Theoretical Study of the Catalytic Mechanism of DNA-(N4-Cytosine)-Methyltransferase from the Bacterium *Proteus vulgaris*

Juan Aranda, Maite Roca,* Violeta López-Canut, and Iñaki Tuñón*

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

Received: November 20, 2009; Revised Manuscript Received: April 28, 2010

In this paper the reaction mechanism for methylation of cytosine at the exocyclic N4 position catalyzed by M.PvuII has been explored by means of hybrid quantum mechanics/molecular mechanics (QM/MM) methods. A reaction model was prepared by placing a single cytosine base in the active site of the enzyme. In this model the exocyclic amino group of the base establishes hydrogen bond interactions with the hydroxyl oxygen atom of Ser53 and the carbonyl oxygen atom of Pro54. The reaction mechanism involves a direct methyl transfer from AdoMet to the N4 atom and a proton transfer from this atom to Ser53, which in turn transfers a proton to Asp96. Different timings for the proton transfers and methylation steps have been explored at the AM1/MM and B3LYP/MM levels including localization and characterization of stationary structures. At our best estimate the reaction proceeds by means of a simultaneous but asynchronous proton transfer from Ser53 to Asp96 and from N4 of cytosine to Ser53 followed by a direct methyl transfer from AdoMet to the exocyclic N4 of cytosine.

1. Introduction

DNA methyltransferases catalyze the transfer of a methyl group from the cofactor S-adenosyl-L-methionine (AdoMet) to a given position of a particular DNA base within a specific DNA sequence. These enzymes attach inheritable information to the DNA that is not encoded in the nucleotide sequence. This so-called epigenetic information has many important biological functions. In prokaryotes, DNA methylation has three major biological roles such as distinction of self- and non-self-DNA, direction of postreplicative mismatch repair, and control of DNA replication and cell cycle. In eukaryotes, DNA methylation contributes to the control of gene expression, protection of the genome against selfish DNA, maintenance of genome integrity, parental imprinting, X-chromosome inactivation in mammals and regulation of development.[1]

DNA methyltransferases differ in the nature of the modification introduced. The members of one class (e.g., M.*Hha*I) methylate C5, a ring carbon of cytosine, yielding 5-methylcytosine, while members of the second class methylate the exocyclic amino group (NH_2) of cytosine (e.g., M.PvuII) or adenine (e.g., M.*Taq*I), yielding N4-methylcytosine or N6-methyladenine, respectively. In general, both types of methyltransferases are two-domain proteins comprising one large and one small domain, with the binding site for the AdoMet and the catalytic center of the enzyme being located at the large domain. The most interesting structural and mechanistic feature of DNA methyltransferases is that they completely flip the target base out of the DNA helix and bind it into a hydrophobic pocket in the large domain of the enzyme that creates the active site for methyl group transfer.[2]

This work is focused on the study of the catalytic mechanism of the methyl transfer from AdoMet to N4-cytosine catalyzed by PvuII methyltransferase (M.PvuII) from *Proteus vulgaris*[3] in its recognition sequence 5′-CAGCTG-3′. This enzyme, part of the restriction−modification system, folds into a structure with a V-shape cleft with enough space to accommodate duplex DNA,[4] and the methyl donor AdoMet binds at the bottom of the cleft. The M.PvuII polypeptide chain is 323 amino acids long and it is crystallized with the methyl donor forming the M.PvuII-AdoMet complex.[4] To our knowledge this is the only resolved structure of a N4-methyltransferase.

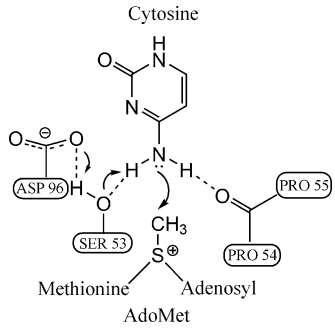
Structures of DNA methyltransferases in complex with target DNA are available, and base flipping is observed in all of them. The dynamics of base flipping have been most intensively studied with the M.*Hha*I cytosine-C5-methyltransferase.[2,5] It has been shown that the base flipping comprises at least two steps: first, the base is rotated out of the DNA helix, but not tightly bound by the enzyme, and then it is inserted into the hydrophobic binding pocket of the enzyme.[6] Base flipping will work with any base at the target position,[7,8] but the methylation reaction is sensitive to the base at the target position. However, to our knowledge, the crystal structure of a M.PvuII−DNA complex has not been determined although in a previous study Gong et al. constructed a model of M.PvuII docked to cognated DNA taken from the structure of the complex with R.PvuII (an endonuclease enzyme).[4] The catalytic loop of the amino methyltransferases contains several conserved residues that correspond to motif IV. Mutations within this region strongly reduce the catalytic activity of different DNA methyltransferases.[9–13] As seen in the docked model obtained by Gong et al.,[4] the target cytosine N4 would have two possible hydrogen bond partners: the hydroxyl group Ser53 and the main chain carbonyl Pro54, the first two amino acids of the conserved motif IV. Moreover, Asp96 may hydrogen bond with the Ser53 hydroxyl group, thereby facilitating proton transfer from the cytosine amino group through Ser and eventually to the Asp (see Scheme 1). Thus, Ser53 and Asp96 appear to belong to a charge relay system analogous to that seen in serine proteinases.[14]

Different reaction mechanisms are in principle possible for the methyl transfer from AdoMet to N4-cytosine. A schematic view of the M.PvuII active site and the elementary chemical steps that take place during the reaction are represented in Scheme 1. Given the pK_a values of the N4 atom of cytosine,

* To whom correspondence should be addressed. E-mail: M.R., M.Teresa.Roca@uv.es; I.T., Ignacio.Tunon@uv.es.

10.1021/jp911036w
Published on Web 06/04/2010

SCHEME 1



aspartic acid, and serine residues, it could be proposed that the proton transfer occurs late in the reaction when the N4−CH_3 bond has been almost completely formed. Therefore, a cationic transition state would be a plausible proposal that could be further stabilized by cation−π interactions with surrounding aromatic amino acid residues.[15] This model is supported by the presence of one or even more aromatic residues that could contact the flipped base in the active sites of *N*-methyltransferases. However, the methylation reaction at the exocyclic amino group of cytosine rings proceeds with inversion of the configuration of the methyl group in an S_N2 reaction, which is not an easy task because cytosine does not display nucleophilicity at the exocyclic amino group since the free electron pair is conjugated with the aromatic system. Then, it is quite unlikely that the reaction proceeds through a small activation barrier if methylation precedes the proton transfer. However, experimental burst kinetic assays have been carried out on this enzyme, showing that the product release and not the chemical reaction is the rate-determining step.[16,17]

To solve these mechanistic issues, we herein present a theoretical study of the reaction that catalyzes the DNA-(N4-cytosine)-methyltransferase from the bacterium *Proteus vulgaris*. Hybrid quantum mechanics/molecular mechanics (QM/MM) and gas phase calculations are carried out to obtain the potential energy profiles. Our results seem to point to a mechanism where the deprotonation of the exocyclic amino group occurs first through a proton relay mechanism that involves Ser53 and Asp96. The amino group of the deprotonated base is then more nucleophilic, facilitating the subsequent methyl transfer step from AdoMet.

2. Methods

The initial coordinates were taken from the X-ray crystal structure of PvuII DNA-(cytosine N4)-methyltransferase complexed with *S*-adenosyl-L-methionine (AdoMet) with PDB code 1BOO.[4] We rebuilt some missing amino acid side chains and they were optimized using the Newton−Raphson method.[18–20] Hydrogen atoms on titratable residues were added using the cluster method, as implemented by Field and co-workers.[21,22] Since there is no X-ray structure crystallized for the enzyme with DNA, we then built a small model of the reacting system that consists of inserting the flipped cytosine into the active site (see Figure 1). For this purpose, we inserted the methyl transfer transition structure obtained from a PCM/B3LYP/6-31G* calculation for a cytosine base and trimethylsulfonium system, with values of 2.49 Å for the breaking bond (S(TMS)−CH_3-(TMS)) and 2.00 Å for the forming bond (CH_3(TMS)−N4(Cyt)). After the trimethylsulfonium moiety was overlapped with the X-ray coordinates of the corresponding atoms in AdoMet, the whole enzymatic system (with AdoMet as methyl donor and cytosine as methyl acceptor) was introduced inside

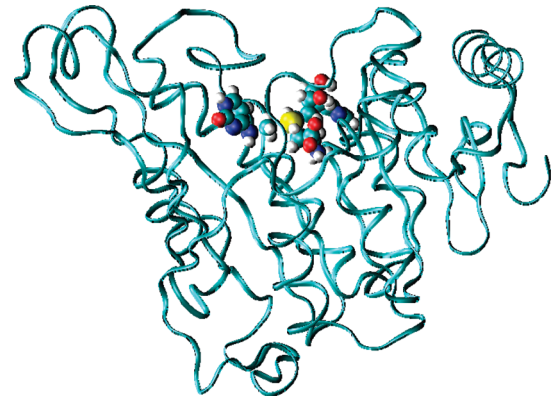


Figure 1. X-ray crystal structure of PvuII DNA-(cytosine N4)-methyltransferase complexed with AdoMet and docked with the cytosine base. The protein folds into a structure with a V-shaped cleft. AdoMet and cytosine, in sphere representations, are bound at the bottom of the cleft.

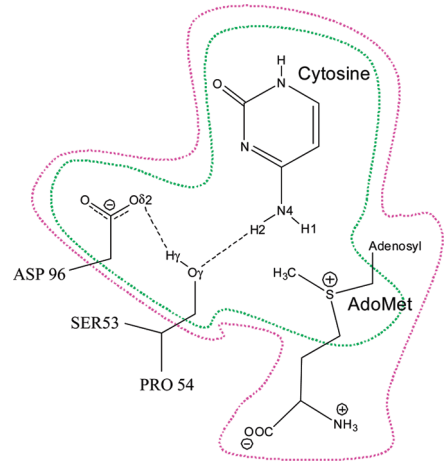


Figure 2. Details of the M.PvuII active site and definition of the different quantum regions used in this work during the QM/MM calculations. The region depicted in pink is the quantum region defined using AM1 Hamiltonian (74 QM atoms) and the region in green is the one used in B3LYP calculations (35 QM atoms).

a cubic box of TIP3P[23] water molecules of side 79.5 Å. Thus the final system consists of the enzyme (4543 atoms), AdoMet, cytosine base, and 15 137 water molecules (50 017 atoms in total).

The QM region consists of AdoMet, cytosine base, and the side chains of Ser53 and Asp96 residues, a total of 74 atoms (see Figure 2) and was initially described by the semiempirical method AM1[24] while the classical atoms were described by means of the OPLS-AA[25,26] and TIP3P[23] force fields as implemented in fDYNAMO.[27,28] To saturate the valence of the QM/MM frontier, we used the link atoms procedure.[29,30] To treat the nonbonding interactions, a switch function with a cutoff distance in the range 14−18 Å was used. The MD simulations were performed on a structure similar to the transition state (TS) structure of the continumm model methylation reaction, thus we applied a harmonic constraint to the distances involved in the methyl transfer reaction, the breaking bond (S(AdoMet)−CH_3(AdoMet)) and the forming bond (CH_3(AdoMet)−N4(Cyt)), to remain in the putative TS region. The umbrella force constants used were of 2500 kJ·mol^{-1}·Å^{-2}. We performed several optimizations and QM/MM MD simulations in the NVT ensemble at 300 K using the DYNAMO program. First, we optimized and equilibrated the water molecules by means of 100 ps MD simulation, and after that, we optimized and

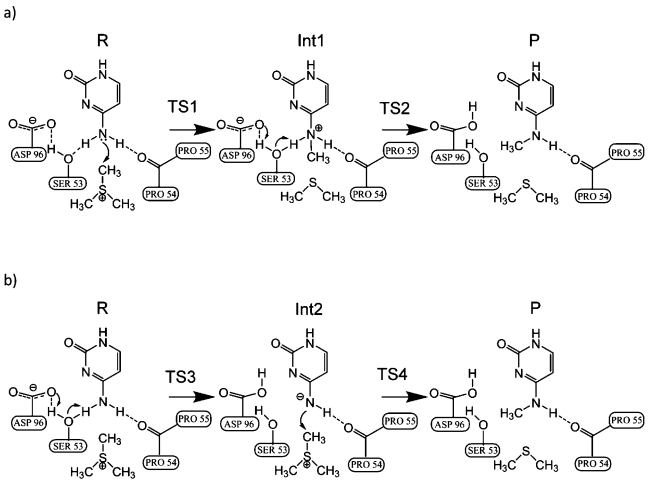


Figure 3. Details of the mechanisms studied in this work, in which (a) the methylation reaction precedes the proton transfer steps and (b) the proton transfers are first and then methylation occurs.

equilibrated the system except the backbone running 100 ps MD simulation. Finally, the entire system was optimized and equilibrated by means of 100 ps MD simulations. Periodic boundary conditions and a time step of 1 fs were used all around the simulations.

After analyzing the dynamics, we found that the most important interactions between amino acids of the active site and the cytosine base are identical to those predicted by Gong et al.[4] The active site structure obtained from MD simulations was fully optimized without any geometrical constraint applied and then taken as the starting point for performing gas phase calculations at the AM1 and B3LYP/6-31G* level using GAUSSIAN03.[31] A PDB file containing the coordinates of those atoms belonging to residues placed less than 10 Å from the substrate is provided as Supporting Information. The gas phase model consisted of the cytosine base, the side chain of Asp96, the side chain of Ser53 and part of its backbone, the backbone of the Pro54 residue, and part of the backbone of the Pro55 residue, while cofactor AdoMet was replaced by trimethylsulfonium (TMS). This model was used to explore different possible reaction mechanisms including different timings for the methyl and proton transfers. During this exploration the position of the C^{α} atoms and N1 and O6 of cytosine were kept frozen to the values found in the enzymatic optimized structure.

The more reliable mechanisms (selected on the basis of the energy barriers) are schematically shown in Figure 3. These two mechanisms were then explored in a model that included the whole enzymatic system and solvation water molecules using DYNAMO. Potential energy surfaces (PES), the location of stationary structures, and the corresponding intrinsic reaction coordinates (IRCs)[32,33] were obtained using a hybrid QM/MM approach, describing the quantum atoms using the AM1 Hamiltonian and B3LYP functional[34,35] with the 6-31G* basis set and the classical atoms by means of the OPLS-AA and TIP3P force fields. QM subsystems (see Figure 2) included the cytosine base, the side chains of Asp96 and Ser53 and also the entire cofactor AdoMet (in the AM1/MM treatment) or just the sulfur atom and the atoms directly bonded to it (by means of B3LYP/MM treatment). The total number of QM atoms was 74 and 35 for the AM1/MM and B3LYP/MM calculations, respectively.

Exploration of the PES was carried out by selection of distinguished coordinates and optimization of the remaining degrees of freedom of the system. The stationary points localization and characterization and the tracing of the IRCs were performed using the micro/macroiteration optimization algorithm.[36–42] This method consists of dividing the coordinate space into two subsets; the control space (that usually matches up with the QM region) and the complementary space (the rest of the system). Optimization steps on the control space (microiterations) make use of a Hessian-based algorithm. At each step of the control space Hessian guided optimization, the complementary space is minimized using gradient vectors (macroiterations).

To perform DFT/MM calculations, we used a combination of DYNAMO/GAUSSIAN03 programs, in which we calculated the wave function for the QM region by means of the density functional theory methodology using GAUSSIAN03. To take advantage of powerful optimization algorithms,[42–44] the QM/MM interaction has been calculated in different ways: using the standard electrostatic interaction of the polarized wave function with the charges of the environment during the QM optimization (or core); and a pure classical Coulombic term during the MM optimization (or environment), making use of a fitted charge distribution (CHELP charges)[45] obtained from the polarized wave function.[42] Starting points for optimization of the stationary structures were taken from exploration of the PES at the DFT/MM level. IRCs were traced from the TSs to the corresponding reactant and products structures. Finally, single-point energy calculations of the stationary points obtained at the B3LYP/MM level with the 6-31G* basis set (reactants, TSs, intermediates and products) were done using a larger 6-311+G** basis set.

Before presenting the results, it is important to discuss the possible limitations of the model built to study the reaction mechanism. The absence of the DNA fragment in our calculations can have at least two effects. First, the DNA sequence could deform the enzymatic structure significantly. However, as shown by Gong et al.,[4] the DNA sequence seems to nicely fit into the V-shaped cleft of the enzyme and then the flipped cytosine is placed close to the methyl donor, establishing some key hydrogen bond interactions. The same interactions (see below) are found in our flexible model of the enzyme with the cytosine base. Second, the presence of the DNA phosphate groups could produce important changes in the electrostatic properties of the active site. Using the X-ray structure of a C5-methyltransferase with DNA (PDB code 2HR1[46]), we measured that the closest phosphate group is placed more than 9 Å from the sulfur atom of AdoMet. Moreover, the electrostatic properties of these phosphate groups can be modulated by the presence of counterions and water molecules. Finally, it must also be taken into account that our present calculations do not include temperature effects, which could play an important role both in the determination of the free energy barrier and in favoring the sampling of some regions of the configurational space of the system. However, it must be noticed that using this simplified model, we have been able to employ DFT/MM techniques that would be otherwise computationally too expensive. We believe that these calculations on this model provide a good reference for future research on this enzyme.

3. Results and Discussion

The gas phase model mentioned in the previous section was used to explore different reaction mechanisms and determine the most reliable one in terms of potential energy barriers. The two trustworthy reaction mechanisms obtained in the gas phase are schematically shown in Figure 3 and have been explored for a cytosine substrate placed in the active site of a solvated

TABLE 1: Relevant Geometrical Parameters (Distances in Å, Angles in deg) Found for the Stationary Structures of the Two Reaction Mechanisms Explored at the AM1/MM Level[a]

	R	TS1	Int1	TS2	TS3	Int2	TS4	P
d(S(AdoMet)−CH_3(AdoMet))	1.827	2.213	3.298	3.332	1.821	1.822	2.157	3.446
d(CH_3(AdoMet)−N4(Cyt))	3.033	1.962	1.485	1.480	3.496	3.416	2.128	1.435
d(N4(Cyt)−H2(N4)(Cyt))	0.999	1.012	1.035	1.113	1.199	1.795	2.046	2.619
d(Oγ(Ser53)−H2(N4)(Cyt))	2.443	2.758	2.237	1.618	1.323	0.995	0.976	0.965
d(Oγ(Ser53)−Hγ(Ser53))	0.973	0.973	0.974	1.421	2.026	2.175	2.083	2.113
d(Oδ2(Asp96)−Hγ(Ser53))	2.002	2.003	1.980	1.068	0.981	0.972	0.973	0.972
d(S(AdoMet)−N4(Cyt))	4.857	4.174	4.478	4.750	5.241	5.157	4.282	4.566
d(O(Pro54)−H1(N4)(Cyt))	1.962	1.912	2.088	2.170	2.463	2.328	2.321	2.238
a(S(AdoMet)−CH_3−N4(Cyt))	175.68	178.84	135.33	160.12	159.55	158.77	175.97	134.35
ν		667.5i		618.5i	1198.5i		634.4i	
ΔE	0.0	29.3	5.7	20.7	20.7	12.5	27.2	−31.6

[a] Imaginary frequency of transition structures are given in cm^{-1}, and relative energies, in $kcal \cdot mol^{-1}$.

TABLE 2: Relevant Geometrical Parameters (Distances in Å, Angles in deg) Found for the Stationary Structures of the Two Reaction Mechanisms Explored at the B3LYP/6-31G*/MM Level[a]

	R	TS1	Int1	TS3	Int2	TS4	P
d(S(AdoMet)−CH_3(AdoMet))	1.835	2.417	3.465	1.839	1.852	2.111	3.562
d(CH_3(AdoMet)−N4(Cyt))	3.213	2.169	1.512	3.239	2.983	2.411	1.465
d(N4(Cyt)−H2(N4)(Cyt))	1.023	1.045	1.120	1.218	1.774	1.901	2.217
d(Oγ(Ser53)−H2(N4)(Cyt))	2.027	1.904	1.611	1.339	1.009	0.993	0.971
d(Oγ(Ser53)−Hγ(Ser53))	1.009	1.012	1.031	1.649	1.844	1.957	2.069
d(Oδ2(Asp96)−Hγ(Ser53))	1.721	1.710	1.623	1.031	0.994	0.989	0.982
d(S(AdoMet)−N4(Cyt))	4.985	4.575	4.956	4.962	4.738	4.503	4.984
d(O(Pro54)−H1(N4)(Cyt))	1.992	1.898	1.748	2.036	1.993	2.034	1.706
a(S(AdoMet)−CH_3−N4(Cyt))	161.05	172.02	168.52	154.43	156.35	169.33	163.51
ν		438.6i		993.8i		346.5i	
ΔE	0	20.2	−7.6	11.1	3.2	7.1	−48.6

[a] Imaginary frequency of transition structures are given in cm^{-1}, and relative energies (obtained from single-point calculation using the 6-311+G** basis set), in $kcal \cdot mol^{-1}$.

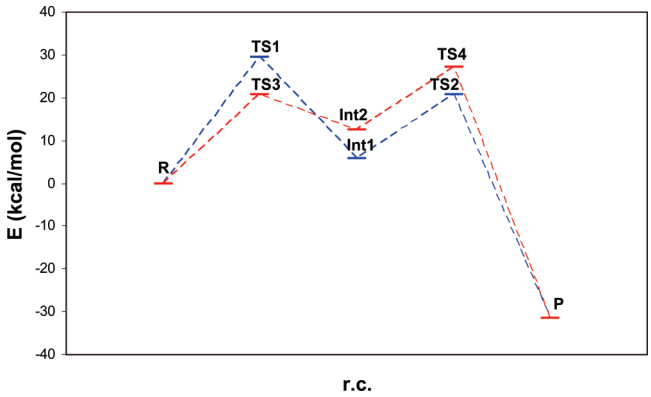


Figure 4. Potential energy profiles obtained for the two reaction mechanisms explored at the AM1/MM level for N4-methylation in M.PvuII.

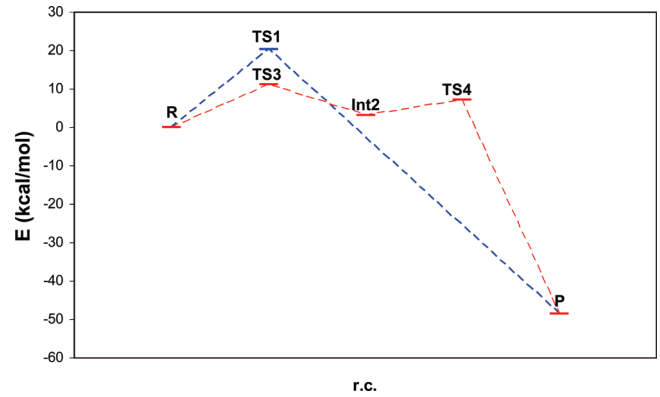


Figure 5. Potential energy profiles obtained for the two reaction mechanisms explored at the B3LYP/MM level for N4-methylation in M.PvuII.

M.PvuII enzyme. In the first mechanism the methyl transfer from AdoMet to the exocyclic nitrogen atom of the cytosine base precedes the proton transfers from this nitrogen atom to Ser53 and from this residue to Asp96. In the second one the proton transfer steps occur first, giving place to a negatively charged N4 atom, and then this atom receives the positively charged methyl group from AdoMet. Other combinations of methyl and single proton transfer were explored in a gas phase model but they produce significantly larger reaction barriers. The relative energies and most relevant geometrical parameters of the stationary structures located at the AM1/MM and B3LYP/MM levels are given in Tables 1 and 2, respectively. Reaction profiles are shown in Figures 4 (AM1/MM) and 5 (B3LYP/MM), while B3LYP/MM transition structures are depicted in Figure 6.

AM1/MM Results. At the AM1/MM level the reactant structure is characterized by the presence of two hydrogen bonds established between hydrogen atoms of the exocyclic amino group of cytosine and the hydroxyl oxygen atom of Ser53 (2.443 Å; see Table 1) and the carbonyl oxygen atom of Pro54 (1.962 Å). Ser53 is also hydrogen bonded to Asp96, the distance between the hydrogen hydroxyl atom of Ser53 and the carboxylate oxygen of Asp96 being 2.002 Å. The distance from the AdoMet methyl group to the exocyclic N4 atom of cytosine is only 3.033 Å. From this reactant structure and exploring the PES using the antisymmetric combination of the breaking bond (S(AdoMet)−CH_3(AdoMet)) and the forming bond (CH_3(AdoMet)−N4(Cyt)), we localized and characterized the transition structure TS1 corresponding to the methyl transfer from AdoMet to the exocyclic nitrogen atom of cytosine. In this transition structure the carbon−N4 distance is shorter than

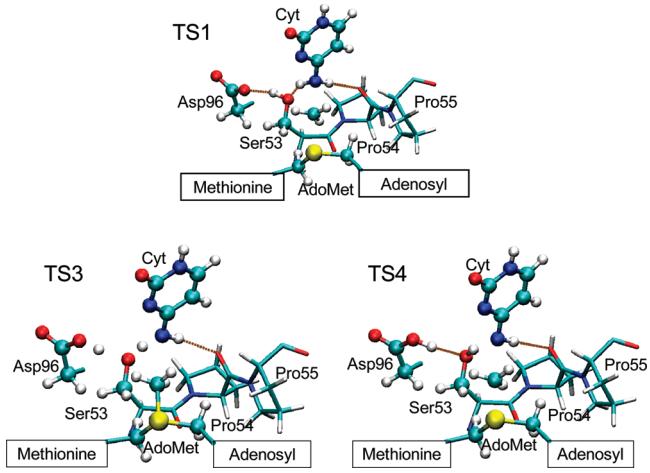


Figure 6. Transition structures located at the B3LYP/MM level for the reaction mechanisms explored for N4-methylation in M.PvuII.

the carbon−S one (1.962 Å and 2.213 Å, respectively). Pauling's bond orders (BOs)[47] show that S−C bond breaking (BO = 0.526) and C−N bond forming (BO = 0.452) are approximately equally advanced in this transition structure. The positioning of the transferred methyl group between the donor and acceptor atoms is stabilized by the carbonyl oxygen atom of Ser53 (the distance from this atom to the closest hydrogen atom of the transferred methyl group is 2.080 Å). The reaction path following from this transition structure leads to a positively charged methylated cytosine (Int1) with strengthened hydrogen bond interactions with Ser53 (2.237 Å). Only a small fraction of the positive charge (0.13 au) is delocalized on the cytosine ring. Furthermore, we did not find any evidence of cation−π or π−π interactions with enzymatic residue. The closest aromatic residue (Phe56) is found more than 7 Å from the N4 atom. From this intermediate structure we explored the PES corresponding to the proton transfers from exocyclic amino group to Ser53 and from this residue to Asp96 using the antisymmetric combination of hydrogen-donor atom and hydrogen-acceptor atom distances. The only reaction path located corresponds to a simultaneous but asynchronous process. Effectively, the transition structure (TS2) shows an advanced proton transfer from Ser53 to Asp96, where the distance from the Ser53 hydrogen atom to the Asp96 carboxylate oxygen atom is 1.068 Å. Otherwise, the proton transfer from the exocyclic amino group to the Ser53 hydroxyl group is in a very early stage in this structure, the distance between the hydrogen and N4 atoms of cytosine being 1.113 Å. From this structure IRC directly leads to the final reaction product, where the neutral methylated base still presents a hydrogen bond interaction with the carbonyl oxygen atom of Pro54 (2.238 Å). As shown in Figure 4, the product is considerably more stable than the reactant (relative energy of −31.6 kcal·mol^{-1}). In this mechanisms, TS1 is the highest transition structure, the relative energy with respect to reactants being 29.3 kcal·mol^{-1}.

The second reaction mechanism begins from the same reactant structure, but now the proton transfers precede the methyl transfer. Transition structure TS3 corresponds again to a simultaneous but asynchronous proton transfer from the exocyclic amino group of cytosine to Ser53 and from this residue to Asp96. In this transition structure proton transfer from Ser53 to Asp96 is nearly completed (distance from hydrogen to carboxylic oxygen is very close to the equilibrium distance, 0.981 Å) while the proton transfer from N4 of cytosine to Ser53 is almost halfway (the hydrogen−N4 distance is 1.199 Å, and the distance between the hydrogen of cytosine and the hydroxyl oxygen of Ser53 is 1.323 Å). From this transition structure we located an intermediate structure (Int2) that presents a protonated Asp96, a neutral Ser53, and a formally negatively charged N4 atom of cytosine. The subsequent transition structure (TS4) corresponds to a methyl group placed nearly exactly halfway between the donor atom (S) and the acceptor one (N4). The donor−acceptor distance has been significantly reduced (4.282 Å). Analysis of BOs shows that C−S bond breaking (BO = 0.572) is more advanced than C−N4 bond forming (BO = 0.315). IRC following from this structure directly leads to the reaction products. In this mechanism the second transition structure (TS4) is the highest one, presenting a relative energy with respect to the reactants of 27.2 kcal·mol^{-1}. The energy barrier associated with this mechanism is 2.1 kcal·mol^{-1} lower than for the first one. Thus, a mechanism in which the exocyclic amino group is first deprotonated and then methylated seems to be kinetically more favorable than a mechanism where the N4 atom of cytosine is first methylated and then deprotonated. However, the difference is small and the conclusions, at this computational level, could change after consideration of temperature effects.

B3LYP/MM Results. At the B3LYP/MM level the reactant structure presents the same hydrogen bond interaction pattern as the reactant structure located at the AM1/MM level. The exocyclic amino group is hydrogen bonded to the carbonyl oxygen atom of Pro54 and to the hydroxyl group of Ser53. In turn, the latter also acts as proton donor to Asp96. The hydrogen bond distance between the hydrogen amino atom of cytosine and the hydroxyl oxygen atom of Ser53 (2.027 Å, see Table 2) and from the hydrogen atom of the Ser53 hydroxyl group to the carboxylate oxygen atom of Asp96 (1.721 Å) are significantly shorter than at the AM1/MM level. Instead, the distance from the carbon atom of the methyl group to N4 atom of cytosine (3.213 Å) is larger than at the AM1/MM level. In the first explored reaction mechanism we located a transition structure (TS1) corresponding to the methyl transfer from AdoMet to the neutral cytosine base. In this structure obtained at the B3LYP/MM level the methyl transfer takes place with a significantly larger donor−acceptor distance than that at the AM1 level (4.575 Å versus 4.174 Å). As a consequence, this transition structure is more dissociative than predicted at the AM1 level. The Pauling's bond orders for the S−C and the C−N4 bonds are 0.379 and 0.334 respectively. This transition structure is stabilized by the interaction established between the transferred positively charged methyl group and the carbonyl oxygen atom of Ser53 (2.667 Å). The hydrogen bonds between the cytosine amino group and the residues of the active site are also strengthened, as reflected in the shorter distances (see Table 2). From this transition structure, and after methylation of N4 of cytosine has been completed, we found that proton transfers from the cytosine amino group to Ser53 and from this residue to Asp95 take place in a concerted way without an energy barrier. PES exploration shows a very flat region corresponding to the positively charged methylated cytosine. This evolves in a barrierless process until the final product of the reaction that presents a protonated Asp96, a neutral Ser53, and a neutral methylated base. This product structure shows hydrogen bond interactions between the N4 atom and the the hydroxyl oxygen atom of Ser53 (2.217 Å), the N4 atom and the carbonyl oxygen atom of Pro54 (1.706 Å), and the hydrogen atom of the hydroxyl group of Ser53 and the hydroxyl oxygen of Ser53 (2.069 Å). As shown in Figure 5, this product is significantly more stable than the reactant (−48.6 kcal·mol^{-1}). The energy barrier

associated with this mechanisms is 20.2 kcal·mol^{-1}, about 9.1 kcal·mol^{-1} lower than the one obtained for the equivalent mechanism at the AM1/MM level.

In the second reaction mechanism we located a transition structure (TS3) in which proton transfer from Ser53 to Asp96 is nearly completed (the distance from the transferred hydrogen to the carboxylate oxygen atom of Asp96 is 1.031 Å) and the proton transfer from N4 of cytosine to the hydroxyl oxygen atom of Ser53 is in a much earlier stage (the distances from the transferred hydrogen to the donor and the acceptor atoms are 1.218 and 1.339 Å, respectively). From this transition structure the reaction proceeds to an intermediate (Int2) where the deprotonated base is stabilized by means of hydrogen bond interactions with Ser53 (1.774 Å) and Pro54 (1.993 Å). In this intermediate the distances from N4 of cytosine to the sulfur atom of AdoMet and from N4 of cytosine to the methyl group are already reduced (4.738 and 2.983 Å, respectively). A second transition structure (TS4) was located corresponding to a methyl transfer where, different from the AM1/MM results, the carbon−N4 distance (2.411 Å) is larger than the carbon−S one (2.111 Å). The BOs calculated at the B3LYP level (0.649 and 0.207, for the C−S and C−N4 bonds respectively) reflect the earlier nature of the transition state, where the bond breaking and bond forming processes are not very advanced. Methylation of the exocyclic N4 atom of cytosine is accompanied by a lengthening of the hydrogen bond distances established between this atom and Ser53 and Pro54 residues. Note that methylation takes place with a larger donor−acceptor (N4(Cyt)−S(AdoMet)) distance and a less linear N4−CH_3−S arrangement than that at the AM1/MM level. At the B3LYP level the rate-determining step in this second mechanism (see Figure 5) corresponds to the proton transfers, with an energy barrier relative to reactant of 11.1 kcal·mol^{-1}, estimated after single-point B3LYP/MM calculations using the 6-311+G** basis set. The rate-determining step for this mechanism at the AM1/MM level was the methyl transfer. It is also important to note that the energy barriers obtained at the DFT/MM level are significantly smaller than the ones obtained by the AM1/MM level, being in better accordance with the range of values expected for an enzymatic reaction.

Reaction Mechanism for N4 Methylation in M.PvuII. According to the results presented above, methylation of N4 atom of cytosine catalyzed by M.PvuII is an exothermic reaction (−31.6 and −48.6 kcal·mol^{-1} at the AM1/MM and B3LYP/MM levels, respectively). The driving force for this reaction seems to be the charge annihilation. Effectively the reaction starts with a positively charged AdoMet and a negatively charged Asp96. As a reference, the calculated reaction energy for the gas phase process $S(CH_3)_3^+ + Cyt + CH_3COO^- \rightarrow S(CH_3)_2 + Cyt\text{-}CH_3 + CH_3COOH$ is −130.7 kcal·mol^{-1} at the same B3LYP level. In the gas phase model of the active site (see Methods) the reaction energy is calculated to be −42.6 kcal·mol^{-1} at the same DFT level. Obviously, these values can be substantially modified when we move to free energies because charge annihilation may be accompanied by a substantial reorganization of the environment. With respect to the mechanism, the catalyzed process can take place, in principle, through two different paths: methylation and then deprotonation or deprotonation and after that methylation. Deprotonation occurs through two residues of the active site Ser53 and Asp96. A hydrogen atom of the amino group of cytosine is transferred to the oxygen atom of the hydroxyl group of Ser53 while the hydrogen of this group is transferred to a carboxylate oxygen atom of Asp96. Interestingly, at both theoretical levels (AM1/MM and B3LYP/MM) and independently of the relative timing with respect to methylation, these proton transfers take place in a single but asynchronous step where the proton transfer from Ser53 to Asp96 clearly precedes the proton transfer from N4 of cytosine to Ser53. This proton transfer is barrierless if it occurs after methylation of cytosine at the B3LYP/MM level. With respect to the methylation step the nature of the transition structure depends on the protonation state of N4 of cytosine. If methylation takes place first, then the transition structure shows a methyl group closer to the acceptor atom than to the donor and a large potential energy barrier. Instead, if the exocyclic amino group of cytosine is deprotonated prior to methylation, then the transition state is geometrically earlier (closer to the corresponding previous intermediate geometry) and presents a significantly lower potential energy barrier measured from the intermediate structure preceding the methyl transfer.

Both theoretical treatments agree with the fact that mechanisms where deprotonation precedes methylation are kinetically favored, presenting an energy barrier lower than the other mechanism (methylation and then deprotonation). The nature of the rate-determining step, however, changes from the semiempirical to the DFT description. In the first case the highest transition structure corresponds to the methyl transfer with an energy relative to the reactants of 27.2 kcal·mol^{-1}. When the B3LYP description is used for the QM subsystem, the highest transition structure corresponds to the deprotonation step of the exocyclic amino group of the base and the proton transfer from Ser53 to Asp96. The energy of this structure relative to reactants is only of 11.1 kcal·mol^{-1}, a value much more reasonable for this enzymatic process taking into account that in M.PvuII the release of products and not the chemical reaction is the rate-limiting step.[16,17]

4. Conclusions

Two possible reaction mechanisms for exocylic nitrogen (N4) methylation of cytosine base in the active site of M.PvuII have been explored using a semiempiririical (AM1) and DFT (B3LYP) hybrid QM/MM methods and a model consisting of the complete enzyme, a cytosine base, and the solvation water molecules. After AM1/MM MD simulations, a reactant structure was obtained after full relaxation and used as a starting point for PES explorations. A gas phase model, where only some atoms of the active site with restricted mobility are considered, was used to explore different reaction mechanisms determining the more reliable possibilities, which were subsequently studied in the enzymatic model. These reaction mechanisms consist of (i) the N4 atom of cytosine being methylated and then deprotonated and (ii) the N4 atom of cytosine being deprotonated and then methylated. In both cases the deprotonation takes place through a proton relay mechanism involving Ser53 and the final proton acceptor Asp96. Our studies point out that the second mechanistic possibility (deprotonation followed by methylation) presents the lower energy barrier using either AM1/MM or B3LYP/MM methods. At the highest theoretical level the rate-determining step of this mechanism seems to be the proton transfer step, with a transition structure where the proton transfer from Ser53 to Asp96 is nearly finished while the proton transfer from the N4 atom of cytosine to Ser53 is at a much earlier stage. The potential energy barrier determined after single-point calculations using the 6-311+G** basis set is 11.1 kcal·mol^{-1}.

Obviously, the consideration of the whole substrate, a DNA sequence, and of the dynamics of the complex are important factors for a full understanding of this important enzymatic process. However, exploration of the more plausible mechanisms

using simplified models allows us to employ higher QM treatments, such as the B3LYP/MM method employed here. These studies should then be useful to determine the reaction mechanism to be explored in future studies and to determine the reliability of faster QM methods that must be used in MD simulations. Further studies in this direction are currently being developed in our laboratory.

Acknowledgment. This work was supported by DGI project CTQ2009-14541-C02-02. V.L.-C. and M.R. thank the Ministerio Ciencia e Innovación for a doctoral grant and a "Juan de la Cierva" contract, respectively. J.A. thanks "La Caixa" for a master fellowship. We acknowledge computational facilities of the Servei d'Informàtica de la Universitat de València in the "Tirant" supercomputer, which is part of the Spanish Supercomputing Network.

Supporting Information Available: PDB file containing the coordinates of all the atoms belonging to residues placed less than 10 Å from the substrate in the AM1/MM reactant structure. This material is available free of charge via the Internet at http://pubs.acs.org.

References and Notes

(1) Jeltsch, A. *Chembiochem* **2002**, *3*, 382.
(2) Cheng, X. D.; Roberts, R. J. *Nucleic Acids Res.* **2001**, *29*, 3784.
(3) Gingeras, T. R.; Greenough, L.; Schildkraut, I.; Roberts, R. J. *Nucleic Acids Res.* **1981**, *9*, 4525.
(4) Gong, W. M.; Ogara, M.; Blumenthal, R. M.; Cheng, X. D. *Nucleic Acids Res.* **1997**, *25*, 2702.
(5) Klimasauskas, S.; Kumar, S.; Roberts, R. J.; Cheng, X. D. *Cell* **1994**, *76*, 357.
(6) Klimasauskas, S.; Szyperski, T.; Serva, S.; Wuthrich, K. *EMBO J.* **1998**, *17*, 317.
(7) Klimasauskas, S.; Roberts, R. J. *Nucleic Acids Res.* **1995**, *23*, 1388.
(8) Yang, A. S.; Shen, J. C.; Zingg, J. M.; Mi, S.; Jones, P. A. *Nucleic Acids Res.* **1995**, *23*, 1380.
(9) Roth, M.; Helm-Kruse, S.; Friedrich, T.; Jeltsch, A. *J. Biol. Chem.* **1998**, *273*, 17333.
(10) Sugisaki, H.; Yamamoto, K.; Takanami, M. *J. Biol. Chem.* **1991**, *266*, 13952.
(11) Willcock, D. F.; Dryden, D. T. F.; Murray, N. E. *EMBO J.* **1994**, *13*, 3902.
(12) Guyot, J. B.; Grassi, J.; Hahn, U.; Guschlbauer, W. *Nucleic Acids Res.* **1993**, *21*, 3183.
(13) Kong, H. M.; Smith, C. L. *Nucleic Acids Res.* **1997**, *25*, 3687.
(14) Warshel, A.; Narayszabo, G.; Sussman, F.; Hwang, J. K. *Biochemistry* **1989**, *28*, 3629.
(15) Schluckebier, G.; Labahn, J.; Granzin, J.; Saenger, W. *Biol. Chem.* **1998**, *379*, 389.
(16) Adams, G. M.; Blumenthal, R. M. *Biochemistry* **1997**, *36*, 8284.
(17) Bheemanaik, S.; Reddy, Y. V. R.; Rao, D. N. *Biochem. J.* **2006**, *399*, 177.
(18) Cerjan, C. J.; Miller, W. H. *J. Chem. Phys.* **1981**, *75*, 2800.
(19) Simons, J.; Jorgensen, P.; Taylor, H.; Ozment, J. *J. Phys. Chem.* **1983**, *87*, 2745.
(20) Nichols, J.; Taylor, H.; Schmidt, P.; Simons, J. *J. Chem. Phys.* **1990**, *92*, 340.
(21) Antosiewicz, J.; McCammon, J. A.; Gilson, M. K. *J. Mol. Biol.* **1994**, *238*, 415.
(22) Field, M. J.; David, L.; Rinaldo, D. Personal communication.
(23) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926.
(24) Dewar, M. J. S.; Zoebisch, E. G.; Healy, E. F.; Stewart, J. J. P. *J. Am. Chem. Soc.* **1985**, *107*, 3902.
(25) Jorgensen, W. L.; Tiradorives, J. *J. Am. Chem. Soc.* **1988**, *110*, 1657.
(26) Pranata, J.; Wierschke, S. G.; Jorgensen, W. L. *J. Am. Chem. Soc.* **1991**, *113*, 2810.
(27) Field, M. J.; Albe, M.; Bret, C.; Proust-De Martin, F.; Thomas, A. *J. Comput. Chem.* **2000**, *21*, 1088.
(28) Field, M. J. *A practical Introduction to the Simulation of Molecular Systems*; Cambridge University Press: Cambridge, U.K., 2007.
(29) Singh, U. C.; Kollman, P. A. *J. Comput. Chem.* **1986**, *7*, 718.
(30) Field, M. J.; Bash, P. A.; Karplus, M. *J. Comput. Chem.* **1990**, *11*, 700.
(31) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, J. A., Jr.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, M.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchain, H. P.; 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.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Danniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaroni, I.; Martin, R. L.; Fox, D. L.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; Pople, J. A. *Gaussian 03*, revision D.02; Gaussian Inc.: Wallingford, CT, 2004.
(32) Gonzalez, C.; Schlegel, H. B. *J. Chem. Phys.* **1989**, *90*, 2154.
(33) Gonzalez, C.; Schlegel, H. B. *J. Phys. Chem.* **1990**, *94*, 5523.
(34) Becke, A. D. *Phys. Rev. A* **1988**, *38*, 3098.
(35) Lee, C. T.; Yang, W. T.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785.
(36) Moliner, V.; Turner, A. J.; Williams, I. H. *Chem. Commun.* **1997**, 1271.
(37) Turner, A. J.; Moliner, V.; Williams, I. H. *Phys. Chem. Chem. Phys.* **1999**, *1*, 1323.
(38) Vreven, T.; Morokuma, K.; Farkas, O.; Schlegel, H. B.; Frisch, M. J. *J. Comput. Chem.* **2003**, *24*, 760.
(39) Prat-Resina, X.; Gonzalez-Lafont, A.; Lluch, J. M. *J. Mol. Struct. (THEOCHEM)* **2003**, *632*, 297.
(40) Monard, G.; Prat-Resina, X.; Gonzalez-Lafont, A.; Lluch, J. M. *Int. J. Quantum Chem.* **2003**, *93*, 229.
(41) Prat-Resina, X.; Bofill, J. M.; Gonzalez-Lafont, A.; Lluch, J. M. *Int. J. Quantum Chem.* **2004**, *98*, 367.
(42) Marti, S.; Moliner, V.; Tunon, I. *J. Chem. Theory Comput.* **2005**, *1*, 1008.
(43) Zhang, Y. K.; Liu, H. Y.; Yang, W. T. *J. Chem. Phys.* **2000**, *112*, 3483.
(44) Lameira, J.; Alves, C. N.; Moliner, V.; Marti, S.; Kanaan, N.; Tunon, I. *J. Phys. Chem. B* **2008**, *112*, 14260.
(45) Chirlian, L. E.; Francl, M. M. *J. Comput. Chem.* **1987**, *8*, 894.
(46) Shieh, F. K.; Youngblood, B.; Reich, N. O. *J. Mol. Biol.* **2006**, *362*, 516.
(47) Pauling's bond orders (BOs) were calculated using the equation; $BO = \exp((R_0 - R_{TS})/c)$ with $c = 0.6$, where R_0 is the reference bond and R_{TS} is the bond at the transition structure.

JP911036W