

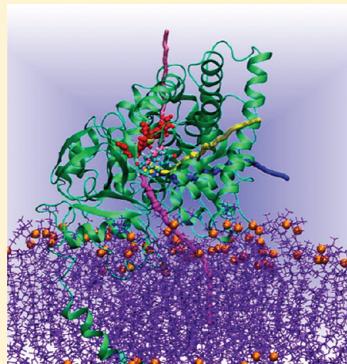
# Influence of the Membrane Lipophilic Environment on the Structure and on the Substrate Access/Egress Routes of the Human Aromatase Enzyme. A Computational Study

Jacopo Sgrignani<sup>†</sup> and Alessandra Magistrato\*,<sup>†</sup>

<sup>†</sup>CNR-IOM-Democritos National Simulation Center c/o International Studies for Advanced Studies (SISSA/ISAS), via Bonomea 265, 34165 Trieste (TS), Italy

## Supporting Information

**ABSTRACT:** Human aromatase (HA), an enzyme located on the membrane of the endoplasmatic reticulum, is of crucial biological importance in the biosynthesis of estrogens. High levels of estrogens are related with important pathologies, conferring to HA a key role as a pharmacological target. In this study we provide, for the first time, an atomistic model of HA embedded on a membrane model to understand the influence of the membrane lipophilic environment on the structural and dynamical properties of HA and on the access/egress pathways of the substrate (androstenedione, ASD) and of the oxygen molecule (involved in the enzymatic process) into/from the HA active site. To this end we used several computational techniques such as force field-based molecular dynamics (MD) simulations, Random Expulsion MD, Steered MD, and Implicit Ligand Sampling. Our results show that the membrane anchoring does not markedly affect the structural properties and the flexibility of the protein, but they clearly point out that the membrane has a marked effect on the access/egress routes of the reactants, stabilizing the formation of different channels for both ASD and O<sub>2</sub> with respect to those observed in pure water solution. Due to the importance of HA in medicine and since access/egress channels may influence its substrate selectivity, a detailed understanding of the role of the membrane in shaping these channels may be of valuable help in drug design.



## 1. INTRODUCTION

Cytochromes P450 are a superfamily of heme monooxygenases of enormous interest for their potential applications in the i) synthesis of new drugs, ii) targeted cancer therapy, iii) biosensor design, and iv) bioremediation. Despite their importance, practical applications of cytochromes P450 are limited since they exhibit a low stability and a low turnover rate.<sup>1</sup> Therefore, there is a great interest in the understanding of the function of these enzymes to design new P450 biocatalysts for potential practical applications in biotechnology, medicine, and bioremediation.

Here, we focused on Human Aromatase (HA) enzyme, a member of the CYP19 family of the P450 cytochromes, which catalyzes the conversion of androgens to estrogens, through three oxidative reactions that require molecular oxygen and electrons (Figure S1 of the Supporting Information). These latter are supplied by a specific reductase, which binds on the side of the protein proximal to the heme moiety.<sup>2</sup>

From the pathological points of view, high levels of estrogens have been related with an abnormal cellular proliferation that is at the origin of several diseases such as breast cancer, endometriosis, and gynecomastia.<sup>3,4</sup> The importance of this enzyme is remarkable as breast cancer is, nowadays, the most diffused cancer type. This makes HA a key target for the therapy of all the estrogen receptor positive cancers, which represent 75% of the total cases.<sup>5</sup>

Moreover, a reduced aromatase expression or activity is the etiological reason of a pathological syndrome leading to an incomplete virilization.<sup>6</sup>

Finally, a neuroprotective role has been recently attributed to HA.<sup>7</sup> In fact, its genetic or pharmacological inhibition may result in neurodegeneration. Thus, HA may also represent an important pharmacological target for therapies aimed at preventing neurodegenerative disorders.<sup>8</sup> Interestingly, an exogenous modulation of the HA enzymatic activity has been ascribed to some highly diffused water pollutants (atrazine, endosulfan, etc).<sup>9</sup> Thus, this enzyme plays a key role from both the pharmacological and toxicological point of view.<sup>10</sup>

HA is topologically located on the surface of the endoplasmic reticulum (ER) membrane, and it is anchored to it via a single N-terminal trans-membrane helix, with the catalytic domain directed toward the cytosol.<sup>11,12</sup>

Due to the key importance of HA as drug target, a large number of experimental<sup>2,12,13</sup> and modeling studies was done in order to unveil its dimensional structure and the interactions with its physiological substrates or drugs.<sup>2,14–23</sup> However, due to the lack of a crystal structure, the theoretical studies were carried out using homology modeling techniques and were, therefore, not fully reliable.<sup>2,14–23</sup>

**Received:** March 19, 2012

**Published:** May 23, 2012



Only recently, Ghosh et al.<sup>12</sup> have successfully crystallized HA in complex with its natural substrate androstanedione (ASD) (pdb code 3EQM). From the structural point of view, HA shows a fold similar to other P450 proteins with its secondary structure composed by 12 helix, identified by a letter code from A to L, and by 10 sheets, labeled from 1 to 10.<sup>12</sup> The catalytic site, deeply buried inside the protein, is characterized by a heme group and an iron ion, which are essential for catalysis.<sup>24,25</sup> A possible orientation of HA with respect to the ER membrane was suggested based on the X-ray structure and later confirmed by other studies.<sup>2,12</sup>

In spite of the atomistic details provided by the X-ray structure, the helix connecting the protein to the ER membrane was not solved, and it is still unclear as to the influence of the membrane on the structural and dynamical properties of HA as well as on the location of the access/egress channels of the molecular species involved in the catalytic process (i.e., the molecular oxygen (O<sub>2</sub>), the ASD, the reaction intermediates, and the product). In fact, the crystallization conditions adopted for membrane-bound CYP450s may induce artifacts in the X-ray structure, preventing the identification of possible channels at the membrane/protein interface.<sup>26</sup> Furthermore, the interaction of CYP450s with the cellular membrane has been demonstrated to affect the binding of liposoluble molecules to their buried active site.<sup>27–30</sup>

Most of the modeling studies carried out in the last ten years on membrane-bound CYPs did not consider the membrane explicitly.<sup>26,31–38</sup> This important aspect has started to be taken into account only recently on the 2C9 and the CYP3A4 cytochromes.<sup>27–29</sup>

Elucidating the influence of CYPs/membrane interactions at the molecular level is of paramount importance to identify the most likely access/egress pathways of the reactants.<sup>26,27,39</sup> This aspect is crucial both from the biochemical and the medicinal chemistry point of view. Indeed, in proteins bearing a deeply buried catalytic site, the substrate selectivity is influenced not only by the structural features of the active site but also by the access/release routes of the ligands.<sup>40,41</sup>

Here, we have addressed for the first time this issue in HA by performing classical molecular dynamics (MD) simulations based on force field of a HA/membrane (*HAm*) model.<sup>42–44</sup> The structural and dynamical properties of *HAm* have been compared with that of HA/water (*HAw*) model to unveil similarities and differences. Both models have been employed to identify the most likely access/release routes of ASD and O<sub>2</sub>, using a variety of computational approaches and performing also a critical comparison of the results obtained by using different methodologies. Finally, our results have been compared with previous computational studies of the access/release routes to/from the catalytic sites of other membrane-bound P450 cytochromes<sup>26,30,32</sup> and with structural<sup>12</sup> and mutagenesis experiments carried out on HA.<sup>17,45–53</sup>

## 2. COMPUTATIONAL METHODS

**2.1. HA in Water.** Before running the MD simulations of the *HAm* model the force field parameters, the intrinsic flexibility of HA, and the protonation state of Asp309 (Asp309 was simulated both in the protonated and deprotonated forms) were evaluated by running two short MD simulations (20–30 ns) of the extra-membrane region of the protein immersed in explicit water.<sup>54,55</sup>

The X-ray structure of HA (pdb code: 3EQM)<sup>12</sup> was solvated with a box of water having a minimum distance from

the protein surface of 10 Å, and the counterions were added using the *leap* module of AMBER11.<sup>56</sup> The total number of atoms was ~54,000. The test MD simulations in water showed that Asp309 must be protonated to correctly reproduce the X-ray structure of the active site region. In particular, the slightly distorted orientation assumed by the substrate in the X-ray structure is conserved only in the presence of a protonated Asp309 as this orientation is stabilized by a H-bond between the substrate and Asp309 (vide infra). Therefore, this residue was considered in its protonated form during all subsequent MD simulations.

**2.2. HA in the Membrane.** The *HAm* model was built starting again from the recently released X-ray structure of HA.<sup>12</sup> The orientation of HA with respect to the membrane was defined according to mutational studies aimed at understanding which surface of HA interacts with its specific reductase and according to the HA regions tagged by monoclonal antibodies.<sup>2,12</sup> The trans-membrane part of HA, corresponding to the first 44 residues, lacks in the X-ray structure. This region was built by homology modeling, using the trans-membrane helix from the M2 protein of the influenza virus (pdb code: 3ELV)<sup>57</sup> as template and the SPDBV software.<sup>58</sup> Moreover, according to the glycosylation data reported in ref 59 the residues from Met1 to Asn12 have been located in the luminal part of the membrane.

Then, using the VMD program,<sup>60</sup> the protein was inserted in a pre-equilibrated double layer of POPC molecules (kindly provided by T. A. Martinek and co-workers<sup>61</sup>), deleting the POPC molecules clashing with the protein atoms. One Cl<sup>-</sup> counterion was added using the *leap* module of AMBER11 to ensure the electric neutrality of the system and the non trans-membrane regions of HA were solvated with 18361 water molecules. The final model was composed by ~81,000 atoms.

**2.3. Force Field Parameters.** All calculations were done at the Molecular Mechanics (MM) level, using the parm99<sup>62</sup> force field for the protein, the TIP3P water model for the solvent,<sup>63</sup> and the Åqvist parametrization for the counterions.<sup>64</sup>

The bonded and the van der Waals parameters available within the AMBER11 distribution<sup>56,65</sup> were employed for the heme moiety, while the ASD and molecules were simulated using the general AMBER force field (gaff).<sup>66</sup> For the POPC we have used the parametrization of Martinek and co-workers.<sup>61</sup>

The atomic charges of the heme moiety as well as that of the molecular oxygen, of the cysteine residue (Cys437), coordinating the iron, of the iron ion, and of the ASD molecule were derived using the Merz–Kollman scheme.<sup>67,68</sup>

In particular, a small model of the active site, containing the heme moiety, O<sub>2</sub>, and Cys437, was initially optimized with the Gaussian03 program<sup>69</sup> at the DFT-B3LYP level of theory with LANL2DZ and 6-31G\* basis sets for the Fe and the remaining atoms, respectively. In this calculation the position of the α-carbon of Cys437 was frozen to conserve the geometry of the gas phase model as close as possible to that of the active site. A calculation of the ESP charges was subsequently done on the optimized geometry, and the atomic RESP charges were obtained using the *resp* module of AMBER11.<sup>56,70,71</sup>

Recently Gant et al.<sup>72</sup> identified the catalytically active specie as a peroxy-ferric intermediate via EPR spectroscopy. In our calculations of the reduced models of the active site the total charge and spin multiplicity were set up according to these data.

**2.4. Classical MD Simulations.** The MD simulations of the *HAm* and *HAw* models were performed using the NAMD

2.7 program<sup>73</sup> with a time step of 2.0 fs. Before starting the productive MD simulations, the systems were minimized for 1000 steps using the conjugate gradient algorithm and fixing the positions of the backbone atoms of the protein. After the minimization, some (15) short MD simulations were run in an NVT ensemble freezing the coordinate of the backbone atoms and slowly increasing the temperature up to 200 K. After this equilibration phase the systems were released and simulated under NPT conditions for 6 ns, slowly increasing the temperature until the target value of 298.5 K. The temperature was regulated using Langevin dynamics, and the pressure was maintained to 1 atm. For *HAm* we used a semi-isotropic Nosé-Hoover Langevin piston pressure control<sup>74,75</sup> where the X,Y area was maintained constant and the Z dimension of the box was free to move independently. The Langevin piston period was set to 200 fs (fs) with a decay period of 50 fs.

When the *HAm* model was fully equilibrated, a production run of 85 ns was performed. The simulation of the *HAw* was instead carried out for 65 ns, due to faster equilibration.

**2.5. Other Analyses.** H-bond, secstruct, watershell, Root Mean Square Fluctuation (RMSF), and cluster analysis were done with the *ptraj* module of AMBER11.<sup>56</sup> In the H-bond analysis the distance cutoff was set to 3.5 Å and the angle to 120°. The cluster analysis was done with the averagelinkage algorithm, using a Root Mean Square Deviation (RMSD) cutoff of 1.5 Å.<sup>76</sup>

The visual inspection of the trajectories was carried out with the VMD software.<sup>60</sup> This software was used also to prepare all the figures in the paper.

**2.6. Computational Investigation of the Access/Egress Channels.** **2.6.1. Analysis of Access/Egress Pathways Using Caver 2.1.1.** The Caver program (<http://www.caver.cz>)<sup>77</sup> implements a computational procedure based on the Dijkstra algorithm<sup>78</sup> to identify the possible pathways connecting a user defined starting point, located inside the protein, with its surface. This algorithm looks for the most accessible pathways, systematically exploring a grid of points describing the protein. This program is accessible thorough a plug-in for the PyMol program (<http://www.pymol.org>).<sup>79</sup> However, in this implementation, the program can only analyze static protein structures. In our study, the Caver 2.1.1 program<sup>77</sup> was employed to explore the access/egress channels connecting the HA active site to the protein surface. The geometrical center of O<sub>2</sub> was used as origin of the channels in all calculations. Instead in the X-ray structure (pdb code 3EQM), in which O<sub>2</sub> is missing, the iron Cartesian coordinates were used. The analyses were of *HAm* and *HAw* performed on the most representative structures selected by a cluster analysis of the conformations sampled during the MD simulation of *HAw* (1 conformation) and of *HAm* (2 conformations).

To carry out an exhaustive search we required as output 10 possible channels. According to the calculations of Cojocaru et al.<sup>26</sup> only the channels with a minimal radius larger than 1.2 Å were retained and analyzed.

**2.6.2. Access/Egress Pathways Analysis by Random Acceleration Molecular Dynamics (RAMD).** RAMD<sup>33</sup> is one of the most used computational methods to study the access/egress pathways of ligands in proteins. This has been recently implemented in the NAMD program.<sup>80</sup>

In RAMD a randomly oriented force is applied to the ligand atoms, for a certain number of steps (nSTP). This force maintains a constant magnitude during the nSTP, but, if the ligand encounters steric hindrance along its expulsion route

(i.e., it is not able to cover a user defined minimal distance ( $R_{\min}$ )), a new random direction is chosen and maintained for other nSTP.

To avoid the sampling of very unlikely pathways, the acceleration should be selected as small as possible. Therefore, several preliminary calculations are necessary to choose the optimal simulation conditions (the nSTP and the acceleration values).

In our simulations the direction of the force was kept constant for 50 MD steps (100 fs) and changed if the ligand was not able to cover a given Rmin (see Table 1). Different

**Table 1. Summary of the Different Operative Conditions Employed in the RAMD Calculations**

ligand	acceleration kcal/mol* Å*amu	Rmin (Å)	total RAMD calculations	expulsion ratio	expulsion (%)
ASD	0.075	0.008	125	18/125	14.5
ASD	0.10	0.008	100	68/100	68
OXX	0.25	0.006	160	14/160	8.7
ASD (in water)	0.075	0.008	100	44/100	44

acceleration values were tested in the range between 0.25 kcal/mol\* Å \*amu and 0.075 kcal/mol\* Å \*amu to find the smallest acceleration value that guarantees an expulsion rate (i.e., expulsion events/RAMD calculations) of 8–15%.

The RAMD simulations were considered unsuccessful and stopped if the dissociation was not detected in 50 ps, while they were considered successful if the ligand reached a distance of 35 Å from the protein surface.

We used RAMD to investigate the access/egress paths of both OXX and ASD. In the calculations carried out on OXX, to avoid the expulsion of ASD, due to the pushing of OXX, we used a weak harmonic restraint between ASD and the iron atom.

As in other stochastic methodologies, also in RAMD a large number of independent calculations is necessary to explore all the possible access/release routes. Therefore, several single RAMD calculations were carried out, specifying a different random seed number to generate different trajectories. The number of simulations in the different operative conditions (i.e., acceleration and  $R_{\min}$ ), for a total simulation time of ~24 ns, is summarized in Table 1.

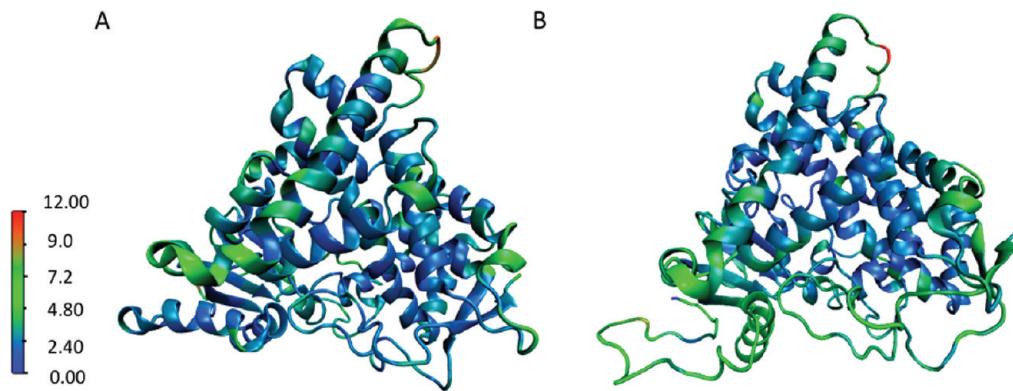
The coordinates of the center of mass (COM) have been printed during all successful expulsion trajectories and visualized together with the protein. This procedure allowed performing a visual cluster analysis, reducing the expulsion routes only to some representative pathways.

**2.6.3. Steered Molecular Dynamics (SMD).** SMD is a computational technique, which induces the dissociation of a ligand by applying an external force.<sup>33,41,80–82</sup>

Here, we applied SMD to study the dissociation of ASD along the routes identified by RAMD. In these simulations the COM of a molecule is connected, via a spring of known force constant, to a dummy atom moving in a predefined direction, and the pulling force is applied according to eq 1

$$\vec{F}(t) = K[\vec{v}t - (\vec{r} - \vec{r}_0) \cdot \vec{n}] \quad (1)$$

where  $K$  is the force constant of the spring,  $\vec{v}$  is the pulling velocity,  $\vec{n}$  is the pulling direction normal, and  $\vec{r}$  and  $\vec{r}_0$  are the positions of the ligand COM at the time  $t$  and at the initial time. In our simulations a spring with a force constant of 7



**Figure 1.** Most representative HA structures resulting from the MD simulation in water (A) and in the membrane model (B). The structures are colored according to the average RMSD per residue with respect to the X-ray structure (calculated considering only the  $C\alpha$ ).

kcal/mol was used. Moreover, all SMD simulations were run in a NVT ensemble, using a time step of 1 fs and the same settings described for the standard MD simulations.

To avoid a drift of the protein during the simulation, the positions of the  $C\alpha$  of the trans-membrane region (the first 44 residues) were fixed using a harmonic restraint with force constant of 5 kcal/mol.

After visual inspection of the RAMD simulations, considering the statistical significance of the expulsion directions and a reasonable compromise between accuracy and computational cost, 11 pulling directions were considered as preliminary guess for the SMD calculations. These simulations were run in constant velocity conditions ( $0.025 \text{ \AA}/\text{ps}$ ). This allowed selecting three paths (SMD\_1, SMD\_2, SMD\_3), considering as selection criteria the choice of different pulling directions, which were requiring the smallest work to dissociate ASD. Then, to sample these three directions more accurately, 8 SMD simulations were performed for each expulsion path, reassigning the initial velocities at the temperature of 298.5 K and using a slower constant pulling velocity (i.e.,  $0.0125 \text{ \AA}/\text{ps}$ ).

**2.6.4. Implicit Ligand Sampling (ILS) Calculations.** We also applied the ILS technique to study the access path of OXX to the active site. ILS has been extensively validated in this kind of applications.<sup>83–87</sup> This method relies in the assumption that the gas migration does not happen along permanent channels but following pathways connecting transient cavities. In fact, cavities formed by thermal fluctuations of proteins are able to temporarily host gas molecules.<sup>85</sup> In an ILS analysis the free energy to place a probe, in this case a probe mimicking the properties of the oxygen molecule, is iteratively calculated on discrete points, located on an equally spaced grids over the protein.

Here, the ILS calculations were done considering 4300 snapshots sampled during the *HAm* MD simulation and 3300 snapshots sampled in the *HAw* MD simulation. In all the calculations a grid resolution of  $0.5 \text{ \AA}$  was used. All the ILS calculations were performed using the VMD plugin.<sup>85</sup>

### 3. RESULTS AND DISCUSSION

#### 3.1. MD Simulations of the HA/Membrane and Water Models.

We initially built a model of HA embedded in model membrane according to the proposals of Ghosh et al.<sup>12</sup> and Hong et al.<sup>2</sup> In our model the A, A', G', and K' helices and the sheet 1, 2, 8, and 9 are in contact with the membrane, while residues Lys108 and Lys420, which are essential for the HA interaction with its specific reductase,<sup>2</sup> are water exposed. In

this model HA emerges by  $\sim 45 \text{ \AA}$  with respect to the plane of the membrane bilayer, assuming an orientation similar to that adopted by Berka et al.<sup>28</sup> for the 2C9 cytochrome, and in agreement with the experimental studies of Bayburt et al.<sup>88</sup>

The *HAm* model underwent 85 ns of classical MD to evaluate both the stability of the protein fold and the effect of the membrane apolar environment on the protein structural and dynamical features. The results of this simulation have been compared with those obtained by a 65 ns long MD simulation of *HAw*.<sup>63</sup>

*HAm* appears to be stable during the 85 ns. In fact, the RMSD, calculated on the backbone atoms of the residues solved in X-ray structure (i.e., Ser45-Asn496), reaches a constant average value ( $3.01 \pm 0.05 \text{ \AA}$ ) after 30 ns, and it oscillates around this value during the last 55 ns (Figure S2). A similar trend is observed for *HAw*, which reaches a stable RMSD average value of  $1.64 \pm 0.08 \text{ \AA}$  after 30 ns (Figure S3).

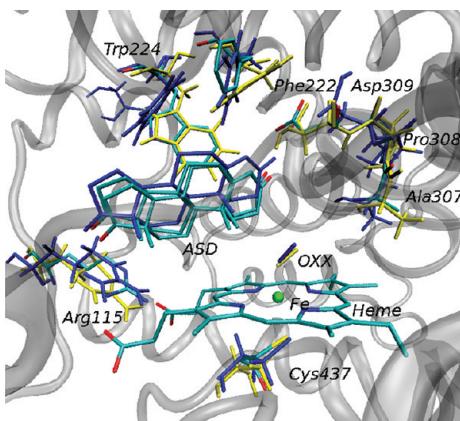
In *HAm*, the protein conserves its overall fold, although helix A becomes unfolded (a complete classification of the HA helices is available in Table S1). However, this phenomenon does not alter the position of the protein with respect to the membrane plane. The unfolding of helix A is most likely due to the influence of the lipophilic environment of the membrane. In fact, this change of secondary structure does not take place in the simulation of *HAw*.

To better evaluate how the membrane environment influences the different regions of the protein, we calculated the average RMSD per residue, with respect to the X-ray structure. As expected this analysis shows that the most significant structural perturbations occur at the protein/membrane interface (Figure 1A-B).<sup>12</sup> In fact, the RMSD has a high value in the N-terminal region (around Ser45), accordingly with the observed unfolding of helix A. The stability of the protein/membrane model has been checked calculating the distance between the centers of mass of the extra membrane domain of the protein and the membrane, resulting in an average distance of  $34.2 \pm 0.5 \text{ \AA}$ .

We have also calculated the RMSF per residues for both the MD simulations in water and in the membrane. The RMSFs show (Figure S4) a similar and small flexibility of HA in both simulation conditions, with the only significant differences localized in the loop Ser269-Met276 that connects helices H and G and in the loop formed by Val383-Tyr386. This latter is directly in contact with the membrane surface. Thus, its flexibility is clearly enhanced in *HAm* with respect to *HAw*.

Interestingly, the Ser269-Met276 loop, which reaches an RMSF value of 7 Å in *HAm*, is characterized by a high B-factor also in the X-ray structure (a picture of crystal structure colored according to the B-factor value is available in Figure S5), confirming the reliability of our findings.

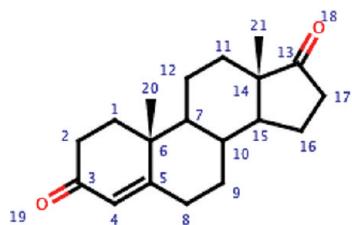
In *HAm* ASD maintains the same position occupied in the X-ray structure (Figure 2), and the RMSD for all the ASD heavy



**Figure 2.** Superposition of the most representative structures of *HAm* (colored in yellow), of *HAw* (colored in blue), and of the X-ray structure (colored by atom type). The protein backbone is shown as gray ribbons.

atoms (measured after the alignment of the protein backbone atoms comprised in a shell of 30 Å from ASD at the beginning of the simulation) is of  $1.05 \pm 0.22$  Å.

Since Asp309 is known to play a key role in the substrate binding and recognition,<sup>89,90</sup> we have accurately checked its behavior during the simulation. In the last 60 ns of the *HAm* MD simulation Asp309 H-bonds with O19 (for atom numbering see Figure 3) of ASD (occupancy: 90%, average



**Figure 3.** Structure and atom numbering of ASD.

distance:  $2.8 \pm 0.1$  Å, average angle:  $15 \pm 9^\circ$ ), retaining the same orientation displayed in the X-ray structure. This is consistent with the role of Asp309 in stabilizing the HA/ASD complex hypothesized on the basis of the X-ray structure and site-directed mutagenesis experiments.<sup>12,89,90</sup>

A similar geometry is observed also in the simulation of *HAw*. However, in this case, the H-bond between ASD and Asp309 is less frequent (occupancy: 30%, average distance:  $2.9 \pm 0.3$  Å, average angle:  $30^\circ \pm 13^\circ$ ).

This behavior may be induced by the rotation of Trp224 that occurs only in the *HAw* simulation (Figure 2). Actually this movement brings the aromatic ring of Trp224 far apart from ASD in a position different with respect to the crystallographic one, changing the local environment around O19 of ASD.

An analysis of the first solvent shell of Oδ1@Asp309 indicates an average number of water molecules of  $1.45 \pm 1.04$

during the MD simulation of *HAm*, confirming the presence and the stability of the water molecule solved in the X-ray structure.<sup>12</sup>

The OXX was instead not solved in the X-ray structure. Therefore, we have reconstructed it in our model, and we have verified that its binding to the heme moiety does not alter the stability and the orientation of ASD. In our simulations, OXX H-bonds with the OH group of Thr310 (occupancy: 79%, average distance:  $2.8 \pm 0.2$  Å, average angle:  $25 \pm 14^\circ$ ).

### 3.2. ASD and OXX Pathways to/from the Catalytic Site.

We have also evaluated the influence of the membrane on the access/egress pathways of ASD and OXX to/from the catalytic site. To this end we have employed different computational techniques ranging from those based on "channel detecting" algorithms<sup>78</sup> to those grounded on the MD simulations.<sup>91,92</sup> Moreover, to obtain also a qualitative energetic rank of the possible pathways we have used SMD simulations and ILS calculations, for ASD and OXX respectively. In order to evaluate the influence of the lipophilic environment of the membrane on the number and location of the access/release pathways, all calculations, except the computationally expensive SMDs, were carried out for both *HAm* and *HAw* models.

**3.2.1. Channel Detection Using Caver.** In the last years, systematic studies on several P450 cytochromes were done with the help of channel detecting methodologies on static X-ray structures or on selected conformations obtained from short MD simulations in explicit solvent.<sup>26,30,32,35,36,93</sup>

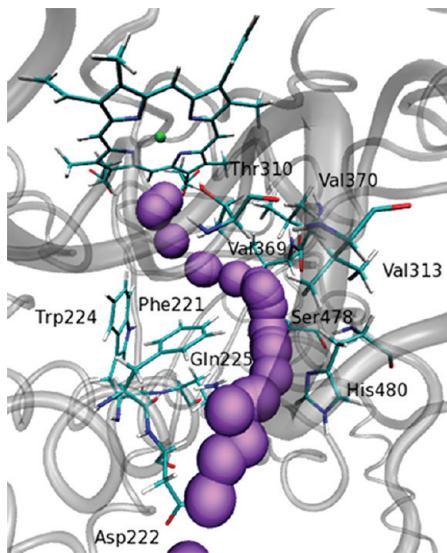
Here, a first analysis to detect the access/egress channels of HA has been carried out on *HAm* by employing the Caver program.<sup>26,94</sup> In an attempt to take into account the protein dynamics,<sup>95,96</sup> this analysis was done on the two most representative structures (the first accounting for the 49% and the second for the 25% of the protein conformations) obtained from a cluster analysis of the *HAm* MD trajectory.

Ten different channels have been obtained, although they all converge to a unique pathway with its entrance close to the membrane bilayer and its end close to the heme moiety. This channel is delimited by residues Phe221, Trp224, Gln225, Val313, Thr320, Val369, Val370, Ser478, and His480 (Figure 4) and approximately corresponds to the 2d and 2f channels according to the nomenclature proposed by Cojocaru et al. (Figure S6).<sup>26</sup> The influence of Gln225<sup>45</sup> and Ser478<sup>52</sup> on ligand binding was investigated by site directed mutagenesis experiments, which indicated Gln225 as a possible gatekeeper in a putative access/egress channel to the catalytic site.<sup>45</sup> Indeed, the Q225A mutant, where the side chain size is reduced, showed an increased apparent affinity (lower  $K_m$ ) and an increased apparent  $V_{max}$  consistently with the fact that Gln225 may play an important role in regulating the access/egress of the substrate to/from the catalytic site. In line with these experimental data, our analysis of the channel radius shows two bottlenecks: the first in proximity Gln225 and the second between Val369 and Ser478.

We have then evaluated the influence of the membrane environment on the channels location, by repeating this analysis on the simulation of *HAw*. The Caver analysis, carried out on the most populated cluster sampled along the *HAw* trajectory (~90% of the conformations), resulted into the same channel observed for the membrane-bound model.

### 3.2.3. Random Acceleration (Expulsion) MD Simulations.

Subsequently, we have applied RAMD simulations to study the reactants dissociation path from the active site. RAMD has been



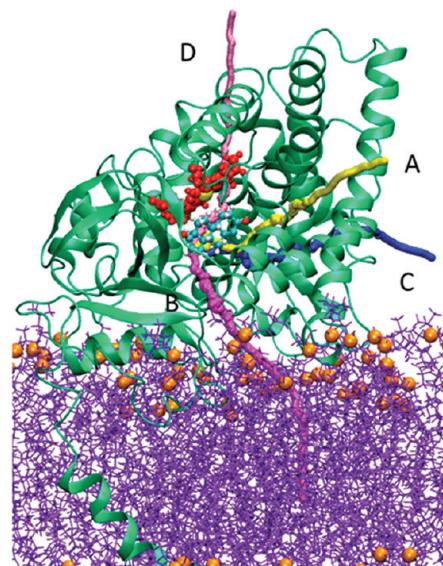
**Figure 4.** The access/egress channel of ASD to/from the catalytic site as identified by the Caver. The channel is depicted as violet van der Waals spheres. The protein backbone is depicted in gray tubes. The heme moiety and the residues delimiting the channel are depicted in licorice and colored according to the atom name. The iron ion appears as a green sphere.

frequently applied to study the ligand expulsion pathways from the active site of P450 cytochromes.<sup>26,33,34</sup> Here, after having selected the optimal simulation conditions to obtain an expulsion rate of 10–15%, we have run 125 RAMD simulations for ASD and 160 for OXX (see Table 1).

For ASD 18 expulsion events have been sampled. All paths start inside the active site, with the substrate exiting from HA in different positions. The majority of these (50%) involved a route with the exit (entrance) located between helices F and G (path A in Figure 5). Moreover, three additional expulsion directions have been identified (paths B, C, and D in Figure 5). Path B (22%) had the exit (entrance) in contact with the membrane, path C (17%) was located between the A' and A helices, and path D (11%) was placed between helixes H and C. Path A roughly corresponds to the solvent (S) channel of the nomenclature introduced by Cojocaru et al. (Figure S6A).<sup>26</sup> B, instead, corresponds to the zone occupied by the 2f, 2a, and 2ac channels, and C corresponds to the 2c channel. D corresponds to a small percentage of the expulsion events, and it does not have a clear correspondence with any of the channels labeled by Cojocaru et al.<sup>26</sup> The most probable pathway (A) involves residues Ala212, Ile217, Tyr220, Tyr221, Ile255, Val257, Arg264, Asp309, and Leu310. Consistently with these data, mutagenesis experiments demonstrated that Asp309 and Leu310 have a remarkable influence on the HA catalytic activity.<sup>45,49,89</sup> However, the proximity of these two residues to the catalytic site does not allow to clearly establish if they may be involved in the access/egress pathway of the substrate or in the enzymatic reaction.

Notably, recent studies, performed by Di Nardo et al.<sup>90</sup> on the influence of the pH in regulating the substrate binding to HA, suggested the presence of an ionizable residue, such as Asp309, at the entrance of the ASD binding tunnel, supporting our findings that path A may be the preferential route for the substrate binding.<sup>90</sup>

Path B, still representative from a statistical point of view, is very similar to the channel detected by Caver. A visual



**Figure 5.** The four possible expulsion paths A (yellow), B (magenta), C (blue), and D (pink), as obtained from RAMD simulations. The protein backbone is depicted in green. The heme moiety and the iron ion are depicted as red and yellow van der Waals spheres, respectively. ASD is depicted as ball and sticks colored by atom name. The lipids of the membrane are represented in violet lines with the phosphate atoms highlighted as orange spheres.

inspection of this path pointed out that, in addition to Gln225, also His480 may be involved in a gate regulating the entrance/egress of ASD to/from the catalytic site. Consistently with this, site-directed mutagenesis experiments<sup>52</sup> showed that two H480K and H480Q mutants display reduced  $K_m$  with respect to the wild type enzyme. Then, considering the RAMD calculations and the results of the site-directed mutagenesis experiments, we hypothesize that the region delimited by Phe221, Asp222, Trp224, Gln225, Asp476, Ser478, His480, Asp482, and Glu483 may form a flexible gate, regulating the access/release of the substrate to/from the catalytic site. Although, this pathway does not seem to be the most favorable for ASD.

RAMD simulations have been also performed to study the dissociation of ASD in HA<sub>w</sub>, using the same simulation conditions of HA<sub>m</sub>. In this case, the expulsion events were very frequent (expulsion percentage of 44%), and no preferential expulsion direction was identified (Figure S7). Therefore, these results suggest that the membrane lipophilic environment helps in stabilizing some specific pathways.

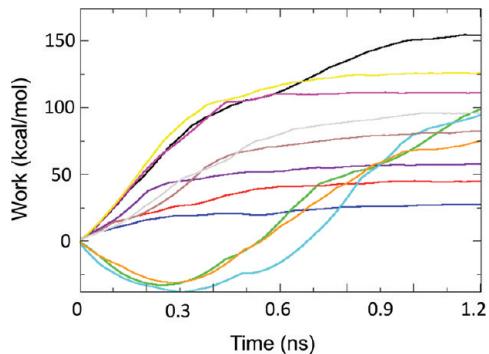
**3.2.4. SMD Simulations of ASD Dissociation.** To estimate which is the most likely dissociation pathway of ASD, among those identified by RAMD, we have also carried out SMD simulations, a computational method frequently applied to P450 cytochromes.<sup>33,41,80,81,97</sup>

If a larger number of independent slow simulations is carried out, these kinds of simulations can potentially give a reliable measure of the free energy associated with the dissociation of the ligand by using the Jarzinsky equality.<sup>98</sup> However, we have employed SMD here only to calculate the work necessary to dissociate ASD along the paths identified by RAMD simulations.<sup>82</sup> This allows us to obtain only a qualitative energetic rank of the different paths.

Before starting the SMD calculations, we have visually checked all the expulsion events sampled in the RAMD

calculations, and, after having discarded very similar directions, we have defined 11 pulling vectors to allow the ligand to move in the direction of the selected RAMD path.

The simulations in which ASD was pulled along path C (yellow and magenta lines in Figure 6) and along the path D



**Figure 6.** Work necessary to dissociate ASD form the catalytic site of HA during the preliminary SMD simulations.

(black line in Figure 6) resulted in a larger work than the other pulling directions (Figure 6). Moreover, several other pulling directions were characterized by a large work.

Therefore, we selected three directions (the cyan, orange, and green lines of Figure 6, hereafter named SMD\_1, SMD\_2, and SMD\_3, respectively), using as selection criteria that fact that they had to correspond to different pulling directions and they had to require the smallest possible work to dissociate ASD. SMD\_1 roughly corresponds to path A as identified by RAMD, while both SMD\_2 and SMD\_3 correspond to path B, although following slightly different routes. Remarkably, all these directions have entrances/exits located on HA sides not involved in the binding of the specific reductase.<sup>2</sup>

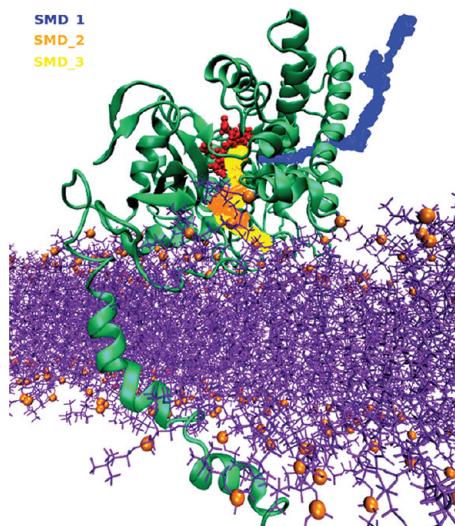
Then, to obtain a more accurate sampling and a reliable estimate of the work for the three selected directions, we repeated the simulations with a reduced pulling velocity to dissociate ASD from the binding site (8 simulations for every pulling directions).

These simulations resulted in average works of  $77 \pm 28$  kcal/mol,  $184 \pm 38$  kcal/mol, and  $134 \pm 25$  kcal/mol and for SMD\_1, SMD\_2, and SMD\_3, respectively (Figures 7 and 8).

A large standard deviation is observed in the values of the work. This may be mainly due to the inclusion of the membrane environment in the simulation. In fact, the lipophilic environment enhances the structural protein flexibility of the protein, changing the molecular environment of the access/egress channels in the different simulations.

This large standard deviation may be possibly reduced using a slower velocity to pull the ligand out of the protein and/or a larger force constant for the coupling with the dummy atoms. However, an accurate estimation of the binding energy is beyond the goals of this work, and the values obtained by our simulation intend only to qualitatively rank the accessibility of the different access/egress channels. Notably, these values are in line with the same calculations performed on other cytochromes.<sup>97</sup>

In spite of the hypothesis formulated by Ghosh et al.<sup>12</sup> and of the results of the Caver calculations, the calculated works, as well as the average profile of the forces (Figure S8), confirm SMD\_1 as the most likely egress route from the catalytic site. Notably, SMD\_1 corresponds also to the most frequent



**Figure 7.** Coordinates of the ASD center of mass during the SMD simulations along the SMD\_1 (blue), SMD\_2 (orange), and SMD\_3 (green) directions. The protein backbone is shown in green. The heme moiety and the iron ion are depicted as red and yellow van der Waals spheres, respectively. The lipids of the membrane are represented in violet lines with the phosphate atoms highlighted as orange spheres.

expulsion path of the RAMD simulations performed on the HAm model.

Moreover, the average force profile along SMD\_1 is smoother than those obtained along the other two directions (Figure S8). Indeed, a large force, necessary to reorient the ligand inside the catalytic pocket, is measured at the beginning of the simulation, but afterward the average force profile is rather smooth and ASD can easily travel outside the protein.

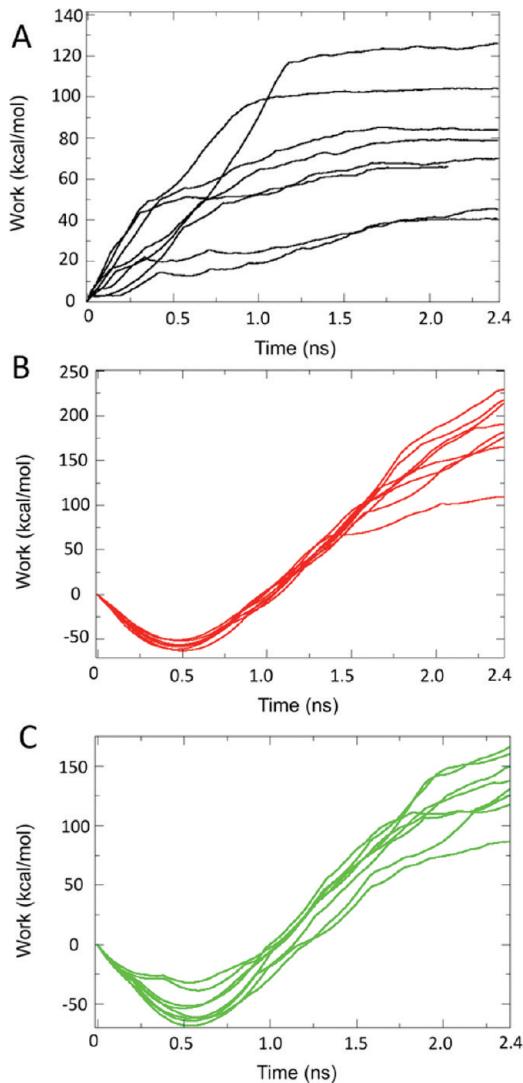
The other two sampled pathways (SMD\_2 and SMD\_3) lead to the membrane surface. Among these, SMD\_3, where ASD travels along a hydrophobic channel formed by Ile70, Ile89, Phe134, Phe221, Trp224, Ile229, and Leu477, is more favorable than SMD\_2.

### 3.2.5. Oxygen Access/Egress Pathway to the Catalytic Site.

We have initially used RAMD to investigate the access/egress paths of OXX to/from the catalytic site. Most of the expulsion trajectories (93%) followed the B channel identified for ASD (blue channel in Figure 9, named A<sub>ox</sub>), while a second possible expulsion route (7%, magenta channel in Figure 9, named B<sub>ox</sub>) passed through the helices F and B. Since, RAMD has been never applied to study the diffusion of small gas molecules inside proteins, we have also considered ILS to verify the reliability of these results.

ILS, in fact, has been extensively applied to study the diffusion of small molecules inside proteins.<sup>83–86</sup> Therefore, we here have applied it to study the OXX path through HA, and we have compared its results with those of RAMD.

The ILS calculations were performed on equally spaced snapshots extracted from the HAm MD simulation with the aim of detecting the pathway connecting the most frequent transient cavities formed during the MD simulation. This analysis has pointed out the presence of a channel delimited by the residues Arg193, Leu190, Met195, Met446, Thr198, Leu191, and Ile450 (Figure 10A). This channel slightly differs from A<sub>ox</sub> although its entrance is located on the HA side in contact with the membrane, supporting the idea that, having the gas diffusion through the membrane a very flat free energy



**Figure 8.** Work necessary to dissociate ASD form the catalytic site of HA during SMD\_1 (black), SMD\_2 (red), and SMD\_3 (green) directions.

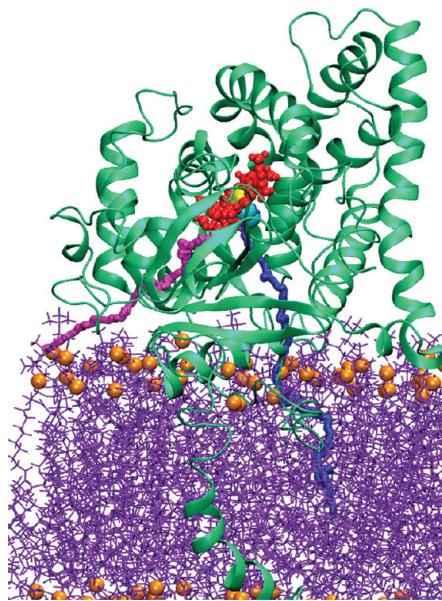
profile,<sup>99</sup> and the gas permeation occurs from the membrane and not from the cytosol.

The ILS calculations have been repeated also considering snapshots extracted from the MD simulation of *HAw*. In this case the preferential channel was less defined, and its entrance is located in the cytosolic side of the protein between the C helix and the protein loop between Phe390 and Lys400 (Figure 10B). The relative energetic values associated with these channels, although highly qualitative, suggest that the membrane stabilizes the oxygen diffusion channel and that this most probably takes place through the membrane.

#### 4. SUMMARY AND CONCLUSIONS

In this study we have applied a variety of simulation techniques to verify the influence of a membrane-mimicking environment on the structural and dynamical properties of HA as well as on the access/egress pathways of the reactants to/from its catalytic site.

Our simulations of *HAw* and *HAm* revealed the most marked differences in the flexibility and in the RMSD of the protein are in the N-terminal region of HA, which is embedded in the



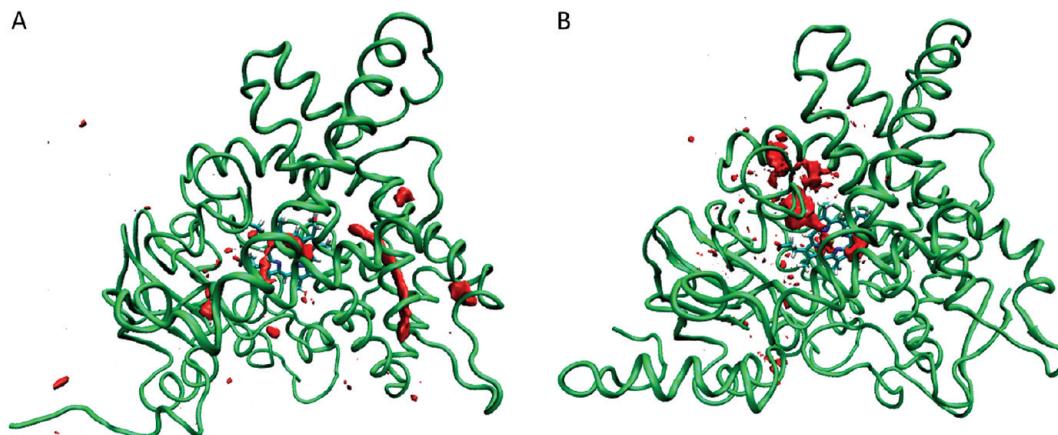
**Figure 9.** Paths A<sub>ox</sub> (blue) and B<sub>ox</sub> (magenta) for OXX access/egress as identified by RAMD simulations. The protein backbone is depicted in green. The heme moiety and the iron ion are depicted as red and yellow van der Waals spheres, respectively. The lipids of the membrane are represented in violet lines with the phosphate atoms highlighted as orange spheres.

membrane, consistently with the computational studies of other cytochromes.<sup>27,28</sup> In addition, our results indicate that Asp309, a residue known to play a key role for HA activity, has to be protonated to conserve the position determined in the X-ray structure.<sup>90</sup>

Then, we have identified possible access/egress routes to/from the catalytic site for both ASD and OXX. Comparing the results of the RAMD simulations in water and in the lipophilic membrane environment we saw that the access/release pathways of ASD are stabilized by the membrane. In fact, clear preferential pathways for ASD binding/dissociation are not detectable if HA is embedded in water.

We have also performed SMD simulations to identify which is the most energetically favored dissociation route for ASD in *HAm*. The route SMD\_1 (Figures 7 and 8), connecting the active site to a fully solvated protein surface, which corresponds to the S channel of the Cojocaru et al.'s nomenclature (Figure S6),<sup>26</sup> appears as the most favored pathway as it displays the lowest work necessary to dissociate ASD from the active site. These data are consistent with experimental findings appearing in the literature for HA<sup>90</sup> and with theoretical investigations carried out on other cytochromes.<sup>30</sup>

We have also studied the access/egress paths of OXX to/from the active site by performing RAMD and ILS simulations. From RAMD calculations two possible and slightly different pathways were identified for *HAm*, both ending at the protein side in contact with the membrane. Instead, in the ILS simulations different channels were identified for *HAm* and *HAw*, pointing toward the membrane and the cytosol, respectively. The relative energetic values associated with the detected channels, although highly qualitative, suggest that the channel observed in *HAm* is energetically favored. Therefore, the lipophilic environment of the membrane may help in also stabilizing the formation of the OXX channel.



**Figure 10.** OXX access/release channels identified by ILS in the *HAm* (A) and in *HA<sub>w</sub>* (B). For the sake of clarity, the ILS iso-surface has been cut to an upper interaction energy value of  $-2.4 \text{ kcal/mol}$  and  $-1.2 \text{ kcal/mol}$  for *HAm* and *HA<sub>w</sub>*, respectively. Although not clearly visible in this picture, we remark that the channels connect the external surface of the protein with the catalytic site.

In summary, in this study we point out that the membrane environment has a role in stabilizing the access/egress paths of the reactants, especially when they point toward the membrane.<sup>100</sup> [When this manuscript was ready for submission Jiang and Ghosh reported about normal-mode analysis studies about HA in both membrane-free and membrane-integrated models. Interestingly they found results similar to those reported in this paper, concerning both the membrane-protein orientation and the influence of the membrane on the intrinsic mechanical properties of the protein. (See ref 100.)] Our findings may be shared also by other membrane-bound P450 cytochromes. Due to the influence of the channels on substrate selectivity of this family of proteins, and due to the importance of P450s in biotechnology, medicine, and bioremediation, a detailed understanding of the role of membrane in shaping entrance/exit channels would help to exploit these enzymes in proteins and pathways engineering and for drug design.

## ■ ASSOCIATED CONTENT

### Supporting Information

Scheme of the catalytic reaction performed by human aromatase. RMSD value of the C $\alpha$  atoms calculated with respect to the X-ray structure during the 85 ns of the HA/membrane (*HAm*) (MD) simulation. RMSD value of the C $\alpha$  atoms calculated with respect to the X-ray structure during the 65 ns of MD simulation of HA in water (*HA<sub>w</sub>*).

RMSF values per residues calculated for the residues ranging from Ser45 to Asn496 during the 65 ns and the 85 ns of the MD simulations of *HA<sub>w</sub>* and *HAm*. HA structure solved by X-ray (pdb code: 3EQM) colored according to the B-factor. General scheme of the channels identified during the Random Acceleration MD (RAMD) and the steered MD (SMD) simulations. Expulsion directions sampled during the RAMD calculations of *HA<sub>w</sub>*. Average profile of the force necessary to dissociate the ASD molecule. Classification of the helices in the HA structure. Force field parameters not available in the AMBER11 distribution for ASD, OXX, HEME, and Cys437. This material is available free of charge via the Internet at <http://pubs.acs.org>.

## ■ AUTHOR INFORMATION

### Corresponding Author

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

### Notes

The authors declare no competing financial interest.

## ■ ACKNOWLEDGMENTS

We acknowledge the CINECA award N. HP10BMPFH0, 2011 for the availability of high performance computing resources and support.

We thank T. A. Martinek for kindly supplying the force field parameters of the membrane model. J.S. thanks M. Marti, F. Forti, and J. P. Bustamante for the important help with the ILS calculations. J.S. also thanks F. Colizzi and D. Franco for the interesting discussions.

## ■ REFERENCES

- (1) Kumar, S. Engineering cytochrome P450 biocatalysts for biotechnology, medicine and bioremediation. *Expert Opin. Drug Metab. Toxicol.* **2010**, *6*, 115–131.
- (2) Hong, Y.; Li, H.; Ye, J.; Miki, Y.; Yuan, Y.-C.; Sasano, H.; Evans, D. B.; Chen, S. Epitope Characterization of an Aromatase Monoclonal Antibody Suitable for the Assessment of Intratumoral Aromatase Activity. *PLoS One* **2009**, *4*, e8050.
- (3) Czajka-Oraniec, I.; Simpson, E. R. Aromatase research and its clinical significance. *Endokrynol. Pol.* **2010**, *61*, 126–134.
- (4) 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.
- (5) Altundag, K.; Ibrahim, N. K. Aromatase Inhibitors in Breast Cancer: An Overview. *The Oncologist* **2006**, *11*, 553–562.
- (6) Jones, M. E. E.; Boon, W. C.; McInnes, K.; Maffei, L.; Carani, C.; Simpson, E. R. Recognizing rare disorders: aromatase deficiency. *Nat. Clin. Pract. Endocrinol. Metab.* **2007**, *3*, 414–421.
- (7) Saldanha, C. J.; Duncan, K. A.; Walters, B. J. Neuroprotective actions of brain aromatase. *Front. Neuroendocrinol.* **2009**, *30*, 106–118.
- (8) Garcia-Segura, L. M.; Veiga, S.; Sierra, A.; Melcangi, R. C.; Azcoitia, I. Aromatase: a neuroprotective enzyme. *Prog. Neurobiol.* **2003**, *71*, 31–41.
- (9) Laville, N.; Balaguer, P.; Brion, F.; Hinfray, N.; Casellas, C.; Porcher, J.-M.; Ait-Aissa, S. Modulation of aromatase activity and mRNA by various selected pesticides in the human choriocarcinoma JEG-3 cell line. *Toxicology* **2006**, *228*, 98–9108.

- (10) Banting, L.; Ahmed, S. Aromatase: the enzyme and its inhibition. *Anticancer Agents Med. Chem.* **2009**, *9*, 627–641.
- (11) Lala, P.; Higashiyama, T.; Erman, M.; Griswold, J.; Wagner, T.; Osawa, Y.; Ghosh, D. Suppression of human cytochrome P450 aromatase activity by monoclonal and recombinant antibody fragments and identification of a stable antigenic complex. *J. Steroid Biochem. Mol. Biol.* **2004**, *88*, 235–245.
- (12) Ghosh, D.; Griswold, J.; Erman, M.; Pangborn, W. Structural basis for androgen specificity and oestrogen synthesis in human aromatase. *Nature* **2009**, *457*, 219–223.
- (13) Maurelli, S.; Chiesa, M.; Giamello, E.; Di Nardo, G.; Ferrero, V. E.; Gilardi, G.; Van Doorslaer, S. Direct spectroscopic evidence for binding of anastrozole to the iron heme of human aromatase. Peering into the mechanism of aromatase inhibition. *Chem. Commun.* **2011**, *47*, 10737–10379.
- (14) Graham-Lorence, S.; Khalil, M. W.; Lorence, M. C.; Mendelson, C. R.; Simpson, E. R. Structure-function relationships of human aromatase cytochrome P-450 using molecular modeling and site-directed mutagenesis. *J. Biol. Chem.* **1991**, *266*, 11939–11946.
- (15) Koymans, L. M. H.; Moereels, H.; Vanden Bossche, H. A molecular model for the interaction between vorozole and other non-steroidal inhibitors and human cytochrome P450 19 (P450 aromatase). *J. Steroid Biochem. Mol. Biol.* **1995**, *53*, 191–197.
- (16) Laughton, C. A.; Zvelebil, M. J.; Neidle, S. A detailed molecular model for human aromatase. *J. Steroid Biochem. Mol. Biol.* **1993**, *44*, 399–407.
- (17) Graham-Lorence, S.; Amarneh, B.; White, R. E.; Peterson, J. A.; Simpson, E. R. A three-dimensional model of aromatase cytochrome P450. *Protein Sci.* **1995**, *4*, 1065–1080.
- (18) Chen, S.; Kao, Y. C.; Laughton, C. A. Binding characteristics of aromatase inhibitors and phytoestrogens to human aromatase. *J. Steroid Biochem. Mol. Biol.* **1997**, *61*, 107–115.
- (19) Lewis, D. F. V.; Lee-Robichaud, P. Molecular modelling of steroidogenic cytochromes P450 from families CYP11, CYP17, CYP19 and CYP21 based on the CYP102 crystal structure. *J. Steroid Biochem. Mol. Biol.* **1998**, *66*, 217–233.
- (20) Chen, S.; Zhang, F.; Sherman, M. A.; Kijima, I.; Cho, M.; Yuan, Y. C.; Toma, Y.; Osawa, Y.; Zhou, D.; Eng, E. T. Structure-function studies of aromatase and its inhibitors: A progress report. *J. Steroid Biochem. Mol. Biol.* **2003**, *86*, 231–237.
- (21) Loge, C.; Le Borgne, M.; Marchand, P.; Robert, J. M.; Le Baut, G.; Palzer, M.; Hartmann, R. W. Three-dimensional model of cytochrome P450 human aromatase. *J. Enzyme Inhib. Med. Chem.* **2005**, *20*, 581–585.
- (22) Favia, A. D.; Cavalli, A.; Masetti, M.; Carotti, A.; Recanatini, M. Three-dimensional model of the human aromatase enzyme and density functional parameterization of the iron-containing protoporphyrin IX for a molecular dynamics study of heme-cysteinato cytochromes. *Proteins: Struct., Funct. Bioinf.* **2006**, *62*, 1074–1087.
- (23) Karkola, S.; Höltje, H. D.; Wähälä, K. A three-dimensional model of CYP19 aromatase for structure-based drug design. *J. Steroid Biochem. Mol. Biol.* **2007**, *105*, 63–70.
- (24) Conner, K. P.; Woods, C. M.; Atkins, W. M. Interactions of cytochrome P450s with their ligands. *Arch. Biochem. Biophys.* **2011**, *507*, 56–65.
- (25) Denisov, I. G.; Makris, T. M.; Sligar, S. G.; Schlichting, I. Structure and Chemistry of Cytochrome P450. *Chem. Rev.* **2005**, *105*, 2253–2278.
- (26) Cojocaru, V.; Winn, P. J.; Wade, R. C. The ins and outs of cytochrome P450s. *Biochim. Biophys. Acta* **2007**, *1770*, 390–401.
- (27) Cojocaru, V.; Balali-Mood, K.; Sansom, M. S. P.; Wade, R. C. Structure and Dynamics of the Membrane-Bound Cytochrome P450 2C9. *PLoS Comput. Biol.* **2011**, *7*, e1002152.
- (28) Berka, K.; Hendrychova, T.; Anzenbacher, P.; Otyepka, M. Membrane Position of Ibuprofen Agrees with Suggested Access Path Entrance to Cytochrome P450 2C9 Active Site. *J. Phys. Chem. A* **2011**, *115*, 11248–11255.
- (29) Denisov, I. G.; Shih, A. Y.; Sligar, S. G. Structural differences between soluble and membrane bound cytochrome P450s. *J. Inorg. Biochem.* **2011**, *108*, 150–158.
- (30) Otyepka, M.; Berka, K.; Anzenbacher, P. Is there a relationship between the substrate preferences and structural flexibility of cytochromes p450? *Curr. Drug Metab.* **2012**, *13*, 130–142.
- (31) Schleinkofer, K.; Sudarko; Winn, P. J.; Ludemann, S. K.; Wade, R. C. Do mammalian cytochrome P450s show multiple ligand access pathways and ligand channelling? *EMBO Rep.* **2005**, *6*, 584–589.
- (32) Wade, R. C.; Winn, P. J.; Schlichting, I.; Sudarko. A survey of active site access channels in cytochromes P450. *J. Inorg. Biochem.* **2004**, *98*, 1175–1182.
- (33) Ludemann, S. K.; Lounnas, V.; Wade, R. C. How do substrates enter and products exit the buried active site of cytochrome P450cam? 2. Steered molecular dynamics and adiabatic mapping of substrate pathways. *J. Mol. Biol.* **2000**, *303*, 813–830.
- (34) Winn, P. J.; Ludemann, S. K.; Gauges, R.; Lounnas, V.; Wade, R. C. Comparison of the dynamics of substrate access channels in three cytochrome P450s reveals different opening mechanisms and a novel functional role for a buried arginine. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 5361–5366.
- (35) Yaffe, E.; Fishelovitch, D.; Wolfson, H. J.; Halperin, D.; Nussinov, R. MolAxis: efficient and accurate identification of channels in macromolecules. *Proteins: Struct., Funct., Bioinf.* **2008**, *73*, 72–86.
- (36) Krishnamoorthy, N.; Gajendrara, P.; Thangapandian, S.; Lee, Y.; Lee, K. W. Probing possible egress channels for multiple ligands in human CYP3A4: a molecular modeling study. *J. Mol. Model.* **2010**, *16*, 607–614.
- (37) Li, W.; Shen, J.; Liu, G.; Tang, Y.; Hoshino, T. Exploring coumarin egress channels in human cytochrome p450 2a6 by random acceleration and steered molecular dynamics simulations. *Proteins: Struct., Funct., Bioinf.* **2011**, *79*, 271–281.
- (38) Skopalik, J.; Anzenbacher, P.; Otyepka, M. Flexibility of Human Cytochromes P450: Molecular Dynamics Reveals Differences between CYPs 3A4, 2C9, and 2A6, which Correlate with Their Substrate Preferences. *J. Phys. Chem. B* **2008**, *112*, 8165–8173.
- (39) Cojocaru, V.; Winn, P. J.; Wade, R. C. Multiple, ligand-dependent routes from the active site of cytochrome P450 2C9. *Curr. Drug Metab.* **2011**, *13*, 143–154.
- (40) Branduardi, D.; Gervasio, F. L.; Cavalli, A.; Recanatini, M.; Parrinello, M. The role of the peripheral anionic site and cation-pi interactions in the ligand penetration of the human AChE gorge. *J. Am. Chem. Soc.* **2005**, *127*, 9147–9155.
- (41) Shen, J.; Li, W.; Liu, G.; Tang, Y.; Jiang, H. Computational insights into the mechanism of ligand unbinding and selectivity of estrogen receptors. *J. Phys. Chem. B* **2009**, *113*, 10436–10444.
- (42) Lindahl, E.; Sansom, M. S. P. Membrane proteins: molecular dynamics simulations. *Curr. Opin. Struct. Biol.* **2008**, *18*, 425–431.
- (43) Tozzini, V. Multiscale modeling of proteins. *Acc. Chem. Res.* **2009**, *43*, 220–230.
- (44) Kandt, C.; Ash, W. L.; Tieleman, P. D. Setting up and running molecular dynamics simulations of membrane proteins. *Methods* **2007**, *41*, 475–488.
- (45) Amarneh, B.; Corbin, C. J.; Peterson, J. A.; Simpson, E. R.; Graham-Lorence, S. Functional domains of human aromatase cytochrome P450 characterized by linear alignment and site-directed mutagenesis. *Mol. Endocrinol.* **1993**, *7*, 1617–1624.
- (46) Auvray, P.; Nativelle, C.; Bureau, R.; Dallemagne, P.; Seralini, G. E.; Sourdaine, P. Study of substrate specificity of human aromatase by site directed mutagenesis. *Eur. J. Biochem.* **2002**, *269*, 1393–1405.
- (47) Zhou, D.; Cam, L. L.; Laughton, C. A.; Korzekwa, K. R.; Chen, S. Mutagenesis study at a postulated hydrophobic region near the active site of aromatase cytochrome P450. *J. Biol. Chem.* **1994**, *269*, 19501–19508.
- (48) Zhou, D. J.; Pompon, D.; Chen, S. A. Structure-function studies of human aromatase by site-directed mutagenesis: kinetic properties of mutants Pro-308---Phe, Tyr-361---Phe, Tyr-361---Leu, and Phe-406---Arg. *Proc. Natl. Acad. Sci. U.S.A.* **1991**, *88*, 410–414.

- (49) Zhou, D. J.; Korzekwa, K. R.; Poulos, T.; Chen, S. A. A site-directed mutagenesis study of human placental aromatase. *J. Biol. Chem.* **1992**, *267*, 762–768.
- (50) Conley, A.; Mapes, S.; Corbin, C. J.; Greger, D.; Graham, S. Structural determinants of aromatase cytochrome p450 inhibition in substrate recognition site-1. *Mol. Endocrinol.* **2002**, *16*, 1456–1468.
- (51) Kao, Y. C.; Zhou, C.; Sherman, M.; Laughton, C. A.; Chen, S. Molecular basis of the inhibition of human aromatase (estrogen synthetase) by flavone and isoflavone phytoestrogens: A site-directed mutagenesis study. *Environ. Health Perspect.* **1998**, *106*, 85–92.
- (52) Kao, Y. C.; Korzekwa, K. R.; Laughton, C. A.; Chen, S. Evaluation of the mechanism of aromatase cytochrome P450. A site-directed mutagenesis study. *Eur. J. Biochem.* **2001**, *268*, 243–251.
- (53) Kao, Y. C.; Cam, L. L.; Laughton, C. A.; Zhou, D.; Chen, S. Binding characteristics of seven inhibitors of human aromatase: a site-directed mutagenesis study. *Cancer Res.* **1996**, *56*, 3451–3460.
- (54) Simona, F.; Magistrato, A.; Vera, D. M. A.; Garau, G.; Vila, A. J.; Carloni, P. Protonation state and substrate binding to B2 metallo-beta-lactamase CphA from *Aeromonas hydrofila*. *Proteins: Struct., Funct., Bioinf.* **2007**, *69*, 595–605.
- (55) Hong, R.; Magistrato, A.; Carloni, P. Anthrax Lethal Factor Investigated by Molecular Simulations. *J. Chem. Theory Comput.* **2008**, *4*, 1745–1756.
- (56) Case, D. A.; Darden, T. A.; Cheatham, T. E., III; Simmerling, C. L.; Wang, J.; Duke, R. E.; Luo, R.; Walker, R. C.; Zhang, W.; Merz, K. M.; Roberts, B.; Wang, B.; Hayik, S.; Roitberg, A.; Seabra, G.; Kolossaváry, I.; Wong, K. F.; Paesani, F.; Vanicek, J.; Liu, J.; Wu, X.; Brozell, R. 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.; Saguí, C.; Babin, V.; Luchko, T.; Gusrarov, S.; Kovalevko, A.; Kollman, P. A. AMBER11; University of California: San Francisco, CA, 2010.
- (57) Carnevale, V.; Fiorin, G.; Levine, B. G.; Degrado, W. F.; Klein, M. L. Multiple Proton Confinement in the M2 Channel from the Influenza A Virus. *J. Phys. Chem. C* **2010**, *114*, 20856–20863.
- (58) Guex, N.; Peitsch, M. C. SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. *Electrophoresis* **1997**, *18*, 2714–2723.
- (59) Shimozawa, O.; Sakaguchi, M.; Ogawa, H.; Harada, N.; Mihara, K.; Omura, T. Core glycosylation of cytochrome P-450(arom). Evidence for localization of N terminus of microsomal cytochrome P-450 in the lumen. *J. Biol. Chem.* **1993**, *268*, 21399–21402.
- (60) Humphrey, W.; Dalke, A.; Schulten, K. VMD: visual molecular dynamics. *J. Mol. Graphics* **1996**, *14*, 33–38.
- (61) Jojart, B.; Martinek, T. A. Performance of the general amber force field in modeling aqueous POPC membrane bilayers. *J. Comput. Chem.* **2007**, *28*, 2051–2058.
- (62) 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.
- (63) 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.
- (64) Aqvist, J. Ion-water interaction potentials derived from free energy perturbation simulations. *J. Chem. Phys.* **1990**, *94*, 8021–8024.
- (65) Giannoni, D. A. Ph.D. Thesis, University of California Davis, CA, 1984.
- (66) Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. Development and testing of a general amber force field. *J. Comput. Chem.* **2004**, *25*, 1157–1174.
- (67) Bayly, C. I.; Cieplak, P.; Cornell, W. D.; Kollman, P. A. A well-behaved electrostatic potential based method using charge restraints for determining atom-centered charges: the RESP model. *J. Phys. Chem.* **1993**, *97*, 10269–10280.
- (68) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Kollmann, P. A. Application of RESP charges to calculate conformational energies, hydrogen bond energies, and free energies of solvation. *J. Am. Chem. Soc.* **1993**, *115*, 9620–9631.
- (69) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, J. A.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazayev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Laham, A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; Pople, J. A. Gaussian 03, Revision C.02; 2003.
- (70) Sgrignani, J.; Magistrato, A. The structural role of mg(2+) ions in a class I RNA polymerase ribozyme: a molecular simulation study. *J. Phys. Chem. B* **2012**, *116*, 2259–2268.
- (71) Vargiu, A. V.; Magistrato, A. Detecting DNA mismatches with metallo-insertors: a molecular simulation study. *Inorg. Chem.* **2012**, *51*, 2046–2057.
- (72) 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.
- (73) Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kale, L.; Schulten, K. Scalable molecular dynamics with NAMD. *J. Comput. Chem.* **2005**, *26*, 1781–1802.
- (74) Martyna, G.; Tobias, D.; Klein, M. Constant pressure molecular dynamics algorithms. *J. Chem. Phys.* **1994**, *101*, 4177–4189.
- (75) Feller, S.; Zhang, Y.; Pastor, R.; Brooks, B. Constant pressure molecular dynamics simulation: The Langevin piston method. *J. Chem. Phys.* **1995**, *103*, 4613–4621.
- (76) Shao, J.; Tanner, S. W.; Thompson, N.; Cheatham, T. E. Clustering Molecular Dynamics Trajectories: 1. Characterizing the Performance of Different Clustering Algorithms. *J. Chem. Theory Comput.* **2007**, *3*, 2312–2334.
- (77) Petrek, M.; Otyepka, M.; Banas, P.; Kosinova, P.; Koca, J.; Damborsky, J. CAVER: a new tool to explore routes from protein clefts, pockets and cavities. *BMC Bioinf.* **2006**, *7*, 316–316.
- (78) Dijkstra, E. W. A note on two problems in connexion with graphs. *Numerische Math.* **1959**, *1*, 269–271.
- (79) The PyMOL Molecular Graphics System, V. r. p., Schrödinger, LLC: *The PyMOL Molecular Graphics System, Version 1.2r3pre*, Schrödinger, LLC.
- (80) Vashisth, H.; Abrams, C. F. Ligand escape pathways and (un)binding free energy calculations for the hexameric insulin-phenol complex. *Biophys. J.* **2008**, *95*, 4193–4204.
- (81) Colizzi, F.; Perozzo, R.; Scapozza, L.; Recanatini, M.; Cavalli, A. Single-molecule pulling simulations can discern active from inactive enzyme inhibitors. *J. Am. Chem. Soc.* **2010**, *132*, 7361–7371.
- (82) Yang, L.-J.; Zou, J.; Xie, H.-Z.; Li, L.-L.; Wei, Y.-Q.; Yang, S.-Y. Steered molecular dynamics simulations reveal the likelier dissociation pathway of imatinib from its targeting kinases c-Kit and Abl. *PLoS One* **2009**, *4*, e8470.
- (83) Cohen, J.; Arkhipov, A.; Braun, R.; Schulten, K. Imaging the migration pathways for O<sub>2</sub>, CO, NO, and Xe inside myoglobin. *Biophys. J.* **2006**, *91*, 1844–1857.
- (84) Cohen, J.; Schulten, K. O<sub>2</sub> Migration Pathways Are Not Conserved across Proteins of a Similar Fold. *Biophys. J.* **2007**, *93*, 3591–3600.
- (85) Cohen, J.; Olsen, K. W.; Schulten, K. Finding gas migration pathways in proteins using implicit ligand sampling. *Methods Enzymol.* **2008**, *437*, 439–457.
- (86) Forti, F.; Boechi, L.; Estrin, D. A.; Marti, M. A. Comparing and combining implicit ligand sampling with multiple steered molecular

- dynamics to study ligand migration processes in heme proteins. *J. Comput. Chem.* **2011**, *32*, 2219–2231.
- (87) Forti, F.; Boechi, L.; Bikiel, D.; Marti, M. A.; Nardini, M.; Bolognesi, M.; Viappiani, C.; Estrin, D. O.; Luque, F. J. Ligand Migration in Methanosaerina acetivorans Protogobin: Effects of Ligand Binding and Dimeric Assembly. *J. Phys. Chem. B* **2011**, *115*, 13771–13780.
- (88) Bayburt, T. H.; Sligar, S. G. Single-molecule height measurements on microsomal cytochrome P450 in nanometer-scale phospholipid bilayer disks. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 6725–6730.
- (89) Kadohama, N.; Yarborough, C.; Zhou, D.; Chen, S.; Osawa, Y. Kinetic properties of aromatase mutants Pro308Phe, Asp309Asn, and Asp309Ala and their interactions with aromatase inhibitors. *J. Steroid. Biochem. Mol. Biol.* **1992**, *43*, 693–701.
- (90) Di Nardo, G.; Breitner, M.; Sadeghi, S. J.; Nicolai, E.; Mei, G.; Bandino, A.; Ghosh, D.; Panzica, G.; Gilardi, G. A pH-dependent structural rearrangement of the human aromatase active site: implications in catalysis. *FEBS J.* **2010**, *277*, 175–175.
- (91) Arroyo-Manez, P.; Bikiel, D. n. E.; Boechi, L.; Capece, L.; Di Lella, S.; Estrin, D. A.; Martí, M. A.; Moreno, D. M.; Nadra, A. D.; Petruk, A. A. Protein dynamics and ligand migration interplay as studied by computer simulation. *Biochim. Biophys. Acta* **2011**, *1814*, 1054–1064.
- (92) Scorciapino, M. A.; Robertazzi, A.; Casu, M.; Ruggerone, P.; Ceccarelli, M. Heme proteins: the role of solvent in the dynamics of gates and portals. *J. Am. Chem. Soc.* **2010**, *132*, 5156–5163.
- (93) Carrillo, O.; Orozco, M. GRID-MD-A tool for massive simulation of protein channels. *Proteins: Struct., Funct., Bioinf.* **2008**, *70*, 892–899.
- (94) Otyepka, M.; Skopalík, J.; Anzenbacherov, E.; Anzenbacher, P. What common structural features and variations of mammalian P450s are known to date? *Biochim. Biophys. Acta* **2007**, *1770*, 376–389.
- (95) Sgrignani, J.; Bonaccini, C.; Grazioso, G.; Chioccioli, M.; Cavalli, A.; Gratteri, P. Insights into docking and scoring neuronal alpha4beta2 nicotinic receptor agonists using molecular dynamics simulations and QM/MM calculations. *J. Comput. Chem.* **2009**, *30*, 2443–2454.
- (96) Amaro, R. E.; Baron, R.; McCammon, J. A. An improved relaxed complex scheme for receptor flexibility in computer-aided drug design. *J. Comput.-Aided Mol. Des.* **2008**, *22*, 693–705.
- (97) Fishelovitch, D.; Shaik, S.; Wolfson, H. J.; Nussinov, R. Theoretical characterization of substrate access/exit channels in the human cytochrome P450 3A4 enzyme: involvement of phenylalanine residues in the gating mechanism. *J. Phys. Chem. B* **2009**, *113*, 13018–13025.
- (98) Jarzynski, C. Nonequilibrium Equality for Free Energy Differences. *Phys. Rev. Lett.* **1997**, *78*, 2690–2693.
- (99) Wang, Y.; Cohen, J.; Boron, W. F.; Schulter, K.; Tajkhorshid, E. Exploring gas permeability of cellular membranes and membrane channels with molecular dynamics. *J. Struct. Biol.* **2007**, *157*, 534–544.
- (100) Jiang, W.; Ghosh, D. Motion and Flexibility in Human Cytochrome P450 Aromatase. *PLoS One* **2012**, *7* (2), e32565.