JOURNAL OF CHEMICAL INFORMATION AND MODELING

Article

pubs.acs.org/jcim

Role of Water in the Puzzling Mechanism of the Final Aromatization Step Promoted by the Human Aromatase Enzyme. Insights from QM/MM MD Simulations

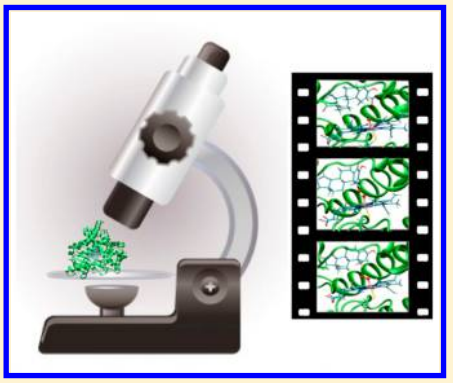
Jacopo Sgrignani,[†,§] Marcella Iannuzzi,[‡] and Alessandra Magistrato*[,†]

[†]CNR-IOM-Democritos National Simulation Center c/o International School for Advanced Studies (SISSA/ISAS), via Bonomea 265, Trieste, Trieste, Italy

[‡]Physical Chemistry Institute, University of Zurich, Winterthurerstrasse 190, CH-8057 Zurich, Zurich, Switzerland

Supporting Information

ABSTRACT: The enzyme human aromatase (HA) catalyzes the conversion of androgens to estrogens via two hydroxylation reactions and a final unique aromatization step. Despite the great interest of HA as a drug target against breast cancer detailed structural and spectroscopic information on this enzyme became available only in the past few years. As such, the enigmatic mechanism of the final aromatization step is still a matter of debate. Here, we investigated the final step of the HA enzymatic cycle via hybrid quantum-classical (QM/MM) metadynamics and blue-moon ensemble simulations. Our results show that the rate-determining step of the aromatization process is the nucleophilic attack of the distal oxygen of a peroxo-ferric species on the formyl carbon of the enol-19-oxo-androstenedione, which occurs with a free energy barrier ($\Delta F^{\#}$) of ~16.7 ± 1.9 kcal/mol, in good agreement with experimental data. This reaction is followed by a water mediated 1β-hydrogen abstraction ($\Delta F^{\#}$ = 7.9 ± 0.8 kcal/mol) and by the formation of a hydroxo-ferric moiety. This latter may be finally protonated by a hydrogen delivery channel involving Asp309 and Thr310, both residues pointed out as crucial for HA activity. In the absence of the catalytic water in the active site the substrate does not assume a position suitable to undergo the nucleophilic attack. Our data not only reveal a novel possible mechanism for the aromatization process consistent with some of the spectroscopic and kinetic data available in the literature, complementing current knowledge on the mechanism of this enzyme, but also point out a remarkable influence of the level of theory used on the calculated free energy barriers. The structural information obtained in this study may be used for the rational structure-based drug design of HA inhibitors to be employed in breast cancer therapy.



1. INTRODUCTION

Cytochrome P450 19A1 (CYP19), commonly known as human aromatase (HA), is an enzyme, located on the endoplasmic reticulum membrane of estrogen-producing cells, converting androgens (androstenedione (ASD) and testosterone) to estrogens (estrone (EST) and estradiol).[1] HA, characterized by a prosthetic heme group and an androgen-specific cleft,[2] promotes the aromatization of steroids via a three-steps reaction with each step requiring molecular oxygen and a specific NADPH cytochrome reductase.[1,3] The final step of the estrogens biosynthesis consists of the aromatization of the "a" ring of the androgen precursor (such as ASD in Figure 1A, 1B and S1 of the Supporting Information (SI)), which is converted to a phenol ring.[1,4] From the physiopathological point of view HA is involved in 75% of the breast cancer genesis and progression. Therefore, HA is of paramount importance as a drug target for breast cancer therapy in postmenopausal women.[5,6] Several HA inhibitors, some of which are in clinical trials, were shown to block estrogens production.[4,6−9] However, the discovery of novel inhibitors is mandatory as current therapies are plagued by considerable side effects (heart problems, osteoporosis, joint pain) and resistance problems.[4,6−9] What is more, over the years, HA has also been related with the onset of other neoplastic pathologies such as endometrium, ovary, prostate, lung, and colon cancers.[4]

The first crystal structure of HA was solved in 2009,[2] providing the basis for an atomistic understanding of the reaction mechanism,[10] substrate binding,[11,12] possible novel inhibition routes,[13] and HA interaction with the membrane.[14−16]

To date, despite HA has been studied for almost 40 years, the aromatization step of the enzymatic cycle is still a subject of active debate.[3,10] Recent experimental EPR studies suggested that the most abundant species accumulating prior to the enzymatic final step is a peroxo-ferric intermediate in a low spin state[17] (Figure 1C). Akhtar et al. proposed that the C2,C3 double bond of the "a" steroid ring may be formed before the release of the formate molecule and the formation of compound I (CpdI) (Figure 1C and S2).[3,18]

Received: April 30, 2015
Published: September 18, 2015

DOI: 10.1021/acs.jcim.5b00249
J. Chem. Inf. Model. 2015, 55, 2218−2226

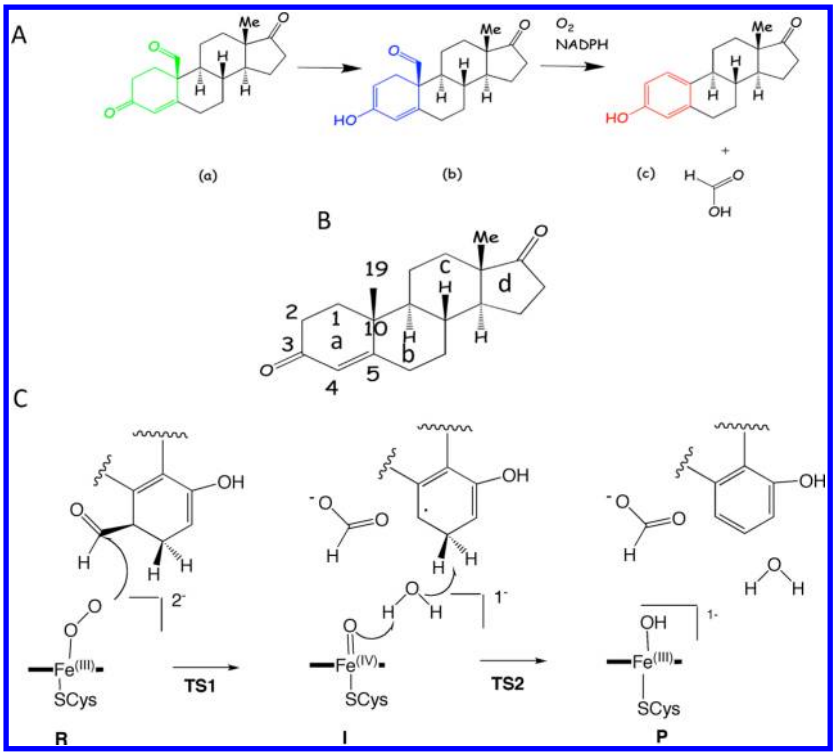


Figure 1. (A) The final step of the aromatization reaction: (a) 19-oxo-ASD, (b) 3-enol-19-oxo-ASD, (c) EST. The steroid "a" ring of the three compounds is highlighted in green, blue, and red, in figures (a), (b), (c), respectively; (B) ASD numbering and ring classification, and (C) general reaction scheme for the aromatization reaction, as suggested by this study.

In contrast to this proposal[2] a computational study, carried out with classical molecular dynamics (MD) and static quantum-classical (QM/MM) calculations, proposed a reaction mechanism based on the initial barrier-less formation of a peroxo-hemiacetal species, followed by the rate-determining breakage of the O−O bond ($\Delta G^{\#} \sim 15$ kcal/mol) (Figure S3).[10] Moreover, a different theoretical study reported a reaction mechanism involving the initial enolization of the substrate and suggested that this occurs via a Asp309-water-ASD proton delivery path, leading to the formation of compound 0 (Cpd0).[19,20] Despite this path requiring a low free energy cost ($\Delta G^{\#} \sim 5$ kcal/mol),[19] this finding is not consistent with Gantt et al.[17] who suggested that an deprotonated peroxo-ferric intermediate as the dominant species generated by cryoreduction. Moreover, the rest of the path has a rate-determining free energy barrier of 3−6 kcal/mol, which is remarkably lower than that measured experimentally for HA.[20] All these findings have been recently questioned by novel experimental data[21−24] and by a recent study based on static QM/MM calculations, suggesting an aromatization reaction path involving CpdI and a 19,19-gemdiol-3-enol-ASD.[21,24]

Here we investigated the final step of the aromatization of ASD to estrone (EST) via QM/MM metadynamics (MTD) simulations. MTD is a computational technique aimed at accelerating rare events in MD simulations and mapping the corresponding free energy surface (FES).[25−27] As further refinement of the reaction path we employed the 'blue-moon' ensemble simulations.[25,28,29] Both methodologies have been already successfully applied in a QM/MM MD framework to unravel the reaction mechanism of several metal-containing enzymes.[26,29−32] Our results pinpoint the initial nucleophilic attack as the rate-determining step of the reaction, followed by a water mediated H1β transfer to CpdI. The final resting state of the enzyme is most likely reached by a proton transfer from bulk water to the hydroxy-ferric species mediated by Asp310 and Thr310, both highly conserved residues in CYP450s and pointed out as crucial for the enzymatic activity.[19]

2. COMPUTATIONAL DETAILS

2.1. Model Systems. Starting from the X-ray structure of HA (pdb code: 3EQM) we have reconstructed molecular oxygen and a 3-enol-19-oxo-ASD in the active site. We have considered three possible models of the initial Michaelis complex (see Figure S4): Model A in which Asp309 is protonated according to our previous findings[13,15] and a recent experimental evidence[33] and a water (Wat-R) is present in the active site; Model B in which Asp309 is deprotonated and Wat-R is in the active site; Model C in which Asp309 is protonated and no water is in the active site. The overall systems solvated with explicit water molecules comprised ∼64055 atoms.

2.2. Classical MD Simulations. Force field (FF) based MD simulations were done using parm99[34] FF for the protein, the TIP3P[35] water model for the solvent, and the Åqvist parameters for counterions.[36] See the SI for additional details on heme moiety parameters. Each HA/3-enol-19-oxo-ASD complex (Model A-C) was simulated for 15 ns in an NPT ensemble (T = 298.5 K and P = 1 atm), with a Nosè-Hoover thermostat[37,38] and Parrinello-Rahaman barostat[39] as implemented in Gromacs 4.5.3[40] using a time step of 2 fs.

2.3. QM/MM MD Simulations. The QM/MM Born−Oppenheimer (BO) MD and MTD simulations were run with the CP2K package.[41−43] See the SI for details on the QM part description, which is based on Density Functional Theory (DFT). The MM part of the system was treated using the same FF parameters employed for the classical MD simulations.

DOI: 10.1021/acs.jcim.5b00249
J. Chem. Inf. Model. 2015, 55, 2218−2226

Only Models A and C were considered in the QM/MM MD simulations (vide infra). In these simulations the heme moiety, the "a" and "b" (Figure 1) rings of the 3-enol-19-oxo-ASD, the oxygen molecule (OXX), the iron ion, the side chain of Cys437, and a Wat-R were included in the QM part (63 atoms). To relax the HA structure at the beginning of the simulation the system was annealed, slowly increasing the temperature up to 298.5 K. The simulation was run for 2.5 ps with a time step of 0.50 fs in NVT ensemble regulating the temperature with a Nosè-Hoover thermostat[37,38] to equilibrate the system.

From Electron Paramagnetic Resonance (EPR) experiments[17] and previous theoretical calculations,[44,45] the peroxo-ferric species appears to be in a doublet spin state. Therefore, we verified that this was the ground state of our system, using different XC functionals (Table S1), and we considered only this state in the QM/MM MD runs. Then, we verified a posteriori which was the spin state of the intermediate and product obtained from the MTD simulations. Moreover, we have tested the $\langle S^2 \rangle$ for all the species investigated in this work using different XC functionals (Table S2), and we have checked the localization of the spin density at different stages of the reaction (see the SI).

2.4. QM/MM Metadynamics Simulations. Then, the last configuration obtained from the QM/MM MD trajectory was used as a starting point for MTD calculations. A first explorative MTD run was done considering two collective variables (CVs).[46] The first, CV1, is expressed as the difference of the distance between the distal oxygen atom of the peroxo group (OXX2) and C19 and the distance between the two peroxo-oxygens (OXX1 and OXX2). The second one, CV2, is expressed as the difference of the distance between the oxygen of Wat-R (OW) and the hydrogen atom to be abstracted from C1 (1β-H) and the distance between one of the hydrogens of Wat-R (H1) and the proximal oxygen of the peroxo group (OXX1). A graphical scheme about the CVs is reported in Figure S5. In this simulation, Gaussian hills with a height of 3.0 kcal/mol and a width of 0.250 Bohr for CV1 and 0.325 for CV2, respectively, were added every 10 fs of QM/MM MD simulation. The explored FES was restricted using harmonic position restraints placed at 4.0 Bohr and at 9.0 Bohr for CV1 and CV2, respectively. Then, a second type of MTD simulations was done, using only one CV at a time, in order to focus on the rate-determining step of the reaction mechanism. In this case also the MTD parameters were modified. Namely, Gaussian potentials with a height of 1.1 kcal/mol and a width of 0.10 Bohr were added every 10 fs, and the R basin was filled in ~3 ps, adding ~315 hills. Each simulation was done three times, reassigning differently the initial velocities, in order to improve the statistics and estimate the error.[47,48]

2.5. Transition State Localization and Dependence of Free Energy Barriers on Exchange Correlation Functional. We verified that the structure corresponding to the transition state (TS) of the initial nucleophilic attack (TS1) was a real TS by performing commitment probability analysis.[49] For several putative TS1 structures selected from the analysis of the FES obtained by the QM/MM MTD simulations, we started several (5−6) independent short QM/MM MD simulations to identify the most representative structure of the TS. The most promising candidate structure was validated by performing 10 additional QM/MM MD simulations, verifying that the TS structure was evolving with equal probability toward reactant and product.

The ground state and the selected TS structures were then employed to obtain small models of the catalytic site (these models included only ASD, the heme porphyrin ring, Cys437, and the iron atom). These model systems were optimized, and the zero point energy corrections were calculated using BLYP[50,51] exchange correlation functional (XC) only. Single point calculations on the BLYP optimized models were done using different XC functionals aiming at estimating the influence of the DFT level of theory on the $\Delta G^{\#}$ (TS1) of the rate-determining step. In particular, BLYP[50,51] PBE,[52−54] M062X,[55] OLYP,[51,56−58] and B3LYP[51,59,60] and in the case of BLYP and M062X also the dispersion correction proposed by Grimme[61] et al. were used. These calculations were run using the Gaussian09 program[62] and considering a 6-31G** basis set and LANL2DZ for the iron atom. All the calculations were done in vacuum after having verified the influence of an implicit solvent model like PCM was of the order of 3.7 kcal/mol.[63]

2.6. "Blue-Moon" Ensemble Simulations. The reactive path obtained from MTD simulations was also investigated by performing "blue-moon" ensemble simulations. To further check the reliability of our findings and be aware of many successful QM/MM MD studies available in the literature on heme proteins[64−67] we have employed a different computational approach. Namely, we used the QM(Car-Parrinello)/MM MD scheme developed by Laio et al.,[68] as part of the CPMD code.[69] Here, we treated at the QM level the same region considered above using, however, the BP XC functional,[50,70] following an established computational protocol.[64−67] See the SI for details of the calculations of the QM part. The MM part of the complex is treated using the same set up of the previous calculations.

The procedure to refine the free energy profile (FEP) consisted of selecting 17 snapshots along the MTD reaction path and using them as initial configurations for annealing runs with constrained CV. The FES of the reaction is obtained by thermodynamic integration.[28] Each point was simulated for 2 ps, taking the average of the resulting constraint force at each step on the last half part of the trajectory. The error was calculated by propagation analysis.[29] In these simulations we have used CV1, while for the second step we used the difference between H1β and O-WatR and OWatR-H1, as CV2 was not available in this code. This CV will be referred to as CV3.

3. RESULTS AND DISCUSSIONS

3.1. Classical and QM/MM MD Equilibration of the Reactant. Based on the mechanism proposed by Akhtar et al., we started our calculations considering the substrate as the 3-enol-19-oxo-ASD ((b) in Figure 1).[3,71]

Before starting QM/MM MD calculations, a 15 ns force field based MD simulation was run to equilibrate the structure of the 3-enol-19-oxo-ASD/HA adduct. In the adduct in which no water molecule is present in the active site (Model C) and Asp309 was considered in its neutral form, according to our previous studies[10,15] and recent pK_a measurements,[33] the formyl group of the substrate and the oxygen moiety of the peroxo moiety intermediate do not assume a conformation favorable for the nucleophilic attack (Figure S4c). Thus, since the importance of water content in the active site has been elucidated for different metallic enzymes,[29] even performing a very similar catalytic function,[72,73] we placed a water molecule

DOI: 10.1021/acs.jcim.5b00249
J. Chem. Inf. Model. 2015, 55, 2218−2226

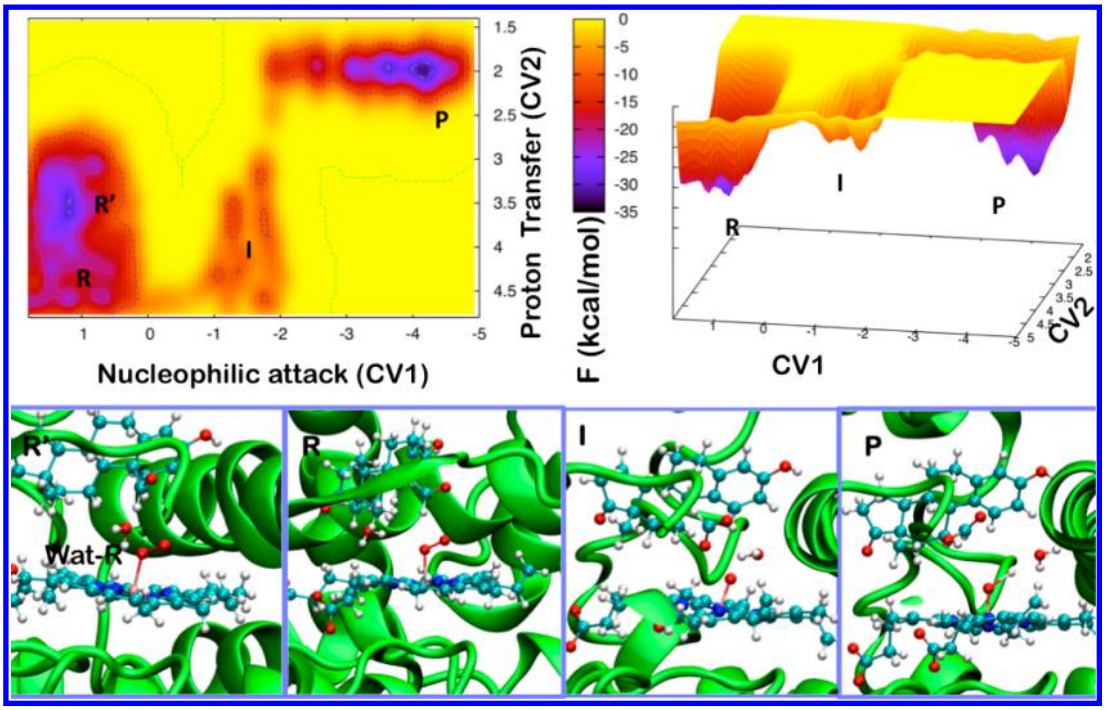


Figure 2. Top: top and side view of the Free Energy Surface (FES) of the preliminary MTD simulations with CV1 and CV2. Bottom: structures corresponding to the nonproductive reactant (R′), the reactant (R), the intermediate (I), and the product (P) states.

in the vicinity of the peroxide (Model A). During the classical MD this water (Wat-R) remains in the active site, forming a H-bond bridge between the peroxide and the Hβ of the substrate (Figure S4a) as well as H-bonding to Thr310. In order to check if the presence of this water was consistent also with the ionized form of Asp309, we considered an additional model of the active site (Model B). Also in this case the water remained in the same position in 15 ns classical MD (Figure S4b). This latter model was considered only to check if the protonation state of Asp309 was affecting substrate mobility of the active site residues and of the substrate (Figure S6).

Only Model A was considered to fully explore the FES in simulations, since it is consistent with a reactive configuration of the substrate. Moreover, Model B should not strongly affect the reaction, as it differs in the protonation of Asp309, which was not directly involved in the reaction step investigated here. Model C instead was not reactive in preliminary MTD simulations (vide infra).

Before starting the QM/MM MTD simulations Model A was equilibrated using QM/MM BO MD for 2.0 ps. In the equilibrated state the iron coordination sphere displayed the following distances: 1.89 ± 0.05 Å for Fe-OXX1, 2.51 ± 0.05 Å for Fe−S@Cys437, and 2.00 ± 0.01 Å for the average distance between Fe and the nitrogen atoms of the heme group. In addiction, the peroxo group had an OXX1-OXX2 bond length of 1.38 ± 0.03 Å. These distances are in agreement with the results reported by Wang et al. obtained with B3LYP functional.[44] Notably, Wat-R H-bonds with OXX1 (H@(Wat-R)-OXX1 = 1.30 ± 0.03 Å), thus stabilizing the peroxo-ferric species (Figure S7).[74]

3.2. Metadynamics Simulations. After the initial equilibration, MTD was employed to investigate the aromatization reaction (Figure 1C). In the first exploratory run two CVs are used (CV1 and CV2, Figure S5)). We observed that the reaction occurs in a "step-wise" manner. As shown in Figure 2, in order to have the nucleophilic attack, the water molecule first is slightly displaced from the peroxo-ferric moiety. Only afterward, the reactant state (R) evolves via a first transition state (TS_1) into the intermediate (I), corresponding to CpdI (Figure 3). The displacement of the water is consistent with what was observed by Vidossich et al. in horseradish peroxidase, in which a minor populated state is the reactive conformer promoting catalysis.[72] The second step to the final product hydroxy-ferric species (P) occurs via a second transition state (TS_2). This path also shows that a direct protonation of the peroxide is not viable, in contrast to what was proposed by Kramos et al.[19]

We also considered whether the reaction might occur without the assistance of Wat-R (considering Model C). However, in this case we did not observe the proton transfer from ASD to OXX1 with a barrier compatible with the kinetic data reported in the literature.[75,76] This finding is consistent with early QM studies of Hackett et al., reporting a very high free energy barrier for direct proton abstraction of H1β by Cpd0.[77] Thus, these results pinpoint a crucial role of Wat-R in the aromatization reaction. Since from the preliminary MTD simulation it was clear that the reaction proceeds in two steps, each driven only by a single CV, it is possible to separate the reaction in two independent MTD runs each performed with CV only, to refine better the topology of the FES, using more conservative MTD parameters and exploring more carefully the configurational space.

In the simulations performed with CV1, once the nucleophilic attack occurred, a formate molecule was formed. In the typical simulations time-scale of QM/MM MD it is unfeasible to observe the backward reaction from I to R using the same CV. In fact, the newly formed formate molecule departs from the ASD, and a recrossing event becomes very unlikely in a few ps. Thus, to accurately estimate the free energy barrier ($\Delta F^{\#}$) we performed three independent MTD simulations, stopping the calculation once the nucleophilic attack has occurred. The resulting average $\Delta F^{\#}$ was 23.5 ± 1.2

DOI: 10.1021/acs.jcim.5b00249
J. Chem. Inf. Model. 2015, 55, 2218−2226

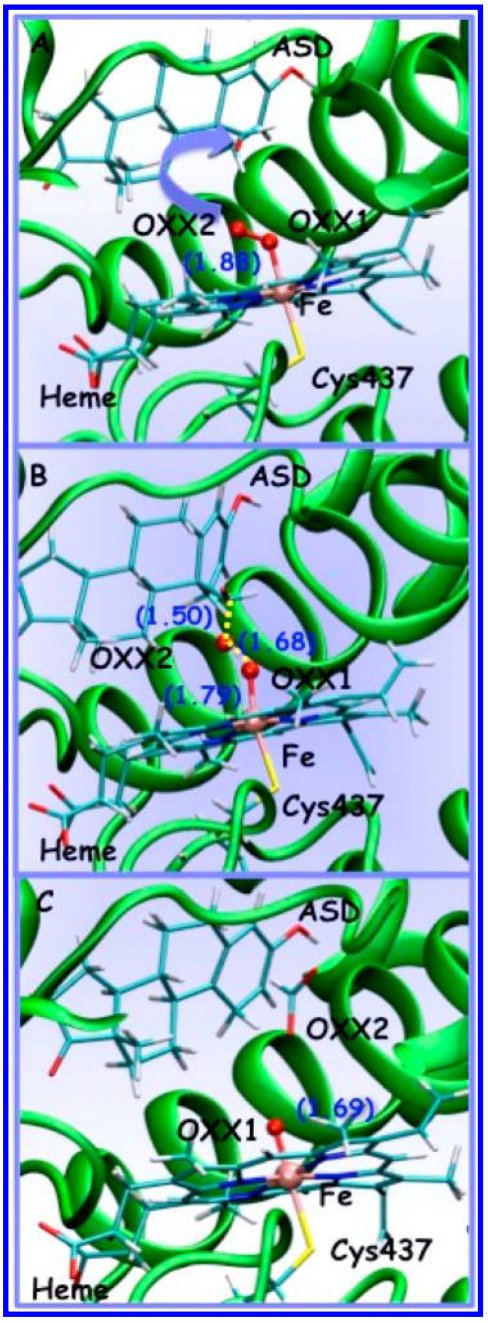


Figure 3. Structures, geometric distances (Å) for the R (A), TS_1 (B), and I (C). The protein backbone is depicted in green ribbons, the heme moiety and ASD are depicted in licorice, the iron and the oxygen molecule are represented as a ball and sticks.

kcal/mol, which is ∼6 kcal/mol higher than the experimental value (17.1 kcal/mol).[75] However, considering (1) the short time scale accessible to QM(BO)/MM MD simulations and (2) the inaccuracies deriving from the use of a gradient corrected XC functional such as BLYP,[50,51,78] the estimated $\Delta F^{\#}$ can be considered qualitatively a good approximation. Moreover, the use of only one CV does not allow to control the position of Wat-R, which H-bonding with the peroxide moiety may increase the reaction free energy barrier. Since the second step is not rate limiting, we have not performed additional MTD simulations with CV2 only.

3.2. Rate-Limiting Transition State Validation. In order to confirm that the structure corresponding to TS_1 (characterized by a CV1 value of −0.40) was really a transition state, we also did a committor analysis (as described in the methods section), verifying that the TS structure was evolving with equal probability toward R and P in different unbiased QM/MM MD simulations. We remark that this only a posteriori validation of the TS structure.

We also calculated $\Delta G^{\#}$ on small model systems of the active site minimized at the BLYP level to evaluate the influence of the XC functional on the free energy barrier. The BLYP,[50,51] PBE,[52−54] M062X,[55] OLYP,[51,56−58] and B3LYP[51,59,60] functionals were considered as well as in the case of BLYP and M062X also the dispersion correction proposed by Grimme.[61] From these calculations, the $\Delta G^{\#}$ corresponding to TS_1 is ∼7−23 kcal/mol (Table 1), suggesting that the transition state is highly sensitive to the choice of the XC functional, as shown by Kramos et al.[19,20] However, the BLYP functional is in good agreement with the experimental value of 17.1 kcal/mol.[75] For the BLYP functional we also estimated that the zero point energy corrections were of only 0.6 kcal/mol, while in the influence of the implicit solvent was of 3.7 kcal/mol (Table 1).

Table 1. $\Delta G^{\#}_{R\text{-}TS1}$ (kcal/mol) Evaluated with Different XC Functionals

XC functional	$\Delta G^{\#}_{R\text{-}TS1}$
BLYP	15.8
BLYP-D	12.0
BLYP-PCM	12.1
PBE	7.2
M062X	22.3
M062X-D	21.6
B3LYP	26.2
OLYP	22.8
experimental value[75]	17.1

3.3. Blue-Moon Ensemble Calculations. As a further check to validate our findings we used a different computational setup and different computational approach (thermodynamic integration) to estimate the FEP, employing a computational protocol already validated for other iron enzymes.[31,65−67,72,73,79,80] These blue-moon ensemble simulations show that the first reaction step occurs with a $\Delta F^{\#}$ of 16.7 ± 1.9 kcal/mol, leading to the formation of I, which is exothermic by 35.4 ± 2.3 kcal/mol (Figure 4). From I the reaction proceeds with a second step, which, crossing a barrier of a $\Delta F^{\#}$ of 7.9 ± 0.8 kcal/mol, leads to P, confirming that the initial nucleophilic attack is the rate-limiting step.

3.4. Ending the Enzymatic Cycle. Since the enzyme resting-state bears a water molecule coordinated to the iron, we looked for a H-bond network around the hydroxo-ferric moiety of P, which could regenerate the resting state of the enzyme via a proton transfer. After 2.5 ps long QM/MM MD simulation of P we observed a H-bond network involving Wat-R and a second water (Wat-A), which connected Asp309 to Thr310 and, finally, to OXX1. Therefore, a possible path to complete the enzymatic cycle may be a proton transfer from Asp309 to OXX1 through this network (Figure 5). This is consistent with the protonation path observed in other CYP450s and is fully in line with the functional role ascribed by mutational studies to Asp309 and Thr310 for the enzymatic activity.[19]

Consistently with mutagenesis studies the observed proton network, mediated by few water molecules, extends toward Arg192 and Glu483. These residues form a salt bridge at the

DOI: 10.1021/acs.jcim.5b00249
J. Chem. Inf. Model. 2015, 55, 2218−2226

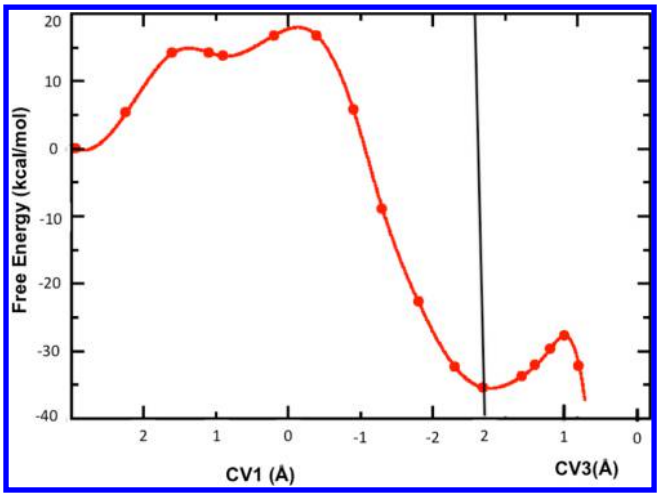


Figure 4. FEP (kcal/mol) obtained from blue moon ensemble simulations. Black vertical line divides the first step (the nucleophilic attack) and the second step (the deprotonation of C1β of ASD). The two steps have been done using CV1 and CV3 as reaction coordinate, respectively.

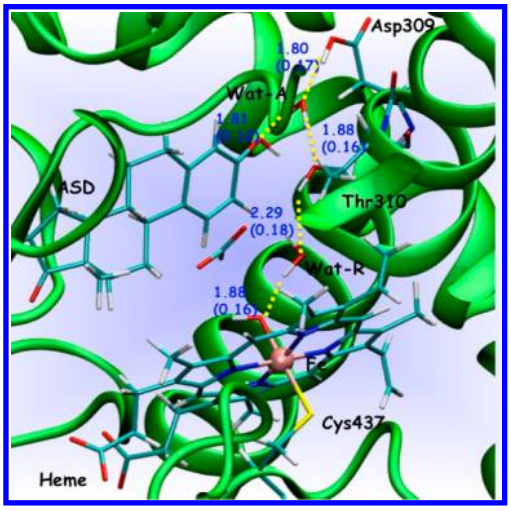


Figure 5. Representative structure (selected by cluster analysis) of P as resulting from a 2.5 ps long QM/MM MD simulation. The H-bond network, involving Wat-A and Wat-R, is depicted in red dashed lines. The protein backbone is depicted in green ribbons, the heme moiety and ASD are depicted in licorice, the iron is represented as a pink ball. Average distances (Å) and standard deviation (within parentheses) are also reported.

entrance of the catalytic site, lining the access channel, and they act as gatekeeper playing a regulatory role for the entrance of the solute/solvent molecules. Their mutations most probably affect the proton network observed in our study, determining reduction of the enzymatic activity as observed in mutagenesis experiments.

4. SUMMARY AND CONCLUSIONS

Here we have investigated the final aromatization step of the estrogen biosynthesis promoted by the enzyme HA, using a MTD simulations and thermodynamics integration. The results have been validated also by employing different XC functionals, basis set and different types of QM/MM MD approaches based both on Born−Oppenheimer and Car−Parrinello MD for the QM region.

First, MTD simulations allowed dynamically exploring the FES to identify a novel reaction path. Certainly, this computational approach provides more flexibility with respect to traditional QM/MM studies of reaction mechanisms based only on structural optimizations.[25] The MTD trajectories were subsequently employed as a guess of the reaction path for "blue-moon" ensemble calculations and to better refine the reaction free energy barrier.

The main outcome of this work is a novel aromatization reaction path in which (i) the peroxo-ferric species is the active species responsible for the deformylation reaction in HA; after formation of the 3-enol-19-oxo-ASD, (ii) the reaction occurs in two steps with the nucleophilic attack being the rate-determining one; (iii) the second step of the aromatization process is mediated by a water molecule (Wat-R).

This study confirms the well-known role of water molecules in enzymatic processes.[81] Moreover, our study highlights the presence of a proton network involving Thr310, Asp309, Arg192, and Glu483, mediated by a few water molecules, which may stabilize the initial peroxo-ferric species and restore the resting-state of the enzyme at the end of the enzymatic cycle. All the mentioned residues have been observed to regulate catalysis, consistently with our results. Remarkably, in the absence of Wat-R the substrate does not assume a conformation suitable to undergo the nucleophilic attack. Thus, our results provide important new structural insights for future structure-based drug discovery efforts against breast cancer, as new inhibitors may be designed as transition state analogues on the basis of our findings.

Consistently with other studies we also observe a large dependence of the free energy barriers on the choice of the XC functionals.[19,20] Furthermore, recent experimental findings proposed a new aromatization mechanism based on CpdI,[21] which requires further computational investigation.[20,24]

Thus, the different computational and experimental studies available to date suggest that in HA a plethora of possible competitive paths may coexist, which are difficult to discriminate both experimentally and computationally. Our and other studies demonstrate that this may be very challenging at a computational level as Fe-dependent enzymes possess an intricate electronic structure, so that the typical errors of DFT based calculations may be of the same order of magnitude of the free energy differences among the different possible paths.[20]

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jcim.5b00249.

Details about previous computational studies of human aromatase (HA) reaction mechanism, Computational details about MM and QM/MM MD and MTD simulations, Figures S1−S8, and Tables S1 and S2 (PDF)

■ AUTHOR INFORMATION

Corresponding Author

*Phone: 39 040 3787 529. Fax. 39 040 3787 528. E-mail: alessandra.magistrato@sissa.it.

Present Address

§Institute of Research in Biomedicine, Via Vincenzo Vela 6, CH-6500, Bellinzona, Switzerland.

DOI: 10.1021/acs.jcim.5b00249
J. Chem. Inf. Model. 2015, 55, 2218−2226

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We acknowledge the CINECA award N. HP10BO043N, 2010 for the availability of high performance computing resources and support. J.S. thanks Dr. S. Piccinin and Dr. D. Branduardi for the useful discussion about metadynamics. A.M. thanks Dr. P. Vidossich for discussions and suggestions.

■ REFERENCES

(1) Waterman, M. R. Structural Biology Anticancer Drug Target Pictured. *Nature* **2009**, *457*, 159−160.
(2) Ghosh, D.; Griswold, J.; Erman, M.; Pangborn, W. Structural Basis for Androgen Specificity and Oestrogen Synthesis in Human Aromatase. *Nature* **2009**, *457*, 219−223.
(3) Akhtar, M.; Wright, J. N.; Lee-Robichaud, P. A Review of Mechanistic Studies on Aromatase (Cyp19) and 17α-Hydroxylase-17,20-Lyase (Cyp17). *J. Steroid Biochem. Mol. Biol.* **2011**, *125*, 2.
(4) Dai, Y.; Wang, Q.; Zhang, X.; Jia, S.; Zheng, H.; Feng, D.; Yu, P. Molecular Docking and Qsar Study on Steroidal Compounds as Aromatase Inhibitors. *Eur. J. Med. Chem.* **2010**, *45*, 5612−5620.
(5) Czajka-Oraniec, I.; Simpson, E. R. Aromatase Research and Its Clinical Significance. *Endokrynol. Polym.* **2010**, *61*, 126−134.
(6) Santen, R. J.; Brodie, H.; Simpson, E. R.; Siiteri, P. K.; Brodie, A. History of Aromatase: Saga of an Important Biological Mediator and Therapeutic Target. *Endocr. Rev.* **2009**, *30*, 343−375.
(7) Liu, J.; Flockhart, P. J.; Lu, D.; Lv, W.; Lu, W. J.; Han, X.; Cushman, M.; Flockhart, D. A. Inhibition of Cytochrome P450 Enzymes by the E- and Z-Isomers of Norendoxifen. *Drug Metab. Dispos.* **2013**, *41*, 1715−20.
(8) Kesmodel, S. B.; Sabnis, G. J.; Chumsri, S.; Brodie, A. M. Combined Cancer Therapy: Strategies to Overcome Acquired Aromatase Inhibitor Resistance. *Curr. Pharm. Des.* **2014**, *20*, 6575−83.
(9) Favia, A. D.; Nicolotti, O.; Stefanachi, A.; Leonetti, F.; Carotti, A. Computational Methods for the Design of Potent Aromatase Inhibitors. *Expert Opin. Drug Discovery* **2013**, *8*, 395−409.
(10) Sen, K.; Hackett, J. C. Coupled Electron Transfer and Proton Hopping in the Final Step of Cyp19-Catalyzed Androgen Aromatization. *Biochemistry* **2012**, *51*, 3039−3049.
(11) Narayana, B. L.; Pran Kishore, D.; Balakumar, C.; Rao, K. V.; Kaur, R.; Rao, A. R.; Murthy, J. N.; Ravikumar, M. Molecular Modeling Evaluation of Non-Steroidal Aromatase Inhibitors. *Chem. Biol. Drug Des.* **2012**, *79*, 674−682.
(12) Dai, Y.; Wang, Q.; Zhang, X.; Jia, S.; Zheng, H.; Feng, D.; Yu, P. Molecular Docking and Qsar Study on Steroidal Compounds as Aromatase Inhibitors. *Eur. J. Med. Chem.* **2010**, *45*, 5612−5620.
(13) Sgrignani, J.; Bon, M.; Colombo, G.; Magistrato, A. Computational Approaches Elucidate the Allosteric Mechanism of Human Aromatase Inhibition: A Novel Possible Route to Small-Molecule Regulation of Cyp450s Activities? *J. Chem. Inf. Model.* **2014**, *54*, 2856−68.
(14) Jiang, W.; Ghosh, D. Motion and Flexibility in Human Cytochrome P450 Aromatase. *PLoS One* **2012**, *7*, e32565.
(15) Sgrignani, J.; Magistrato, A. Influence of the Membrane Lipophilic Environment on the Structure and on the Substrate Access/Egress Routes of the Human Aromatase Enzyme. A Computational Study. *J. Chem. Inf. Model.* **2012**, *52*, 1595−1606.
(16) Sgrignani, J.; Colombo, G.; Cavalli, A.; Magistrato, A. *Mini Rev. Med. Chem.* **2015**, In press.
(17) Gantt, S. L.; Denisov, I. G.; Grinkova, Y. V.; Sligar, S. G. The Critical Iron-Oxygen Intermediate in Human Aromatase. *Biochem. Biophys. Res. Commun.* **2009**, *387*, 169−173.
(18) Furthermore, the crystal structure of HA in complex with ASD suggests that the formation of the C2,C3 double bond, coming along with the formation of the hydroxyl group at C3, may be catalyzed by a protonated Asp309, which H-bonds to the 3-keto oxygen of ASD, while the abstraction of the 2β-H may be catalyzed by the carbonyl of Ala306, with participation of Thr310 and, possibly, of a catalytic water.
(19) Kramos, B.; Olah, J. Enolization as an Alternative Proton Delivery Pathway in Human Aromatase (P450 19a1). *J. Phys. Chem. B* **2014**, *118*, 390−405.
(20) Krámos, B.; Oláh, J. The Mechanism of Human Aromatase (Cyp 19a1) Revisited: DFT and QM/MM Calculations Support a Compound I-Mediated Pathway for the Aromatization Process. *Struct. Chem.* **2015**, *26*, 279−300.
(21) Yoshimoto, F. K.; Guengerich, F. P. Mechanism of the Third Oxidative Step in the Conversion of Androgens to Estrogens by Cytochrome P450 19a1 Steroid Aromatase. *J. Am. Chem. Soc.* **2014**, *136*, 15016−25.
(22) Khatri, Y.; Luthra, A.; Duggal, R.; Sligar, S. G. Kinetic Solvent Isotope Effect in Steady-State Turnover by Cyp19a1 Suggests Involvement of Compound 1 for Both Hydroxylation and Aromatization Steps. *FEBS Lett.* **2014**, *588*, 3117−22.
(23) Mak, P. J.; Luthra, A.; Sligar, S. G.; Kincaid, J. R. Resonance Raman Spectroscopy of the Oxygenated Intermediates of Human Cyp19a1 Implicates a Compound I Intermediate in the Final Lyase Step. *J. Am. Chem. Soc.* **2014**, *136*, 4825−8.
(24) Xu, K.; Wang, Y.; Hirao, H. Estrogen Formation Via H-Abstraction from the O−H Bond of Gem-Diol by Compound I in the Reaction of Cyp19a1: Mechanistic Scenario Derived from Multiscale Qm/Mm Calculations. *ACS Catal.* **2015**, *5*, 4175−4179.
(25) Ensing, B.; De Vivo, M.; Liu, Z.; Moore, P.; Klein, M. L. Metadynamics as a Tool for Exploring Free Energy Landscapes of Chemical Reactions. *Acc. Chem. Res.* **2006**, *39*, 73−81.
(26) Barducci, A.; Bonomi, M.; Parrinello, M. Metadynamics. *WIREs Comput. Mol. Sci.* **2011**, *1*, 826−843.
(27) Laio, A.; Parrinello, M. Escaping Free-Energy Minima. *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99*, 12562−12566.
(28) Carter, E. A.; Ciccotti, G.; Hynes, J. T.; Kapral, R. Constrained Reaction Coordinate Dynamics for the Simulation of Rare Events. *Chem. Phys. Lett.* **1989**, *156*, 472−477.
(29) Simona, F.; Magistrato, A.; Dal Peraro, M.; Cavalli, A.; Vila, A. J.; Carloni, P. Common Mechanistic Features among Metallo-Beta-Lactamases: A Computational Study of Aeromonas Hydrophila Cpha Enzyme. *J. Biol. Chem.* **2009**, *284*, 28164−71.
(30) Petersen, L.; Ardevol, A.; Rovira, C.; Reilly, P. J. Molecular Mechanism of the Glycosylation Step Catalyzed by Golgi α-Mannosidase: A QM/MM Metadynamics Investigation. *J. Am. Chem. Soc.* **2010**, *132*, 8291−8300.
(31) Alfonso-Prieto, M.; Biarnes, X.; Vidossich, P.; Rovira, C. The Molecular Mechanism of the Catalase Reaction. *J. Am. Chem. Soc.* **2009**, *131*, 11751−11761.
(32) Sgrignani, J.; Magistrato, A. First-Principles Modeling of Biological Systems and Structure-Based Drug-Design. *Curr. Comput.-Aided Drug Des.* **2013**, *9*, 15−34.
(33) Di Nardo, G.; Breitner, M.; Bandino, A.; Ghosh, D.; Jennings, G. K.; Hackett, J. C.; Gilardi, G. Evidence for an Elevated Aspartate Pka in the Active Site of Human Aromatase. *J. Biol. Chem.* **2015**, *290*, 1186−1196.
(34) Duan, Y.; Wu, C.; Chowdhury, S.; Lee, M. C.; Xiong, G.; Zhang, W.; Yang, R.; Cieplak, P.; Luo, R.; Lee, T.; Caldwell, J.; Wang, J.; Kollman, P. A Point-Charge Force Field for Molecular Mechanics Simulations of Proteins Based on Condensed-Phase Quantum Mechanical Calculations. *J. Comput. Chem.* **2003**, *24*, 1999−2012.
(35) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, L. M. Comparison of Simple Potential Functions for Simulating Liquid Water. *J. Chem. Phys.* **1983**, *79*, 926−935.
(36) Aaqvist, J. Ion-Water Interaction Potentials Derived from Free Energy Perturbation Simulations. *J. Phys. Chem.* **1990**, *94*, 8021−8024.
(37) Nosè, S. A Unified Formulation of the Constant Temperature Molecular-Dynamics Method. *J. Chem. Phys.* **1984**, *81*, 511−519.
(38) Hoover, W. G. Canonical Dynamics: Equilibrium Phase-Space Distributions. *Phys. Rev. A: At., Mol., Opt. Phys.* **1985**, *31*, 1695.

DOI: 10.1021/acs.jcim.5b00249
J. Chem. Inf. Model. 2015, 55, 2218−2226

(39) Parrinello, M.; Rahman, A. Polymorphic Transitions in Single Crystals: A New Molecular Dynamics Method. *J. Appl. Phys.* **1981**, *52*, 7182−7190.
(40) Hess, B.; Kutzner, C.; van der Spoel, D.; Lindahl, E. Gromacs 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation. *J. Chem. Theory Comput.* **2008**, *4*, 435−447.
(41) VandeVondele, J.; Krack, M.; Mohamed, F.; Parrinello, M.; Chassaing, T.; Hutter, J. r. Quickstep: Fast and Accurate Density Functional Calculations Using a Mixed Gaussian and Plane Waves Approach. *Comput. Phys. Commun.* **2005**, *167*, 103−128.
(42) Laino, T.; Mohamed, F.; Laio, A.; Parrinello, M. An Efficient Real Space Multigrid QM/MM Electrostatic Coupling. *J. Chem. Theory Comput.* **2005**, *1*, 1176−1184.
(43) Laino, T.; Mohamed, F.; Laio, A.; Parrinello, M. An Efficient Linear-Scaling Electrostatic Coupling for Treating Periodic Boundary Conditions in QM/MM Simulations. *J. Chem. Theory Comput.* **2006**, *2*, 1370−1378.
(44) Wang, D.; Thiel, W. The Oxyheme Complexes of P450cam: A QM/MM Study. *J. Mol. Struct.: THEOCHEM* **2009**, *898*, 90−96.
(45) Hackett, J. C.; Brueggemeier, R. W.; Hadad, C. M. The Final Catalytic Step of Cytochrome P450 Aromatase: A Density Functional Theory Study. *J. Am. Chem. Soc.* **2005**, *127*, 5224−5237.
(46) The two CVs used were selected after a few attempts to observe the reaction with different CVs.
(47) Bisha, I.; Laio, A.; Magistrato, A.; Giorgetti, A.; Sgrignani, J. A Candidate Ion-Retaining State in the Inward-Facing Conformation of Sodium/Galactose Symporter: Clues from Atomistic Simulations. *J. Chem. Theory Comput.* **2013**, *9*, 1240−1246.
(48) Sgrignani, J.; Grazioso, G.; De Amici, M.; Colombo, G. Inactivation of Tem-1 by Avibactam (Nxl-104): Insights from Quantum Mechanics/Molecular Mechanics Metadynamics Simulations. *Biochemistry* **2014**, *53*, 5174−85.
(49) Bolhuis, P. G.; Chandler, D.; Dellago, C.; Geissler, P. L. Transition Path Sampling: Throwing Ropes over Rough Mountain Passes, in the Dark. *Annu. Rev. Phys. Chem.* **2002**, *53*, 291−318.
(50) Becke, A. D. Density-Functional Exchange-Energy Approximation with Correct Asymptotic Behavior. *Phys. Rev. A: At., Mol., Opt. Phys.* **1988**, *38*, 3098.
(51) Lee, C.; Yang, W.; Parr, R. G. Development of the Colle-Salvetti Correlation-Energy Formula into a Functional of the Electron Density. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1988**, *37*, 785.
(52) Perdew, J. P.; Burke, K.; Wang, Y. Generalized Gradient Approximation for the Exchange-Correlation Hole of a Many-Electron System. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1996**, *54*, 16533−16539.
(53) Perdew, J. P.; Burke, K.; Ernzerhof, M. Generalized Gradient Approximation Made Simple. *Phys. Rev. Lett.* **1996**, *77*, 3865.
(54) Perdew, J. P.; Burke, K.; Ernzerhof, M. Generalized Gradient Approximation Made Simple. *Phys. Rev. Lett.* **1997**, *78*, 1396−1396.
(55) Zhao, Y.; Truhlar, D. The M06 Suite of Density Functionals for Main Group Thermochemistry, Thermochemical Kinetics, Noncovalent Interactions, Excited States, and Transition Elements: Two New Functionals and Systematic Testing of Four M06-Class Functionals and 12 Other Functionals. *Theor. Chem. Acc.* **2008**, *120*, 215−241.
(56) Hoe, W.-M.; Cohen, A. J.; Handy, N. C. Assessment of a New Local Exchange Functional Optx. *Chem. Phys. Lett.* **2001**, *341*, 319−328.
(57) Molawi, K.; Cohen, A. J.; Handy, N. C. Left−Right and Dynamic Correlation. *Int. J. Quantum Chem.* **2002**, *89*, 86−93.
(58) Miehlich, B.; Savin, A.; Stoll, H.; Preuss, H. Results Obtained with the Correlation Energy Density Functionals of Becke and Lee, Yang and Parr. *Chem. Phys. Lett.* **1989**, *157*, 200−206.
(59) Becke, A. D. Density-Functional Thermochemistry. Iii. The Role of Exact Exchange. *J. Chem. Phys.* **1993**, *98*, 5648−5652.
(60) Stephens, P. J.; Devlin, F. J.; Chabalowski, C. F.; Frisch, M. J. Ab Initio Calculation of Vibrational Absorption and Circular Dichroism Spectra Using Density Functional Force Fields. *J. Phys. Chem.* **1994**, *98*, 11623−11627.
(61) Grimme, S.; Ehrlich, S.; Goerigk, L. Effect of the Damping Function in Dispersion Corrected Density Functional Theory. *J. Comput. Chem.* **2011**, *32*, 1456−1465.
(62) 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.
(63) Miertuš, S.; Scrocco, E.; Tomasi, J. Electrostatic Interaction of a Solute with a Continuum. A Direct Utilizaion of Ab Initio Molecular Potentials for the Prevision of Solvent Effects. *Chem. Phys.* **1981**, *55*, 117−129.
(64) Alfonso-Prieto, M.; Biarnes, X.; Vidossich, P.; Rovira, C. The Molecular Mechanism of the Catalase Reaction. *J. Am. Chem. Soc.* **2009**, *131*, 11751−61.
(65) Alfonso-Prieto, M.; Vidossich, P.; Rodriguez-Fortea, A.; Carpena, X.; Fita, I.; Loewen, P. C.; Rovira, C. Electronic State of the Molecular Oxygen Released by Catalase. *J. Phys. Chem. A* **2008**, *112*, 12842−8.
(66) Vidossich, P.; Alfonso-Prieto, M.; Carpena, X.; Loewen, P. C.; Fita, I.; Rovira, C. Versatility of the Electronic Structure of Compound I in Catalase-Peroxidases. *J. Am. Chem. Soc.* **2007**, *129*, 13436−46.
(67) Vidossich, P.; Magistrato, A. Qm/Mm Molecular Dynamics Studies of Metal Binding Proteins. *Biomolecules* **2014**, *4*, 616−45.
(68) Laio, A.; VandeVondele, J.; Rothlisberger, U. A Hamiltonian Electrostatic Coupling Scheme for Hybrid Car-Parrinello Molecular Dynamics Simulations. *J. Chem. Phys.* **2002**, *116*, 6941−6947.
(69) http://www.cpmd.org/ (accessed Sept 1, 2015).
(70) Perdew, J. P. Density-Functional Approximation for the Correlation Energy of the Inhomogeneous Electron Gas. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1986**, *33*, 8822−8824.
(71) Wright, J. N.; Akhtar, M. Studies on Estrogen Biosynthesis Using Radioactive and Stable Isotopes. *Steroids* **1990**, *55*, 142−151.
(72) Vidossich, P.; Florin, G.; Alfonso-Prieto, M.; Derat, E.; Shaik, S.; Rovira, C. On the Role of Water in Peroxidase Catalysis: A Theoretical Investigation of Hrp Compound I Formation. *J. Phys. Chem. B* **2010**, *114*, 5161−5169.
(73) Derat, E.; Shaik, S.; Rovira, C.; Vidossich, P.; Alfonso-Prito, M. The Effect of a Water Molecule on the Mechanism of Formation of Compound 0 in Horseradish Peroxidase. *J. Am. Chem. Soc.* **2007**, *129*, 6346−6347.
(74) This is different with respect to other P450 cytochromes, where the protonation of the distal peroxo oxygen (OXX2) is propedeutic to the formation of Cpd0 and then of the very reactive CpdI.
(75) Sohl, C. D.; Guengerich, F. P. Kinetic Analysis of the Three-Step Steroid Aromatase Reaction of Human Cytochrome P450 19a1. *J. Biol. Chem.* **2010**, *285*, 17734−17743.
(76) This is probably due to the distance between 1β-H@Asd and Oxx1 (average distance 2.92 ± 0.27 Å), which is longer than that (∼2 Å) measured for the 1β-hydrogen abstraction from camphor performed by Cyp10. (See: Shaik, S.; Cohen, S.; Wang, Y.; Chen, H.; Kumar, D.; Thiel, W. P450 Enzymes: Their Structure, Reactivity, and Selectivity-Modeled by Qm/Mm Calculations. *Chem. Rev.* **2010**, *110*, 949−941017)..
(77) Hackett, J. C.; Brueggemeier, R. W.; Hadad, C. M. The Final Catalytic Step of Cytochrome P450 Aromatase: A Density Functional Theory Study. *J. Am. Chem. Soc.* **2005**, *127*, 5224−5237.

DOI: 10.1021/acs.jcim.5b00249
J. Chem. Inf. Model. 2015, 55, 2218−2226

(78) Shaik, S.; Cohen, S.; Wang, Y.; Chen, H.; Kumar, D.; Thiel, W. P450 Enzymes: Their Structure, Reactivity, and Selectivity-Modeled by QM/MM Calculations. *Chem. Rev.* **2010**, *110*, 949−941017.
(79) Loewen, P. C.; Carpena, X.; Vidossich, P.; Fita, I.; Rovira, C. An Ionizable Active-Site Tryptophan Imparts Catalase Activity to a Peroxidase Core. *J. Am. Chem. Soc.* **2014**, *136*, 7249−52.
(80) Vidossich, P.; Loewen, P. C.; Carpena, X.; Fiorin, G.; Fita, I.; Rovira, C. Binding of the Antitubercular Pro-Drug Isoniazid in the Heme Access Channel of Catalase-Peroxidase (Katg). A Combined Structural and Metadynamics Investigation. *J. Phys. Chem. B* **2014**, *118*, 2924−31.
(81) Simona, F.; Magistrato, A.; Dal Peraro, M.; Cavalli, A.; Vila, A. J.; Carloni, P. Common Mechanistic Features among Metallo-Beta-Lactamases: A Computational Study of Aeromonas Hydrophila Cpha Enzyme. *J. Biol. Chem.* **2009**, *284*, 28164−28171.

DOI: 10.1021/acs.jcim.5b00249
J. Chem. Inf. Model. 2015, 55, 2218−2226