Multistep Drug Intercalation: Molecular Dynamics and Free Energy Studies of the Binding of Daunomycin to DNA

Matthieu Wilhelm,[†] Arnab Mukherjee,*,[‡] Benjamin Bouvier,[†] Krystyna Zakrzewska,[†] James T. Hynes,*,[§,‖] and Richard Lavery*,[†]

[†]Bioinformatics: Structures and Interactions, Bases Moléculaires et Structurales des Systèmes Infectieux, Univ. Lyon I/CNRS UMR 5086, IBCP, 7 Passage du Vercors, Lyon 69367, France
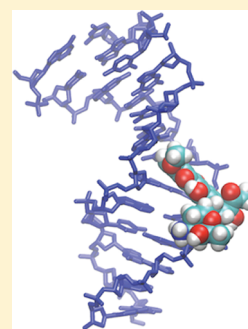
[‡]Chemistry Department, Indian Institute of Science Education and Research, Pune, 411021, India

[§]Department of Chemistry and Biochemistry, University of Colorado, Boulder, Colorado 80309-0215, United States

[‖]Chemistry Department, Ecole Normale Supérieure, CNRS UMR 8640, 24 rue Lhomond, 75005 Paris, France

S *Supporting Information*

ABSTRACT: Atomic-scale molecular dynamics and free energy calculations in explicit aqueous solvent are used to study the complex mechanism by which a molecule can intercalate between successive base pairs of the DNA double helix. We have analyzed the intercalation pathway for the anticancer drug daunomycin using two different methods: metadynamics and umbrella sampling. The resulting free energy pathways are found to be consistent with one another and point, within an equilibrium free energy context, to a three-step process. Daunomycin initially binds in the minor groove of DNA. An activated step then leads to rotation of the drug, coupled with DNA deformation that opens a wedge between the base pairs, bends DNA toward the major groove, and forms a metastable intermediate that resembles structures seen within the interfaces between DNA and minor-groove-binding proteins. Finally, crossing a small free energy barrier leads to further rotation of daunomycin and full intercalation of the drug, reestablishing stacking with the flanking base pairs and straightening the double helix.

INTRODUCTION

An important class of pharmacologically active molecules function by intercalation between the base pairs of DNA, subsequently interfering with DNA function or replication, often by blocking the action of topoisomerases, either by stabilizing the cleaved DNA produced by topoisomerase II[1] or by preventing topoisomerase I from binding to DNA.[2] Some intercalators can also act as dual topoisomerase I/II inhibitors.[3,4] Topoisomerase inhibition subsequently blocks DNA replication and leads to cell death. Intercalators are characterized by planar, conjugated, fused ring systems, whose dimensions resemble DNA base pairs. Insertion between consecutive base pairs (BPs) shields the hydrophobic rings of the intercalator from water, while favorable π-stacking is maintained within the double helix.[5,6] Despite the medical applications of intercalators as anticancer drugs,[7,8] and many years of study, a detailed mechanism of the intercalation process has not yet been determined.[9,10] As structural studies have shown, intercalation must clearly lengthen DNA, doubling the inter-BP rise at the intercalation site (from approximately 3.4 to 6.8 Å) and generally reducing twist, but the pathway (or pathways) leading to this state from the free intercalator in solution is unknown. Kinetic studies point to a complex multistep process.[11−19] Notably, for the final intercalation step, current evidence suggests that the intercalator is involved in creating the intercalation site, rather than passively waiting for such a site to occur spontaneously,[20−22] despite the fact that

significant fluctuations in DNA rise are known to occur at room temperature.[23]

Here, we use atomic-resolution molecular dynamics in aqueous solution, coupled with two different free energy evaluation procedures, to study the details of the intercalation process for daunomycin (also know as daunorubicin),[24,25] an anthracyclin anticancer drug[8] that binds to DNA and blocks its replication. The structure of DNA—daunomycin complexes has been resolved by crystallography[26−29] and by NMR spectroscopy[30,31] and shows the characteristic features of most intercalating drugs. Kinetic studies of daunomycin by Chaires et al.[16] suggested a three-step model for intercalation, involving an initial "outside" binding to DNA and a final "reshuffling" of the drug within the intercalation site. Other studies by Rizzo et al. suggested a five-step process.[19] Using detailed simulations, we hope to be able to clarify the structural and energetic features of the intercalation pathway and to identify any intermediate states, providing information that cannot easily be obtained with current experimental techniques.

This work builds on our previous umbrella-sampling studies of daunomycin binding to DNA that revealed a stable, minor-groove-bound state reached by a barrierless route from the free intercalator state.[32] This state did not require significant DNA deformation and could explain the outside bound step

postulated by Chaires and co-workers.[16] Our former study also indicated that the intercalator was involved in creating the intercalation site and enabled us to identify some key motions of the DNA−daunomycin−water system (e.g., DNA rise, roll, and bend variations) involved in the transition from the groove-bound to the intercalated state. Although this study provided an estimate of the free energy barrier, the methodology employed could not provide a detailed description of the transition. The present study fills this gap.

Since intercalation is a conformationally complex process, involving significant DNA deformation coupled with a major repositioning of the drug molecule, we have derived optimal free energy pathways with two distinct methods. The first uses nonequilibrium, metadynamics simulations[33] in a specifically designed collective variable space. Metadynamics enables a system to sample large volumes of conformational space in a reasonable computational time by regularly adding local destabilizing potentials to the natural energy hypersurface. It has already been successfully used to study a variety of binding and unbinding processes.[34−36] The second method, in common with our earlier studies,[32] uses equilibrium umbrella sampling with the aid of a biasing potential to drive the system between the desired states in regular steps. In order to study the separation of two molecules using this technique, we have recently developed a new restraint that avoids having to specify a separation pathway. This approach has already been successfully used to study the mechanism of protein−DNA unbinding.[37,38] The results of these two independent approaches are found to be consistent and reveal a passage from the groove to the intercalation site involving a metastable intermediate and two transition states. In the metastable intermediate state, daunomycin is partially intercalated at a 5′-CpG-3′ step, opening up a wedge angle between the two base pairs on the minor-groove side and bending DNA toward the major groove. This structure is similar to those seen at the interface between minor-groove-binding proteins and DNA,[39] where hydrophobic amino acid side chains are also found to partially intercalate into the double helix.

■ METHODOLOGY

Simulation Setup. Following the procedure set out in our earlier study,[32] the initial conformation of daunomycin and of its DNA intercalation site were drawn from crystallographic results,[28] using the PDB structure 1D11. Since this structure contains two daunomycin molecules, we extracted a single intercalation site at a CpG step, including one daunomycin, contained within a three nucleotide pair 5′-ACG-3′ segment. The 5′-AT base pair was retained because the charged sugar moiety of daunomycin (see Figure 1) lies in the minor
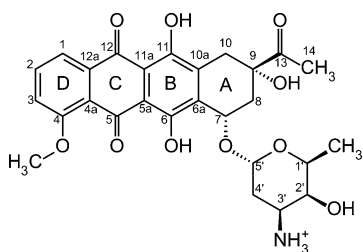

Figure 1. Chemical structure and notation of daunomycin. The rings B, C, and D constitute the planar, conjugated anthraquinone moiety that intercalates between adjacent DNA base pairs. The A-ring is nonplanar and flexible and carries the cationic daunosamine sugar at position C7.

groove, close to this base pair. The ACG segment was then extended using JUMNA,[40] by adding canonical B-DNA arms on either side to form a double-stranded dodecamer with the sequence d-(GCGCACGTGCGC)$_2$, where the underline indicates the position of the central C6pG7 intercalation site. Note that a CpG step (that is, a CG base pair followed by a GC base pair in the 5′−3′ sense) has been shown to be the strongest binding site for daunomycin.[41−43] Note also that the dodecamer has a palindromic base sequence. This implies that the orientations of the drug created by 180° rotations around the C5−C12 axis of the anthraquinone segment (see Figure 1), either in the intercalation site or centrally positioned in a DNA groove, will be symmetrically equivalent.

Simulation Protocol. DNA was simulated using the AMBER parm99/bsc0 force field[44,45] that corrects earlier problems connected with unusual backbone states and leads to conformational distributions in good agreement with experiment.[45,46] Daunomycin was represented using the general AMBER force field (GAFF).[47] Its atomic charges were calculated with the Merz−Kollman electrostatic potential potential-fitting procedure[48] in the Gaussian quantum chemistry package,[49] using a Hartee-Fock wave function obtained with a 6-31G* basis set for compatibility with the partial charges used for DNA in the AMBER force field (for further details see the Supporting Information).

The initial conformation of the daunomycin−DNA complex was simulated in a cubic box with periodic boundary conditions. The solute was surrounded by roughly 9000 TIP3P water molecules,[50] leaving at least 10 Å space between the solute and the faces of the box. The solute charges were compensated in two possible ways, either by adding a minimal number of ions to neutralize the system (22 Na$^+$ and 1 Cl$^-$) or by adding sufficient additional salt (NaCl) to reach a concentration of 0.15 M, close to physiological conditions. In both cases, ion parameters from Dang et al. were used.[51] The system was equilibrated using energy minimization followed by careful heating to 300 K with harmonic restraints on the solute atom positions. These restraints were relaxed in a series of energy minimizations, followed by short molecular dynamics simulations. The system was then equilibrated for 5 ns, with no positional restraints, and using an isothermal−isobaric ensemble. Subsequent production simulations were carried out using either the GROMACS[52] or NAMD[53] program suites. All simulations used a 2 fs time step. For the metadynamics simulations, iterative restraints were applied to all bond lengths, while for the umbrella-sampling dynamics simulations only X−H bond lengths (where X is any heavy atom) were restrained.

Metadynamics Free Energy Pathways. The first pathway generation approach uses metadynamics, which is a nonequilibrium sampling method, carried out using chosen collective variables.[33,54] In order to generate a pathway between the intercalated and groove-bound states of daunomycin, we use three collective variables: X, which describes the distance of the drug from DNA; Y, which describes its movement in the direction of the DNA helical axis; and θ, which describes the orientation of the planar anthraquinone moiety (which is required to describe the reorientation of the drug between the groove-bound and intercalated states). X and θ are the same variables used in our earlier study.[32] The geometrical definition of these variables is given in the Supporting Information. We also follow the procedure of this publication in examining the intercalation process in the reverse sense. In order to make the drug move away from the intercalation site as the simulation progresses, destabilizing Gaussian potentials are added to the region of collective variable space currently being sampled, thus forcing the system to explore higher energy regions. When convergence is reached, the Gaussian potentials allow the underlying free energy surface to be reconstructed. Metadynamics was carried out with GROMACS using the GROMETA implementation.[55] Gaussian potentials were added every 0.2 ps with a height of 0.024 kcal mol^{-1} and widths of 0.1 Å in the X and Y coordinates and of 0.2 radians in the θ coordinate. Several independent simulations were carried out in 1D (using the X coordinate), in 2D (using the X and θ coordinates), and in 3D (using the X, θ, and Y coordinates). We present here only the 3D simulations, that required of the order of 100 ns of sampling.

Umbrella Sampling Free Energy Pathways. The second pathway generation approach uses umbrella sampling. This is an equilibrium technique that constrains the system at a series of intermediate steps along a chosen coordinate using a biasing potential. Providing the sampling at successive steps overlaps, the impact of the biasing potential can be removed using the weighted-histogram analysis method (WHAM).[56] In order to separate daunomycin from DNA without imposing a specific spatial coordinate, we replaced the specific coordinates used in our earlier work[32] with a more general approach developed for studying macromolecular dissociation processes.[37] This involves a restraint on the minimal approach distance between any pair of non-hydrogen atoms at the interface between two molecules. Any atom pair falling below the chosen distance limit leads to a quadratic penalty energy and consequently to atomic forces pushing the pair apart to the desired minimal distance. To avoid the two molecules drifting apart once they no longer interact strongly, we also impose a similar quadratic penalty on the closest atom pair at the interface, if this pair lies beyond the desired distance (see Supporting Information for further details). By increasing the restraint distance in small steps (0.15 Å) and adequately sampling the corresponding conformational fluctuations of the system (1–2 ns sampling before changing the restraint distance, followed by 2–4 ns of further simulation) we can buildup an accurate view of the free energy pathway.

We also used umbrella sampling to analyze the internal flexibility of daunomycin and, in particular, of its flexible A-ring. This was sampled using the C7–C8–C9–C10 dihedral angle (hereafter termed τ) to scan the possible conformations of this ring. The dihedral was changed in 5° steps and restrained with a force constant of 0.15 kcal mol^{-1} deg^{-1}. We performed 1 ns of sampling before each change. Each point was then sampled for a further 3 ns. The 30 sampling windows along each pathway required 120 ns of simulation. The corresponding free energy curves were again obtained using the WHAM procedure.[56]

Conformational Analysis. A variety of tools were used for conformational analysis. The overall characteristics of molecular dynamics trajectories were analyzed using structures ("snapshots") extracted at regular 1 ps intervals. These structures were superposed[57] and compared using root-mean-square difference (rmsd) calculations restricted to non-hydrogen atoms. Representative conformations at given points along the free energy pathways were chosen as the structure closest to the average structure of the snapshots sampled at this point, again using rmsd calculations. DNA structure was analyzed using the Curves+ program,[58] which was extended for the present study by including the possibility of locating a DNA-bound ligand. This enables a reference frame on the ligand to be located with respect to a reference frame on the DNA helical axis. Here, the position of daunomycin with respect to the center of the DNA helical axis is simplified to two parameters, the overall translation (L-Tra) and the overall rotation (L-Rot) between the two frames (see Supporting Information for details).

All molecular representations in this work were prepared using VMD,[59] and 2D and 3D graphics were prepared with Gnuplot, Mathematica, and MatLab.

■ RESULTS

Metadynamics Intercalation Pathway. Following a 110 ns metadynamics trajectory in the three collective variables described above (X, θ, Y) and using minimal salt conditions, we extracted a 2D free energy surface by averaging over the values along the Y coordinate. The results in Figure 2 show several local free energy minima that we have labeled IC, IM, MG1, MG2, and MG3. A dotted line indicates the lowest free energy path between these minima and indicates the associated transition states, labeled TS1–TS4. IC corresponds to the intercalated state of daunomycin, while MG1, MG2, and MG3 all have the drug lying within the minor groove. MG1 is the deepest of these minima, and we will not discuss the other groove-bound minima further, since drug movement within the
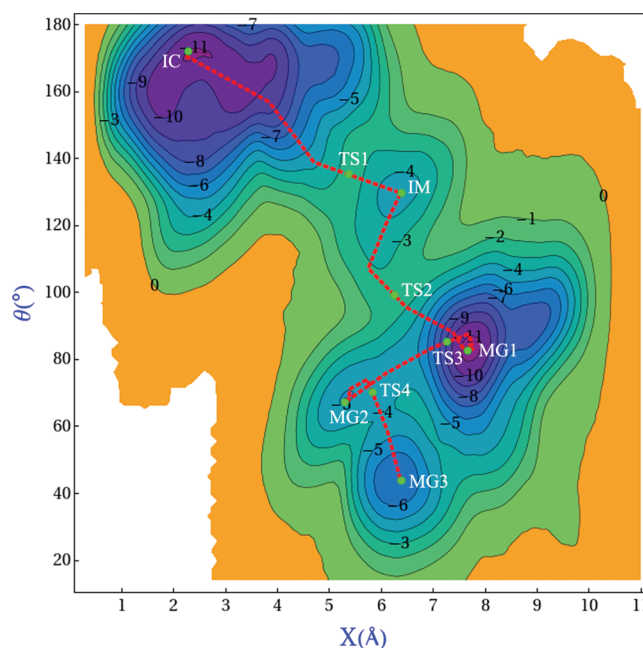

Figure 2. 2D contour plot of the metadynamics free energy trajectory obtained using three collective variables $(X, \theta, \text{and } Y)$. Values corresponding to displacements along the Y-axis (roughly aligned with the DNA helical axis) have been Boltzmann-averaged. Free energy contours are drawn at 1 kcal mol^{-1} intervals and colored from yellow to purple as the system becomes more stable. The lowest free energy path through the most important local minima and the corresponding transition states is shown.

minor groove is still under investigation and is not our focus here.

The most striking feature of Figure 2 is a well-defined free energy basin corresponding to an intermediate state IM, lying between the highest transition states. This minimum is reached after daunomycin has moved roughly 4.5 Å along X, away from the intercalated state IC, and has rotated by roughly 40° in the θ coordinate. As shown in Table 1, IM lies 7.3 kcal mol^{-1} above

Table 1. Free Energy Differences (kcal mol^{-1}) with Respect to the Intercalated State along the Metadynamics and Umbrella-Sampling Paths toward the Minor-Groove-Bound State

pathway	IC	TS1	IM	TS2	MG1
metadynamics	0	9.9	7.3	9.7	0.4
umbrella sampling	0	9.0	7.7	9.0	2.7

IC and 6.9 kcal mol^{-1} above the best minor-groove-bound state MG1. The intermediate state is flanked by two transition states, TS1, 2.6 kcal mol^{-1} above IM, leading to the intercalation site, and TS2, 2.4 kcal mol^{-1} above IM, leading to the minor-groove site.

Figure 3 (top) illustrates representative structures along the metadynamics intercalated to minor-groove-bound pathway. These were obtained as the average structures of clusters[60] lying within 2 Å rmsd of each significant point (transition state or free energy minimum) along the pathway. Looking again at Figure 2, we see that the drug initially moves mainly in the X-direction, since its anthraquinone moiety is still held between the DNA base pairs. Once this interaction weakens, the drug begins to rotate (decreasing θ) and DNA bends away from
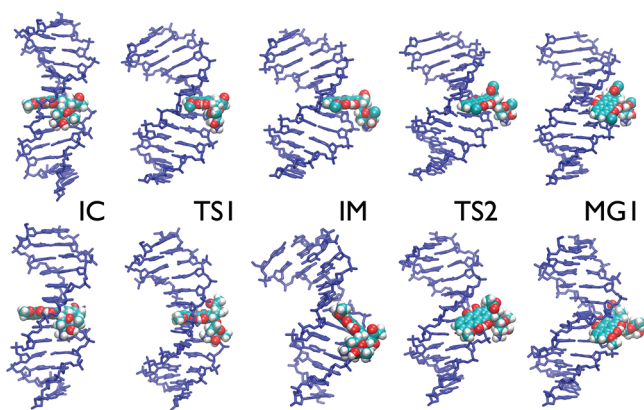
Figure 3. Representative structures leading from the intercalation site to the minor-groove-bound site for the metadynamics (above) and umbrella-sampling (below) pathways. DNA is shown as a blue Dreiding model and daunomycin as a CPK model with standard chemical coloring.

daunomycin, creating a large positive roll at the CpG step in order to regain some favorable stacking between these base pairs (see the top central image of Figure 3). This is the situation at the intermediate IM state. As the drug moves slightly further away and continues to rotate, DNA can relax by recovering full CpG stacking, while at $\theta \approx 80°$, the drug finally locates its optimal position within the minor groove. The main conformational parameters of the states along the metadynamics pathway are given in Table 2A.

Table 2. Structural Comparison of Significant Points along the Free Energy Pathways from the Intercalated State (IC), through the Intermediate (IM) State, to the Minor-Groove-Bound State (MG1)[a]

Parameters	IC	TS1	IM	TS2	MG1
(A) Metadynamics					
rise C6pG7 (Å)	7.4	7.3	6.0	5.5	3.8
roll C6pG7 (deg)	−2	28	44	42	8
$\sum$roll A5-T8 (deg)	−2	40	45	44	7
twist C6pG7 (deg)	38	25	28	29	39
$\sum$twist A5-T8 (deg)	81	74	84	84	99
minor-groove width (Å)	9.5	10.4	10.9	10.5	6.9
A-ring pucker	eq.	eq.	eq.	ax.	ax.
(B) Umbrella Sampling					
rise (deg)	7.4	7.2	4.2	3.6	3.8
roll C6pG7 (deg)	0	6	39	2	−4
$\sum$roll A5-T8 (deg)	−9	8	32	6	−4
twist C6pG7 (deg)	34	24	14	35	41
$\sum$twist A5-T8 (deg)	87	77	74	95	99
minor-groove width (Å)	7.0	9.0	11.3	6.0	4.9
A-ring pucker	eq.	strained	ax.	ax.	eq.

[a]All parameters were calculated using Curves+.[58] The minor-groove width is measured midway between the C6 and G7 base pairs.

Umbrella Sampling of Daunomycin Conformations. Since the metadynamics pathway showed that the conformation of daunomycin changes as the drug moves from the intercalation site to the groove (see Table 2A), we began the umbrella sampling studies by analyzing the internal flexibility of daunomycin and, notably, the properties of its nonplanar A-ring (see Figure 1). It has been shown experimentally that this ring favors two pucker states that can be characterized by the torsion

angle τ (C7−C8−C9−C10). $\tau \approx −60°$ places the C13 acetyl group in an axial position, while $\tau \approx +50°$ places it in an equatorial position.[61−65] The most recent experiments find that the equatorial state is favored in solution, although the axial state is observed in the presence of divalent ions.[66]

We have studied the repuckering of the A-ring in three different environments: free in aqueous solution, in the intercalated state (IC), and in the minor-groove-bound state (MG1). The corresponding free energy curves were obtained by umbrella sampling of the torsion angle τ, initially using minimal salt conditions as for the metadynamics studies described in the Metadynamics Intercalation Pathway section. As expected, they all show an axial minimum close to $\tau = −60°$ and an equatorial minimum close to $\tau = +50°$. When daunomycin is isolated (Figure 4A), these two states are almost equally stable and are separated by a barrier of ≤3 kcal mol^{-1}. In contrast, the minor-groove-bound state (Figure 4B) shows a clear preference for the axial conformation (1.5 kcal mol^{-1}), coupled with a larger barrier separating the two states (≥4 kcal mol^{-1}). Lastly, in the intercalated site, the DNA environment strongly favors an equatorial conformation (Figure 4C), in agreement with crystallographic observations,[31] and any values of τ <+20° are disfavored by at least 6 kcal mol^{-1}.

During this study, we noticed that changing the A-ring puckering in the minor-groove-bound state led to a strong coupling with the minor-groove width, which increased by roughly 3 Å as daunomycin transited to the wider axial state (Supporting Information, Figure S2A). This led us to wonder whether minimal salt conditions were exaggerating the role of interactions between DNA and the charged daunosamine sugar of the drug. We consequently repeated the free energy calculations using a 0.15 M NaCl concentration, close to physiological conditions. This indeed led to a decoupling of the minor-groove width (see Figure S2B, Supporting Information) and significant change in the free energy profile (see Figure 4B, blue curve), which now favored the equatorial conformation of the A-ring. The free energy profiles for free or intercalated daunomycin with a physiological salt concentration showed less change and did not affect the preferred A-ring pucker. However, given the impact on the minor-groove width, we decided to use physiological salt conditions for the umbrella-sampling studies of the daunomycin−DNA intercalation pathway.

Umbrella Sampling Intercalation Pathway. Umbrella sampling was used to construct a single pathway between the intercalated state and the groove-bound state from two separate pathways, one removing daunomycin from its intercalation site (the IC pathway) and the other removing daunomycin from its minor-groove-bound site (the MG1 pathway). These two pathways were found to intersect and can thus be used to build a single IC−MG1 free energy profile.

The IC pathway removed daunomycin from its intercalation site using a separation restraint acting on the rigid part of the A-ring (atoms C7−C6a−C10a−C10 in Figure 1) and the atoms of the two base pairs forming the intercalation site. This choice ensured movement toward the minor-groove side of the intercalation site, while leaving the ligand free to reorient once it was no longer stacked with the flanking bases. This restraint had a value of 3.2 Å for the IC state. Umbrella sampling was carried out over the range 3.1−12.5 Å in 0.15 Å steps with 4−8 ns of sampling at each point (more sampling was used in
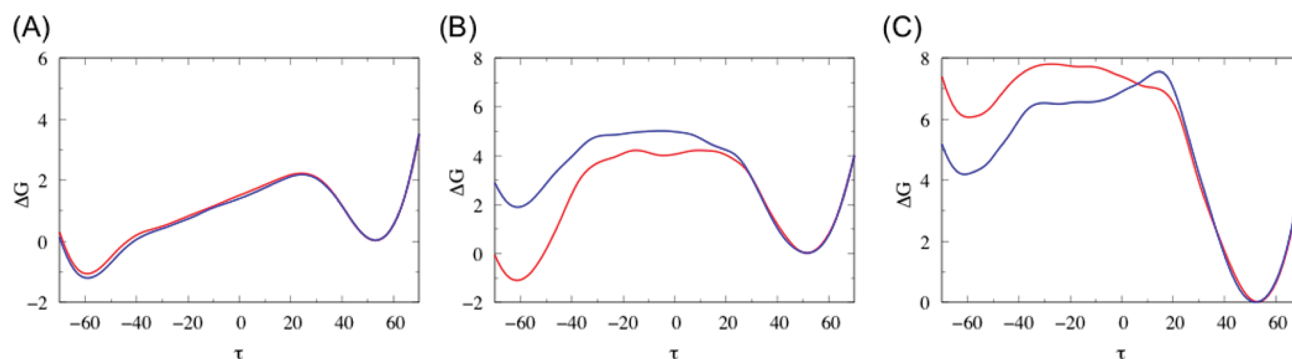
Figure 4. Free energy profiles (kcal mol^{-1}) for repuckering the A-ring of daunomycin using the dihedral τ = C7−C8−C9−C10 (deg). Local free energy minima are found at $\tau \approx 50°$ (corresponding to an equatorial position of the C13 acetyl group) and at $\tau \approx -60°$ (corresponding to an axial position of the acetyl group): (A) daunomycin in aqueous solution, (B) daunomycin bound to the minor groove of DNA, and (C) daunomycin intercalated within at the C6pG7 step of DNA. In each case, the red curve was obtained in minimal salt conditions and the blue curve in "physiological" conditions (0.15 M NaCl).
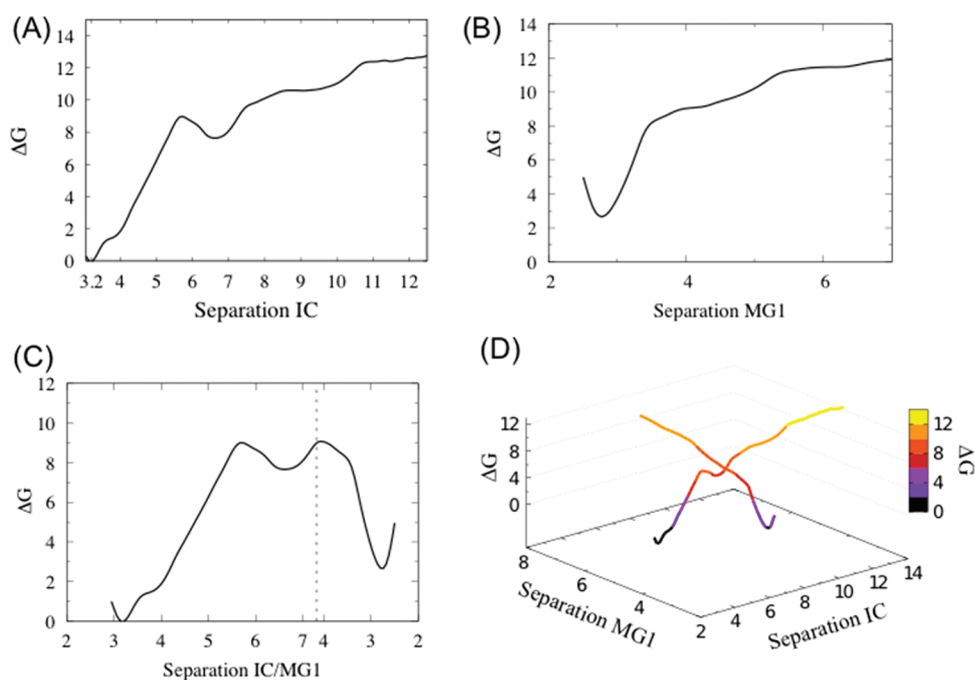

Figure 5. Umbrella-sampling free energy pathways (energies in kcal mol^{-1} and separation distances in Å): (A) separation from the intercalated state (IC), showing the intermediate state (IM); (B) separation from the minor-groove-bound state (MG1); (C) assembly of a single pathway between IC and MG1 at the point indicated by the vertical dotted line; and (D) 3D view of the intersection of the two free energy pathways.

regions where large conformational changes occurred either in DNA or in the orientation of the drug).

The resulting free energy profile (Figure 5A) shows a sharp, linear free energy rise as the anthraquinone ring system is removed from the DNA base stack. At 5.8 Å, daunomycin begins to rotate, reflected by a local free energy maximum 9 kcal mol^{-1} above the free energy of the intercalated state, followed by a metastable intermediate state at 6.8 Å, with a free energy of 7.7 kcal mol^{-1}. This intermediate lies within a quadratic free energy well that ends at roughly 7.4 Å. Further separation of the drug leads to a slower, more or less linear free energy rise. The final free energy is 13 kcal mol^{-1} above the IC state, but it is important to note that complete sampling is very difficult once the ligand moves away from DNA and is free to tumble.

A 2D rmsd plot of representative snapshots drawn from this free energy profile (Figure 6A) confirms the existence of five conformational clusters along the pathway: (a) intercalated or partially intercalated, (b) first transition state, (c) metastable intermediate, (d) second transition state, and (e) separation. Note that the intermediate state flanked by two transition states was not observed in our earlier, less detailed study of the reaction pathway.[32]

A second umbrella-sampling pathway, the MG1 pathway, was used to remove daunomycin from the minor-groove-bound state that we constructed during our earlier studies by manually placing the drug in the center of the minor groove and then running unrestrained molecular dynamics.[32] In order to remove the drug from the groove we used a separation restraint acting on all heavy atoms of the ligand and on all base atoms of DNA. The restraint had a value of 2.8 Å for the MG1 state. Umbrella sampling was carried out over the range 2.5−7.1 Å in 0.15 Å steps with 6 ns of MD sampling at each step. Figure 5B shows an approximately quadratic free energy profile around the
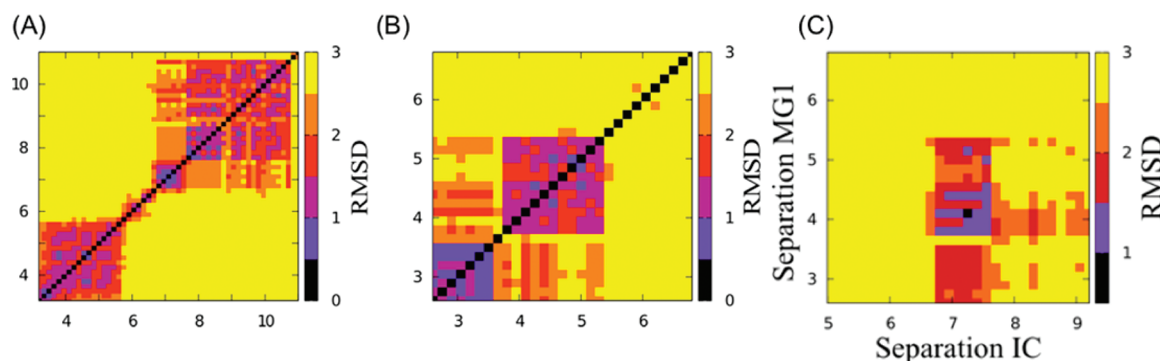
Figure 6. 2D rmsd plots (Å) between DNA−daunomycin conformations drawn from the umbrella-sampling pathways. The numbers along the axes show the increasing separation restraints, and smaller rmsd differences are shown by darker colors. Dark patches centered on the diagonal indicate similar structures that persist for a range of separation distances: (A) conformations along the pathway from the intercalated state (IC), (B) conformations along the pathway from the minor-groove-bound state (MG1), and (C) between conformations from the two pathways, showing the common structures sampled at 7.25 Å along the IC pathway and at 4.1 Å along the MG1 pathway.
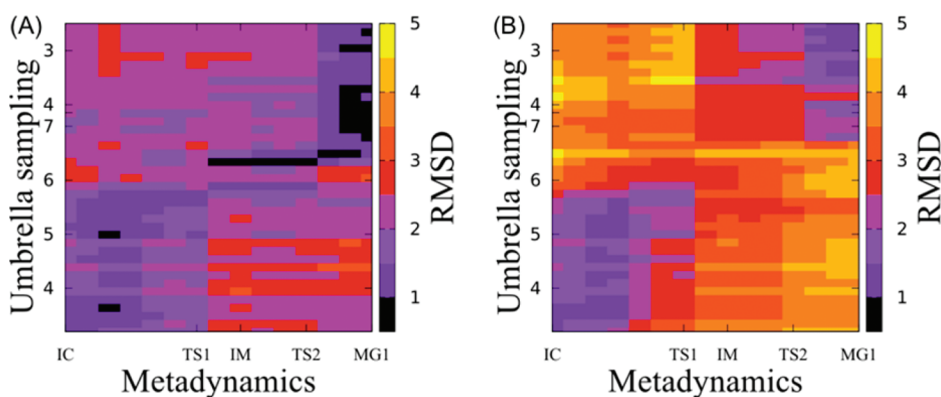

Figure 7. 2D rmsd plots (Å) between conformations drawn from the metadynamics pathway (horizontal axis) and the umbrella-sampling pathway (vertical axis): (A) for DNA alone and (B) for the DNA−daunomycin complex. See the caption to Figure 6 for further explanations.

minimum, followed by a more linear free energy rise starting from a separation of 3.5 Å, when attractive interactions between daunomycin and the sides of the minor groove are lost. The changes that occur at this point are reflected in the 2D rmsd plot shown in Figure 6B.

As shown by the 2D rmsd plot in Figure 6C, it is in fact possible to link the IC and MG1 pathways into a single IC−MG1 pathway. Indeed, we find a very close correspondence between the structures at a separation of 7.25 Å on the IC pathway and those at a separation of 4.1 Å on the MG1 pathway. We recall that these two distances differ significantly because they refer to different choices of interface atoms for the restraint.

Joining the two free energy pathways allows us to place the minor-groove minimum at a free energy 2.7 kcal mol^{-1} relative to the intercalated state (see Figures 5C,D). Representative structures along the IC−MG1 pathway, including the intermediate free energy minimum IM and the two flanking transition states, TS1 and TS2, are illustrated in Figure 3 (bottom line), and their main conformational parameters are given in Table 2B.

Comparison of Metadynamics and Umbrella Results. A first comparison of the metadynamics and umbrella sampling IC−MG1 pathways can be made using the free energy values in Table 1 and the conformational parameters in Table 2. The free energies of the significant points along both pathways are very similar to one another. The differences are almost certainly below the accuracy of the calculations, with the possible

exception of the minor-groove-bound state MG1. The umbrella-sampling pathway finds this state to be 2.7 kcal mol^{-1} less stable than the intercalated state (compared to only 0.4 kcal mol^{-1} using metadynamics). We attribute this to the use of a physiological salt concentration for the umbrella-sampling simulations, since these conditions diminish the importance of the ionic interactions between the daunosamine sugar and DNA backbone. This result also seems more in line with the fact that the minor-groove-bound state is only an intermediate preceding intercalation. The physiological salt concentration may also be responsible for slightly lowering the free energies of the transition states on either side of the IM intermediate.

The conformational parameters for the states along the two pathways are given in Table 2, and as for the free energies discussed above, they follow similar trends. The IC and MG1 states are similar (except for a narrower minor groove for the umbrella-sampling MG1 state). Some differences occur for IM and the two flanking transition states, although these are not very evident from the conformational snapshots shown in Figure 3, apart from differences in the daunomcycin orientation. For the IM state, the orientation of daunomycin varies significantly during the sampling along both pathways, notably in terms of the orientation of the drug (as shown in Figure S3, Supporting Information). Despite differences in the representative structures shown in Figure 3, similar structures are visited along both pathways (as shown in Figure S4, Supporting Information). Moreover, it should be noted that

despite movements of the drug, the DNA structures sampled are very similar with minimal and mean rmsd values of 0.3 and 1.4 Å for the nucleotides constituting the intercalation site. The differences in daunomycin orientation at IM may be linked to the change in the internal conformation of the A-ring (which is axial along the metadynamics pathway and equatorial along the umbrella sampling pathway; see Table 2) or may be linked to the different salt conditions used for the two pathways.

For a more global comparison of the metadynamics and umbrella-sampling results, we have performed two further analyses. The first uses 2D rmsd plots comparing representative conformations drawn from the two pathways. As Figure 7 shows, the agreement is good for the significant points along the pathway, both for the DNA alone (rmsd <2 Å) and for the DNA−daunomycin complex (rmsd <2.5 Å), with somewhat larger values for the IM state, where the inclination of the drug fluctuates significantly as discussed above.

The second, and more continuous, view of the pathways is given in the 3D plot in Figure 8 that combines information on
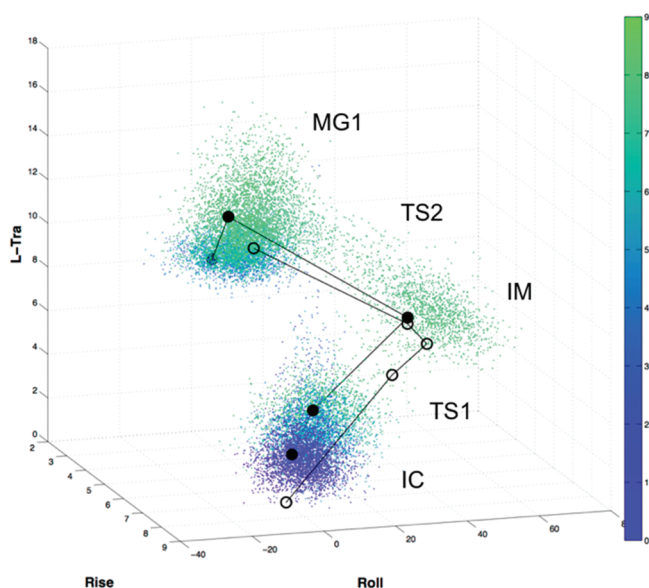

Figure 8. Pathway sampling from the minor-groove site to the intercalation site. The 3D plot shows the rise and roll of DNA at the C6pG7 intercalation step (horizontal plane) and the drug to center of DNA axis distance, L-Tra (vertical axis). Snapshots from the umbrella-sampling pathway are shown as points colored from blue to green (increasing free energy with respect to the intercalated state, see color bar in kcal mol^{-1}). The main states along the pathway are labeled IC (high rise and low roll), IM (high positive roll), and MG1 (low rise and high L-Tra), and TS1 (separating IC and IM) and TS2 (separating IM and MG1) are the less densely sampled arms joining the main clusters. Representative structures from the metadynamics and umbrella-sampling pathways are shown as open and filled circles, respectively.

DNA structure (namely, the rise and roll at the C6pG7 intercalation step), information on the ligand position (the distance from the center of the DNA helical axis), and free energy (colored blue to green as free energy increases). This scatter plot contains more than 15 000 snapshots from the umbrella-sampling IC−MG1 path which clearly visualize the clusters of points associated with the intercalated state IC (bottom), the intermediate state IM (center right), and the groove-bound state MG1 (top). The lines between the

significant points along the pathway show that the metadynamics path (open circles) and the umbrella-sampling path (filled circles) are very similar in the chosen variables. The differences in the IM state due to daunomycin orientation mentioned above do not show up here, as only DNA−drug distance is plotted. These results also show that the transition states, TS1 and TS2, occur closer to IM in the metadynamics pathway than in the umbrella-sampling pathway. However, conformational changes occur very rapidly during these transitions, as witnessed by the sparsity of sampling between the IC, IM, and MG1 regions.

■ DISCUSSION AND CONCLUSIONS

By using two different free energy sampling methods we have obtained a consensus view of how the drug daunomycin intercalates within the base stack of a DNA double helix. The optimal pathway involves initial binding to the minor groove with minimal DNA deformation (with a binding free energy of roughly −10 kcal mol^{-1} using the umbrella-sampling simulations with a physiological salt concentration). An activated process (requiring roughly 6.5 kcal mol^{-1} under the same conditions) causes the drug to rotate and to partially insert its planar anthraquinone moiety into a wedge formed between two successive base pairs. This intermediate state lies in a shallow free energy well (with a depth of roughly 1.5−2 kcal mol^{-1}). Crossing a small free energy barrier is therefore sufficient to achieve full intercalation.

The existence of an intermediate state, where the drug is partially intercalated and DNA is wedged open on the minor-groove side and bent away from the drug, is of particular interest when we consider other DNA interactions. Similar states are seen in with number of minor-groove-binding proteins including high-mobility group proteins, such as SRY[67] and HMGB1,[68] with Sac7d,[69] and also with the TATA-box binding protein.[70] In each of these cases, the protein partially intercalates a hydrophobic side chain between successive base pairs, causing a positive roll angle and significant bending toward the major groove (see Figure S5, Supporting Information). A similar state occurs in the crystal structure of a partially intercalating polypyridyl ruthenium complex with DNA[71] and, again, in very recent simulation studies of the interaction of Δ-[Ru(phenanthroline)$_3$]$^{2+}$ with DNA in aqueous solution. The interesting contrast between all these cases and the IM state is that the overall conformation of the proteins, or of the ligands, prevents full intercalation. In contrast, for daunomycin, the IM state is only a shallow local minimum on the intercalation pathway. However, in the light of this discussion, it is perhaps not so surprising that the IM state exists.

From a therapeutic point of view, it is interesting to note that when DNA kinking toward the major groove is stabilized, for example, by cis-Pt binding to successive bases in the major groove, this can enhance protein binding in the minor groove, notably by HMGB1, which can then block DNA repair and induce cell apoptosis.[68,72] This suggests that if it were possible to significantly stabilize the IM state by chemically modifying daunomycin, or related drugs, this might provide an interesting new direction for drug development.

Lastly, we consider the kinetic implications of the intercalation pathway. Following the direction in which the simulations were performed, we consider the daunomycin dissociation reaction. The simulations indicate that crossing a 9 kcal mol^{-1} barrier is necessary to reach the IM intermediate

state. A second small barrier of 1.5−2 kcal mol^{-1} separates this state from the minor-groove-bound site. Roughly 10 kcal mol^{-1} is then required to escape into solution. This result is partially compatible with the three-state model of Chaires.[16,73] If we associate the "outside-bound" state (Chaires state 1) with our minor-groove-binding state, and the final state (Chaires state 3) with the intercalated state, then the activation energies (estimated from the temperature dependence of the corresponding rate constants) for escaping from the intercalation site (14.3 kcal mol^{-1}) and then from the minor-groove site (10.7 kcal mol^{-1}) fit with our data reasonably well. The second barrier to escape our intermediate state, which does not appear in the 3-state model,[16] might contribute to the larger single activation energy seen for this step. However, we see no process related to the unimolecular step between Chaires states 2 and 3 corresponding to "a conformational rearrangement of the intercalated drug molecule or the DNA molecule" (with a high activation energy for escaping from this third state of 16.2 kcal mol^{-1}).[16] It is possible that this step corresponds to preequilibrium repositioning of the drug at another intercalation site between different base pairs.[32] In this respect, it should be stressed that, just as in our previous work,[32] the present PMF of daunomycin/DNA interactions excludes any nonequilibrium features, since the aim is to equilibrate the system at each step along the reaction pathway. This could result in suppressing the rearrangement responsible for the final step in the Chaires model and could also hide intermediate steps in the five-state model of Rizzo et al. (who notably break down the Chaires step 2 → 3 into two constituent steps). Such pathway features could be investigated using nonequilibrium simulations in the future, in which case the present work would be a valuable starting point. Further studies will also be necessary to characterize drug movements along the minor groove and to analyze the role of water and ion dynamics around the daunomycin−DNA complex.

ASSOCIATED CONTENT

Ⓢ Supporting Information

Description of metadynamics collective variables, simulation protocols and error analysis, conformational analysis of ligand−DNA complexes, daunomycin conformational fluctuations, conformational details of the metastable partially intercalated intermediate state. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author

arnab.mukherjee@iiserpune.ac.in; chynes43@gmail.com; r.lavery@ibcp.fr

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

REFERENCES

(1) Tewey, K. M.; Rowe, T. C.; Yang, L.; Halligan, B. D.; Liu, L. F. *Science* **1984**, *226*, 466−468.
(2) Crow, R. T.; Crothers, D. M. *J. Med. Chem.* **1994**, *37*, 3191−3194.
(3) Denny, W. A. *Expert Opin. Invest. Drugs* **1997**, *6*, 1845−1851.
(4) Denny, W. A.; Baguley, B. C. *Curr. Top. Med. Chem.* **2003**, *3*, 339−353.
(5) Lerman, L. S. *J. Mol. Biol.* **1961**, *3*, 18−30.
(6) Strekowski, L.; Wilson, B. *Mutat. Res.* **2007**, *623*, 3−13.
(7) Weiss, R. B. *Semin. Oncol.* **1992**, *19*, 670−686.
(8) Minotti, G.; Menna, P.; Salvatorelli, E.; Cairo, G.; Gianni, L. *Pharmacol. Rev.* **2004**, *56*, 185−229.
(9) Müller, W.; Crothers, D. M. *J. Mol. Biol.* **1968**, *35*, 251−290.
(10) Chaires, J. B.; Dattagupta, N.; Crothers, D. M. *Biochemistry* **1982**, *21*, 3933−3940.
(11) Li, H. J.; Crothers, D. M. *J. Mol. Biol.* **1969**, *39*, 461−477.
(12) Chaires, J. B. *Biochemistry* **1983**, *22*, 4204−4211.
(13) Graves, D. E.; Krugh, T. R. *Biochemistry* **1983**, *22*, 3941−3947.
(14) Forster, W.; Stutter, E. *Int. J. Biol. Macromol.* **1984**, *6*, 114−124.
(15) Wilson, W. D.; Krishnamoorthy, C. R.; Wang, Y. H.; Smith, J. C. *Biopolymers* **1985**, *24*, 1941−1961.
(16) Chaires, J. B.; Dattagupta, N.; Crothers, D. M. *Biochemistry* **1985**, *24*, 260−267.
(17) Krishnamoorthy, C. R.; Yen, S. F.; Smith, J. C.; Lown, J. W.; Wilson, W. D. *Biochemistry* **1986**, *25*, 5933−5940.
(18) Macgregor, R. B. J.; Clegg, R. M.; Jovin, T. M. *Biochemistry* **1987**, *26*, 4008−4016.
(19) Rizzo, V.; Sacchi, N.; Menozzi, M. *Biochemistry* **1989**, *28*, 274−282.
(20) Ramstein, J.; Dourlent, M.; Leng, M. *Biochem. Biophys. Res. Commun.* **1972**, *47*, 874−882.
(21) Bresloff, J. L.; Crothers, D. M. *J. Mol. Biol.* **1975**, *95*, 103−110.
(22) Corin, A. F.; Jovin, T. M. *Biochemistry* **1986**, *25*, 3995−4007.
(23) Leger, J. F.; Robert, J.; Bourdieu, L.; Chatenay, D.; Marko, J. F. *Proc. Natl. Acad. Sci. U. S. A.* **1998**, *95*, 12295−12299.
(24) DiMarco, A.; Gaetani, M.; Orezzi, P.; Scarpinato, B. M.; Silvestrini, R.; Soldati, M.; Dasdia, T.; Valentini, L. *Nature* **1964**, *201*, 706−707.
(25) Myers Jr, C. E.; Chabner, B. A. *Cancer Chemotherapy: Principles and Practice*; Lippincott: Philadelphia, PA, 1990.
(26) Tsai, C. C.; Jain, S. C.; Sobell, H. M. *Proc. Natl. Acad. Sci. U. S. A.* **1975**, *72*, 628−632.
(27) Quigley, G. J.; Wang, A. H.; Ughetto, G.; Van Der Marel, G.; Van Boom, J. H.; Rich, A. *Proc. Natl. Acad. Sci. U. S. A.* **1980**, *77*, 7204−7208.
(28) Wang, A. H. J.; Ughetto, G.; Quigley, G. J.; Rich, A. *Biochemistry* **1987**, *26*, 1152−1163.
(29) Nunn, C. M.; Van Meervelt, L.; Zhang, S.; Moore, M. H.; Kennard, O. *J. Mol. Biol.* **1991**, *222*, 167−177.
(30) Davies, D. B.; Eaton, R. J.; Baranovsky, S. F.; Veselkov, A. N. *J. Biomol. Struct. Dyn.* **2000**, *17*, 887−901.
(31) Barthwal, R.; Sharma, U.; Srivastava, N.; Jain, M.; Awasthi, P.; Kaur, M.; Barthwal, S. K.; Govil, G. *Eur. J. Med. Chem.* **2006**, *41*, 27−39.
(32) Mukherjee, A.; Lavery, R.; Bagchi, B.; Hynes, J. T. *J. Am. Chem. Soc.* **2008**, *130*, 9747−9755.
(33) Laio, A.; Parrinello, M. *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99*, 12562−12566.
(34) Gervasio, F. L.; Laio, A.; Parrinello, M. *J. Am. Chem. Soc.* **2005**, *127*, 2600−2607.
(35) Pietrucci, F.; Marinelli, F.; Carloni, P.; Laio, A. *J. Am. Chem. Soc.* **2009**, *131*, 11811−11818.
(36) Vargiu, A. V.; Ruggerone, P.; Magistrato, A.; Carloni, P. *Nucleic Acids Res.* **2008**, *36*, 5910−5921.
(37) Bouvier, B.; Lavery, R. *J. Am. Chem. Soc.* **2009**, *131*, 9864−9865.
(38) Bouvier, B.; Zakrzewska, K.; Lavery, R. *Angew. Chem., Int. Ed.* **2011**, 6516−6518.

(39) Bewley, C. A.; Gronenborn, A. M.; Clore, G. M. *Annu. Rev. Biophys. Biomol. Struct.* **1998**, *27*, 105−131.

(40) Lavery, R.; Zakrzewska, K.; Sklenar, H. *Comput. Phys. Commun.* **1995**, *91*, 135−158.

(41) Chen, K. X.; Gresh, N.; Pullman, B. *J. Biomol. Struct. Dyn.* **1985**, *3*, 445−466.

(42) Chaires, J. B.; Herrera, J. E.; Waring, M. J. *Biochemistry* **1990**, *29*, 6145−6153.

(43) Roche, C. J.; Thomson, J. A.; Crothers, D. M. *Biochemistry* **1994**, *33*, 926−935.

(44) Wang, J. M.; Cieplak, P.; Kollman, P. A. *J. Comput. Chem.* **2000**, *21*, 1049−1074.

(45) Pérez, A.; Marchán, I.; Svozil, D.; Sponer, J.; Cheatham, T. E.; Laughton, C. A.; Orozco, M. *Biophys. J.* **2007**, *92*, 3817−3829.

(46) Lavery, R.; Zakrzewska, K.; Beveridge, D.; Bishop, T. C.; Case, D. A.; Cheatham, T.; Dixit, S.; Jayaram, B.; Lankas, F.; Laughton, C.; Maddocks, J. H.; Michon, A.; Osman, R.; Orozco, M.; Perez, A.; Singh, T.; Spackova, N.; Sponer, J. *Nucleic Acids Res.* **2010**, *38*, 299−313.

(47) Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. *J. Comput. Chem.* **2004**, *25*, 1157−1174.

(48) Besler, B. H.; Merz, K. M., Jr; Kollman, P. A. *J. Comput. Chem.* **1990**, *11*, 431−439.

(49) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, J. A., Jr.; Vreven, T.; Kudin, K. N.; Burant, J. C. *Gaussian*; Gaussian, Inc.: Wallingford, CT, 2003.

(50) Jorgensen, W. L.; Jenson, C. *J. Comput. Chem.* **1998**, *19*, 1179−1186.

(51) Dang, L. X. *J. Am. Chem. Soc.* **1995**, *117*, 6954−6960.

(52) Hess, B.; Kutzner, C.; van der Spoel, D.; Lindahl, E. *J. Chem. Theory Comput* **2008**, *4*, 435−447.

(53) Bhatele, A.; Kumar, S.; Mei, C.; Phillips, J. C.; Zheng, G.; Kale, L. V. 2009, .

(54) Iannuzzi, M.; Laio, A.; Parrinello, M. *Phys. Rev. Lett.* **2003**, *90*, 238302−238306.

(55) Camilloni, C.; Provasi, D.; Tiana, G.; Broglia, R. A. *Proteins: Struct., Funct. Bioinf.* **2008**, *71*, 1647−1654.

(56) Kumar, S.; Bouzida, D.; Swendsen, R. H.; Kollman, P. A.; Rosenberg, J. M. *J. Comput. Chem.* **1992**, *13*, 1011−1021.

(57) McLachlan, A. D. *Eur. J. Biochem.* **1979**, *100*, 181−187.

(58) Lavery, R.; Moakher, M.; Maddocks, J. H.; Petkeviciute, D.; Zakrzewska, K. *Nucleic Acids Res.* **2009**, *37*, 5917−5929.

(59) Humphrey, W.; Dalke, A.; Schulten, K. *J. Mol. Graph.* **1996**, *14*, 33−38.

(60) Daura, X.; Gademann, K.; Jaun, B.; Seebach, D.; Van Gunsteren, W. F.; Mark, A. E. *Angew. Chem., Int. Ed.* **1999**, *38*, 236−240.

(61) Nuss, M. E.; James, T. L.; Apple, M. A.; Kollman, P. A. *Biochim. Biophys. Acta-Nucleic Acids Protein Synth.* **1980**, *609*, 136−147.

(62) Lown, J. W.; Chen, H. H. *Can. J. Chem.* **1981**, *59*, 3212−3217.

(63) Mondelli, R.; Ragg, E.; Fronza, G.; Arnone, A. *J. Chem. Soc., Perkin Trans. 2* **1987**, 15−26.

(64) Barthwal, R.; Mujeeb, A.; Srivastava, N.; Sharma, U. *Chem.−Biol. Interact.* **1996**, *100*, 125−139.

(65) Barthwal, R.; Srivastava, N.; Sharma, U.; Govil, G. *J. Mol. Struct.* **1994**, *327*, 201−220.

(66) Myung, J. M.; Jhon, M. S.; Kang, Y. K. *Bull. Korean Chem. Soc.* **1987**, *8*, 39−45.

(67) King, C. Y.; Weiss, M. A. *Proc. Natl. Acad. Sci. U. S. A.* **1993**, *90*, 11990−11994.

(68) Ohndorf, U. M.; Rould, M. A.; He, Q.; Pabo, C. O.; Lippard, S. J. *Nature* **1999**, *399*, 708−712.

(69) Robinson, H.; Gao, Y. G.; McCrary, B. S.; Edmondson, S. P.; Shriver, J. W.; Wang, A. H. *Nature* **1998**, *392*, 202−205.

(70) Kim, Y.; Geiger, J. H.; Hahn, S.; Sigler, P. B. *Nature* **1993**, *365*, 512−520.

(71) Hall, J. P.; O'Sullivan, K.; Naseer, A.; Smith, J. A.; Kelly, J. M.; Cardin, C. J. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108*, 17610−17614.

(72) Wang, D.; Zhu, G.; Huang, X.; Lippard, S. J. *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107*, 9584−9589.

(73) Chaires, J. B. *Biophys. Chem.* **1990**, *35*, 191−202.