Supporting Information

Unraveling the Reaction Mechanism of Enzymatic C5-Cytosine Methylation of DNA. A Combined Molecular Dynamics and QM/MM Study of wild type and Gln119 variant

Juan Aranda[1], Kirill Zinovjev[1], Katarzyna Świderek[1,2], Maite Roca[1*], Iñaki Tuñón[1*]

[1]Departament de Química Física, Universitat de València, 46100 Burjassot, Spain

[2]Institute of Applied Radiation Chemistry, Lodz University of Technology, 90-924 Lodz, (Poland)

Ignacio.tunon@uv.es

M.Teresa.Roca@uv.es

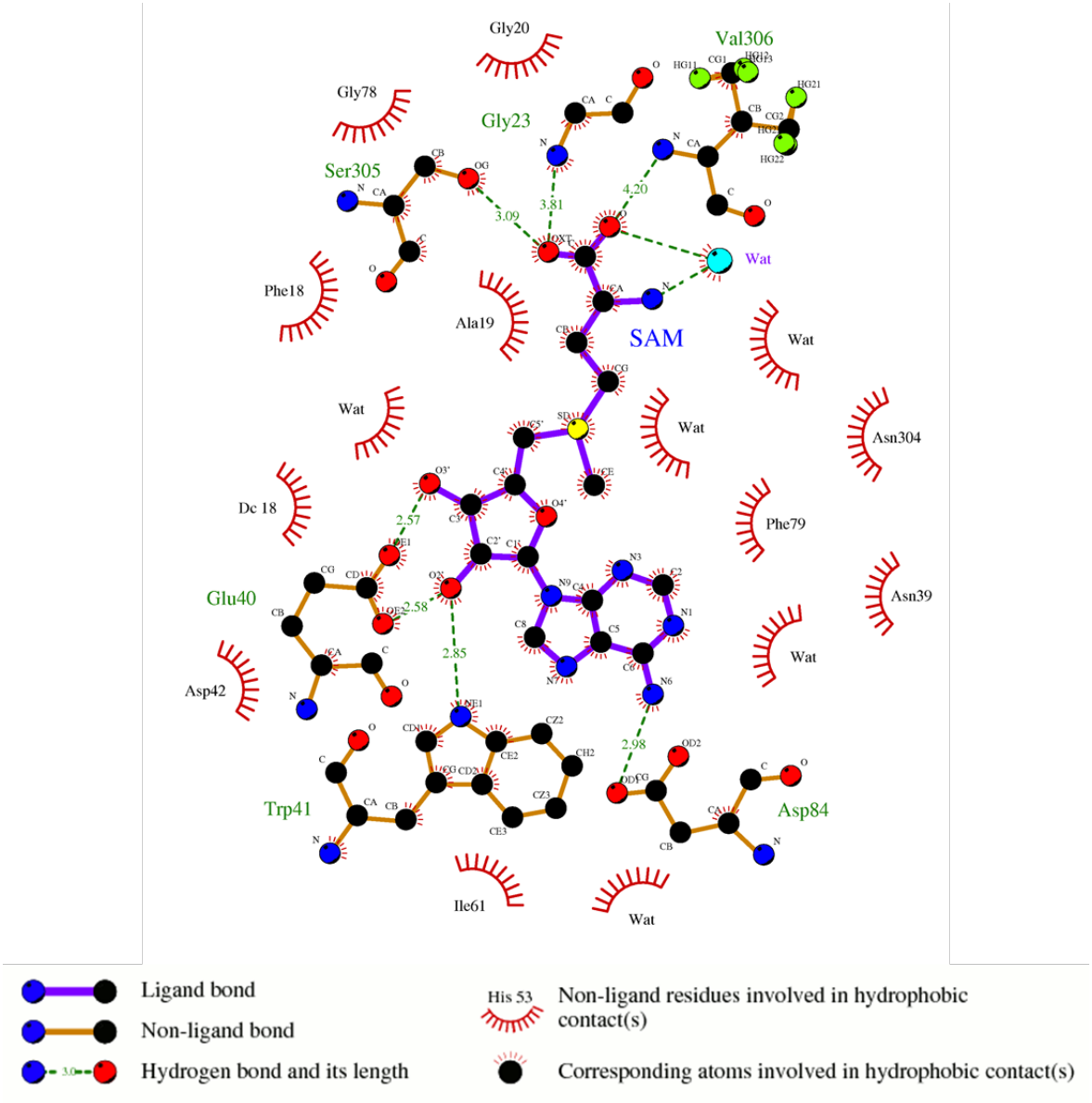
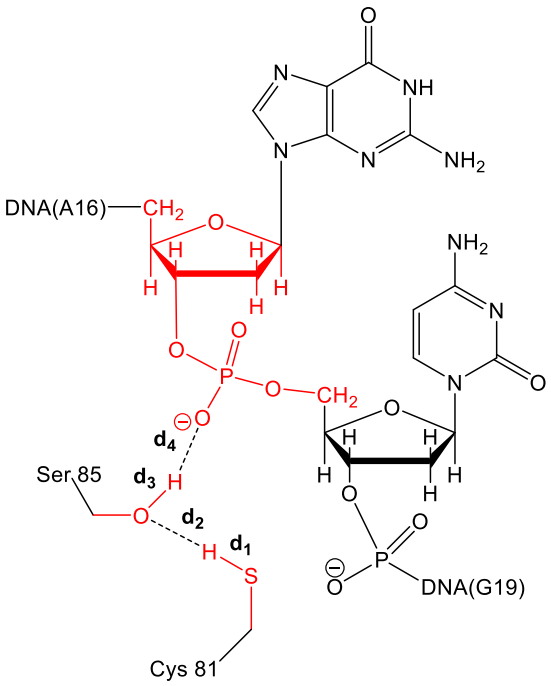


Figure S1. Schematic representation of the interactions established between SAM and surrounding amino acids in the PDB structure 2HR1. Averaged distances obtained during the MD simulation of the system with unprotonated Cys81 are given (Å).

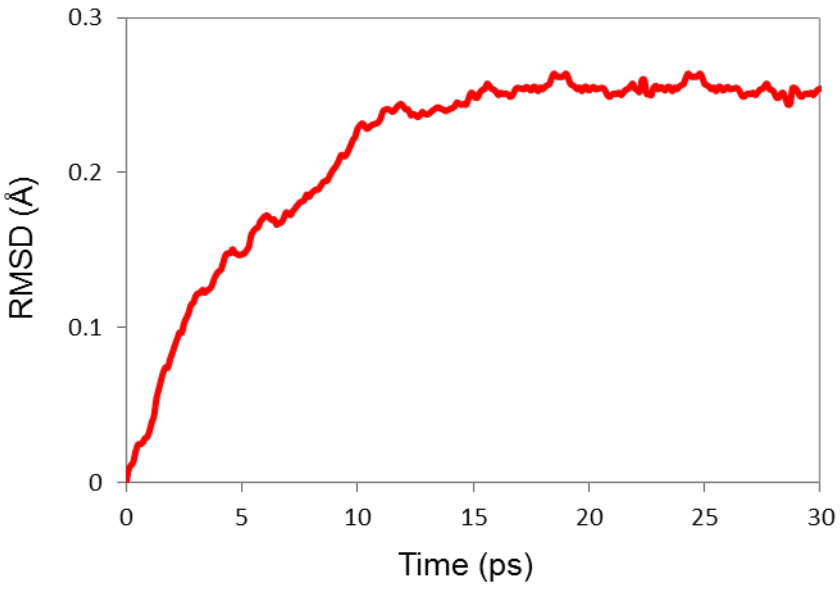
QM subsystems and active space selected to trace the MFEPs with the string method

A) Cys81 deprotonation step.

a)

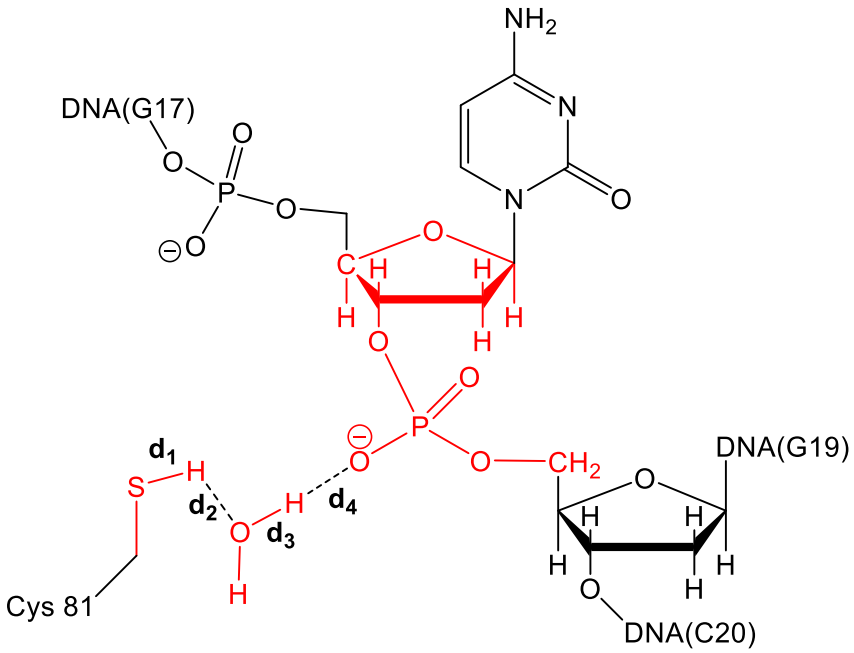


b)

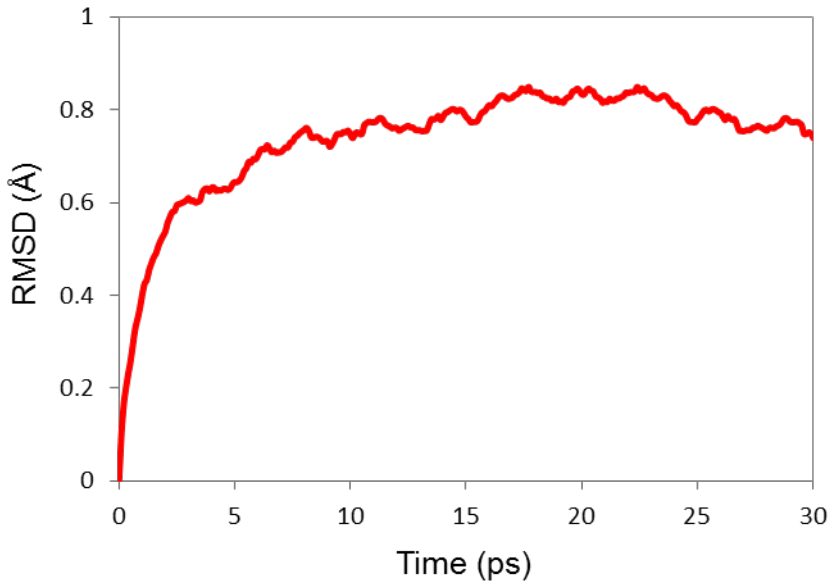


Scheme S1. QM subsystem for the Cys81 deprotonation mediated by the Ser85 residue highlighted in red. The active space selected to trace the MFEP was formed by distances d1, d2, d3 and d4. b) RMSD evolution of the nodes during the application of the string method.

a)



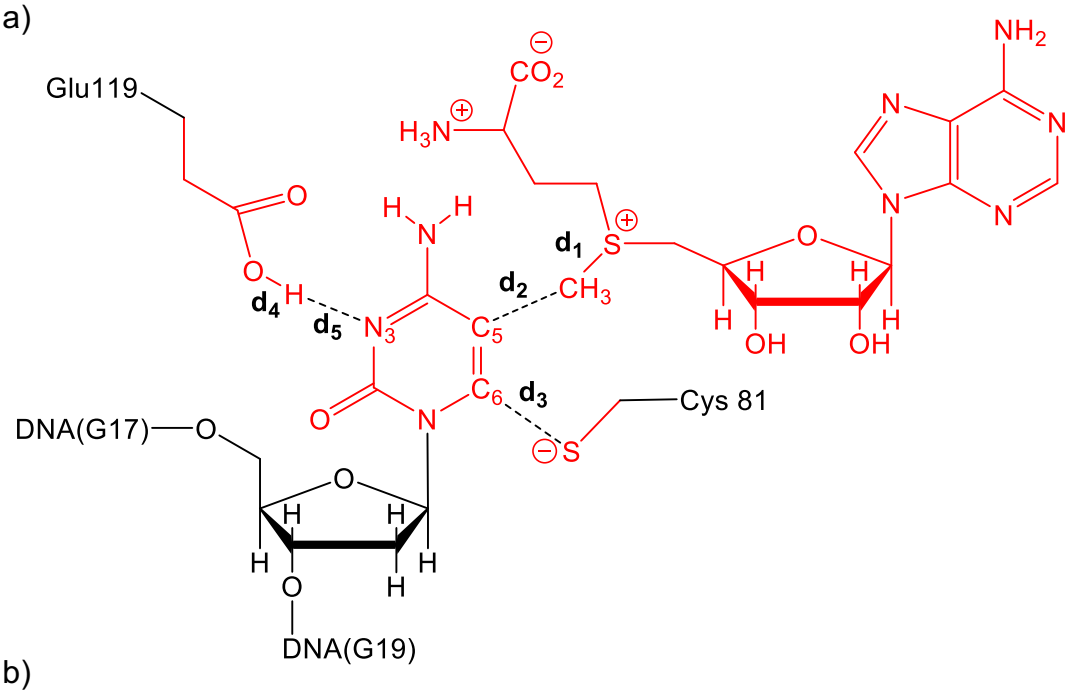
b)



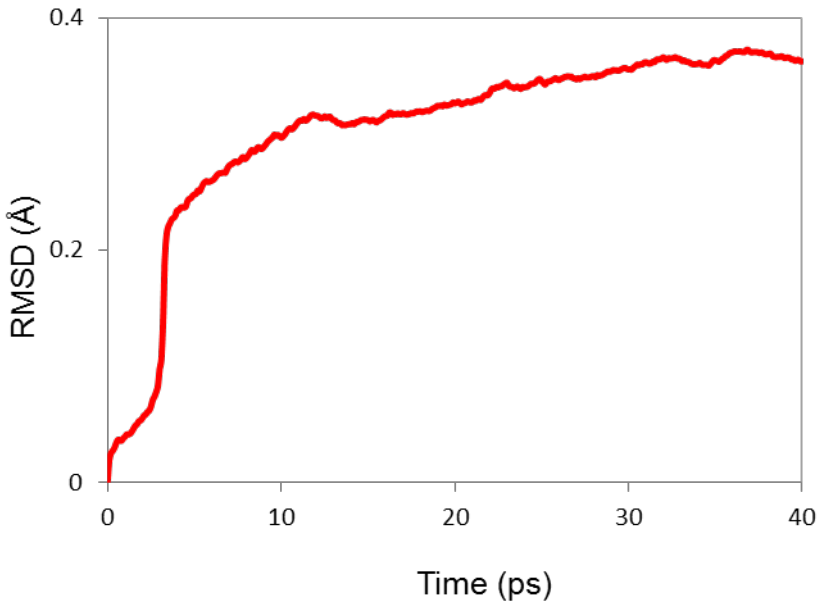
Scheme S2. a) QM subsystem for the Cys81 deprotonation mediated by a water molecule highlighted in red. The active space selected to trace the MFEP was formed by distances d1, d2, d3 and d4. b) RMSD evolution of the nodes during the application of the string method.

B) Methylation step

a)



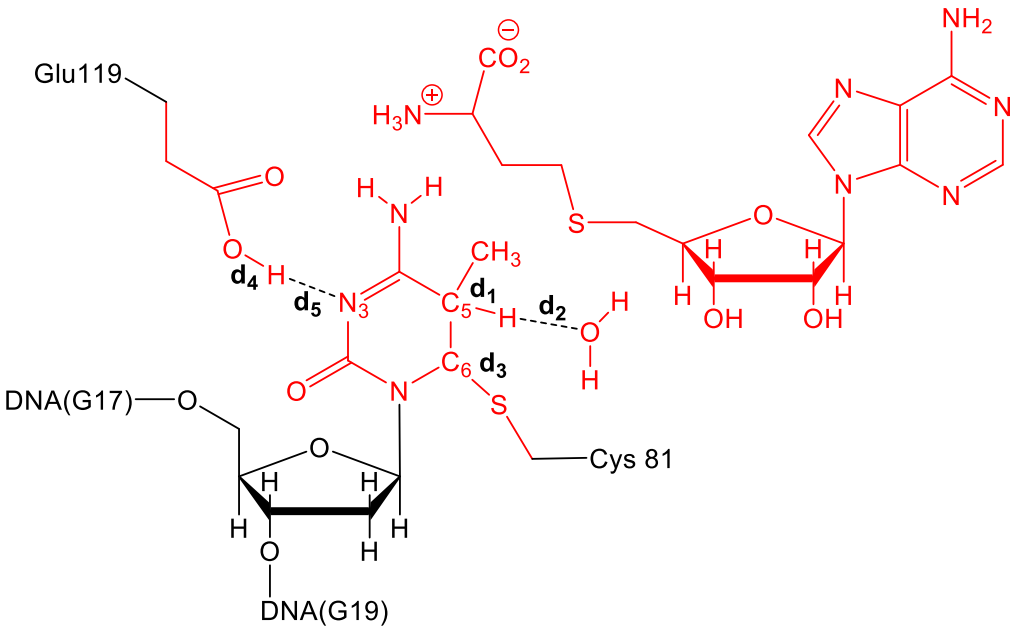
b)



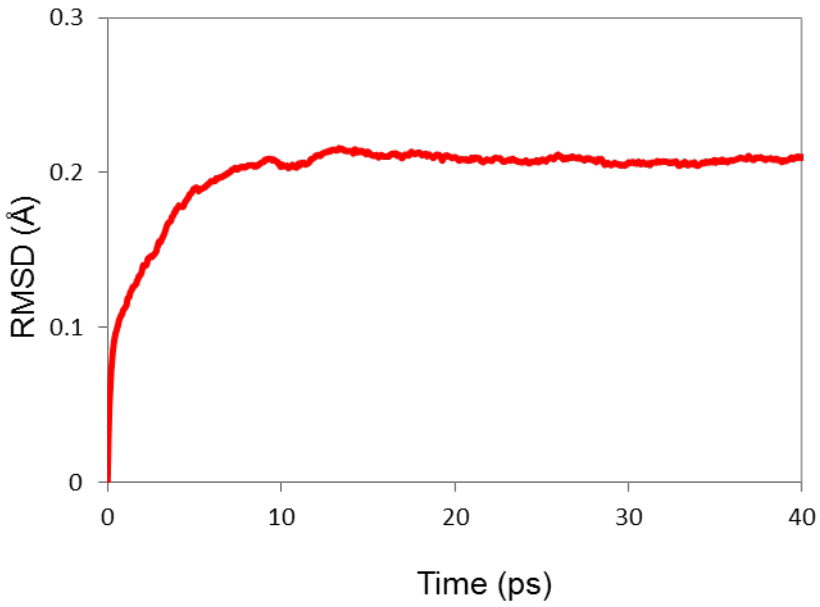
Scheme S3. a) QM subsystem for the methylation step highlighted in red. The active space selected to trace the MFEP was formed by distances d1, d2, d3, d4 and d5. b) RMSD evolution of the nodes during the application of the string method.

C) β-Elimination step

a)

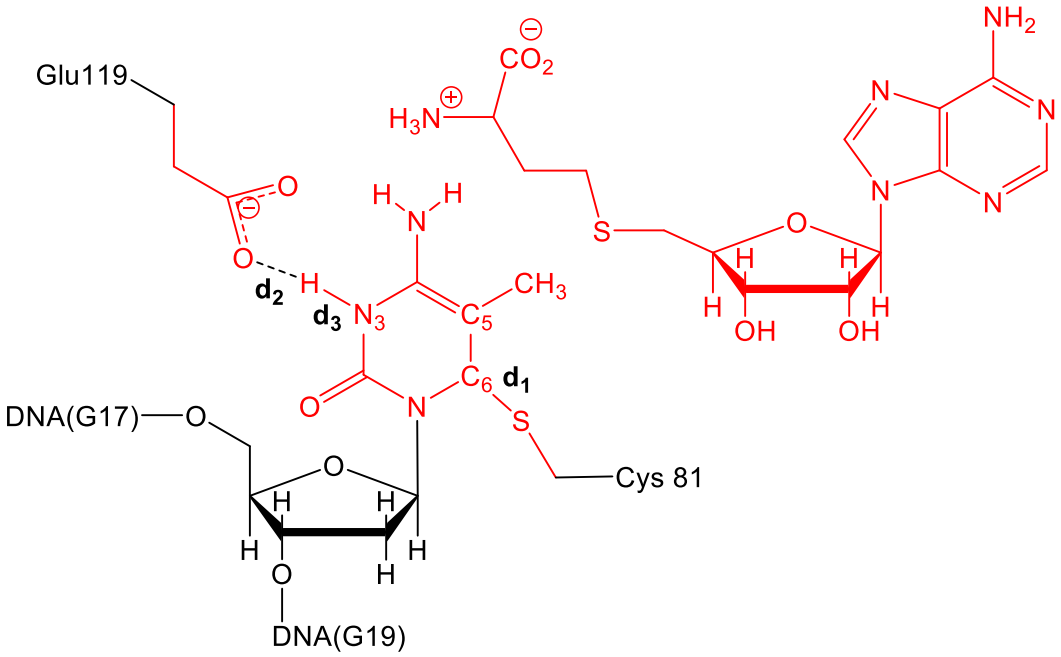


b)

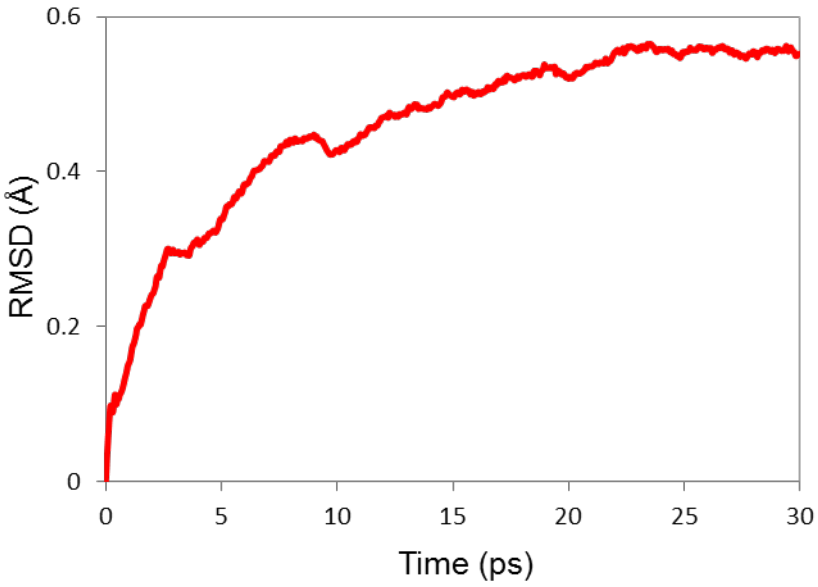


Scheme S4. a) QM subsystem for the β-elimination step where Glu119 residue donates a proton to the N_3 atom of cytosine and a water molecule abstracts the leftover proton H_5, highlighted in red. The active space selected to trace the MFEP was formed by distances d1, d2, d3, d4 and d5. b) RMSD evolution of the nodes during the application of the string method.

a)



b)



Scheme S5. a) QM subsystem highlighted in red for the elimination of Cys81 from the cytosine base and deprotonation of the N_3 atom of cytosine by Glu119. The active space selected to trace the MFEP was formed by distances d1, d2 and d3. b) RMSD evolution of the nodes during the application of the string method.

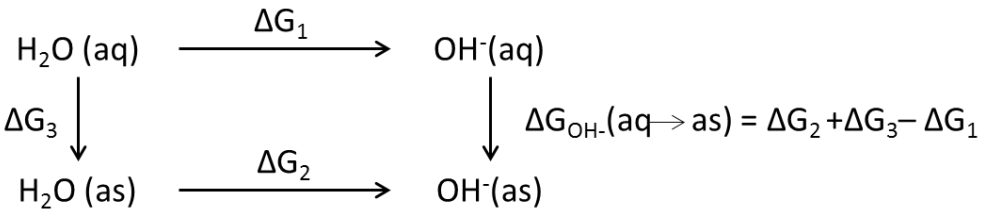
Free Energy Perturbation Protocol

We have estimated the free energy cost of moving a hydroxide anion from bulk water (aqueous solution, *aq*) to the active site (*as*) as the free energy difference of transforming one water molecule into that anion in both environments (see Scheme S6). To perform this we carried out two simulations. One where a water molecule was transformed into a hydroxide ion in a periodic box of water molecules, therefore calculating ΔG_1. And another one where a water molecule remains inside the active site of the solvated enzyme-DNA-SAM system, therefore calculating ΔG_2. Subtracting ΔG_1 from ΔG_2 we were able to obtain the free energy cost for the overall process (ΔG_{OH^-} (aq→as)) (see Scheme S6).

For the process in aqueous solution a periodic box of size ~100x100x100 Å was employed (similar to that of the enzymatic system). Simulations were performed with a time step of 1 fs with the NAMD program employing the AMBER force field. Periodic boundary conditions were applied using the particle mesh Ewald method with a cutoff of 12 Å in the real part of the sum. Short range non-bonded interactions were calculated at every step using a smooth switching function with a cutoff from 10 to 12 Å. Simulations were carried out in the NVT ensemble. Langevin-Verlet dynamics were performed using a damping coefficient of 10 ps^{-1} at a target temperature of 300 K. Initial structures for each window of the FEP simulation were taken from a previous FEP simulation carried out in the forward direction with 100 windows and with a total simulation time of 11 ns.

For the process in the active site, in order to keep the transformed water molecule into the active site, a semi-harmonic potential restraint of 100 $kcal \cdot mol^{-1}$ was applied when the distance between the oxygen atom and the C5 atom of the flipped out cytosine was larger than 5 Å. For both transformations, in aqueous solution (*aq*) and in the active site (*as*), a soft-core potential[1] and a dual topology approach[2,3] were used. The force field was smoothly changed from that of a water molecule to that of the hydroxide anion using a coupling parameter λ. In both cases we carried out a total of 100 windows between λ=0 and λ=1 and for each window 110 ps of MD simulation were performed, being the total sum of the simulation of 22 ns for each transformation. Forward and

backward paths were simulated to evaluate the hysteresis, with full convergence (the differences were in both cases smaller than 0.1 kcal·mol^{-1}). Initial structures for each window of the FEP simulation were taken from a previous FEP simulation carried out in the forward direction with 100 windows and with a total simulation time of 11 ns. The statistical error was calculated through the parseFEP plugin for VMD.[4]



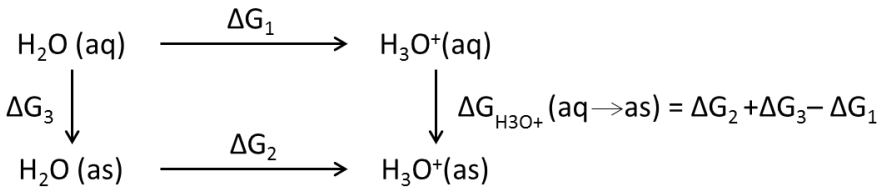
Scheme S6. Thermodynamic cycle for the calculation of the free energy cost of moving a hydroxide anion from the bulk (*aq*) to the active site (*as*) of M.*Hha*I using Free Energy Perturbation.

For the calculation of ΔG_3 we considered that the chemical potentials of water molecules in the bulk and in the active site are equal. Then the free energy difference was evaluated from the ratio between the concentration of water in the bull (~ 55 M) and in the active site:

$$\Delta G_3 = RT \cdot \ln \frac{[H_2O]_{aq}}{[H_2O]_{as}}$$

The concentration of water in the active site was evaluated from the average number of water molecules observed during the 100 ns simulation (6.0) and its average volume (367 Å^3, as determined using the program CASTp).[5] This resulted in a concentration in the active site of 27.2 M and then ΔG_3 at 300 K is equal to 0.4 kcal·mol^{-1}.

To estimate the free energy of moving a hydronium cation from the active site to the bulk we follow the same protocol and scheme as the one described above. The thermodynamic cycle is now as follows:



Scheme S7. Thermodynamic cycle for the calculation of the free energy cost of moving a hydronium cation from the bulk (*aq*) to the active site (*as*) of M.*Hha*I using Free Energy Perturbation.

MD Simulation Analysis

A) Cys81 Desprotonated system

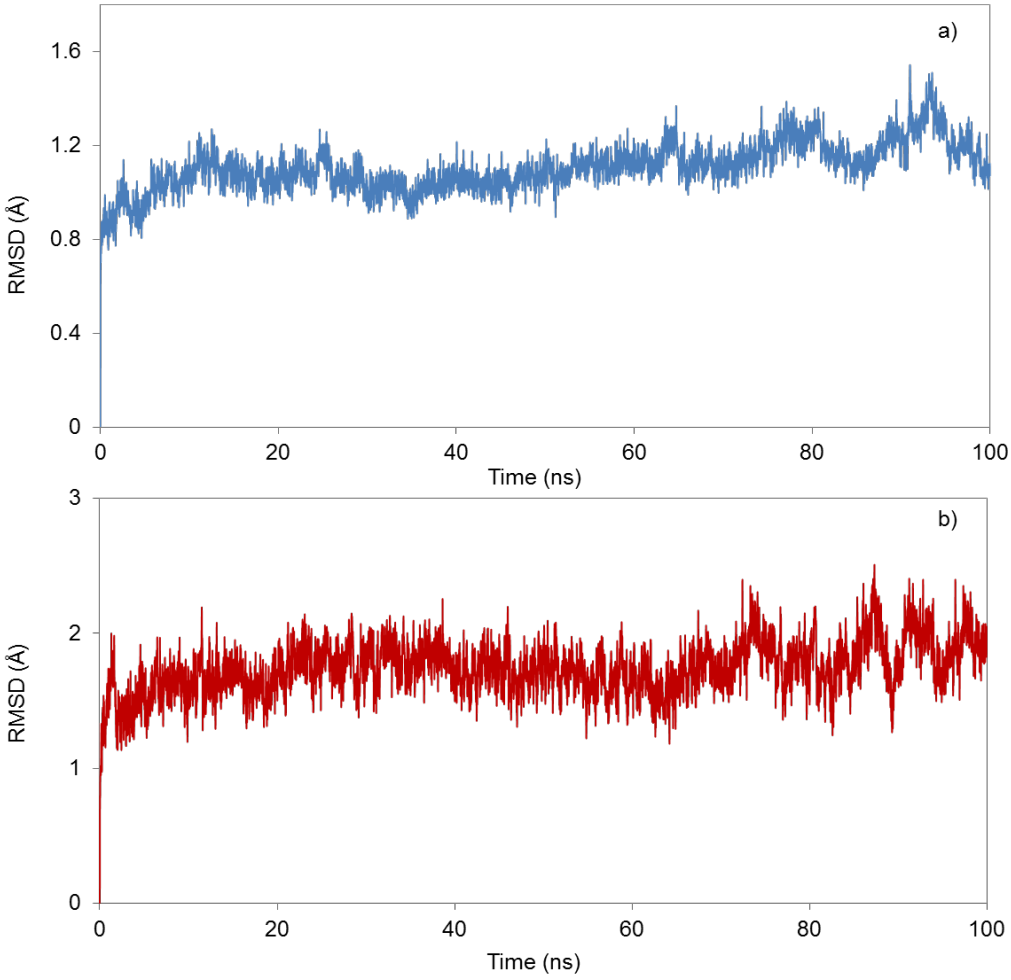


Figure S2. Root-Mean-Square Deviations calculated for for the backbone atoms of the protein (a) and the DNA (b) chains, relative to the structure obtained after equilibration for the Cys81 deprotonated system.

Protein-DNA interactions

There are two kinds of hydrogen bond interactions that can be established between M.*Hha*I and the DNA dodecamer: direct hydrogen bonding between donor/acceptor atoms of protein and DNA, and hydrogen bonds where a water molecule bridges between donor/acceptor atoms of these two entities. These water-bridged hydrogen bonds have been reported to be involved in the recognition and the stabilization of protein-DNA complexes.[6] In our analysis the geometrical criteria used to identify a hydrogen bond interaction were a donor-acceptor distance smaller than 3.2 Å and a donor-hydrogen-acceptor angle larger than 140 degrees. Figure S3 shows the number of direct and water-bridged protein-DNA hydrogen bonds during our simulation. The total number of hydrogen bonds was 59.6 ± 5.2, being the average values for direct protein-DNA and water-bridged protein-DNA hydrogen bonds of 30.7 ± 2.6 and 28.9 ± 4.4, respectively. The number of direct protein-DNA interactions presents a lower standard deviation, showing that these interactions are rather specific. The number of water-bridged protein-DNA interactions displays a larger standard deviation, reflecting the mobility of these water molecules at the protein-DNA interface.

a)

b)

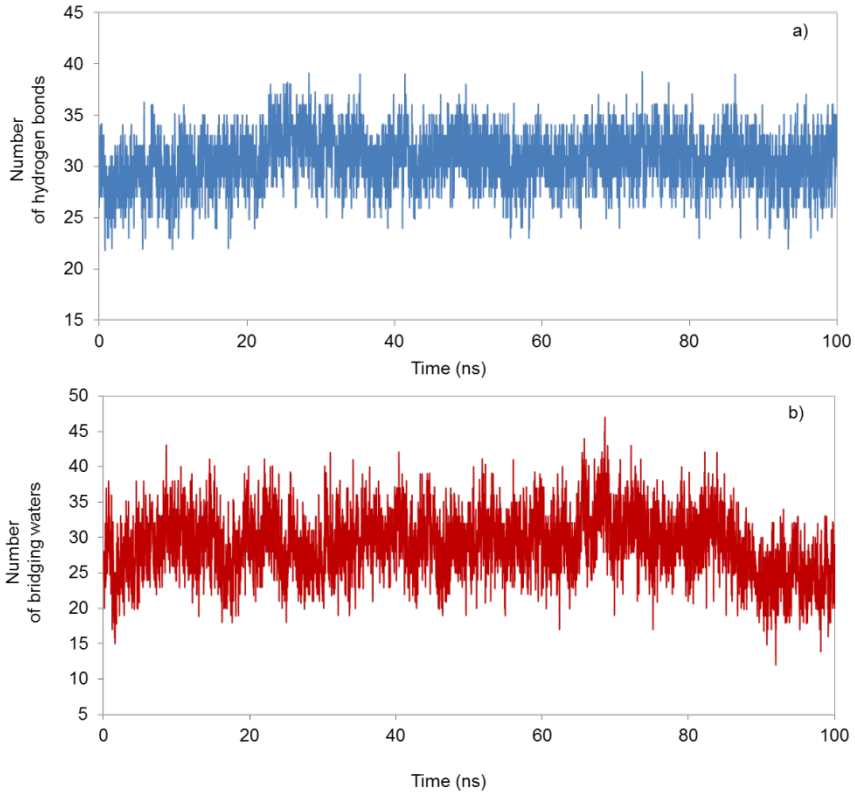
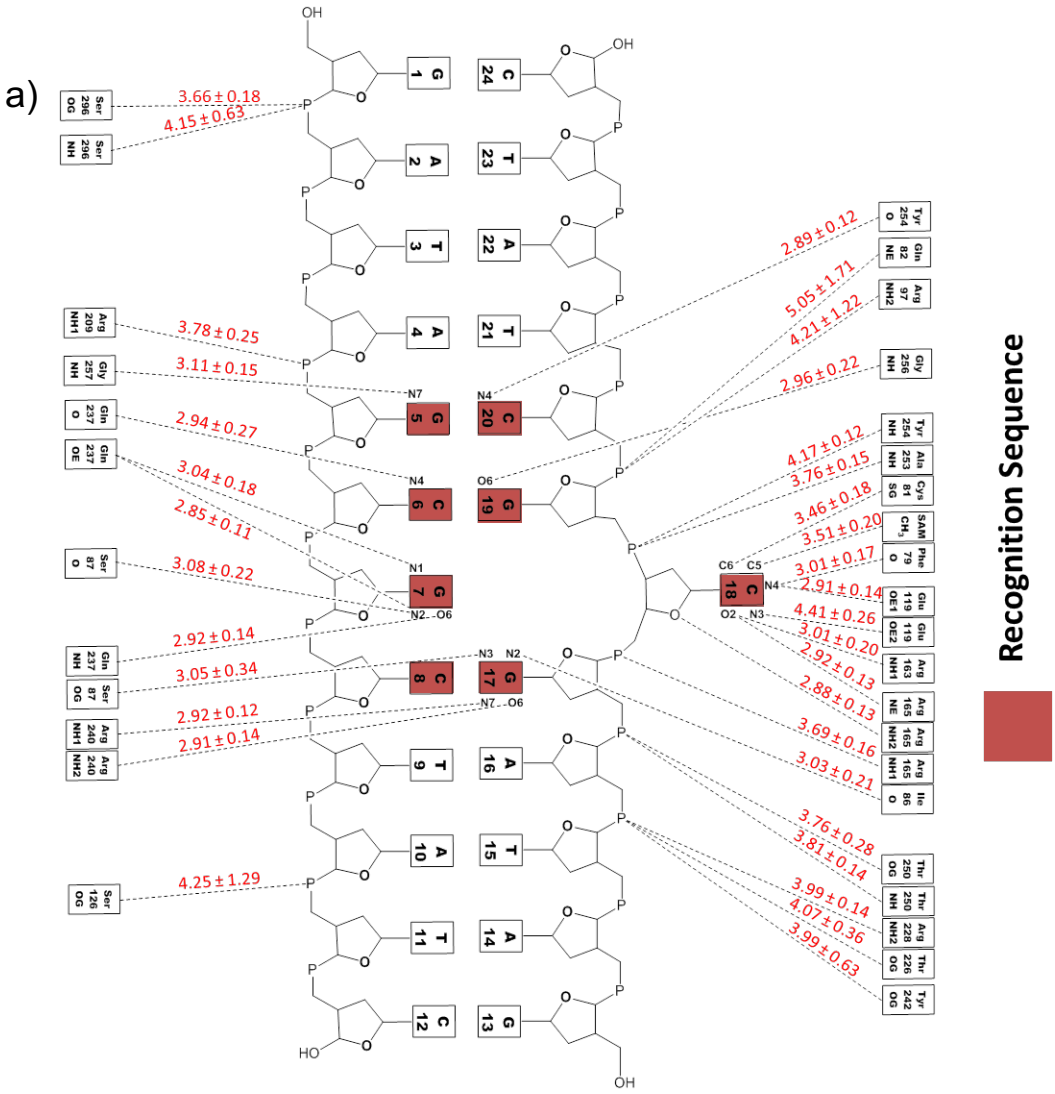


Figure S3. Number of direct DNA-protein hydrogen bonds (a) and number of water mediated protein-DNA hydrogen bonds (b) observed during the 100 ns simulation of the protein-DNA complex.


a)
Recognition Sequence
3.66 ± 0.18
4.15 ± 0.63
3.78 ± 0.25
3.11 ± 0.15
2.94 ± 0.27
3.04 ± 0.18
2.85 ± 0.11
3.08 ± 0.22
2.92 ± 0.14
3.05 ± 0.34
2.92 ± 0.12
2.91 ± 0.14
4.25 ± 1.29
2.89 ± 0.12
5.05 ± 1.71
4.21 ± 1.22
2.96 ± 0.22
4.17 ± 0.12
3.76 ± 0.15
3.46 ± 0.18
3.51 ± 0.20
3.01 ± 0.17
2.91 ± 0.14
4.41 ± 0.26
3.01 ± 0.20
2.92 ± 0.13
2.88 ± 0.13
3.69 ± 0.16
3.03 ± 0.21
3.76 ± 0.28
3.81 ± 0.14
3.99 ± 0.14
4.07 ± 0.36
3.99 ± 0.63

b)

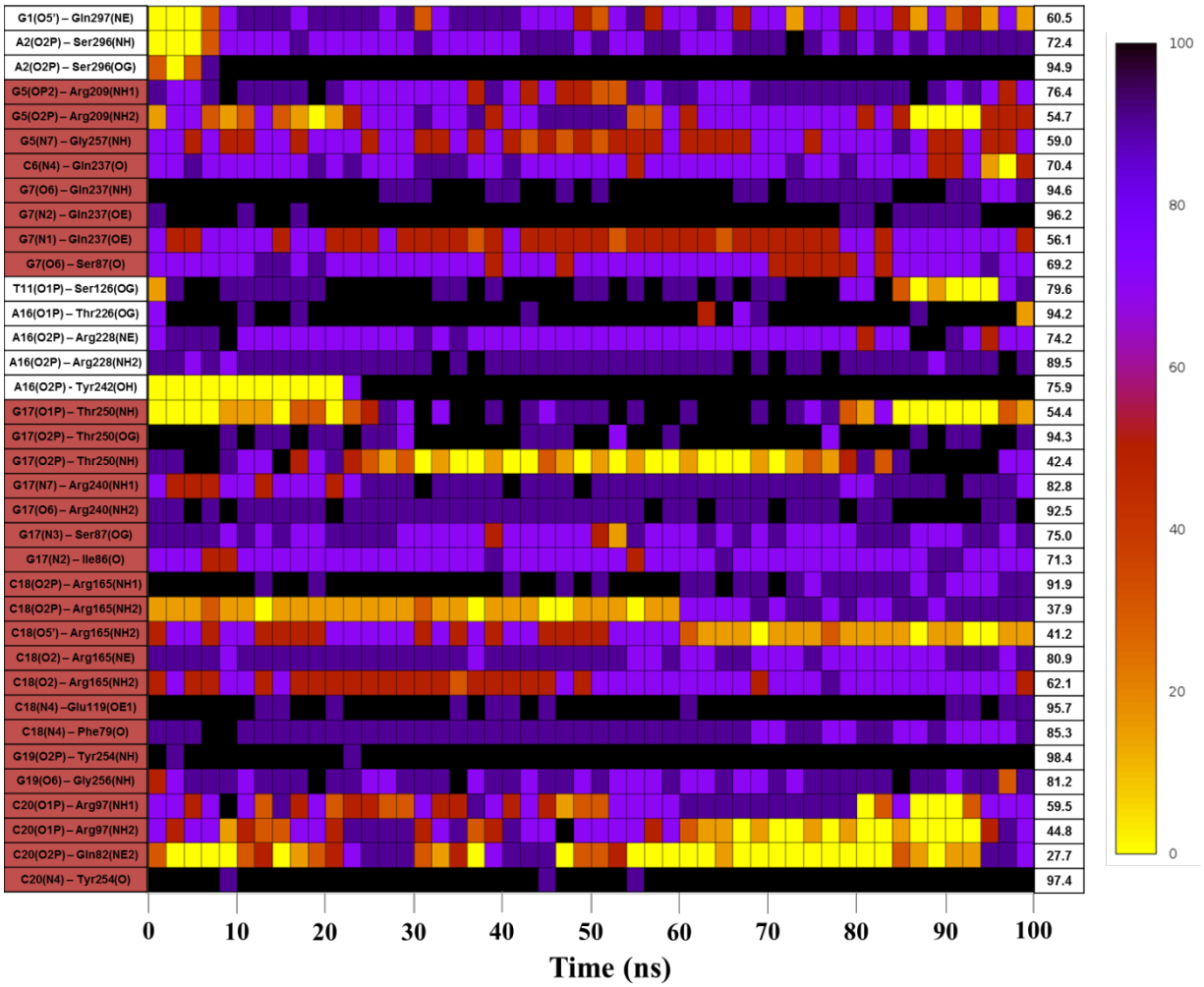


Figure S4. Analysis of important protein-DNA hydrogen bond interactions. a) Average distances and their standard deviations (in Å) between heavy atoms involved in hydrogen bonds. b) Plot of the occupancies for direct protein-DNA hydrogen bonds during the simulation. We only plot those hydrogen bonds observed for at least 20% of the simulation time. Each block of the grid represents occupancies averaged during 2 ns. Cells highlighted in red represent hydrogen bonds formed with the recognition sequence. The last column on the right displays the total occupancy numbers determined from the whole simulation.

The most important direct protein-DNA interactions observed during our MD simulation of 100 ns are gathered in Figure S4. Averaged distances and occupancy rates for direct hydrogen bond interactions are depicted in Figure S4a and S4b, respectively. Shortest average distances and largest occupancy rates correspond to the recognition sequence (highlighted in red in Figure S4). The DNA dodecamer has 22 negative charges which are stabilized by a total of 10 hydrogen bonds involving positively charged residues of the protein. Direct

hydrogen bond interactions established with Arg97, Arg163, Arg165, Arg209, Arg228 and Arg240 display average occupancies larger than 20% while interactions formed with Lys89, Lys122, Lys162 and Lys234 are broken and formed during the simulation displaying smaller average occupancies.

DNA MTases have to stabilize not only the flipped base to be methylated (C18 in the current system), but also the orphan base that remains in the DNA helix (G7) and the two nucleic bases that lose their π-stacking partner (G17 and G19). In our simulations the unpaired guanine base (G7) is hydrogen bonded to Gln237, as seen also in the X-ray structure.[7] An experimental study showed that the Gln237Trp M.*Hha*I mutant has a much smaller catalytic ability than the wild type enzyme, which was attributed to the inability of the mutant enzyme to stabilize the flipped out state of DNA.[8] In our simulation Gln237 establishes two hydrogen bonds with G7: one between the Gln237 O_ε atom and the N_2 atom of G7, and another one between Gln237 backbone nitrogen atom and the O_6 atom of G7 (see Figure S4a). These interactions display short average distances and occupancy rates of 96.2 and 94.6 %, respectively (see Figure S4b). We also found that the unpaired G7 base can be hydrogen bonded to Ser87 during the simulation. This hydrogen bond interaction involves the O_6 atom of the G7 nucleic base and the Ser87 backbone oxygen atom, displaying an occupancy rate of 69.2%.

Interactions between the protein and G17/G19 basis are important both for the recognition of the specific DNA sequence and for the stabilization of the DNA structure once the target cytosine flips out leaving these bases without one π-stacking partner. G17 is doubly hydrogen bonded to Arg240 during the whole simulation (displaying occupancy rates of 92.5 % for the interaction between the N_7 atom of G17 and the $N_{\eta1}$ atom of Arg240 and 82.8 % for the interaction involving the O_6 atom of G17 and $N_{\eta2}$ of the Arg240). This residue has been proposed to be involved in substrate binding, loop closure and stabilization of DNA towards catalysis.[9] A previous study devoted to the calculation of free energy profiles for the base flipping event showed that the largest contribution to DNA binding was the interaction between Arg240 and G17.[10] G17 is also hydrogen bonded to Ile86 during the simulation displaying high occupancy numbers (71.3 %). Ile86 has been also reported in experimental studies[8,11] to

be hydrogen bonded to G17, playing an important role for the stabilization of the extrahelical cytosine. The G19 nucleic base also interacts with Gly256. In our simulation this interaction displays a high occupancy rate (81.2 %).

Previous structural analysis have shown that closure of the mobile loop formed by residues 80-100 (see Figure 1) is the main difference between the structure of the protein with and without DNA.[9] This loop has a crucial role stabilizing the Michaelis complex. All the interactions concerning the residues of this loop are conserved in our simulations, with high occupancies rates and small average distances. In particular, the largest occupancy rates for interactions stablished between DNA and residues belonging to this loop correspond to Cys81, Gln82, Ile86, Ser87 and Arg97 (see Figure S4).

DNA Structure

The binding process and the insertion of C18 in the catalytic pocket of M.*Hha*I provoke structural changes into the canonical structure of DNA. In spite of that, the hydrogen bond interactions between DNA base pairs are not interrupted. As explained in the main text, protein accommodates the flipped out cytosine base stabilizing the interactions lost by G7 with two highly conserved hydrogen bonds with the residue Gln237 and one hydrogen bond with Ser87. G19 and G17 which lose their π-stacking partner are also stabilized with protein interactions. Figure S3 shows the average number of hydrogen bonds between the base pairs in the DNA duplex. Hydrogen bonds are maintained during the simulation with values close to the canonical forms, with the exception of the G7-C18 pair containing the flipped out cytosine and one of the base pair extremes (C12-G13), which displays a lower number of hydrogen bonds and higher standard deviation. The average values of the number of canonical hydrogen bonds for the rest of the nucleic base pairs are 1.8 for the T-A pairs and 2.9 for the C-G pairs.

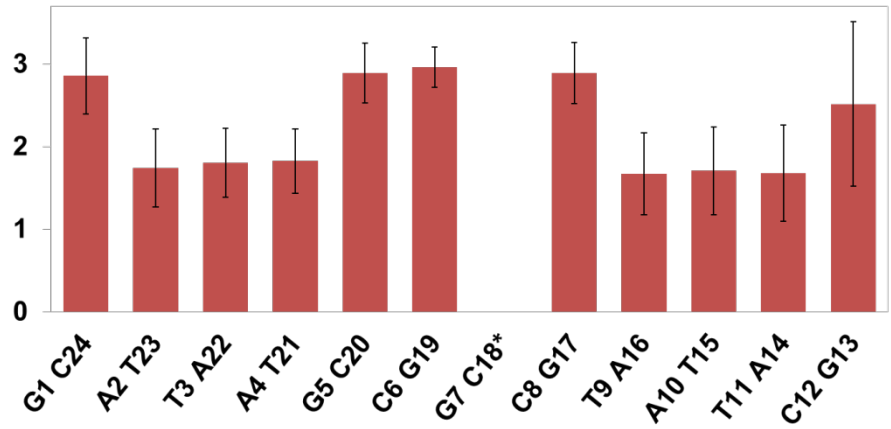


Figure S5. Average number of canonical hydrogen bond interactions established between DNA base pairs for the Cys81 deprotonated system. Standard deviations are also shown as vertical lines.

B) Cys81 Protonated system

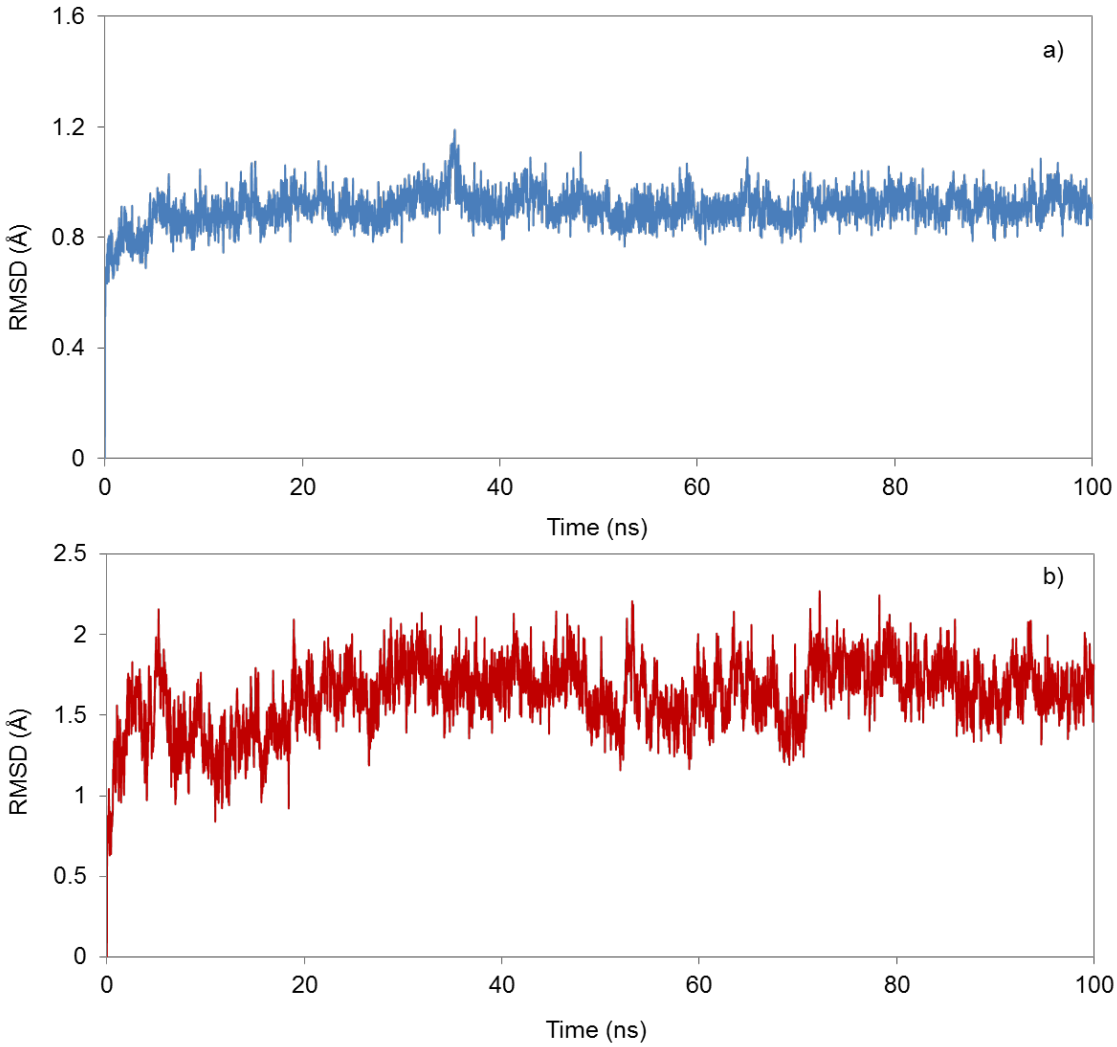


Figure S6. Root-Mean-Square Deviations calculated for the backbone atoms of the protein (a) and the DNA (b) chains relative to the structure obtained after equilibration for the Cys81 protonated system.

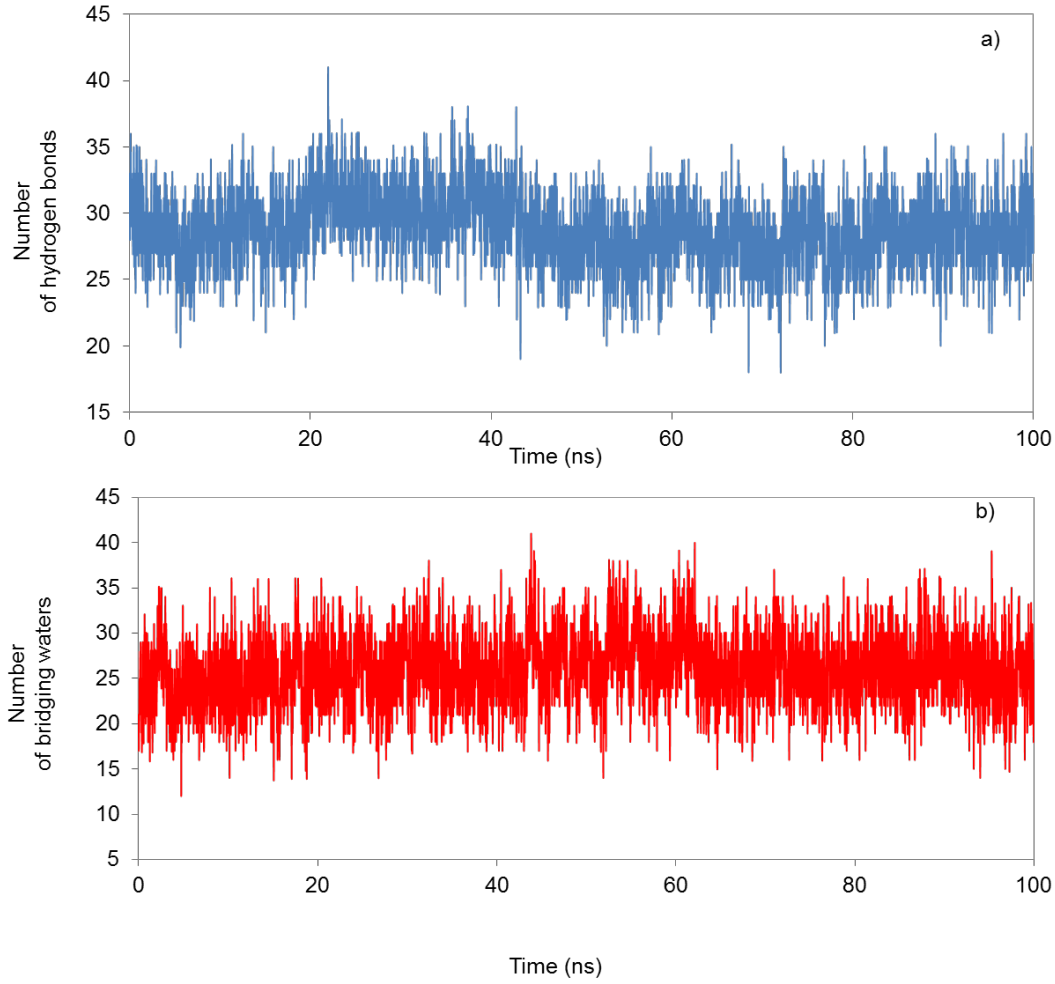


Figure S7. Number of direct DNA-protein hydrogen bonds (a) and number of water-mediated protein-DNA hydrogen bonds (b) observed during the 100 ns simulation of the protein-DNA complex for the Cys81 protonated system.

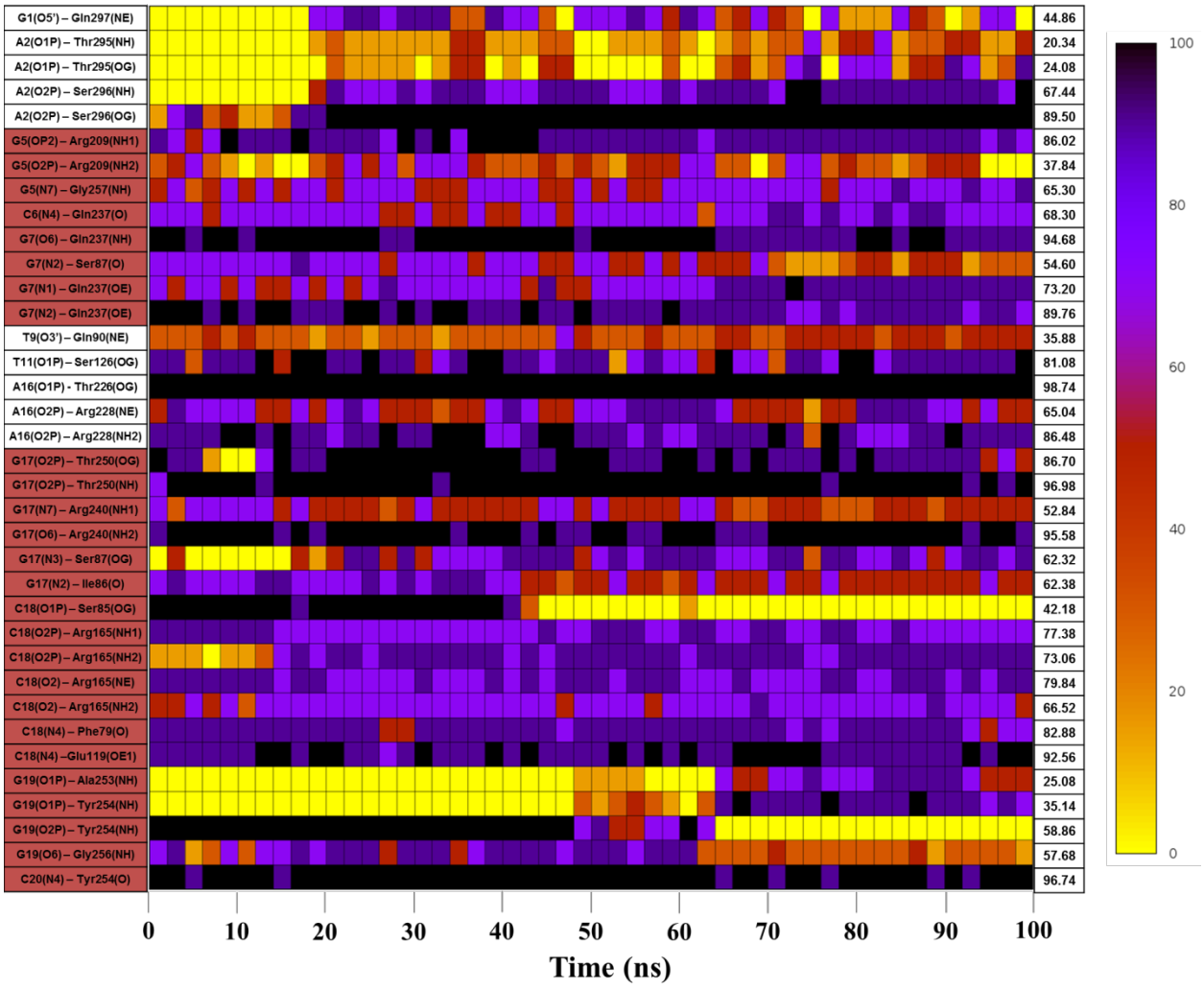


Figure S8. Analysis of important protein-DNA hydrogen bond interactions for the Cys81 protonated system. Each block of the grid represents occupancies for direct protein-DNA hydrogen bonds averaged during 2 ns. We only plot those hydrogen bonds observed for at least 20% of the simulation time. Cells in red represent hydrogen bonds formed with the recognition sequence. The last column on the right side displays the total occupancy numbers determined from the whole simulation.

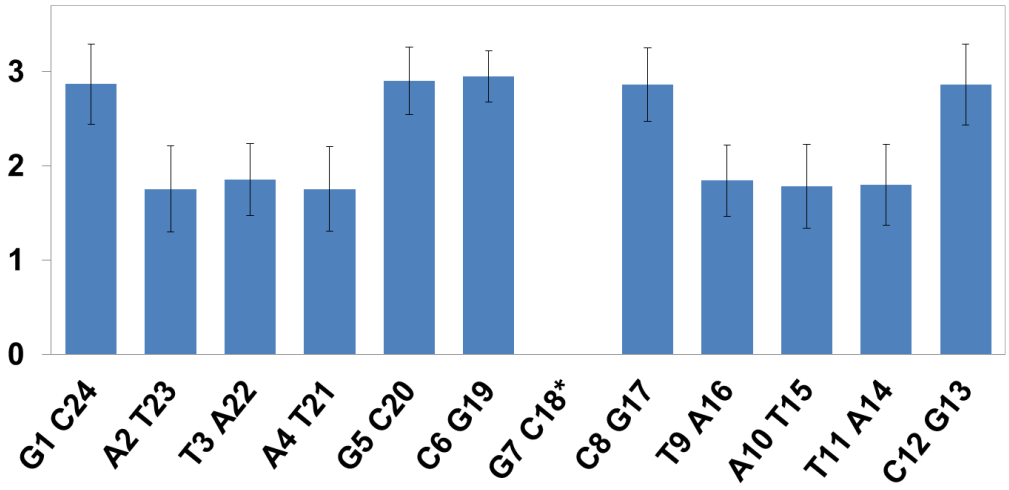


Figure S9. Average number of canonical hydrogen bond interactions established between DNA base pairs for the Cys81 protonated system. Standard deviations are shown as vertical lines.

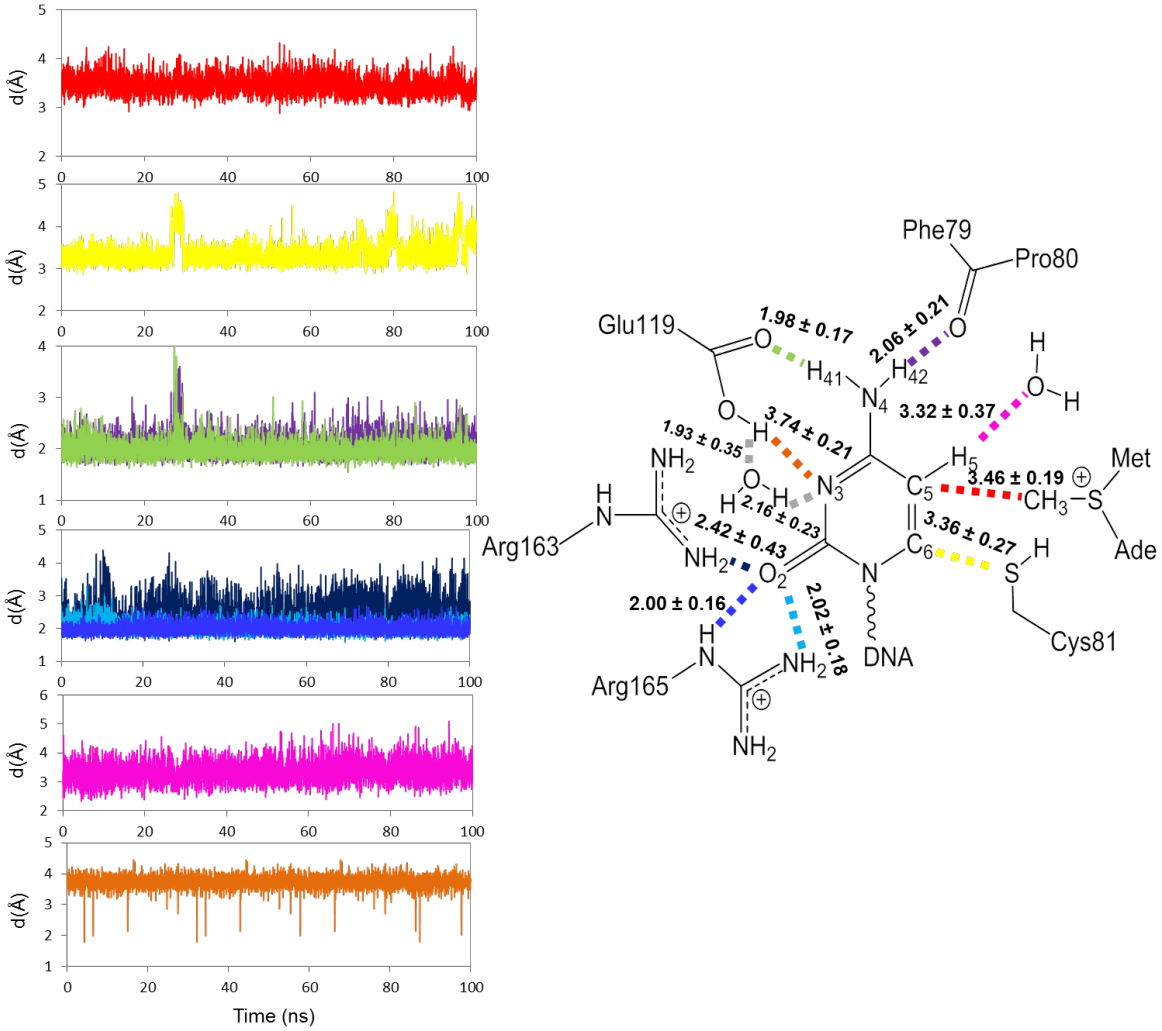


Figure S10. Interactions established between the C18 base and amino acids of M.*Hha*I active site in the Cys81 protonated system. The plots show the instantaneous values observed during the simulations. The values shown in the active site scheme correspond to the averaged values and their standard deviations. Distances are given in Å.

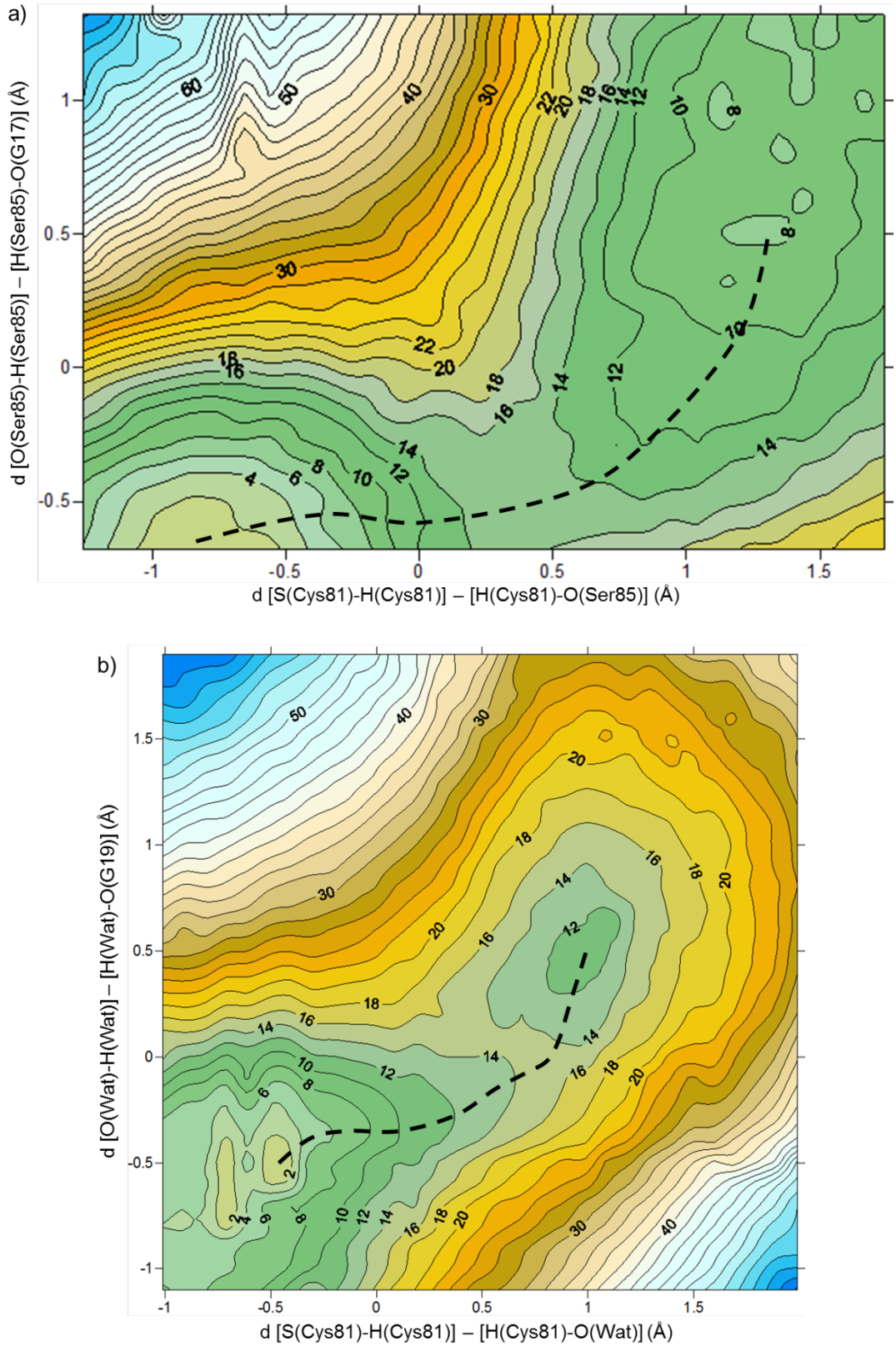
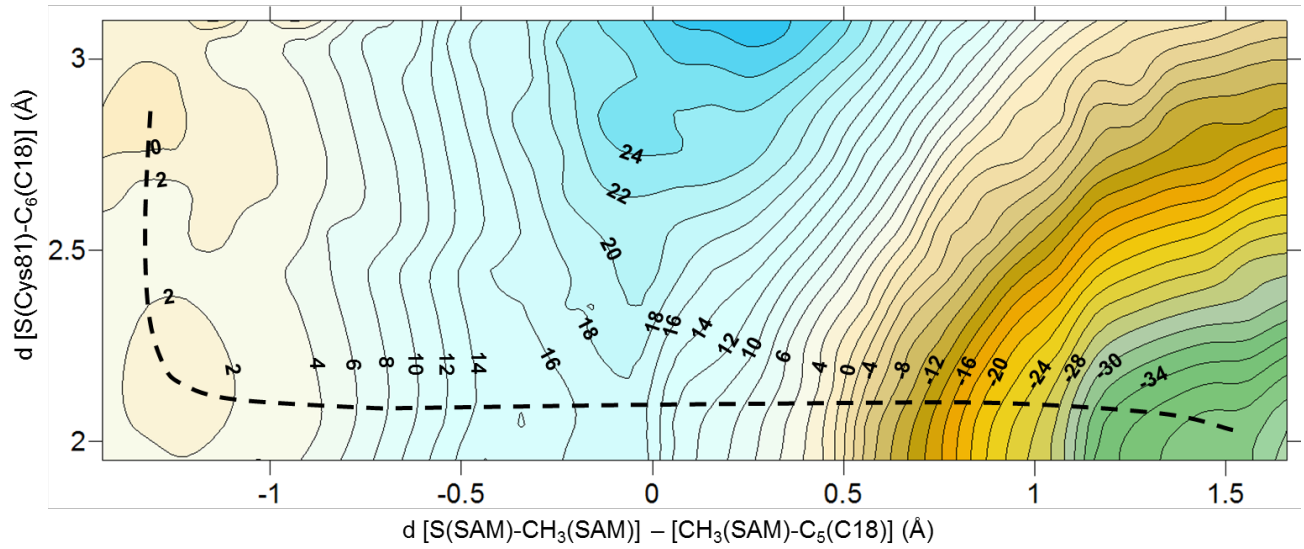


Figure S11. Potential Energy Surface calculated in M.*Hha*I at the M06-2X/6-31G**/MM level for (a) deprotonation of Cys81 by a DNA phosphate group mediated by Ser85 and (b) by a DNA phosphate group mediated by a water molecule. Isoenergetic lines are given in kcal·mol^{-1}.

Table S1. Key averaged distances (in Å) obtained for reactants, intermediate, products and transition states of the two reaction mechanisms for the deprotonation of Cys81 in M.*Hha*I. Standard deviations are included. Free energies obtained after M06-2X/6-311+G**/MM corrections are given in kcal·mol^{-1}.

	1	**TS 1-2**	**2**
d(S(Cys81) -H(Cys81))	1.49 ± 0.04	1.87 ± 0.08	2.00 ± 0.08
d(H(Cys81) - O(Ser85))	1.73 ± 0.05	1.11 ± 0.07	1.01 ± 0.03
d(O(Ser85) - H(Ser85))	0.98 ± 0.03	1.03 ± 0.03	1.62 ± 0.04
d(H(Ser85) - O(G17))	1.90 ± 0.09	1.59 ± 0.07	1.00 ± 0.03
ΔG	0	14.3 ± 0.3	7.6
	1	**TS 1-2'**	**2'**
d(S(Cys81) -H(Cys81))	1.47 ± 0.04	1.92 ± 0.09	2.23 ± 0.09
d(H(Cys81) - O(Wat))	2.02 ± 0.07	1.00 ± 0.03	0.97 ± 0.03
d(O(Wat) - H(Wat))	0.97 ± 0.02	1.23 ± 0.04	2.59 ± 0.13
d(H(Wat) - O(G19))	2.01 ± 0.12	1.15 ± 0.04	0.98 ± 0.03
ΔG	0	14.8 ± 0.2	9.2

a)



b)

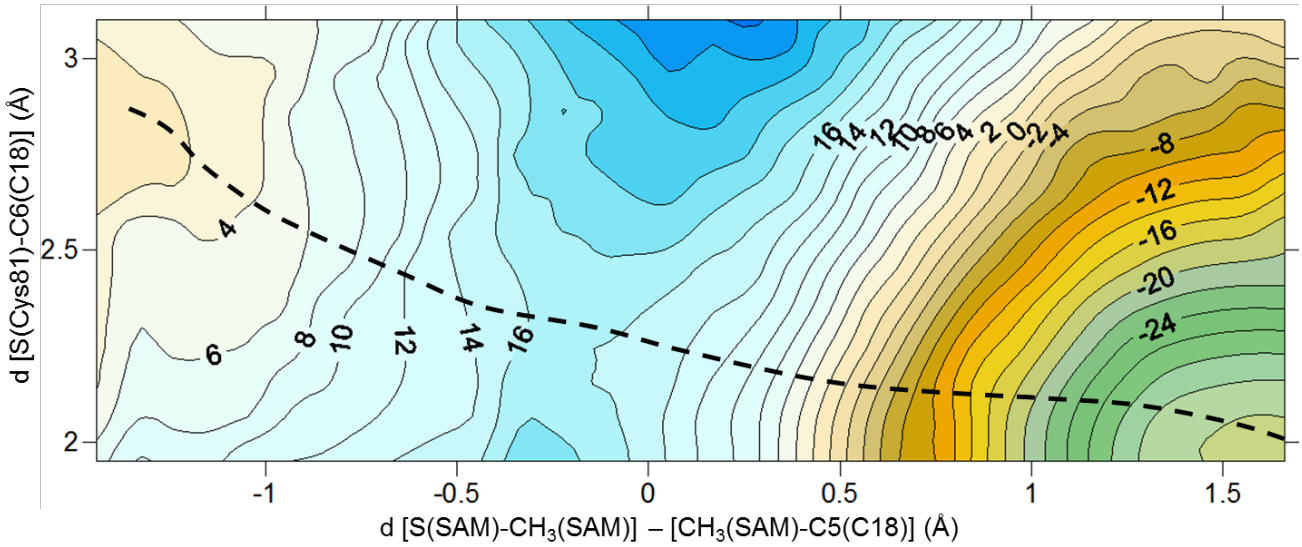


Figure S12. Potential Energy Surfaces for the methyl transfer and the nucleophilic attack of Cys81 in M.*Hha*I calculated at a) M06-2X/6-31G**/MM level and b) B3LYP/6-31G** level. Isoenergetic lines are given in kcal·mol^{-1}.

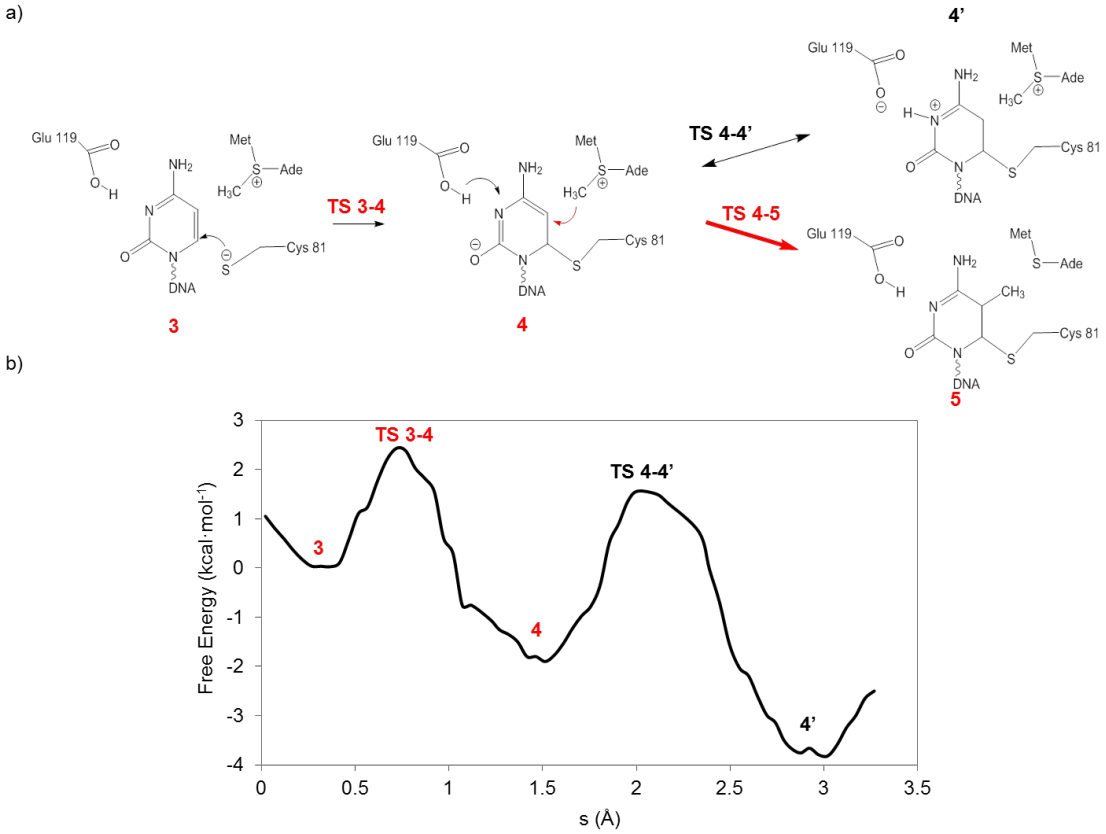


Figure S13. a) Scheme for the addition of the Cys81 to the cytosine cycle followed by i) the side reaction concerning the protonation of the N_3 atom of the cytosine by the Glu119 residue or ii) by methylation without protonation of the N_3 atom. b) AM1/MM PMFs, corrected at the M06-2X/6-311+G** level, obtained for the Cys81 approach (TS 3-4) and proton transfer from the Glu119 residue to the cytosine N_3 atom (TS 4-4') in M.*Hha*I as a function of the collective path coordinate *s*.

Table S2. Key averaged distances (in Å) obtained for reactants, intermediate, products and transition states of the reaction mechanism for the Cys81 addition and proton transfer from the Glu119 residue to the cytosine base in M.*Hha*I. Standard deviations are included. Free energies obtained after M06-2X/6-311+G**/MM corrections are given in kcal·mol^{-1}.

	4	TS 4-4'	4'
d(S(SAM) - CH_3(SAM))	1.83 ± 0.14	1.84 ± 0.18	1.83 ± 0.20
d(CH_3(SAM) - C_5(C18)	2.93 ± 0.12	3.10 ± 0.10	3.14 ± 0.14
d(S(Cys81) - C_6(C18))	2.10 ± 0.08	2.05 ± 0.09	1.98 ± 0.08
d($O\varepsilon_2$(Glu119) - $H\varepsilon_2$(Glu119))	0.97 ± 0.02	1.35 ± 0.08	1.98 ± 0.09
d($HO\varepsilon_2$(Glu119) - N_3(C18))	2.62 ± 0.14	1.25 ± 0.06	1.01 ± 0.07
d($O\varepsilon_1$(Glu119) - H_{41}(C18))	2.32 ± 0.24	2.34 ± 0.33	2.43 ± 0.23
d(O(Phe79) - H_{42}(C18))	2.22 ± 0.17	2.12 ± 0.18	2.16 ± 0.21
d(HNε(Arg165) – O_2(C18))	1.97 ± 0.13	2.10 ± 0.18	1.97 ± 0.11
d($HN\eta_2$(Arg165) – O_2(C18))	2.09 ± 0.16	2.13 ± 0.13	2.12 ± 0.23
d($HN\eta_1$(Arg163) – O_2(C18))	1.99 ± 0.24	2.21 ± 0.14	2.15 ± 0.14
ΔG	-1.91	1.56 ± 0.3	-3.75

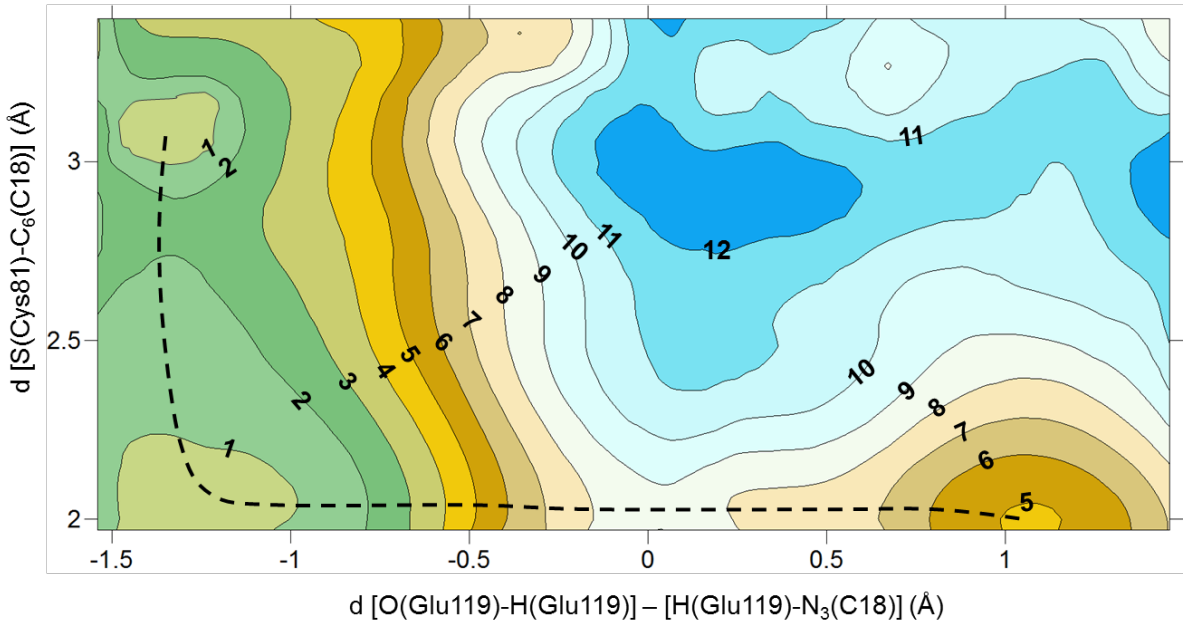


Figure S14. Potential Energy Surface for the addition of Cys81 to the cytosine ring and proton transfer from Glu119 to the N_3 atom of cytosine calculated at the M06-2X/6-31G**/MM level, starting from state **3** in M.*Hha*I. Isoenergetic lines are given in kcal·mol^{-1}.

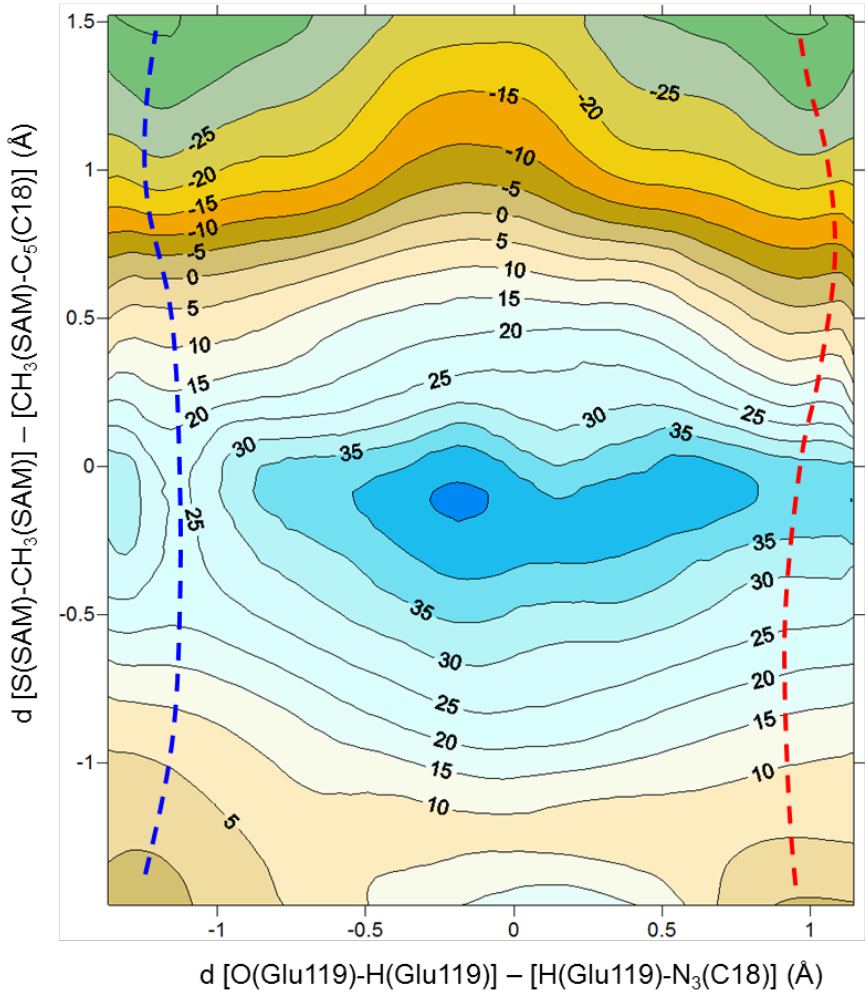
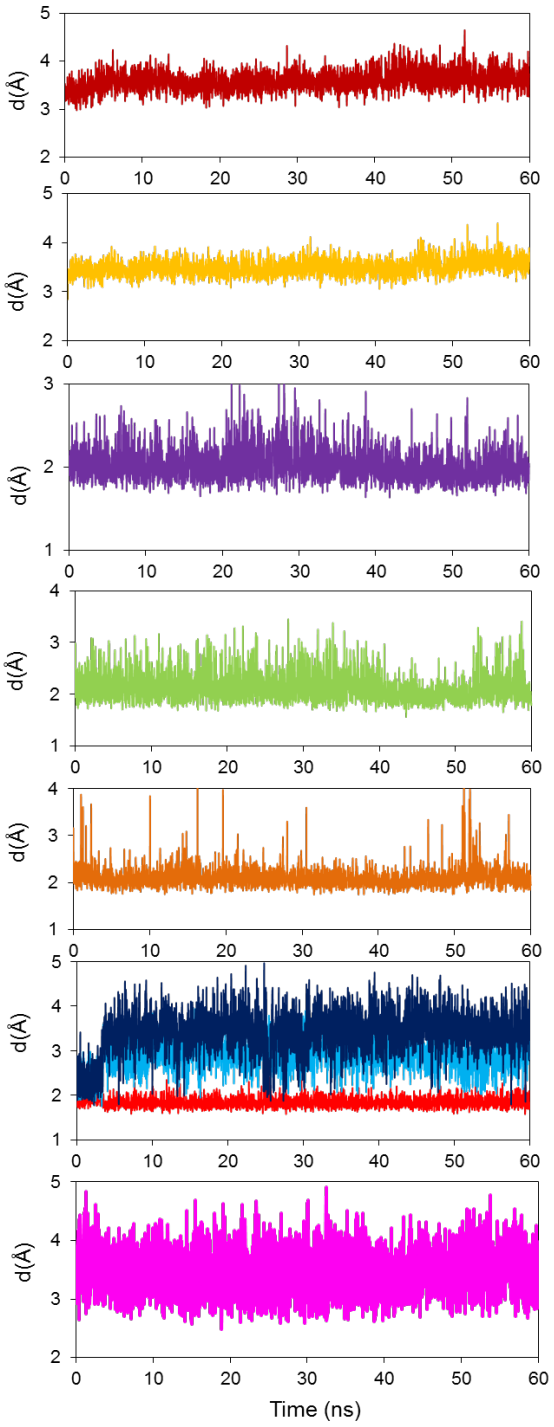


Figure S15. Potential Energy Surface at the M06-2X/6-31G**/MM level in M.*Hha*I for the proton transfer from Glu119 to the N_3 atom of the cytosine ring and the methyl transfer from cofactor SAM to the C_5 position of the cytosine ring starting from intermediate **4**. Isoenergetic lines are given in kcal·mol^{-1}. Note that the barrier for the methylation to a N_3 protonated intermediate (red line) is significantly higher than for an unprotonated one (blue line).



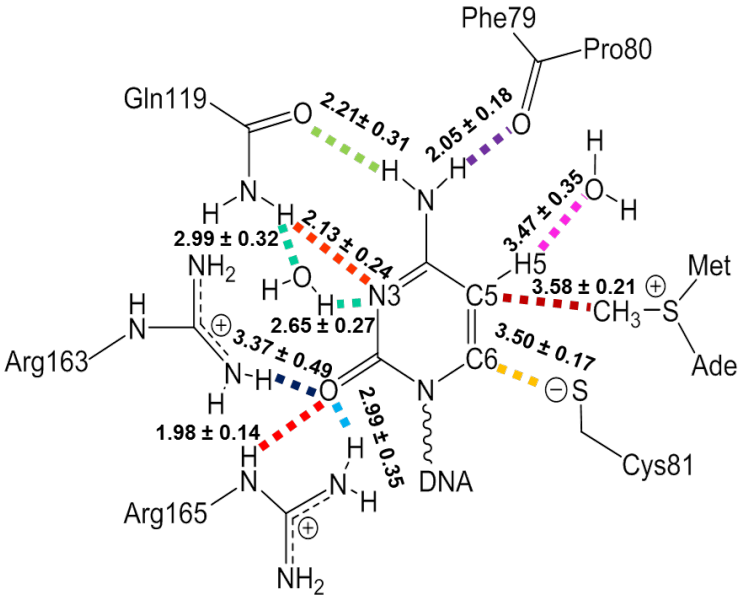
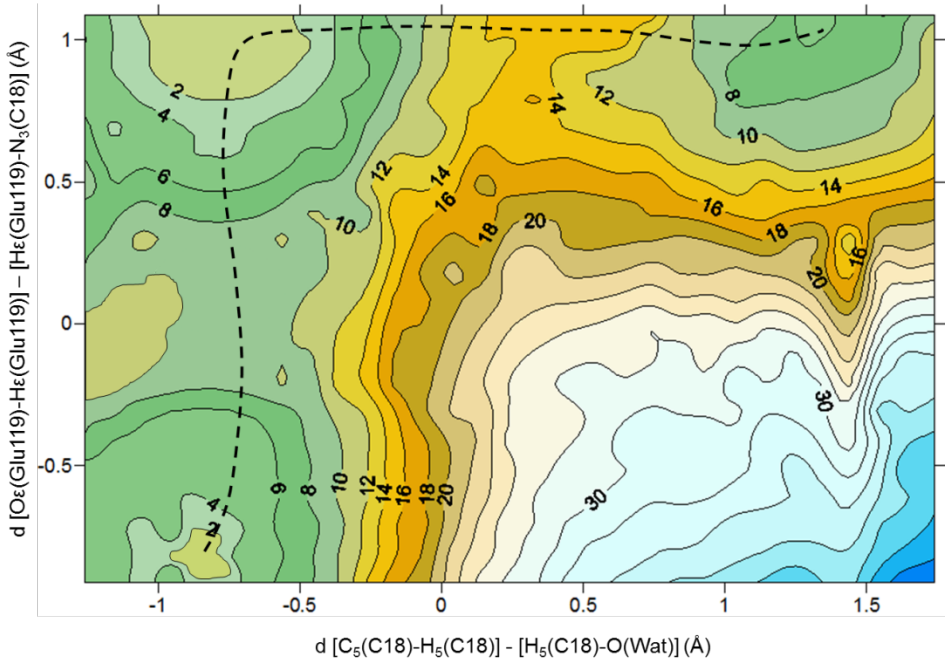


Figure S16. Interactions established between the C18 base and amino acids of the Gln119 M.*Hha*I variant active site. The plots on the left show the instantaneous distances observed during the simulations. The values shown in the active site scheme displayed on the right correspond to the average values and their standard deviations. Distances are given in Å.

a)



b)

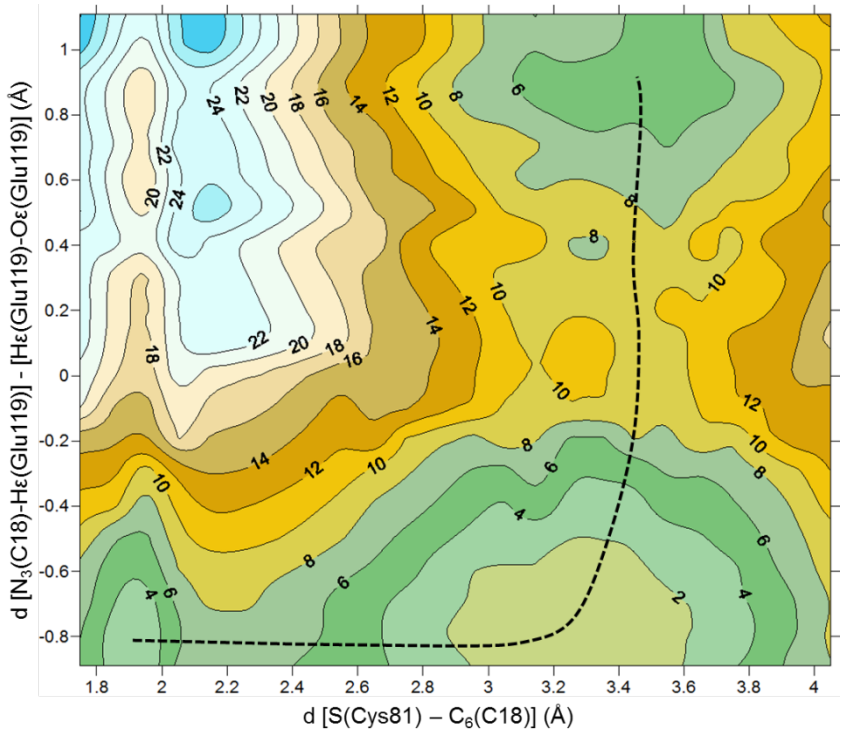


Figure S17. Potential Energy Surfaces for the β-elimination step involving a) the proton transfer between Glu119 and the N_3 atom of cytosine and the proton transfer from the C_5 atom of the cytosine to a water molecule and b) the breaking of the Cys81-cytosine bond and the proton transfer back from N_3 to the Glu119 residue. PESs were calculated at the M06-2X/6-31G**/MM level. Isoenergetic lines are given in kcal·mol^{-1}.

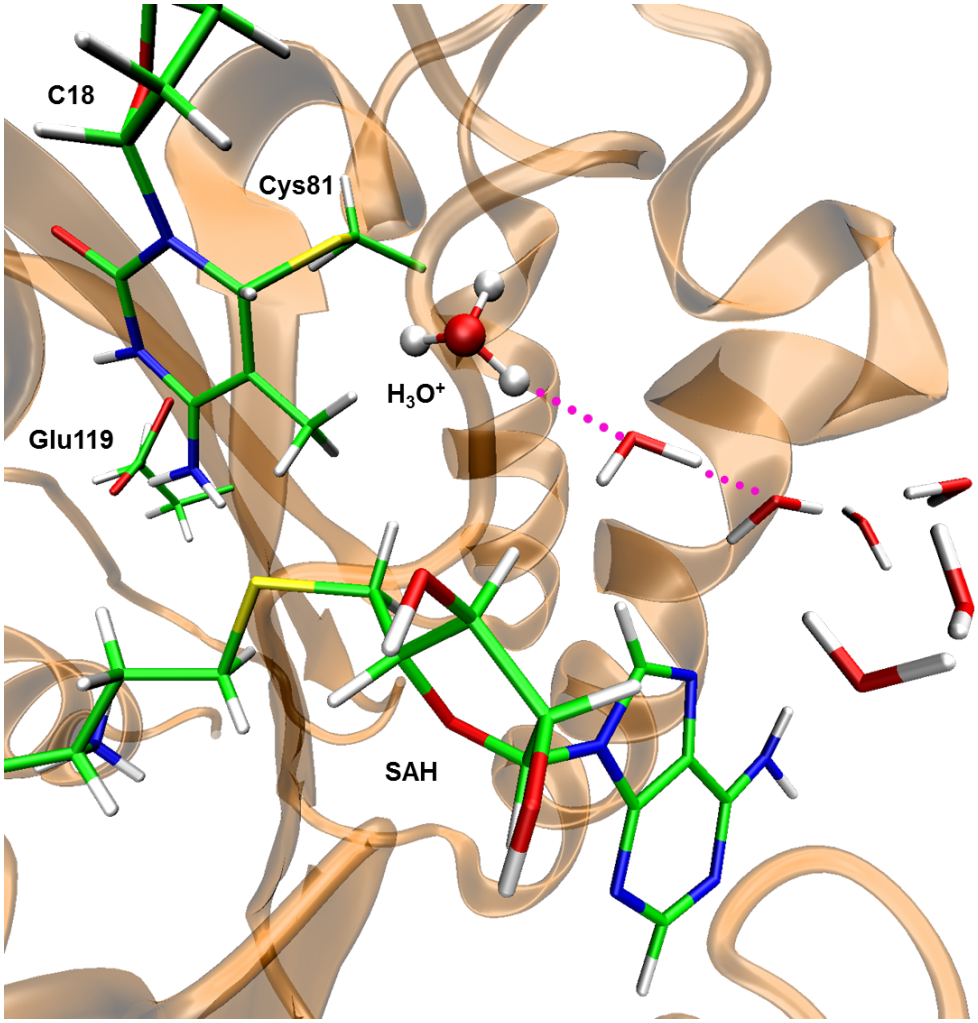


Figure S18. Water channel connecting the hydronium ion formed in the active site after abstraction of the C_5 proton by a water molecule with the bulk.

References

(1) Zacharias, M.; Straatsma, T. P.; McCammon, J. A. *J. Chem. Phys.* **1994**, *100*, 9025-9031.

(2) Gao, J.; Kuczera, K.; Tidor, B.; Karplus, M. *Science* **1989**, *244*, 1069-1072.

(3) Pearlman, D. A. *J. Phys. Chem.* **1994**, *98*, 1487-1493.

(4) Liu, P.; Dehez, F.; Cai, W.; Chipot, C. *J. Chem. Theory Comput.* **2012**, *8*, 2606-2616.

(5) Dundas, J.; Ouyang, Z.; Tseng, J.; Binkowski, A.; Turpaz, Y.; Liang, J. *Nucleic Acids Res.* **2006**, *34*, W116-W118.

(6) Jayaram, B.; Jain, T. *Annu. Rev. Biophys. Biomol. Struct.* **2004**, *33*, 343-361.

(7) Shieh, F. K.; Youngblood, B.; Reich, N. O. *J. Mol. Biol.* **2006**, *362*, 516-527.

(8) Svedružić, Ž. M.; Reich, N. O. *Biochemistry* **2004**, *43*, 11460-11473.

(9) Matje, D. M.; Coughlin, D. F.; Connolly, B. A.; Dahlquist, F. W.; Reich, N. O. *Biochemistry* **2011**, *50*, 1465-1473.

(10) Huang, N.; MacKerell Jr, A. D. *J. Mol. Biol.* **2005**, *345*, 265-274.

(11) Youngblood, B.; Buller, F.; Reich, N. O. *Biochemistry* **2006**, *45*, 15563-15572.

Complete References

(38) Case, D. A.; Darden, T. A.; Cheatham, T. E.; Simmerling, C. L.; Wang, J.; Duke, R. E.; Luo, R.; Walker, R. C.; Zhang, W.; Merz, K. M.; Roberts, B.; Hayik, S.; Roitberg, A.; Seabra, G.; Swails, J.; Goetz, A. W.; Kolossváry, I.; Wong, K. F.; Paesani, F.; Vanicek, J.; Wolf, R. M.; Liu, J.; Wu, X.; Brozell, S. R.; Steinbrecher, T.; Gohlke, H.; Cai, Q.; Ye, X.; Hsieh, M. J.; Cui, G.; Roe, D. R.; Mathews, D. H.; Seetin, M. G.; Salomon-Ferrer, R.; Sagui, C.; Babin, V.; Luchko, T.; Gusarov, S.; Kovalenko, A.; Kollman, P. A.; Amber 12; University of California, San Francisco: **2012**.

(56) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J. A.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, J. M.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. Gaussian 09, Revision C.01; Wallingford CT, **2009**.