Supporting Information

Intercalation and De-intercalation Pathway of Proflavine through the Minor and Major Grooves of DNA: Role of Water and Entropy

Wilbee D. Sasikala and *Arnab Mukherjee

Department of Chemistry, Indian Institute of Science Education and Research, Pune, Maharashtra, 411021, India. Fax: +91 (20) 2589 9790; Tel: +91 (20) 2590 8051; E-mail: *arnab.mukherjee@iiserpune.ac.in*

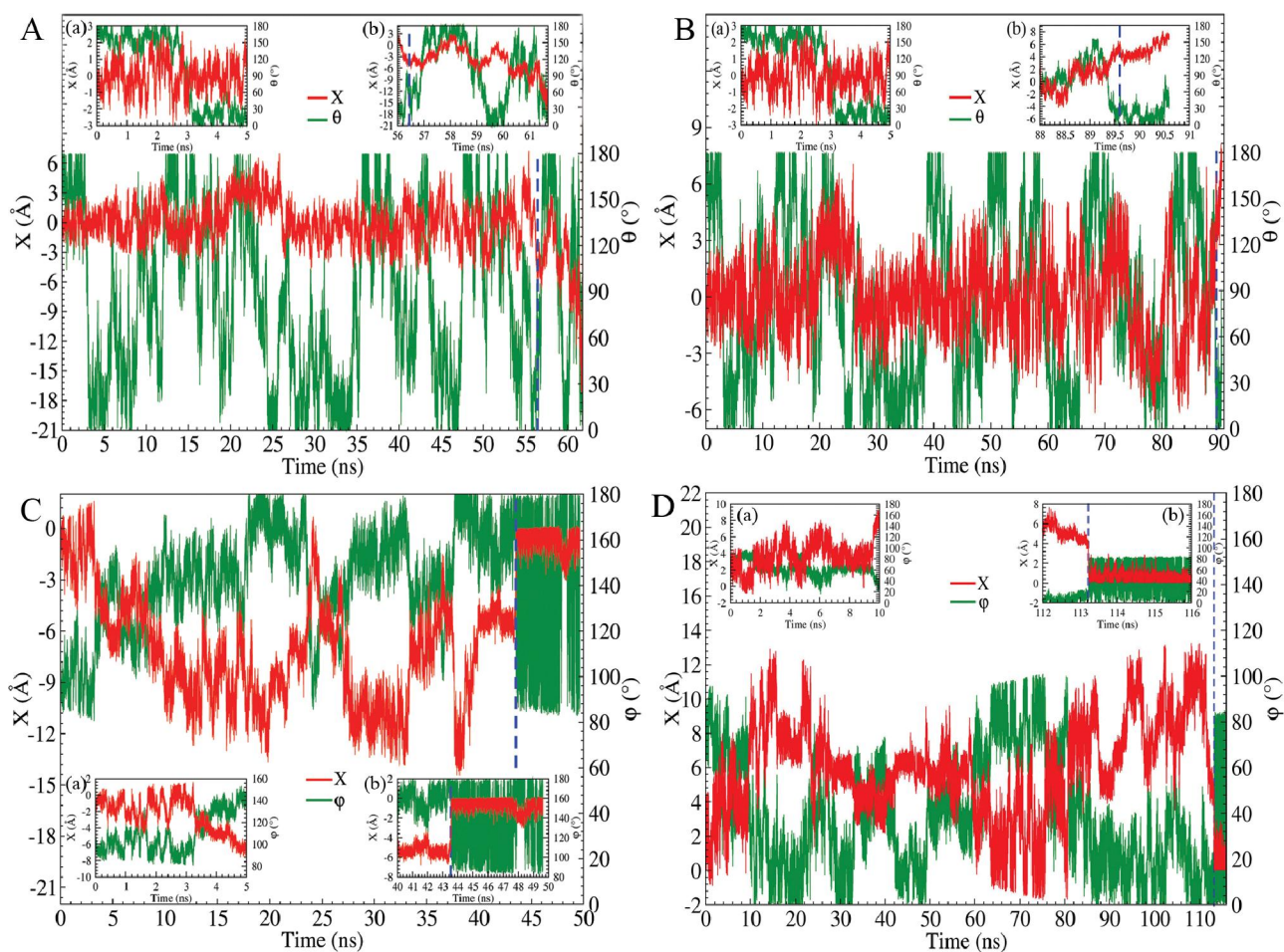## 1. Time dependence of collective variables:

Figure S1: Panel A and B shows the time dependence of the collective variables X and θ for de-intercalation through major groove side and minor groove side respectively. Panel C and D shows the time dependence of collective variables X and φ for intercalation from major groove bound state and minor groove bound state respectively. The inset figures show the magnified portion of the trajectory during the initial stage of simulation and at the transition state (TS), represented by a vertical dashed line. (A) Drug came out through the major groove side at 56.4 ns shown by the larger negative X and (B) it came out through minor groove side at 89.5ns with a larger positive X. (C) The intercalation from major groove bound state happened at 43.5 ns and (D)minor groove bound state happened at 113.2 ns.

2. **Analysis of intercalation from minor and major groove sides:** Figures S2 and S3 show the detailed analysis of intercalation from minor groove and major groove side respectively.
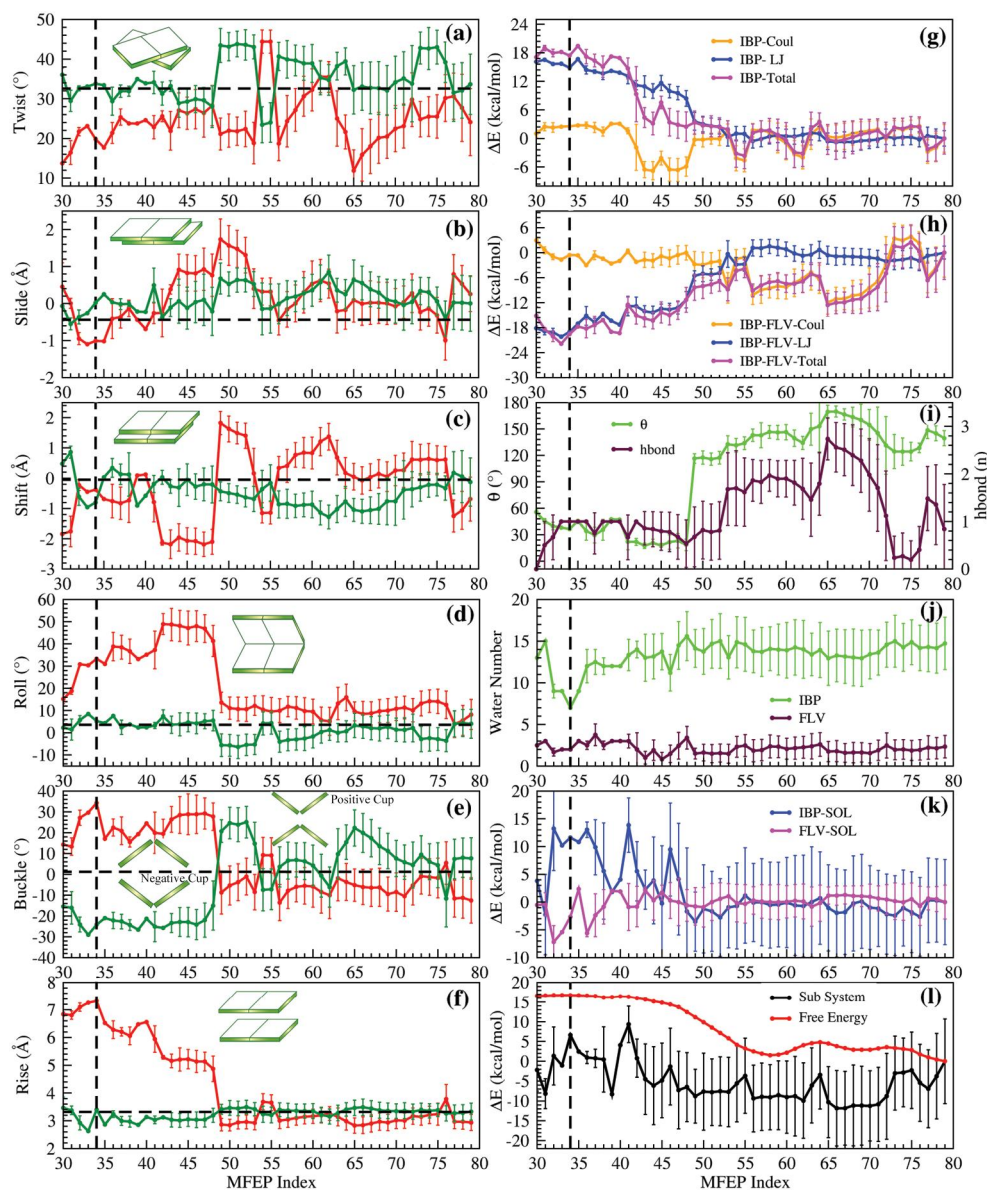

Figure S2: Analysis of intercalation process from minor groove-bound state. Each point shows the average values obtained from the members of the biggest cluster of the structures collected around discrete points along the MFEP. The standard deviation is shown. Left panel shows DNA base pair step parameters (red C6G19 and green G7C18), calculated by Curves+[1] . The vertical black dashed line represents the transition state (MFEP index 34). The horizontal dashed line represents the average value of normal B-DNA parameters[2] . Right panel shows various relative energy parameters along MFEP in (g), (h), (k), (l), which are taken to be zero at the MNS. (i) average θ value and number of hydrogen bond (hbond) between IBP and FLV, (j) number of water within 3.4 Å from heavy atoms of IBP and FLV, (h) interaction between IBP and FLV with water within 3.4 Å from heavy atoms of IBP-FLV complex, (l) Total energy of the subsystem comprising IBP, FLV and largest number of water along the entire MFEP within 3.4 Å from the heavy atoms of IBP-FLV complex. The increased stacking interaction energy at the transition state is compensated by DNA-proflavine Coulombic

interaction energy. The hydrogen bonding between proflavine and DNA is minimum at the transition state. The interaction energy between IBP and solvent increases as desolvation takes place at the transition state.
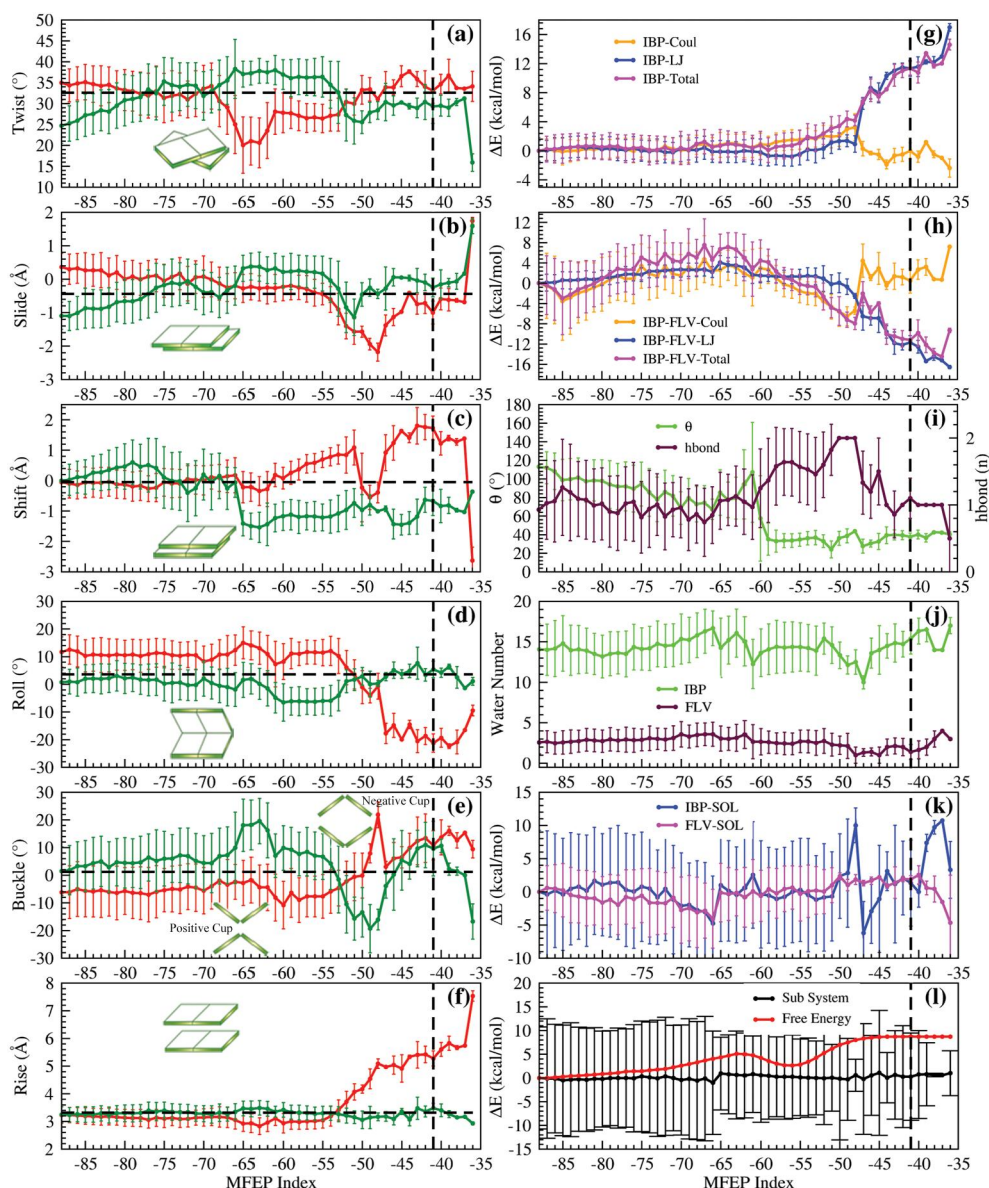

Figure S3: Analysis of intercalation process from major groove-bound state. The vertical dashed line represents TS from MJS (MFEP index -41). The energy values are taken to be zero at major groove bound state (MFEP index -88). Rest is same as Figure S2. At TS, Twist increases and higher Shift of base pair occurs. This leads to opening of the base pair towards major groove indicated by high negative Roll. Gradually the drug slides in by increasing the Rise. Increase in stacking energy (g) between base pairs is compensated by the decrease in energy between drug and base pairs (h). Low value of θ at the transition state (i) indicates that the amine groups of drug faces the bases during intercalation. Number of water molecules around major groove atoms does not change appreciably reflecting in little desolvation cost near TS at -41 and subsequently results in no overall energy cost for intercalation from major groove side.

3. Analysis of de-intercalation from major and minor groove sides: Figures S4 and S5 show the analysis of de-intercalation from major groove and minor groove sides respectively.
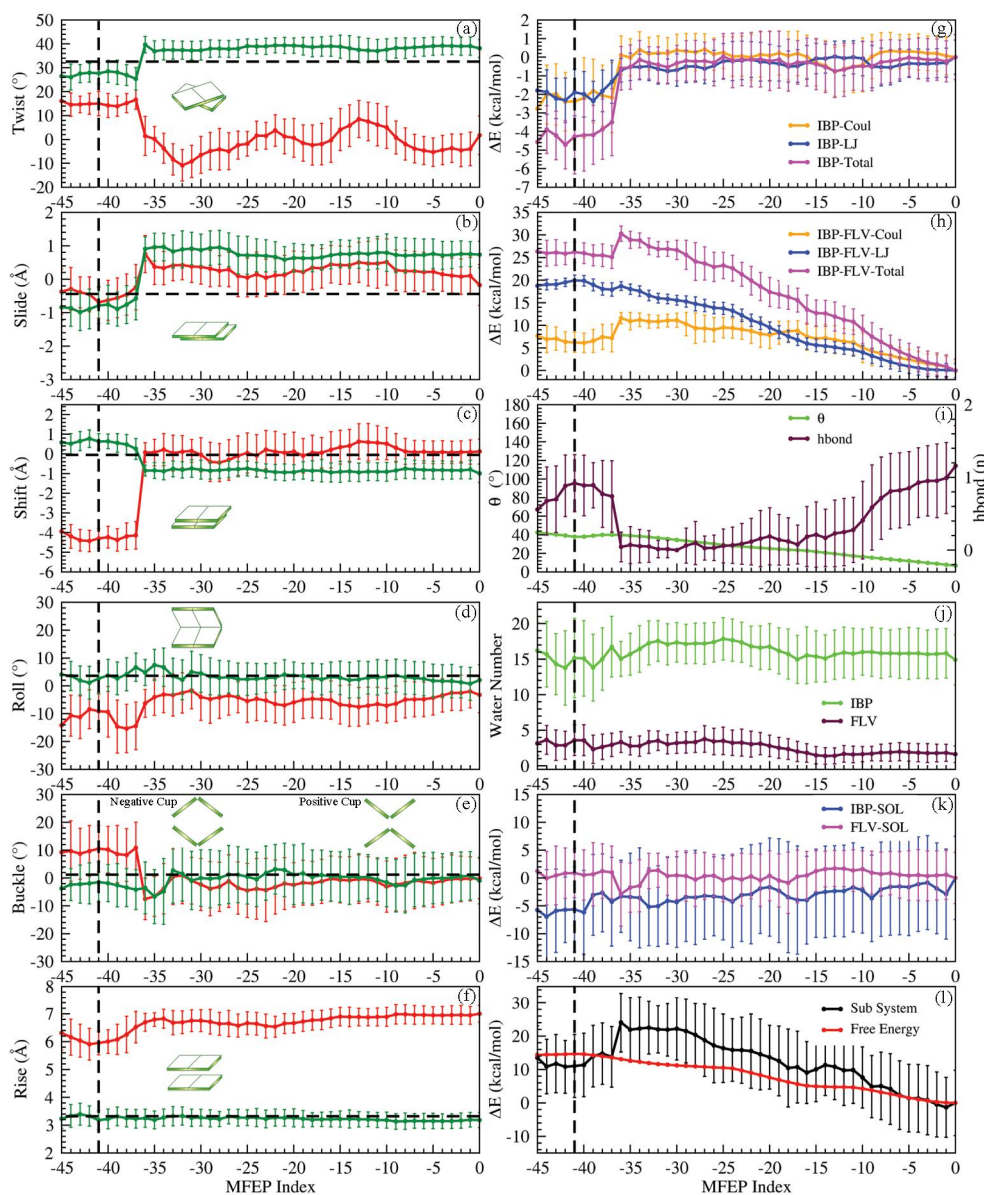

Figure S4: Analysis of de-intercalation process through major groove along MFEP. The vertical black dashed line represents the transition state (MFEP index -41). The energy values are taken to be zero at the intercalated state (MFEP index 0). Rest is same as Figure S2. At the TS, base pair Twist increases and the corresponding Shift of the bases leads to opening of the base pairs towards the major groove side represented by negative Roll. Drug's exit is reflected by the decrease in Rise. As the drug starts coming out, (g) the interaction energy between base pairs decreases and (h) the energy between drug and base pairs increases. The low value of θ at TS indicates de-intercalation happens with amino groups of the proflavine facing the bases. Main contribution of energy barrier come from the loss of stacking interaction between IBP and FLV where other interactions (IBP stacking interaction or IBP-water interaction) lead to energetic stabilization to reduce the energy barrier.
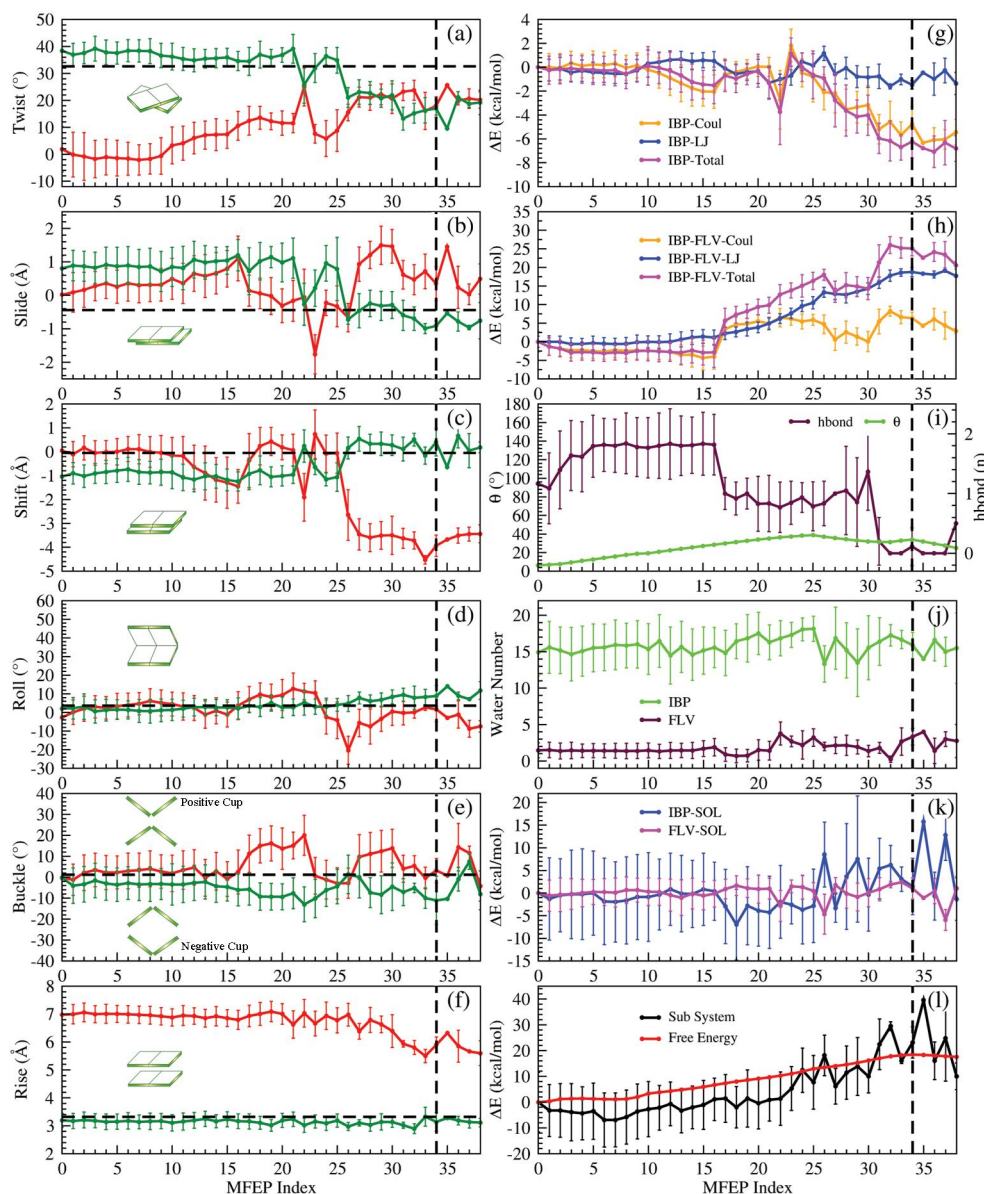
Figure S5: Analysis of de-intercalation process through minor groove along MFEP. Transition state is now at MFEP index 34. The energy values are taken to be zero at the intercalated state (MFEP index 0). Rest is same as Figure S2. DNA parameters change similar to de-intercalation towards major groove. (g) As the drug goes out, the distance between the two base pairs decreases and this leads to decrease in stacking energy where the main contribution comes from Coulombic interactions. (h) The energy between drug and base pairs increases again due to loss of stacking and decrease in hydrogen bond. (i) The low value of θ represents that the amine groups of the drug faces water during its exit from the base pairs. At TS, energy of IBP-water and drug-water is zero indicating negligible desolvation energy cost for de-intercalation through minor groove side. Therefore, the main contribution of the energy cost for de-intercalation through minor groove comes from the decrease in the interaction between the drug and DNA.

4. Analysis of dissociation from major and minor groove-bound states: Figures S6 and S7 show the analysis of dissociation of proflavine from major groove and minor groove bound states respectively.
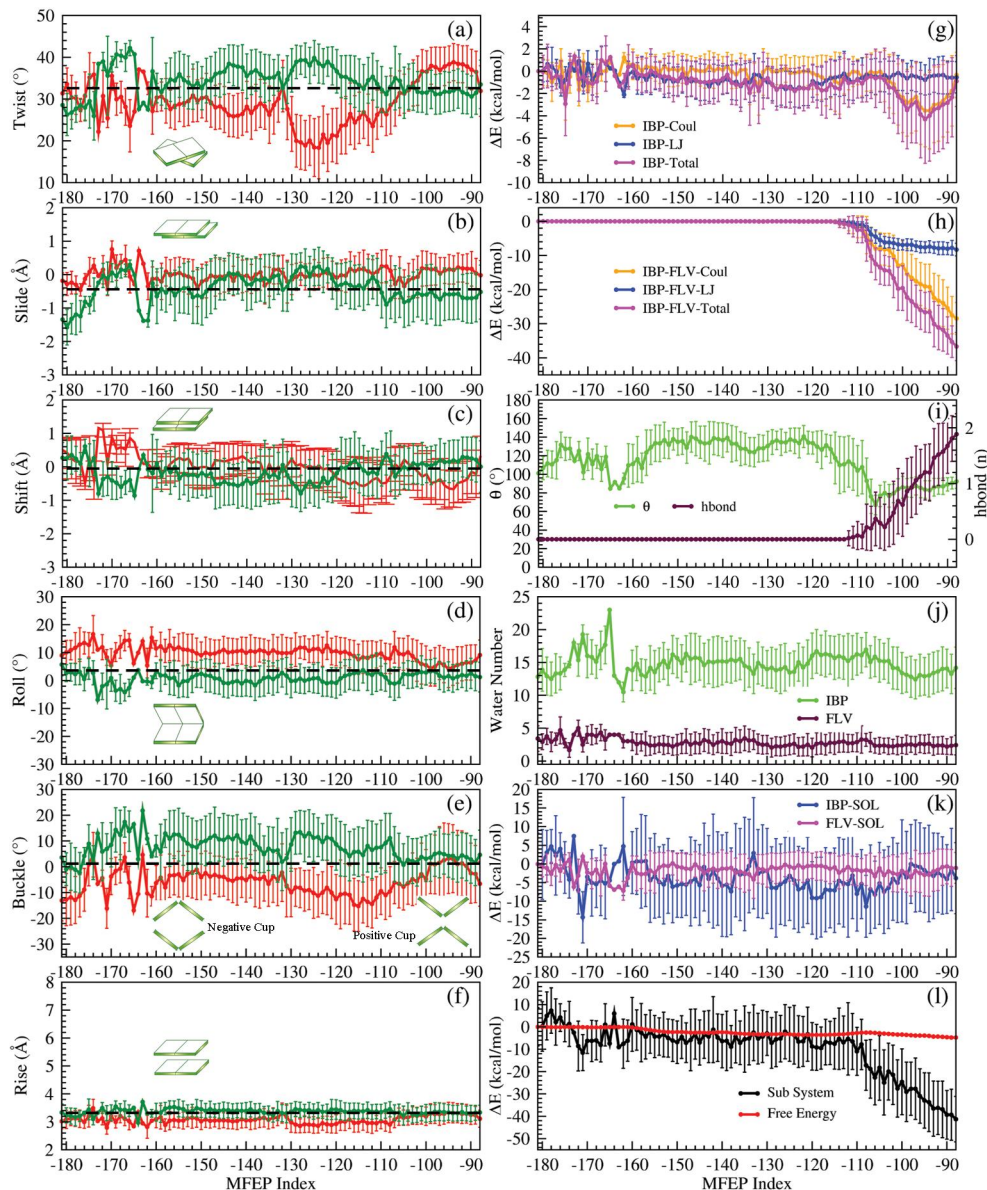

Figure S6: Analysis of dissociation process from major groove-bound state along MFEP. The energy values are taken to be zero at completely dissociated state (MFEP index -181). MFEP index -88 represents major groove bound state. Rest is same as Figure S2. During the dissociation from major groove, there is no significant change in the DNA parameters along MFEP. In the major groove bound state, the energy between drug and base pairs is very low where the main contribution comes from electrostatic interaction through hydrogen bonding. As the drug goes out, it loses hydrogen bonding, thereby increasing the energy between the drug and base pairs.
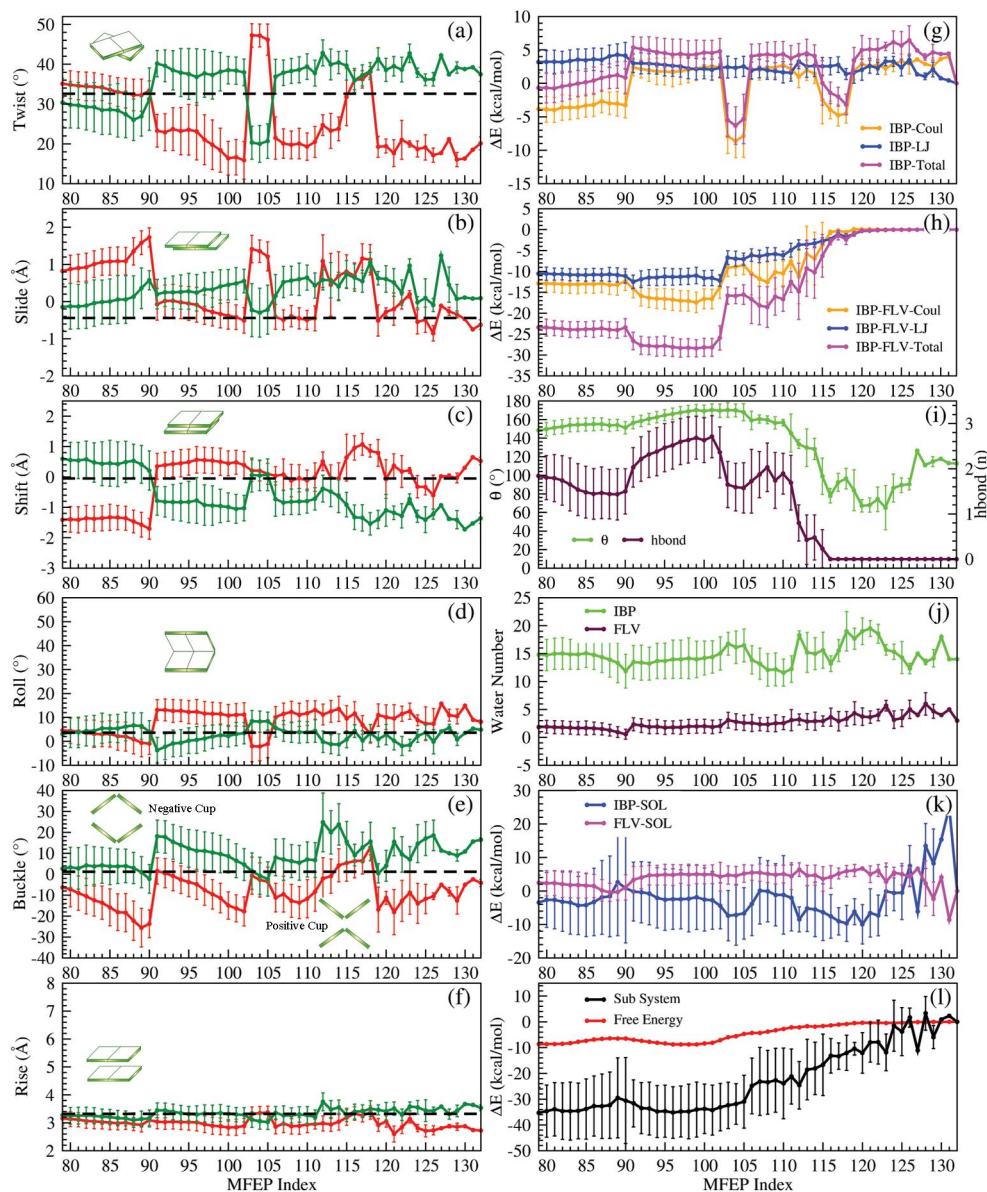
Figure S7: Analysis of dissociation process from minor groove-bound state along MFEP. The energy values are taken to be zero at completely dissociated state (MFEP index 132). MFEP index 79 represents minor groove bound state. Rest is same as Figure S2. As in the case of dissociation from major groove-bound state, the dissociation from minor groove-bound state does not affect the DNA parameters significantly. As the drug moves out of the minor groove, the energy between IBP and drug increases, the main contribution of this comes from the Coulombic energy. The increase in Coulombic energy is due to the loss of hydrogen bonding between IBP and drug, which in turn is due to the change in θ from ~160° to ~60° indicating that the amine groups now face the water.

5. **Estimation of error for intercalation from major groove side:** To find out error for intercalation from major groove side, we performed multiple well-tempered metadynamics simulations, as direct error estimation is not applicable in well-tempered metadynamics[3]. We started the simulation with different initial velocity distributions. Figure S8 shows the additional free energy surface for intercalation of proflavine from major groove side.
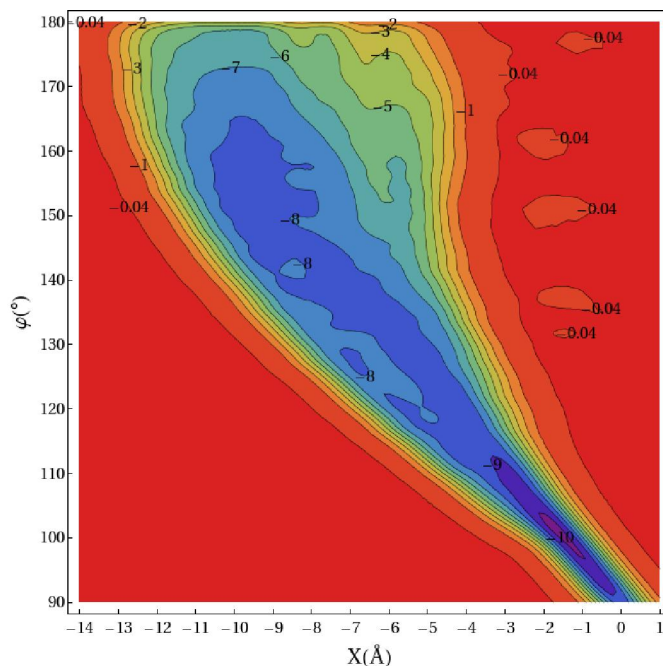

Figure S8. Free energy surface for intercalation from major groove side along X and φ. Free energy values are written on the contour lines in kcal/mol. Height of added Gaussians is 0.048 kcal/mol and width (σ) of Gaussian for X is 0.5 Å and for φ is 0.2 radians with a bias factor 15. The free energy barrier for intercalation is ~ 10 kcal/mol, which is comparable with 8.8 kcal/mol shown in the main article for intercalation barrier from major groove.

6. **Superposition of the equilibrated crystal structure of the intercalating part of DNA with the two stable intercalated states, corresponding to different orientations of proflavine, obtained from the free energy calculation.**

Here non-hydrogen atoms of the base pairs C6-G7 are considered for RMSD calculation. RMSD of stable C6-G7 base pairs where amine groups of intercalated proflavine faces towards major groove is 0.9 Å. RMSD of stable C6-G7 base pairs where amine groups of intercalated proflavine faces towards minor groove is 1.0 Å.
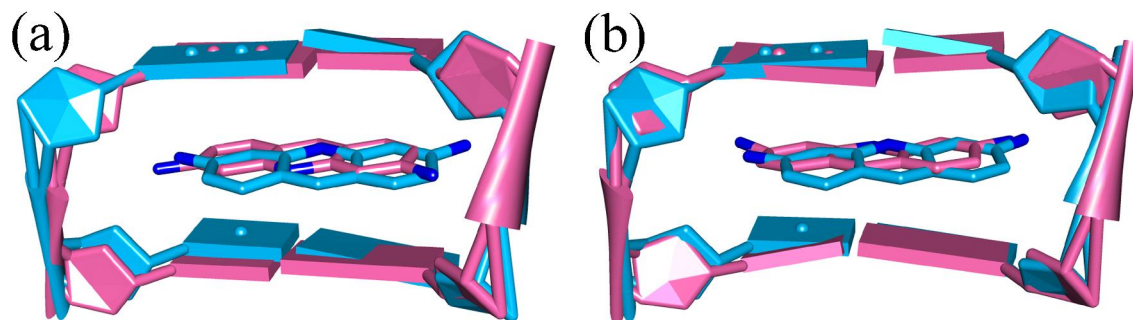
(a) (b)

Figure S9. Superposition of the equilibrated crystal structure of the intercalating part of DNA with the two stable intercalated states. The intercalated relaxed crystal structure of base pairs with proflavine is shown as deep cyan colored ribbon structure with amine groups of proflavine (deep cyan colored stick model) facing major groove side. Figure S9a shows C6-proflavine-G7 complex in pink colored ribbon structure with amine groups of proflavine (stick model with pink color) facing minor groove side which is opposite to that of crystal structure orientation. Figure S9b shows C6-proflavine-G7 complex with pink colored ribbon structure with amine groups of proflavine (pink colored stick model) facing the major groove side. The amine groups of proflavine are shown as element based stick model.

References

1. R. Lavery, M. Moakher, J. H. Maddocks, D. Petkeviciute and K. Zakrzewska, *Nucleic Acids Res*, 2009, **37**, 5917-5929.
2. R. Lavery, K. Zakrzewska, D. Beveridge, T. C. Bishop, D. A. Case, T. Cheatham, S. Dixit, B. Jayaram, F. Lankas, C. Laughton, J. H. Maddocks, A. Michon, R. Osman, M. Orozco, A. Perez, T. Singh, N. Spackova and J. Sponer, *Nucleic Acids Res*, 2010, **38**, 299-313.
3. W. D. Sasikala and A. Mukherjee, *J. Phys. Chem. B*, 2012, **116**, 12208-12212.