JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY
Article
pubs.acs.org/JACS

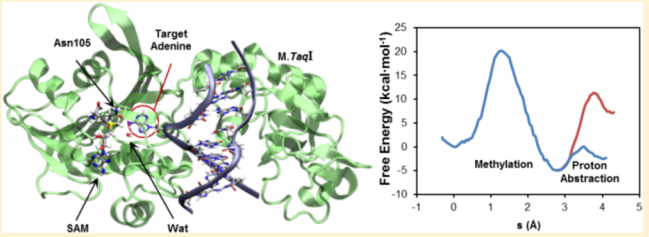
Dynamics and Reactivity in *Thermus aquaticus* N6-Adenine Methyltransferase

Juan Aranda, Kirill Zinovjev, Maite Roca,* and Iñaki Tuñón*

Departament de Química Física, Universitat de València, 46100 Burjassot, Spain

Supporting Information

ABSTRACT: M.*Taq*I is a DNA methyltransferase from *Thermus aquaticus* that catalyzes the transfer of a methyl group from *S*-adenosyl-L-methionine to the N6 position of an adenine, a process described only in prokaryotes. We have used full atomistic classical molecular dynamics simulations to explore the protein−SAM−DNA ternary complex where the target adenine is flipped out into the active site. Key protein−DNA interactions established by the target adenine in the active site are described in detail. The relaxed structure was used for a combined quantum mechanics/molecular mechanics exploration of the reaction mechanism using the string method. According to our free energy calculations the reaction takes place through a stepwise mechanism where the methyl transfer precedes the abstraction of the proton from the exocyclic amino group. The methyl transfer is the rate-determining step, and the obtained free energy barrier is in good agreement with the value derived from the experimental rate constant. Two possible candidates to extract the leftover proton have been explored: a water molecule found in the active site and Asn105, a residue activated by the hydrogen bonds formed through the amide hydrogens. The barrier for the proton abstraction is smaller when Asn105 acts as a base. The reaction mechanisms can be different in other N6-DNA-methyltransferases, as determined from the exploration of the reaction mechanism in the Asn105Asp M.*Taq*I mutant.



1. INTRODUCTION

DNA methylation is an important process in the regulation of gene expression. In eukaryotes DNA methylation in C5 position of cytosine plays a remarkable role in embryonic development, DNA replication, control of differentiation, X-chromosome inactivation, carcinogenesis, and genomic imprinting in epigenetic mechanisms encoding new information in a stable but reversible manner.[1−7] Aberrant cytosine methylation is associated with cancer progression and thus development of inhibitors of C5-methyltransferase activity is an active research field.[8] In prokaryotes N6-methyladenine and N4-methylcytosine are used to differentiate between self- and nonself-DNA by the so-called restriction modification systems.[9] These modified bases are also used as a protection of the host genome against endogenous restriction endonucleases, DNA mismatch repair, and regulation of gene expression and DNA replication. Bacterial DNA methylation plays a fundamental role for the virulence and survival of these organisms.[10,11] Therefore, bacterial DNA methyltransferases (DNA MTases) have been the target for the possible design of antibiotics and antiviral drugs.

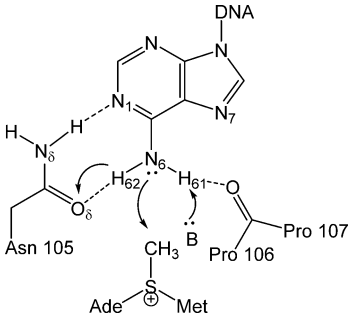
DNA MTases catalyze targeted transfers of methyl groups from *S*-adenosyl-L-methionine (SAM) to adenine residues at N6 position or cytosine residues at C5 or N4 positions, setting up the genetic methylation patterns[12] and being the most common form of postreplicative modification of DNA. Structural and biochemical studies have shown that all DNA MTases flip the target nucleotide base out of the DNA helix, being inserted into the binding pocket, so that the methylation reaction can occur.[13] DNA MTases share similar domains and are formed by one large domain containing the binding sites for the cofactor and the flipped base and one smaller domain that participates in DNA binding and recognition.[12]

In this work we focus on the activity of the N6-adenine DNA MTase from *Thermus aquaticus* (M.*Taq*I), an enzyme that catalyzes the transfer of a methyl group from SAM to the N6 position of adenine within the recognition sequence TCGA.[14] The knowledge about the methylation mechanism of N6-adenine DNA MTase is still incomplete. The mechanism involves the transfer of a methyl group and the abstraction of one of the protons from the exocyclic amine group (Scheme 1). An experimental study on two DNA MTases (*Hha*I and *Eco*RI) showed that the configuration of the transferred methyl group is inverted during the reaction, indicating a single direct transfer from SAM to the DNA base.[15] This could also happen in the process catalyzed by M.*Taq*I.[14] However, it is still unclear if the reaction can proceed with proton abstraction after the methyl transfer or in the opposite way, with proton abstraction preceding methyl transfer. In addition these two processes could take place in a single step (concerted mechanisms) or in two steps (stepwise mechanism), although a first theoretical analysis on a reduced active site model showed preference for the latter.[16] The nature of the base involved in the abstraction

Received: July 29, 2014
Published: October 27, 2014

dx.doi.org/10.1021/ja5077124 | J. Am. Chem. Soc. 2014, 136, 16227−16239

Scheme 1. General Reaction Mechanism for DNA Methylation in M.*Taq*I



of the proton is also unknown in the case of of M.*Taq*I, because there are no basic residues in its active site. This is a characteristic of N6-MTases belonging to group γ, where motif IV is composed by Asn-Pro-Pro-Tyr (residues 105−108 in M.*Taq*I, see Scheme 1).[17] Instead, motif IV of N6-MTases of groups α and β is Asp-Pro-Pro-Tyr.[17] The Asp residue of this motif is correctly placed to act as the base that abstracts the leftover proton from the exocyclic amino group. Because of the absence of a basic residue in the active site, it was proposed that, in the case of M.*Taq*I, a water molecule could be responsible for the proton abstraction step.[16] It has been also proposed that the carbonyl group of the Asn105 side chain could act as the proton acceptor.[14] Therefore, new insights are needed to elucidate these and other mechanistic details of N6-methylation in M.*Taq*I that could be relevant for the design of more specific inhibitors of bacterial activity.[18]

The structure of M.*Taq*I complexed with DNA and a nonreactive cofactor (AETA) was solved by means of X-ray crystallography with a resolution of 2.0 Å.[14] This structure shows the target adenine (A6) flipped out in the active site and kept there by means of hydrogen-bond interactions with residues Asn105, Pro106, and Tyr108 (see Figure 1). Asn105 is forming two hydrogen bonds through Nδ and Oδ atoms with distances of 2.98 and 3.15 Å to N1 and N6 atoms of adenine, respectively. The carbonyl oxygen atom of Pro106 forms a hydrogen bond with the N6 atom of adenine (2.88 Å), and the nitrogen atom belonging to Tyr108 backbone is hydrogen bonded to the N7 atom of the nucleic base with a distance of 3.08 Å. This arrangement is consistent with a direct methyl transfer from the cofactor to the exocyclic amino group of A6 that would be activated by the hydrogen bonds established by this group in the active site. In the X-ray structure the bound DNA shows large distortions compared to B-form DNA. Not only the target adenine A6 is completely rotated but also the partner thymine (T15) is displaced toward the center of the double helix. DNA−protein interactions were identified in the X-ray structure: there were 12 direct contacts between M.*Taq*I and phosphodiester groups in addition to specific interactions with all the bases of the recognition sequence, with the only exception of T15. This structure was used as starting point for a quantum mechanical study that analyzed the methylation reaction by means of a small cluster model.[16] This work employed trimethylsulfonium as methyl donor, while adenine was constrained to keep an arrangement similar to that observed in the crystallographic structure. The effect of hydrogen bonding by Asn105 was explored using a formamide molecule, which showed only a minor effect on the methyl transfer barrier. It was proposed that the reaction could proceed through a stepwise mechanism, with proton abstraction taking place after methylation with the participation of water molecules. The same X-ray structure was used in a molecular dynamics (MD) simulation of the ternary DNA−SAM−protein complex.[19] This study showed that some nonspecific protein−DNA interactions are altered with respect to the crystallographic structure and that an unusual backbone conformation of DNA could assist the flipping of the target base out of the helix.

To our knowledge there are no studies that integrate the dynamics and the reactivity in a full atomistic model of the system. DNA distortion, fluctuations of protein−DNA interactions, and diffusion of solvent molecules can affect the reactivity of the system, favoring conformations different to those observed in the X-ray structure. With the purpose of analyzing the N6-methylation process in a realistic model, we have performed long molecular dynamics (MD) simulations for the whole ternary complex. Then, starting from a relaxed structure we calculated the free energy profiles corresponding to the reaction process, including both the methylation and proton abstraction steps. These profiles are obtained by means of MD simulations using hybrid quantum mechanics/molecular mechanics potentials (QM/MM). Our simulations show extensive rearrangements in the DNA helix. These changes favor the accommodation of the system with the target adenine in the active site of the enzyme, as discussed below. According to our QM/MM calculations the reaction in M.*Taq*I takes place through a stepwise mechanism with methylation preceding

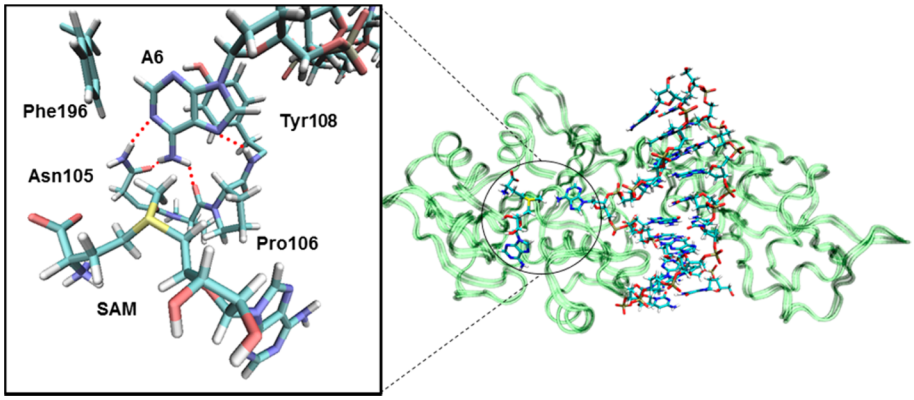


Figure 1. Structure of M.*Taq*I complexed with a decamer of DNA (PDB code 1G38). The inset shows an insight into the active site with cofactor SAM, the target base (A6), and key residues establishing hydrogen bonds with A6. Hydrogen bonds are drawn using red dotted lines.

dx.doi.org/10.1021/ja5077124 | J. Am. Chem. Soc. 2014, 136, 16227−16239

proton abstraction, being the first step the rate-limiting one. Finally, we have analyzed the role of the base that abstracts the excess proton from the amino group of A6, comparing the results obtained for wild-type M.*Taq*I with those obtained for the Asn105Asp mutant. This mutant enzyme should be closer to N6-MTases of groups α and β, which present the same catalytic motif in the active site.[20−24]

2. METHODOLOGY

2.1. Building the System. The starting structure for our simulations was built from the crystal structure of N6-adenine MTase from M.*Taq*I solved in complex with a specific DNA decamer and a nonreactive cofactor AETA (PDB ID: 1G38, resolution of 2 Å).[14] The recognition and methylation target sequence is $\begin{matrix}5'\text{-}TCGA'\text{-}3'\\ 3'\text{-}A''GCT\text{-}5'\end{matrix}$ where the target nucleoside A′ is rotated out of the DNA helix and inserted into the catalytic pocket. The PDB structure contains a dimer of the ternary structure. From the PDB file we selected the chains labeled as A-B-C, as no relevant differences were found by comparison with the other chains. As DNA was hemimethylated, the methyl group added to the adenine labeled as A″ was substituted by a hydrogen atom to study *de novo* methylation in a prechemistry system. Afterward, we superimposed our starting PDB coordinates with those of the crystal structure of the same N6-MTase in complex with its natural cofactor SAM (PDB ID: 2ADM).[25] Structures were superimposed on the basis of the backbone atoms of the protein, and the AETA inhibitor was replaced with SAM. Finally, because the X-ray structure contains only residues 21−413, Val 21 was capped with an acetyl group. In order to validate the resulting structural model we confirmed that all the hydrogen bonds observed between the natural cofactor and the enzyme in 2ADM[25] were preserved in the simulations performed with our model (see Figure S1 in the Supporting Information). It has been shown that some DNA MTases can suffer important conformational changes during the catalytic cycle.[26] In our case we have checked this possibility and its impact on the quality of our initial structure for the simulations overlapping different X-ray structures available in the PDB: 1G38, the structure selected as starting point of our work, which contains the protein in complex with DNA and a cofactor analogue (AETA); 2ADM[25] that contains the protein and the cofactor SAM (without DNA); and 1AQJ[25] in which the protein forms a complex with another cofactor analogue (sinefungin). All these structures overlap very well except for small displacements in some flexible loops. Thus, important conformational changes are not expected in M.*Taq*I due to a modification in the nature of the cofactor or after DNA binding, supporting our choice to model the starting point of the simulations.

The pK_a of titratable residues was calculated with PropKa 3.0.[27−30] Once this was determined, hydrogens were added to the enzyme−SAM−DNA complex assuming a pH value equal to 7, using the LEAP module of the AMBER program.[31] The LEAP module was also used to solvate our system, keeping crystallization water molecules in our initial system. A box of TIP3P water molecules of 108 × 78 × 78 Å size was used, keeping a buffer region of about 15 Å around the protein−DNA system in all the directions. A total of 15 sodium ions were added to neutralize the system. The final system contained 20537 water molecules and a total of 68686 atoms. We then performed several optimization and molecular dynamics simulations using AMBER ff03[32] with the NAMD program.[33] Parameters for the cofactor SAM were taken from Stacklies et al.[34] We carried out energy optimizations of the cofactor in vacuum, and the resulting overall geometry was similar to that obtained with quantum mechanical methods. We also checked the geometry over the MD simulation to verify the reliability of these parameters. In particular, the pattern of hydrogen bonds established between the protein and the cofactor (Figure S1) was preserved during the simulations.

2.2. MD Simulations. Simulations were performed with a time step of 1 fs and periodic boundary conditions with full electrostatics computed using the particle mesh Ewald method.[35] Short range nonbonded interactions were calculated at every step using a smooth switching function with a 12 Å cutoff. The protocol followed to equilibrate the system starts with the relaxation of the water box and counterions by means of 1 ns of MD simulation in the NPT ensemble and then 1 ns of NVT MD simulation. Pressure was maintained at 1 atm using the Nosé−Hoover Langevin method with a period of 100 fs and a decay time scale of 50 fs. Langevin−Verlet dynamics were performed using a damping coefficient of 10 ps^{-1} at a target temperature of 333 K. This is the value at which the catalytic rate constant has been determined.[36]

After relaxation of the water box, the enzyme and DNA atoms were gradually released starting with side chains and nucleotide bases atoms and finishing with backbone atoms. At each step the system was first minimized by means of 10000 steps of conjugate-gradient algorithm followed by MD simulations where the system was heated, rescaling velocities from 0 to 333 K during 500 ps. Additional 500 ps were performed at constant temperature. During this process harmonic constraints were applied to the hydrogen-bond interactions observed between the flipped adenine base (A6) and active site residues (Pro106 and Asn105, see Figure 1) and to the distance between the methyl group of SAM and N6 atom of adenine. These harmonic constraints were gradually removed through several cycles of optimization and MD simulations. After equilibration of the system we performed 100 ns of fully unrestrained MD simulation in the NVT ensemble at 333 K.

2.3. QM/MM Calculations. A snapshot from the last 10 ns of the MD simulation was selected to build up the system for QM/MM-MD simulations with the fDynamo program.[37] The QM subsystem, described at the AM1 semiempirical level,[38] consisted in the SAM cofactor, the base that abstracts the proton (Asn105 side chain or one water molecule located at the active site), and the adenine A6, while the rest of the system was described using the AMBER force field implemented in fDynamo.[39] To saturate the valence of the QM/MM frontier we used the link atoms procedure.[40,41] As the charge of nucleotides and amino acids side chains is not an integer in the AMBER force field, we redistributed the remaining fractional charges between the C1′ and H1′ atoms of the sugar and the CA and HA atoms of Asn105 so that the charge was an integer in the classical subsystem.[42] After minimization, a 200 ps long NVT simulation was performed using the hybrid QM/MM potential. We then explored the potential energy surface (PES) at M06-2X/6-31G**/MM level.[43] These calculations were done with Gaussian09[44] coupled to fDynamo. The reaction catalyzed by M.*Taq*I consists of two processes: the methyl transfer from SAM to the exocyclic amino group of A6 and the abstraction of the excess proton by a surrounding base. We performed calculations where Asn105 or a water molecule encountered in the catalytic pocket act as the proton acceptors.

dx.doi.org/10.1021/ja5077124 | J. Am. Chem. Soc. 2014, 136, 16227−16239

The water molecule was selected after inspection of the MD simulation. The X-ray structure shows a water molecule in the vicinity of the N7 atom of the target adenine.[14] This position is occupied by a water molecule during most of the simulation. As shown below, the distance between the oxygen atom of this water molecule and the proton to be abstracted is drastically reduced after methylation, making this water molecule a good candidate to act as a base. Each PES was explored using two distinguished reaction coordinates that control the advance of the methyl and proton transfers, respectively. For the methyl transfer the coordinate was constructed as the antisymmetric combination of the distances of the carbon atom of the methyl group to the donor and to the acceptor atoms ($S-CH_3$ and $N6-CH_3$). Similarly, the proton transfer was mapped using the antisymmetric combination of the distances to the donor atom (N6) and the acceptor atom (the Oδ atom of the Asn105 residue or the oxygen atom of the closest water molecule, O_w). According to the obtained PESs (see Figure S2 in Supporting Information) the most feasible mechanism turned out to be a stepwise mechanism where the methylation occurs first and then the abstraction of the proton takes place.

2.4. QM/MM Free Energy Calculations. QM/MM MD simulations were performed to obtain the minimum free energy path (MFEP) by means of the on-the-fly string method.[45] After obtaining the MFEP, a path collective variable[46,47] was defined to obtain the potential of mean force (PMF) using umbrella sampling.[48] This methodological combination offers an efficient approach for studying the free energy landscape of complex enzymatic reactions.[49] The active space selected to trace the MFEP was formed by the distances $S-CH_3$, CH_3-N6, N6–H62, and H62–OδAsn105 for the reaction mechanism where the Asn105 abstracts the proton and $S-CH_3$, CH_3-N6, N6–H61, and H61–O_w for the process where the water molecule is the proton acceptor. A total of 30 string nodes was used to trace the MFEP for each reaction mechanism, and the values employed for the force constant and friction in the string method were K = 6000 kJ·mol^{-1}·Å^{-2}, γ = 9000 ps^{-1}, respectively. We performed 50 ps of Langevin dynamics at 333 K with a time step of 1 fs. The strings converged within the first 10 ps of dynamics, and the averaged positions of the string nodes were determined over the last 40 ps. We used different initial guesses in the string method to explore all possible reaction mechanisms. The only ones resulting in a converged string were those corresponding to a stepwise mechanism with methyl transfer preceding proton abstraction in agreement with the aforementioned PESs (see below).

In order to obtain the PMF a set of 100 points were interpolated from the string nodes to be used in the definition of the path coordinate (the s coordinate). This coordinate measures the advance of the system along the MFEP and was parametrized with values ranging from 0 to L, where L corresponds to the total arc length of the path. Negative values and values larger than L correspond to reference structures extrapolated beyond the reactants and products, respectively. This extrapolation procedure is used to have well-defined reactants and products valleys in the PMF. PMFs were obtained using interpolated corrections[50,51] along the s coordinate:

$$E = E^{LL}_{QM} + E^{LL}_{QM/MM} + E_{MM} + Spl[\Delta E^{HL}_{LL}(s)]$$

where Spl denotes a one-dimensional spline function and its argument, ΔE^{HL}_{LL}, is a correction term taken as the difference between the single-point high-level (HL) energy of the QM system and the low level one (AM1). As HL theory we chose M06-2X/6-311+G**.[43] Energies are calculated for structures optimized at the AM1/MM level, constrained at different values of s and considering electrostatic embedding. For the converged strings, the harmonic force constants for the biasing potentials used in 120 umbrella sampling[48] windows were adjusted to flatten the free energy profiles obtained from the string method with values ranging between 2000 and 5000 kJ·mol^{-1}·Å^{-2}. Initial structures for each window were taken from the final structures of the closest string node. Ten ps relaxation followed by 40 ps of production dynamics with a time step of 1.0 fs were performed for each simulation window. Then the PMFs were integrated from the obtained histograms using WHAM.[52]

3. RESULTS

3.1. MD Simulations. We performed 100 ns of NVT MD simulation of the ternary complex (DNA–SAM–enzyme) in order to investigate the most important interactions formed between the enzyme and DNA, the interactions established within the active site, and how the DNA helix accommodates the perturbation that the rotation of the target adenine causes in its structure. Plot of the root-mean-squared deviations (RMSD) for the enzyme and DNA backbone atoms as well as the percentage of canonical α/γ torsions in the latter are shown in Figure S3 in the Supporting Information. Our initial structure is expected to be close to the real Michaelis complex, and thus large conformational changes are not expected, as suggested by the good overlap among different crystal structures discussed above. This is confirmed by the stability of our MD simulations. Otherwise, biased simulations (with an adequate target) would be required to obtain an appropriate initial complex.

3.1.1. Protein–DNA Interactions. First we analyzed the hydrogen-bond network established between the protein and the DNA duplex. There are two main types of hydrogen-bond interactions that can be formed: (i) direct hydrogen bonding between a donor/acceptor atom of the enzyme and an acceptor/donor DNA atom and (ii) water-mediated hydrogen bonds where a water molecule is hydrogen bonded simultaneously to a donor atom of the protein/DNA chain and an acceptor atom of the DNA/protein chain. These water-bridged hydrogen bonds may play a key role in both recognition and stabilization of protein–DNA complexes.[53−55] Figure 2 shows the time evolution in the number of direct and water-bridged enzyme–DNA hydrogen-bond interactions. The criteria used to establish the formation of a hydrogen bond were a donor–acceptor distance smaller than 3.2 Å and a donor-H-acceptor angle higher than 140°. We obtained an averaged value of 16.5 ± 2.9 direct protein–DNA hydrogen bonds and 21.5 ± 4.1 water molecules bridging between both chains through the simulation. The total number of hydrogen-bond interactions between the protein and DNA was then about 38. A total of nine direct hydrogen bonds involve positively charged residues of the protein (Lys116, Lys126, Lys139, Lys199, Arg267, Arg271, Arg296, Arg353, and Arg398). These interactions stabilize the 16 negative charges found in the interacting surface between the DNA duplex and the protein that extends from G1 until G8 and from A12 until C20. It is important to realize that while the averaged numbers are kept during the whole simulation, the instantaneous values present important fluctuations. In some snapshots the number of direct hydrogen bonds can be as low as 6 while in others can be as high as 26. These fluctuations are partially compensated

dx.doi.org/10.1021/ja5077124 | J. Am. Chem. Soc. 2014, 136, 16227−16239

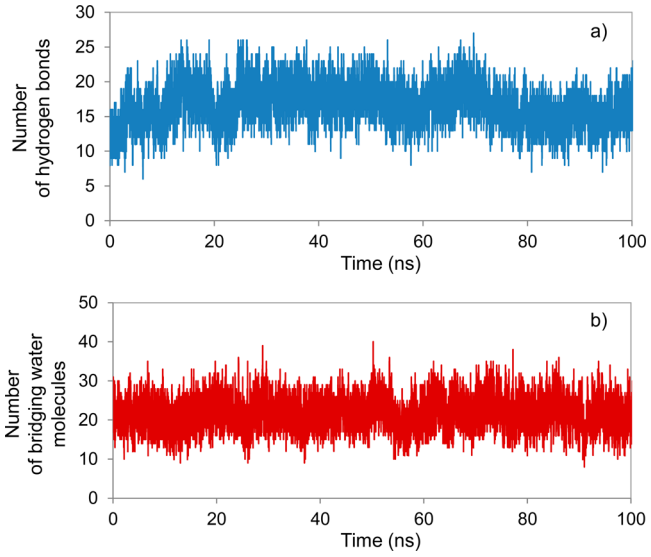


Figure 2. Representation of the number of direct DNA−protein hydrogen bonds (a) and of the number of bridging water molecules between DNA and protein chains (b) established during the simulation time.

by the number of bridging water molecules because some direct hydrogen bonds are lost just to accommodate a water molecule between the donor and the acceptor atoms.

Figure 3 shows the averaged distances (a) and the occupancies (b) for the hydrogen-bond contacts between the protein and the DNA obtained from the 100 ns simulation. As expected, the most stable hydrogen-bond interactions are those established with the recognition sequence (TCGA) highlighted in green in Figure 3. These hydrogen bonds display shorter averaged distances and higher occupancy rates. We note that some direct DNA−protein interactions found in the X-ray structure are lost or partially lost at early stages of our simulation. In particular, those established between DNA and residues Lys130, Glu172, Asp173, Thr167, Lys200, Phe268, Gly295, and Arg323. In contrast, some DNA−protein hydrogen bonds not observed in the X-ray structure appeared during the simulation: Arg267, Arg398, Asn141, Lys199, Arg296, His335, and Ser115.

From the analysis of the X-ray structure it has been proposed that the distortion produced by the flipping of the target A6 base is stabilized by means of protein−DNA interactions established between the 3′- and 5′-phosphates groups of the unpaired T15 base with residues Gly295 and Arg353. This is usually described as a local compression of the DNA helix at the orphanage base backbone position.[14] During our simulation we do not observe any interaction of Gly295 with these phosphate groups, while Arg353 is hydrogen bonded to the 5′-phosphate through a water bridging molecule. We also found two additional protein−DNA interactions in this compression region: His394 is hydrogen bonded to O4 atom of T15, and the 3′ phosphate group of the same base is hydrogen bonded either to Arg353 or to Lys116.

According to Figure 3b there are some hydrogen-bond interactions between DNA and the protein that can be labeled as highly specific, as they are largely conserved during the simulation with high occupancy numbers (purple color in Figure 3b). High occupancy numbers mostly appear for those hydrogen bonds established with the nucleotides of the recognition and methylation sequence (T3, C4, G5, A6, T15, C16, G17, A18). Protein residues displaying occupancies rates higher than 50% over the 100 ns simulation were: Arg398 (78.6), His335 (62.1), Asn105 (81.4), Ser272 (57.1), and Arg273 (91.7). However, most of the protein−DNA hydrogen-bond interactions can be formed, broken, and later recovered during the simulation time. Individually these hydrogen bonds display smaller occupancy numbers, but overall they provide the adequate electrostatic environment to bind the DNA duplex.

3.1.2. DNA Structure. The flipping of the A6 adenine and the interaction of the duplex with the protein induces some changes in the DNA structure. We have analyzed the hydrogen-bond interactions held between base pairs in the DNA duplex and the behavior of the unpaired T15 base. Figure 4 shows the average number of hydrogen bonds established between the canonical pairs. It can be observed that the hydrogen bonds between base pairs are maintained all over the simulation, with the exceptions of the A6-T15 pair (that includes the flipped base), the neighbor T7-A14 pair, and the terminal C10-G11 pair, placed at the extreme of the decamer that is farther to the recognition sequence. The rest of base pairs keep an averaged number of canonical hydrogen bonds close to 1.7 or 2.5 for A-T or C-G base pairs, respectively. These values are close to those observed for simulations of B-DNA in aqueous solution.[56]

The unpaired T15 base is not only stabilized by means of protein−DNA interactions (see Figure 3) but also by means of DNA−DNA interactions appearing in the deformed DNA structure. Figure 5 illustrates the evolution of the distance between the centroid of the unpaired base T15 and the surrounding nucleic bases. It can be observed that, in the ternary complex, the T15 base is shifted toward the complementary strand forming a π-stacking interaction with G5 (see Figure 5a). Moreover, during our simulations T7 loses the hydrogen-bond interactions to A14 and forms a π-stacking interaction with both T15 and A14 (see Figure 5b). We also observed a conformational change in the backbone of DNA, reflected in the RMSD plot presented in Figure S3 of the Supporting Information, reorganizing finally a noncanonical triple π-stacking T15-T7-A14.

3.1.3. Active Site Interactions. The A6 base is inserted into the active site of M.*Taq*I and stabilized by means of hydrogen-bond interactions with residues Asn105, Pro106, and Tyr108 (see Figure 6). The exocyclic amino group is hydrogen bonded to the Oδ oxygen atom of Asn105 and the backbone oxygen of Pro106. In addition, the Nδ nitrogen of Asn105 acts as proton donor to the N1 atom of A6. Tyr108 forms one hydrogen bond through the backbone nitrogen atom with N7 atom of A6 and a π-stacking interaction with the base ring. The average distance between the centroids of the Tyr108 aromatic ring and the six-membered ring of adenine is 3.9 ± 0.2 Å, while the average angle between the planes of these rings is 11 ± 6°. The flipped adenine establishes a T-shaped CH···π interaction with another aromatic residue present in the active site, Phe196. The average distance between the hydrogen attached to position C2 of adenine and the centroid of the Phe196 aromatic ring is 2.7 ± 0.3 Å, while the average angle between the planes defined by the adenine and the Phe196 aromatic ring planes is 74 ± 14°. These T-shaped CH···π interactions have been reported in other biological systems.[57] Our observations agree with the results of mutational studies on these two conserved aromatic residues, which concluded that both Tyr108 and Phe196 establish favorable interactions in the active site with the target

dx.doi.org/10.1021/ja5077124 | J. Am. Chem. Soc. 2014, 136, 16227−16239

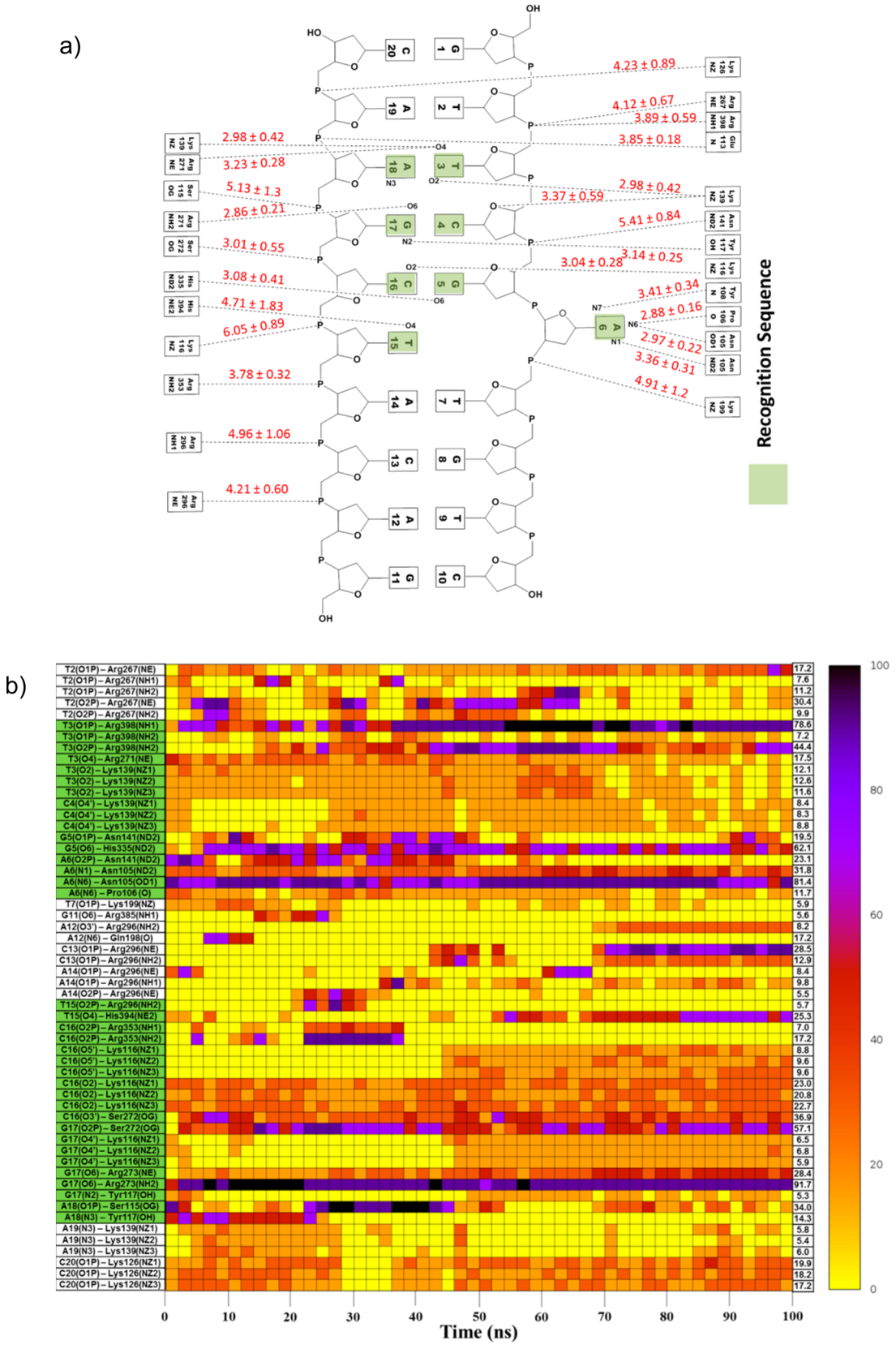


Figure 3. Analysis of important DNA–protein hydrogen-bond interactions. (a) Averaged distances and their standard deviations (in Å) between heavy atoms involved in the hydrogen bond. (b) Plot of the occupancies of DNA–protein hydrogen bonds during the simulation. Each block of the grid represents occupancies averaged during 2 ns. We only plot those hydrogen bonds observed for at least 5% of the simulation time. Cells in green represent hydrogen bonds formed with the recognition sequence. The last column on the right displays the total occupancy numbers determined from the whole simulation.

dx.doi.org/10.1021/ja5077124 | J. Am. Chem. Soc. 2014, 136, 16227–16239

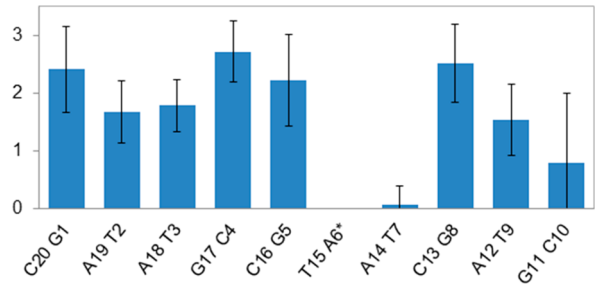


Figure 4. Average number of canonical hydrogen-bond interactions established between DNA base pairs. The standard deviation is also shown as a vertical line.

adenine.[36] It was also suggested that these residues could stabilize the reaction TS by means of cation−π interactions.[36,58] Figure 6 shows the time fluctuations of these hydrogen-bond distances and their averaged values. It must be emphasized that all the distances associated with the interactions of the A6 base within the protein active site remain stable over the whole simulation with small fluctuations around the averaged distances as indicated by the values of the standard deviations. This behavior is an indication of the large affinity of the M.*Taq*I active site for adenine as the target base for methylation.

The interactions established by the target adenine in the active site also assist the change in the hybridization state of the N6 atom from sp^2 toward sp^3. We measured the distance of the exocyclic nitrogen atom to the plane formed by the three substituents. The averaged value corresponding to the non-flipped adenines (A12, A14, A18, and A19) is 0.067 ± 0.050 Å, a small value as expected for sp^2 nitrogen atoms. However, the averaged value obtained for the N6 nitrogen atom of A6 during the MD simulation was 0.174 ± 0.067 Å, indicating a significant pyramidalization. This process would favor the localization of the lone pair and then the methylation process. This effect is more pronounced when the A6 adenine is described at the QM level (the averaged distance being 0.260 ± 0.072 Å), probably because the MM force field is not flexible enough to fully capture this effect.

3.2. M.*Taq*I Reactivity. We have explored different reaction mechanisms and determined the most reliable one in terms of free energy barriers. First we explored the two-dimensional PES corresponding to the methyl transfer and proton abstraction processes at the M06-2X/6-31G**/MM level of theory (see Figure S2 of the Supporting Information). These preliminary calculations showed that the preferred mechanism is a stepwise process where the methyl transfer precedes the abstraction of the excess proton of the exocyclic amino group (see Scheme 2). This last step can be carried out either by the Asn105 side chain or by a water molecule present in the active site. Then, the MFEPs were calculated using the string method.[45] The PMFs were then obtained as a function of the path-based collective coordinate (s) that maps the advance of the system along the MFEPs.[46,47] All the simulations performed to obtain the free energy profiles were carried out at the AM1/MM level, with the PMFs including interpolated corrections at M06-2X/6-311+G** level as a function of the s coordinate (see Section 2, Methodology).

The presence of a water molecule hydrogen bonded to N7 atom of adenine base (see Figure 6, pink dashed line) made this molecule a good candidate to act as the proton acceptor. In our 100 ns simulation the N7 atom of adenine forms in average 0.85 ± 0.26 hydrogen bonds with water molecules. A water molecule is found in that position in the X-ray structure.[14] It is important to emphasize that methylation leads to a positively charged adenine intermediate where the exocyclic nitrogen atom presents an increased acidity. This is required to facilitate proton abstraction when there are no strong bases available. In addition, the basicity of this water molecule is enhanced because it donates a hydrogen bond to the N7 atom of the A6 base. In many of the snapshots analyzed this water molecule was correctly oriented to abstract a proton from the exocyclic amino group. When the positively charged methyl group is transferred to the N6 atom, the water molecule approaches to

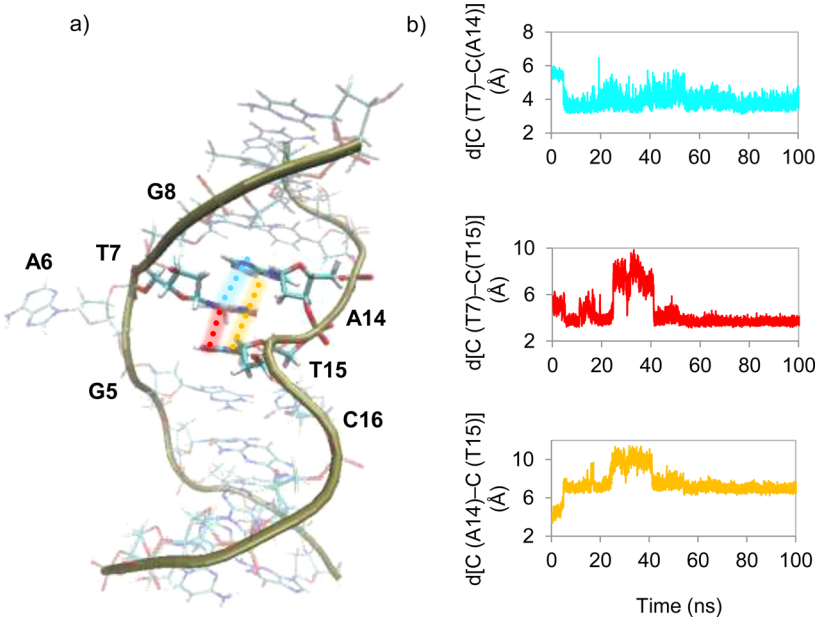


Figure 5. Interactions between unpaired thymine base and surrounding nucleic bases. (a) General view of the triple stacking conformation. (b) Distance between centroids (C) of T7, T15, and A14.

dx.doi.org/10.1021/ja5077124 | J. Am. Chem. Soc. 2014, 136, 16227−16239

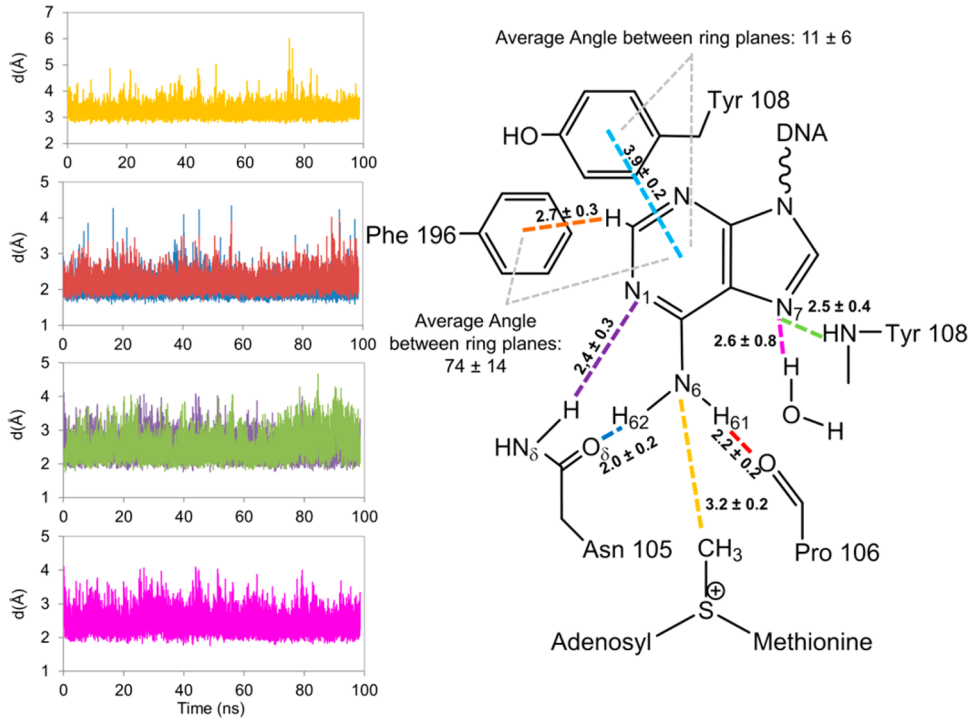
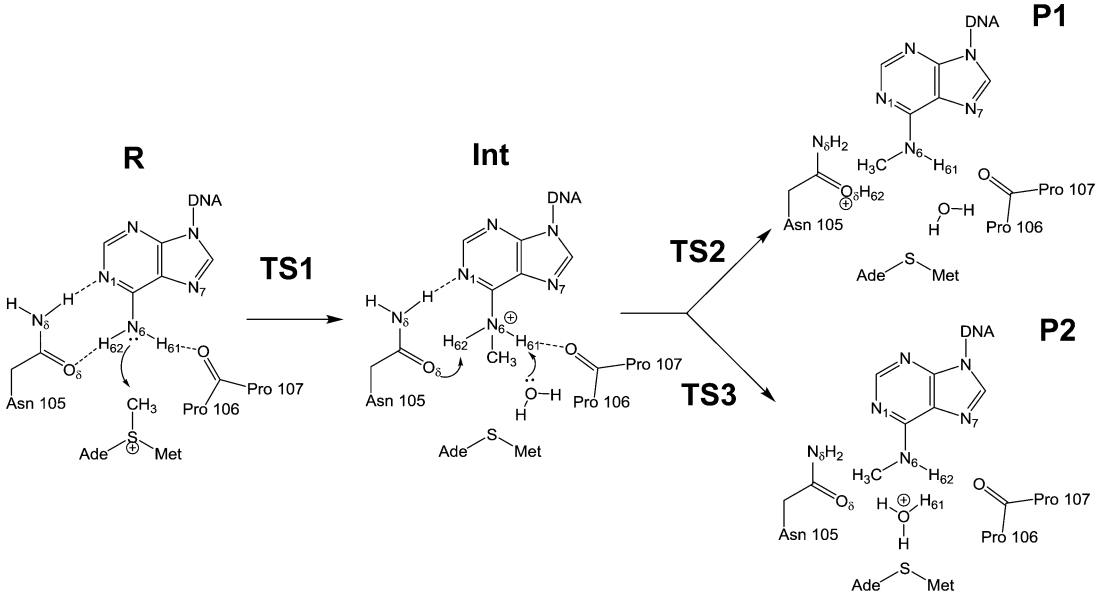


Figure 6. Interactions established between the A6 base and amino acids of M.*Taq*I active site. The graphs show the instantaneous values observed during the simulations. The values shown in the schematic design of the active site correspond to the averaged values and their standard deviations. Distances are given in Å and angles in degrees.

Scheme 2. Reaction Mechanisms for Adenine Methylation in M.*Taq*I



the exocyclic amino group and then can abstract the H61 proton. After methylation the distance between the oxygen atom of this water molecule and the excess proton is reduced up to 2.49 ± 0.06 Å, see below. As can be observed in Figure S4 in the Supporting Information, this water molecule makes part of a water channel connecting the active site with the bulk,[16] which could facilitate the transport of the excess proton. We also investigated the possibility of proton abstraction by the carboxamide oxygen of Asn105 which was proposed as a possible base to deprotonate the 6-methylammonium intermediate.[14] This atom is found to remain hydrogen bonded to the H62 atom of the A6 amino group (see Figure 6). When the methyl group is transferred, the positive charge of this hydrogen atom increases and the distance to the Asn105 Oδ atom decreases. The basicity of the oxygen atom of the carboxamide group is enhanced by the hydrogen bonds formed through the Nδ atom of this residue with the N1 atom of adenine and with the Oγ atom of Thr23. The hydroxyl group of this residue is in turn hydrogen bonded to the carboxylate group of SAM (see Figure S5), strengthening the hydrogen-bond interaction established with Asn105.

dx.doi.org/10.1021/ja5077124 | J. Am. Chem. Soc. 2014, 136, 16227−16239

Figure 7 shows the PMFs corresponding to the reaction mechanism, while key averaged distances obtained for the

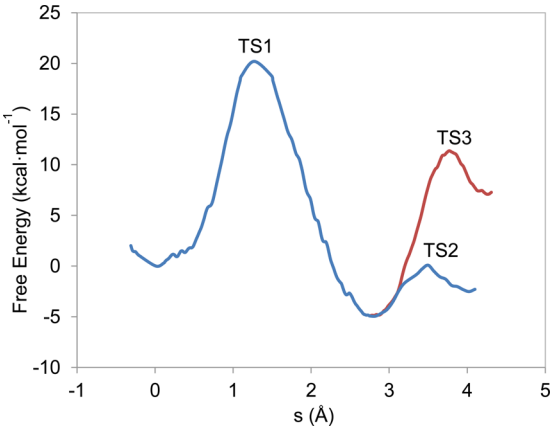


Figure 7. AM1/MM PMFs, corrected at the M06-2X/6-311+G** level, obtained for the adenine methylation (TS1) and proton abstraction reaction (TS2 or TS3) in M.*Taq*I as a function of the collective path coordinate *s*. The most favored mechanism, where the Asn105 abstracts the leftover proton, is depicted in blue, while the mechanism where the proton is accepted by a water molecule is shown in red.

stationary states are given in Table 1. Two mechanisms displaying two different TS were found, identifying a stepwise mechanism. The preferred reaction mechanism is a nucleophilic attack of the N6 exocyclic atom of the adenine base to the methyl group of SAM followed by the abstraction of the excess proton by the Asn105 residue (blue line in Figure 7). In our QM/MM description the carbon atom of the methyl group is located at 3.11 Å from the N6 atom at the reactant state, being well placed for an in-line transfer to the exocyclic nitrogen atom of the target adenine (the average S−CH_3−N6 angle being 167 ± 7°). The exocyclic amino group of the target A6 base forms two hydrogen bonds, one with the carbonyl oxygen atom of Pro106 (average distance of 2.30 Å) and one with the Oδ atom of Asn105 (2.44 Å). These hydrogen bonds increase the nucleophilic character of the amino group.[14] In addition the A6 base is hydrogen bonded through the N1 and N7 atoms to the Nδ atom of Asn105 and to the backbone nitrogen atom of Tyr108, with averaged distances of 2.27 and 2.55 Å, respectively.

Methyl transfer from SAM to the N6 atom of the target adenine base is the rate-limiting step. The calculated activation free energy for this step is 20.1 kcal·mol^{-1}, which is in excellent agreement with the experimental value determined at 333 K (the same temperature used in our simulations), 19.8 ± 0.1 kcal·mol^{-1}.[36] While this agreement can partially result from errors cancellation, we have applied similar computational protocols to other methyltransferases (catechol *O*-methyltransferase,[51,59] glycine *N*-methyltransferase,[60,61] and guanidinoacetate *N*-methyltransferase[47]) obtaining for all these systems good quantitative agreements between the theoretical and experimentally derived free energy barriers. This agreement reinforces us about the goodness of our computational model. A representative structure of this transition state (TS) is presented in Figure 8, while averaged distances are provided

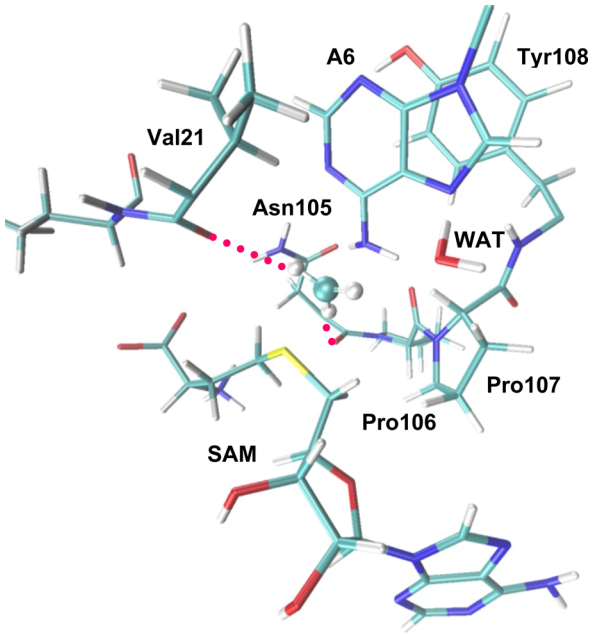


Figure 8. Transition state structure for the rate-limiting step of the reaction (methyl transfer from SAM to the exocyclic amino group, TS1). In pink dotted lines, interactions established between the methyl group and the carbonyl oxygen atoms of Asn105 and Val21 are shown.

in Table 1. In this TS the distance of the carbon atom of the transferred methyl group to the N6 atom of adenine is shorter than to the sulfur atom of SAM (1.93 and 2.25 Å, respectively). The average S−CH_3−N6 angle is close to linearity, 173 ± 4°. The positioning of the transferred methyl group between the

Table 1. Relevant Averaged Distances (in Å) Obtained for Reactant, Intermediate, Products and Transition States of N6-Adenosine Methylation in M.*Taq*I[a]

	R	TS1	INT	TS2	P1	TS3	P2
d(S(SAM)-CH_3(SAM))	1.83 ± 0.04	2.25 ± 0.05	3.72 ± 0.04	4.22 ± 0.21	4.34 ± 0.23	4.11 ± 0.26	4.11 ± 0.29
d(CH_3(SAM)-N6(A6))	3.11 ± 0.04	1.93 ± 0.05	1.49 ± 0.03	1.47 ± 0.04	1.45 ± 0.02	1.47 ± 0.03	1.45 ± 0.03
d(N6(A6)-H62(A6))	1.01 ± 0.04	1.03 ± 0.03	1.04 ± 0.04	1.22 ± 0.04	2.02 ± 0.04	1.04 ± 0.04	1.01 ± 0.02
d(H62(A6)-Oδ(Asn105))	2.44 ± 0.03	2.30 ± 0.03	2.24 ± 0.03	1.38 ± 0.04	1.00 ± 0.03	2.77 ± 0.02	2.48 ± 0.03
d(N6(A6)-H61(A6))	1.01 ± 0.03	1.02 ± 0.03	1.04 ± 0.03	1.03 ± 0.03	1.02 ± 0.03	1.34 ± 0.04	3.00 ± 0.04
d(O(WAT)-H61(A6))	2.49 ± 0.06	2.39 ± 0.05	2.09 ± 0.04	2.41 ± 0.08	3.27 ± 0.21	1.27 ± 0.04	1.01 ± 0.03
d(H61(A6)-O(Pro106))	2.30 ± 0.22	2.26 ± 0.20	1.98 ± 0.20	2.22 ± 0.27	2.52 ± 0.28	2.21 ± 0.19	2.87 ± 0.36
d((HNδ(Asn105)-N1(A6))	2.27 ± 0.20	2.42 ± 0.19	3.01 ± 0.22	2.87 ± 0.21	2.61 ± 0.23	4.04 ± 0.32	3.81 ± 0.35
d((HN(Tyr108)-N7(A6))	2.55 ± 0.25	2.51 ± 0.20	2.86 ± 0.39	2.30 ± 0.17	2.28 ± 0.15	2.40 ± 0.19	2.38 ± 0.23
free energy	0	20.1	−5.5	0.1	−2.5	11.1	7.4

[a]Standard deviations are included. Free energies obtained after M06-2X/6-311+G**/MM corrections are given in kcal·mol^{-1}.

dx.doi.org/10.1021/ja5077124 | J. Am. Chem. Soc. 2014, 136, 16227−16239

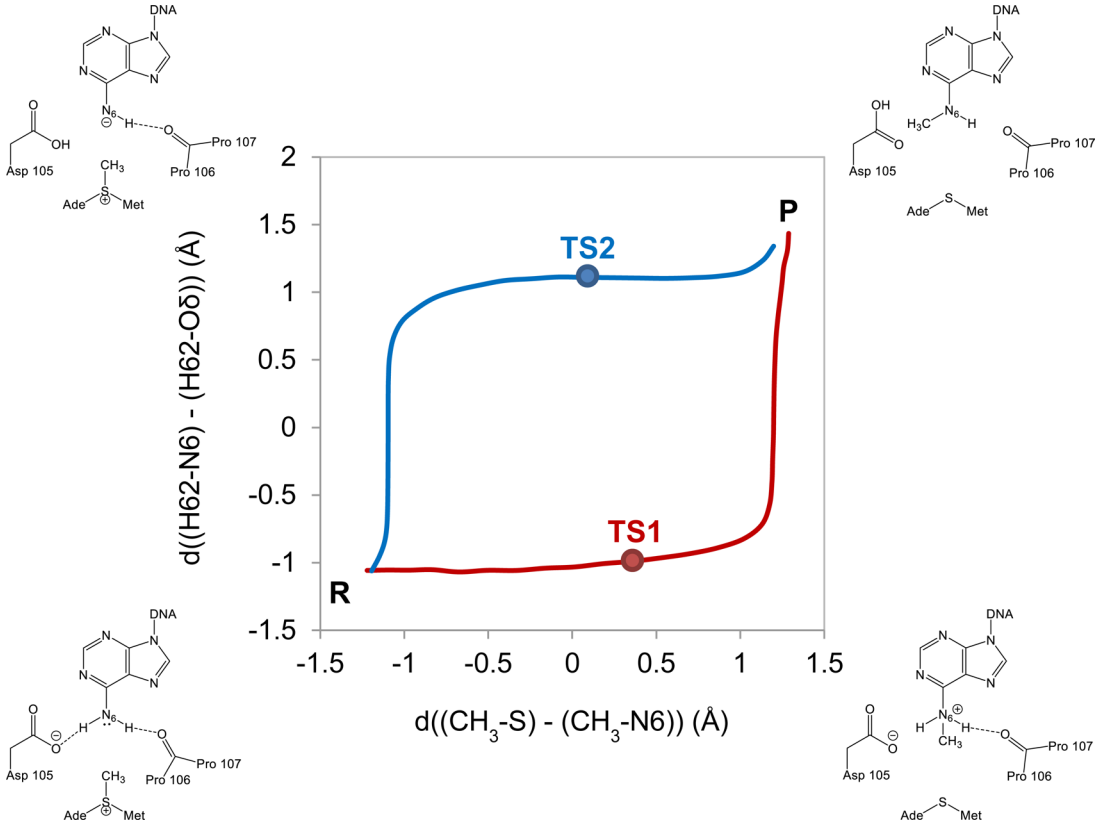


Figure 9. Minimum free energy paths obtained for two possible reaction mechanisms observed for adenine methylation in Asn105Asp M.*Taq*I. The paths are projected on the plane defined by the methyl transfer coordinate ($(CH_3–S) - (CH_3–N6)$ in the *x*-axis, in Å) and the proton transfer coordinate ((H62–N6) − (H62–Oδ) in the *y*-axis, in Å).

donor and acceptor atoms is stabilized by means of C–H···O hydrogen-bond interactions established between the methyl group and the carbonyl oxygen atoms of Asn105 and Val21 (average distances of 3.5 and 3.4 Å). Similar stabilizing interactions have been identified in other methyltransferases.[62] The reaction path following this TS leads to a positively charged methylated adenine intermediate stabilized by strengthened hydrogen-bond interactions with the carbonyl oxygen atom of Pro106 (with an averaged distance of 1.98 Å), with the Oδ atom of Asn105 (2.24 Å) and with a water molecule (2.09 Å). The free energy of this reaction intermediate is found to be 5.5 kcal·mol^{-1} below the reactants.

The second step of the reaction proceeds with the abstraction of the leftover proton from the amino group by the Asn105 residue. At the TS (TS2 in Scheme 2 and Table 1) the distances of the abstracted proton to the Oδ oxygen atom of the Asn105 residue and the N6 atom are 1.38 and 1.22 Å, respectively. The activation free energy barrier determined from the intermediate is 5.6 kcal·mol^{-1}. Then, this second TS is found to be significantly lower in free energy than the first one. The product state determined in our simulations presents a free energy 2.6 kcal·mol^{-1} below the reactant state and corresponds to the formation of *S*-adenosyl-L-homocysteine (SAH), a methylated adenine and a protonated carboxamide oxygen atom.

The possibility of a second step in which the excess proton of the A6 amino group is abstracted by a water molecule has been also investigated (red line in Figure 7). At the TS (TS3 in Scheme 2 and Table 1), the H61 proton is found at 1.34 Å of the N6 atoms and 1.27 Å of the oxygen atom of the closest water molecule. However, this TS presents a free energy 11.0 kcal·mol^{-1} higher than TS2 (when the leftover proton is abstracted by Asn105). Then we conclude that the preferred mechanism for the proton abstraction in M.*Taq*I involves the Asn residue of motif IV.

3.2.1. Asn105Asp M.TaqI Mutant. As explained in the introduction, motif IV of N-MTases is formed by Asn/Asp-Pro-Pro-Tyr (residues 105–108 in M.*Taq*I). N6-MTases of group γ (to which M.*Taq*I belongs) present an asparagine at the beginning of this motif, while N6-MTases of groups α and β contain aspartate.[17] The residues of this motif are involved in hydrogen-bonding interactions established by the flipped adenine. Both Asn and Asp are able to act as hydrogen-bond acceptors with the exocyclic amino group of adenine. The strength of the base abstracting the proton from the exocyclic amino group (weak in the case of group γ and strong for groups α and β) could affect the reaction mechanism, making the process concerted or even reversing the order between the proton and the methyl transfers.

In order to answer the question of the mechanistic implications of the presence of a strong base in the active site of N6-MTases we prepared *in silico* the M.*Taq*I Asn105Asp mutant. With this purpose we changed manually the residue, performed an energy minimization with a conjugate gradient algorithm, and ran a 20 ns long MD simulation with the MM force field using the same conditions than for the wild-type enzyme (see Section 2, Methodology). We then performed a 200 ps of AM1/MM MD simulation where the QM subsystem consisted of the cofactor SAM, the A6 nucleotide base, and the side chain of Asp105. We used the string method to obtain the

dx.doi.org/10.1021/ja5077124 | J. Am. Chem. Soc. 2014, 136, 16227–16239

MFEP on the free energy surface traced as a function of the four distances defining the methyl and proton transfer (CH_3−S, CH_3−N6, H62−N6, and H62−Oδ(Asp105)), following the procedure indicated above. In the case of the M.*Taq*I mutant we obtained two converged MFEPs that correspond to two different reaction mechanisms: one in which the methyl transfer precedes the proton transfer (red line in Figure 9) and another one where the proton transfer precedes the methyl transfer (blue line in Figure 9).

We obtained the PMFs corresponding to both mechanisms in the mutant enzyme using the path-based coordinate *s* including interpolated corrections at the M06-2X/6-311+G** level. The PMFs obtained for the two possible reaction mechanisms in the Asn105Asp mutant are presented in Figure 10, while averaged key distances characterizing the stationary

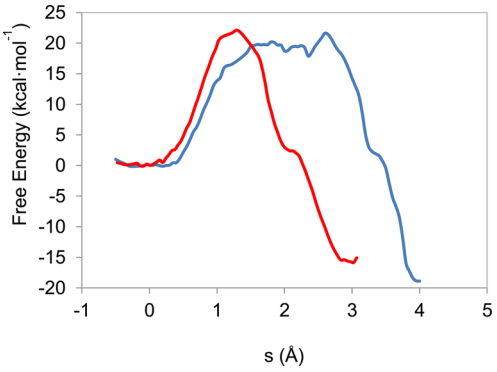


Figure 10. PMFs along the path coordinate corrected at the M06-2X/6-311+G** level. The reaction mechanism where the methyl transfer precedes proton abstraction is shown in red, while the mechanism in which proton transfer precedes methyl transfer is shown in blue.

states are provided in Table S1 of the Supporting Information. The two mechanisms present a single TS and are thus concerted. The reaction mechanism in which proton transfer from N6 to Asp105 precedes methyl transfer from SAM (blue line) presents a free energy barrier of 21.9 kcal·mol^{-1}. At the TS of this mechanism the H62 proton of the exocyclic amino group has been completely transferred from N6 to the Oδ atom of Asp105, being the averaged N6−H62 and Oδ−H62 distances equal to 2.01 ± 0.02 and 0.97 ± 0.03 Å, respectively. Instead, the methyl group is found halfway between the SAM and the amino group, presenting an averaged CH_3−S distance of 2.14 ± 0.04 Å and a CH_3−N6 distance of 2.09 ± 0.05 Å. The opposite mechanism, in which the methyl transfer from SAM to N6 precedes the abstraction of the proton of the exocyclic amino group by Asp105 (red lines in Figures 9 and 10), has a free energy barrier equal to 21.6 kcal·mol^{-1}, and thus according to our simulations, both reaction mechanisms are approximately equally probable. In the TS of this mechanism the methyl transfer is quite advanced, presenting an averaged CH_3−S distance of 2.30 ± 0.05 Å and a CH_3−N6 distance of 1.89 ± 0.06 Å, while the H61 proton remains bonded to N6 (1.03 ± 0 0.03 Å) and at hydrogen-bond distance of the Oδ atom of Asp105 (2.14 ± 0. 08 Å).

Thus, according to our simulations, the reaction mechanism could be different for different groups of N6-MTases. Those enzymes in which the first residue of motif IV is Asn (group γ)[17] do not have in the active site a strong base to abstract the excess proton from the N6 atom of adenine. In this case, the methyl transfer must be completed first to increase the acidity of the amino group forming the 6-methylammonium intermediate. This facilitates the subsequent proton abstraction. Instead, those N6-MTases belonging to groups α and β present an Asp residue in motif IV.[17] The presence of this residue in the active site opens different mechanistic possibilities. Methyl transfer is no longer a prerequisite for the proton abstraction because the presence of a stronger base makes unnecessary to diminish so much the pK_a of the proton donor. In the mutant both reaction mechanisms (methyl transfer followed by proton abstraction and proton abstraction followed by methyl transfer) take place in a single step and present very similar free energy barriers. The mechanism preferred in a particular N6-MTase belonging to groups α or β will then depend on the particular characteristics of the active site that could preferentially stabilize one of the two possible TSs described here. These diverse mechanistic routes reported for the different groups of N6-MTases could be the reason for other differences observed between them as far as different strategies could be needed to optimize the rate of the chemical step depending on the TS of the rate-limiting step. If the catalytic effect of enzymes is essentially due to the electrostatic complementarity between the active site and the TS, then increasing the efficiency of different reaction mechanisms may require different characteristics at the active sites.[63]

4. CONCLUSIONS

M.*Taq*I is a DNA methyltransferase from *Thermus aquaticus* that catalyzes the transfer of a methyl group from *S*-adenosyl-L-methionine to the N6 position of an adenine. N6-methylated adenines have been found only in prokaryotes, and then this enzyme is a potential target for the development of antibacterial drugs. A 100 ns long MD simulation has been carried out for the ternary complex formed by the cofactor, the protein, and a DNA decamer that contains the recognition sequence TCGA. The starting point for our study was the X-ray structure found in the PDB with code 1G38 that shows that the target adenine (A6) is flipped out in the active site and kept there by means of hydrogen-bond interactions to residues Asn105, Pro106, and Tyr108. These interactions are stable during the whole simulation. In addition, the DNA−protein complex is stabilized by means of several specific hydrogen-bond interactions that can also be mediated by bridging water molecules. Finally, the unpaired T15 base is stabilized not only by means of protein−DNA interactions but also by means of π-stacking interactions with G5 and T7.

An equilibrated structure was used as starting point for the exploration of the free energy landscape associated with the reaction mechanism in M.*Taq*I. QM/MM calculations were performed using the AM1 Hamiltonian corrected with single-point calculations at the M06-2X level. Our results point to a stepwise mechanism where the methyl transfer precedes the abstraction of the excess proton from the exocyclic amino group of adenine. The methyl transfer is the rate-determining step, and the calculated free energy barrier (20.1 kcal·mol^{-1}) is in very good agreement with the value derived from the experimental rate constant at 333 K (19.8 kcal·mol^{-1}). The absence of a strong base in the active site able to abstract the excess proton of the amino group determines the relative ordering between this step and the methyl transfer. Once methylated, the acidity of the exocyclic amino group increases significantly. Two possible candidates can accept the excess proton, a water molecule present in the active site and Asn105, an active site residue that is hydrogen bonded to the amino

dx.doi.org/10.1021/ja5077124 | J. Am. Chem. Soc. 2014, 136, 16227−16239

group of A6. This residue, whose basicity is enhanced by the hydrogen bonds formed through its amide group, seems the best candidate according to the activation free energies obtained for the proton abstraction process. Interestingly, the reaction mechanisms can differ if a stronger base is present in the active site. This is the case of N6-MTases of groups α and β that contains aspartate in the position equivalent to Asn105. An exploration of the reaction mechanism in the Asn105Asp M.*Taq*I mutant shows that the reaction mechanism could be concerted, with proton abstraction and methyl transfer happening in a single step. This opens the possibility to develop specific drugs to inhibit the activity of different N6-MTases. The electrostatic characteristics of the transition state and/or possible reaction intermediates are excellent molds to design drugs with high affinity for the protein. In this sense a stepwise mechanism where methylation precedes proton abstraction suggests that the active site could display high electrostatic complementarity to quaternary ammonium compounds with the nitrogen atom bearing a significant fraction of positive charge.

■ ASSOCIATED CONTENT

Supporting Information

A representation of the interactions established between cofactor SAM and the surrounding amino acids over the MD simulation is provided in Figure S1. Figure S2 shows the PES calculated at M06-2X/6-31G**/MM level for the methylation and proton abstraction by a water molecule and by the Asn105 residue. Plots of the RMSD for the enzyme and DNA backbones as well as the percentage of canonical α/γ torsions in the latter are shown in the Figure S3. Figure S4 shows the water channel connecting the hydronium ion formed in the active site with the bulk once the proton abstraction from the exocyclic N6 atom of the A6 base takes place. Figure S5 shows a transition-state structure for the proton abstraction process carried out by Asn105. PES calculated at the M06-2X/6-31G**/MM level for the methylation and proton abstraction by Asp in Asn105Asp M.*Taq*I mutant are depicted in Figure S6. Finally, Table S1 shows the most relevant averaged distances obtained for reactants, intermediate, products, and transition states for the reaction mechanisms found for the mutant Asn105Asp M.*Taq*I. Complete refs 31 and 44 are also provided. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Authors

Ignacio.tunon@uv.es
M.Teresa.Roca@uv.es

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

Authors acknowledge Dr. K. Swiderek for the implementation of AMBER force field in fDynamo. Authors also acknowledge Prof. F. J. Luque for many comments and suggestions and for hosting J. A. at the Departament de Fisicoquimica Universitat de Barcelona during a stay supported by a grant of Ministerio de Economía y Competitividad. The authors gratefully aknowledge financial support from FEDER funds and the Ministerio de Economía y Competitividad (project CTQ2012-36253-C03-03) and Generalitat Valenciana project GV/2012/053. J. A. thanks Ministerio de Economía y Competitividad for FPI fellowship, and K.Z. acknowledges a FPU fellowship of the Ministerio de Educación. The authors also acknowledge the computational facilities of the Servei d'Informàtica de la Universitat de València in the "Tirant" Supercomputer.

■ REFERENCES

(1) Baylin, S. B. *Nat. Clin. Pract. Oncol.* **2005**, *2* (Suppl 1), S4−11.
(2) Okano, M.; Bell, D. W.; Haber, D. A.; Li, E. *Cell* **1999**, *99*, 247−257.
(3) Li, E.; Beard, C.; Jaenisch, R. *Nature* **1993**, *366*, 362−365.
(4) Cheng, X., Hashimoto, H, Horton, J. R., Zhang, X In *Handbook of Epigenetics: The New Molecular and Medical Genetics*; Tollefsbol, T., Ed.; Academic Press: Oxford: 2010; p 9−24.
(5) Jurkowska, R. Z.; Jurkowski, T. P.; Jeltsch, A. *ChemBioChem* **2011**, *12*, 206−222.
(6) Jones, P. A. *Nat. Rev. Genet.* **2012**, *13*, 484−492.
(7) Law, J. A.; Jacobsen, S. E. *Nat. Rev. Genet.* **2010**, *11*, 204−220.
(8) Medina-Franco, J. L.; Caulfield, T. *Drug Discovery Today* **2011**, *16*, 418−425.
(9) Bickle, T. A.; Krüger, D. H. *Microbiol. Rev.* **1993**, *57*, 434−450.
(10) Low, D. A.; Weyand, N. J.; Mahan, M. J. *Infect. Immun.* **2001**, *69*, 7197−7204.
(11) Reisenauer, A.; Kahng, L. S.; McCollum, S.; Shapiro, L. *J. Bacteriol.* **1999**, *181*, 5135−5139.
(12) Jeltsch, A. *ChemBioChem.* **2002**, *3*, 274−293.
(13) Cheng, X. D.; Roberts, R. J. *Nucleic Acid Res.* **2001**, *29*, 3784−3795.
(14) Goedecke, K.; Pignot, M.; Goody, R. S.; Scheidig, A. J.; Weinhold, E. *Nat. Struct. Biol.* **2001**, *8*, 121−125.
(15) Ho, D. K.; Wu, J. C.; Santi, D. V.; Floss, H. G. *Arch. Biochem. Biophys.* **1991**, *284*, 264−269.
(16) Newby, Z. E. R.; Lau, E. Y.; Bruice, T. C. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 7922−7927.
(17) Malone, T.; Blumenthal, R. M.; Cheng, X. *J. Mol. Biol.* **1995**, *253*, 618−632.
(18) Mashhoon, N.; Pruss, C.; Carroll, M.; Johnson, P. H.; Reich, N. O. *J. Biomol. Screen.* **2006**, *11*, 497−510.
(19) Wibowo, F. R.; Rauch, C.; Trieb, M.; Liedl, K. R. *Biopolymers* **2005**, *79*, 128−138.
(20) Thomas, C. B.; Scavetta, R. D.; Gumport, R. I.; Churchill, M. E. *J. Biol. Chem.* **2003**, *278*, 26094−26101.
(21) Liebert, K.; Horton, J. R.; Chahar, S.; Orwick, M.; Cheng, X.; Jeltsch, A. *J. Biol. Chem.* **2007**, *282*, 22848−22855.
(22) Punekar, A. S.; Liljeruhm, J.; Shepherd, T. R.; Forster, A. C.; Selmer, M. *Nucleic Acids Res.* **2013**, *41*, 9537−9548.
(23) Yang, Z.; Horton, J. R.; Zhou, L.; Zhang, X. J.; Dong, A.; Zhang, X.; Schlagman, S. L.; Kossykh, V.; Hattman, S.; Cheng, X. *Nat. Struct. Biol.* **2003**, *10*, 849−855.
(24) Tran, P. H.; Korszun, Z. R.; Cerritelli, S.; Springhorn, S. S.; Lacks, S. A. *Structure* **1998**, *6*, 1563−1575.
(25) Schluckebier, G.; Kozak, M.; Bleimling, N.; Weinhold, E.; Saenger, W. *J. Mol. Biol.* **1997**, *265*, 56−67.
(26) Evdokimov, A. A.; Zinoviev, V. V.; Malygin, E. G.; Schlagman, S. L.; Hattman, S. *J. Biol. Chem.* **2002**, *277*, 279−286.
(27) Li, H.; Robertson, A. D.; Jensen, J. H. *Proteins: Struct., Funct., Bioinf.* **2005**, *61*, 704−721.
(28) Bas, D. C.; Rogers, D. M.; Jensen, J. H. *Proteins: Struct., Funct., Bioinf.* **2008**, *73*, 765−783.
(29) Olsson, M. H. M.; Søndergaard, C. R.; Rostkowski, M.; Jensen, J. H. *J. Chem. Theory Comput.* **2011**, *7*, 525−537.
(30) Søndergaard, C. R.; Olsson, M. H. M.; Rostkowski, M.; Jensen, J. H. *J. Chem. Theory Comput.* **2011**, *7*, 2284−2295.
(31) Case, D. A. et al. *Amber 12*; University of California: San Francisco,CA, 2012.
(32) Duan, Y.; Wu, C.; Chowdhury, S.; Lee, M. C.; Xiong, G.; Zhang, W.; Yang, R.; Cieplak, P.; Luo, R.; Lee, T.; Caldwell, J.; Wang, J.; Kollman, P. *J. Comput. Chem.* **2003**, *24*, 1999−2012.

(33) Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kalé, L.; Schulten, K. *J. Comput. Chem.* **2005**, *26*, 1781−1802.
(34) Stacklies, W.; Xia, F.; Gräter, F. *PLoS Comput. Biol.* **2009**, *5*, e1000574.
(35) Darden, T.; York, D.; Pedersen, L. *J. Chem. Phys.* **1993**, *98*, 10089−10092.
(36) Pues, H.; Bleimling, N.; Holz, B.; Wolcke, J.; Weinhold, E. *Biochemistry* **1999**, *38*, 1426−1434.
(37) Field, M. J.; Albe, M.; Bret, C.; Proust-De Martin, F.; Thomas, A. *J. Comput. Chem.* **2000**, *21*, 1088−1100.
(38) Dewar, M. J. S.; Zoebisch, E. G.; Healy, E. F.; Stewart, J. J. P. *J. Am. Chem. Soc.* **1985**, *107*, 3902−3909.
(39) Krzemińska, A.; Paneth, P.; Moliner, V.; Świderek, K. *J. Phys. Chem. B* **2014** [Online early access]. DOI: 10.1021/jp506119h (accessed October 19, 2014).
(40) Singh, U. C.; Kollman, P. A. *J. Comput. Chem.* **1986**, *7*, 718−730.
(41) Field, M. J.; Bash, P. A.; Karplus, M. *J. Comput. Chem.* **1990**, *11*, 700−733.
(42) The new charges (q_i') of the boundary atoms (i = CA, H. f. a. a. i. C., H1′ for sugars) were obtained redistributing the excess charge (Q) according to the charges of these atoms in the force field (q_i):

$$q_i' = q_i\left(1 + \frac{Q}{\sum_i q_i}\right)$$

.
(43) Zhao, Y.; Truhlar, D. *Theor. Chem. Acc.* **2008**, *120*, 215−241.
(44) Frisch, M. J. et al. *Gaussian 09*, revision C.01; Gaussian, Inc.: Wallingford CT, 2009.
(45) Maragliano, L.; Vanden-Eijnden, E. *Chem. Phys. Lett.* **2007**, *446*, 182−190.
(46) Zinovjev, K.; Martí, S.; Tuñón, I. *J. Chem. Theory Comput.* **2012**, *8*, 1795−1801.
(47) Zinovjev, K.; Ruiz-Pernía, J. J.; Tuñón, I. *J. Chem. Theory Comput.* **2013**, *9*, 3740−3749.
(48) Torrie, G. M.; Valleau, J. P. *J. Comput. Phys.* **1977**, *23*, 187−199.
(49) Roca, M.; Aranda, J.; Moliner, V.; Tuñón, I. *Curr. Opin. Chem. Biol.* **2012**, *16*, 465−471.
(50) Ruiz-Pernía, J. J.; Silla, E.; Tuñón, I.; Martí, S.; Moliner, V. *J. Phys. Chem. B* **2004**, *108*, 8427−8433.
(51) Ruiz-Pernía, J. J.; Silla, E.; Tuñón, I.; Martí, S. *J. Phys. Chem. B* **2006**, *110*, 17663−17670.
(52) Kumar, S.; Bouzida, D.; Swendsen, R. H.; Kollman, P. A.; Rosenberg, J. M. *J. Comput. Chem.* **1992**, *13*, 1011−1021.
(53) Spyrakis, F.; Cozzini, P.; Bertoli, C.; Marabotti, A.; Kellogg, G.; Mozzarelli, A. *BMC Struct. Biol.* **2007**, *7*, 1−18.
(54) Schwabe, J. W. R. *Curr. Opin. Struct. Biol.* **1997**, *7*, 126−134.
(55) Jayaram, B.; Jain, T. *Annu. Rev. Biophys. Biomol. Struct.* **2004**, *33*, 343−361.
(56) Pérez, A.; Luque, F. J.; Orozco, M. *J. Am. Chem. Soc.* **2007**, *129*, 14739−14745.
(57) Kadam, R. U.; Garg, D.; Schwartz, J.; Visini, R.; Sattler, M.; Stocker, A.; Darbre, T.; Reymond, J.-L. *ACS Chem. Biol.* **2013**, *8*, 1925−1930.
(58) Schluckebier, G.; Labahn, J.; Granzin, J.; Saenger, W. *Biol. Chem.* **1998**, *379*, 389−400.
(59) Roca, M.; Moliner, V.; Ruiz-Pernía, J. J.; Silla, E.; Tuñón, I. *J. Phys. Chem. A* **2006**, *110*, 503−509.
(60) Castillo, R.; Roca, M.; Soriano, A.; Moliner, V.; Tuñón, I. *J. Phys. Chem. B* **2008**, *112*, 529−534.
(61) Soriano, A.; Castillo, R.; Christov, C.; Andrés, J.; Moliner, V.; Tuñón, I. *Biochemistry* **2006**, *45*, 14917−14925.
(62) Horowitz, S.; Dirk, L. M. A.; Yesselman, J. D.; Nimtz, J. S.; Adhikari, U.; Mehl, R. A.; Scheiner, S.; Houtz, R. L.; Al-Hashimi, H. M.; Trievel, R. C. *J. Am. Chem. Soc.* **2013**, *135*, 15536−15548.
(63) Aranda, J.; Roca, M.; Tuñón, I. *Org. Biomol. Chem.* **2012**, *10*, 5395−5400.