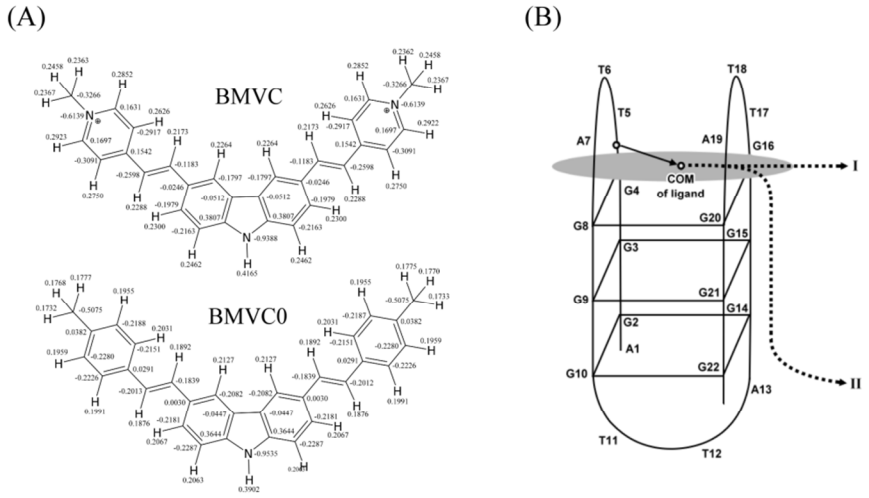

Fig. 1 **Structures.** (A) Structures and partial charges of BMVC and BMVC0. (B) The schematic bound-state G-quadruplex/ligand complex structure. The ligand binding site is indicated in the shadow region. The vector (solid line) connects a pair of atoms: the origin point is at the O5' atom of base T5 in the G-quadruplex, and the acting point is at the ligand center of mass. The two unbinding pathways are indicated by labeling I and II (dashed line); this defines the reaction coordinate for SMD simulation and PMF calculation.

C Results and Discussion

The root mean squared displacement (RMSD) analysis of the G-quadruplex/BMVC complex structure showed that during the 20 ns MD simulation the RMSDs of the G-quartets remain at a constant value of ~ 1.5 Å, whereas those of the loops are greater than 3.0 Å (Supplementary Fig. S1-A), implying that the G-quadruplex structure is relatively stable during the simulation. The force constant and pulling velocity were calibrated during the SMD simulation. Comparison of the pulling distance and the constrained distance measured in the SMD simulation of the G-quadruplex/BMVC complex performed with the force constant 7.2 kcal/(mol Å^2) and the velocity 0.01 Å/ps showed that the stiff-spring approximation was satisfied (Supplementary Fig. S1-B).

Rupture Process. We performed the rupture of BMVC from a G-quadruplex bound state to explore the mechanism by which BMVC enters the end-stacked G-quartet binding sites and forms a stable complex. Indeed, there are apparent multiple pathways for ligand binding to the G-quadruplex. Based

on the experiment and simulation results[52,55,56,86] that predicted BMVC binding sites, we performed PMF computations of BMVC unbinding from the lateral loops and the diagonal loop end-stacked binding sites (Supplementary Fig. S2). The result showed that the most appropriate pathway for BMVC was via a route beneath the loop region. Furthermore, two possible pathways for ligand unbinding from the ligand-bound G-quadruplex were proposed. Pathway I involves ligand escape from the bound state via the loop region followed by ligand entry into the bulk. Pathway II involves ligand movement via the loop region and then via the groove surface, followed by entry into the bulk (Fig. 1B).

Fig. 2A-(a) presents the profile of the rupture force associated with unbinding pathway I of the ligands BMVC and BMVC0 from the G-quadruplex along the reaction coordinate. Note that BMVC0 is an analogue of BMVC in which the N atom in the pyridinium group is replaced with a C atom. All snapshot structures obtained by SMD simulation are shown in Supplementary Fig. S3. Two relatively broad peaks of the rupture forces are present at $2 < r < 10$ Å and $10 < r < 28$ Å. The rupture force decreases to zero at $r = 10$ Å and $r > 28$ Å. It is obviously shown that the maximum rupture force corresponds to a stable structure of the G-quadruplex/ligand complex in which the structures at $r = 3$ and 21 Å have better stacking form than those at $r = 10$ Å (Fig. 2B). Our structural analysis revealed that the carbazole and pyridinium portions of BMVC could stack with the adjacent G-quartet and the TTA loops, where this structural stacking contributes to the rupture force in the process. A greater rupture force is required for BMVC than for BMVC0.

Coulomb interaction and van der Waals interaction. In the bound state, the ligand stacks well on the G-quartet (bases G4, G8, G16 and G20) and is surrounded by various loops (bases T5, T6, A7, T17, T18 and A19); the G-quartet structure remains nearly intact throughout the simulation. However, in the SMD simulation, the hydrogen bonds in the G-quartet are disrupted, and the bases become flexible, as shown in Fig. 2B. In this scenario, it is important to examine the effect of the electrostatic and van der Waals interactions due to the ligand during the SMD simulation. Our analysis can estimate the strength of the influence of Coulomb interaction and van der Waals interaction on the structural distortion.

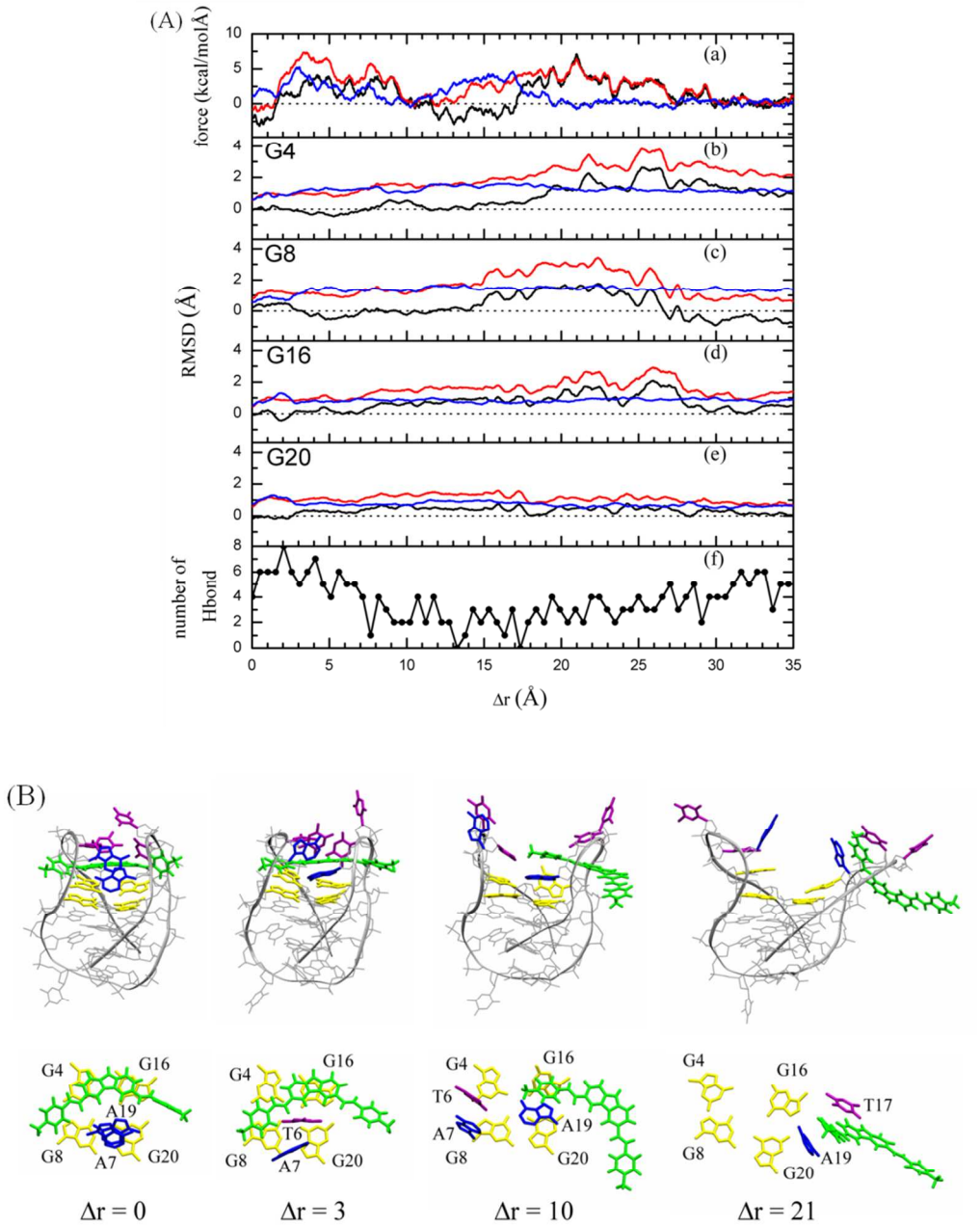

Fig. 2 **Rupture process of the G-quadruplex/ligand complex.** (A) (a) The effective force *vs*. the reaction coordinate. RMSD values for the heavy atoms of bases (b) G4, (c) G8, (d) G16, and (e) G20 along the reaction coordinate for BMVC (pathway I) and BMVC0 (pathway III) comparing the G-quadruplex/BMVC complex (red), the G-quadruplex/BMVC0 complex (blue), and the RMSD difference between the two complexes (black). (f) The number of hydrogen bonds among bases G4, G8, G16 and G20. (B) Snapshots of the G-quadruplex/BMVC complex as a side view (upper) and top view (bottom) from the SMD simulations. BMVC, guanine, adenine, thymine and the DNA backbone are shown in green, yellow, blue, purple, and gray, respectively.

The RMSD of bases G4, G8, G16 and G20 along the reaction coordinate versus the rupture force is shown in Fig. 2A-(b-e). In the BMVC0 system, these bases remain relatively intact, with only a small fluctuation in the RMSD of 1.0 Å throughout the process. However, in the BMVC system, at r < 10 Å, these bases have an RMSD value of 1.0 Å, whereas at r > 10 Å, the BMVC slides above the contacted bases G4, G8 and G16, leading to an increase in their RMSD values. At r < 10 Å, the RMSD difference between the two ligand complex systems is very small (Fig. 2A-(b-e)). Because BMVC remain stacked with the G-quartet in which the bases are neutral, the charge effect could be negligible. Upon pulling, van der Waals interaction is a major contributor to the pulling force. At 15 < r < 28 Å, the RMSD values of these contacted bases increase considerably to 3 Å only for the BMVC system. This increase is attributable to the increased importance of electrostatic interactions as the positively charged BMVC moves closer to the negatively charged phosphodiester backbone; Coulomb interaction becomes the major force involved in breaking apart the planar G-quartet structure. Notably, no dramatic RMSD change is observed for base G20 because BMVC does not stack directly above it. After leaving the G-quartet (r = 21 Å), BMVC stacks with the bases T17 and A19, which act as a clamp to hold BMVC. Coulomb interactions pull the bases in the loop, causing it to depart from its original position along the pulling force, while the G-quartet structure twists to a limited degree relative to the loop (Fig. 2B). The RMSD values of the loops are shown in Supplementary Fig. S4. Remarkably, the displacements of the loops are relatively large. During the unbinding process, BMVC may stick to the nucleotides that are close to it and make them toward a maximum displacement, but they finally bounce back at r > 30 Å. In addition, the number of hydrogen bonds among bases G4, G8, G16 and G20 was measured (Fig. 2A-(f)). The H-bond number decreases upon BMVC leaving. When BMVC is at the separated state (r = 35 Å), the H-bond number is restored to its original value and the G-quartet planar structure can be reformed. These analyses demonstrated that electrostatic interactions contribute significantly to the stabilization of the complex structure, whereas the RMSD difference between the BMVC and BMVC0 complex systems is primarily caused by Coulomb interaction.

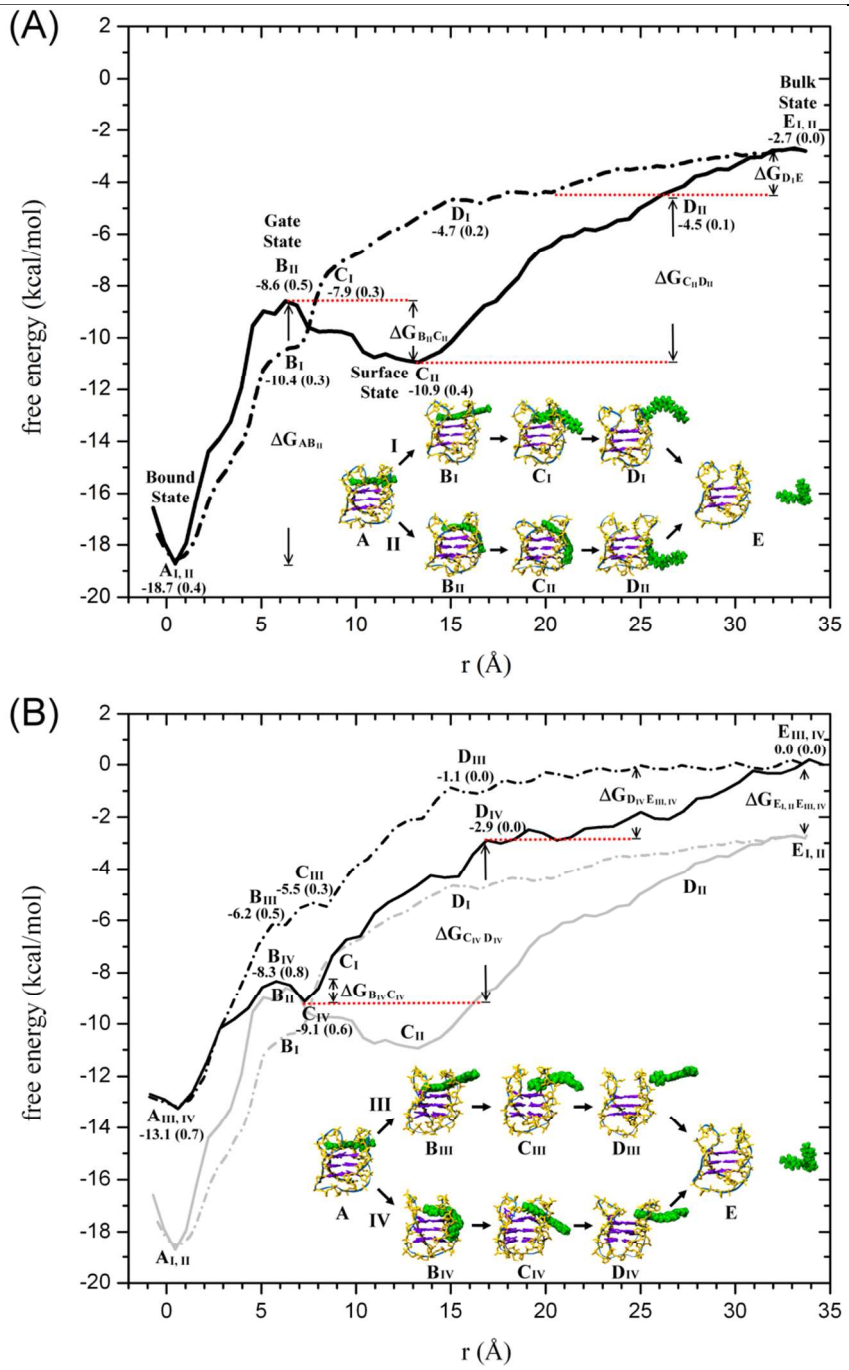

Fig. 3 **Free energy profiles.** (A) For BMVC, two pathways, I (black dashed dot) and II (black solid), are shown. The states are labeled A_I-E_I, etc. $\Delta G_{A_I B_I}$, for example, denotes the free energy change for the step $A_I \rightarrow B_I$. Inset: Snapshots of the G-quadruplex/BMVC complex structures are depicted from the umbrella sampling simulations. The standard error is shown in parenthesis. (B) For BMVC0, two pathways, III (black dashed dot) and IV (black solid), are shown. All notations are the same as in (A). Pathways I (gray dashed dot) and II (gray solid) for BMVC are also shown for comparison.

PMF calculations along the Ligand Unbinding Process. The unbinding processes of both ligands (BMVC and BMVC0) from the G-quadruplex were further used to investigate the free energy (or PMF) profiles using WHAM analysis of the umbrella sampling simulations (Fig. 3). The convergence of the PMF calculations was examined using five repeats. When BMVC0 and the G-quadruplex are separated at r = 35 Å, the free energy has a zero value as a reference in the bulk state. According to our simulations, the solvation free energy difference between BMVC and BMVC0 can be obtained by calculating the free energy difference of the two ligands in the bulk state, $\Delta G_{E_{(BMVC)}E_{(BMVC0)}} = 2.7$ kcal/mol. Because the charged BMVC is more likely to dissolve in aqueous solution compared with the neutral BMVC0, the solvation free energy for BMVC is, thereby, more negative by ~2.7 kcal/mol than that for BMVC0. This difference makes the calibration of the referenced free energy in the bulk state for BMVC.

(a) Pathways I and II for Charged BMVC. In Fig. 3A, pathways I and II are matched in the bound state and in the bulk state, as would be expected; however, pathway II is relatively steeper than pathway I. Pathway I for BMVC is characterized by the following stages: $A_I \rightarrow B_I \rightarrow C_I \rightarrow D_I \rightarrow E_I$ (Fig. 3A). First, BMVC is in the bound state (A_I) and forms a stable G-quadruplex/BMVC complex in which the carbazole and two pyridinium rings of BMVC stack on the G-quartet. In the process $A_I \rightarrow B_I$, when the ligand is pulled away from the bound state to the loop region (TTA loop), the free energy change is $\Delta G_{A_IB_I} \approx 8.3 \pm 0.5$ kcal/mol. The process $B_I \rightarrow C_I \rightarrow D_I$ indicates that BMVC subsequently passes through the loop region and enters the hydration shell; the free energy change is approximately 5.7 ± 0.4 kcal/mol. Finally, in the process $D_I \rightarrow E_I$, BMVC leaves the D_I state and enters the bulk state, assisted by thermal fluctuation to overcome the free energy of the surface hydration shell. The properties of the hydration water molecules that are nearby DNA differ from those of bulk water,[88] and it is difficult to measure the hydration shell free energy experimentally. In accordance with our analysis, the free energy of the surface hydration shell can be estimated as $\Delta G_{D_IE_I} \approx 2.0 \pm 0.2$ kcal/mol. This DNA hydration shell free energy is very close to that of a protein[89] and is approximately three times the thermal fluctuation energy.

Pathway II can be distinguished as the process $A_{II} \rightarrow B_{II} \rightarrow C_{II} \rightarrow D_{II} \rightarrow E_{II}$, in which BMVC leaves the bound state (A_{II}) via the loop region (B_{II}), succeeded by sliding on the groove surface (C_{II}),

penetrating the hydration shell and, finally, entering the bulk state (E_{II}) (Fig. 3A, inset). As the results shown, $\Delta G_{A_{II}B_{II}}$ is greater than $\Delta G_{A_IB_I}$ because the BMVC moiety is solvated by water molecules in pathway I and ΔG_{solv} plays a key role in lowering the free energy change during this process. In contrast, for pathway II, the same BMVC moiety kinks and adsorbs on the groove surface in the loop region (B_{II}). When BMVC arrives at the groove surface (C_{II}) via the loop region (B_{II}), there is an apparent barrier height with $\Delta G_{A_{II}B_{II}} \approx 10.1 \pm 0.6$ kcal/mol, implying that the loop region acts as a gate to prevent the ligand from drifting into the bulk or allowing it to enter the bound state; this is called the gate state. The motion of the loop region regulates the gate dynamics. BMVC at the groove surface state has a relatively local minimum free energy of -10.9 kcal/mol, and the free energy change for the process $B_{II} \rightarrow C_{II}$ is $\Delta G_{B_{II}C_{II}} = -2.3 \pm 0.6$ kcal/mol. Furthermore, in the process $C_{II} \rightarrow D_{II} \rightarrow E_{II}$, the corresponding groove surface unbinding free energy change is $\Delta G_{C_{II}E_{II}} = 8.2 \pm 0.4$ kcal/mol.

These two pathways have an overall free energy change of approximately 16.0 kcal/mol that is associated with BMVC unbinding from the bound state to the bulk state. The docking results showed that the G-quadruplex groove surface is the most accessible binding site. Accordingly, a two-step ligand binding scheme is proposed. First, a ligand arrives at the G-quadruplex surface and then shuffles on the groove region. Eventually, the ligand enters the gate part, i.e., the loop region, and finally forms a stable end-stacked state with the G-quartet (bound state). In this scenario, a barrier height exists between the bound state and the groove surface state.

(b) Pathways III and IV for Neutral BMVC0. To further study the ligand charge effect on the unbinding process, we performed the same procedures for BMVC0 to generate pathways III and IV (Fig. 3B). BMVC0 is a neutral molecule and thereby has weaker electrostatic interactions with the G-quadruplex and a lower solvation free energy. Thus, the PMF values of pathways III and IV for BMVC0 are relatively less negative than those of pathways I and II for BMVC. Several features of the simulations are notably discussed (Fig. 3).

(1) $\Delta G_{A_{III(IV)}E_{III(IV)}} < \Delta G_{A_{I(II)}E_{I(II)}}$. Here, $\Delta G_{A_{i(j)}E_{i(j)}}$ indicates the ΔG values from A_i to E_i and from A_j to E_j, respectively. The $\Delta G_{A_{III(IV)}E_{III(IV)}}$ required to rupture BMVC0 from the bound state ($A_{III(IV)}$) into the bulk ($E_{III(IV)}$) is approximately 13.1 ± 0.7 kcal/mol, which is 2.9 kcal/mol lower than the $\Delta G_{A_{I(II)}E_{I(II)}}$ value for

BMVC (16.0 ± 0.4 kcal/mol). This difference is primarily due to the ligand charge effect.

(2) $\Delta G_{A_{IV}B_{IV}} < \Delta G_{A_{II}B_{II}}$. In the region $A_{II(IV)} \rightarrow B_{II(IV)}$, BMVC and BMVC0 are in the same conformation on the G-quadruplex surface, and the solvation effect has not yet been established. In comparison, the nonbonding interaction between BMVC0 and the G-quadruplex is smaller than that of BMVC.

(3) At r < 3 Å, $PMF_{A_{III}B_{III}} \sim PMF_{A_{IV}B_{IV}}$. Here, $PMF_{A_iB_i}$ indicates the PMF value in the region from A_i to B_i and so on. In the processes $A_{III(IV)} \rightarrow B_{III(IV)}$ inside the loop region, the charge effect is weak. At r > 3 Å, $PMF_{A_{III}B_{III}} > PMF_{A_{IV}B_{IV}}$, the BMVC0 moiety enters the hydration shell region, and the solvation effect is initiated in the process $A_{III} \rightarrow B_{III}$. Note that the solvation effect is weak for BMVC0. However, in the process $A_{IV} \rightarrow B_{IV}$, BMVC0 slides on the G-quadruplex surface, where the nonbonding interaction between BMVC0 and the G-quadruplex dominates the free energy change and the solvation effect is negligible. Hence, the difference of $\Delta G_{A_{III}B_{III}} - \Delta G_{A_{IV}B_{IV}} \approx 2.1$ kcal/mol is mainly due to ΔG_{solv}. Similarly, for $PMF_{A_{II}B_{II}} > PMF_{A_IB_I}$, the solvation effect plays a key role in lowering the PMF value in the process $A_I \rightarrow B_I$.

(4) $\Delta G_{B_{III(I)}C_{III(I)}} > 0 > \Delta G_{B_{IV(II)}C_{IV(II)}}$. The groove surface binding interactions and the ligand charge effect make a significant difference in the free energy changes in these processes $B_{III(I)} \rightarrow C_{III(I)}$ and $B_{IV(II)} \rightarrow C_{IV(II)}$.

(5) $PMF_{C_{IV}} > PMF_{C_{II}}$ and $r(C_{IV}) < r(C_{II})$. Due to electrostatic interactions, the ligand binding modes are remarkably different between the C_{II} and C_{IV} states (Figs. 3A and 3B, inset). In the C_{II} state, the entire structure of BMVC remains bound to the G-quadruplex. In the C_{IV} state, the BMVC0 moiety binds to the G-quadruplex, and the remainder is solvated by water molecules. At this stage, only half of BMVC0 is solvated by water molecules, and the solvation effect can be rewritten as $\Delta G_{C_{IV}C_{II}} \sim \frac{1}{2} G_{E_{IV}E_{II}}$.

(6) $\Delta G_{C_{III}E_{III}} \sim \Delta G_{C_IE_I}$ and $\Delta G_{C_{IV}E_{IV}} \sim \Delta G_{C_{II}E_{II}}$. These relationships imply that the ligand charge effect is

weak in the processes $C_{I(II)} \rightarrow E_{I(II)}$ and $C_{III(IV)} \rightarrow E_{III(IV)}$. The groove surface unbinding free energy changes for BMVC0 and BMVC are approximately $\Delta G_{C_{IV}E_{IV}} = 9.1 \pm 0.6$ and $\Delta G_{C_{II}E_{II}} = 8.2 \pm 0.4$ kcal/mol, respectively.

Consequently, we have computed all of the free energy changes associated with the ligand unbinding process. The binding constant of BMVC with the G-quadruplex was obtained as 1.12×10^9 M^{-1} in Na^+ solution;[52] hence, the corresponding binding free energy change is estimated -12.42 kcal/mol, which is in agreement with our computed end-stacking binding free energy change $\Delta G_{A_{I(II)}E_{I(II)}}$ (-16.0 ± 0.4 kcal/mol). In addition, our computational results are also supported by experimental evidence. Koirala et al.[43] measured a groove binding free energy change ΔG_{bind} of -8.6 kcal/mol for pyridostatin binding to the G-quadruplex. Di Leva et al.[45] reported a groove binding free energy change ΔG_{bind} of -9.4 ± 1.4 kcal/mol for a ligand of interest. Typically, the experimental ΔG_{bind} values of lead compounds for drug development are -6.5 to -13.0 kcal/mol.[90] Consistent with the similar chemical structures of BMVC and pyridostatin, the computed groove surface binding free energy changes $-\Delta G_{C_{II}E_{II}} = -8.2$ kcal/mol for BMVC and $-\Delta G_{C_{IV}E_{IV}} = -9.1$ kcal/mol for BMVC0 are similar to that for pyridostatin (-8.6 kcal/mol).[43] Notably, our computed ΔG_{bind} values fall in a reasonable range of experimental values. Furthermore, when the finite concentration and ionic contribution free energy corrections[91] are considered, we can correct the binding free energy change as $\Delta G_{corrected} = \Delta G_{bind} + \Delta G_{ion} - \Delta G_{conc}$, where ΔG_{bind}, ΔG_{conc} and ΔG_{ion} are the binding free energy change, the concentration and ionic contribution free energy corrections, respectively. In the present study, the corrected term is $\Delta G_{ion} - \Delta G_{conc} = 0.4$ kcal/mol (see the Supplementary information, Table S1).

Conformational Effect. The same approach was applied to other human G-quadruplex structures, such as propeller (PDB: 1KF1),[92] hybrid-1 (PDB: 2HY9)[93] and hybrid-2 (PDB: 2JPZ)[94] types, and the PMF result was used to compare with that of the basket type. First, we searched for possible ligand-binding sites using a docking method. The binding sites (i.e., end-stacked G-quartet, side loop and groove surface), that have higher ranking, and their docking score are shown in Supplementary Fig. S5. We found that the end-stacked G-quartet binding site has higher docking score among the binding sites and our docking results are in good agreement with the results obtained by Alcaro et al.[95] Remarkably, the PMF calculation of BMVC unbinding process for each model displays a similar mechanism

involving the end-stacked G-quartet (A state), the groove surface (C state) and the bulk (E state) (Fig. 4). For the three G-quadruplex structures, we found that the A state is the most stable G-quadruplex-BMVC complex structure that the ligand stacks well on the G-quartet and the ΔG_{AE} value is slightly less (~2.0 kcal/mol) than that in the basket type, where BMVC not only is end-stacked on the G-quartet but also is bound by the lateral loops. In the propeller type, BMVC is initially intercalated in between the groove and the side loop binding site (S state) that has the highest docking score; however, after pulling BMVC to the A state, the PMF change is of about -3.7 kcal/mol, indicating that the end-stacked binding is superior to the groove-side loop binding. There is a local minimum at the C state where BMVC is warped by the side loop. In comparison, the ΔG_{AB} values and the associated groove widths are approximately: propeller (7.8, 14.4), hybrid-2 (8.8, 11.9), basket (10.1, 10.0) and hybrid-1 (14.8, 16.4) (kcal/mol, Å). In the case of hybrid-1, while BMVC moves on the wide groove surface, there is no apparent conformation change for BMVC until it reaches the C state and the barrier ΔG_{AB} is larger than that in the other conformations. This is due to the larger contact area of BMVC with the wide groove surface. On the contrary, the lower barrier ΔG_{AB} appears when BMVC moiety leaves the G-quartet region and kinks its conformation fit toward the medium groove surface in the propeller and hybrid-2 that is similarly observed in the basket type. Significantly, the groove width and the loop type are distinct for the various G-quadruplex structures; the PMF change of the ligand unbinding is groove pathway-dependent. The present studies demonstrate that the loops and groove pathway make important contributions to the ligand binding process.

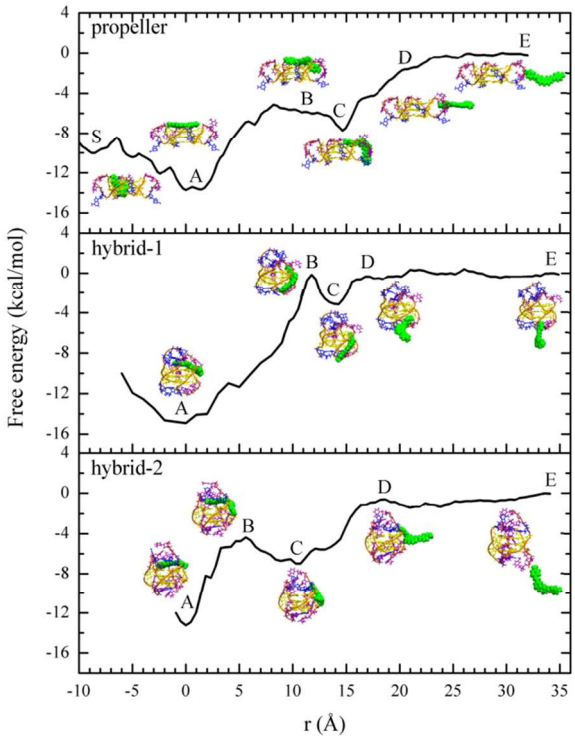

Fig. 4 **Free energy profile of BMVC unbinding process.** The G-quadruplex-BMVC complex structures are propeller, hybrid-1 and hybrid-2 types. The states are labeled A-E, etc. Inset: Snapshots of the complex structures are depicted from the umbrella sampling simulations.

Hydration Shell Thickness. To investigate the dynamic motion of the hydration shell water molecules near the G-quadruplex/BMVC complex surface, the mean square displacement Δr^2 *vs*. time of these water molecules was measured (Supplementary Fig. S6-A). The result showed that these water molecules exhibit abnormal Brownian motion and follow a fractional diffusion process, which obeys the relationship, $\Delta r^2 \propto t^{\alpha}$ and $\alpha \neq 1$. The temperature-dependent Δr^2 within a distance from the G-quadruplex surface of 15 Å is also shown (Supplementary Fig. S6-B). Thus, a scaling power α value across the different regions was fitted (Table 1). When a water molecule is very close to the G-quadruplex surface, the α value becomes smaller that reflects an aberrant Brownian motion for water molecule within the hydration shell. As the temperature increases, the hydration shell water molecule ultimately exhibits a normal Brownian motion. Notably, the computed α value for the G-quadruplex system is smaller than that for a protein system such as myoglobin.[89] This discrepancy occurs because the G-quadruplex DNA is a polyelectrolyte, while a protein is a polyamphiphilic macromolecule.

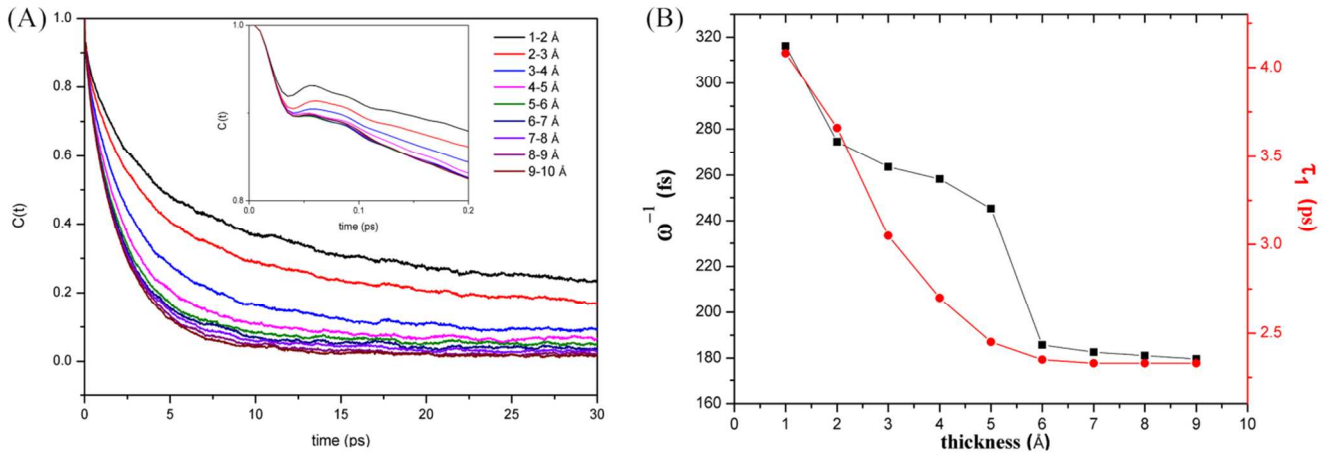

Fig. 5 (A) **Dipole-dipole autocorrelation function of the water molecule.** The water dipole-dipole autocorrelation as a function of the distance between water molecules and the G-quadruplex surface is plotted *vs*. time. Inset: An enlarged correlation function within the time 0.2 ps is shown. (B) **Liberation time (ω^{-1}) and dipole relaxation time τ_1 of the water molecules.** The fitted liberation time (the solid square using the left scale) and the dipole relaxation time (the solid circle using the right scale) are shown *vs*. the distance (in Å) between the water molecules and the G-quadruplex surface.

Table 1. Scaling power α parameters from the linear fitting.

	Scaling power α					
	$\overline{R}$ = 4 Å		$\overline{R}$ = 6 Å		$\overline{R}$ = 15 Å	
T (K)	0-0.2 (ps)	1-50 (ps)	0-0.2 (ps)	1-50 (ps)	0-0.2 (ps)	1-50 (ps)
277	1.57±0.03	0.60±0.01	1.60±0.03	0.75±0.01	1.63±0.03	0.91±0.01
300	1.58±0.03	0.62±0.01	1.61±0.03	0.76±0.02	1.65±0.02	0.92±0.01
350	1.62±0.03	0.69±0.01	1.64±0.02	0.79±0.01	1.68±0.02	0.93±0.01
Pure water						
300	1.65±0.02	0.97±0.01				

Moreover, we computed the water dipole-dipole autocorrelation function $C(t)$ relative to the

normal distance from the G-quadruplex surface (Fig. 5A). The autocorrelation function was fitted to the following typical expression:

$$C(t) = \frac{<\sum_i \vec{\mu_i}(t) \cdot \vec{\mu_i}(0)>}{<\sum_i \vec{\mu_i}(0) \cdot \vec{\mu_i}(0)>} = N^{-1} \sum_i <\cos\ \theta_i(t)> = A_1 e^{-\omega^2 t^2/2} + A_2 e^{-t/\tau_1} \quad (3)$$

where $\vec{\mu_i}(t)$ is the dipole moment at time t, $\theta_i(t)$ is the angle between $\vec{\mu_i}(t)$ and $\vec{\mu_i}(0)$, N is the number of water molecules, ω describes the liberational motion of the free water molecule, the longer dipole relaxation time τ_1 is an orientation autocorrelation time, and A_1 and A_2 are prefactors; all of these variables are hydration shell thickness-dependent. It is able to extract the liberation time and dipole relaxation time of the water molecules from the $C(t)$. For bulk water the liberation time ω^{-1} and the longer dipole relaxation time τ_1 are approximately 179 fs and 2.3 ps, respectively (Fig. 5B). As shown in Fig. 5B, there is an abrupt decrease in the dipole relaxation time at around the thickness = 6.0 Å, which relates to the hydration shell thickness of the G-quadruplex.

In comparison, a seminal study of water relaxation around the coumarin 343 ion by Fleming et al. showed a strong oscillation of the time correlation function within 100 fs.[96] In the G-quadruplex system, the present study also shows oscillation behavior similar to that obtained by Fleming et al.; however, the amplitude of the present oscillation is relatively small and damps within 50 fs (Fig. 5A, inset). This is because DNA is a polyelectrolytic ion and has a heterogeneous structure. Meanwhile, Zewail et al. reported that in a drug-duplex DNA system the water molecule has a liberation time of 290 fs and a dipole relaxation time of ~1.7 ps.[50] Surprisingly, both our computed liberation and relaxation time scales are consistent with Zewail's experimental data.

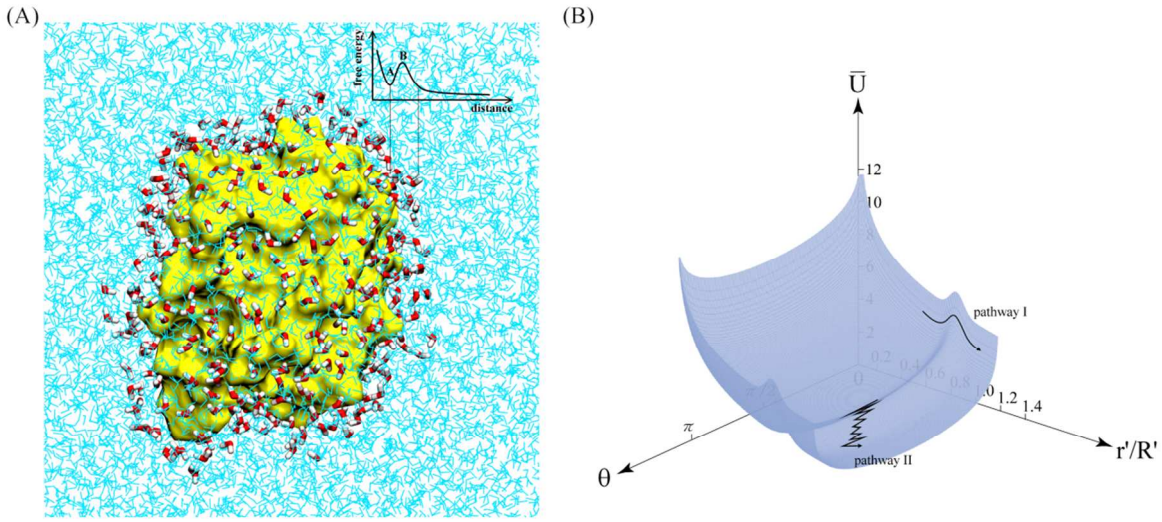

Fig. 6 **Model of the surface hydration shell free energy.** (A) A solvated G-quadruplex system. The G-quadruplex, the hydration shell and the bulk water molecules are shown in yellow, red and cyan, respectively. Inset: the hydration shell free energy model. "**A**" and "**B**" denote the local minimum energy state and the energy barrier B, respectively. (B) Entropy potential and hydration shell free energy around the G-quadruplex surface.

Theoretical Model of Hydration Shell Free Energy. The hydration shell barrier plays an important role for a ligand entering the biomolecule. Based on the preceding result, the free energy change in the last stage, $D_{I(II)} \rightarrow E_{I(II)}$, corresponds to the hydration shell barrier, $\Delta G_{D_{II}E_{II}}$. Furthermore, we developed a theoretical model of the free energy surface to describe the unbinding of the ligand from the bound state to the separated state by passing through the surface hydration shell. A hydration shell free energy model of the G-quadruplex is depicted in Fig. 6A. In this model, the G-quadruplex itself is assumed as a hard sphere potential, whereas the hydration shell water molecules are stable and unable to penetrate the G-quadruplex. Outside the G-quadruplex, a local minimum "**A**" is located on the G-quadruplex surface, the energy barrier top "**B**" is defined at the first hydration shell from the G-quadruplex surface, and the free energy decays to an asymptotic value of zero while it approaches the bulk water (Fig. 6A, inset).

The inside of G-quadruplex is modeled as a spherical cavity with a gate on its surface. Outside the cavity, we can use a cubic polynomial function to describe the hydration shell free energy:

$$V(r') = -ar'^3/3 + br' \qquad (4)$$

Furthermore, the free energy can be decomposed according to the following equation:

$$\overline{U} = W + U, \quad (5)$$

where $W(r') = -k_B T \ln(r'/R')^2 H(R'-r') + V(r')H(r'-R')$ and $U = -k_B T \ln\sin\theta$.[89] Here, $H(x)$ is the Heaviside step function, r' is the distance from the cavity center, R' is the radius of the cavity, θ is the polar angle of the gate part, k_B is the Boltzmann constant, and T is the temperature (Fig. 6B). Inside the cavity, the angular part of the surface entropy potential is a logarithm of the sine function, U. On the cavity surface, once the ligand arrives at the gate parts, θ = 0 or π, the surface entropy potential approaches infinity. The radial part of the entropy potential is a logarithm of the r'^2 function inside the cavity, but outside, it is replaced by the function $V(r')$. When the ligand motion overcomes the top of the hydration shell barrier, it can escape. Based on our estimated barrier height and the hydration shell thickness, we obtained a = 0.11 kcal/(mol Å^3) and b = 1.01 kcal/(mol Å) in $V(r')$. In Fig. 6B, pathways I and II indicate that the ligand moves in a restrained manner on the entropy potential surface. Pathway I shows that the ligand crosses over the barrier top and directly escapes. Pathway II shows that the ligand diffuses on the G-quadruplex surface, and then escapes. This theory confirms that the ligand unbinding process through the hydration shell in biomolecules is both enthalpy-dependent and entropy driven.

Relaxation Process. Based on the results from pathway II, we proposed that the BMVC unbinding process follows a two-step reaction mechanism and that there is one intermediate state C_{II}. Thus, we outline the following feasible reaction scheme:

$$A_{II} \underset{k_{C_{II}A_{II}}}{\overset{k_{A_{II}C_{II}}}{\rightleftarrows}} C_{II} \underset{k_{E_{II}C_{II}}}{\overset{k_{C_{II}E_{II}}}{\rightleftarrows}} E_{II} + L$$

where A_{II}, C_{II} and E_{II} denote the bound state, the groove state and the separated state, respectively, L is the ligand, and $k_{A_{II}C_{II}}$ describes the transition rate from state A_{II} to C_{II}, etc. In the separated state, the association rate constant of a receptor and a ligand depends on the ligand concentration. The process A_{II} □ C_{II} is a unimolecular reaction, which is independent of the ligand concentration. The association process of E_{II} and L is a bimolecular reaction. Here, we assumed that the ligand concentration, [L], is nearly constant, $[L] \approx [L_0]$, where $[L_0]$ is the initial concentration. The kinetic equation obeys Eq. (6), as follows:

$$\frac{\partial}{\partial t}\phi = M\ \phi \quad (6)$$

where ϕ=[[A_{II}], [C_{II}], [E_{II}]] and M is given in Eq. (7), as follows:

$$M = \begin{bmatrix} -k_{A_{II}C_{II}} & k_{C_{II}A_{II}} & 0 \\ k_{A_{II}C_{II}} & -(k_{C_{II}A_{II}} + k_{C_{II}E_{II}}) & k_{E_{II}C_{II}} \\ 0 & k_{C_{II}E_{II}} & -k_{E_{II}C_{II}}[\mathrm{L}_0] \end{bmatrix} \quad (7)$$

Note that $k_{E_{II}C_{II}}$ is an association rate constant. In the present study, we estimated the second-order bimolecular association rate of the G-quadruplex and the ligand using the simulation of the diffusional association method.[84] The calculated rate constant $k_{E_{II}C_{II}}$ is 2.3 × 10^9 M^{-1} s^{-1} for BMVC and 9.3 × 10^5 M^{-1} s^{-1} for BMVC0, illustrating that the charged ligand plays a key role in dictating the ligand association rate. By solving the eigenmodes of Eq. (7), ϕ(t) can be exactly obtained according to the following equation:

$$\phi(t) = c_1 e^{-t/\lambda_1} + c_2 e^{-t/\lambda_2} + c_3 e^{-t/\lambda_3}, \quad (8)$$

where c_i are constants and λ_i are relaxation times. Combined with the PMF results, we were able to obtain three relaxation times, $\lambda_1 = 333$ s, $\lambda_2 = 11$ s and $\lambda_3 = \infty$ (Table 2), where the infinite relaxation time corresponds to a constant term. These relaxation times are also within the range of experimental data.[97] Hence, we suggest further experiments to investigate these processes.

Table 2. Contributions to the relaxation times of BMVC in pathway II[(a)].

term	$\Delta G_{A_{II}C_{II}}$ (kcal/mol)	$\Delta G_{C_{II}A_{II}}$ (kcal/mol)	$\Delta G_{C_{II}E_{II}}$ (kcal/mol)	$k_{A_{II}C_{II}}$ (s^{-1})	$k_{C_{II}A_{II}}$ (s^{-1})	$k_{C_{II}E_{II}}$ (s^{-1})	$k_{E_{II}C_{II}}[L_0]$ (s^{-1})	λ_1 (s)	λ_2 (s)	λ_3 (s)
value	15.1	3.4	12.2	7.5×10^{-7}	0.1	3.1×10^{-6}	2.3×10^{-3}	333	11	∞

(a) ΔG_{ij} is the activation energy from state *i* to *j*, k_{ij} is the transition rate from state *i* to *j*, k_{EC} is an association rate constant, $[\mathrm{L}_0]$=1 M is the initial concentration of ligand, λ_i are relaxation times, and $k_{ij} = Ae^{-\beta\Delta G_{ij}}$, where A is a pre-factor and $\beta = 1/(k_B T)$ (k_B is the Boltzmann constant and T is the temperature).

D Conclusion

In this study, we have investigated the mechanical properties and kinetic pathways of the ligands (BMVC and BMVC0) unbinding from the G-quadruplex by using SMD and umbrella sampling simulations. By the structural analyses and the PMF calculations, we have shown detailed atomistic insights and relevant free energies involved in the unbinding processes, which have not been determined in ensemble assays. The simulations for both these ligands clearly reflect their distinct rupture force and PMF profiles, which can be correlated with G-quadruplex/ligand binding affinities, suggesting that this protocol could be applied to other G-quadruplex ligands to assess its robustness and rank a series of derivatives with similar potencies. Specifically, while a positively charged ligand could stack with the G-quartet, it might impact the interactions between two molecules. Our findings confirm the seminal experimental data, highlighting the importance of using computer-based simulations to predict and complement experiments. This demonstration of simulations combined with theoretical analyses is valuable for further investigations on various G-quadruplex systems and could be helpful in structure-based drug design.

Acknowledgments

This study was supported by grants from the National Science Council of Taiwan. DYY and SYS gratefully acknowledge Grants NSC-99-2113-M-001-019 and NSC-102-2113-M-010-002-/MOST-103-2113-M-010-004-, respectively. The authors would like to thank the National Center for High-Performance Computing of Taiwan for computer time and facilities.

Notes

[a] *Department of Life Sciences and Institute of Genome Sciences, National Yang-Ming University, Taipei 112, Taiwan*

[b] *Institute of Atomic and Molecular Sciences, Academia Sinica, Taipei 106, Taiwan*

[c] *Institute of Biomedical Informatics, National Yang-Ming University, Taipei 112, Taiwan*

[†] *These authors contributed equally to this work.*

**To whom correspondence should be addressed: Sheh-Yi Sheu,* Fax: 886-2-28202449; Tel: 886-2-28267233; *Email: sysheu@ym.edu.tw; Dah-Yen Yang,* Fax: 886-2-23620200; Tel: 886-2-23668239; *Email: dyyang@po.iams.sinica.edu.tw*

†Electronic Supplementary Information (ESI) available: [details of any supplementary information available should be included here]. See DOI: 10.1039/b000000x/

References:

1. J. B. Chaires, *FEBS J.*, 2010, **277**, 1098-1106.
2. S. Neidle, *FEBS J.*, 2010, **277**, 1118-1125.
3. J. L. Huppert and S. Balasubramanian, *Nucleic Acids Res.*, 2005, **33**, 2908-2916.
4. J. E. Johnson, J. S. Smith, M. L. Kozak and F. B. Johnson, *Biochimie*, 2008, **90**, 1250-1263.
5. Y. Qin and L. H. Hurley, *Biochimie*, 2008, **90**, 1149-1171.
6. L. H. Hurley, *Nat. Rev. Cancer*, 2002, **2**, 188-200.
7. A. De Cian, L. Lacroix, C. Douarre, N. Temime-Smaali, C. Trentesaux, J. F. Riou and J. L. Mergny, *Biochimie*, 2008, **90**, 131-155.
8. R. K. Moyzis, J. M. Buckingham, L. S. Cram, M. Dani, L. L. Deaven, M. D. Jones, J. Meyne, R. L. Ratliff and J. R. Wu, *Proc. Natl. Acad. Sci. U. S. A.*, 1988, **85**, 6622-6626.
9. C. B. Harley, A. B. Futcher and C. W. Greider, *Nature*, 1990, **345**, 458-460.
10. C. B. Harley and B. Villeponteau, *Curr. Opin. Genet. Dev.*, 1995, **5**, 249-255.
11. V. Lundblad and J. W. Szostak, *Cell*, 1989, **57**, 633-643.
12. N. W. Kim, M. A. Piatyszek, K. R. Prowse, C. B. Harley, M. D. West, P. L. Ho, G. M. Coviello, W. E. Wright, S. L. Weinrich and J. W. Shay, *Science*, 1994, **266**, 2011-2015.
13. W. C. Hahn, S. A. Stewart, M. W. Brooks, S. G. York, E. Eaton, A. Kurachi, R. L. Beijersbergen, J. H. M. Knoll, M. Meyerson and R. A. Weinberg, *Nat. Med.*, 1999, **5**, 1164-1170.
14. B. S. Herbert, A. E. Pitts, S. I. Baker, S. E. Hamilton, W. E. Wright, J. W. Shay and D. R. Corey, *Proc. Natl. Acad. Sci. U. S. A.*, 1999, **96**, 14276-14281.
15. S. Neidle and L. R. Kelland, *Anticancer Drug Des.*, 1999, **14**, 341-347.
16. J. L. Mergny and C. Helene, *Nat. Med.*, 1998, **4**, 1366-1367.
17. H. Han and L. H. Hurley, *Trends Pharmacol. Sci.*, 2000, **21**, 136-142.
18. A. M. Zahler, J. R. Williamson, T. R. Cech and D. M. Prescott, *Nature*, 1991, **350**, 718-720.
19. Q. Li, J. Zhang, L. Yang, Q. Yu, Q. Chen, X. Qin, F. Le, Q. Zhang and J. Liu, *J. Inorg. Biochem.*, 2014, **130**, 122-129.
20. D. Sun, Y. Liu, D. Liu, R. Zhang, X. Yang and J. Liu, *Chem. Eur. J.*, 2012, **18**, 4285-4295.
21. S. Zhang, Y. Wu and W. Zhang, *ChemMedChem*, 2014, **9**, 899-911.

22. M. Ebrahimi and T. Khayamian, *Med. Chem. Res.*, 2014, **23**, 1327-1339.
23. S. A. Ohnmacht, E. Varavipour, R. Nanjunda, I. Pazitna, G. Di Vita, M. Gunaratnam, A. Kumar, M. A. Ismail, D. W. Boykin, W. D. Wilson and S. Neidle, *Chem. Commun.*, 2014, **50**, 960-963.
24. S. A. Ohnmacht and S. Neidle, *Bioorg. Med. Chem. Lett.*, 2014, **24**, 2602-2612.
25. S. Bhattacharya, P. Chaudhuri, A. K. Jain and A. Paul, *Bioconjugate Chem.*, 2010, **21**, 1148-1159.
26. M. Gellert, M. N. Lipsett and D. R. Davies, *Proc. Natl. Acad. Sci. U. S. A.*, 1962, **48**, 2013-2018.
27. H. M. Wong, L. Payet and J. L. Huppert, *Curr. Opin. Mol. Ther.*, 2009, **11**, 146-155.
28. S. Neidle and G. Parkinson, *Nat. Rev. Drug Discov.*, 2002, **1**, 383-393.
29. S. Balasubramanian and S. Neidle, *Curr. Opin. Chem. Biol.*, 2009, **13**, 345-353.
30. G. N. Parkinson, F. Cuenca and S. Neidle, *J. Mol. Biol.*, 2008, **381**, 1145-1156.
31. N. H. Campbell, G. N. Parkinson, A. P. Reszka and S. Neidle, *J. Am. Chem. Soc.*, 2008, **130**, 6722-6724.
32. N. H. Campbell, N. H. A. Karim, G. N. Parkinson, M. Gunaratnam, V. Petrucci, A. K. Todd, R. Vilar and S. Neidle, *J. Med. Chem.*, 2011, **55**, 209-222.
33. M. Micco, G. W. Collie, A. G. Dale, S. A. Ohnmacht, I. Pazitna, M. Gunaratnam, A. P. Reszka and S. Neidle, *J. Med. Chem.*, 2013, **56**, 2959-2974.
34. T. Wilson, P. J. Costa, V. Félix, M. P. Williamson and J. A. Thomas, *J. Med. Chem.*, 2013, **56**, 8674-8683.
35. X.-f. Zhang, H.-j. Zhang, J.-f. Xiang, Q. Li, Q.-f. Yang, Q. Shang, Y.-x. Zhang and Y.-l. Tang, *J. Mol. Struct.*, 2010, **982**, 133-138.
36. J. B. Chaires, *Arch. Biochem. Biophys.*, 2006, **453**, 26-31.
37. I. Haq, *Arch. Biochem. Biophys.*, 2002, **403**, 1-15.
38. B. Pagano, C. A. Mattia, A. Virno, A. Randazzo, L. Mayol and C. Giancola, *Nucleos. Nucleot. Nucl.*, 2007, **26**, 761-765.
39. L. Petraccone, I. Duro, E. Erra, A. Randazzo, A. Virno and C. Giancola, *Nucleos. Nucleot. Nucl.*, 2007, **26**, 675-679.
40. K. J. Breslauer, D. P. Remeta, W. Y. Chou, R. Ferrante, J. Curry, D. Zaunczkowski, J. G. Snyder and L. A. Marky, *Proc. Natl. Acad. Sci. U. S. A.*, 1987, **84**, 8922-8926.
41. L. A. Marky and K. J. Breslauer, *Proc. Natl. Acad. Sci. U. S. A.*, 1987, **84**, 4359-4363.
42. C. Giancola and B. Pagano, *Top. Curr. Chem.*, 2012, **128**, 347.

43. D. Koirala, S. Dhakal, B. Ashbridge, Y. Sannohe, R. Rodriguez, H. Sugiyama, S. Balasubramanian and H. Mao, *Nature Chemistry*, 2011, **3**, 782-787.
44. B. Corry and N. M. Smith, *Chem. Commun.*, 2012, **48**, 8958-8960.
45. F. S. Di Leva, E. Novellino, A. Cavalli, M. Parrinello and V. Limongelli, *Nucleic Acids Res.*, 2014, **42**, 5447-5455.
46. M. K. Gilson and H.-X. Zhou, *Annu. Rev. Biophys. Biomol. Struct.*, 2007, **36**, 21-42.
47. E. Westhof, *Annu. Rev. Biophys. Biophys. Chem.*, 1988, **17**, 125-144.
48. G. G. Prive, U. Heinemann, S. Chandrasegaran, L. S. Kan, M. L. Kopka and R. E. Dickerson, *Science*, 1987, **238**, 498-504.
49. R. H. Austin, K. W. Beeson, L. Eisenstein, H. Frauenfelder and I. C. Gunsalus, *Biochemistry*, 1975, **14**, 5355-5373.
50. X. Qu, C. Wan, H. C. Becker, D. Zhong and A. H. Zewail, *Proc. Natl. Acad. Sci. U. S. A.*, 2001, **98**, 14212-14217.
51. C.-C. Chang, J.-Y. Wu and T.-C. Chang, *J. Chin. Chem. Soc.*, 2003, **50**, 185-188.
52. C. C. Chang, C. W. Chien, Y. H. Lin, C. C. Kang and T. C. Chang, *Nucleic Acids Res.*, 2007, **35**, 2846-2860.
53. C. C. Chang, I. C. Kuo, J. J. Lin, Y. C. Lu, C. T. Chen, H. T. Back, P. J. Lou and T. C. Chang, *Chem. Biodivers.*, 2004, **1**, 1377-1384.
54. C.-C. Chang, J.-Y. Wu, C.-W. Chien, W.-S. Wu, H. Liu, C.-C. Kang, L.-J. Yu and T.-C. Chang, *Anal. Chem.*, 2003, **75**, 6177-6183.
55. S.-W. Liu, J.-F. Chu, C.-T. Tsai, H.-C. Fang, T.-C. Chang and H.-W. Li, *Anal. Biochem.*, 2013, **436**, 101-108.
56. D. Y. Yang, T. C. Chang and S. Y. Sheu, *J. Phys. Chem. A*, 2007, **111**, 9224-9232.
57. Y. Wang and D. J. Patel, *Structure*, 1993, **1**, 263-282.
58. B. K. Mai and M. S. Li, *Biochem. Bioph. Res. Commun.*, 2011, **410**, 688-691.
59. G. Miño, M. Baez and G. Gutierrez, *Eur. Biophys. J.*, 2013, **42**, 683-690.
60. Z. Shen, F. Cheng, Y. Xu, J. Fu, W. Xiao, J. Shen, G. Liu, W. Li and Y. Tang, *PLoS One*, 2012, **7**, e33500.
61. J.-L. Zhang, Q.-C. Zheng, Z.-Q. Li and H.-X. Zhang, *PLoS One*, 2013, **8**, e53811.
62. S. Kalyaanamoorthy and Y.-P. P. Chen, *Phys. Chem. Chem. Phys.*, 2014, **16**, 3777-3791.
63. J. S. Patel, A. Berteotti, S. Ronsisvalle, W. Rocchia and A. Cavalli, *J. Chem. Inf. Model.*, 2014, **54**, 470-480.

64. E. Paci, A. Caflisch, A. Pluckthun and M. Karplus, *J. Mol. Biol.*, 2001, **314**, 589-605.

65. F. Colizzi, R. Perozzo, L. Scapozza, M. Recanatini and A. Cavalli, *J. Am. Chem. Soc.*, 2010, **132**, 7361-7371.

66. H. Grubmüller, B. Heymann and P. Tavan, *Science*, 1996, **271**, 997-999.

67. S. Park and K. Schulten, *J. Chem. Phys.*, 2004, **120**, 5946-5961.

68. M. S. Li and B. K. Mai, *Curr. Bioinform.*, 2012, **7**, 342-351.

69. S. Izrailev, S. Stepaniants, M. Balsera, Y. Oono and K. Schulten, *Biophys. J.*, 1997, **72**, 1568-1581.

70. G. M. Torrie and J. P. Valleau, *J. Comput. Phys.*, 1977, **23**, 187-199.

71. M. Souaille and B. t. Roux, *Comput. Phys. Commun.*, 2001, **135**, 40-57.

72. S. Stepaniants, S. Izrailev and K. Schulten, *J. Mol. Model.*, 1997, **3**, 473-475.

73. E. L. Florin, V. T. Moy and H. E. Gaub, *Science*, 1994, **264**, 415-417.

74. S. Park, F. Khalili-Araghi, E. Tajkhorshid and K. Schulten, *J. Chem. Phys.*, 2003, **119**, 3559-3566.

75. B. Isralewitz, J. Baudry, J. Gullingsrud, D. Kosztin and K. Schulten, *J. Mol. Graphics Model.*, 2001, **19**, 13-25.

76. B. Isralewitz, M. Gao and K. Schulten, *Curr. Opin. Struc. Biol.*, 2001, **11**, 224-230.

77. B. R. Brooks, R. E. Bruccoleri, B. D. Olafson, D. J. States, S. Swaminathan and M. Karplus, *J. Comput. Chem.*, 1983, **4**, 187-217.

78. A. D. MacKerell, D. Bashford, M. Bellott, R. L. Dunbrack, J. D. Evanseck, M. J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, D. Joseph-McCarthy, L. Kuchnir, K. Kuczera, F. T. K. Lau, C. Mattos, S. Michnick, T. Ngo, D. T. Nguyen, B. Prodhom, W. E. Reiher, B. Roux, M. Schlenkrich, J. C. Smith, R. Stote, J. Straub, M. Watanabe, J. Wiorkiewicz-Kuczera, D. Yin and M. Karplus, *J. Phys. Chem. B*, 1998, **102**, 3586-3616.

79. N. Foloppe and J. A. D. MacKerell, *J. Comput. Chem.*, 2000, **21**, 86-104.

80. M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery Jr., J. E. Peralta, F. Ogliaro, M. J. Bearpark, J. Heyd, E. N. Brothers, K. N. Kudin, V. N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. P. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, N. J. Millam, M. Klene, J. E. Knox, J. B.

Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, Ö. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski and D. J. Fox, *Gaussian 09*, (2009) Gaussian, Inc., Wallingford, CT, USA.

81. W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey and M. L. Klein, *J. Chem. Phys.*, 1983, **79**, 926-935.
82. U. Essmann, L. Perera, M. L. Berkowitz, T. Darden, H. Lee and L. G. Pedersen, *J. Chem. Phys.*, 1995, **103**, 8577-8593.
83. J. P. Ryckaert, G. Ciccotti and H. J. C. Berendsen, *J. Comput. Phys.*, 1977, **23**, 327-341.
84. R. R. Gabdoulline and R. C. Wade, *Methods*, 1998, **14**, 329-341.
85. S. Kumar, D. Bouzida, R. H. Swendsen, P. A. Kollman and J. M. Rosenberg, *J. Comput. Chem.*, 1992, **13**, 1011-1021.
86. C.-C. Chang, J.-F. Chu, H.-H. Kuo, C.-C. Kang, S.-H. Lin and T.-C. Chang, *J. Lumin.*, 2006, **119–120**, 84-90.
87. R. V. Gorbachev, A. K. Geim, M. I. Katsnelson, K. S. Novoselov, T. Tudorovskiy and I. V. Grigorieva, *Nat. Phys.*, 2012, **8**, 896-901.
88. V. P. Chuprina, U. Heinemann, A. A. Nurislamov, P. Zielenkiewicz, R. E. Dickerson and W. Saenger, *Proc. Natl. Acad. Sci. U. S. A.*, 1991, **88**, 593-597.
89. S. Y. Sheu and D. Y. Yang, *J. Phys. Chem. B*, 2010, **114**, 16558-16566.
90. E. Freire, *Chem. Biol. Drug Des.*, 2009, **74**, 468-472.
91. M. K. Gilson, J. A. Given, B. L. Bush and J. A. McCammon, *Biophys. J.*, 1997, **72**, 1047-1069.
92. G. N. Parkinson, M. P. H. Lee and S. Neidle, *Nature*, 2002, **417**, 876-880.
93. J. Dai, C. Punchihewa, A. Ambrus, D. Chen, R. A. Jones and D. Yang, *Nucleic Acids Res.*, 2007, **35**, 2440-2450.
94. J. Dai, M. Carver, C. Punchihewa, R. A. Jones and D. Yang, *Nucleic Acids Res.*, 2007, **35**, 4927-4940.
95. S. Alcaro, C. Musetti, S. Distinto, M. Casatti, G. Zagotto, A. Artese, L. Parrotta, F. Moraca, G. Costa, F. Ortuso, E. Maccioni and C. Sissi, *J. Med. Chem.*, 2013, **56**, 843-855.
96. R. Jimenez, G. R. Fleming, P. V. Kumar and M. Maroncelli, *Nature*, 1994, **369**, 471-473.
97. J. B. Chaires, N. Dattagupta and D. M. Crothers, *Biochemistry*, 1985, **24**, 260-267.

