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

Angelo Danilo Favia,[1] Andrea Cavalli,[2]* Matteo Masetti,[2] Angelo Carotti,[1] Maurizio Recanatini[2]
[1]*Department of Medicinal Chemistry, University of Bari, Via E. Orabona 4, I-70124 Bari, Italy*
[2]*Department of Pharmaceutical Sciences, University of Bologna, Via Belmeloro 6, I-40126 Bologna, Italy*

ABSTRACT **Mammalian cytochromes P450 (CYP) are enzymes of great biological and pharmacotoxicological relevance. Due to their membrane-bound nature, the structural characterization of these proteins is extremely difficult, and therefore computational techniques, such as comparative modeling, may help obtaining reliable structures of members of this family. An important feature of CYP is the presence of an iron-containing porphyrin group at the enzyme active site. This calls for quantum chemical calculations to derive charges and parameters suitable for classical force field-based investigations of this proteins family. In this report, we first carried out density functional theory (DFT) computations to derive suitable charges for the Fe^{2+}-containing heme group of P450 enzymes. Then, by means of the homology modeling technique, and taking advantage of the recently published crystal structure of the human CYP2C9, we built a new model of the human aromatase (CYP19) enzyme. Furthermore, to study the thermal stability of the new model as well as to test the suitability of the new DFT-based heme parameters, molecular dynamics (MD) simulations were carried out on both CYP2C9 and CYP19. Finally, the last few ns of aromatase MD trajectories were investigated following the essential dynamics protocol that allowed the detection of some correlated motions among some protein domains. Proteins 2006;62:1074–1087.**

Key words: mammalian cytochromes P450 (CYP); DFT parameterization; heme-cysteinato cytochromes

INTRODUCTION

Mammalian cytochromes P450 (CYP) are involved in the biosynthesis of several important hormones, mainly of a steroidal nature and in the metabolism of more than 90% of the drugs in current clinical use.[1] In addition, they participate in both toxification and detoxification processes of many xenobiotics. The primary sequences of CYP are extremely diverse with levels of identity as low as 16%,[2] whereas the overall folding, comprising 12 helices and loops denoted as A-L (Fig. 1), has remained unaltered through the evolution. Furthermore, CYP share a common prosthetic group constituted by an iron-containing protoporphyrin IX, which is linked to the protein through both a covalent bond between the metal and the sulphur atom of a proximal cysteine, and a network of hydrogen bonds among its propionate groups and the side chains of positively charged amino acids. The heme group is located between helices I and L, while helices B, C, F, and G contribute to the substrate recognition.

Although the key steps of the catalytic mechanism seem to be similar among all isoforms, CYP show a surprising degree of selectivity versus their own specific substrates.[3] The cytochrome P450 19 (CYP19; EC 1.14.14.1), commonly called aromatase (AR), metabolizes a wide variety of important substrates in many species of bacteria, plants, and animals. In the human body, AR catalyzes the conversion of androgens into estrogens, through the aromatization of the A ring of substrates like testosterone and androstenedione.[4] In the anti-cancer therapy, AR is an important pharmacological target because of its critical role in the progression of post-menopausal breast cancer.[5] Actually, the reduction of the levels of circulating estrogens in women with the disease has been demonstrated to be clinically effective,[6,7] and AR inhibitors (ARIs) such as anastrozole, vorozole, letrozole, and fadrozole are widely used, even as the first-line drugs in the therapy of breast cancer.[8,9] However, some important side effects associated with the prolonged clinical use of ARIs[10,11] called for the search of new, potent, more selective, and less toxic CYP19 inhibitors.

In the era of the so-called "molecular targeted therapy," there is a strong need to develop reliable and validated 3D models of P450 enzymes of therapeutic interest, enabling the application of the structure-based design (SBD) of potent and selective inhibitors. Unfortunately, the SBD of P450 inhibitors has been strongly hampered by the ab-

Grant sponsor: MIUR (Ministero dell'Istruzione dell'Università e della Ricerca, Rome, Italy).

*Correspondence to: Andrea Cavalli, Department of Pharmaceutical Sciences, University of Bologna, Via Belmeloro 6, I-40126 Bologna, Italy. E-mail: andrea.cavalli@unibo.it

Received 13 May 2005; Accepted 5 October 2005

Published online 4 January 2006 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/prot.20829

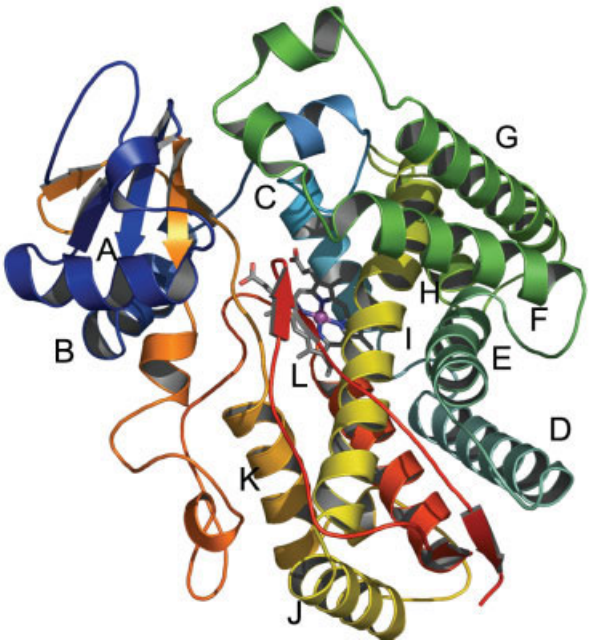

Fig. 1. X-ray 3D structure of the cytochrome P450 2C9. Helices from A to L are highlighted to help the interpretation. The heme group is depicted as stick model. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

sence of sound and consistent experimental and/or theoretical 3D models of many P450 enzymes of therapeutic interest. In fact, the structural characterization of human enzymes has been extremely difficult, mostly because of their membrane-bound nature. In 1985, the first 3D bacterial structure ($P450_{cam}$) pertaining to the family of P450 was reported by Poulos and co-workers.[12] At the end of the past century, no more than ten X-ray structures of four CYP ($P450_{BM-3}$,[13] $P450_{terp}$,[14] $P450_{eryF}$,[15] $P450^{15}_{nor}$) were solved. During the last three years, an actual breakthrough[16] was observed in the field, as structures of some rabbit and human P450's were solved, making available, for the first time, 3D coordinates of mammalian enzymes.[17–19]

Another important drawback in the SBD of P450 inhibitors is the lack of an appropriate parameterization of both bonding and non-bonding interactions of the iron-carrying protoporphirin IX that is fundamental to perform accurate docking and virtual screening simulations. Therefore, the construction of a reliable 3D model of CYP, and a quantum chemical heme parameterization, might provide the right tool for a rational approach to the design of new P450 inhibitors.

In the light of the above considerations, we herein present the construction and the validation of a 3D model of human aromatase that might be useful in guiding the SBD of hARIs with improved characteristics, as well as for studying quantum mechanically the hAR enzymatic reaction mechanism. The 3D features of the model were obtained by a comparative procedure using the recently solved crystal structure of the human cytochrome P450 2C9[18] as a template. To properly describe the electrostatic properties of the heme-cysteinato complex, and to parameterize the iron-containing protoporphyrin IX, which are both critical for reliable molecular dynamics (MD), docking, and virtual screening simulations of P450, density functional theory (DFT) calculations were also carried out. Then, the hAR 3D model and the new cysteinato-heme parameters were tested by performing both docking studies, and 10 ns of molecular dynamics simulations, which were conducted in parallel on both structures (model and template). Finally, the MD trajectories were further investigated according to the essential dynamics protocol,[20] to detect possible correlated motions of hAR domains.

METHODS

DFT Parameterization of the Heme-Thiolate Complex

The parameterization of the heme–thiolate complex was performed in two steps: (1) full geometry optimization of the complex at the DFT level of theory, and (2) calculation of its charges using the restrained electrostatic potential (RESP) fitting procedure.[21] To deal with the iron atom, the B3LYP hybrid functional,[22] which has been proven to accurately describe transition metal-carrying systems, was chosen.[23] For both the optimization and the electrostatic potential (ESP) calculation, the following locally dense basis set was used: 6-31G* was assigned to the porphyrin as well as to the axial ligand, except those atoms directly involved in the iron coordination. In fact, to better describe charge transfer phenomena, the basis set on these atoms was augmented to 6-31+G*. Conversely, to limit the amount of computational time, the protoporphyrin IX side chains, as well as the acetyl (ACE) and N-methyl (NME) capping residues (see below), were treated with the smaller 3-21G*. Moreover, the ten innermost electrons of the iron ion were replaced by the Stuttgart-Dresden (SDD) energy consistent effective core potential,[24] which deals with the Dirac-Fock relativistic correction, according to a well-tested approach,[25] while the valence electrons were explicitly treated by means of the consistent SDD basis set (of the quality of triple-zeta). All of the first principles calculations were carried out using the Gaussian03 package.[26]

The starting geometry of the prosthetic group was obtained by the crystallographic template used in the comparative modeling procedure (PDB code 1OG5, see below).[18] In this crystal, the iron approximately lies on the porphyrin plane (see Table I), suggesting an octahedral low spin coordination complex, although the sixth axial ligand seems to be missing. In this respect, the X-ray crystallography analysis was unable to univocally determine both the correct iron spin and redox state.[18] Such features are needed to properly describe the system using quantum chemical calculations. To our knowledge, non-covalent hARIs bind the heme group in the ferrous and low spin state (S = 0). A suitable hexa-coordinated system was therefore built, where the sixth axial ligand was retrieved from another crystallographic complex (PDB code 1SUO).[27] In details, the imidazole moiety of the ligand 4-(4-chlorophenyl)imidazole of 1SUO was merged into 1OG5 after a rigid superposition of the two porphyrin rings

TABLE I. A Comparison of the Main Structural Parameters of the Heme in the Crystal Structure, the Intermediate Geometry, and the Final Optimized System†

	Crystal	Intermediate	Final
Distances			
Fe-S	2.304	2.415	2.414
Fe-N_0	2.145	2.113	2.107
Fe-N_{21}	2.014	2.032	2.042
Fe-N_{22}	2.105	2.024	2.037
Fe-N_{23}	2.002	2.020	2.019
Fe-N_{24}	2.100	2.022	2.016
Fe-p-plane	0.013	0.043	0.077
RMSD porphyrin atoms-p-plane	0.031	0.027	0.034
2′-p-plane	0.030	0.050	0.083
3′-p-plane	−0.056	−0.017	−0.041
7′-p-plane	0.072	0.017	−0.026
8′-p-plane	0.012	0.016	0.130
12′-p-plane	−0.036	−0.071	−0.011
13′-p-plane	−0.047	−0.028	−0.003
17′-p-plane	0.046	0.034	−0.012
18′-p-plane	0.090	0.010	−0.174
Angles			
S-Fe-N_0	177.68	174.41	173.95
S-Fe-N_{21}	90.98	90.45	89.51
S-Fe-N_{22}	85.30	85.03	85.07
S-Fe-N_{23}	89.52	92.25	93.69
S-Fe-N_{24}	94.34	96.29	97.04
N_0-Fe-N_{21}	88.94	88.50	88.20
N_0-Fe-N_{22}	97.02	89.49	89.33
N_0-Fe-N_{23}	90.45	88.81	88.58
N_0-Fe-N_{24}	83.34	89.20	88.58
S-Fe-n (δ1)	4.74	5.97	7.15
N_{IM}-Fe-n (δ2)	7.04	0.39	1.10
Dihedrals			
2-3-3′-3″	−14.21	−25.42	−23.04
7-8-8′-8″	138.18	148.88	150.57
4-3-2-2′	179.41	176.87	176.59
1-2-3-3′	−179.78	−178.38	−179.08
9-8-7-7′	−179.98	179.16	178.861
6-7-8-8′	179.68	−177.64	−177.36
14-13-12-12′	−179.98	179.56	179.93
11-12-13-13′	−179.44	−179.21	−176.37
19-18-17-17′	179.91	−179.88	177.18
16-17-18-18′	−179.60	−179.75	177.29
Φ	57.56	45.25	45.02
6-22-24-19 (τa)	−0.76	−0.21	1.72
1-21-23-14 (τb)	2.16	0.72	3.19

†The data refer to heme complexed with the imidazole as the sixth coordinating group. Distances are given in Å, angles in degrees. The porphyrin mean plane is referred to as p-plane.

(RMSD equal to 0.132 Å). The optimization was initially performed on a truncated system (intermediate system of Table I), in which the protoporphyrin IX propionates were replaced by two methyl groups, as they do not affect the electronic structure of similar complexes,[28–30] while the cysteine was treated as a simple ethane-thiolate anion. During the geometry optimization, the external carbons of each truncated residue were kept fixed to simulate the protein frame. The model system was optimized till reaching the default Gaussian03 convergence criterion. The model system extent was then increased. First, the truncated residues were restored to their original dimension. Second, to properly describe the electrostatic protein environment, the N and C termini of the coordinating cysteine were capped with NME and ACE residues, respectively. Hence, the system was again geometrically optimized constraining the external atoms, and using the previously reported locally dense basis set.

The RESP charges were then calculated on the geometrically minimized heme–thiolate complex after removing of the imidazole ring (total formal charge = −3). Actually, since the aim of the present heme parameterization was to obtain point charges suitable to both MD simulations and docking studies, a penta- instead of hexa-coordinated complex was eventually employed to compute the quantum ESP. The RESP charges were obtained by a two-stage fitting procedure,[21,31] fitting first the polar areas by using weak hyperbolic restraints (0.0005 a.u.), and then fitting the remaining areas imposing equivalencies and by using a stronger hyperbolic restraint (0.001 a.u.). In each step, the charges of the standard residues ACE and NME were constrained to their AMBER force field value.[32] The electrostatic potential used as an input by the RESP program was sampled adapting the Merz-Singh-Kollman scheme,[33,34] namely using 9 concentric layers at the default level of spacing, a surface density of 6 points/Å^2, and supplying the covalent radius of 1.25 Å for the iron ion. Actually, the increase of the sampling resolution should result in a better quality of the charge fitting.[35]

To dock the substrate molecule testosterone, an oxygen molecule coordinating the sixth position of the prosthetic group was added. This required a further DFT-based parameterization to provide a new set of partial atomic charges. In details, a geometry optimization of the heme-oxygen complex was performed as above reported, keeping the spin multiplicity equal to 3. This required the use of the unrestricted exchange and correlation functional UB3LYP (the 6-31+G* basis set was assigned to the oxygen atoms). Then, RESP charges were derived adopting the previously reported computational criteria.

Homology Modeling of hAR

The 3D model of hAR was produced using as references the spatial features extracted from the crystallographic data of a human cytochrome P450 available at PDB.[36] Recently, several X-ray structures of human cytochromes were determined both alone and complexed with small ligands. The template selection was done on the basis of sequence similarity, residues completeness, and crystal resolution. Among the available structures, the cytochrome P450 2C9 with bound warfarin (PDB code 1OG5) resolved at 2.55 Å[18] was eventually chosen on the basis of the above-mentioned criteria.

The hAR amino acids sequence was downloaded from the SWISS-PROT database (code P11511), and aligned onto the template sequence by means of the LALIGN server.[37] The first 52 residues, which are part of the transmembrane domain of the enzyme, were not taken into account. Several scoring matrixes were applied to

sample the quality of the overall alignment looking to local matching, and preserving the correspondences of very conserved residues among CYP.[38] However, the percentage of identity criterion did not lead to a unique hypothesis of alignment. To further assess the quality of each binary alignment, the primary structures of both P450 2C9 and hAR were analyzed by PSIPRED.[39] This web server allows one to predict secondary structures of amino acid sequences. The predictions were compared to the experimentally reported secondary structures of CYP2C9 in the PDB file 1OG5.

Finally, several models were generated by means of Modeller 6v2 software.[40] The outcomes were ranked on the basis of the internal scoring function of the program, and validated by means of the PROCHECK procedure.[41]

Molecular Dynamics Simulations

MD simulations (10 ns) on both CYP19 and CYP2C9 were carried out, and the root mean square deviation (RMSD) of parts and whole of both enzymes were monitored vs. the simulations time. Either structure was immersed in a box, whose edges were located 10 Å apart from the closest atom of the protein, and contained ~ 16,800 water molecules. Then, 1 Cl^- (model) and 4 Na^+ (template) counterions were added to the solvent bulk of the protein-water complexes to maintain neutrality on the systems. Prior to starting the MD simulations, a minimization of the entire ensembles was performed setting a convergence criterion on the gradient of 0.001 kcal mol^{-1} $Å^{-1}$. Then, water shells and counterions were equilibrated for 75 ps at 300 K. This was followed by 10 ns of MD simulations in the NPT ensemble (constant temperature and pressure) on both enzymes. The parm99 version[42] of the all atom Amber force field[32] was used for the protein and the counterions, whereas the TIP3P model[43] was employed to explicitly represent water molecules. For the iron containing protoporphyrin IX, force field parameters were adopted from previously published works[44,45] with proper modifications according to the present DFT calculations. The systems were simulated in periodic boundary conditions. Van der Waals and short-range electrostatic interactions were estimated within a 10 Å cutoff, whereas the long-range electrostatic interactions were assessed by using the particle mesh Ewald (PME) method,[46] with ~1 Å charge grid spacing interpolated by fourth-order B-spline, and by setting the direct sum tolerance to 10^{-5}. Bonds involving hydrogens were constrained by using the SHAKE algorithm[47] with a relative geometric tolerance for coordinate resetting of 0.00001 Å. To maintain directionality between gamma sulphur and the iron ion in the heme-thiolate complex, distance constraints were applied in analogy to previous works dealing with metal coordination complexes.[48,49] Berendsen's coupling algorithms were employed to maintain constant temperature and pressure[50] with the same scaling factor for both solvent and solutes and with the time constant for heat bath coupling maintained at 1.5 ps. The pressure for the isothermal-isobaric ensemble was regulated by using a pressure relaxation time of 1 ps in the Berendsen's algorithm. The simulations of the solvated protein models were performed using constant pressure of 1 atm and constant temperature of 300 K. A time step of 1.5 fs was used in all simulations, which were carried out with the AMBER 8.0 program suite.[51]

In the last 3.5 ns of the simulations, the hAR model showed a good thermal stability, and the trajectories within this simulation interval were further studied applying the essential dynamic protocol.[20] Essential dynamics allows one to separate harmonic atomic displacements, which are supposed to be poorly relevant for the protein functionality, from the anharmonic ones, whose dynamic behavior could in principle explain the main dynamic features characteristic of the studied protein. In this way, essential dynamics studies permit a better investigation of relationships among motions of structural protein domains, which could be of crucial importance for the enzyme characterization. The covariance matrix of the protein's Cα atoms displacement was analyzed by means of the ptraj module of AMBER 8. The essential movements were visualized by means of a plugin[52] of the VMD suite.[53] Finally, a snapshot from the last 3.5 ns of simulations was further minimized, submitted to the PROCHECK[41] validation procedure, and compared to the starting model.

All calculations were performed on a Linux cluster employing an openMosix® architecture.

RESULTS

DFT Parameterization of Protoporphyrin IX

The heme b group consists of an iron-protoporphyrin IX complex, where the four pyrrolic nitrogen atoms coordinate the central metallic ion (Fig. 2). Due to the wide aromatic π-electron delocalization, the isolated porphyrin systems show a planar conformation, while in the protein environment relevant ruffling phenomena of the group have been observed.[54,55]

The geometry optimization of the crystal structure was carried out with an additional axial ligand, namely an imidazole, to force the system to relax in a proper geometrical low-spin hexa-coordinated state, to closely resemble the crystallographic geometry of the cofactor. It turned out that the RMSD between the DFT optimized hexa-coordinated complex and the crystallographic one was as low as 0.128 Å. The main geometric features of the optimized system are summarized in Table I. It should be noted that the orientation of the sixth ligand (i.e., the added imidazole ring) was displaced during the optimization. In particular, the dihedral angle ϕ measured between the ligand plane and the plane orthogonal to the porphyrin and containing the $Fe\text{-}N_{24}$ bond ranged from an initial value of 57.56 to 45.02 degrees after the minimization [for the atom numbering see Fig. 2(A)]. This means that the projection of the imidazole plane on the porphyrin one reached the bisector of the N-Fe-N angle, likely as a consequence of a decrease of steric hindrance between the imidazole and the porphyrin. Finally, the hexa-coordinated Fe^{2+} ion kept its coplanarity with an overall height above plane of 0.077 Å. This behavior is in good agreement with

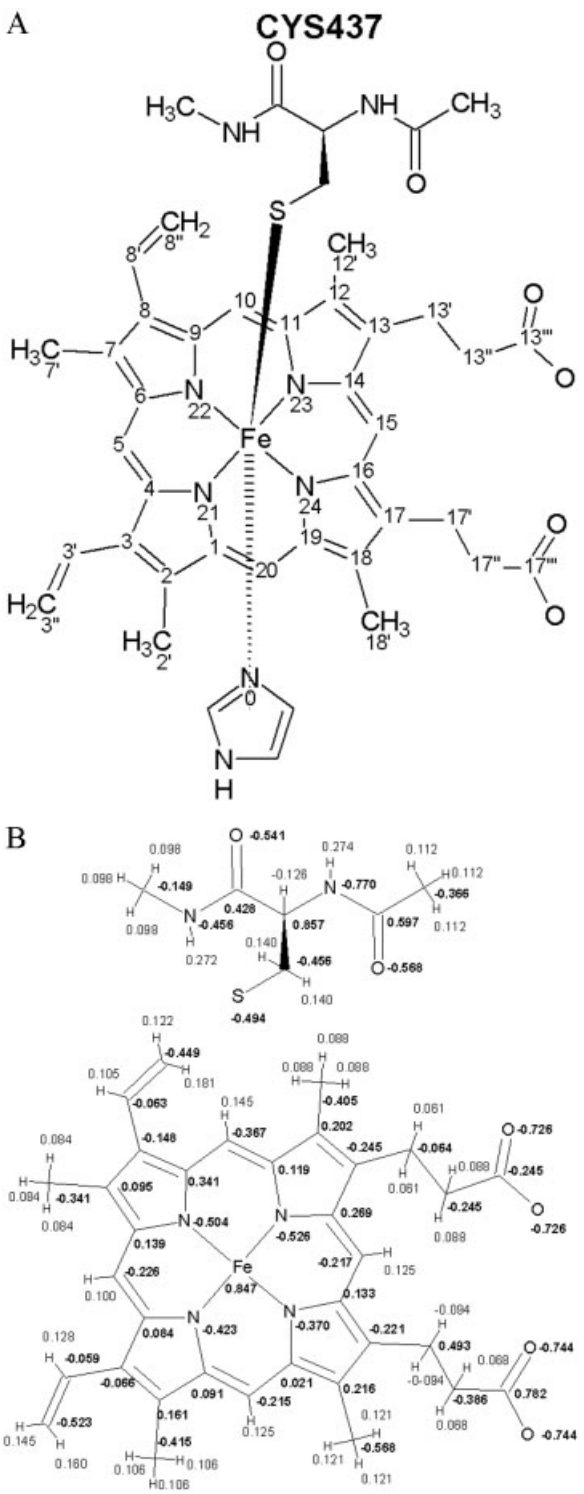

Fig. 2. **A:** Schematic representation of the heme model system submitted to DFT calculations. The atoms numbering is according to the IUPAC convention. **B:** RESP charges adopted in the docking of the inhibitors and MD simulations.

earlier reported experimental data about the displacement of the iron atom in hexa-coordinated low-spin complexes.[56]

It is well known that potential-derived charges show a strong dependence on the chosen sampling method.[35] In the Merz-Singh-Kollman scheme, the electrostatic potential is calculated on points in the Cartesian space located on sets of layers around the molecule. These layers are an overlay of shells obtained scaling the van der Waals radius by a previously set factor. This approach avoids sampling in close proximity of the nuclei, where the electrostatic potential rapidly changes relative to small displacements, thus worsening the fitting quality.[33,34] The default scheme employs 4 layers ranging from 1.4 to 2.0 times the van der Waals radii, and a surface density of 6 points/Å^2. However, a rather higher resolution would be needed to obtain better results relative to the reproducibility of multipole moments.[35] Furthermore, the fitting procedure quality strongly depends on the chosen van der Waals radii. Sigfridsson and Ryde have already discussed this issue, particularly when dealing with transition metals.[35] Actually, for such elements the van der Waals radius is not a priori known, and it has to be defined by the user. Regrettably, the derived charges strongly depend on the metal radius employed in the ESP calculation,[35] and, usually, the iron radius issue is not explicitly accounted for in RESP model charges of heme.[45,57]

To investigate how the charge-fitting procedure depended on the metal radius size, a systematic approach was undertaken. The ESP was evaluated modifying the iron radius in the range 0.5–2.5 Å with a step size of 0.05 Å. All calculations were performed at the highest possible resolution level in the system bearing the greatest number of sampling points, namely the one having an iron radius of 2.5 Å. The trend of the RESP charges for the metal, and for its five ligands (the four porphyrin nitrogen atoms and the sulphur), is shown in Figure 3. Eventually, we adopted the empirical iron covalent radius (1.25 Å), as reported in the WebElements server,[58] to evaluate the DFT-based ESP, as this provided reliable charges for docking studies and MD simulations of binary protein-inhibitor complexes. The final set of RESP charges employed in the present study is reported in Figure 2(B).

Homology Model of hAR

The 3D model of hAR was generated using the cytochrome P450 2C9 with bound warfarin[18] as the main template. All experimental mammalian structures were initially taken into account as potential templates, mainly focusing on human data. Multiple alignments were carried out to characterize the structurally conserved regions among the mammalian members of the P450 family. Such outcomes were further exploited as hints to guide the binary alignment with P450 2C9 (Fig. 4). The alignment was then manually refined, according to both the multiple alignment outcomes and the secondary structure predictions. The secondary structure predictions obtained for the template sequence and compared to the actual secondary structures as reported in the template PDB file 1OG5 are also shown in Figure 4. A good correlation was identified between the experimental and the predicted motifs, leading us to exclude other alignment hypotheses, whose secondary structure correspondences were

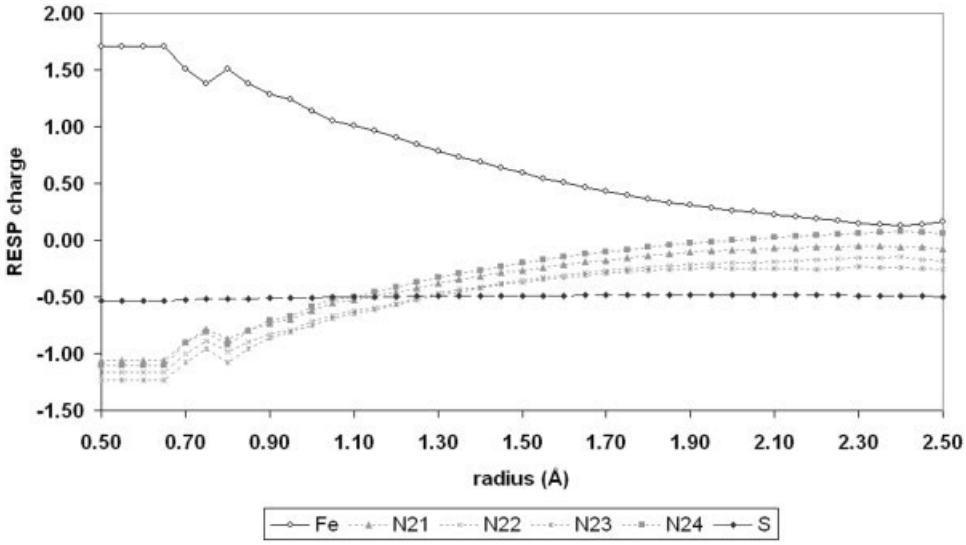

Fig. 3. Dependence of the RESP charges on the iron radius in the penta-coordinated systems. Only the point charges on the atoms directly involved in the iron coordination are reported. The atom numbering is that of Figure 2(A).

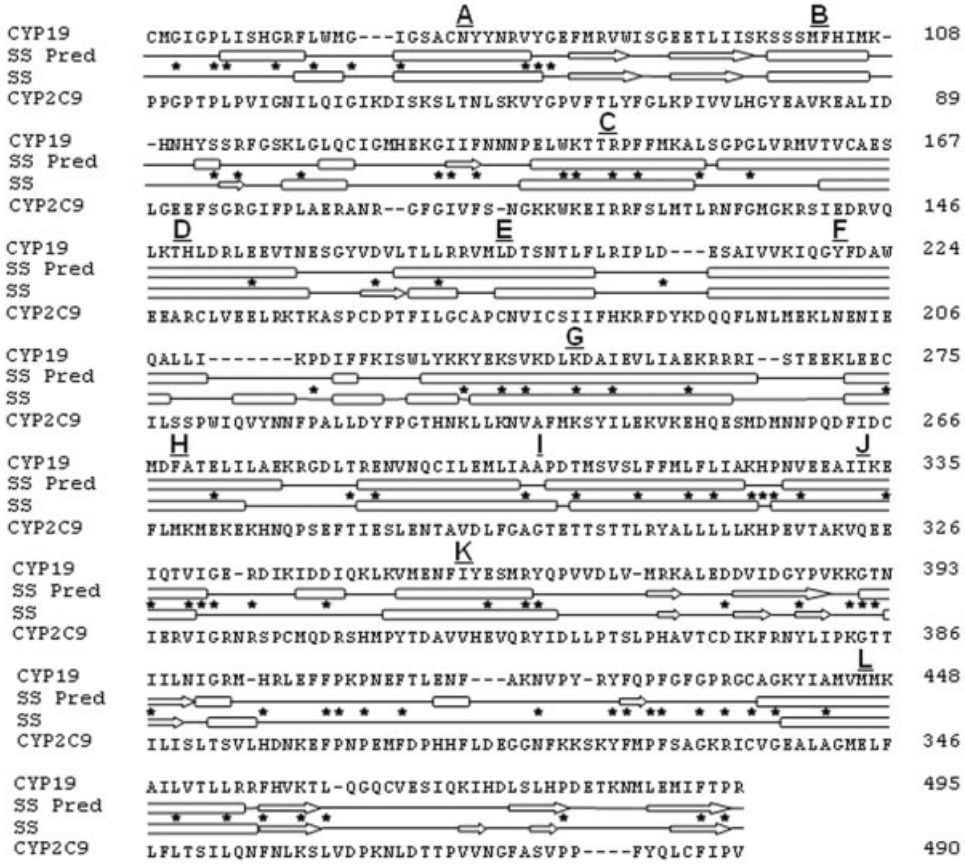

Fig. 4. Pair-wise sequence alignment between human aromatase (CYP19) and CYP2C9 sequences. Actual and predicted secondary structure features are reported for the template and the model, respectively. Helices from A to L are highlighted. The identical residues between the sequences are labeled with asterisks.

not assured. In the final alignment (Fig. 4), the overall sequence identity was 20%. However, in the core of the protein (i.e., taking into account residues in a sphere of 4 radius Å from the heme), where the highly conserved residues were located, the percentage of identical amino acids turned out to be 54%.

Fig. 5. 3D arrangement of the model of the human aromatase enzyme. Helices I, L, F, and G are highlighted to help the interpretation. The heme group is depicted as stick model. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

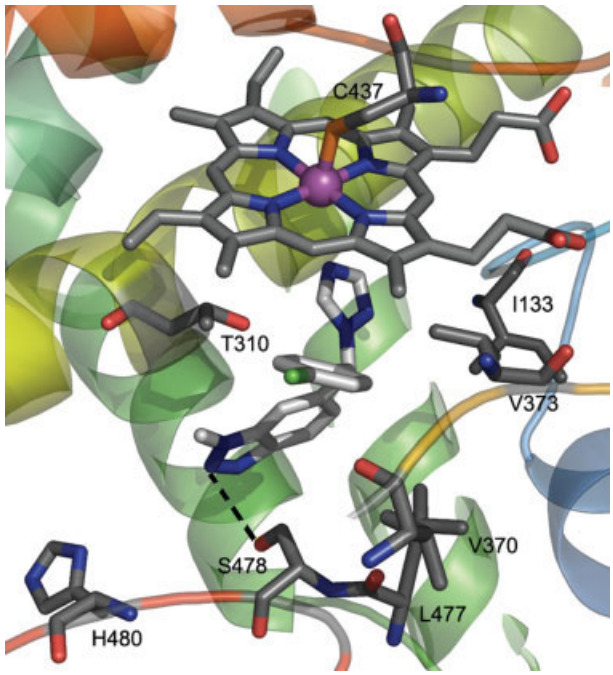

Fig. 6. Docking of the nonsteroidal inhibitor vorozole at the 1TQA active site. Residues I133, T310, V370, V373, C437, V477, S478, H480, and heme prosthetic group are depicted as stick model, colored according to the atom-code (C atoms in gray). Vorozole is depicted as stick model, colored according to the atom-code (C atoms in white). The dashed line is drawn between the atoms likely involved in an H-bond interaction. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

A good quality model, selected on the basis of the internal Modeller scoring function values and the PROCHECK outcomes (see below), was validated by means of the ADIT! server,[59] and submitted to the PDB in the theoretical model section associated with the code 1TQA. The quality of the model was assessed by the PROCHECK validation. The overall G-factor was −0.21, as a consequence of a low percentage (2.5%) of residues in disallowed and generously allowed regions of the Ramachandran plot. Ten residues whose phi–psi angle values were outside the allowed regions of the plot belonged to peripheral loops of the protein, whose geometry was further relaxed by means of MD simulations, thus improving its geometrical quality. The overall 3D arrangement of the hAR enzyme (Fig. 5) is characteristic of this family of CYP enzymes that share the common orthogonal bundle architecture. The core of the protein can be approximately located around the heme group, which is comprised between the helices I and L. The I helix, which goes throughout the protein's width, presents a distortion approximately in its center, near the highly conserved T310 residue, which seems to take part in the catalytic mechanism.[60] Seventy percent of the protein is assembled as α-helices, while 20% of its residues are stabilized as β-sheets. The main difference between the theoretical model and the template was due to a different length of the F–G loop, which is a peculiar feature of the template, and within the P450 family seems to be involved in the substrate recognition.

1TQA model was then exploited for a docking study of the nonsteroidal drug vorozole. The simulations were carried out by means of the AutoDock 3.0.5 suite.[61,62] Ad hoc parameters were added to properly account for the imidazole:N-heme:Fe coordination interaction. One hundred runs were carried out, and a minimum energy conformation binding is reported in Figure 6. Besides the primary interaction, namely the coordination bond between the imidazole of vorozole and the heme group of hAR, the ligand interacted directly, through an H-bond, also with S478 and lied into a hydrophobic binding cleft defined by V370, V373, and L477. Some of these residues have been previously demonstrated to be part of the hAR active site by means of mutagenesis experiments.[4,63]

Molecular Dynamics Simulations

MD simulations were performed with the aim of both validating the results of our DFT-based parameterization and assessing the quality of the proposed theoretical model. The analyses of the RMSD values vs. simulation time for both structures are shown in Figure 7. RMSD values were calculated as the difference between the position of each Cα atom in the starting structure and in every sampled conformation. The template RMSD stabilized around 2.5 Å after a couple of ns, whereas the hAR homology model required as many as 6.5 ns to reach RMSD stability around 4 Å. A snapshot taken from the last 3.5 ns of simulations was further minimized, and finally submitted to the PROCHECK analysis (the 3D model of the minimized structure is available as pdb file in the Supporting Information). The final model had an overall G-factor of −0.15 due to a very low percentage (2%) of residues in the generously allowed and disallowed regions of the Ramachandran plot. A comparison between

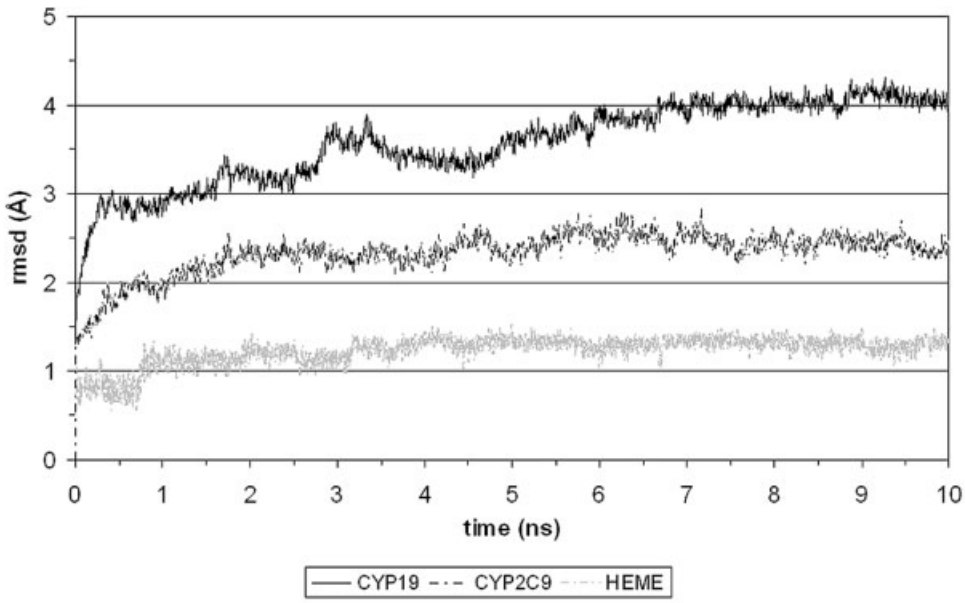

Fig. 7. RMSD of the aromatase (CYP19) and template (CYP2C9) Cα atoms, and of the heme nonhydrogen atoms of CYP19 plotted as a function of the simulation time.

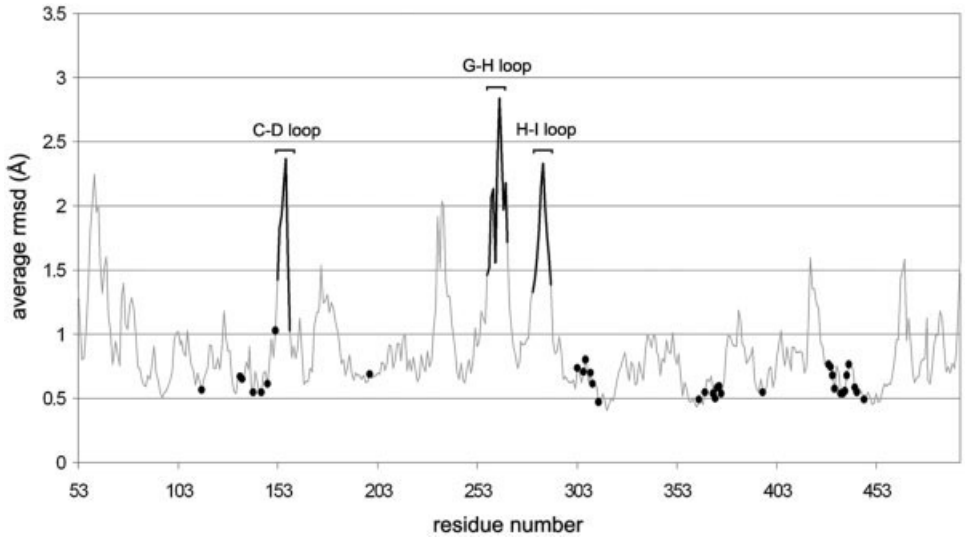

Fig. 8. Mean RMSD calculated per residue within the simulation interval 2.5–3.5 ns. The amino acids 5 Å far from the docked inhibitor and the loops labeled in the picture are highlighted as black dots and bold lines, respectively.

the minimized snapshot and the starting hAR model 1TQA revealed an RMSD value of 4.2 Å.

Within the simulations interval of 2.5–3.5 ns, the RMSD of the hAR Cα atoms rapidly increased and decreased, likely as a consequence of a conformational rearrangement of some regions of the models. This prompted us to further investigate the dynamical behavior of the system within such an interval, aimed at likely detecting protein conformational movements. Therefore, the mean RMSD values of the Cα atoms within 2.5–3.5 ns were monitored. In Figure 8, the atomic fluctuations per residue are plotted vs. the residues number. It can be seen that external loops with a low degree of sequence identity with the template were the most flexible regions of the hAR model. In particular, the segments comprising residues connecting helices C with D, G with H, and H with I (bold lines in Fig. 8), whose sequence identity was as low as 10% (see Fig. 4), were found to be the most flexible regions of the protein. In contrast, the heme group and the protein core, whose sequence identity was about 50% showed very low RMSD values throughout the MD simulations (black dots in Fig. 8). In Figure 7, the RMSD values of the heme vs. the simulation time are also reported. The position of the prosthetic group was stabilized during the dynamics by the coordination bond between the iron atom and the gamma sulphur atom of C437, and by a network of

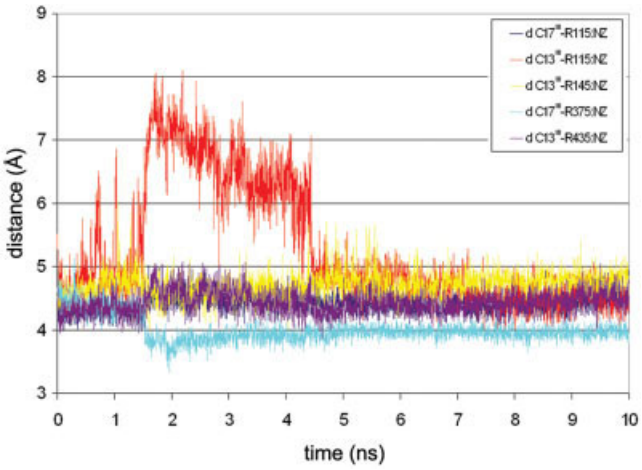

Fig. 9. Distances between the CZ of the arginine residues (R115, R145, R375, and R435) and the closest CG ($C13^{III}$ and $C17^{III}$) [see Fig. 2(A)] of the propionate group plotted as a function of the simulation time. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

charge-reinforced H-bond interactions among heme propionate groups and four arginine residues of the enzyme (R115, R145, R375, and R435). In particular, R115 played a central role in the heme stabilization as it interacted with both propionates during the MD runs. However, this property was clearly evident in the stable simulation interval (i.e., after 6 ns). Actually, as shown in Figure 9, where the distances between the R115:CZ and the propionate:CGA and :CGD vs. simulation time are reported, the interaction between R115 and one propionate was kept all simulation long, while the other was lost during the 1.5–4.5-ns simulation interval. After 4.5 ns, such an interaction was again formed and kept until the end of the simulation. The above results clearly show that more than 6 ns were needed to allow the hAR model to completely relax and to reach a stable conformation.

Another important issue to be addressed concerned the heme group distortion during the MD simulations. Actually, while the isolated heme was previously shown to be planar, a slight distortion of the porphyrin plane could be required to modulate the biological activity of hemeproteins.[54,55] Here, the heme deformation was monitored sampling two opposite dihedral angles [τa and τb, see Table I and Fig. 2(A)] of the pyrrole rings, as previously reported by others.[64,65] In Figure 10, the heme deformation defined by τa and τb dihedral angles is shown as function of the simulation time. While the heme of the starting conformation was almost planar (τa = −0.66° and τb = 1.47°), some out of plane distortions were identified during the MD, the τa and τb average values being of −7.53° and 10.91°, respectively, during the last 5 ns of simulations. This slight ruffling might be related to a modulating effect of the prosthetic group on the biological properties of the hAR enzyme.

The trajectory of the last 3.5 ns of simulations was further investigated applying the essential dynamics protocol. The covariance matrix of the displacements of the protein Cα atoms is reported in Figure11(A), while in Figure11(B), the main protein movements detected according to the essential dynamics protocol are shown. It can be seen that hAR experienced harmonic motions, as a consequence of a good degree of stability within that simulations interval. However, a rearrangement of the protein was observed in a region comprising the helices E, G, H, and I along with the loop connecting B and C helices. A deeper investigation showed that the latter loop showed a high degree of movement correlation with helices G and I. Interestingly enough, the H helix, whose amino acidic sequence is comprised between those of helices G and I, did not seem to get directly involved with their motions but appeared to be strictly related to the motion of the E helix. In particular, the B–C connecting loop, due to its peculiar spatial position, seemed to play a central role in the dynamic behavior of the protein. Its spatial proximity to helices G and I allowed the concerted motions of a big part of the protein, modulating the exposure to the solvent of the prosthetic heme group and the enzyme active site. As a speculation, we may advance the hypothesis that such a motion might be responsible for the substrate entrance and the product release.

DISCUSSION

A new model of the hAR based on the very recent crystal structure of human cytochrome P450 2C9 was built and validated through both docking and MD simulations. Since both types of calculations rely upon force field parameters for the energy estimation, accurate parameterization of the hAR prosthetic group heme was earlier addressed. To this aim, DFT calculations were carried out using the B3LYP hybrid functional and the SDD basis set, which both have been demonstrated to be very suitable to describe transition metal complexes (see Methods). To our knowledge, here for the first time parameters at a DFT level of theory have been developed for heme-cysteinato complexes with a hexacoordinated Fe^{2+}. Since most of the cytochromes of biological, pharmacological, toxicological, and metabolic interest belong to the P450 family, we are confident that the present set of DFT parameters for the Fe^{2+} heme group along with those for the Fe^{3+} heme very recently reported by Oda et al.[57] will be exploited in future molecular mechanics studies of P450. In the following, we will discuss the relevance of both the Fe^{2+} heme parameterization and the homology built model in the light of previous computational and mutagenesis studies.

Heme Parameterization

The P450 heme group was extracted from the crystallographic structure of CYP2C9 and geometrically optimized by means of DFT-based computations. During the minimizations, both vinyl dihedral angles of the crystallographic heme (angles 2-3-3′-3” and 7-8-8′-8”; see Fig. 2 and Table I) varied only slightly, revealing a fairly good stability of these (heme) side chains. Remarkably, both of the prosthetic groups of the homo-dimer of P450 2C9 crystal structure bear the 3-vinyl in a *cis* conformation, while the 8-vinyl adopts the *trans* one. It is well known that the *cis*

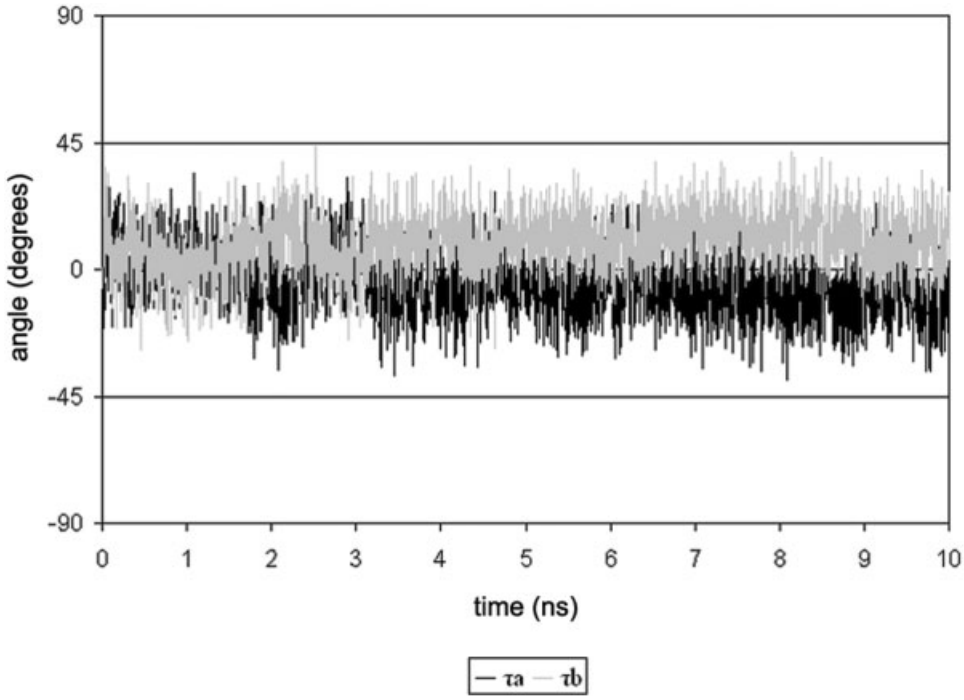

Fig. 10. τa (black lines) and τb (gray lines) defined according to Table I and Figure 2(A), monitored as a function of the simulation time. The complete heme planarity is observed when τa and τb are equal to zero.

conformation of the vinyl groups in porphyrins is more stable than the *trans* one.[27] However, an experimental and theoretical study for a set of vinyl-hemins has revealed that those isomers may easily interconvert at room temperature,[66] and more importantly, the side chain conformations do not affect the electronic properties of the metal coordination system.[27] Moreover, it should be remarked that isolated heme groups are quite different from those embedded in biological systems, as the protein environment may stabilize peripheral substituents in a less stable conformation.

Besides the heme side chain conformations, the planarity of the porphyrin is a further geometrical issue of pivotal importance that has to be accounted for when dealing with quantum chemical study of heme groups. In particular, the porphyrin conformation should be kept as much as possible close to the crystallographic one for further classical mechanics-based modeling studies.[55] Actually, while isolated porphyrin rings are perfectly planar,[54] those embedded within protein environment might display relevant distortions. Besides the protein environment, the main factors responsible for the porphyrin conformations are also the axial coordination of the central iron atom,[55] and radius, spin, and redox states of the metal center.[64,67] To assess the overall distortion of the porphyrin ring (24 atoms), we monitored the RMSD of its non-hydrogen atoms from the porphyrin mean plane (Table I). As reported in the literature, four main conformations for the porphyrin ring (ruffled, saddled, domed, and waved) have been characterized, the ruffled one being the most commonly observed.[54,55] In the present study, the amount of ruffling distortions was evaluated employing the definition earlier introduced by Vangberg and Ghosh,[64] namely, monitoring the torsional angles Cα-N-N-Cα (dihedral angels τa and τb) [see Table I and Fig. 2(A)] of the heme moiety. As shown in Figure 10, a slight ruffling was observed during the MD; however, τa and τb dihedral angles kept close to planarity values throughout the simulations.

In conclusion, our DFT calculations on the heme group retrieved from the P450 2C9 crystallographic structure provided a reliable Fe^{2+} containing heme geometry, as also confirmed by analyzing the splitting level of the frontier molecular orbitals in the coordination complex. Actually, we could assert that the complex here identified was characteristic of a low-spin iron ion in an octahedral ligand field, as the three highest occupied molecular orbitals were essentially the d_{xy}, d_{yz}, d_{xz}, whilst the d_{z^2} and $d_{x^2-y^2}$ orbitals lay among the unoccupied ones.

Model of hAR

A new hAR homology model was presented and deposited in the theoretical section of the PDB server (PDB code 1TQA). Since the model was mainly built for drug design purposes and docking simulations, the starting homology built model, rather than an equilibrated MD snapshot, was deposited in the PDB server. Actually, a high percentage of sequence identity in the ligand-binding pocket of hAR might provide a rather suitable conformation of the active site amino acids for target-based computational studies. The new homology model was, in fact, built using the human P450 2C9 as template and turned out to be significantly different from earlier models[60,68–70] based on bacterial P450 structures. In particular, the present model

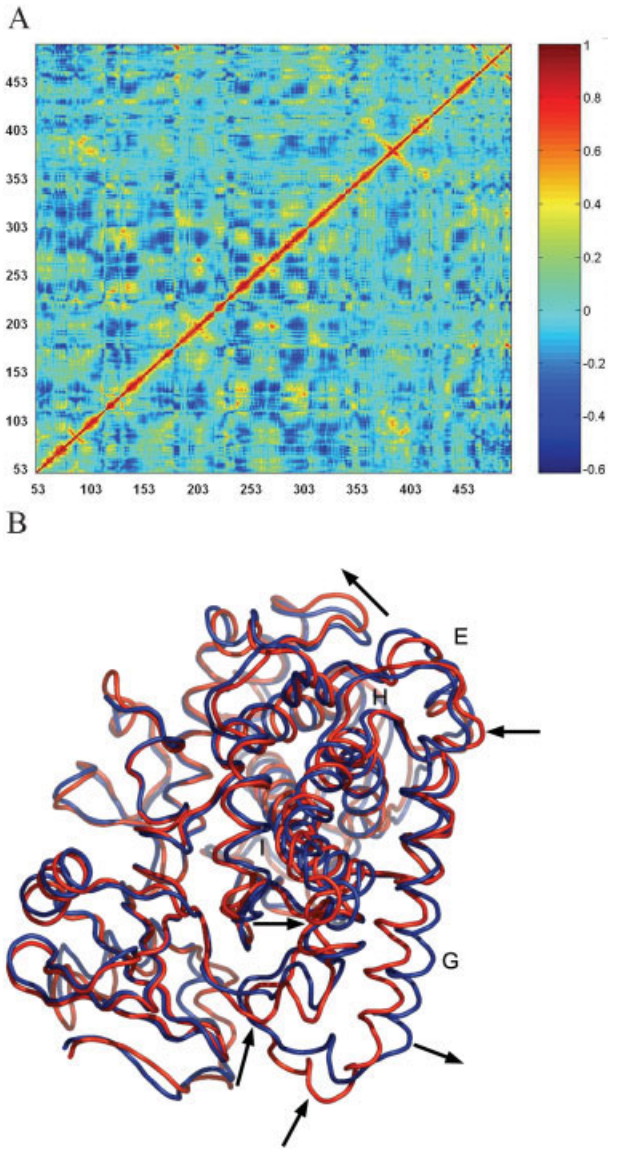

Fig. 11. **A:** Correlation matrix of the motions of aromatase protein Cα atoms, calculated during the last 3.5 ns of MD simulations. Correlation coefficients are color coded as shown in the right bar. **B:** 3D representation of the main movement of human aromatase, detected according to the essential dynamics protocol. The protein trace is shown. Helices E, G, H, and I are highlighted to help the interpretation. The starting and final conformations are colored red and blue, respectively.

differs in the residues previously assigned to the B–C loop and the F and G helices that participate to the active site composition. In the present model, all of these have been reassigned due to the differences in the alignments. Conversely, as expected, the present model might be quite similar to the hAR structure modeled on the basis of the crystallographic rabbit cytochrome P450 2C5 (PDB code 1DT6) (see Chen et al.[73]). However, slight differences between the human and rabbit enzymes might be reflected on the homology models based on the two templates. In particular, 1DT6 misses 11 residues belonging to the F–G loop, making the modeling of this important region of the enzyme somewhat hard to be properly carried out. Moreover, a difference between the two protein binding sites concerned the presence, in P450 2C9, of a valine, namely V113 instead of an alanine. In our study, this residue was used to model I133 an important residue (see below) of the AR active site (Fig. 6).

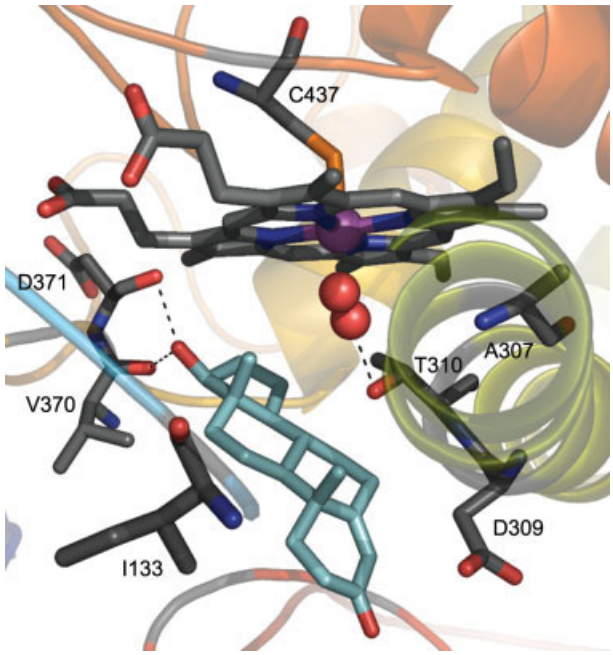

Fig. 12. Docking of the natural substrate testosterone at the 1TQA active site. Residues I133, A307, D309, T310, V370, D371, C437, and heme prosthetic group are represented in stick, colored according to the atom-code (C atoms in gray). Testosterone is represented in stick, colored according to the atom-code (C atoms in light blue). Molecular oxygen is represented in CPK. Dashed lines connect atoms involved in H-bond interactions.

Similarly to some other P450s, hAR is a target for drug design,[14,17] and, therefore, it might be very interesting to accurately determine the characteristic and composition of the active site. In our model, we detected 30 amino acids forming its active site pocket (i.e., the amino acids lying 7 Å from vorozole in the docking complex reported in Fig. 6). In particular, E302, D309, T310, H475, S478, H480, and a set of hydrophobic amino acids (I133, F134, I305, A307, V313, P368, V373, P429, F430, and L477) were found in the enzyme active site. As shown by the docking complex hAR-vorozole (Fig. 6), besides the coordination bond with the heme iron, nonsteroidal inhibitors could interact both with some hydrophobic residues of the active site (i.e., I305, V313, V370, V373, and L477) and also with S478 (by means of an H-bond interaction).

In the literature, one can find several studies dealing with mutagenesis experiments mainly aimed at establishing the hAR active site composition and the role of some of the above-mentioned residues. For instance, E302, an amino acid belonging to the I-helix, was supposed to be a key residue of the AR active site. Actually, E302L mutant was found to be inactive,[71] while the E302D had only a 30% of activity of the wild type enzyme.[72] In our model, E302 was aligned onto D293 of the I-helix of CYP2C9 (Fig. 4), and was thus positioned in the hAR active site. How-

ever, inspecting the docking model of testosterone at the hAR active site (Fig. 12), we could not detect a direct interaction between E302 and the substrate, although E302 was found to be not far from the docked molecule. Conversely, E302 was clearly identified to take part in a salt-bridge with K130, as this interaction was quite stable throughout 10 ns of MD simulations. This suggested an electrostatic role for E302 in keeping the overall active site folding, in good agreement with a previous proposal by Graham-Lorence et al.[60] Another fundamental I-helix residue is D309. Interestingly, D309A mutant was resistant to a great number of hARIs[73,74] as the mutation is expected to cause the collapse of the enzyme active site. Further MD simulations on D309A mutant could help verify this hypothesis. For the aromatization mechanism, the role played by T310, a very conserved amino acid throughout the P450 family, is very important. T310 belongs to the I-helix and here it was aligned onto T300 of P450 2C9. In our model, T310 is located in the core of the active pocket, under the area occupied by the prosthetic group, supporting the hypotheses dealing with its peculiar role during the aromatization reaction.

To better investigate the role of T310 and other active site amino acids in the aromatization reaction, docking studies of the substrate testosterone at the enzyme active site were carried out. In Figure 12, carbons 1 and 2 of the A ring of testosterone are shown to be close to D309 that was demonstrated to be a fundamental active site residue, whereas the C19 methyl is facing the molecular oxygen, and is 2.8 Å far from T310. Moreover, the substrate is stabilized, into the binding site cavity, by the formation of H bonds involving its OH group and the backbone oxygen atoms of V370 and E371. Although the docking model was not able to account for all of the mutations so far reported in the literature, this allowed us a better understanding of some previously reported experimental results. In particular, the key role in the catalysis of D309 and T310 is confirmed by the present docking model.

Two important polar residues likely involved in the interaction with some hAR ligands, are S478 and H480, which in some early hAR models were located far from the enzyme active site.[60,74] Mutagenesis experiments clearly showed that such amino acids should be located within the hAR active site and involved in interactions with ligands.[4] In the present model, both residues are within 5.4 Å from vorozole, and S478 particularly interacted with the nonsteroidal inhibitor by means of an H-bond, as shown by the docking complex reported in Figure 6.

Most of the above-mentioned residues have been recently reported as key residues in the enzymatic reaction catalyzed by hAR.[4] Interestingly, D309, S478, and H480 have been hypothesized to participate in a charge relay system analogous to the catalytic triad of serine proteases.[75] Monitoring the distances between the heteroatoms likely involved in such a relay system, some snapshots matching geometrical requirements of a triad-like system were identified (data not shown). Starting from a reliable conformation, ab initio calculations might help to unravel the issue studying the charge and proton transfer process among these amino acids. In general, quantum chemical calculations might help to confirm a previously suggested reaction mechanism.[4] In this respect, very recently Hackett et al. investigated at the DFT level the final catalytic step of the hAR enzyme.[76] The model system investigated in the study was composed of the heme and the cycles A and B of the androgen substrate. In the light of this pioneering report, future QM/MM calculations that require a reliable starting hAR model will help to estimate the effects of the protein molecular fields on the aromatization mechanism disclosed by Hackett et al.[76] We are confident that the present model of the hAR enzyme could be exploited in future atomistic studies, first-principles quantum chemical calculations of the aromatization reaction.

Also the role of hydrophobic residues of the hAR active site has been assessed by means of mutagenesis experiments. For instance, A307 has been shown to affect substrate binding affinity as much as 50%.[72] Concerning I133, recent mutagenesis experiments pointed to this residue as a key amino acid of hAR active site.[77] Actually, substitutions of I133 with valine, alanine, or tyrosine led to a decrease of efficiency of the aromatization reaction. Moreover, a comparative analysis based on the crystal structure of P450 2C5 showed that A113 of P450 2C5 corresponding to I133 of hAR is located in the active site in close proximity to the heme group.[77] Actually, the docking of testosterone at the hAR active site (Fig. 12) pointed out a role for I133 as the residue located in the binding pocket in close proximity to the substrate molecule.

CONCLUSIONS

The mammalian cytochrome P450 enzymes family comprises several validated and promising pharmacological targets. Very recently, an overview on the role of modeling approaches to the study of P450 appeared in the literature.[78] The computational methodologies applied to P450 span from historical ligand-based QSAR and 3D QSAR approaches to target-based homology modeling, MD, and quantum chemical calculations. Due to their membrane-bound nature, the purification of mammalian CYP proteins is extremely difficult, and therefore the experimental structural characterization of mammalian P450 is proceeding very slowly. Recently, rabbit and human P450 crystal structures have been solved, opening a new way in the field of homology modeling of P450. Particularly, the human crystallographic structures provide an excellent template for the construction of other human P450, whose models might be exploited for both basic and applied computational works. Actually, studying P450 reaction mechanisms by means of DFT or other QM-based approaches, as well as docking simulations to discover new potential bioactive agents, need a very accurate starting model. While for the basic studies, MD simulations could further refine the model on which ab initio calculations will be performed, for the biological and pharmacological applications, scoring function parameters of appropriate quality will provide reliable (free) energy of ligand-target binding. In this scenario, we thought it relevant to develop

a suitable protocol to build P450 homology models, and to validate it through docking and MD simulations. For the latter point, a set of ad hoc parameters for the iron-carrying protoporphyrin group had to be earlier determined. In the light of the above considerations, in this report we proposed a new homology model of the hAR based upon the recently reported crystal structure of the human P450 2C9, together with a DFT-based heme parameterization. We believe that the present protocol and the set of heme parameters might be exploited in future docking and MD simulations of hAR and other heme-cysteinato cytochromes.

ACKNOWLEDGMENTS

We thank Ms. Alessandra Barozzino for technical assistance and Dr. Walter Rocchia for useful discussions.

REFERENCES

1. Hodgson J. ADMET: turning chemicals into drugs. Nat Biotechnol 2001;19:722–726.
2. Meunier B, de Visser SP, Shaik S. Mechanism of oxidation reactions catalyzed by cytochrome p450 enzymes. Chem Rev 2004;104:3947–3980.
3. de Groot MJ, Ekins S. Pharmacophore modeling of cytochromes P450. Adv Drug Deliv Rev 2002;54:367–383.
4. Kao YC, Korzekwa KR, Laughton CA, Chen S. Evaluation of the mechanism of aromatase cytochrome P450. A site-directed mutagenesis study. Eur J Biochem 2001;268:243–251.
5. Brueggemeier RW, Hackett JC, Diaz-Cruz ES. Aromatase inhibitors in the treatment of breast cancer. Endocr Rev 2005;26:331–345.
6. Chen S. Aromatase and breast cancer. Front Biosci 1998;3:d922–933.
7. Seralini G, Moslemi S. Aromatase inhibitors: past, present and future. Mol Cell Endocrinol 2001;178:117–131.
8. Recanatini M, Cavalli A, Valenti P. Nonsteroidal aromatase inhibitors: recent advances. Med Res Rev 2002;22:282–304.
9. Allred DC, Baum M, Buzdar AU, Carlson RW, Dowsett M, Elledge RM, Gradishar WJ, Grana G, Howell A, Mamounas EP. A roundtable discussion of aromatase inhibitors as therapy for breast cancer. Breast J 2003;9:213–222.
10. Osborne C, Tripathy D. Aromatase inhibitors: rationale and use in breast cancer. Annu Rev Med 2005;56:103–116.
11. Jackson J, Miller WR, Dixon JM. Safety issues surrounding the use of aromatase inhibitors in breast cancer. Expert Opin Drug Safe 2003;2:73–86.
12. Poulos TL, Finzel BC, Howard AJ. High-resolution crystal structure of cytochrome P450cam. J Mol Biol 1987;195:687–700.
13. Ravichandran KG, Boddupalli SS, Hasermann CA, Peterson JA, Deisenhofer J. Crystal structure of hemoprotein domain of P450BM-3, a prototype for microsomal P450's. Science 1993;261:731–736.
14. Hasemann CA, Ravichandran KG, Peterson JA, Deisenhofer J. Crystal structure and refinement of cytochrome P450terp at 2.3 A resolution. J Mol Biol 1994;236:1169–1185.
15. Cupp-Vickery JR, Poulos TL. Structure of cytochrome P450eryF involved in erythromycin biosynthesis. Nat Struct Biol 1995;2:144–153.
16. Williams PA, Cosme J, Sridhar V, Johnson EF, McRee DE. Microsomal cytochrome P450 2C5: comparison to microbial P450s and unique features. J Inorg Biochem 2000;81:183–190.
17. Williams PA, Cosme J, Sridhar V, Johnson EF, McRee DE. Mammalian microsomal cytochrome P450 monooxygenase: structural adaptations for membrane binding and functional diversity. Mol Cell 2000;5:121–131.
18. Williams PA, Cosme J, Ward A, Angove HC, Matak Vinkovic D, Jhoti H. Crystal structure of human cytochrome P450 2C9 with bound warfarin. Nature 2003;424:464–468.
19. Scott EE, He YA, Wester MR, White MA, Chin CC, Halpert JR, Johnson EF, Stout CD. An open conformation of mammalian cytochrome P450 2B4 at 1.6-A resolution. Proc Natl Acad Sci USA 2003;100:13196–13201.
20. Amadei A, Linssen AB, Berendsen HJ. Essential dynamics of proteins. Proteins 1993;17:412–425.
21. Bayly CI, Cieplak P, Cornell WD, Kollman PA. 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.
22. Becke AD. A new mixing of Hartree–Fock and local density-functional theories. J Chem Phys 1993;98:1372–1377.
23. Siegbahn PE. Mechanisms of metalloenzymes studied by quantum chemical methods. Q Rev Biophys 2003;36:91–145.
24. Dolg M, Wedig U, Stoll HP, H. Energy-adjusted ab initio pseudopotentials for the first row transition elements. J Chem Phys 1987;86:866–872.
25. Liu YP. Applications of effective core potentials and density functional theory to the spin states of iron porphyrin. J Chem Inf Comput Sci 2001;41:22–29.
26. Frisch MJ, Trucks GW, Schlegel HB, Scuseria GE, Robb MA, Cheeseman JR, Zakrzewski VG, Montgomery JA Jr, Stratmann RE, Burant JC, Dapprich S, Millam JM, Daniels AD, Kudin KN, Strain MC, Farkas O, Tomasi J, Barone V, Cossi M, Cammi R, Mennucci B, Pomelli C, Adamo C, Clifford S, Ochterski J, Petersson GA, Ayala PY, Cui Q, Morokuma K, Malick DK, Rabuck AD, Raghavachari K, Foresman JB, Cioslowski J, Ortiz JV, Stefanov BB, Liu G, Liashenko A, Piskorz P, Komaromi I, Gomperts R, Martin RL, Fox DJ, Keith T, Al-Laham MA, Peng CY, Nanayakkara A, Gonzalez C, Challacombe M, Gill PMW, Johnson BG, Chen W, Wong MW, Andres JL, Head-Gordon M, Replogle ESP, J. A. Gaussian 03. Pittsburgh, PA: Gaussian, Inc.; 2003.
27. Scott EE, White MA, He YA, Johnson EF, Stout CD, Halpert JR. Structure of mammalian cytochrome P450 2B4 complexed with 4-(4-chlorophenyl)imidazole at 1.9-A resolution: insight into the range of P450 conformations and the coordination of redox partner binding. J Biol Chem 2004;279:27294–27301.
28. Rovira C, Carloni P, Parrinello M. The iron-sulphur bond in cytochrome c. J Phys Chem B 1999;103:7031–7035.
29. Rovira C, Kunc K, Hutter J, Ballone P, Parrinello M. Equilibrium geometries and electronic structure of iron-porphyrin complexes: a density functional study. J Phys Chem A 1997;101:8914.
30. Rovira C, Parrinello M. Factors influencing ligand-binding properties of heme models: a first principle study of picket-fence and protoheme complexes. Chem Eur J 1999;5:250–256.
31. Cornell WD, Cieplak P, Bayly CI, Kollman PA. Application of RESP charges to calculate conformational energies, hydrogen bond energies, and free energies of solvation. J Am Chem Soc 1993;115:9620–9631.
32. Cornell WD, Cieplak P, Bayly CI, Gould IR, Merz KM, Ferguson DM, Spellmeyer DC, Fox T, Caldwell JW, Kollman PA. A second generation force field for the simulation of proteins, nucleic acids, and organic molecules. J Am Chem Soc 1995;117:5179–5197.
33. Singh UC, Kollman PA. An approach to computing electrostatic charges for molecules. J Comput Chem 1983;5:129–145.
34. Besler BH, Merz KM, Kollman PA. Atomic charges derived from semiempirical methods. J Comput Chem 1990;11:431–439.
35. Sigfridsson E, Ryde U. Comparison of methods for deriving atomic charges from the electrostatic potential and moments. J Comput Chem 1998;4:377–395.
36. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE. The Protein Data Bank. Nucleic Acids Res 2000;28:235–242.
37. Huang X, Miller W. A time-efficient, linear-space local similarity algorithm. Adv Appl Math 1991;12:337–357.
38. Mestres J. Structure conservation in cytochromes P450. Proteins 2005;58:596–609.
39. McGuffin LJ, Bryson K, Jones DT. The PSIPRED protein structure prediction server. Bioinformatics 2000;16:404–405.
40. Sali A, Blundell TL. Comparative protein modelling by satisfaction of spatial restraints. J Mol Biol 1993;234:779–815.
41. Laskowski RA, McArthur MW, Moos DS, Thornton JMJ. PROCHECK: a program to check the stereochemical quality of protein structures. Appl Cryst 1993;26:283–291.
42. Wang J, Cieplak P, Kollman PA. How well does a restrained electrostatic potential (RESP) model perform in calculating conformational energies of organic and biological molecules? J Comput Chem 2000;21:1049–1074.
43. Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein LM. Comparison of simple potential functions for simulating liquid water. J Chem Phys 1983;79:926–935.

44. Collins JR, Camper DL, G.H. L. Valproic acid metabolism by cytochrome P450: A theoretical study of stereoelectronic modulators of product distribution. J Am Chem Soc 1991;113:2736–2743.
45. Autenrieth F, Tajkhorshid E, Baudry J, Luthey-Schulten Z. Classical force field parameters for the heme prosthetic group of cytochrome c. J Comput Chem 2004;25:1613–1622.
46. Essmann U, Perera L, Berkowitz ML, Darden T, Lee H, Pedersen LG. A smooth particle mesh Ewald method. J Chem Phys 1995;103: 8577–8593.
47. Ryckaert JP, Ciccotti G, Berendsen HJC. Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. J Comput Phys 1977;23:327–341.
48. Banci L, Schroder S, Kollman PA. Molecular dynamics characterization of the active cavity of carboxypeptidase A and some of its inhibitor adducts. Proteins 1992;13:288–305.
49. Cavalli A, Dezi C, Folkers G, Scapozza L, Recanatini M. Three-dimensional model of the cyclin-dependent kinase 1 (CDK1): Ab initio active site parameters for molecular dynamics studies of CDKS. Proteins 2001;45:478–485.
50. Berendsen HJC, Postma JPM, Van Gunsteren WF, Di Nola A, Haak JR. Molecular dynamics with coupling to an external bath. J Chem Phys 1984;81:3684–3690.
51. Case DA, Darden TE, Cheatham TEI, Simmerling CL, Wang J, Duke RE, Luo R, Merz KM, Wang B, Pearlman DA, Crowley M, Brozell S, Tsui V, Gohlke H, Mongan J, Hornak V, Cui G, Beroza P, Schafmeister C, Caldwell JW, Ross WS, Kollman PA. AMBER 8. San Francisco: University of California; 2004.
52. Mongan J. Interactive essential dynamics. J Comput Aided Mol Des 2004;18:433–436.
53. Humphrey W, Dalke A, Schulten K. VMD: visual molecular dynamics. J Mol Graph 1996;14:33–38.
54. Jentzen W, Ma JG, Shelnutt JA. Conservation of the conformation of the porphyrin macrocycle in hemoproteins. Biophys J 1998;74: 753–763.
55. Shelnutt JA, Song XZ, Ma JG, Jia SL, Jentzen W, Medforth CJ. Nonplanar porpyrins and their significance in proteins. Chem Soc Rev 1998;27:31–41.
56. Scheidt WR, Reed CA. Spin-state/stereochemical relationships in iron porphyrins: implications for the hemoproteins. Chem Rev 1981;81:543–555.
57. Oda A, Yamaotsu N, Hirono S. New AMBER force field parameters of heme iron for cytochrome p450s determined by quantum chemical calculations of simplified models. J Comput Chem 2005;26:818–826.
58. Winter M. WebElements™ Periodic Table. United Kingdom: University of Sheffield. http://www.webelements.com/.
59. RCSB PDB Protein Data Bank. http://deposit.pdb.org/adit.
60. Graham-Lorence S, Amarneh B, White RE, Peterson JA, Simpson ER. A three-dimensional model of aromatase cytochrome P450. Protein Sci 1995;4:1065–1080.
61. Morris GM, Goodsell DS, Halliday RS, Huey R, Hart WE, Belew RK, Olson AJ. Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function. J Comput Chem 1998;19:1639–1662.
62. Bursulaya BD, Totrov M, Abagyan R, Brooks CL 3rd. Comparative study of several algorithms for flexible ligand docking. J Comput Aided Mol Des 2003;17:755–763.
63. Ludwig M, Beck A, Wickert L, Bolkenius U, Tittel B, Hinkel K, Bidlingmaier F. Female pseudohermaphroditism associated with a novel homozygous G-to-A (V370-to-M) substitution in the P-450 aromatase gene. J Pediatr Endocrinol Metab 1998;11:657–664.
64. Vangberg T, Ghosh AA. First-principles quantum chemical analysis of the factors controlling ruffling deformations of porphyrins: insights from the molecular structures and potential energy surfaces of silicon, phosphorus, germanium, and arsenic poprphyrins and of a peroxidase compound I model. J Am Chem Soc 1999;121:12154–12160.
65. Kiefl C, Sreerama N, Haddad R, Sun L, Jentzen W, Lu Y, Qiu Y, Shelnutt JA, Woody RW. Heme distortions in sperm-whale carbonmonoxy myoglobin: correlations between rotational strengths and heme distortions in MD-generated structures. J Am Chem Soc 2002;124:3385–3394.
66. Kalsbeck WA, Ghosh A, Pandey RK, Smith KM, Bocian DF. Determinants of the vinyl stretching frequency in protoporphyrins. Implications for cofactor-protein interactions in heme proteins. J Am Chem Soc 1995;117:10959–10968.
67. Munro OQ, Bradley JC, Hancock RD, Marques HM, Marsicano F, Wade PW. Molecular mechanics study of the ruffling of metalloporhyrins. J Am Chem Soc 1992;114:7218–7230.
68. Laughton CA, Zvelebil MJ, Neidle S. A detailed molecular model for human aromatase. J Steroid Biochem Mol Biol 1993;44:399–407.
69. Koymans LM, 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.
70. Cavalli A, Greco G, Novellino E, Recanatini M. Linking CoMFA and protein homology models of enzyme-inhibitor interactions: an application to non-steroidal aromatase inhibitors. Bioorg Med Chem 2000;8:2771–2780.
71. Zhou DJ, Korzekwa KR, Poulos T, Chen SA. A site-directed mutagenesis study of human placental aromatase. J Biol Chem 1992;267:762–768.
72. Graham-Lorence S, Khalil MW, Lorence MC, Mendelson CR, Simpson ER. Structure-function relationships of human aromatase cytochrome P-450 using molecular modeling and site-directed mutagenesis. J Biol Chem 1991;266:11939–11946.
73. Chen S, Kao YC, Laughton CA. Binding characteristics of aromatase inhibitors and phytoestrogens to human aromatase. J Steroid Biochem Mol Biol 1997;61:107–115.
74. Kao YC, Cam LL, Laughton CA, Zhou D, Chen S. Binding characteristics of seven inhibitors of human aromatase: a site-directed mutagenesis study. Cancer Res 1996;56:3451–3460.
75. Lewis DF, 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.
76. Hackett JC, Brueggemeier RW, Hadad CM. The final catalytic step of cytochrome P450 aromatase: a density functional theory study. J Am Chem Soc 2005;127:5224–5237.
77. Conley A, Mapes S, Corbin CJ, Greger D, Graham S. Structural determinants of aromatase cytochrome p450 inhibition in substrate recognition site-1. Mol Endocrinol 2002;16:1456–1468.
78. de Graaf C, Vermeulen NP, Feenstra KA. Cytochrome p450 in silico: an integrative modeling approach. J Med Chem 2005;48: 2725–2755.