ACS Catalysis

Research Article

pubs.acs.org/acscatalysis

QM/MM MD Simulations on the Enzymatic Pathway of the Human Flap Endonuclease (hFEN1) Elucidating Common Cleavage Pathways to RNase H Enzymes

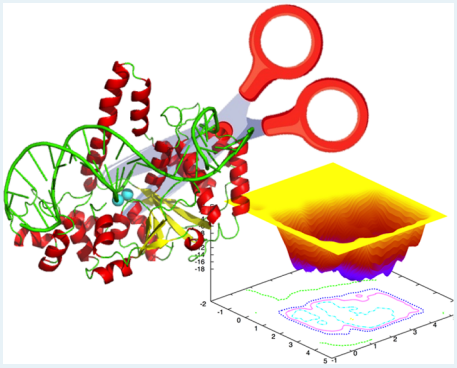
Jacopo Sgrignani*,† and Alessandra Magistrato*,‡

†Institute of Research in Biomedicine (IRB), Via Vincenzo Vela, 6500 Bellinzona, Switzerland

‡CNR-IOM-Democritos National Simulation Center c/o International School for Advanced Studies (SISSA/ISAS), Via Bonomea 265, 34136 Trieste, Italy

Supporting Information

ABSTRACT: Flap endonucleases (FENs) are nucleic acid hydrolyzing enzymes in charge of excising 5′-small DNA and RNA fragments (flaps) protruding from nucleic acid structures during the lagging strand DNA replication or the long-patch base excision repair (LP-BER) processes. In this work we report, for the first time, an atomistic and energetic rendering of the enzymatic catalysis promoted by the human FEN1. After reconstruction of a reactive hFEN/double strand (ds) DNA adduct we employed mixed quantum-classical (QM/MM) metadynamics and umbrella sampling free energy calculations, with the QM part treated with the AM1/d-PhoT Hamiltonian, to perform an extensive characterization of all possible reaction pathways underlying the enzymatic cycle. Our extensive investigation points to a most likely reaction pathway very similar to that recently proposed for ribonuclease H, in which the rate-determining step is the nucleophilic attack of a water to the scissile phosphate, which occurs concomitantly with its activation by the pro-*R*p oxygen of the nucleobase flanking the scissile phosphate. This step requires a free energy barrier in good agreement with experimental data ($\Delta G^{\ddagger}_{exp}$ = 16.1 kcal/mol vs $\Delta F^{\ddagger}_{calc}$ = 16 ± 2 kcal/mol). Due to the important role of FENs in maintaining nucleic acid fidelity and cell proliferation, a detailed understanding of its enzymatic mechanism has broad interest to elucidate a key enzymatic biological process for preserving genome integrity and has implications for medical and biotechnological applications.



KEYWORDS: *flap endonucleases, QM/MM, umbrella sampling, metadynamics, catalytic mechanism*

INTRODUCTION

All cells have genomes encoded in double-stranded (ds) DNA, which hence needs accurate replication and repair. Flap endonucleases 1 (FEN1) are metal-dependent and structure-specific nucleic acid hydrolyzing enzymes. These specifically cleave one of the two P−O phosphodiester bonds of the 5′-flaps (single stranded protrusions) of DNA and RNA, resulting in flap removal, regardless of the nucleic acid sequence.

DNA polymerase can form flaps during the DNA replication of the lagging strand. This latter is composed by RNA-primed discontinuous fragments (Okazaki fragments), later joining to form the novel DNA filament. Remarkably, DNA replication in humans generates 50 million Okazaki primers at each cell cycle. Since each primer has to be removed to form the novel DNA strand, FEN1 must be an extremely efficient enzyme, speeding the catalytic rate of phosphate hydrolysis up to 10^{17} times.[1,2]

Similarly, during long-patch base excision repair (LP-BER), single-strand gaps in dsDNA molecules, originating from the excision of damaged bases, are repaired by adding 2−10 newly synthesized nucleotides. In this process a small flap, protruding from the DNA structure, is formed and has to be cleaved by FEN1 before ligation of the repaired strand can take place.[3]

Importantly, for replication and repair, FENs must ensure that incision takes place efficiently and specifically at the correct location in order to produce connectable nucleic acid fragments. Failure to precisely remove the primers creates gaps or overlaps, impairing cell division and the maintenance of genome fidelity.[1]

Consistent with its important role in DNA replication and repair, FEN1 is overexpressed in all proliferative tissues so that high levels of FEN1 are believed to support cancer cell hyperproliferation.[3] Hence, FEN1 inhibitors have been proposed as molecules with potential chemotherapeutic activity.[4−7]

In addition to flaps, other aberrant DNA structures require the action of structure-specific nuclease enzymes, which belong to a conserved FEN1 5′-superfamily.[1,8,9] It is therefore a challenge to understand how sequence-related proteins are capable of recognizing structurally diverse aberrant DNA structures. Concerning FEN1, a detailed understanding of its

Received: January 28, 2015
Revised: May 4, 2015
Published: May 11, 2015

DOI: 10.1021/acscatal.5b00178
ACS Catal. 2015, 5, 3864−3875

DNA recognition process and its the nucleotide excision reaction catalysis has been so far limited by a lack of structural information. Recently, Tsutakawa et al.[10] reported three-dimensional structures of human FEN1 (hFEN1) complexed with DNA for both the reactant and the products (PDB codes 3Q8L and 3Q8K, respectively), enabling key advancements in the comprehension of the biological mechanism of this enzyme.[1,10] This structure revealed that the catalytic site selectively recognizes the 5′-flaps thanks to a so-called "bind−push−pull−unpair" mechanism.[8,10,11]

Once the 5-flap enters into the catalytic site, the hydrolytic reaction is promoted by two Mg^{2+} ions (Figure 1). According to experimental and computational studies,[2,12−15] the cleavage of the phosphodiester bond is believed to occur via a general acid−base catalysis where the nucleophile is activated by a basic group. The reaction then proceeds as an S_N2 nucleophilic substitution in three steps: (1) nucleophile attack, (2) formation of a putative pentavalent negatively charged intermediate (phosphorane), and (3) cleavage of the scissile bond (Figure 2).[16]

Figure 1. Molecular model of the hFEN1/DNA precatalytic adduct as a result of our MD simulations. The scissile nucleotide is depicted in orange van der Waals spheres (vdW), and the two magnesium ions are shown as magenta vdW spheres.

As in other nuclease enzymes, the presence of two Mg^{2+} ions in the catalytic site appears to be mandatory for the enzymatic activity of hFEN1.[12] In the past decade, quantum mechanics−molecular mechanics (QM/MM)[17] MD simulations have become established valuable tools to study enzymatic reactions,[18−33] including those involved in the formation and cleavage of nucleic acid filaments.[13,14,34−41] Here, we have employed this method to characterize the structure of the hFEN1/DNA adduct in the presence of catalytically active Mg^{2+} ions, which were replaced by Sm^{3+} ions in the crystals, and to elucidate its enzymatic reaction mechanism. To this aim we performed both force-field and QM/MM MD simulations, where the QM part of the system was treated at the AM1/d-PhoT level of theory,[42,43] along with metadynamics (MTD)[44−46] and umbrella sampling (US) simulations to gather a picture of the free energy landscape underlying the enzymatic process.[47−49]

The choice of a semiempirical approach for the QM part is due to the lack of accurate structural information on the reactive FEN1/dsDNA adduct in the presence of Mg^{2+} ions. In fact, our experience in performing QM/MM MD simulations in different biological systems[29,33,50,51] teaches us that it is very challenging to study a chemical reaction in a biosystem when, as in the present case, accurate structural information about the reactive adduct is lacking. Thus, a semiempirical Hamiltonian, properly tuned for this kind of enzymatic reaction,[43] is preferable in this case as its moderate computational cost allows (i) performing a long sampling of the starting adduct, resulting in a good relaxation of the active site, (ii) investigating a plethora of possible reaction routes, and (iii) considering very large QM regions. Here, in fact, we performed a total of ∼1 ns of QM/MM MD simulations, which would be barely affordable with density functional theory (DFT) based methods and would certainly be too risky to be carried out at a more accurate level of theory in the absence of an accurate starting structure.

Despite the limitations in accuracy of the theoretical approach, our results suggest that the most likely reaction path of the enzymatic cycle involves nucleophilic attack on the scissile phosphate carried out by a $P-O_{proRp}$ oxygen activated water (the rate-determining step), followed by the formation of a stable phosphorane intermediate and by the protonation of the O3′ leaving group promoted by a protonated Asp86. This mechanism is fully consistent with that recently proposed by Rosta et al.[14] for RNase H enzymes, pointing for the first time to common mechanistic features between these two families of endonuclease enzymes.

METHODS

Model Building and Force Field Parameters for Classical MD Simulations. The starting structure of our simulation was that of hFEN1 in complex dsDNA substrate (PDB code 3Q8L).[10] We replaced the two Sm^{3+} ions, present in the crystal structure,[10] with Mg^{2+} ions, as hFEN1 is a Mg^{2+}-dependent enzyme. The sulfate bridging the two Sm^{3+} ions was deleted, while the K^+ ion, potentially having a structural role in stabilizing the binding of the substrate,[10] was retained. Since in the X-ray structure of the hFEN1/substrate complex the scissile phosphate is located ∼7 Å away from the metal ions, we constructed a reactive adduct on the basis of the structural information on the enzyme−product complex (PDB code 3Q8K[10]), in which the scissile phosphate of the product is placed between the two Sm^{3+} ions.

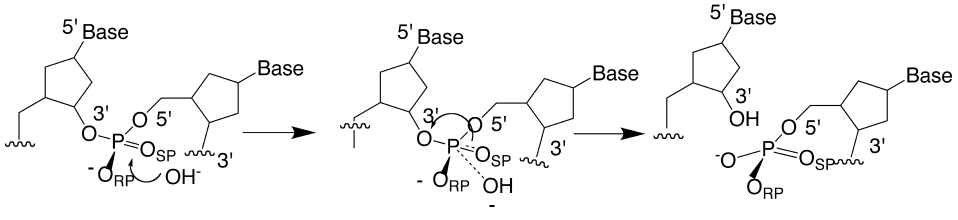


Figure 2. General reaction mechanism of nuclease enzymes: (a) reactant; (b) phosphorane intermediate; (c) product.

DOI: 10.1021/acscatal.5b00178
ACS Catal. 2015, 5, 3864−3875

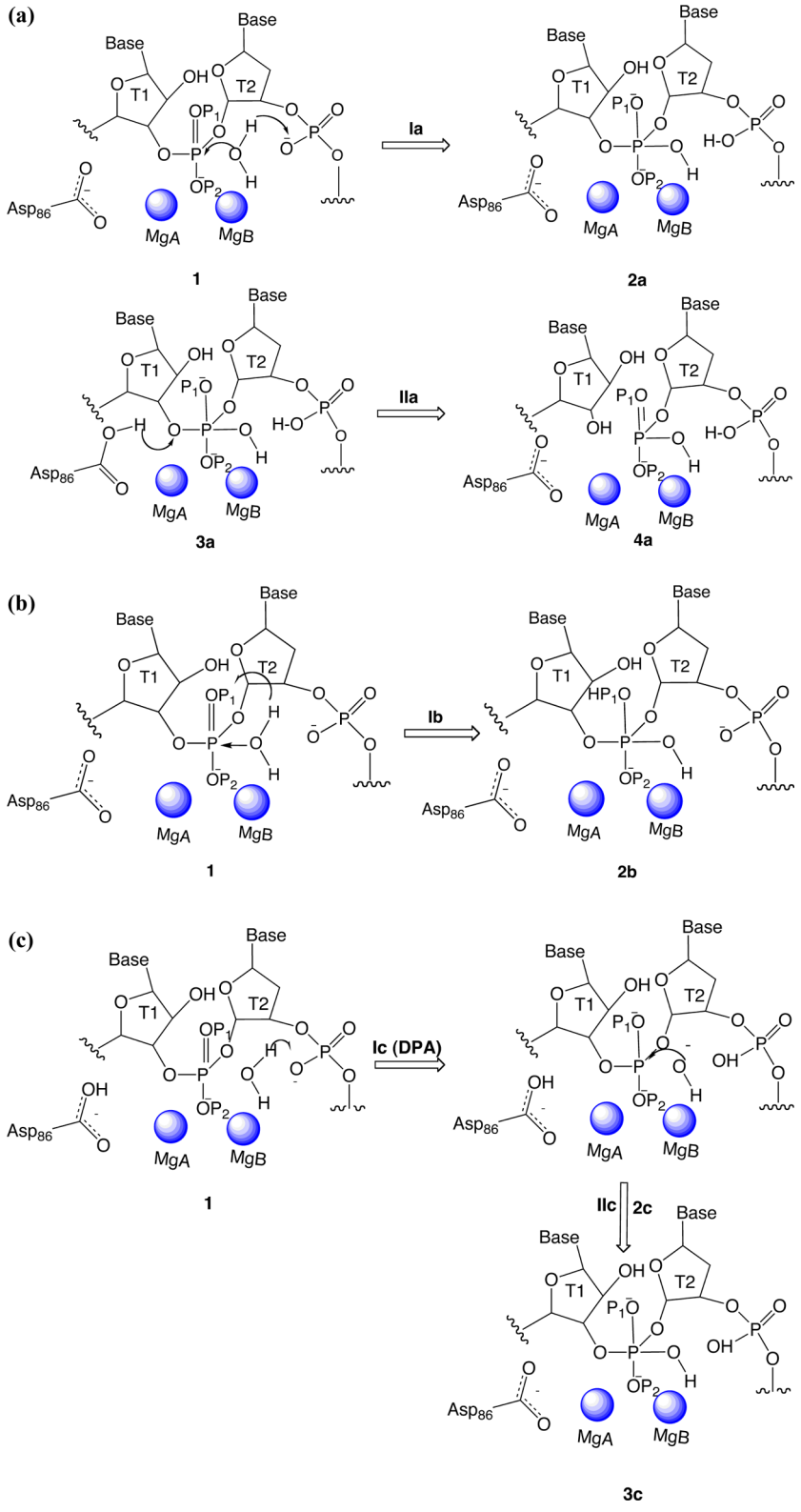


Figure 3. Scheme of the reaction mechanisms investigated in this study: (a) reaction path A; (b) reaction path B; (c) reaction path C. In the last two cases the enzymatic cycle ends as in step IIa.

The force-field (FF) parameters were assigned according to Amber-ff10,[52,53] using the leap module of Amber12.[54] During FF-based MD simulations the two Mg^{2+} ions were described using the dummy cation model developed by Wharshel and co-workers[55] that has been proven to well describe complexes involving nucleic acids.[35,56]

The complex was inserted in a TIP3P water box with a minimum distance of 12 Å from the nearest atom of the protein−DNA complex, and its overall charge was neutralized by adding 36 or 35 Na^+ ions, depending on the protonation state of the active site residues considered in the simulation. For

DOI: 10.1021/acscatal.5b00178
ACS Catal. 2015, 5, 3864−3875

those ions we used the parameters of Cheatham and co-workers.[57,58]

In order to obtain a reactive enzyme/substrate adduct, the system was first minimized and equilibrated using some distance restraints based on the enzyme/product complex (PDB code 3Q8K). These restraints were removed after 5 ns, and the entire system was simulated in an NPT ensemble (pressure 1 atm; temperature 298.5 K) for 50 ns, with a time step of 1.5 fs, using the PMEMD code in its GPU accelerated version.[59,60] The pressure and the temperature were regulated using the Berendsen pressure and temperature coupling algorithms.[61] The nonbonded interactions were evaluated using a cutoff of 10 Å, whereas long-range electrostatic interactions were assessed using the particle mesh Ewald method.[62]

QM/MM MD. *Model Systems.* All QM/MM MD simulations were run using the sander code available in the Amber12[54] package. The MM part of the system was described using Amber-ff10[52,53] as in MM calculations. The QM part, formed by the two nucleotides flanking the scissile phosphate (i.e Thymine1 and Thymine2, T1 and T2), the residues coordinating the Mg^{2+} ions (Asp86, Glu158, Glu160, Asp179, Asp181, and Asp233), and a water molecule (WatR), identified during FF-MD simulation as potentially involved in the enzymatic reaction, was described using the AM1/d-PhoT Hamiltonian,[43] while the two Mg^{2+} ions were described with the AM1/d model.[42] The QM region was saturated with capping H atoms. The QM part comprised 145 atoms; the remaining MM part was composed of 67000 atoms. In some steps the QM region was enlarged. In particular, Guanine3 (G3) was included in the QM part of the system only in the simulations of step Ia (the nucleophilic water deprotonation; Figures 3 and 4), while we included in the QM region a protonated Asp86 or Lys93 for the protonation of the leaving group (steps IIa, IId, and IIe; Figures 3 and 4).

Notably, Crehuet and co-workers reported good performances for the AM1/d-PhoT Hamiltonian in describing a pentacoordinated phosphorus also in comparison with DFT calculations.[63] Moreover, this level of theory has been successfully used in the study of enzymatic reactions promoting the formation and the cleavage of bonds in nucleic acids, such as those investigated here,[64−66] and in other biological systems.[64,67−69] Despite this, it is certainly true that the accuracy of this semiempirical approach strongly depends on the system investigated.[31,70]

The last frame of the FF-MD simulations was considered as a starting point for QM/MM MD. The whole system was simulated in the NVT ensemble, and the temperature was regulated using a Langevin thermostat.[71] For this reason all subsequent energetic values are reported as Helmholtz free energies. The nonbonded interactions were calculated considering a cutoff of 10 Å. Initially, the system was slowly heated using a time step of 0.5 fs for 36 ps. Then, considering the lack of an experimentally solved Michaelis complex, the system was simulated for an additional 100 ps with the aim of obtaining a well-equilibrated starting structure for the reactivity studies. In this latter simulation a larger time step (2 fs) was used for the force integration; consequently, the QM part of the system was constrained using the SHAKE algorithm.[72] Short- and long-range electrostatic interactions were estimated as in FF-MD.[62,73]

QM/MM Metadynamics. Since the first step of phosphodiester hydrolysis has been subject of debate in previous

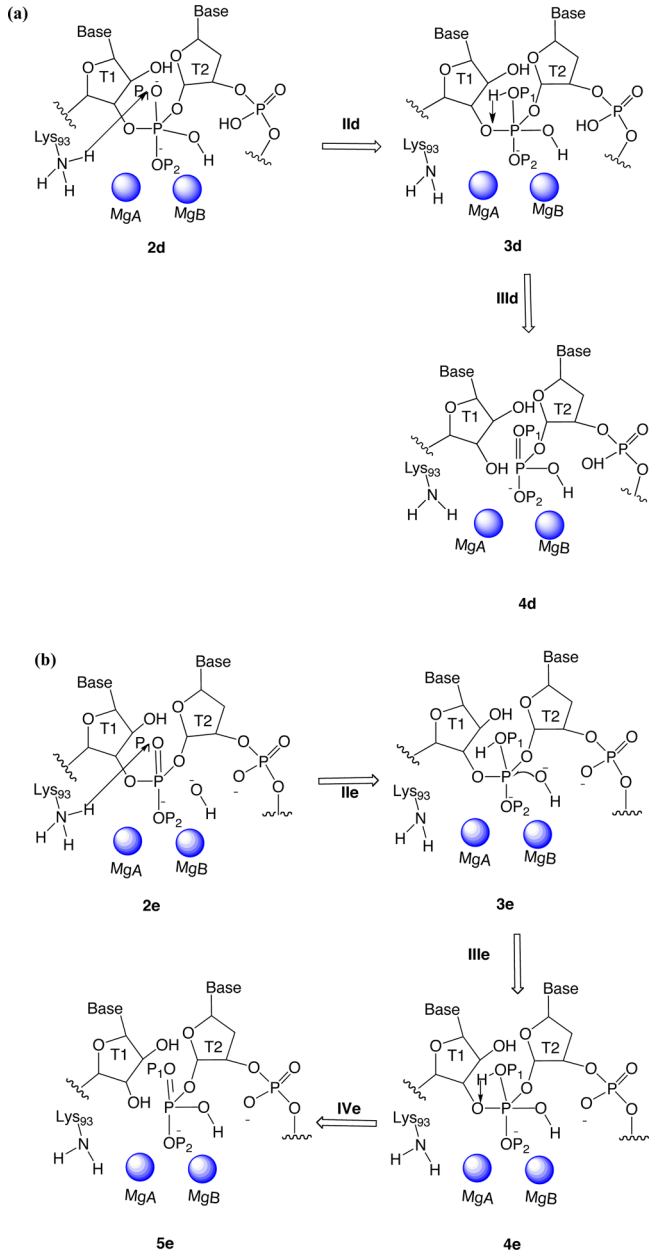


Figure 4. Scheme of alternative reaction paths involving Lys93: (a) reaction path D (only steps II and III have been investigated); (b) reaction path E (step Ie is not shown, as it is identical with step Ic).

computational studies,[13,14] we have investigated this step (Ia and Ib, Figure 3) using metadyamics (MTD) simulations. MTD is a computational technique, first introduced by Laio and Parrinello,[46] aimed at accelerating rare events. This method has been successfully applied to study chemical reactivity in several biological systems.[18,28,45,74−77] Hence, two collective variables (CVs) have been selected: CV1, the difference in the distance between one of the WatR hydrogen atoms and the nearest oxygen of the T2 phosphate group (pro-*Rp*O@T2) group and the distance between the abstracted hydrogen atom of WatR and the O@WatR, and CV2, the difference between the O@WatR−P@T1 distance and the O3′−P@T1 distance (graphical representations of all CVs used in MTD and US simulations are reported in the Figures S1−S6 in the Supporting Information). The simultaneous use of the two CVs in MTD simulations was chosen to understand their

DOI: 10.1021/acscatal.5b00178
ACS Catal. 2015, 5, 3864−3875

interdependence, as according to Rosta et al. these two CVs are correlated.[13]

In a second attempt (step Ib, Figure 3b) we also considered the mechanism proposed by De Vivo et al.[13] To this aim we used as reaction coordinates the difference in the distance between the OP1@T2 and the H1@WatR and the distance between H1@WatR and O@WatR (CV3) as well as the distance between O@WatR and P@T2 (CV4) (Figure S2 in the Supporting Information).

Repulsive Gaussian potentials with a height of 0.70−0.50 kcal/mol and a width of 0.15 Å were placed in the CV space every 15 fs. In order to increase the computational efficiency of the calculations, we used MTD in its multiple walker formulation.[78] During this type of calculation more replicas of the system (six to eight in this case) are simultaneously run and at a predefined time interval information about bias is exchanged. Consequently, the time necessary to explore the CV space is proportionally reduced with respect to number of replicas. During our simulation we extended the calculations until reaction was observed in a reasonable number of replicas (two to three). MTD was employed to define the model reaction path with the minimum activation free energy approach; consistently no energetic data on the reaction free energy derived from MTD simulations are given. In short, the activation free energy barriers ($\Delta F^{\ddagger}$) were estimated following the recipe suggested in ref 44 and successfully applied in other cases.[75,79] $\Delta F^{\ddagger}$ was measured by considering the bias added just after the system exit from the stable minimum corresponding to the reactant state. We performed three independent MTD simulations with initial velocities randomly assigned from a Maxwell−Boltzmann distribution at 298.5 K. Finally, the reported $\Delta F^{\ddagger}$ was the average over the three simulations and its standard deviation was taken as a measure of the accuracy of the barriers. MTD simulations were run using Amber12 patched with the PLUMED 1.3 code.[80] The total length of each MTD simulation was ∼10 ps/replica for the step I−II (total simulation time 60 ps) and 18 ps/replica for the alternative mechanism proposed by De Vivo et al.[13] (total simulation time 144 ps).

QM/MM Umbrella Sampling. In order to obtain an accurate estimate of reaction free energy profile (FEP), when the reaction path could be described by only one CV, we performed umbrella sampling (US) calculations on the paths hypothesized in previous computational studies.[13,14]

A frame from QM/MM MTD simulations in which the phosphorane intermediate was formed was used as the starting structure for QM/MM US simulations of step II (Figures 3 and 4). Instead, a representative frame of product taken from the last window of the US simulations was the starting structure of the following steps. All US simulations were run using Amber12 patched with the PLUMED 1.3 code.[80]

We studied 8 reaction steps with US performing a total of 75 restrained QM/MM MD simulations (total simulation time 329 ps). The data from US calculations were elaborated using the weighted histogram analysis method as implemented in the program of Grossfield (http://membrane.urmc.rochester.edu/content/wham).[81] The statistical uncertainties of the PMF profile were estimated using Monte Carlo bootstrap error analysis. Additional details about calculations for the single reaction steps are available as Supporting Information. Interatomic distances monitored during simulations analysis, root-mean-square deviation (RMSD), values and the images reported in this work were elaborated using VMD[82] or Pymol software (www.pymol.org).

Preactivation of the Nucleophilic Water. We have also considered a path involving a preactivation of the nucleophilic water. In order to calculate the energy gain of the binding of an OH^- in the vicinity of the active site or to the metal ion, we performed AM1/d-PhoT and DFT/B3LYP[83−85] calculations on reduced model systems of the active site comprising the two Mg^{2+} ions, the side chains of Asp86, Glu158, Glu160, Asp179, and Asp181, the backbone of Thymines 1 and 2, a water molecule, and/or an hydroxide ion.

DFT calculations have been performed with the Gaussian09 code[86] using a 6-31G(d,p) basis set. Calculations on reduced models were done using implicit solvent. Generalized-Born[87] was used in AM1/d-PhoT calculations and PCM[88] in DFT calculations. The binding energy (ΔE_{bind}) was calculated as $E_{enzOH/H_2O} - E_{enz} - E_{OH/H_2O}$, where the first is the total energy of the enzyme with hydroxyl group/water bound to Mg^{2+}, the second is the energy of the enzyme without hydroxyl/water ligand, and the third is the energy of the OH^-/water. The water substitution energy ΔE_{sob} is reported as $\Delta E_{bind}^{WAT} - \Delta E_{bind}^{OH}$. The binding energy has been corrected by the free energy cost of water dissociation (20.1 kcal/mol), and we have considered the experimental free energy barrier $\Delta G^{\ddagger}$ of water dissociation (23.9 kcal/mol).[89] Finally we corrected the ΔG value of the OH/water exchange reaction by the entropy cost of binding confining an OH^- ion from the bulk to a restricted volume of 1−3 water molecules following the formula

$$\Delta S_{tr} = -K_b T \ln C_0 \Delta V$$

where C_0 is the standard concentration (for $[OH^-] = 1 \times 10^{-7}$ M) and ΔV is the variation of the volume (here the volume of one to three waters).[90,91]

RESULTS AND DISCUSSION

Model Building and Equilibration. We initially performed 50 ns of classical MD of the potentially reactive hFEN1/DNA adduct. The time evolution of the system was monitored to check both the stability of the structure and to identify the water molecules potentially involved in the enzymatic reaction. The global structure of the complex was well conserved during the whole simulation; actually, the RMSD calculated on Cα atoms of the protein and on the P atoms of the DNA was 2.1 ± 0.3 Å. A visual inspection of the trajectory revealed that one water molecule (WatR) completed the MgA coordination shell (Figure 5). In particular, although this molecule at times exchanges with other bulk water molecules, one water always replaces this position, maintaining a proper geometry to carry out the nucleophile attack on the P atom of the scissile phosphate (i.e., the distance between O@WatR and P@T1 is ∼3.5 Å, in line with the values previously measured in other hydrolytic enzymes).[14,21,25,92,93]

Consistent with mutagenesis studies,[10] the Michaelis complex stability is reinforced by interactions with Lys93 and Arg100, which act as electrostatic clamps. After preliminary FF-MD simulations we equilibrated the system via QM/MM MD. Our simulation revealed that both MgA and MgB coordination distances (Table S1 in the Supporting Information) were in line with the structural parameters observed for this type of biological system.[35] Interestingly, also during QM/MM MD equilibration, WatR conserves a good orientation to carry out the nucleophile attack on the scissile phosphate, always

DOI: 10.1021/acscatal.5b00178
ACS Catal. 2015, 5, 3864−3875

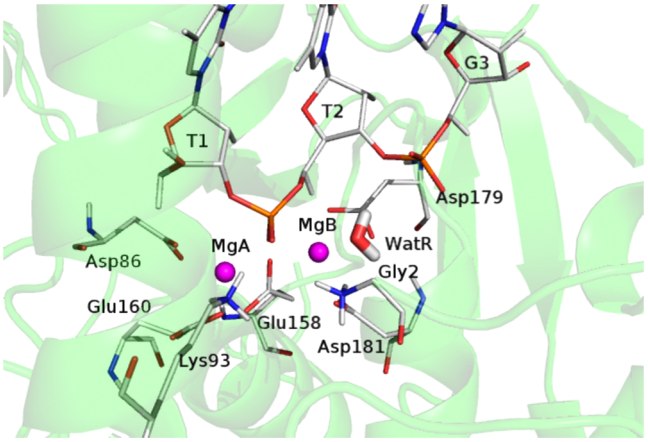


Figure 5. QM/MM equilibrated structure of the hFEN/DNA adduct. Nonpolar hydrogens are omitted for clarity. The Mg ions are shown as VdW spheres and the other atoms as sticks; the nucleophilic water is displayed as thick sticks.

coordinating MgA (the O@WatR−MgA average distance was 2.34 ± 0.07 Å, while the distance between P@T1 and O@WatR was 3.62 ± 0.15 Å (Figure 5)). WatR H-bonds to the pro*Rp*-O atom of the phosphate group belonging to the flanking nucleobase T2 (O1@T2). In the QM/MM MD the MgA−MgB distance relaxed to 3.78 ± 0.11 Å, consistent with a catalytically active bimetallic site.[35,94]

Investigation of Reaction Mechanism by QM/MM MD. *Simulations of Step I.* A consensus view in the scientific literature establishes that enzymatic phosphate hydrolysis proceeds as an S_N2-like nucleophilic attack on the scissile phosphate performed by an hydroxide ion, which is typically formed upon water activation.[13,14,95] Analyzing the initial interactions of WatR and according to the mechanism proposed by Rosta et al. for RNase H,[14] the more promising candidate to activate WatR is O1P@T2 (Figure 3).

Since Rosta et al. suggested that phosphate hydrolysis is initiated by the simultaneous attack of the water and its deprotonation, we performed MTD (CV1 and CV2; Figure S1 in the Supporting Information) simulations of step Ia (Figure 3). This revealed an overall $\Delta F^{\ddagger}_{TSIa}$ value of 16 ± 2 kcal/mol for both the deprotonation and nucleophilic attack (Figure 6 and Figure S7 in the Supporting Information), in good agreement with an experimental $\Delta G^{\ddagger}$ value of 16.1 kcal/mol measured for the rate-determining step.[10]

In two out of the three MTD simulations the nucleophile attack partially occurs with formation of a metastable state (Min2) (Figure 6 and Figure S7 in the Supporting Information), which is followed by the proton transfer, consistent with what was observed by Rosta et al.[14] Notably during MTD simulations the minimum corresponding to the product state was not exhaustively sampled; hence, the ΔF value between **1** and **2** cannot be estimated accurately by MTD simulations. We have therefore calculated this to be −28.4 ± 0.2 kcal/mol by US simulations.

We also verified by MTD simulations (CV3 and CV4; Figure S2 in the Supporting Information) the mechanism proposed by DeVivo et al.,[13] in which WatR is activated by the apical oxygen of the scissile phosphate (Figure 3b). However, these simulations resulted in higher free energy barriers, $\Delta F^{\ddagger}_{TSIb} \approx$ 33 ± 2.6 kcal/mol (Figure 7 and Figure S8 in the Supporting Information), with respect to that estimated for TSIa.

For the sake of completeness we have also investigated the case in which the nucleophilic attack is carried out by a preactivated nucleophile. Water can be activated by a proteic/nucleic acid moiety or by a hydroxide coming from the bulk (deprotonation paths A (DPA) and B (DPB), respectively), or

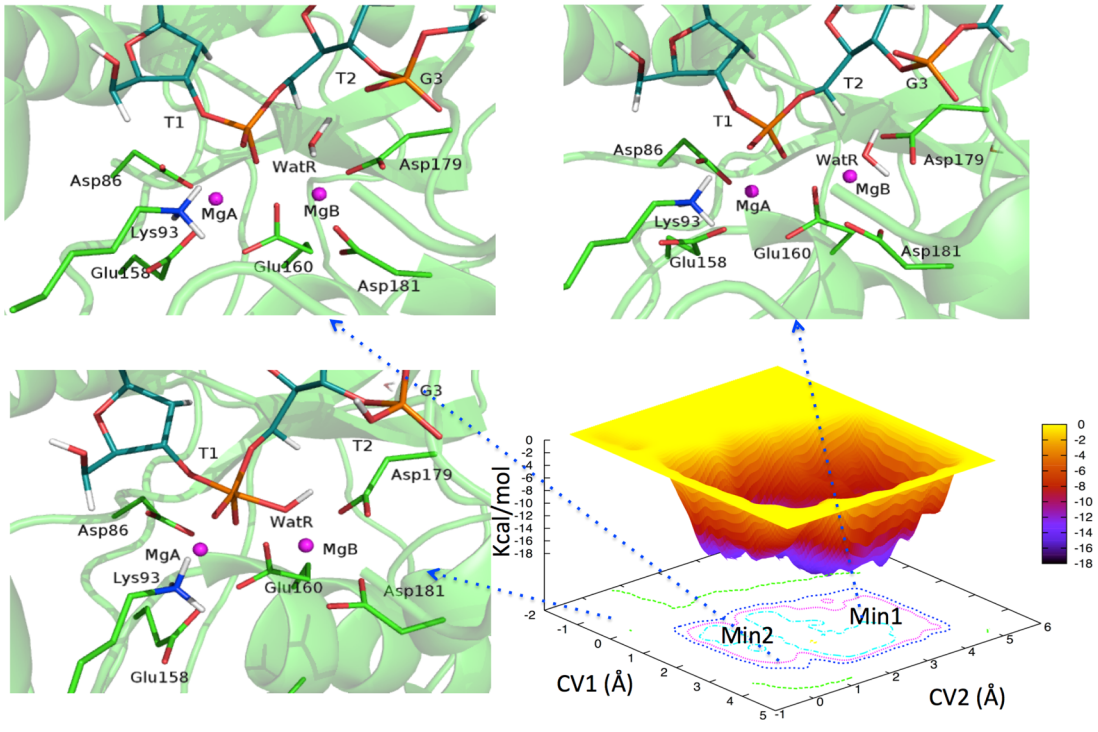


Figure 6. Three-dimensional free energy surface resulting from one of three independent MTDs for step Ia. The structural features of Min1, Min2, and TS_{Ia} are shown.

DOI: 10.1021/acscatal.5b00178
ACS Catal. 2015, 5, 3864−3875

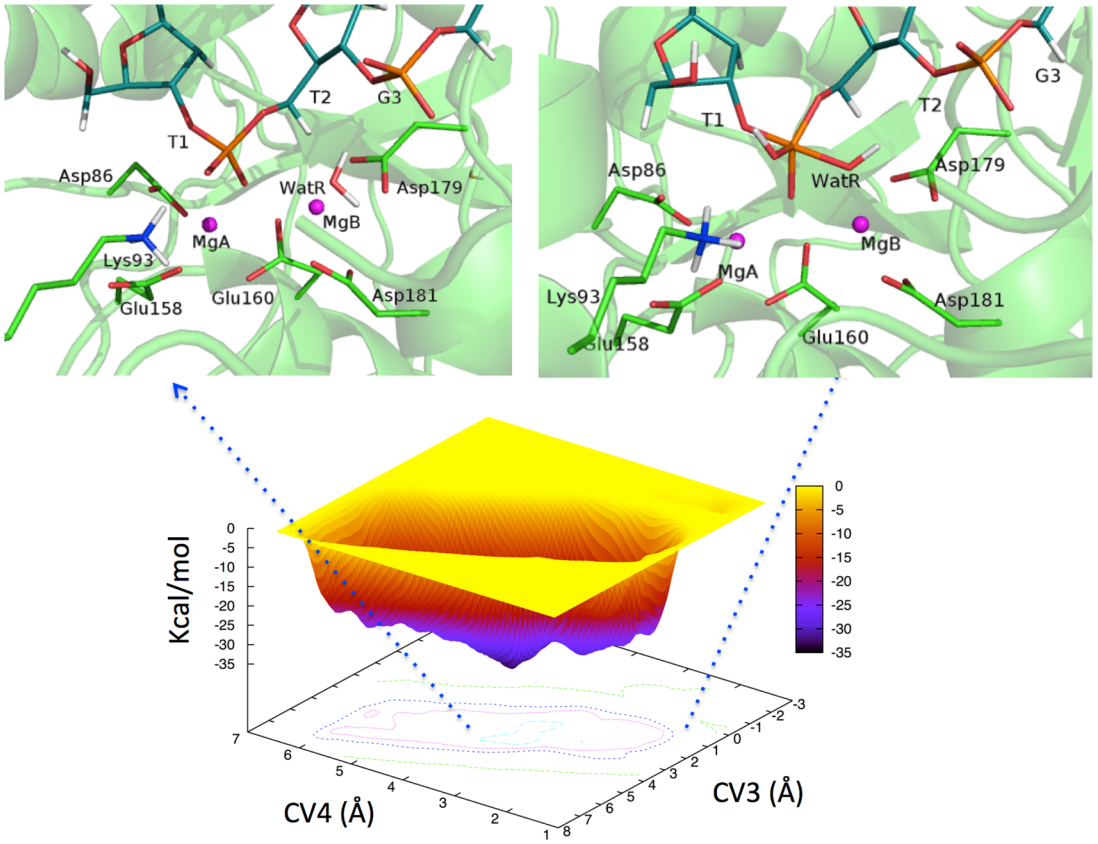


Figure 7. Three-dimensional free energy surface resulting from one of three independent MTD simulations for step Ib. The structural features of **1** and TS_{Ib} are shown.

a hydroxide coming from the bulk can coordinate the metal ion in an active site of an enzyme (DPC).[91]

We have calculated DPA in which the nucleophilic water is activated by OP1@T1. Since this involves a single CV (CV1 in Figure S1 in the Supporting Information), we have done US calculations. This activation requires a $\Delta F^{\ddagger}_{TSIc\ (DPA)}$ value of 13.7 ± 0.8 kcal/mol (Figure S9 in the Supporting Information) and is exothermic by $\Delta F_{Ic(DPA)} = -10.3 \pm 0.2$ kcal/mol.

Concerning the two other possible deprotonation paths, we initially built a system with a OH^- in the vicinity of the nucleophilic water as proposed in ref 91. The binding of OH^- close to the nucleophilic water either deprotonated the nucleophilic water spontaneously or the OH^- replaced the nucleophilic water, both in barrierless processes, with either QM/MM MD or DFT calculations of active site models. Hence, the binding of OH^- in the vicinity of the active site is not a stable state, and DPB and DPC both converge to a unique path. At physiological pH a number of hydroxide ions in solution are present at a low concentration of 10^{-7} mol/dm^3; hence, the translational entropy lost in the confinement of the ion from its volume in bulk solution to a volume of restricted around the active site (one to three water molecules) is known to range between 11.9 and 11.3 kcal/mol.[90,91] Moreover, we should consider the ΔG value of water dissociation in solution.[89] We calculated the energy gain in binding OH^- to the Mg^{2+} by either AM1/d-PhoT MD or DFT/B3LYP[83−85] calculations using reduced model systems of the active site in implicit solvent. The binding energy of OH^- is −66.5 (AM1) and −56.5 kcal/mol (DFT), while the binding of water is exothermic by −17.1 (AM1) and −12.3 kcal/mol (DFT). Thus, considering the entropic cost for the confinement of the ion (12.4 kcal/mol) and the free energy cost of water deprotonation (20.1 kcal/mol)[89] the overall energy gain in replacing the nucleophilic water by an OH^- is −17.3 (AM1) and −12.1 kcal/mol (DFT). However, this path will be kinetically controlled by the free energy cost of water dissociation ($\Delta G^{\ddagger}_{exp}$ = 23.9 kcal/mol).[89] Hence, we explored step IIc (Figure 3c) by US calculations (CV5; Figure S3 in the Supporting Information) considering a preactivated OH^- ion instead of WatR as nucleophilic agent. This reaction step was exothermic ($\Delta F = -18.1 \pm 0.1$ kcal/mol) with a $\Delta F_{TSIIc}^{\ddagger}$ value of 7.4 ± 0.2 kcal/mol (Figure S10 in the Supporting Information), which is in line with the free energy barriers calculated at the DFT level of theory by DeVivo et al. for the reaction step occurring with RNase H.[12]

Simulations of Step II. The remaining possible paths completing the enzymatic cycle (Figures 3 and 4) were investigated by performing US calculations along one selected CV. The formation of a phosphoranic intermediate (**2a**) during enzymatic phosphate hydrolysis has been highly debated in the literature.[13,14,39] Its stability was demonstrated for trans-esterification reactions of RNA filaments,[41,64,96,97] but only metastable structures have been observed in DFT calculations where the hydrolysis is performed by water or hydroxide nucleophilic agent.[13] Thus, we carefully analyzed the structural stability of this compound.

Rosta et al. demonstrated that a protonated aspartate residue may protonate the O3′ atom of the scissile phosphodiester bond during RNAase H catalysis.[14] A structural comparison of FEN1 and RNaseH catalytic sites pinpoints Asp86 as potentially involved in the reaction of FEN1. Thus, we checked the stability of the Michaelis adduct in the presence of a

DOI: 10.1021/acscatal.5b00178
ACS Catal. 2015, 5, 3864−3875

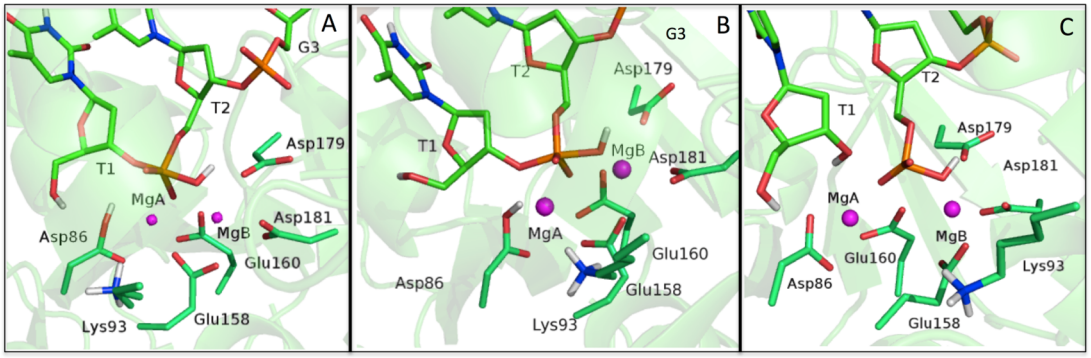


Figure 8. Graphical representation of the structure of reactant (**3a**, A), transition state (TS_{IIa}, B), and product (**4a**, C) for proton transfer from Asp86-OH to the leaving oxygen of the phosphorane. Active site residues are depicted as sticks and are colored by atom type; Mg^{2+} ions are shown as magenta vdW spheres. Nonpolar hydrogens are omitted for clarity.

protonated Asp86. However, classical MD simulation resulted in a catalytically incompetent state. This probably depends on the transient nature of the protonated Asp86, which may simply act as a proton shuttle. Then, this residue was considered as protonated only for step IIa (Figure 3a), as it has been demonstrated in previous studies that the protonation state of this aspartate does not affect the free energy barrier of the first step.[14]

The phosphorane intermediate was also stable in the presence of a protonated Asp86 for 60 ps of unbiased QM/MM MD. Hence, we investigated the formation of product promoted by a proton transfer from Asp86 to the leaving O3′@T1 (IIa, Figures 3 and 8). The selected CV (CV4; Figure S4 in the Supporting Information) was in this case the distance between Asp86@Oδ H and O3′@T1. US calculations pinpoint a low $\Delta F^{\ddagger}_{TSIIa}$ value of 7.4 ± 0.3 kcal/mol (Figure S11 in the Supporting Information). This second reaction step was highly exothermic ($\Delta F \approx -29$ kcal/mol). Both the stability of the phosphorane intermediate and the high exothermicity calculated for this reaction could be due to the level of theory considered. In fact, benchmark calculations showed that the AM1/d-PhoT Hamiltonian slightly overestimates reaction free energies in this type of reaction.[43]

Alternative Catalytic Pathways. Aware of the importance of the catalytic site protonation state for enzymatic reactions[21,25,92,98] and of the influence of the protein environment on pK_a values of ionizable residues,[99] we have also considered other catalytic mechanisms operative with less common protonation states of the active site residues.

Mutagenesis studies[10,100,101] highlighted that Lys93 and Arg100 are essential for the hFEN1 enzymatic activity and/or for substrate recognition; in particular Qiu at al.[100] reported that R100A and K93A mutations do not influence or only moderately influence hFEN1 affinity for the substrate, while both mutations completely abolish its enzymatic activity.

A possible involvement of a Lys residue in the endonuclease catalytic mechanism has been proposed on the basis of mutagenesis studies.[102] In addition, Garcia-Moreno and coworkers[103,104] recently reported that the pK_a values of Lys residues are strongly influenced by the environment and can be reduced up to 5.3, while Arg residues always retain their positive charge. Computational studies also suggested that Lys residues might act as acidic residues during catalysis.[14,105] Thus, we also explored an alternative reaction path in which Lys93 acts as a proton shuttle (Figure 4a)[99] in the phosphorane decomposition step (path D).

In this case US calculations using as CV the distance between Lys93@NzH and OP1@T1 (CV5; Figure S5 in the Supporting Information) report a $\Delta F^{\ddagger}_{TSIId}$ value of 16.8 ± 0.8 kcal/mol and a ΔF value between **2d** and **3d** of ∼1 kcal/mol (Figure S12 in the Supporting Information). After step IId a biprotonated phosphorane intermediate was formed (**3d** in Figure 4). Considering the peculiarity of this molecular species, already described by Elsasser and Fels in ribonuclease H,[39] we tested its stability by unbiased QM/MM MD simulation of 50 ps. This species was stable over the entire simulation time, with the distances between O3′@T1 and P@T1 and between O@WatR and P@T1 being 1.69 ± 0.04 and 1.57 ± 0.03 Å, respectively.

For the sake of completeness, we have also considered the possibility after WatR is preactivated as in step Ic, Lys93 may protonate the OP1@T1 (IIe), followed by the nucleophilic attack of the hydroxide, which therefore could occur on the resulting protonated phosphate group (IIIe) (path E; Figure 4b).

As reported above, the binding of a free OH^- from the bulk to the vicinity of the active site or to Mg^{2+} is the most likely process (step IcDPB/C or step IeDPB/C). Instead, Lys93 deprotonation attained a $\Delta F^{\ddagger}_{TSIIe}$ value of 16 ± 0.4 kcal/mol, which is very similar to the value calculated for step IId (Figure S13).

The $\Delta F^{\ddagger}_{TSIIIe}$ value for the nucleophilc attack of the hydroxide (IIIe) on a protonated phosphate was 7.4 ± 0.2 kcal/mol (Figure S14 in the Supporting Information). The last two reactions were exothermic by ∼−29 kcal/mol.

We then investigated the proton transfer from OP1-H to O3′@T1, which is common to both paths D and E (IIId, Figure 4) using as CV the distance between T1@OP1-H and O3′@T1 (CV6; Figure S6 in the Supporting Information). In this case the $\Delta F^{\ddagger}$ value is 22.3 ± 0.5 kcal/mol and the ΔF value between **3d** and **4d** is −13.3 ± 0.2 kcal/mol (Figure S15 in the Supporting Information). This proton transfer results in a breakage of the O3′@T1 to P@T1 bond, leading to the formation of the final product. However, this barrier is higher that that of the rate-determining step measured experimentally, suggesting that the mechanism in which a protonated Asp89 induces phosphorane decomposition may be more viable. This latter process is the final step of paths A−C.

SUMMARY AND CONCLUSIONS

In an alternative to QM(DFT)/MM studies which have been extensively used in the characterization of biological systems,[25,29,33] we applied AM1/d-PhoT based QM/MM

DOI: 10.1021/acscatal.5b00178
ACS Catal. 2015, 5, 3864−3875

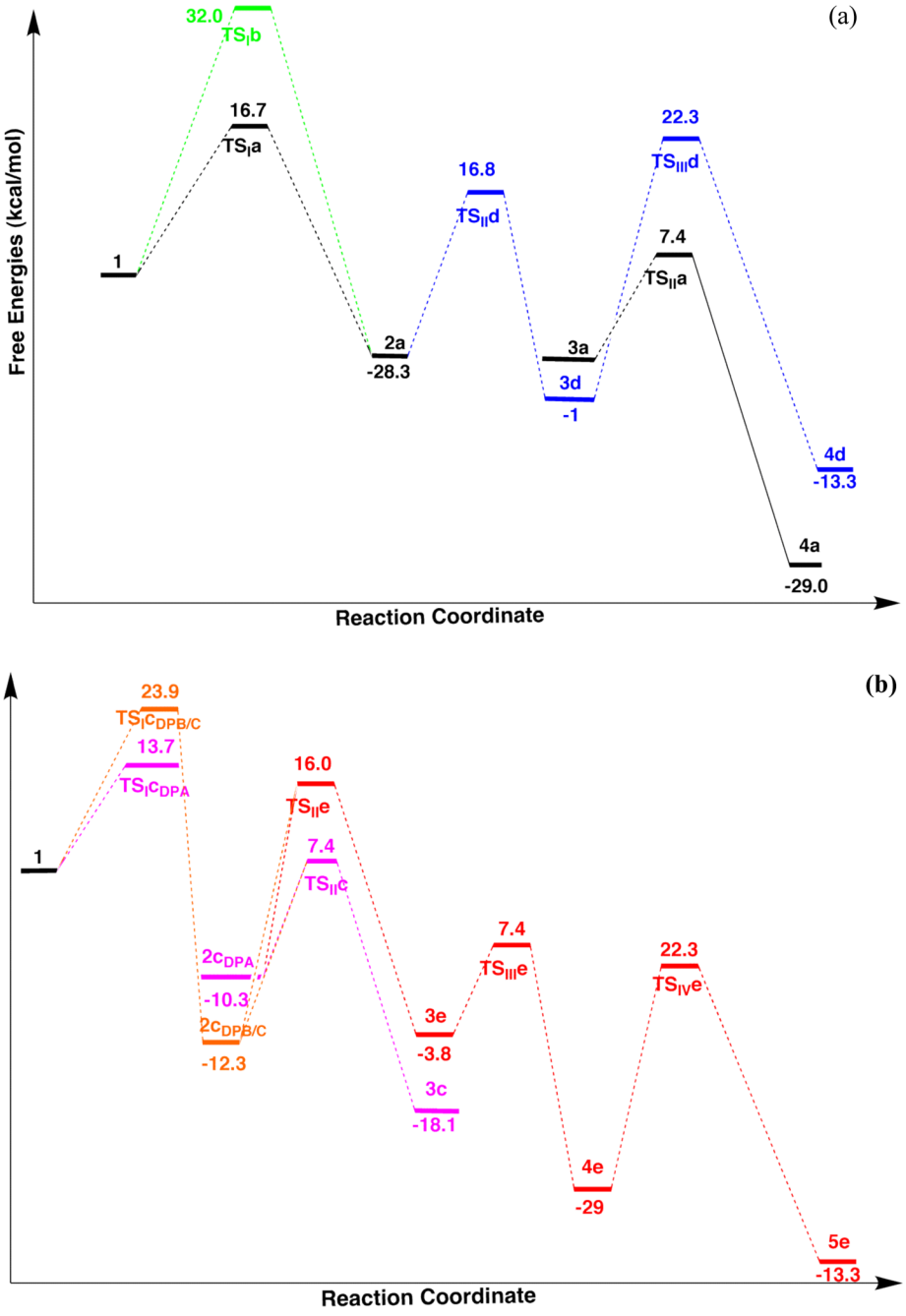


Figure 9. Sketch of the different reaction paths investigated here. ΔF and $\Delta F^{\ddagger}$ (kcal/mol) are reported with respect to the preceding stable intermediate. (a) Water performs the nucleophilic attack. Black lines refer to path A, green to path B, and blue to path D. In the black path **2a** and **3a** are not connected, as they differ in the protonation of Asp86. (b) A preactivated nucleophile performs the nucleophilic attack. Magenta and orange lines refer to deprotonation path A and path B/C, respectively. $TS_{IC}DPB/C$ refers to the free energy of water dissociation as estimated experimentally.[89] The red line refers to path E. Steps Ic and Ie are the same. For path C the final step is the same as that of paths A and B. Energy values refer to the intermediate preceding each state considered.

simulations to completely characterize the enzymatic mechanism of the hFEN1 reaction. This method has been proven to provide reaction free energies within few kilocalories per mole with respect to DFT calculations in biosystems promoting phosphate hydrolysis,[43,65,106] representing therefore a valid alternative to the more computationally expensive DFT/MM calculations when detailed experimental structural information about a reactive Michaelis complex is lacking, as in the present case. This method allows, in fact, a longer sampling of possible reactive adducts and transition states; most importantly, due to its moderate computational cost, this method allows the exploration of several possible alternative reaction paths within an affordable computational computational cost and time. Although a recent study debates the accuracy of AM1/d in determining either reaction free energy barriers or in determining the cooperativity/consequentiality of chemical events in biocatalysis,[41] there are successful examples of the application of this methodology to enzymatic systems.[43,65] To support our findings, here we observe the same order of events occurring in the reaction mechanism proposed for RNaseH,

DOI: 10.1021/acscatal.5b00178
ACS Catal. 2015, 5, 3864−3875

with free energy barriers in good agreement with experimental data and with previous DFT studies.[14]

Our theoretical description of the system in combination with enhanced sampling and free energy calculations gathered for the first time insights on a reactive FEN1/dsDNA complex and most notably an atomistic and energetic characterization of most viable catalytic paths (Figure 9), which is fully consistent mechanistically and energetically with the recent proposed mechanism of RNAase H enzymes, pinpointing common mechanistic features between RNA and DNA endonuclease enzymes.[14]

In particular, the free energy values obtained for the different paths suggest that the initial nucleophilic attack occurring simultaneously with WatR deprotonation is the likely the rate-determining step of the enzymatic cycle reaction with an estimated $\Delta F^{\ddagger}_{\mathrm{TSIa}}$ value of 16 ± 2 kcal/mol, in good agreement with the data reported by Grasby and co-workers ($k_{\mathrm{cat}} \approx 10\ \mathrm{s}^{-1}$ and $\Delta G^{\ddagger}$ = 16.1 kcal/mol),[11,107] as well as with other theoretical studies of RNase H.[13,14] In addition, the reaction mechanism (path C) involving the nucleophilic attack carried out by a preactivated nucleophile cannot be ruled out ($\Delta F^{\ddagger}_{\mathrm{TSIIc}}$ = 7.4 ± 0.2 kcal/mol). In fact, it is likely that a free hydroxide ion coming from the bulk can deprotonate the nucleophilic water or simply replace it in barrierless processes, with a large overall energy gain of $\Delta F_{\mathrm{Ic(DPB/C)}}$ = −12.3 (DFT), −17.1 kcal/mol (AM1-dPhoT), although in this case the process will be kinetically controlled by water autodissociation.

Our study discards instead the internal proton transfer from the attacking water molecule to the apical oxygen of the scissile phosphate (path B) ($\Delta F^{\ddagger}_{\mathrm{TSIb}}$ = 33.0 ± 2.6 kcal/mol), as well as the spontaneous decomposition of the phosphorane intermediate.[13]

A direct involvement of Lys93 having in the catalytic mechanism seems unlikely, as step IIId requires a high free energy barrier ($\Delta F^{\ddagger}_{\mathrm{TSIId}}$ = 22.3 ± 0.5 kcal/mol). Lys93, pointed out as being important for enzymatic activity, most likely stabilizes the negative charge forming at the TS, during formation of the phosphorane intermediate. However, considering the level of theory of accuracy of this theoretical description applied here,[41] we cannot completely rule out this possible alternative path. Joint experimental and higher level theoretical efforts should be performed in the future to further validate the most likely enzymatic path proposed here. Due to the importance of FENs in cell proliferation, for maintaining nucleic acid fidelity, and in turn for cancer risk, these enzymes are potential targets for therapeutic intervention.[108] FEN1 is also a powerful tool for serial invasive signal amplification reactions to form sensitive detection assays for DNA and RNA, which have applications for single nucleotide polymorphism detection.[10,108] As such, the atomistic level characterization of its enzymatic mechanism provided by this study has broad implications for life sciences, cellular and molecular biology, drug design and discovery purposes, as well as biotechnological applications.[10]

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acscatal.5b00178.

Details about CVs used during MTD simulations and the computational setup employed during the US calculations concerning all the reaction steps, free energy profiles from MTD and US simulations, and average QM/MM coordination distances calculated during unbiased QM/MM MD (PDF)

■ AUTHOR INFORMATION

Corresponding Authors

*J.S.: e-mail, jacopo.sgrignani@irb.usi.ch.

*A.M.: tel, +39 040 3787 529; fax, +39 040 3787 528; e-mail, alessandra.magistrato@sissa.it.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

Computational resources were supplied by CINECA (ISCRA grant HP10BNVJZO). We thank Prof. J. Grasby for the critical reading of the manuscript and Dr. D. Franco for suggestions on this study.

■ REFERENCES

(1) Finger, L. D.; Atack, J. M.; Tsutakawa, S.; Classen, S.; Tainer, J.; Grasby, J.; Shen, B. *Subcell. Biochem.* **2012**, *62*, 301−326.

(2) Balakrishnan, L.; Bambara, R. A. *Annu. Rev. Biochem.* **2013**, *82*, 119−138.

(3) Zheng, L.; Jia, J.; Finger, L. D.; Guo, Z.; Zer, C.; Shen, B. *Nucleic Acids Res.* **2011**, *39*, 781−794.

(4) Dorjsuren, D.; Kim, D.; Maloney, D. J.; Wilson, D. M.; Simeonov, A. *Nucleic Acids Res.* **2011**, *39*, e11 DOI: 10.1093/nar/gkq1082.

(5) Tumey, L. N.; Bom, D.; Huck, B.; Gleason, E.; Wang, J.; Silver, D.; Brunden, K.; Boozer, S.; Rundlett, S.; Sherf, B.; Murphy, S.; Dent, T.; Leventhal, C.; Bailey, A.; Harrington, J.; Bennani, Y. L. *Bioorg. Med. Chem. Lett.* **2005**, *15*, 277−281.

(6) Tumey, L. N.; Huck, B.; Gleason, E.; Wang, J.; Silver, D.; Brunden, K.; Boozer, S.; Rundlett, S.; Sherf, B.; Murphy, S.; Bailey, A.; Dent, T.; Leventhal, C.; Harrington, J.; Bennani, Y. L. *Bioorg. Med. Chem. Lett.* **2004**, *14*, 4915−4918.

(7) McWhirter, C.; Tonge, M.; Plant, H.; Hardern, I.; Nissink, W.; Durant, S. T. *J. Biomol. Screening* **2013**, *18*, 567−575.

(8) Williams, R. S.; Kunkel, T. A. *Cell* **2011**, *145*, 171−172.

(9) In addition to hFEN1, the superfamily comprises EXO1 (exonuclease 1), which acts on nicked or gapped DNA, GEN1 (gap endonuclease 1), which acts on holyday junctions and XPG (xeroderma pigmentosum complementation group G), which acts on bubble structures during nucleotide excision repair.

(10) Tsutakawa, S. E.; Classen, S.; Chapados, B. R.; Arvai, A. S.; Finger, L. D.; Guenther, G.; Tomlinson, C. G.; Thompson, P.; Sarker, A. H.; Shen, B.; Cooper, P. K.; Grasby, J. A.; Tainer, J. A. *Cell* **2011**, *145*, 198−211.

(11) Patel, N.; Atack, J. M.; Finger, L. D.; Exell, J. C.; Thompson, P.; Tsutakawa, S.; Tainer, J. A.; Williams, D. M.; Grasby, J. A. *Nucleic Acids Res.* **2012**, *40*, 4507−4519.

(12) Yang, W.; Lee, J. Y.; Nowotny, M. *Mol. Cell* **2006**, *22*, 5−13.

(13) De Vivo, M.; Dal Peraro, M.; Klein, M. L. *J. Am. Chem. Soc.* **2008**, *130*, 10955−10962.

(14) Rosta, E.; Nowotny, M.; Yang, W.; Hummer, G. *J. Am. Chem. Soc.* **2011**, *133*, 8934−8941.

(15) Rosta, E.; Yang, W.; Hummer, G. *J. Am. Chem. Soc.* **2014**, *136*, 3137−3144.

(16) Yang, W. *Q. Rev. Biophys.* **2011**, *44*, 1−93.

(17) Warshel, A.; Levitt, M. *J. Mol. Biol.* **1976**, *103*, 227−249.

(18) Ensing, B.; De Vivo, M.; Liu, Z.; Moore, P.; Klein, M. L. *Acc. Chem. Res.* **2006**, *39*, 73−81.

(19) Yang, W.; Cui, Q.; Min, D.; Li, H. In *Annual Reports in Computational Chemistry*; Ralph, A. W., Ed.; Elsevier: Amsterdam, 2010; Vol. *6*, pp 51−62.

(20) Grazioso, G.; Legnani, L.; Toma, L.; Ettari, R.; Micale, N.; Micheli, C. *J. Comput.-Aided Mol. Des.* **2012**, *26*, 1035−1043.

DOI: 10.1021/acscatal.5b00178
ACS Catal. 2015, 5, 3864−3875

(21) Sgrignani, J.; Magistrato, A.; Dal Peraro, M.; Vila, A. J.; Carloni, P.; Pierattelli, R. *J. Comput.-Aided Mol. Des.* **2012**, *26*, 425−435.
(22) Cavalli, A.; Carloni, P.; Recanatini, M. *Chem. Rev.* **2006**, *106*, 3497−3519.
(23) Dal Peraro, M.; Ruggerone, P.; Raugei, S.; Gervasio, F. L.; Carloni, P. *Curr. Opin. Struct. Biol.* **2007**, *17*, 149−156.
(24) Senn, H. M.; Thiel, W. *Angew. Chem., Int. Ed.* **2009**, *48*, 1198−1229.
(25) Simona, F.; Magistrato, A.; Dal Peraro, M.; Cavalli, A.; Vila, A. J.; Carloni, P. *J. Biol. Chem.* **2009**, *284*, 28164−28171.
(26) Banáš, P.; Jurecka, P.; Walter, N. G.; Sponer, J.; Otyepka, M. *Methods* **2009**, *49*, 202−216.
(27) Otyepka, M.; Banas, P.; Magistrato, A.; Carloni, P.; Damborsky, J. *Proteins: Struct., Funct., Genet.* **2008**, *70*, 707−717.
(28) Petersen, L.; Ardevol, A.; Rovira, C.; Reilly, P. J. *J. Am. Chem. Soc.* **2010**, *132*, 8291−8300.
(29) Vidossich, P.; Magistrato, A. *Biomolecules* **2014**, *4*, 616−645.
(30) Warshel, A. *Angew. Chem., Int. Ed.* **2014**, *53*, 10020−10031.
(31) Lonsdale, R.; Harvey, J. N.; Mulholland, A. J. *Chem. Soc. Rev.* **2012**, *41*, 3025−3038.
(32) Lodola, A.; Mulholland, A. J. *Methods Mol. Biol.* **2013**, *924*, 67−89.
(33) Sgrignani, J.; Magistrato, A. *Curr. Comput.-Aided Drug Des.* **2013**, *9*, 15−34.
(34) Ho, M.-H.; De Vivo, M.; Dal Peraro, M.; Klein, M. L. *J. Am. Chem. Soc.* **2010**, *132*, 13702−13712.
(35) Sgrignani, J.; Magistrato, A. *J. Phys. Chem. B* **2012**, *116*, 2259−2268.
(36) Wong, K.-Y.; Gu, H.; Zhang, S.; Piccirilli, J. A.; Harris, M. E.; York, D. M. *Angew. Chem., Int. Ed.* **2012**, *51*, 647−651.
(37) Wong, K.-Y.; Lee, T.-S.; York, D. M. *J. Chem. Theory Comput.* **2011**, *7*, 1−3.
(38) Palermo, G.; Stenta, M.; Cavalli, A.; Dal Peraro, M.; De Vivo, M. *J. Chem. Theory Comput.* **2013**, *9*, 857−862.
(39) Elsasser, B.; Fels, G. *Phys. Chem. Chem. Phys.* **2010**, *12*, 11081−11088.
(40) Palermo, G.; Cavalli, A.; Klein, M. L.; Alfonso-Prieto, M.; Dal Peraro, M.; De Vivo, M. *Acc. Chem. Res.* **2015**, *48*, 220.
(41) Mlýnský, V.; Banáš, P.; Šponer, J.; van der Kamp, M. W.; Mulholland, A. J.; Otyepka, M. *J. Chem. Theory Comput.* **2014**, *10*, 1608−1622.
(42) Imhof, P.; Noè, F.; Fischer, S.; Smith, J. C. *J. Chem. Theory Comput.* **2006**, *2*, 1050−1056.
(43) Nam, K.; Cui, Q.; Gao, J.; York, D. M. *J. Chem. Theory Comput.* **2007**, *3*, 486−504.
(44) Laio, A.; Gervasio, F. L. *Rep. Prog. Phys.* **2008**, *71*, 126601.
(45) Barducci, A.; Bonomi, M.; Parrinello, M. *Wiley Interdiscip. Rev. Comput.l Mol. Sci.* **2011**, *1*, 826−843.
(46) Laio, A.; Parrinello, M. *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99*, 12562−12566.
(47) Torrie, G. M.; Valleau, J. P. *J. Comput. Phys.* **1977**, *23*, 187−199.
(48) Kästner, J. *Wiley Interdiscip. Rev. Comput.l Mol. Sci.* **2011**, *1*, 932−942.
(49) Abrams, C.; Bussi, G. *Entropy* **2013**, *16*, 163−199.
(50) Spiegel, K.; Magistrato, A. *Org. Biomol. Chem.* **2006**, *4*, 2507−2517.
(51) Colombo, M. C.; Guidoni, L.; Laio, A.; Magistrato, A.; Maurer, P.; Piana, S.; Rohrig, U.; Spiegel, K.; Sulpizi, M.; VandeVondele, J.; Zumstein, M.; Röthlisberger, U. *Chimia* **2002**, *56*, 13.
(52) Hornak, V.; Abel, R.; Okur, A.; Strockbine, B.; Roitberg, A.; Simmerling, C. *Proteins: Struct., Funct., Genet.* **2006**, *65*, 712−725.
(53) Perez, A.; Marchan, I.; Svozil, D.; Sponer, J.; Cheatham, T. E.; Laughton, C. A.; Orozco, M. *Biophys. J.* **2007**, *92*, 3817−3829.
(54) 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.; Wang, J.; 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. *Amber12*; University of California, San Francisco, CA, 2012.
(55) Oelschlaeger, P.; Klahn, M.; Beard, W. A.; Wilson, S. H.; Warshel, A. *J. Mol. Biol.* **2007**, *366*, 687−701.
(56) Joseph, T. T.; Osman, R. *PLoS Comput. Biol.* **2012**, *8*, e1002693.
(57) Joung, I. S.; Cheatham, T. E. *J. Phys. Chem. B* **2008**, *112*, 9020−9041.
(58) Joung, I. S.; Cheatham, T. E. *J. Phys. Chem. B* **2009**, *113*, 13279−13290.
(59) Goetz, A. W.; Williamson, M. J.; Xu, D.; Poole, D.; Le Grand, S.; Walker, R. C. *J. Chem. Theory Comput.* **2012**, *8*, 1542−1555.
(60) Case, D. A.; Cheatham, T. E., III; Simmerling, C. L.; Wang, J.; Duke, R. E.; Zhang, W.; Merz, K. M.; Roberts, B.; Wang, B.; Hayik, S.; Roitberg, A.; Seabra, G.; Wong, K. F.; Paesani, F.; Vanicek, J.; Liu, J.; Wu, X.; Brozell, S. R.; Steinbrecher, T.; Cai, Q.; Ye, X.; Wang, J.; Hsieh, M.-J.; Cui, G.; Roe, D. R.; Mathews; Sagui, C.; Babin, V.; Luchko, T.; Gusarov, S.; Kovalenko, A.; Kollman, P. A. University of California, 2010.
(61) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; DiNola, A.; Haak, J. R. *J. Chem. Phys.* **1984**, *81*, 3684−3690.
(62) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. *J. Chem. Phys.* **1995**, *103*, 8577−8593.
(63) Marcos, E.; Anglada, J. M.; Crehuet, R. *Phys. Chem. Chem. Phys.* **2008**, *10*, 2442−2450.
(64) Wong, K.-Y.; Lee, T.-S.; York, D. M. *J. Chem. Theory Comput.* **2010**, *7*, 1−3.
(65) Radak, B. K.; Harris, M. E.; York, D. M. *J. Phys. Chem. B* **2012**, *117*, 94.
(66) Nam, K.; Gao, J.; York, D. M. *J. Am. Chem. Soc.* **2008**, *130*, 4680−4691.
(67) López-Canut, V.; Roca, M.; Bertrán, J.; Moliner, V.; Tuñón, I. *J. Am. Chem. Soc.* **2010**, *132*, 6955−6963.
(68) López-Canut, V.; Martí, S.; Bertrán, J.; Moliner, V.; Tuñón, I. *J. Phys. Chem. B* **2009**, *113*, 7816−7824.
(69) Arafet, K.; Ferrer, S.; Martí, S.; Moliner, V. *Biochemistry* **2014**, *53*, 3336−3346.
(70) van der Kamp, M. W.; Mulholland, A. J. *Biochemistry* **2013**, *52*, 2708−2728.
(71) Loncharich, R. J.; Brooks, B. R.; Pastor, R. W. *Biopolymers* **1992**, *32*, 523−535.
(72) Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. C. *J. Comput. Phys.* **1977**, *23*, 327−341.
(73) Walker, R. C.; Crowley, M. F.; Case, D. A. *J. Comput. Chem.* **2008**, *29*, 1019−1031.
(74) Cantu, D. C.; Ardèvol, A.; Rovira, C.; Reilly, P. J. *Chem. - Eur. J.* **2014**, *20*, 9045−9051.
(75) Sgrignani, J.; Grazioso, G.; De Amici, M.; Colombo, G. *Biochemistry* **2014**, *53*, 5174.
(76) Franco, D.; Vargiu, A. V.; Magistrato, A. *Inorg. Chem.* **2014**, *53*, 7999−8008.
(77) Vargiu, A. V.; Ruggerone, P.; Magistrato, A.; Carloni, P. *Nucleic Acids Res.* **2008**, *36*, 5910−5921.
(78) Raiteri, P.; Laio, A.; Gervasio, F. L.; Micheletti, C.; Parrinello, M. *J. Phys. Chem. B* **2005**, *110*, 3533−3539.
(79) Bisha, I.; Laio, A.; Magistrato, A.; Giorgetti, A.; Sgrignani, J. *J. Chem. Theory Comput.* **2013**, *9*, 1240−1246.
(80) Bonomi, M.; Branduardi, D.; Bussi, G.; Camilloni, C.; Provasi, D.; Raiteri, P.; Donadio, D.; Marinelli, F.; Pietrucci, F.; Broglia, R. A.; Parrinello, M. *Comput. Phys. Commun.* **2009**, *180*, 1961−1972.
(81) Grossfield, A. *WHAM: the weighted histogram analysis method, version 2.0.7*; http://membrane.urmc.rochester.edu/content/wham.
(82) Humphrey, W.; Dalke, A.; Schulten, K. *J. Mol. Graphics* **1996**, *14*, 33−38.
(83) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648−5652.
(84) Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1988**, *37*, 785.
(85) Stephens, P. J.; Devlin, F. J.; Chabalowski, C. F.; Frisch, M. J. *J. Phys. Chem.* **1994**, *98*, 11623−11627.

DOI: 10.1021/acscatal.5b00178
ACS Catal. 2015, 5, 3864−3875

(86) 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., Jr.; 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, Ö.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. *Gaussian09*; Gaussian, Inc., Wallingford, CT, 2009.
(87) Onufriev, A.; Bashford, D.; Case, D. A. *Proteins: Struct., Funct., Genet.* **2004**, *55*, 383−394.
(88) Miertuš, S.; Scrocco, E.; Tomasi, J. *Chem. Phys.* **1981**, *55*, 117−129.
(89) Sprik, M. *Chem. Phys.* **2000**, *258*, 139−150.
(90) Wang, J.; Deng, Y.; Roux, B. *Biophys. J.* **2006**, *91*, 2798−2814.
(91) Carvalho, A. T. P.; Fernandes, P. A.; Ramos, M. J. *J. Chem. Theory Comput.* **2011**, *7*, 1177−1188.
(92) Hong, R.; Magistrato, A.; Carloni, P. *J. Chem. Theory Comput.* **2008**, *4*, 1745−1756.
(93) Simona, F.; Magistrato, A.; Vera, D. M. A.; Garau, G.; Vila, A. J.; Carloni, P. *Proteins: Struct., Funct., Genet.* **2007**, *69*, 595−605.
(94) Magistrato, A.; DeGrado, W. F.; Laio, A.; Rothlisberger, U.; VandeVondele, J.; Klein, M. L. *J. Phys. Chem. B* **2003**, *107*, 4182−4188.
(95) Cassano, A. G.; Anderson, V. E.; Harris, M. E. *Biopolymers* **2004**, *73*, 110−129.
(96) Elsässer, B.; Fels, G.; Weare, J. H. *J. Am. Chem. Soc.* **2014**, *136*, 927−936.
(97) Elsässer, B.; Valiev, M.; Weare, J. H. *J. Am. Chem. Soc.* **2009**, *131*, 3869−3871.
(98) Sgrignani, J.; Magistrato, A. *J. Chem. Inf. Model.* **2012**, *52*, 1595−1606.
(99) Isom, D. G.; Castaneda, C. A.; Cannon, B. R.; GarcÌa-Moreno, B. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108*, 5260−5265.
(100) Qiu, J.; Liu, R.; Chapados, B. R.; Sherman, M.; Tainer, J. A.; Shen, B. *J. Biol. Chem.* **2004**, *279*, 24394−24402.
(101) Patel, N.; Exell, J. C.; Jardine, E.; Ombler, B.; Finger, L. D.; Ciani, B.; Grasby, J. A. *J. Biol. Chem.* **2013**, *288*, 34239−34248.
(102) Garforth, S. J.; Ceska, T. A.; Suck, D.; Sayers, J. R. *Proc. Natl. Acad. Sci. U. S. A.* **1999**, *96*, 38−43.
(103) Isom, D. G.; Castaneda, C. A.; Cannon, B. R.; Garcia-Moreno, B. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108*, 5260−5265.
(104) Harms, M. J.; Schlessman, J. L.; Sue, G. R.; Garcia-Moreno, E. B. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108*, 18954−18959.
(105) Gu, H.; Zhang, S.; Wong, K.-Y.; Radak, B. K.; Dissanayake, T.; Kellerman, D. L.; Dai, Q.; Miyagi, M.; Anderson, V. E.; York, D. M.; Piccirilli, J. A.; Harris, M. E. *Proc. Natl. Acad. Sci. U. S. A.* **2013**, *110*, 13002−13007.
(106) Lee, T.-S.; Silva LÛpez, C.; Giambasu, G. M.; Martick, M.; Scott, W. G.; York, D. M. *J. Am. Chem. Soc.* **2008**, *130*, 3053−3064.
(107) Syson, K.; Tomlinson, C.; Chapados, B. R.; Sayers, J. R.; Tainer, J. A.; Williams, N. H.; Grasby, J. A. *J. Biol. Chem.* **2008**, *283*, 28741−28746.
(108) McWhirter, C.; Tonge, M.; Plant, H.; Hardern, I.; Nissink, W.; Durant, S. T. *J. Biomol. Screening* **2013**, *18*, 567−575.

DOI: 10.1021/acscatal.5b00178
ACS Catal. 2015, 5, 3864−3875