Biophysical Chemistry 187–188 (2014) 14–22

Contents lists available at ScienceDirect

Biophysical Chemistry

journal homepage: http://www.elsevier.com/locate/biophyschem

Dual base-flipping of cytosines in a CpG dinucleotide sequence

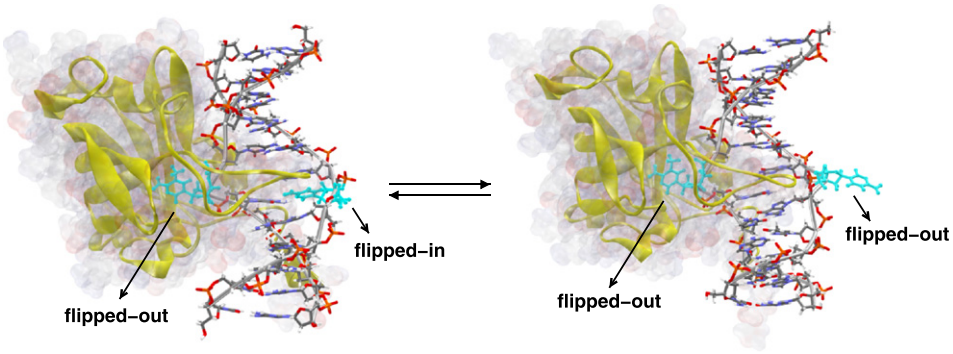
Caterina Bianchi [a], Ronen Zangi [a,b,*]

[a] Department of Organic Chemistry I, University of the Basque Country UPV/EHU, Avenida de Tolosa 72, 20018 San Sebastian, Spain
[b] IKERBASQUE, Basque Foundation for Science, 48011 Bilbao, Spain

HIGHLIGHTS

- Several X-ray structures of DNAs display dual-base flipped-out conformation.
- We investigate the penalty for dual base-flipping of cytosines in hemimethylated CpG.
- We find facilitation of the base-flipping when the other cytosine is extra-helical.
- This is demonstrated for, both, unbound and protein-bound DNAs.

GRAPHICAL ABSTRACT



ARTICLE INFO

Article history:
Received 15 November 2013
Received in revised form 18 December 2013
Accepted 18 December 2013
Available online 29 December 2013

Keywords:
DNA double-helix
Extra-helical conformation
Hemi-methylated CG site
Molecular dynamics simulation
Free energy calculation

ABSTRACT

Simultaneous flipped-out conformation of two neighboring bases on opposite strands of DNAs has been observed in several X-ray structures. It has also been detected for two cytosines on opposite strands in different contexts of CpG sites. In this paper, we study by MD simulations the dual base flipping of the two cytosines in hemi-methylated CpG site. We calculate the potential of mean force of flipping-out the unmethylated cytosine in three model systems. The first is for DNA bound to the regulatory protein UHRF1. In this case, the methyl-cytosine on the complementary strand is flipped-out into the binding pocket of the SRA domain of the protein. The other two systems are for unbound DNAs in which the methyl-cytosine is either intra-helical or extra-helical. We find that when the methyl-cytosine is flipped-out it is easier to flip-out the other (unmethylated) cytosine on the opposite strand by about 14–16 kJ/mol. This lower penalty for dual-base flipping is observed for both the bound and unbound states of the DNA. Analyses of the hydrogen bond network and stacking interactions within the CpG site indicate that the lower penalty is due to stabilization of the dual-base flipped-out conformation via interactions involving the orphan guanines. The results presented in this paper suggest that the extra-helical conformation of the methyl-cytosine recognized by UHRF1 can facilitate the base-flipping process of the target cytosine to be methylated by Dnmt1.

1. Introduction

Base flipping is a process in which a nucleotide breaks its pairing with the base on the complementary strand and rotates away from the double-helix axis. This extra-helical conformation of the base is associated with the loss of Watson–Crick hydrogen bonds and the π-stacking interactions. Thus, if base-flipping takes place in solution it is accompanied by a substantial free energy penalty [1–7]. However,

* Corresponding author at: Department of Organic Chemistry I, University of the Basque Country UPV/EHU, Avenida de Tolosa 72, 20018 San Sebastian, Spain. Tel.: +34 943018112.
E-mail address: r.zangi@ikerbasque.org (R. Zangi).

0301-4622/$ – see front matter © 2014 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.bpc.2013.12.005

several crucial processes in the cell, such as catalysis and recognition, utilize the flipped-out conformation of a specific base. To render the base-flipping a favorable event, the enzyme or the regulatory protein establishes extensive interactions with the DNA. These interactions include strong, Watson–Crick type, hydrogen bonds between the flipped-out base and the binding pocket of the protein. In some cases, the lost π-stackings are compensated by similar interactions that are established between the flipped-out base and aromatic amino-acids of the protein. Furthermore, the lost hydrogen bonds of the orphan base on the complementary strand can also be compensated by hydrogen bonds with a domain of the protein that intrudes the double helix. Examples of extra-helical conformation of a nucleotide in DNA-enzyme complexes can be found in cytosine [8] and adenine [9,10] methyltransferases, thymine-dimer [11,12] and 8-oxoguanine [13] repair enzymes, uracil-DNA glycosylase [14] and endonuclease IV [15]. Furthermore, flipped-out bases can participate in RNA splicing and ribozyme reactions [16,17], in maintaining epigenetic marks [18–21] and in signaling the termination of mitochondrial transcription [22].

Base flipping is observed not only by binding to a protein but also for DNA double-helices free in solutions. The latter have been predominantly studied by NMR imino proton exchange [23–28]. In these cases, however, the stability of the flipped-out base is very low (the life-time is on the order of nanoseconds) as evidenced by small values (around 10^{-6}–10^{-7}) of the equilibrium constant for the opening process. The life-time of the flipped-in state correlates with the strength of the interaction between the paired bases. Within the canonical B-DNA structure, the life-time of a G:C base pair (10–50 ms) is found to be about 10 times longer than that for an A:T base pair (1–5 ms). For the same reason, base-flipping is more probable for mismatched or damaged bases than for the usual base pairs [29,3,30–33] and nucleotides in RNA bulges exhibit larger propensity for base-flipping than bases that are paired [3,4,34,35]. Additional factors that influence the stability of the flipped-out conformation are chemical modifications of the bases. For example, derivatives of adenine and cytosine tethering a phenyl or naphthyl group exhibit larger propensity for extra-helical conformations relative to the unmodified nucleotides [36] whereas methylation of cytosine at position 5 suppresses base-flipping [7]. In the latter, it was found that the smaller propensity for flipping-out the modified bases is due to larger energetic penalty for the extra-helical conformation rather than stronger base-pairing in the flipped-in state. It is argued that the tendency of a base-pair to open is only moderately dependent on the sequence of the DNA [37]. However, sequences of A-tracts are known to retard the opening rate of thymine [25,38], whereas, tracts of G:C base-pairs accelerate the opening rate of guanine [39].

Base-flipping is crucial for repairing damaged bases of the DNA because the repairing enzymes perform their catalytic activity while the damaged base is in its extra-helical conformation. One type of damage bases is induced by the radiation of ultraviolet light on two adjacent pyrimidine bases. This initiates a photochemical reaction in which covalent bonds between the two bases are formed. In the repair process of this pyrimidine dimer, both bases flip-out into the active site of the repair enzyme [11,40–42] and it is often also accompanied by the base-flipping of at least one of the orphan purines on the complementary strand. This simultaneous flip-out conformation of two bases on opposite strands, which is sometimes termed 'dual base-flipping', is not only limited to pyrimidine dimers but has also been observed in the interrogation of other types of damaged bases [43].

Base-flipping is also an important process in the establishment and/or maintenance of DNA methylation. In particular, dual base-flipping has been observed in the X-ray structure of the enzyme Dnmt1 complexed to a hemimethylated DNA strand containing the sequence mCGC base-paired to GfCG (mC denotes 5-methylcytosine and fC denotes 5-fluorocytosine) [44]. In this case, two cytosines, the third and the middle on the complementary strand, adopt extra-helical conformations. As a consequence, the two orphan guanines interact with each other by forming a hydrogen bond between their purine bases (O6 of the middle guanine accepts a hydrogen from N2 of the third guanine on the complementary strand). This interaction is likely to reduce the penalty for the dual extra-helical conformation. Furthermore, in another study the protein SUVH5 has been demonstrated to bind 5mC-containing DNA duplexes in a ratio of 2:1 by a dual flip-out mechanism [45]. These DNA double-strands were characterized by one of the following contexts: hemimethylated CG, fully methylated CG, or methylated CHH (where H = A, T, or C) sites. In these bindings, both the 5mC and a base (C, 5mC, or G, respectively) from the partner strand are simultaneously extruded from the DNA duplex and positioned within the binding pockets of individual SRA (SET and RING-associated) domains of two SUVH5 units.

The protein SUVH5 is unique in the sense that it can bind methylated DNA in several different contexts. In contrast, other proteins with SRA methyl-binding domain bind very specific DNA sequences. For example, the regulatory protein UHRF1 (ubiquitin-like, containing PHD and RING finger domains 1) binds with very high specificity hemimethylated CG (from hereof, CpG) sites [46–49]. This recognition is essential to the epigenetic machinery for maintaining the DNA methylation patterns. In binding to a hemimethylated CpG site, the recognized methylcytosine flips-out of the DNA into the SRA domain of UHRF1 [18–21]. To compensate for the loss of the Watson–Crick hydrogen bonds of the partner guanine, a domain of the protein (NKR) intrudes into the DNA double-helix and forms several hydrogen bonds with this orphan base. UHRF1 binds also the N-terminal domain of Dnmt1 and, therefore, guides the latter to methylate the proper target cytosine on the complementary strand. The catalytic reaction by Dnmt1 also involves the base-flipping of the target cytosine. It is not yet known whether the binding of Dnmt1 occurs before or after URHF1 unbinds the DNA. However, if the binding of Dnmt1 occurs after UHRF1 dissociates from the DNA and the methylcytosine is flipped-in there is the possibility that Dnmt1 can slide along the DNA and lose the positioning of the target cytosine. Thus, it is reasonable to assume that at least the methylcytosine, recognized by UHRF1, is in extra-helical conformation when Dnmt1 induces the base-flipping of the target cytosine, i.e., a dual flipped-out conformation of the two cytosines of the CpG site.

In this paper, we calculate the potential of mean force (PMF) of flipping-out the target cytosine while the methylcytosine on the complementary strand is flipped-out into the binding pocket of UHRF1. The results are then compared to the PMF of flipping the same target cytosine in unbounded DNAs, when the methylcytosine on the complementary strand is either flipped-in or flipped-out. We find that the penalty for flipping-out the target cytosine is smaller when the methylcytosine on the complementary strand is also flipped-out, whether the DNA is bound or unbound to UHRF1. The reason for the smaller penalty is a stabilization of the dual-base flipped-out conformation. In the unbound case, this stabilization results from intra-DNA hydrogen bonds one of the orphan guanines forms with the backbone of the opposite strand, as well as, by stacking interactions between the two purine bases. In the bound case, the stabilization is a consequence of direct or water-bridged hydrogen bonds between Asn489 and the orphan guanine or the backbone atoms on the complementary strand to which the target cytosine is connected to. The results presented in this paper suggest that the flipped-out conformation of the methylcytosine recognized by UHRF1 can facilitate the base-flipping process of the target cytosine to be methylated by Dnmt1.

2. Methods

The sequence of the model DNA used in this study is GGGCC**mCG**CAGGG which is base-paired to the complementary strand, CCCTG**CG**GGCCC. We investigate the base-flipping of the cytosine base C7′ (numbers with the prime symbol indicate numbering of the complementary strand starting from the 3′-end) which is part of a

hemi-methylated CpG site (boldfaced letters in the sequences above). This DNA dodecamer is simulated either bound to the regulatory protein UHRF1 or free in solution. The base-flipping processes are investigated by the construction of PMFs (see below). In the DNA–UHRF1 complex, the nucleotide that is recognized, mC6, flips-out of the DNA helix into the binding pocket of the SRA domain of UHRF1, and therefore, we constructed the base-opening process of C7′ in the bound state under this condition. In the unbound state, the calculation of the PMF is performed when mC6 is in a flipped-in conformation, as well as, when it is extra-helical. In Fig. 1 we illustrate these three model systems.

The DNA sequence indicated above is taken from the crystallographic structure of the bound state (Protein Data Bank accession code: 3CLZ) where the SRA domain contains 204 amino acids [18]. From the different X-ray structures deposited we chose the model with the lowest number of missing atoms. The missing atoms, eight in number belonging to three lysine residues, were then built by the software PyMOL [50]. The side-chains of arginine and lysine were protonated whereas those of glutamate and aspartate were deprotonated. Histidine was simulated in its neutral form in which the δ-position was protonated because this tautomer exhibits a slightly larger pKa value [51]. Asp469 located at the binding pocket of the SRA domain and interacting with the extra-helical mC6 was simulated in its protonated form. This is because the distance in the X-ray structure between one of its carboxylate oxygens and N3 of methylcytosine suggests the presence of a proton either on the oxygen or on the nitrogen. Quantum mechanical calculations indicate that the proton is more likely to reside on the carboxylate group [52] and, therefore, we considered Asp469 to be protonated. The N and the C termini of the protein were taken to be protonated and deprotonated, respectively. The assignment of these protonation states to the amino acid residues yielded a net charge of $+7\,e$ for the protein. In addition, the dodecamer double-stranded DNA contributes a charge of $-22\,e$ due to the phosphate groups. All these charges were neutralized by 7 chlorides and 22 sodium cations added at random positions in the

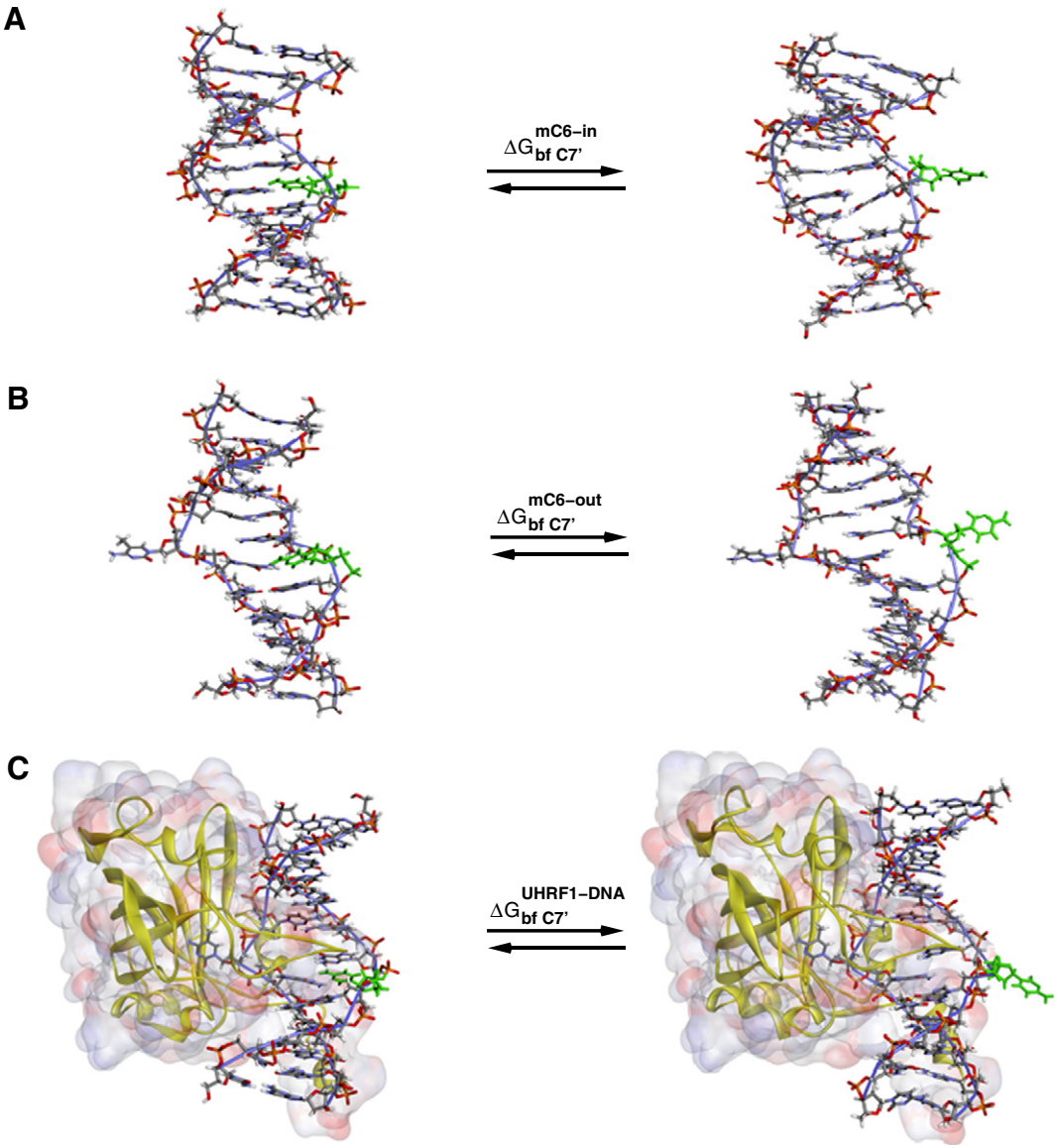


Fig. 1. Base-flipping of a cytosine base, C7′ (denoted in green), in hemi-methylated CpG sites with different conformations and bound states. In (A), the CpG site is unbound and the methycytosine on the complementary strand, mC6, is flipped-in and base-paired to G6′. In (B), the CpG site is also unbound, however, mC6 is flipped-out. In (C), the CpG site is bound to UHRF1 and the recognized base, mC6, is flipped-out into the binding pocket of the protein.

simulation box. The oxygen atoms of waters given in the X-ray structure (124 in total) were built into water molecules. The dimensions of the cubic simulation box were determined by a minimum distance of 0.8 nm between the DNA-protein complex and each of the box edges. The system was then solvated by additional waters resulting in a total of 10,510 water molecules.

The initial structure for the simulations in which the DNA is free in solution was an ideal B-DNA double helix conformation built using the PREDICTOR software [53]. In this case, the negative charge of the DNA was neutralized by 22 sodium cations and the system solvated by 8356 water molecules (which resulted from imposing the 0.8 nm minimum distance between the DNA and the box edge). A starting conformation in which mC6 is also flipped-out was taken from previous simulations in which a pseudo-dihedral angle, θ_{mC6}, is defined by the coordinates of the atoms C4(C6)–P(C6)–P(G7)–C2(G7) [7]. To maintain the extra-helical position of mC6 during the PMF calculations of flipping C7′, we applied a restraining force on θ_{mC6} in these simulations with a magnitude of 3000 kJ/(mol·nm^2).

The DNA was represented by the parmbsc0 force-field [54], the protein by the AMBER-99SB force-field [55] and the water molecules by the TIP3P model [56]. The partial charges of the 5-methyl-cytosine, which are not available in the parmbsc0 (or in AMBER-99) force-field, were taken from the work of Rauch et al. [57]. These charges were obtained from an ab-initio calculation using the Restrained ElectroStatic Potential (RESP) charge fitting procedure [58].

The molecular dynamics package GROMACS [59] version 4.5.5 was used to perform all simulations, with a time step of 0.002 ps and periodic boundary conditions applied in all three dimensions. The electrostatic forces were evaluated by the Particle-Mesh-Ewald method [60,61] with a real-space cutoff of 1.0 nm, grid spacing of 0.12 nm, and quadratic interpolation. The Lennard-Jones forces were calculated using a 1.0 nm cutoff. The simulations employed the velocity rescaling thermostat [62], with a coupling time of 0.1 ps, to maintain a constant temperature of 300 K. In addition, the Berendsen barostat [63] with a compressibility of $5 \cdot 10^{-5}$ 1/bar and a coupling time of 1.0 ps was also employed to maintain the system at a constant isotropic pressure of 1 bar. Water bond distances and angles were constrained using the SETTLE algorithm [64], whereas the distances of the protein and DNA covalent bonds were constrained using the LINCS algorithm [65].

The preparatory steps for all three systems include energy minimization using the steepest descent algorithm followed by a 2 ns simulation in which the positions of the DNA and protein heavy atoms were restrained by a harmonic potential with a force constant of 1000 kJ/(mol·nm^2). Then, a 10 ns of unrestrained simulation was performed. The configurations emerged from these simulations were used as an input for subsequent slow-growth simulations. The slow-growth simulations were performed to prepare a series of starting conformations for the PMF in which C7′ is characterized by different base-opening angles. To induce the base-flipping of C7′ we define another pseudo-dihedral angle, $\theta_{C7'}$, by the position of the atoms C4(C7′)–P(C7′)–P(G6′)–C2(G6′). This dihedral angle was changed in 10^5 steps (200 ps) from 0° to ±180° using a restrained potential of 4000 kJ/(mol·nm^2). To ensure that the base-flipping process of C7′ does not disrupt the base-pairing hydrogen bonds between the base-pairs above (C8:G8′) and below (mC6:G6′), we applied in this preparatory stage position restraints, with a force constant of 1000 kJ/(mol·nm^2) on the heavy atoms that are associated with these Watson–Crick hydrogen bonds. However, in all of the PMF calculations these position restraints were removed.

The reaction coordinate for flipping C7′ out of the double-helix is defined by the pseudo-dihedral, $\theta_{C7'}$, described above. This base-opening angle was changed from 0° to ±180° via 25 $\theta_{C7'}$ points in each direction. Negative and positive values of $\theta_{C7'}$ represent opening via the major- and minor-grooves of the DNA, respectively. For each $\theta_{C7'}$ point the starting configuration was taken from the slow-growth simulations, equilibrated for 5 ns and then data collected for additional 45 ns. Different base-opening angles correspond to different values of the coupling parameter λ and the reported values are averages at each λ-point over the data-collection segment of the trajectory. The average force needed to restrain θ to a particular value (using a force constant of 3000 kJ/(mol·nm^2)), i.e. $\langle \partial H/\partial\lambda \rangle$, is then integrated as a function of λ to obtain the PMF. At λ-points in which $\langle \partial H/\partial\lambda \rangle$ did not exhibit a smooth curve we extended the simulations up to 80 ns and/or added up to four λ-points in each direction of the base-opening. As the PMF represents only relative values, it was shifted such that the free energy of the equilibrium flipped-in state corresponds to zero.

The estimation of the errors of the free energy changes for base-flipping was obtained from [66],

$$\delta\Delta G = \left[\sum_{\lambda_{eq}}^{\lambda=1} \left(\delta\langle \partial H/\partial\lambda \rangle\right)^2\right]^{1/2}, \quad (1)$$

where $\delta\langle \partial H/\partial\lambda \rangle$ is the error in determining the average force at each λ-point, and λ_{eq} is the value of λ that corresponds to the equilibrium flipped-in state. The value of $\delta\langle \partial H/\partial\lambda \rangle$ at each λ-point was evaluated by the block averaging method [67]. To increase sampling and to allow small magnitude fluctuations, the analyses of the structural properties including the calculation of hydrogen bonds of the intra-helical and extra-helical states included the smallest four and the largest four λ-points along the reaction coordinate, respectively.

Note that because the simulations applied in this work is only at the classical mechanics level, the stacking of the DNA bases observed is not due to the interactions between the delocalized π-electrons of the aromatic rings, but a result of direct and solvent-induced (hydrophobic) interactions (electrostatic and Lennard-Jones). Nevertheless, hereafter we maintain the terminology and describe the piling of the bases on top of each other as stacking.

3. Results and discussion

The potentials of mean force of flipping-out C7′ for the three systems shown in Fig. 1 are displayed in Fig. 2. Flipping pathways via, both, the major-and minor-grooves were performed. In Table 1 we extract the penalty in the free energy when C7′ adopts an extra-helical conformation in these three systems. In principle, the penalties for these complete flipped-out conformations, $|\theta_{C7'}| = \pm 180°$, obtained via both routes

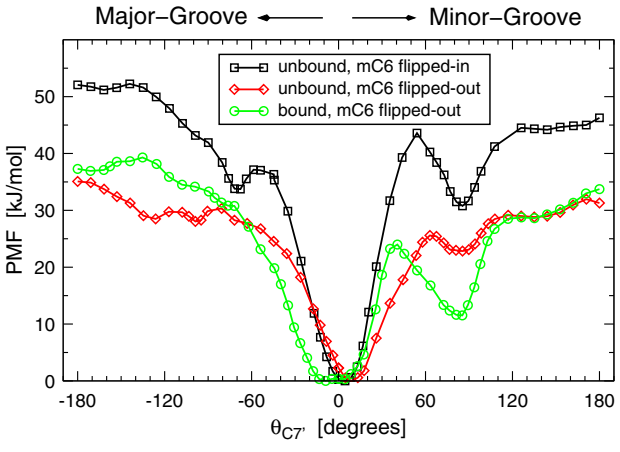


Fig. 2. The potentials of mean force of flipping-out C7′ as a function of the base-opening angle of this nucleotide under the three different setups shown in Fig. 1. The black and the red curves correspond to unbound DNAs in which mC6 is intra-helical and extra-helical, respectively. The green curve corresponds to a DNA that is bound to the protein UHRF1, a complex that is characterized by an extra-helical conformation of mC6. Positive values of the base-opening angle, $\theta_{C7'}$, correspond to flipping via the minor-groove, whereas negative values correspond to flipping via the major-groove.

Table 1
The change in the free energy of flipping-out the target cytosine C7′ under the three situations shown in Fig. 1. The results were extracted from the PMFs presented in Fig. 2 as the difference between the value at the equilibrium flipped-in conformation (around $\theta_{C7'} \approx 0^\circ$) and the value at the flipped-out conformation $|\theta_{C7'}| = \pm 180^\circ$. All values are given in kJ/mol.

	Minor-groove	Major-groove	Average
$\Delta G_{bf-C7'}^{\text{unbound;mC6-in}}$	46.3 ± 5.7	52.1 ± 3.5	49.2 ± 4.6
$\Delta G_{bf-C7'}^{\text{unbound;mC6-out}}$	31.3 ± 4.6	35.1 ± 3.5	33.2 ± 4.1
$\Delta G_{bf-C7'}^{\text{bound;mC6-out}}$	33.7 ± 4.3	37.3 ± 4.0	35.5 ± 4.2

Table 2
The average number of hydrogen bonds observed (for the unbound DNA double-helix) between the two orphan guanine nucleotides, G7 and G6′, when both of the cytosines of the CpG site, mC6 and C7′, are extra-helical. The contributions from the base, phosphate, and sugar of each nucleotide are given separately.

G7	G6′	$\langle N_{HB} \rangle$
Phosphate	Base	0.38 ± 0.07
Base	Phosphate	0.32 ± 0.06
Sugar	Base	0.26 ± 0.08
Base	Sugar	0.30 ± 0.07
Base	Base	0.06 ± 0.01
Sugar	Sugar	<0.01
Nucleotide	Nucleotide	1.32 ± 0.12

should be the same. The discrepancies we observe are with magnitudes of up to 5.8 kJ/mol and they are all within the estimated errors. It is clear that it is harder to flip-out C7′ when mC6 is intra-helical compared with the case when mC6 is extra-helical. The lower free energy change for flipping-out C7′ when mC6 is also flipped-out is true, both, when the DNA is unbound (free in solution) and when it is bound to UHRF1. The decrease in $\Delta G_{bf-C7'}$ (averaged over both pathways) is 16.0 and 13.7 kJ/mol, respectively.

For the case in which the DNAs are free in solution (unbound states), the shapes of the PMFs suggest that the free energy barriers for flipping C7′ via the major- and minor-grooves are comparable when mC6 is flipped-out, whereas, they exhibit a slight preference for opening via the major-groove when mC6 is flipped-in. Previous studies in the literature reported either no preference for base-flipping via these two routes [68,1,2] (argued to be a result of the relatively small size of the pyrimidine base) or a preference for flipping via the major-groove path [5–7]. However, when the DNA is bound to UHRF1, Fig. 2 suggests that flipping via the minor-groove path is preferred (due to smaller barriers for crossing from the equilibrium intra-helical state). The reason for this is that in the minor-groove path a local minimum is observed at around $\theta_{C7'} = 85^\circ$. This minimum is a result of an intra-strand hydrogen bond between the partially flipped-out C7′, oriented parallel to the DNA double-helix axis, and the thymine base at position 9′. More specifically, the NH_2 (i.e., N4) group of C7′ donates a hydrogen to O2 of T9′. A snapshot exhibiting this hydrogen bond is given in Fig. S1 in the Supporting Information.

This type of intra-strand hydrogen bond formed by C7′ in a parallel to the helix orientation is not unique only for the DNA-UHRF1 complex but was also observed for the unbound DNA, in both conformations of mC6, at around the same base-opening angle. In these cases, the NH_2 group of C7′ can also donate the hydrogen to other acceptors, such as the phosphate group or the amine group of G8′ and N3 of G10. For the DNA free in solution when mC6 is flipped-out the corresponding local minimum is shallow. Similar local minima were also observed in some of the major-groove paths but they are less pronounced. The stabilization of a partially-opened base via an intra-strand hydrogen bond in the minor-groove path has been seen by previous computational studies of flipping-out a cytosine base in unbound DNA [69,7].

3.1. Stabilization of the dual-base flipped-out conformation

Why is it easier to flip-out C7′ when the methyl-cytosine, mC6, is also flipped-out compared with the case when the latter is flipped-in? In principle, when C7′ adopts an extra-helical conformation the losses of the base-pairing hydrogen bonds and the stacking (hydrophobic) interactions are the same in both cases and one does not expect to find any difference in the penalties for base-flipping. The analyses we performed (see below) indicate that the reason for the lower penalties is due to excess stabilization of the DNA conformation when both cytosines are flipped-out (dual base-flipping). However, this excess stabilization is driven by different factors for the unbound and bound (to UHRF1) systems.

First we address the lower penalty for opening C7′ when mC6 is flipped-out in the unbound case. In the dual flipped-out conformation the two guanines, G7 and G6′, are no longer base-paired. Assuming that the formation of an intra-DNA hydrogen bond in aqueous solution is associated with a *net* gain in free energy, then the loss of the Watson–Crick pairing also contributes (in addition to the loss of the stacking of the bases) to the penalty of the flipped-out state. To check whether these orphan guanines interact with each other we calculate in Table 2 the average (over the data collection segment of the trajectory) number of hydrogen bonds between the two nucleotides decomposed into contributions from different parts of each nucleotide. The results indicate there is an excess of 1.3 intra-DNA hydrogen bonds which almost entirely arises from the hydrogen bonds one of the orphan guanines forms with either the phosphate or the sugar group of the second guanine on the complementary strands. In particular, the negatively charged oxygen atoms of the phosphate group or the oxygen atoms of the sugars attached to the phosphate accept a hydrogen from the NH_2 (i.e. N2) or NH (N1) groups of the guanine base on the complementary strands. Examples of these interactions are given by two instantaneous conformations of the DNA double helix in the dual flipped-out state shown in Fig. 3. The interactions in the first structure (A and C) includes a hydrogen bond between the NH_2 group of G6′ and the oxygen of phosphate of G7, as well as, stacking of the purine rings of the two orphan guanines. The ability of the two guanine bases to stack onto each other is due to the loss of their base-pairing which allows larger conformational change. In the second structure (B and D) there are two hydrogen bonds; the NH_2 and the NH groups of G7 donate hydrogens to the oxygen of G6′ and G5′ sugar groups. However, there is no interaction due to the stacking of the bases. The results shown in Table 2 indicate that G7 and G6′ nucleotides can play the role of either the acceptor or the donor. The difference in the number of hydrogen bonds in each case is likely to be insignificant. Further analysis indicates that the probability to find the simultaneous existence of two hydrogen bonds in which each of the bases of G7 and G6′ donates a hydrogen to the sugar/phosphate groups on the corresponding complementary strand is small. Note that also the number of hydrogen bonds between the bases of the two orphan guanines is very small, thus, it does not play a role in stabilizing the dual-base flipped-out conformation. In addition to the interactions listed in Table 2, we find that a small number of hydrogen bonds (0.12) is also formed between the base of G7 and the phosphate of C7′.

In the absence of stacking between the two orphan guanines (as shown in Fig. 3 panels B and D), it is typical to find two hydrogen bonds donated by one of the guanine bases to the sugar/phosphate on the complementary strand. This is reasonable because the loss of hydrophobic interaction is compensated by a gain of additional intra-DNA hydrogen bond. In order to characterize the extent of the hydrophobic stacking interactions we plot in Fig. 4 the distribution of the distance between the center of mass of the two guanine bases. The distribution for the dual-base flipped-out case displays a pronounced peak at 3.8 Å and a shoulder, with lower intensity, at around 5.0 Å. In fact, the distances of the instantaneous conformation shown in Fig. 3A

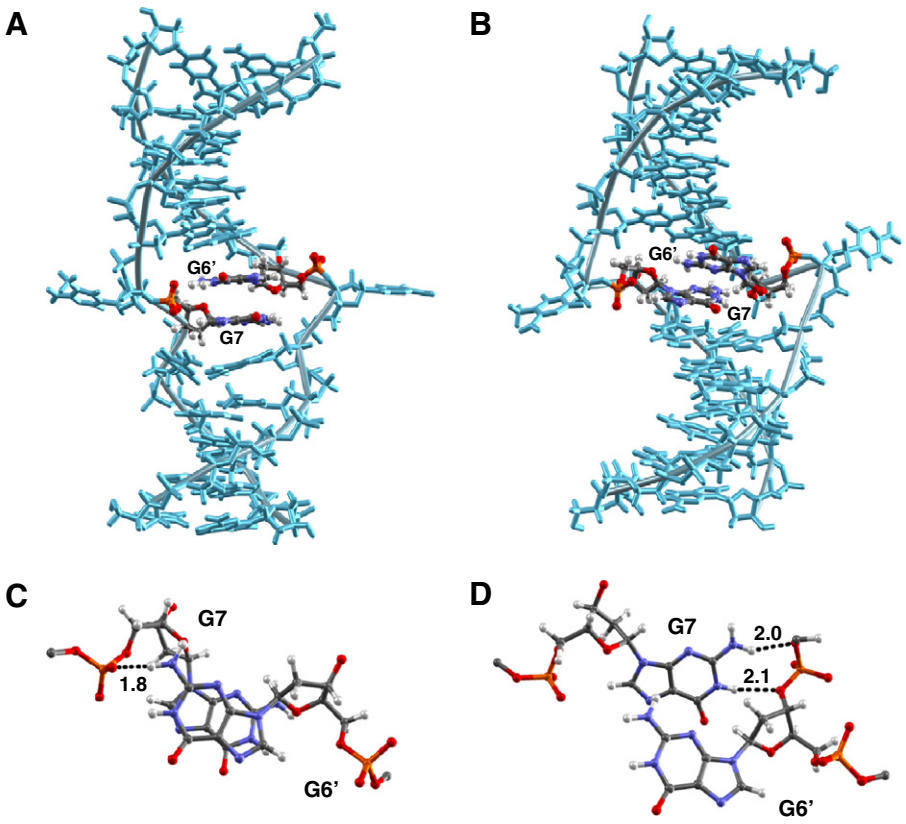


Fig. 3. Two types of interactions between the two orphan guanines, G7 and G6′, when both of their paired cytosines are extra-helical, for the simulations in which the DNA is free in solution. In (A) the interaction between the two guanines (emphasized by ball and stick representation and different colors) is driven by hydrophobic stacking of their purine rings and inter-strand hydrogen bonds between the NH_2 group of G6′ and the oxygen of phosphate of G7. In (B) the interactions include predominantly two hydrogen bonds between one of the guanines (G7) and the oxygens of the sugars on the complementary strand. The conformation shown in A is observed more often than that in B. In both cases, the DNA double-helix underwent some rearrangements, relative to its B-form conformation, to render these interactions possible. In the lower panel, (C) and (D) display top views of the hydrogen bonds arrangement and base-stacking snapshotted in (A) and (B), respectively. The numbers indicate the instantaneous hydrogen–acceptor distance in Å.

and B represent these two peaks; they equal 3.8 Å and 5.2 Å, respectively. Thus, the population of the conformation shown in Fig. 3 panels A and C, in which the two guanines are stacked onto each other, is much larger than that shown in panels B and D.

Fig. 4 also displays the distribution of the distance when both cytosines are flipped-in (as in a normal B-DNA helix). In this case, the distance between the two guanines is the largest because they are both base-paired with their complementary cytosines. The distance between G7 and G6′ gets smaller by about 0.4 Å when only one of the cytosines is flipped-out, and by about 1.2 Å when both of the cytosines are flipped-out, relative to that in a B-DNA structure.

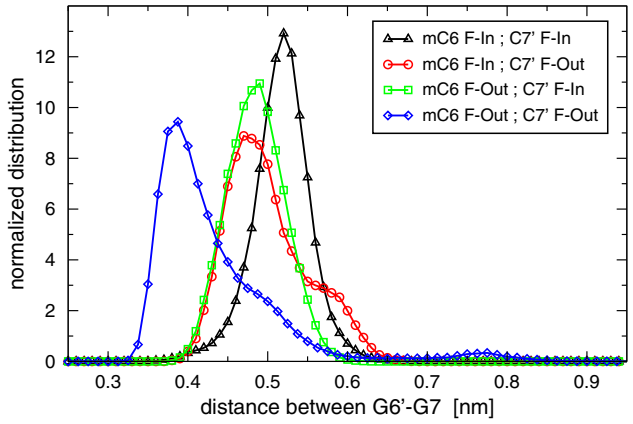


Fig. 4. The normalized distribution of the distance between the center of mass of the two guanine bases in the CpG site (G7 and G6′) for unbound double-stranded DNAs. The graph displays this distribution for all possible cases in which each of the two cytosines (mC6 and C7′) is either flipped-in (F-In) or flipped-out (F-Out).

The major peaks displayed in Fig. 4 for the two cases in which only one of the cytosines is flipped-out occur at the same distance. In contrast to the dual-base flipped-out conformation, here there are no hydrogen bonds between any of the orphan bases and the backbone of the complementary strand as well as no stacking interactions. Nevertheless, for the case in which only mC6 is extra-helical, the orphan guanine G6′ does forms (0.68) hydrogen bonds with the pyrimidine ring of a neighboring base C5, however, this is on the expense of the Watson–Crick hydrogen bonds of C5:G5′ base-pair (2.5 hydrogen bonds compared with 3.1 when all nucleotides are intra-helical). For the case in which C7′ is extra-helical, the orphan guanine G7 hardly forms any intra-DNA hydrogen bonds (0.16). Thus, in the case in which only one base is flipped-out the conformation of the DNA is likely to be too

Table 3
The change in the average number of hydrogen bonds between different groups for the base-opening process of C7′ for the three systems studied. U and B denote unbound (free in solution) and bound (to the protein UHRF1) double stranded DNA, respectively.

	U:mC6-flipped-in	U:mC6-flipped-out	B:mC6-flipped-out
Strand(A)–Strand(B)	−3.5 ± 0.4	−1.7 ± 0.8	−3.3 ± 0.5
mC6pG7–C7′pG6′	−3.2 ± 0.1	−1.5 ± 0.6	−3.1 ± 0.1
mC6pG7 + C7′pG6′–SOL	+5.1 ± 0.9	+1.8 ± 0.7	+6.3 ± 0.6
UHRF1–SOL	–	–	+1.7 ± 0.1
UHRF1–DNA	–	–	−1.0 ± 0.4
UHRF1–mC6pG7 + C7′pG6′	–	–	−1.0 ± 0.4

Table 4
The number of water molecules residing, as well as those forming bridged hydrogen bonds, between the protein (UHRF1) and the DNA. For the former, the values shown are the average number of water molecules simultaneously bound (thus, within a distance of 4 Å) to any atom in a specified group of the protein (UHRF1) and a specified group of the DNA. In the last column of each category, the change in the number of waters for the process of flipping C7′ is given. The CpG site includes the four nucleotides on both strands. The phosphate group of G6′ taken for the calculation includes also the two oxygens of the sugar groups attached to the phosphate atom.

UHRF1	DNA	Bound waters			Bridged HB		
		B:C7′-flipped-in	B:C7′-flipped-out	Δ	B:C7′-flipped-in	B:C7′-flipped-out	Δ
NKR-domain	CpG site	4.88 ± 0.07	6.81 ± 0.10	+1.93 ± 0.17	0.13 ± 0.05	0.65 ± 0.09	+0.52 ± 0.14
−C═O of Asn489	−O of G7	0.00 ± 0.00	0.94 ± 0.04	+0.94 ± 0.04	0.00 ± 0.00	0.25 ± 0.02	+0.25 ± 0.02
$-NH_2$ of Asn489	Phosphate of G6′	0.17 ± 0.10	1.17 ± 0.05	+1.00 ± 0.15	0.04 ± 0.02	0.36 ± 0.06	+0.32 ± 0.08

rigid to allow excess stabilization as for the case of a dual-base flipped-out state. Snapshots of two representative conformations in which only one of the cytosines is flipped-out are shown in Fig. S2 in the Supporting Information.

Note that when only C7′ is flipped-out, the center of mass distribution between G7 and G6′ shown in Fig. 4 (red curve) displays a shoulder at around 6.0 Å. Analysis of the hydrogen bonds around and including the orphan guanine, G7, did not reveal a different network than those corresponding to the major peak at around 4.8 Å. However, we do find that the conformation of the backbone is different in these two cases. In particular, for conformations representing the major peak at 4.8 Å the χ-dihedral angle is in the (canonical) anti state, however, the conformations representing the shoulder are characterized by χ-dihedral that is in the syn state. The distributions of this backbone dihedral are shown in Fig. S3 in the Supporting Information.

In Table 3 we calculate the change of the average number of hydrogen bonds for the process in which C7′ adopts an extra-helical conformation. When mC6 is flipped-in for the DNA free in solution, and when mC6 is flipped-out for the bound DNA, the number of hydrogen bonds that are lost between the two DNA strands is, as expected, around 3. And this readily arises from breaking the G7:C7′ base-pair. However, when mC6 is flipped-out the corresponding number is only 1.7 (and again the majority is represented by the changes of the hydrogen bonding of the CpG site) because of the (1.44) intra-DNA hydrogen bonds between the orphan guanines and the DNA backbone on the complementary strand. When a paired base flips-out, it is expected that six hydrogen bonds between the DNA and the solvent water molecules will be formed. This is exactly the case for the bound DNA, however, in both cases in which the DNA is free in solution this number is smaller. Calculating the change for each nucleotide in the CpG site separately (data not shown) indicate that when mC6 is flipped-in, the extra-helical C7′, as expected, adds three hydrogen bonds with the surrounding water molecules. However, the corresponding change for the orphan G7 is only 1.8. We conjecture this behavior to be due to the limited ability of the bulk waters to enter the DNA duplex and completely substitute the Watson–Crick hydrogen bonds. This incomplete solvation of the orphan base in the mC6 flipped-in conformation is probably also true when mC6 is flipped-out in the unbound state. Note however, that in the latter there is significant intra-DNA hydrogen bonding involving G7 and G6′ (Table 2) which reduces the number of potential sites on the bases that can hydrogen bond to the solvent.

Table 3 also displays the change in the number of hydrogen bonds between the DNA and UHRF1 for the process of flipping C7′ out. When bound to the DNA, the NKR domain of UHRF1 intrudes the double helix and makes one hydrogen bond (via Asn489) with C7′. Obviously, when C7′ flips-out this hydrogen bond is lost. However, this is partially compensated by newly formed interactions between Asn489 and the backbone atoms of C7′. In addition, the process of flipping-out C7′ also triggers a loss of 0.5 hydrogen bonds between the active site of UHRF1 and the flipped-out mC6. In total, there is a loss of one hydrogen bond between UHRF1 and DNA and the entire change comes from the interaction of the protein with the CpG site.

The weakened interaction between mC6 and UHRF1 is not obvious because the binding pocket where mC6 flips into is not in the immediate vicinity of C7′ even when it is intra-helical and paired to G7. More specifically, in two out of eight trajectories we observe that when C7′ is flipped-out, two water molecules enter the binding site and form hydrogen bonds with N3 of mC6 and the protonated carboxyl oxygen of Asp469, two groups that are hydrogen bonded to each other in the flipped-in conformation. This intrusion of water molecules into the binding site of the SRA domain is also observed in the mechanism by which UHRF1 discriminates hemi-methylated from fully-methylated CpG sites [49]. It is tempting to speculate that if the target cytosine to be methylated by Dnmt1, C7′, flips-out while UHRF1 is still bound to the DNA, this flipping process weakens the UHRF1–DNA interaction and, therefore, can render dissociation easier.

Whether or not the UHRF1–DNA complex are weakened upon the flip-out of C7′, there is a local stabilization of the dual-base flipped-out conformation of the DNA via the establishment of bridged hydrogen bonds of the CpG site with the NKR domain of the protein. When C7′ flips-out, the volume occupied by this base in its intra-helical conformation can be taken up by water molecules from the bulk. In Table 4 we calculate the number of water molecules that are simultaneously bound to the protein and the DNA. When considering the entire region of the NKR domain of UHRF1 and the CpG site of the DNA, there are about two water molecules that enter this interface. These waters substitute (at least partially) the hydrogen bonds that are lost when C7′ flips-out. One of these waters resides between the backbone carbonyl of Asn489 and the oxygen (O6) of G7 whereas the second resides between the NH_2 of Asn489 and the phosphate group of G6′. For both cases, we calculate the distance between these two groups. The profile of the former does not display a large difference between the flipped-in and flipped-out conformations of C7′, however, the latter does and its distribution is shown in Fig. 5. In the extra-helical conformation the

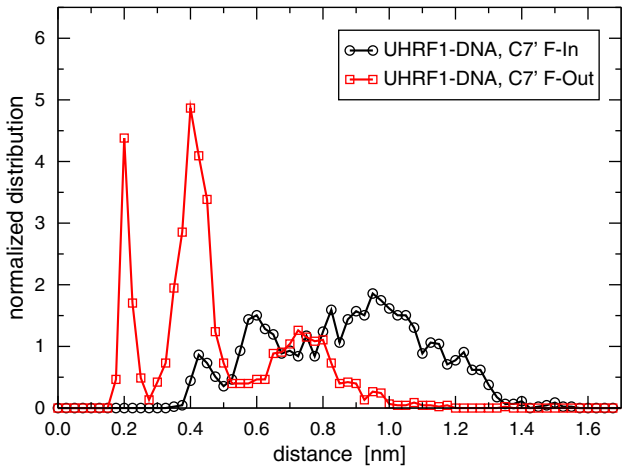


Fig. 5. The normalized distribution of the shortest distance between the oxygens of phosphate of G6′ and the $-NH_2$ group of Asn489 of the NKR domain of UHRF1 for the bound complex (thus, when mC6 is flipped-out). The black curve is calculated for a conformation in which C7′ is flipped-in whereas the red curve is for its flipped-out conformation.

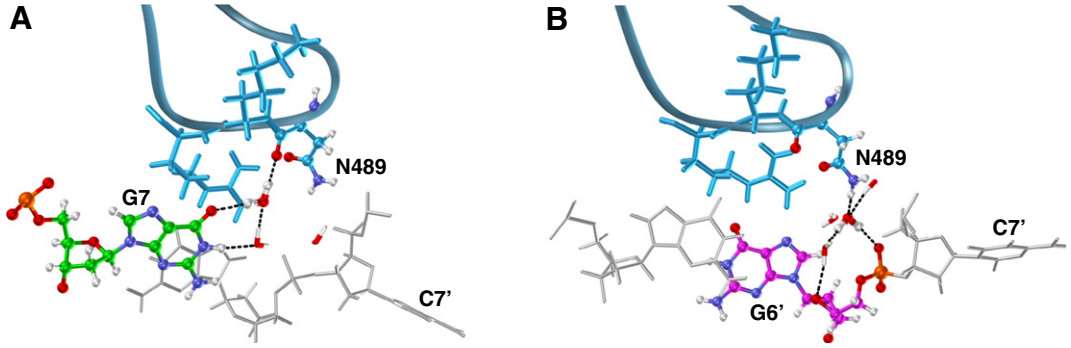


Fig. 6. Water bridged hydrogen bonding between the two orphan guanines and Asn489 of UHRF1 in the dual-base flipped-out conformation of the CpG site. In (A) the O6 of G7 (and occasionally the NH via additional water) is bridged hydrogen bonded to the backbone carbonyl of Asn489. In (B) bridged hydrogen bonds are established between the oxygen of the phosphate group of G6′ (and occasionally the O4′ of the sugar) and the NH_2 group of Asn489. In both panels, the interacting orphan guanine, Asn489, and the bridging water are emphasized by ball and stick representation. The other two residues of the NKR domain, the lysine and arginine are shown by blue stick. In addition, all waters residing within a radius of 0.4 nm from the bridging water are also shown by sticks.

profile exhibits distances that are much shorter and two distinct peaks appear at around 2 Å and 4 Å. The first is a result of direct hydrogen bonds (average number of 0.2) donated from the NH_2 group of Asn489 to the oxygens surrounding the phosphate of G6′. The second is a result of indirect hydrogen bonds, between these two groups, bridged by a water molecule that enters the space previously occupied by C7′. In Table 4 we also present the calculations of the number of these water-bridged hydrogen bonds. It indicates that between the NKR domain and the CpG site there is an increase of about 0.5 indirect hydrogen bonds when C7′ flips out and this increase arises from the two pairs of groups specified above. The combined effect of the increase in the direct and water-bridged hydrogen bonds between the protein and the DNA is likely to explain the stabilization of the flipped-out conformation of C7′ in the bound state (i.e., in dual-base flipped-out conformation) compared with the conformation in which only C7′ is flipped-out in the unbound state. In Fig. 6, we display these two types of water-bridged hydrogen bonds between Asn489 and G7 or the backbone of G6′. Note that in both cases, we also observed in few of the trajectories additional bridged hydrogen bonds via two water molecules (shown in Fig. 6).

Besides the changes described above we find that the behavior of the backbone dihedral angles also displays some changes between the flipped-in and flipped-out conformations of C7′. The most pronounced difference appears for the γ dihedral of G7 where a substantial increase in the non-canonical population is observed for the flipped-out conformation (see Fig. S4 in the Supporting Information). Nevertheless, the associated dihedral angle, α, does not exhibit a significant difference.

4. Conclusions

In this work we examined the flipping-out process of a nucleotide in an unbound DNA double helix when a neighboring base on the opposite strand adopts either intra- or extra-helical conformation. We find, via the construction of potentials of mean force, that the penalty for the base-flipping is smaller when the nucleotide on the complementary strand is also flipped-out. This smaller penalty is a result of stacking interactions established between the two orphan bases and intra-DNA hydrogen bonds one of the two orphan bases forms with the backbone atoms of the complementary strand. We examined this dual base flipping for a biologically relevant system, namely for flipping-out the two cytosines in a hemi-methylated CpG. In this case, it is known that the methyl-cytosine flips-out when it binds to the regulatory protein UHRF1. The latter recruits Dnmt1 to methylate the target unmethylated cytosine on the opposite strand, a reaction that also involves the flipping-out of this cytosine. Nevertheless, it is not known whether the flipping-out of these two cytosines are strictly sequential or whether the flipped-out conformations can also exist simultaneously. Simulations of flipping-out the target cytosine while the DNA is bound to UHRF1, with the methyl-cytosine in an extra-helical conformation, also resulted in a lower free energy penalty similar to the simulations for the unbound states. In this case, the asparagine in the NKR domain of UHRF1 forms direct and water-bridged hydrogen bonds with the orphan guanines. This may suggest that the extra-helical conformation of the methyl-cytosine can facilitate the flipping-out of the target cytosine to be methylated.

Is this reduction in the penalty for base-flipping general and sequence independent? We think that the ability of the two orphan bases to form effective stacking interactions and to reach the backbone on the complementary strand may require the size of purine bases. Thus, we conjecture that the dual base-flipping of guanines or adenines can exhibit qualitatively different behavior.

Acknowledgments

This work has been funded with support from the European Commission, Marie Curie International Reintegration Grant, project number 247485 and from the Spanish Ministry of Science and Innovation, MICINN, grant number CTQ2010-20297. Technical and human support provided by SGIker (USED SERVICES) (UPV/EHU, MICINN, GV/EJ, ESF) is gratefully acknowledged.

Appendix A. Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.bpc.2013.12.005.

References

[1] N.K. Banavali, A.D. MacKerell, Free energy and structural pathways of base flipping in a DNA GCGC containing sequence, J. Mol. Biol. 319 (2002) 141–160.
[2] E. Giudice, P. Várnai, R. Lavery, Base pair opening within B-DNA: free energy pathways for GC and AT pairs from umbrella sampling simulations, Nucleic Acids Res. 31 (2003) 1434–1443.
[3] E. Seibert, J.B.A. Ross, R. Osman, Contribution of opening and bending dynamics to specific recognition of DNA damage, J. Mol. Biol. 330 (2003) 687–703.
[4] P. Várnai, M. Canalia, J.-L. Leroy, Opening mechanism of G:T/U Pairs in DNA and RNA duplexes: A combined study of imino proton exchange and molecular dynamics simulation, J. Am. Chem. Soc. 126 (2004) 14659–14667.
[5] U.D. Priyakumar, A.D. MacKerell, NMR imino proton exchange experiments on duplex DNA primarily monitor the opening of purine bases, J. Am. Chem. Soc. 128 (2006) 678–679.
[6] G. de Marco, P. Várnai, Molecular simulation of conformational transitions in biomolecules using a combination of structure-based potential and empirical valence bond theory, Phys. Chem. Chem. Phys. 11 (2009) 10694–10700.
[7] C. Bianchi, R. Zangi, Base-flipping propensities of unmethylated, hemi-methylated, and fully-methylated CpG sites, J. Phys. Chem. B 117 (2013) 2348–2358.
[8] S. Klimasauskas, S. Kumar, R.J. Roberts, X. Cheng, Hhal methyltransferase flips its target base out of the DNA helix, Cell 76 (1994) 357–369.

[9] B. Holz, S. Klimasauskas, S. Serva, E. Weinhold, 2-Aminopurine as a fluorescent probe for DNA base flipping by methyltransferases, Nucleic Acids Res. 26 (1998) 1076–1083.
[10] E.Y. Bonnist, K. Liebertc, D.T. Drydena, A. Jeltsch, A.C. Jones, Using the fluorescence decay of 2-aminopurine to investigate conformational change in the recognition sequence of the EcoRV DNA-(adenine-N6)-methyltransferase on enzyme binding, Biophys. Chem. 160 (2012) 28–34.
[11] D.G. Vassylyev, T. Kashiwagi, Y. Mikami, M. Ariyoshi, S. Iwai, E. Ohtsuka, K. Morikawa, Atomic model of a pyrimidine dimer excision repair enzyme complexed with a DNA substrate: structural basis for damaged DNA recognition, Cell 83 (1995) 773–782.
[12] X. Zhao, J. Liu, D.S. Hsu, S. Zhao, J.-S. Taylor, A. Sancar, Reaction mechanism of (6-4) photolyase, J. Biol. Chem. 272 (1997) 32580–32590.
[13] A. Banerjee, W. Yang, M. Karplus, G.L. Verdine, Structure of a repair enzyme interrogating undamaged DNA elucidates recognition of damaged DNA, Nature 434 (2005) 612–618.
[14] G. Slupphaug, C.D. Mol, B. Kavli, A.S. Arvai, H.E. Krokan, J.A. Tainer, A nucleotide-flipping mechanism from the structure of human uracil-DNA glycosylase bound to DNA, Nature 384 (1996) 87–92.
[15] D.J. Hosfield, Y. Guan, B.J. Haas, R.P. Cunningham, J.A. Tainer, Structure of the DNA repair enzyme endonuclease IV and its DNA complex, Cell 98 (1999) 397–408.
[16] N.J. Reiter, H. Blad, F. Abildgaard, S.E. Butcher, Dynamics in the U6 RNA intramolecular stem-loop: a base flipping conformational change, Biochemistry 43 (2004) 13739–13747.
[17] A.A. Andersen, R.A. Collins, Rearrangement of a stable RNA secondary structure during VS ribozyme catalysis, Mol. Cell 5 (2000) 469–478.
[18] G.V. Avvakumov, J.R. Walker, S. Xue, Y. Li, S. Duan, C. Bronner, C.H. Arrowsmith, S. Dhe-Paganon, Structural basis for recognition of hemi-methylated DNA by the SRA domain of human UHRF1, Nature 455 (2008) 822–825.
[19] K. Arita, M. Ariyoshi, H. Tochio, Y. Nakamura, M. Shirakawa, Recognition of hemi-methylated DNA by the SRA protein UHRF1 by a base-flipping mechanism, Nature 455 (2008) 818–821.
[20] H. Hashimoto, J.R. Horton, X. Zhang, M. Bostick, S.E. Jacobsen, X. Cheng, The SRA domain of UHRF1 flips 5-methylcytosine out of the DNA helix, Nature 455 (2008) 826–829.
[21] C. Qian, S. Li, J. Jakoncic, L. Zeng, M.J. Walsh, M.-M. Zhou, Structure and hemimethylated CpG binding of the SRA domain from human UHRF1, J. Biol. Chem. 283 (2008) 34490–34494.
[22] E. Yakubovskaya, E. Mejia, J. Byrnes, E. Hambardjieva, M. Garcia-Diaz, Helix unwinding and base flipping enable human MTERF1 to terminate mitochondrial transcription, Cell 141 (2010) 982–993.
[23] C. Mandal, N.R. Kallenbach, S.W. Englander, Base-pair opening and closing reactions in the double helix: a stopped-flow hydrogen exchange study in poly(rA)·poly(rU), J. Mol. Biol. 135 (1979) 391–411.
[24] M.K. Maurice Guéron, J.-L. Leroy, A single mode of DNA base-pair opening drives imino proton exchange, Nature 328 (1987) 89–92.
[25] J.L. Leroy, M. Kochoyan, T. Huynh-Dinh, M. Guéron, Characterization of base-pair opening in deoxynucleotide duplexes using catalyzed exchange of the imino proton, J. Mol. Biol. 200 (1988) 223–238.
[26] M. Leijon, A. Gräslund, Effects of sequence and length on imino proton exchange and base pair opening kinetics in DNA oligonucleotide duplexes, Nucleic Acids Res. 20 (1992) 5339–5343.
[27] I.G. Lyakhov, P.N. Hengen, D. Rubens, T.D. Schneider, The P1 phage replication protein RepA contacts an otherwise inaccessible thymine N3 proton by DNA distortion or base flipping, Nucleic Acids Res. 29 (2001) 4892–4900.
[28] K. Snoussi, J.-L. Leroy, Imino proton exchange and base-pair kinetics in RNA duplexes, Biochemistry 40 (2001) 8898–8904.
[29] M. Fuxreiter, N. Luo, P. Jedlovszky, I. Simon, R. Osman, Role of base flipping in specific recognition of damaged DNA by repair enzymes, J. Mol. Biol. 323 (2002) 823–834.
[30] L.L. O'Neil, A. Grossfield, O. Wiest, Base flipping of the thymine dimer in duplex DNA, J. Phys. Chem. B 111 (2007) 11843–11849.
[31] L.L. O'Neil, O. Wiest, Structures and energetics of base flipping of the thymine dimer depend on DNA sequence, J. Phys. Chem. B 112 (2008) 4113–4122.
[32] E. Fadda, R. Pomès, On the molecular basis of uracil recognition in DNA: comparative study of T–A versus U–A structure, dynamics and open base pair kinetics, Nucleic Acids Res. 39 (2011) 767–780.
[33] S.M. Law, M. Feig, Base-flipping mechanism in postmismatch recognition by MutS, Biophys. J. 101 (2011) 2223–2231.
[34] A. Barthel, M. Zacharias, Conformational transitions in RNA single uridine and adenosine bulge structures: a molecular dynamics free energy simulation study, Biophys. J. 90 (2006) 2450–2462.
[35] P. Kumar, J. Lehmann, A. Libchaber, Kinetics of bulge bases in small RNAs and the effect of pressure on it, PLoS ONE 7 (2012) e42052.
[36] S.-I. Nakano, H. Oka, Y. Uotani, K. Uenishi, M. Fujii, N. Sugimoto, Dynamics and energetics of the base flipping conformation studied with base pair-mimic nucleosides, Biochemistry 48 (2009) 11304–11311.
[37] M. Nakanishi, Y. Mitane, M. Tsuboi, A hydrogen exchange study of the open segment in a DNA double helix, Biochim. Biophys. Acta 798 (1984) 46–52.
[38] K. Snoussi, J.-L. Leroy, Alteration of A·T base-pair opening kinetics by the ammonium cation in DNA A-tracts, Biochemistry 41 (2002) 12467–12474.
[39] U. Dornberger, M. Leijon, H. Fritzsche, High base pair opening rates in tracts of GC base pairs, J. Biol. Chem. 274 (1999) 6957–6962.
[40] A. Mees, T. Klar, P. Gnau, U. Hennecke, A.P.M. Eker, T. Carell, L.-O. Essen, Crystal structure of a photolyase bound to a CPD-like dna lesion after in situ repair, Science 306 (2004) 1789–1793.
[41] F.C. Clement, N. Kaczmarek, N. Mathieu, M. Tomas, A. Leitenstorfer, E. Ferrando-May, H. Naegeli, Dissection of the xeroderma pigmentosum group C protein function by site-directed mutagenesis, Antioxid. Redox Signal. 14 (2011) 2479–2490.
[42] E.M. Meulenbroek, C.P. Cane, I. Jala, S. Iwai, G.F. Moolenaar, N. Goosen, N.S. Pannu, UV damage endonuclease employs a novel dual-dinucleotide flipping mechanism to recognize different DNA lesions, Nucleic Acids Res. 41 (2013) 1363–1371.
[43] C. Yi, B. Chen, B. Qi, W. Zhang, G. Jia, L. Zhang, C.J. Li, A.R. Dinner, C.-G. Yang, C. He, Duplex interrogation by a direct DNA repair protein in search of base damage, Nat. Struct. Mol. Biol. 19 (2012) 671–676.
[44] J. Song, M. Teplova, S. Ishibe-Murakami, D.J. Patel, Structure-based mechanistic insights into DNMT1-mediated maintenance DNA methylation, Science 335 (2012) 709–712.
[45] E. Rajakumara, J.A. Law, D.K. Simanshu, P. Voigt, L.M. Johnson, D. Reinberg, D.J. Patel, S.E. Jacobsen, A dual flip-out mechanism for 5mC recognition by the *Arabidopsis* SUVH5 SRA domain and its impact on DNA methylation and H3K9 dimethylation in vivo, Genes Dev. 25 (2011) 137–152.
[46] M. Bostick, J.K. Kim, P.-O. Estève, A. Clark, S. Pradhan, S.E. Jacobsen, UHRF1 plays a role in maintaining DNA methylation in mammalian cells, Science 317 (2007) 1760–1764.
[47] J. Sharif, M. Muto, S. ichiro Takebayashi, I. Suetake, A. Iwamatsu, T.A. Endo, J. Shinga, Y. Mizutani-Koseki, T. Toyoda, K. Okamura, S. Tajima, K. Mitsuya, M. Okano, H. Koseki, The SRA protein Np95 mediates epigenetic inheritance by recruiting Dnmt1 to methylated DNA, Nature 450 (2007) 908–913.
[48] C. Bianchi, R. Zangi, How to distinguish methyl-cytosine from cytosine with high fidelity, J. Mol. Biol. 424 (2012) 215–224.
[49] C. Bianchi, R. Zangi, UHRF1 discriminates against binding to fully-methylated CpG-sites by steric repulsion, Biophys. Chem. 171 (2013) 38–45.
[50] DeLano Scientific, LLC, The PyMOL Molecular Graphics System, Version 1.2r1, 2009.
[51] M. Tanokura, ^{1}H-NMR study on the tautomerism of the imidazole ring of histidine-residues.1. microscopic pK values and molar ratios of tautomers in histidine-containing peptides, Biochim. Biophys. Acta 742 (1983) 576–585.
[52] R. Zangi, A. Arrieta, F.P. Cossío, Mechanism of DNA methylation: the double role of DNA as a substrate and as a cofactor, J. Mol. Biol. 400 (2010) 632–644.
[53] J. Farwer, M.J. Packer, C.A. Hunter, PREDICTOR: A web-based tool for the prediction of atomic structure from sequence for double helical DNA with up to 150 base pairs, In Silico Biol. 7 (2007) 595–600.
[54] A. Pérez, I. March'an, D. Svozil, J. Sponer, T.E. Cheatham, C.A. Laughton, M. Orozco, Refinement of the AMBER force field for nucleic acids: improving the description of α/γ conformers, Biophys. J. 92 (2007) 3817–3829.
[55] V. Hornak, R. Abel, A. Okur, B. Strockbine, A. Roitberg, C. Simmerling, Comparison of multiple Amber force fields and development of improved protein backbone parameters, Proteins 65 (2006) 712–725.
[56] W.L. Jorgensen, J. Chandrasekhar, J.D. Madura, R.W. Impey, M.L. Klein, Comparison of simple potential functions for simulating liquid water, J. Chem. Phys. 79 (1983) 926–935.
[57] C. Rauch, M. Trieb, B. Wellenzohn, M. Loferer, A. Voegele, F.R. Wibowo, K.R. Liedl, C5-methylation of cytosine in B-DNA thermodynamically and kinetically stabilizes BI, J. Am. Chem. Soc. 125 (2003) 14990–14991.
[58] P. Cieplak, W.D. Cornell, C. Bayly, P.A. Kollman, Application of the multimolecule and multiconformational RESP methodology to biopolymers: charge derivation for DNA, RNA, and proteins, J. Comput. Chem. 16 (1995) 1357–1377.
[59] B. Hess, C. Kutzner, D. van der Spoel, E. Lindahl, GROMACS 4: algorithms for highly efficient, load-balanced, and scalable molecular simulation, J. Chem. Theory Comput. 4 (2008) 435–447.
[60] T. Darden, D. York, L. Pedersen, Particle mesh Ewald: an N-log(N) method for Ewald sums in large systems, J. Chem. Phys. 98 (1993) 10089–10092.
[61] U. Essmann, L. Perera, M.L. Berkowitz, T. Darden, H. Lee, L.G. Pedersen, A smooth particle mesh Ewald method, J. Chem. Phys. 103 (1995) 8577–8593.
[62] G. Bussi, D. Donadio, M. Parrinello, Canonical sampling through velocity rescaling, J. Chem. Phys. 126 (2007) 014101.
[63] H.J.C. Berendsen, J.P.M. Postma, W.F. van Gunsteren, A. DiNola, J.R. Haak, Molecular dynamics with coupling to an external bath, J. Chem. Phys. 81 (1984) 3684–3690.
[64] S. Miyamoto, P.A. Kollman, SETTLE: an analytical version of the SHAKE and RATTLE algorithms for rigid water models, J. Comput. Chem. 13 (1992) 952–962.
[65] B. Hess, H. Bekker, H.J.C. Berendsen, J.G.E.M. Fraaije, LINCS: a linear constraint solver for molecular simulations, J. Comput. Chem. 18 (1997) 1463–1472.
[66] E. Guàrdia, R. Rey, J.A. Padró, Statistical errors in constrained molecular dynamics calculations of the mean force potential, Mol. Simul. 9 (1992) 201–211.
[67] H. Flyvbjerg, H.G. Petersen, Error estimates on averages of correlated data, J. Chem. Phys. 91 (1989) 461–466.
[68] E. Giudice, P. Várnai, R. Lavery, Energetic and conformational aspects of A:T base-pair opening within the DNA double helix, ChemPhysChem 2 (2001) 673–677.
[69] P. Várnai, R. Lavery, Base flipping in DNA: pathways and energetics studied with molecular dynamic simulations, J. Am. Chem. Soc. 124 (2002) 7272–7273.