*Protein Science* (1995), *4*:1065–1080. Cambridge University Press. Printed in the USA.


# A three-dimensional model of aromatase cytochrome P450*

SANDRA GRAHAM-LORENCE,[1] BILAL AMARNEH,[2,3] RONALD E. WHITE,[4] JULIAN A. PETERSON,[1] AND EVAN R. SIMPSON[1,2,3]

[1] Department of Biochemistry, The University of Texas Southwestern Medical Center at Dallas, Dallas, Texas 75235
[2] The Cecil H. and Ida Green Center for Reproductive Biology Sciences, The University of Texas Southwestern Medical Center at Dallas, Dallas, Texas 75235
[3] Department of Obstetrics/Gynecology, The University of Texas Southwestern Medical Center at Dallas, Dallas, Texas 75235
[4] Department of Metabolism and Pharmacokinetics, Bristol-Myers Squibb Pharmaceutical Research Institute, Princeton, New Jersey 085434

(RECEIVED December 22, 1994; ACCEPTED March 20, 1995)

**Abstract**

P450 hemeproteins comprise a large gene superfamily that catalyzes monooxygenase reactions in the presence of a redox partner. Because the mammalian members are, without exception, membrane-bound proteins, they have resisted structure-function analysis by means of X-ray crystallographic methods. Among P450-catalyzed reactions, the aromatase reaction that catalyzes the conversion of C19 steroids to estrogens is one of the most complex and least understood. Thus, to better understand the reaction mechanism, we have constructed a three-dimensional model of P450arom not only to examine the active site and those residues potentially involved in catalysis, but to study other important structural features such as substrate recognition and redox-partner binding, which require examination of the entire molecule (excepting the putative membrane-spanning region). This model of P450arom was built based on a "core structure" identified from the structures of the soluble, bacterial P450s (P450cam, P450terp, and P450BM-P) rather than by molecular replacement, after which the less conserved elements and loops were added in a rational fashion. Minimization and dynamic simulations were used to optimize the model and the reasonableness of the structure was evaluated. From this model we have postulated a membrane-associated hydrophobic region of aliphatic and aromatic residues involved in substrate recognition, a redox-partner binding region that may be unique compared to other P450s, as well as residues involved in active site orientation of substrates and an inhibitor of P450arom, namely vorozole. We also have proposed a scheme for the reaction mechanism in which a "threonine switch" determines whether oxygen insertion into the substrate molecule involves an oxygen radical or a peroxide intermediate.

**Keywords:** model; P450arom; reaction mechanism; sequence alignment; structure

P450 hemeproteins comprise a large gene superfamily with over 300 known members belonging to 36 gene families (Nelson et al., 1993). Members of this family contain an absolutely conserved cysteine residue that serves as the fifth coordinating ligand of the heme iron. The mammalian members are, without exception, membrane-bound proteins requiring detergent treatment for solubilization and thus have resisted structure-function analysis by means of X-ray crystallographic methods because of their intransigence to solubilization, and hence, to crystallization.

In 1987, Poulos et al. (1987) determined the first structure of a P450 enzyme, that of P450cam, a soluble, bacterial P450 isolated from *Pseudomonas putida*. Recently, in the laboratories of Peterson and Deisenhofer, two more structures have been determined, that of P450terp from *Pseudomonas* (Haseman et al., 1994), which requires both an iron-sulfur protein and an NADH reductase for catalysis, as does P450cam (Class II), and that of P450BM-P from *Bacillus megaterium* (Ravichandran et al., 1993), which requires only an NADPH P450-reductase, similar to most mammalian P450s (Class I). Comparing these three proteins, one finds that an overall structural fold of P450s is conserved, with the most highly conserved region, the "core

*This paper is dedicated to Cecil H. Robinson on the occasion of his retirement.

Reprint requests to: Evan R. Simpson, Green Center for Reproductive Biology Sciences, UT Southwestern Medical Center, 5323 Harry Hines Boulevard, Dallas, Texas 75235-9051; e-mail: simpson@grnctr.swmed.edu.

structure," consisting essentially of a four-helix bundle, two β-sheets, and the heme-binding region.

Generally, P450s along with their redox partner(s) catalyze the monooxygenation of hydrophobic compounds such as fatty acids, polycyclic aromatic hydrocarbons, and steroids. One of the more interesting and mechanistically complex reactions catalyzed by a member of the P450 superfamily is the conversion of C19 steroids to estrogens in the endoplasmic reticulum by aromatase P450 (P450arom), the product of the CYP19 gene (Nelson et al., 1993). This reaction, known as aromatization, is the only reaction in vertebrates that results in the formation of an aromatic phenyl ring. Together with its ubiquitous redox partner NADPH-P450 reductase, P450arom catalyzes the aromatization of the A-ring of androgens to form the phenolic ring characteristic of estrogens, with concomitant loss of the carbon-19 (C19) angular methyl group (Fig. 1). Thompson and Siiteri (1974) established that the stoichiometry of the aromatase reaction was such that for every mole of substrate C19 steroid consumed, 3 mol of oxygen and 3 mol of NADPH were utilized. Although there is now general agreement that the first two sites of attack of the C19 steroid substrate by molecular oxygen involve sequential hydroxylations on the C19 angular methyl group (Fig. 1), uncertainty has lingered as to the nature and site of the third oxygen attack. Goto and Fishman (1977) proposed some years ago the formation of a hydroxyl at the 2β-position, resulting in an unstable intermediate, which could then break down nonenzymatically to give the phenolic A ring characteristic of estrogens. However, this mechanism was challenged following the recognition that the site of incorporation of the third oxygen is the C19 methyl group, which is oxidized to formic acid. On this basis then, as shown in Figure 1, it has been suggested by Cole and Robinson (1988) that the third step could involve a peroxidative attack on C19, combined with 1β-hydrogen elimination to give a 1,10-double bond. This was based on an original concept of Akhtar and colleagues (Akhtar et al., 1982, 1993). In addition, provision of a general base for the 2β-hydrogen abstraction, such as a carboxylate, together with protonation of the 3-keto group, could allow enolization of this group to form the corresponding hydroxyl. The proton for enolization could conceivably come from a lysine, arginine, tyrosine, or histidine. Reprotonation of the hydrogen donor and removal of the proton from the carboxylate group could subsequently occur in a facile manner. Thus, the steps necessary for converting the A ring into a phenol derivative are in principle complete.

Further studies that support this mechanism have been conducted by Akhtar and colleagues (1993). These investigators provided a rationale whereby a P450–peroxide intermediate is trapped by the substrate at the active site during the third oxygen attack, resulting in the reaction outlined in Figure 1, instead of breaking down to the iron-oxo (compound I-like) intermediate characteristic of normal hydroxylation reactions. The theoretical basis of this scheme has been validated by Korzekwa and colleagues (1993), based on molecular orbital calculations and computer modeling.

In previous work, we have undertaken studies to model the three-dimensional structure of the substrate-binding pocket of P450arom (Graham-Lorence et al., 1991). This initial model was based on a comparison with P450cam, the only known P450 structure at the time. It was possible to match the corresponding residues in the region of the I-helix between P450cam and P450arom (Graham-Lorence et al., 1991) such that T310 of P450arom was found to correspond to the conserved T252 in the I-helix of P450cam, and C437 in P450arom corresponded to the absolutely conserved C357 in the heme-binding region of P450cam. However, it was apparent to us at the time that attempting to extend this model of P450arom based solely on comparison with P450cam would be unwise due to poor over-

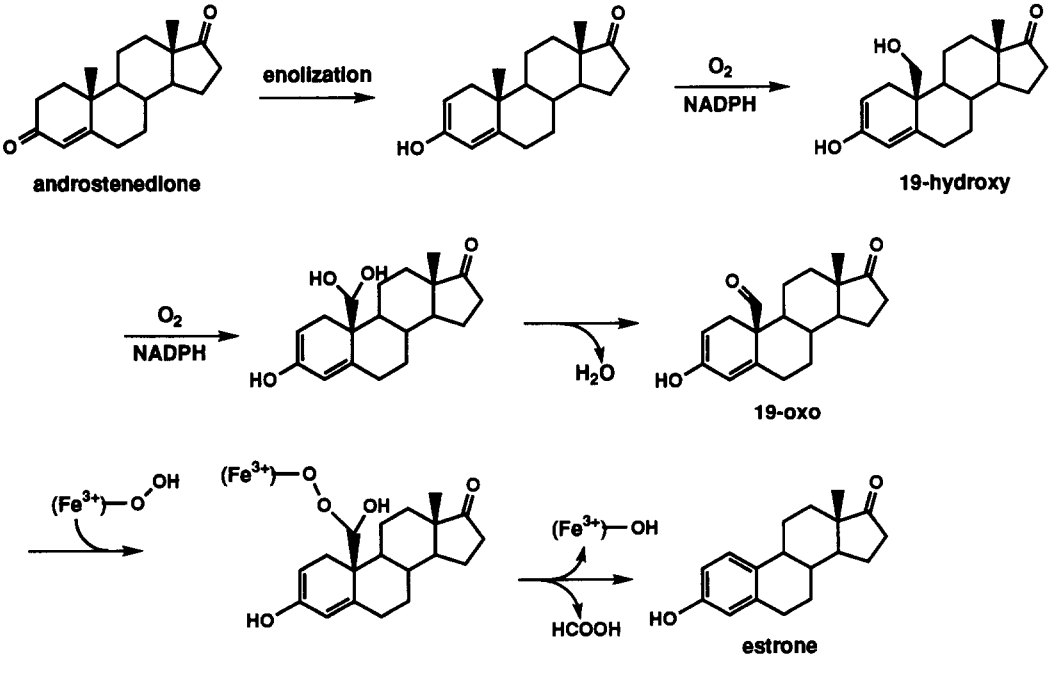


**Fig. 1.** An overall view of the reaction mechanism of aromatase showing the conversion of androstenedione to estrone.

all homology, and thus, uncertain alignments of the linear sequences in the remaining portion of the proteins.

The resolution of two additional P450 structures has revolutionized our ability to align the linear amino acid sequences of unknown P450s with those of the structurally known P450s (Ravichandran et al., 1993; Haseman et al., 1994). In particular, this is the case with the P450 domain of P450BM3 (P450BM-P), which is the first Class II member of the P450 superfamily whose structure is known, i.e., one that interacts directly with an NADPH-P450 reductase flavoprotein, similar to P450arom. Therefore, we have utilized the knowledge gained from the structural alignments of P450cam, P450terp, and P450BM-P to create new linear alignments with P450arom. Additionally, the structures of these three proteins have been utilized to identify the conserved "core structure," and then used as a basis to construct a three-dimensional model of P450arom.

## Results and discussion

### *The general model for P450s*

One of the purposes in building a model is to have a better understanding of the features that are important for catalytic activity. Some of the events in the catalytic cycle of P450s are well understood and others are surmised. As alluded to earlier, in the catalytic cycle of P450s, the enzyme recognizes a substrate, which is then bound in the active site. A redox partner binds the P450 at the redox-partner binding site and an electron is transferred to the heme, reducing the heme iron from ferric (III) to ferrous (II). Molecular oxygen is then able to bind the iron and is presumed to be hydrogen bonded to the adjacent conserved threonine. (In P450arom, this is T310.) A second electron is then donated from the redox partner, the oxygen is activated, i.e., the oxygen–oxygen bond in molecular oxygen is split forming an iron–oxo compound, most likely with the oxygen as a radical species. The oxygen radical attacks the closest reactive carbon, binding a hydrogen, forming a hydroxyl radical bound to the heme iron, and leaving a carbon-centered radical. Subsequently, the hydroxyl radical recombines with the carbon radical to form a hydroxyl group (Gerber & Sligar, 1994a, 1994b). In P450arom, there are three steps involving attack by molecular oxygen, but, in addition, the A ring of the steroid is aromatized. Therefore, in a model of P450arom, one needs to account for recognition of androstenedione (A) or testosterone (T) at the mouth of the access channel, substrate binding in the active site, redox-partner binding, oxygen activation, and positioning of the C19 of A or T within about 2 Å of the oxygen for the three monooxygenation steps. Additionally, one needs to account for general acid/base catalysis at the C2 position, where a hydrogen is abstracted enabling the enolization of the 3-keto group followed by subsequent aromatization of the steroid A ring.

Two types of information can be gleaned from the known three-dimensional structures of P450s. First, there are the "structural alignments" or those sequence alignments obtained by superimposing the three-dimensional structures of P450cam, P450terp, and P450BM-P, and aligning the sequences on the basis of the conserved structural elements. This structural alignment then can be used as a basis for alignments of unknown P450s. Secondly, a conserved "core structure" can be obtained from superimposing these proteins, which then may be used as a basis for construction of a three-dimensional model of membrane-associated P450s. Using the model that we have constructed, we will discuss those features required for membrane association, substrate recognition, redox-partner binding, and catalysis.

### *The sequence alignments*

The structural alignments for P450cam, P450terp, and P450BM-P were obtained both by combining the structural alignments from Haseman et al. (1994) and Ravichandran et al. (1993) and by overlaying the three proteins using the atoms of the heme to align the proteins. The superimposition was generated by aligning the C$\alpha$ backbone as presented in Haseman et al. (1995), but because the regions most obviously conserved in P450s are the heme-binding region and the I-helix, we chose as the focal point the superimposition of the hemes. Thus, there are some small differences between the structural sequence alignments given by Ravichandran et al. (1993) and Haseman et al. (1994, 1995) and that presented here because the ultimate goal of this alignment is the building of a three-dimensional model. This is best exemplified in the case of the G-helix because those authors aligned the G-helices strictly on a structural basis. Here, however, as shown in Figure 2, the G-helix is aligned from the C-terminal end such that the basic charges at the end of this amphipathic helix and the hydrophobic residues throughout it are aligned. This implies that although the charged residues will be pointing toward solvent and the hydrophobic patch will be pointing inward toward the bulk of the molecule in a conserved motif in all P450s, the placement of the G-helix will vary from P450 to P450 and will depend on where the intramolecular hydrophobic interaction exists and how long the N-terminal portion of the G helix is.

To obtain a sequence alignment of P450arom with the general structural alignment, a sequence alignment of more than 100 P450 proteins was generated by computer using the University of Wisconsin GCG "Pileup" program. This program clusters proteins using a pairwise alignment to determine the proteins with the highest sequence similarity and then generates the alignment using the Needleman–Wunsch algorithm (Needleman & Wunsch, 1970). Following the computer-generated sequence alignment, the alignment of P450arom with the structural alignment of P450cam, P450BM-P, and P450terp was further refined by a hand alignment resulting in the alignment shown in Figure 2. Details of these alignments are discussed more fully elsewhere (Amarneh et al., 1993); however, from the structural alignments, it is apparent that the C-terminal portion of P450s starting at the I-helix is easily aligned using almost exclusively the GCG Pileup program. Contained in this half of the molecule is the structurally conserved hydrophobic I-helix, which contains the consensus sequence (A/G)-G-X-(E/D)-T in the center of the helix and which lies directly over pyrrole ring B of the heme, except in P450arom, in which the sequence is A-A-P-D-T. (This will be discussed later in the text and shown in Figure 4.) Further along in the K-helix is the absolutely conserved sequence E-X-X-R where the E and R form a salt bridge and face into a region termed the "meander." The J′-helix, which is present in P450BM-P but not in Class I P450s (e.g., P450cam and P450terp), is immediately N-terminal of the K-helix. The length of the helix was decided partly on the basis of the $i + 4$ rule and partly by the length of the J′-helix in P450BM-P; however, it may extend two more residues N-terminal, thus encompassing R343 and D344. The heme-binding region contains the absolutely con-

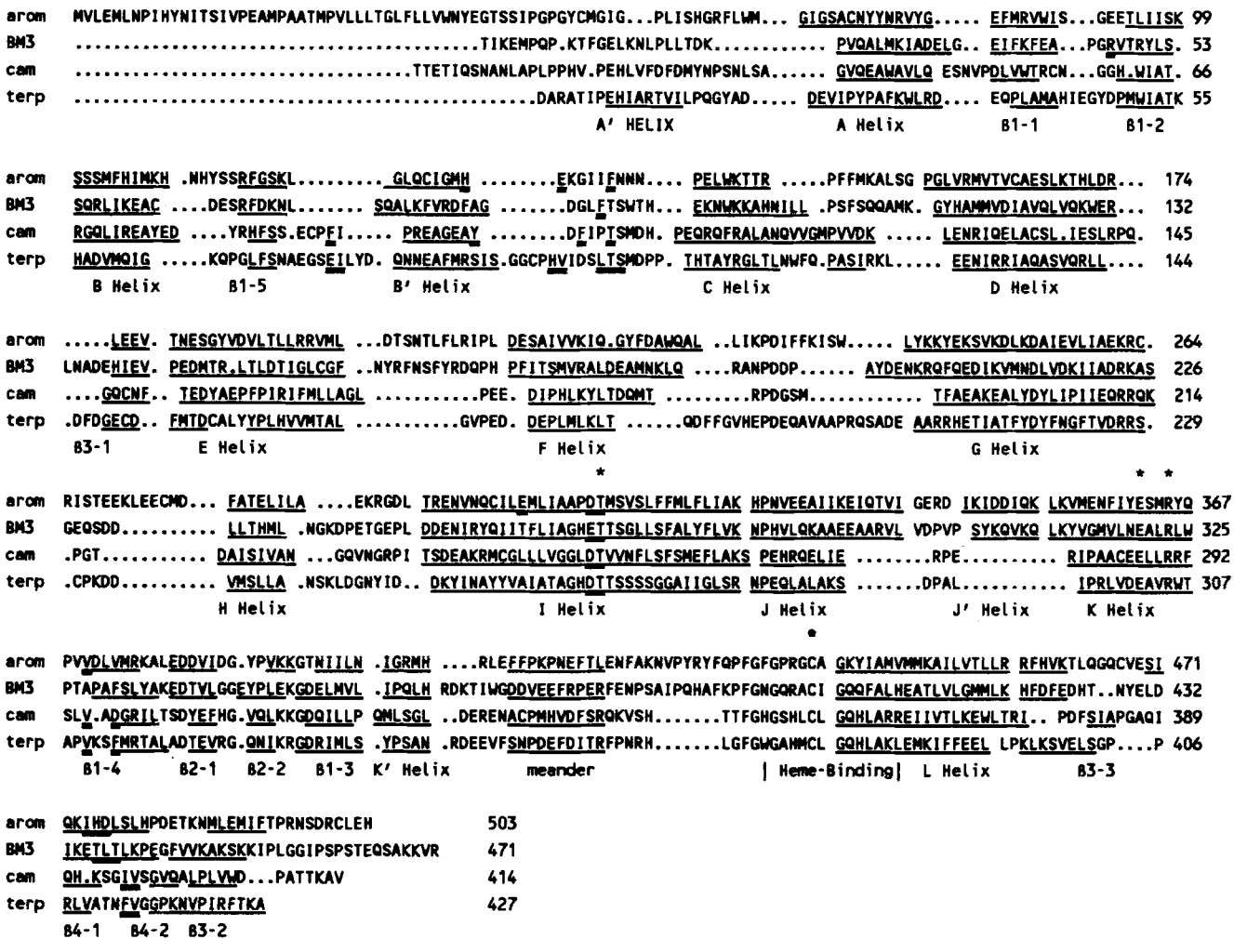


**Fig. 2.** Sequence alignment of P450arom with the structural alignments of P450cam, P450terp, and P450BM-P. $\alpha$-Helices and $\beta$-sheets are underlined and named using the same convention as that used for P450BM-P and P450terp. The highly conserved threonine involved in oxygen activation and three absolutely conserved residues are indicated by asterisks. Residues that protrude into the active site of P450cam, P450terp, and P450BM-P are indicated by double underlining. The conserved heme-binding region with the absolutely conserved cysteine that forms the fifth ligand of the heme iron is bracketed.

served cysteine, which is the fifth coordinating ligand of the heme iron and responsible for the characteristic absorption spectrum of P450s, namely an absorption band with a 450-nm maximum upon binding of CO. Along with the cysteine (C) in the heme-binding region is the highly conserved consensus sequence F(G/S)XGX(H/R)XCXGX(I/L/F)A. Between the K-helix and the heme-binding region are the $\beta$-sheets $\beta$1-3/4, $\beta$2, the K′ helix, and the "meander." $\beta$-Sheets 1 and 2 are highly conserved structures as will be discussed. The alignments of these regions are based on sequence similarity, whereas their propensity to form secondary and tertiary structures, i.e., their ability to form $\beta$-strands, $\beta$-sheets, and turns, was determined by their environment and their potential to hydrogen bond. For example, D380 and D381 of $\beta$2-1 hydrogen bonds with R115 of $\beta$1-5 and K108 of helix B, respectively; D384 in the $\beta$2 turn hydrogen bonds with K99 of $\beta$1-2; and E379 in $\beta$2-1 with K390 in $\beta$2-2. Finally, in the C-terminus is a conserved coil, termed the "meander," that does not possess a high sequence identity among the three known structures but does possess a high structural similarity. For P450arom, we chose to pair the P450arom sequence with that of P450BM-P starting from the highly conserved heme-binding region and going N-terminal to the K′-helix.

Unfortunately, the N-terminal portion of P450s has fewer conserved residues and therefore is more difficult to align. It is known from the crystal structures of P450cam, P450terp, and P450BM-P that the N-terminal region encodes a large portion of the sequences involved in substrate access to, and substrate binding in, the buried active site/heme pocket. Therefore, the alignment of the N-terminus was started from the structurally conserved I-helix, with structural elements being defined initially by making use of the sequence similarity of P450arom to P450BM-P, then by virtue of the amphipathic nature of the helices, and finally by the potential for turns and by helix capping as a result of the presence of P or G, and less frequently D, N, S, or T (Richardson & Richardson, 1988; MacArthur & Thornton, 1991; Hutchinson & Thornton, 1994). The most difficult

regions were identified and aligned last. As mentioned above, one of the signposts for alignment in the N-terminus is the G-helix, which is an amphipathic helix containing at least two basic residues at the C-terminus and with strategically located acidic residues toward the center. The N-terminus of the helix is more difficult to define because the N-capping is not as clear. It appears that there may be a turn at S-W-L; therefore, L was arbitrarily chosen as the beginning of the helix. The F-G loop is divergent except that all four P450s contain a turn defined in part by a P and a D. The F helix of P450arom was started at P-L-D-E because of its potential to make a turn; specifically it was started with the D because this is the case in P450terp and P450cam. The helix was ended with L but may contain an additional L-I. The length was chosen by its $i + 3/i + 4$ nature. Generally, the helices in the N-terminus were chosen using this type of logic. Another guide is the presence of a positively charged residue in the C-helix that interacts with one of the heme propionates.

Because of the diversity in sequence and length, some of the most difficult portions to align are the B-, B′-, and C-helices and their loops, namely the B–B′ and B′–C loops, which interact with the substrate and therefore are highly variable. Fortunately, because of the 30% similarity and 18% identity of P450arom with P450BM-P (high among P450s from different gene families), the alignment of P450arom has been made easier; however, this does not necessarily mean that the helices can be superimposed in three-dimensional space.

*The core structure*

Generally, in the three P450 proteins whose structure is now known, one finds that they all have two domains: a predominately $\alpha$-helical domain, accounting for approximately 70% of the protein and a predominately $\beta$-sheet domain, accounting for 22% of the protein. When the crystal structures of the three proteins are superimposed by overlaying the heme groups, one can identify a conserved "core structure" for P450s. The most highly conserved structures are a four-helix bundle consisting of helices D, I, and L, and the anti-parallel helix E, helices J and K, $\beta$-sheets 1 and 2, the heme-binding region, and the "meander" just N-terminal of the heme-binding region, as is shown in Figure 3 and Kinemage 1. These conserved structures may in fact play a role in folding and certainly in heme-binding. We have used this "core structure" as a basis for building P450arom.

*Computer modeling*

Modeling of P450arom was conducted using the program InsightII from Biosym on a Silicon Graphics workstation. Gen-

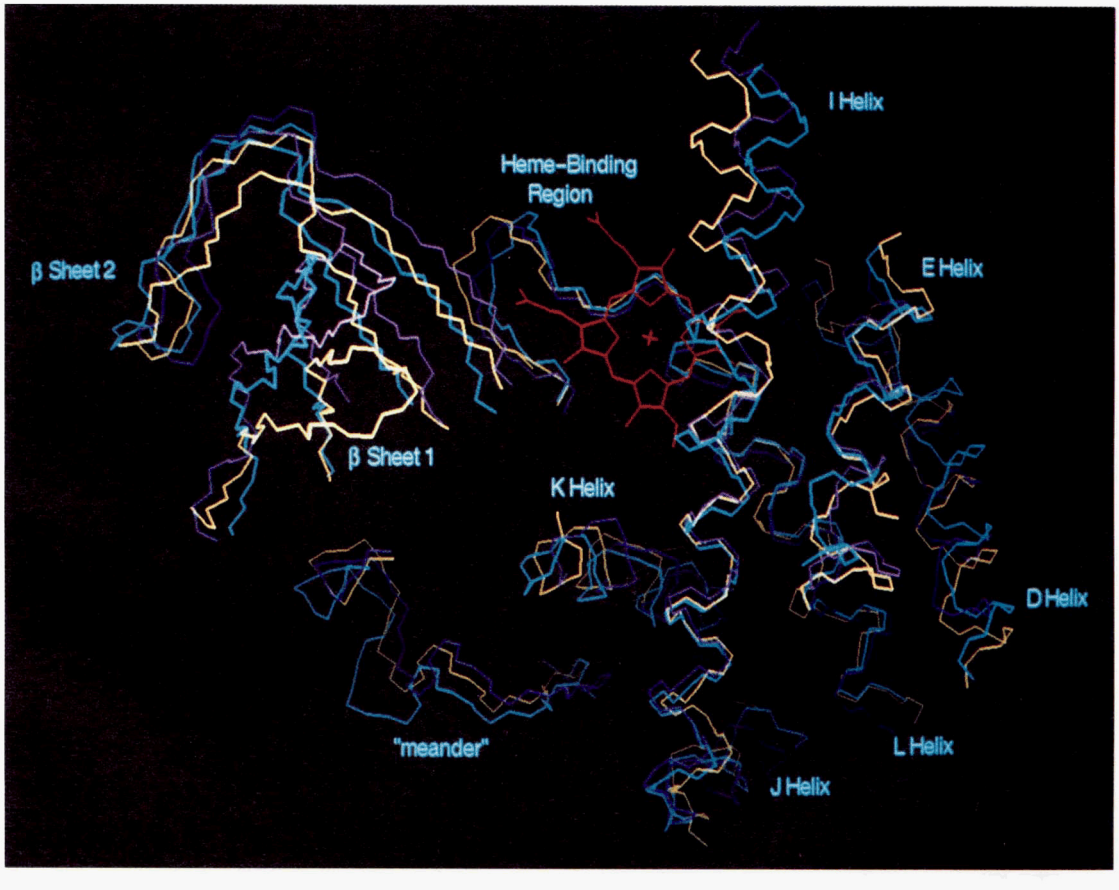


**Fig. 3.** Core structures of P450s. Core structures of P450cam (amber), P450terp (purple), and P450BM-P (blue) are overlaid and shown along with the heme (red). The four-helix bundle on the right is composed of the I-, L-, D-, and E-helices with the heme-binding region and "meander" extending from the N-terminus of the L-helix. Conserved $\beta$-sheets on the left are composed of $\beta$-strands 1 and 2.

erally, the model of P450arom was not constructed by molecular replacement of P450BM-P residues; rather a large portion of it was generated by building a structure de novo. The strategy was to build the most structurally conserved regions first, i.e., the core structure, then "fit in" the less spatially conserved structural elements in a complementary fashion, and finally build the loops, if possible similar to those of one of the three structurally known P450s; however, the membrane-spanning region was not included in this exercise. The completed model is shown in Figure 4 and Kinemage 2. Comparison with the experimentally determined structures of P450BM-P (Ravichandran et al., 1993) and P450cam (Poulos et al., 1987) reveals that they are all strikingly similar.

*The conserved regions*

From the alignments, the four helices in the conserved four-helix bundle of P450arom were identified, individually constructed, and finally overlaid on the corresponding peptide backbone of P450BM-P. The J-helix was also built at this time because of its sequence homology to P450BM-P. These were then minimized using the Discover program from Biosym to remove any side-chain bumping. Secondly, the conserved sheets in the $\beta$-sheet domain were constructed. In all three known structures, the $\beta$1- and $\beta$2-sheet structures are highly conserved and, because of the difficulty in building sheets, these were constructed by molecular replacement of P450BM-P; however, at the turn between $\beta$2-1 and $\beta$2-2, there appeared to be a four-residue turn rather than a three-residue turn. This was accommodated in the structure. In this assemblage, the less conserved A-helix and B-helix were also constructed at the same time for logistic purposes but were constructed de novo. This structure was then minimized. Thirdly, the heme and the heme-binding region were added. On subsequent minimizations, the conserved cysteine and the heme were held fixed.

*The less conserved regions*

The J′-helix and the more conserved K-helix were then built and added using molecular replacement. The region from, and containing, the K′-helix extending to the heme binding region, i.e., the "meander," which is a structurally conserved random coil, was then also built by molecular replacement of the corresponding regions. Both of these regions contain inserts in the sequence alignments of P450BM-P with respect to P450cam and P450terp and correspond to additional structures in P450BM-P. These inserts are most likely highly conserved among Class II P450s, which require only NADPH reductase as a redox partner as determined from sequence alignments.

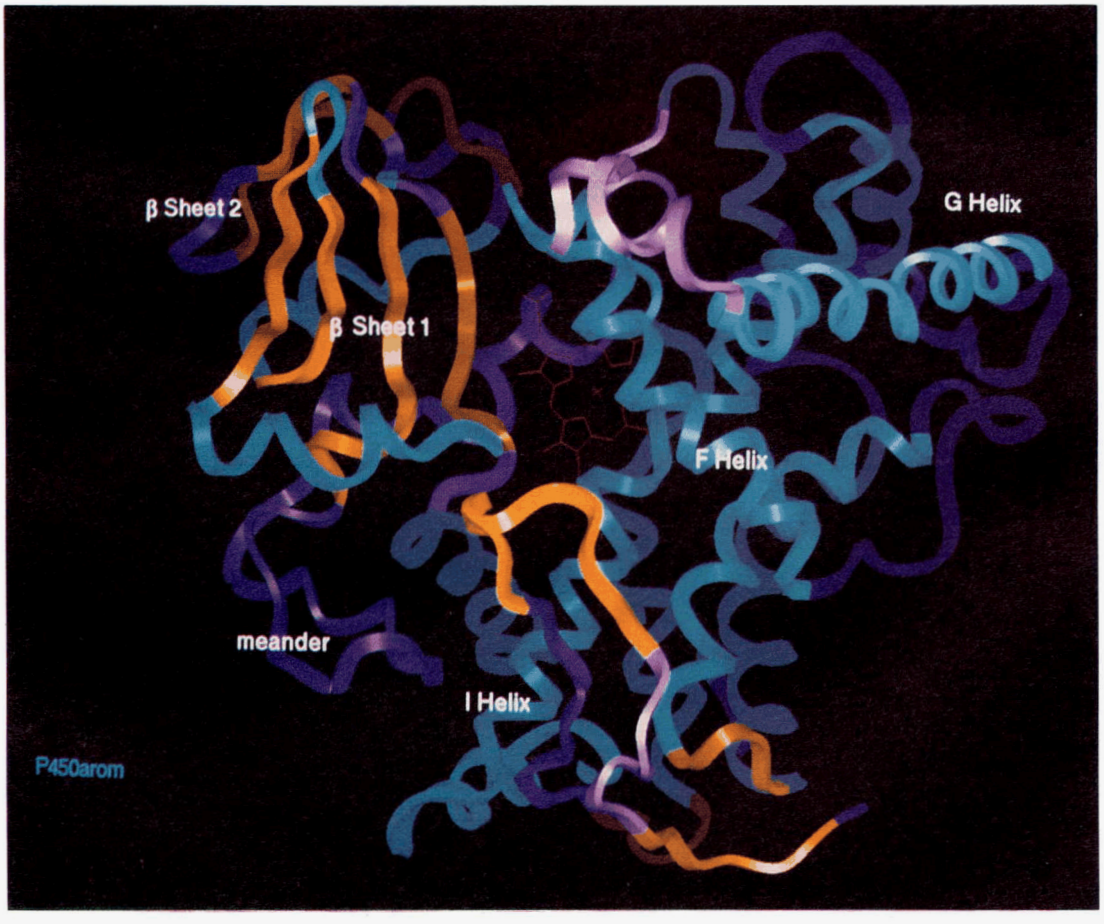


**Fig. 4.** Modeled structure of P450arom viewed from the side distal to the redox partner binding region. In this ribbon diagram, the $\beta$-sheets are in orange, the $\alpha$-helices are in blue, and the loops are in purple.

*The least conserved regions and the loops*

The least conserved regions were constructed and positioned last. Initially, it was hoped that the loop program of Insight would allow us to find structures representative of those loop structures in P450arom; however, it has become obvious with modeling that the best representative for P450arom is P450BM-P because of the sequence similarities. Therefore, replacements were conducted usually using P450BM-P as a template. Thus, the E–F loop (the E helix to F helix loop), the C–D loop, and the J–J′ loop were constructed using a combination of molecular replacement and addition of residues to the carboxylate group followed by adjustments in the torsion angles. The F–G loop, however, was generated without a template in a more random fashion because it was found by Haseman et al. (1994) and Ravichandran et al. (1993) that, in solving the crystal structure, there was a large temperature factor for this loop. That is, it is a random loop in P450BM-P, and in fact, this loop could not be defined in P450terp.

The region from the G-helix to the I-helix containing the H-helix was built by generally following the peptide backbone of P450BM-P; however, there is an extra appendix-like structure composed of six residues protruding from one side of the I-helix just N-terminal of the H-helix in P450arom.

The region from the B-helix to the C–D loop was built first by replacement of the B–B′ loop, then by the addition of the helical structure corresponding to the B′-helix. The B′–C loop, which is located in the active site, was built so that the hydrophobic residues, I132-I133-F134, are either adjacent to other hydrophobic residues or protruding into the active site. The N135-N136-N137 sequence was assumed to be in a coil-type structure to allow for the appropriate turns to join up with the C-helix, and the E129-K130-G131 sequence was built so that there is charge pairing and turned so that it is shielded from hydrophobic residues. In P450cam, the B′ to C region loops in and out of the active site three times; however, in P450BM-P, and by analogy in P450arom, not enough of the residues commonly found in turns are present to permit this type of weaving of the active site. Finally, the C-helix was built and placed in a location similar to P450BM-P and joined with the C–D loop.

The most difficult and speculative regions are the $\beta$-sheets at the C-terminal end of the protein containing the $\beta$3- and $\beta$4-sheets. In the three proteins thus far analyzed, neither of these overlay in a similar fashion, although they are in a similar location. To construct this 35-residue loop, which in the three known structures is in the form of these two $\beta$-sheets, we first determined which regions appear to be juxtaposed in the sheet by charge pairing or hydrophobic interactions and sequence alignment. Based on this, each strand was built residue by residue and adjusted for the environment and the complementary antiparallel strand, trying to maintain the shortest distance between the beginning of $\beta$3-1 and the $\beta$4 turn in the active site. We believe this to be the region for which the structural modeling is the least reliable.

*Minimization and molecular dynamics simulations*

During the construction of the model, minimizations were carried out to relieve steric hindrance; however, this did not seem to optimize the structure in certain loops. We therefore carried out molecular dynamics simulations for 2, 4, and 6 ps at 300 K in vacuum followed by minimizations to further refine the model. In these simulations, we placed androstenedione in the active site in four different orientations, with the C19 methyl in approximately the same starting position. Two of these were almost planar with the heme and parallel to the I-helix. In the third orientation, the steroid was perpendicular to the position of the first and is shown in Figure 5 and Kinemage 3. The fourth orientation of androstenedione was rotated 90° on the long axis as if it were standing on end. After the dynamics simulations, the first and third orientations maintained the C19 of androstenedione close enough to the oxygen and heme iron to potentially allow hydroxylation. Of these, the first orientation was the planar orientation with the steroid A ring close to E302 as had been proposed previously (Graham-Lorence et al., 1991). In the orientation shown in Figure 5, the C1 and C2 of the A ring of androstenedione are close to D309, with the C19 methyl equidistant between the T310 and the molecular oxygen that is bound to the heme iron. The steroid D ring is almost coplanar with H128, suggesting that the C17-keto group may hydrogen bond to one of the imidazole nitrogens. Furthermore, both K473 and H475 in $\beta$-sheet 4 are relatively near the C3-keto group and the C4 position, and incidentally, both H128 and K473 are at the bottom of the access channel and must be passed to enter the active site.

To evaluate this model after dynamics, both Procheck (Laskowski et al., 1993) and 3-D Profile (Luthy et al., 1992) were used for analysis. After 6 ps of dynamics, 65.5% of the residues were in the "most favored regions" in Ramachandran plots, and 27.2% were in "additional allowed regions"; thus, a total of 93% of the residues of the modeled structure after 6 ps of dynamics were in the allowed regions. Although greater than 90% of residues in a crystal structure are expected to be in the most favored regions, in a modeled structure built de novo, the values obtained appear reasonable. Residues scored in disallowed regions were generally found at the beginning of helices or $\beta$-strands and in the G–H loop, the H–I loop, the B′–C loop (where we believe there to be charge pairing), and in the meander. In 3-D Profile (Luthy et al., 1992), the individual residues are characterized by their environment as determined from the area of buried residues, the fraction of side-chain area that is covered by polar atoms, and the local secondary structure. The initial structures and the structures from simulations for up to 6 ps at 300 K fell within limits of an acceptable score similar to that of P450BM-P. With increasing times of simulation, the values first fell further from the optimum value and then returned to acceptable values, similar to that found for P450BM-P.

In earlier test simulations, we found that at 700 K the heme-binding region was unstable, causing a great deal of movement in the protein generally, and specifically of the heme in the active site. Therefore, in simulations longer than 6 ps or higher than 300 K, the heme iron was tethered to the molecular oxygen and the molecular oxygen was tethered to the conserved T310. Figure 6 shows the RMS difference for the starting coordinates and for simulations of 2, 4, 6, 8, 10, 12, 14, 16, 18, 60, and 120 ps. As can be seen, the greatest amount of movement of the peptide backbone was in the loops between helices C and D, F and G, and G and H, as well as $\beta$-strands 1-1 and 1-2; yet, there was relatively little movement in regions associated with the active site (B′–C loop, the I-helix, $\beta$1-4, etc.) or redox-partner binding. With time, the B′–C, C–D, G–H, and H–I loops attained minima. The increased movement in the $\beta$-sheet 1-1 and 1-2 and the F–G loop regions are due, we believe, to a

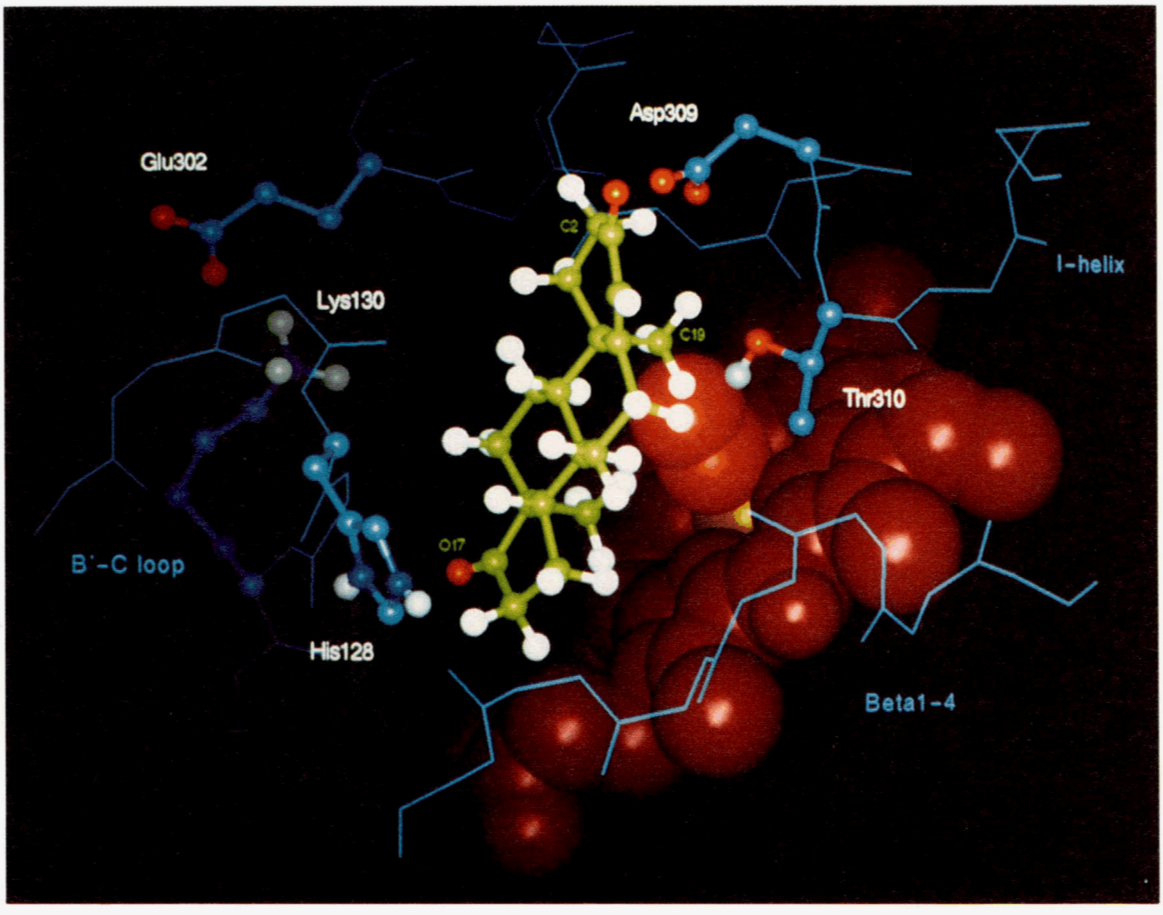


**Fig. 5.** Active site model of P450arom after 6 ps of dynamic simulation at 300 K in vacuum followed by minimization. In this model, the androstenedione is in green and the heme is in red in a CPK representation. The conserved T310, along with D309 and E302 are indicated on the I-helix, and H128 and K130 on the B′–C loop in ball-and-stick representations. Oxygens in the side chains of these residues are represented in red, nitrogens in dark blue, and hydrogens in gray.

jaw-like motion associated with the mouth of the access channel. Even after 120 ps of simulation at 300 K, the tertiary structure of these regions still fluctuates, whereas the loops between G and H and between C and D appeared to reach an optimum confirmation. We conclude that all elements involved in the overall framework of this P450 are in a reasonable orientation with the exception of $\beta$1-1, $\beta$1-2, and the F–G loop region, which may move normally to accommodate substrate binding; otherwise, the fluctuations in the structure have reached a minimum.

Molecular dynamics simulations using slow cooling from 1,000 K to 0 K were also attempted and displayed similar movement to the 300 K runs, especially the F–G loop and $\beta$-strands 1-1 and 1-2. We compared these results with dynamic simulations of P450BM-P and found similar RMS differences for P450BM-P in the F–G loop on short dynamics simulations of 6 ps at 700 K (not shown). This type of movement was also seen by Ornstein and colleagues with P450BM-P after 200 ps of simulation at 300 K (Paulsen & Ornstein, 1995). Because this region is obviously flexible in P450BM-P and P450terp (the F–G loop could not be determined in P450terp because of the high *B* factor) and may also be so in other P450s, and because this region appears to be involved in substrate binding, one may propose that movement in these two regions may have important consequences, i.e., conformational changes may regulate substrate binding.

## *Tertiary features of the model and implications for the catalytic mechanism*

### *The substrate access channel*

It is believed from the crystal structures of P450terp and P450BM-P that the hydrophobic substrates are recognized on the surface of the protein by a hydrophobic patch of amino acids adjacent to the substrate access channel (Ravichandran et al., 1993; Haseman et al., 1994). In P450arom, this would correspond to residues M85, V87, W88, and I89 in the outermost strand of the access channel, $\beta$1-1 (Fig. 7). The substrate then enters the mouth of the access channel, which is formed by the F–G loop (above) and $\beta$-sheet 1 (below), and partitions down into the active site as a result of hydrophobic interactions. As discussed above, in the molecular dynamics simulations of P450arom at 300 K and using slow cooling from 1,000 K to 0 K, one of the most flexible regions in the model was the F–G loop region that aligned itself with the $\beta$-strands 1-1 and 1-2, as seen

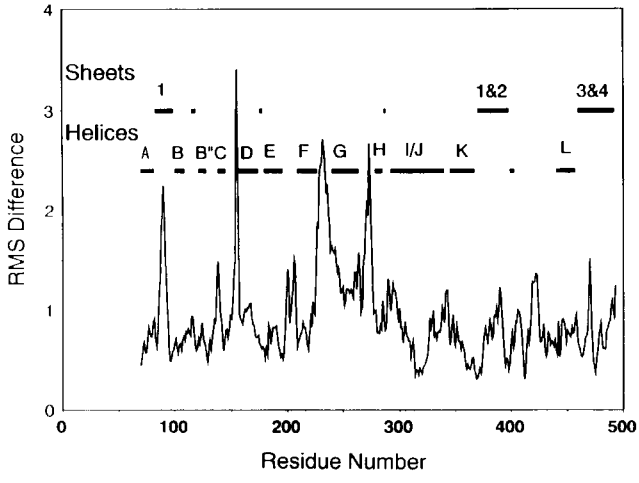


**Fig. 6.** Plot of the RMS difference of the P450arom coordinates after dynamics simulation at 300 K in vacuum were calculated for 2, 4, 6, 8, 10, 12, 14, 16, 18, 60, and 120 ps. Also designated in these plots are the putative $\alpha$-helices and $\beta$-strands for reference. The $x$-axis corresponds to the linear amino acid sequence of the protein; the $y$-axis is the RMS difference in angstroms. The largest movement is seen in the loops between helices C and D, F and G, and G and H, as well as $\beta$-strands 1-1 and 1-2. There was relatively little movement in regions associated with the active site (e.g., the B′–C loop, the I-helix, $\beta$1-4).

in Figure 7. The residues that appear to interact are I89 and S90 at the $\beta$-turn and L227, L228, and I229 in the F–G loop. This indicates that there may be some type of modulation or regulation of substrate binding at the access channel due to conformational changes caused, for example, by the lipid content or fluidity of the associated membrane, or by temperature as described by French et al. (1980) for $P450_{LM2}$. Also, within the F–G loop and the adjacent helical sequences, there are eight aromatic residues, a higher concentration than anywhere else in the primary sequence except the heme-binding region. In Figure 7, these residues appear to align with a cluster of aromatic residues in the $\beta$-sheet domain, possibly forming a lock-and-key fit, which might then be able to occlude the access channel. In mouse CYP2A5, Lindberg and Negishi (1989) have found that, by mutating F209L in the F–G loop, they were able to change substrate specificity from testosterone to coumarin. In our alignments, this aromatic region of P450arom, and in particular Y220, aligns with F209 in CYP2A5, implicating this region as not only involved in membrane association, but in substrate specificity.

Not surprisingly, the access channel in the model of P450arom is shorter than that in P450BM-P because palmitate, one of the substrates for P450BM-P, is considerably longer than the long axis for androstenedione. A path of long-chain hydrophobic amino acids (I96, I395, M374, and L372) runs from the outermost $\beta$ strand to H128 at the entrance to the active site, whereas on the opposite side of the access channel, the eight aromatic residues line the channel starting with Y220. We propose that the aromatic residues on one side and the aliphatic residues on the opposite allow for orientation of the androstenedione molecule as it enters the active site pocket such that the C18 and C19 methyls that protrude from the $\beta$-face are oriented toward the $\beta$-sheets comprised largely of aliphatic residues, and the $\alpha$-face of the steroid is oriented toward the aromatic residues of the F–G loop. The only charged residues are H128 and K473, which appear as a constricture to the active site as seen in Figure 7 (their surface charge is shown in dark blue). We believe H128 may be involved in orientation of the substrate in the active site via the 17-keto group, and K473 may be involved in the enolization of the 3-keto group. H128 was mutated to Ala and Gln by Zhou et al. (1992). The modified protein showed little change in $K_m$ using testosterone as a substrate, but there was an approximately 30% increase in $K_m$ with androstenedione as substrate. With each substrate, the $V_{max}$ was approximately 20% that of wild type. The authors concluded that there was no effect of H128 on substrate binding; however, H128 may affect the reaction in an indirect fashion. From this current model, we believe H128 may stabilize the substrate by hydrogen bonding with the 17-keto group of androstenedione or the 17-hydroxyl of testosterone, thus holding the C19 methyl near the heme iron in the active site and increasing the frequency of productive collisions with the Fe-O· monooxygenating species. However, it may not affect substrate binding greatly because that may be controlled largely by hydrophobic interactions in the access channel.

*The substrate-binding pocket or active site*

As indicated in Figure 5 and Kinemage 3, the active site/substrate-binding pocket is bounded at the bottom by the heme, over pyrrole-ring B by the I helix, over ring C by the B′–C loop, over ring D by $\beta$1-4, and over ring A by $\beta$4 (not shown for clarity). Some of the residues that are in the active site of P450arom in our model are similar to those predicted by Laughton et al. (1993) when they modeled the active site of P450arom using P450cam as a template. However, in their model, it was proposed that the B–B′ loop is also in the active site of P450arom as it is in P450cam. We, on the other hand, have not placed this loop in the active site. In the crystal structure of P450BM-P, this loop does not penetrate into the interior of the molecule, most probably because of the lack of residues that can induce turns in loops; thus, it is not able to thread in and then out of the active site. Similarly, in P450arom, there do not appear to be any potential turns in our alignments, and thus we chose to model this loop after P450BM-P.

The B′–C loop appears to play an important role in substrate orientation in those P450s that have been crystallized, and we believe this to be the case here. Because this region is a loop, its structure is more speculative. We have modeled it such that H128, E129, and K130 are located in or adjacent to the active site. With dynamics simulation, E302 in the I-helix, which we originally proposed to play a role in the aromatization of the A ring (Graham-Lorence et al., 1991), appears to be too far from the substrate to participate in acid/base catalysis, but rather appears to have formed a salt bridge with K130, as shown in Figure 5 and Kinemage 3.

The I-helix is largely hydrophobic, cutting across pyrrole ring B in the three P450s that have been crystallized and appearing to contain a kink produced by a G-G pair in P450cam and an A-G pair in P450terp and P450BM-P. In P450arom, the sequence in this potential kink region is A-A-P. Previously, three mutants were constructed (Graham-Lorence et al., 1991): A307G, P308V, and GAGV, which changed -I305-A306-A307-P308- to -G-A-G-V- (the corresponding sequence found in 17$\alpha$-hydroxylase P450). When the former three proteins were expressed in COS1 cells, it was found that the activity of P308V was approx-

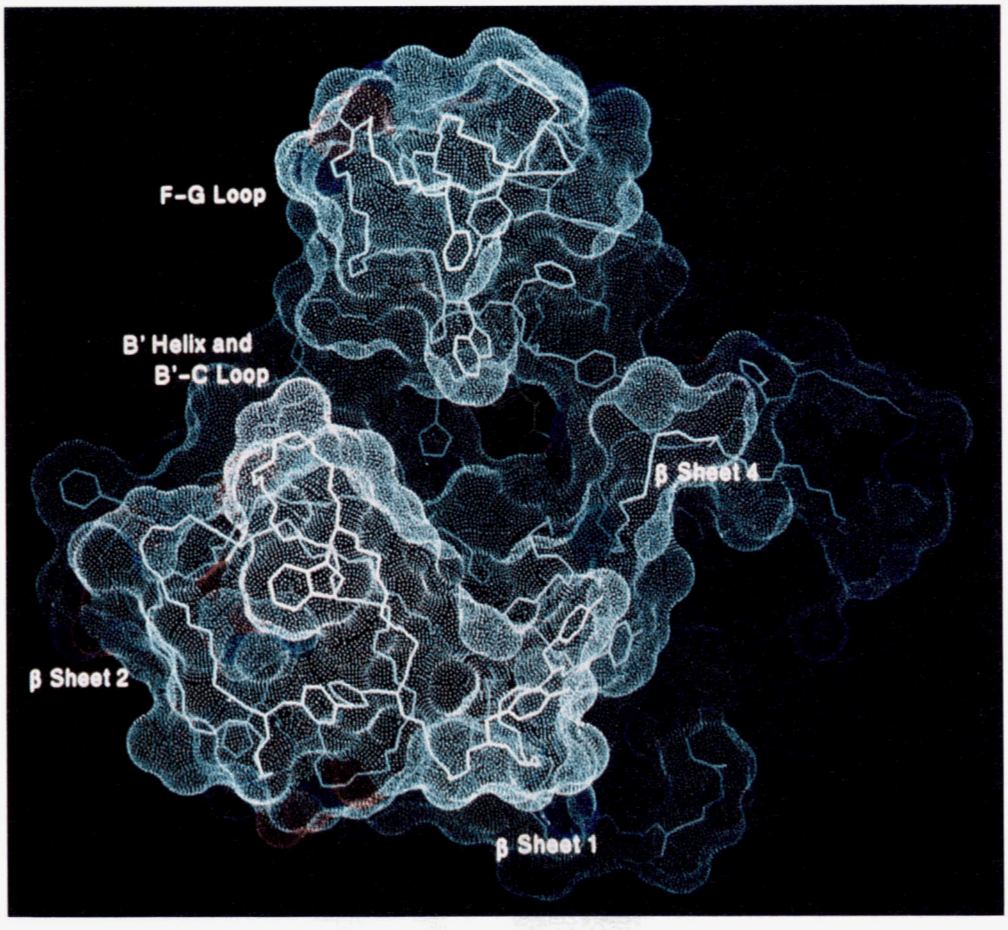


**Fig. 7.** Access channel of P450arom viewed using a solvent-accessible surface. Lower lip of the channel is composed of $\beta$-sheets 1 and 2; upper lip is composed of the F–G loop. The peptide backbone is shown, along with the side chains of the aromatic residues. Surfaces for the side-chain nitrogens of Lys, Arg, and His are shown in blue, and those for the carbonyl oxygens of Asp and Glu in red. The side chain nitrogens and oxygens of Asn, Gln, Thr, and Ser are shown in light blue and pink. Androstenedione is in green and parts of it can be seen through the tiny opening to the active site.

imately one-third that of the wild type, whereas the activities of A307G and the GAGV mutants were little different from that of wild-type P450arom. Presumably the presence of two glycines has a similar effect to the presence of a proline, in terms of ability to induce a bend in the $\alpha$-helix. On the other hand, the GAGV mutant was devoid of 17$\alpha$-hydroxylase activity. These observations are consistent with the concept that P308 causes a bend in the I-helix of P450arom, similar to that observed in P450cam, which is believed to be important in forming the substrate-binding pocket and may play a role in the folding or breathing of the protein.

Directly adjacent to the kink in the I-helix of P450arom is the highly conserved acidic residue D309, proposed to be involved in reprotonation of the conserved threonine (T310) that is involved in oxygen activation (Imai et al., 1989; Raag & Poulos, 1989; Gerber & Sligar, 1994a, 1994b). Both of these residues (shown in ball-and-stick representation in Fig. 5; see also Kinemage 3), when mutated to apolar amino acids in P450arom, produce an essentially inactive protein (Chen & Zhou, 1992; Zhou et al., 1992; Amarneh et al., 1993), implying that both of these residues are important in maintaining catalytic activity. With the new orientation for the substrate in the active site, C1 and C2 are approximately 2 Å from the carboxylate of D309, whereas T310 still remains close to the C19 methyl group.

N-terminal to the I-helix kink is E302, which we originally believed to be involved in the stereospecific abstraction of the C2 hydrogen from the $\beta$-face of androstenedione; however, in the present complete model for P450arom, E302 is too far away from C2 and may form a salt bridge with K130. E302 has been mutated to Ala, Val, Gln, and Asp, and it was found that mutations in which the carboxylic acid was replaced were essentially devoid of activity (Graham-Lorence et al., 1991). On the other hand, changing E302 to Asp resulted in a two-thirds reduction in the apparent $V_{max}$, suggesting that the carboxylate at 302 does play an important role, perhaps by forming a salt bridge with K130 as proposed above or possibly aiding in the reprotonation of D309 and T310.

One other important region bounding the active site is sheet $\beta$4, which protrudes above pyrrole ring A (not shown). In constructing this sheet and $\beta$-sheet 3, it must be remembered that the residues that are located in the active site will be related to the number of residues in a $\beta$-sheet conformation needed to make a path from the L-helix, over the C-terminal portion of the I-helix, up to the active site, and back to the C-terminus. Al-

though in the three crystallized proteins the paths may twist in slightly different ways, they all seem to be relatively distance dependent. Thus, in modeling $\beta 4$ in P450arom, K473-I474-H475-D476 enter the substrate pocket. Zhou et al. (1994) have mutated I474 to Phe and found a difference in inhibitor activity that they believe is a result of steric hindrance in binding of bulky inhibitors. From the present model (Fig. 5; Kinemage 3), it is possible that K473 or H475 may be transiently hydrogen bonded to the 3-keto of androstenedione or testosterone, thus facilitating enolization. Most likely it is K473 because it is conserved across species in P450arom.

*The proximal face of the protein*

Knowledge of the structures of the bacterial P450s has proven to be important in analyzing the residues that interact with the redox partner of P450. A comparison of the alignments of more than 100 microsomal and 17 mitochondrial P450 proteins with the three known structures has implicated a number of residues in docking of these partners (see Fig. 8). Furthermore, there are marked differences between the mitochondrial Class I P450s and microsomal Class II P450s in terms of the residues employed, related to the fact that the former interact with an iron-sulfur protein, and the latter interact directly with a flavoprotein reductase. Very interestingly, this analysis of charge distribution indicates that P450arom falls into neither camp but rather is somewhere in between.

Regions on the proximal face of P450s that appear to be involved in redox-partner binding are those in helices B, C, J, J', K, L, the region C-terminal of the K'-helix, and the heme-binding region. P450BM-P has two insertions in the proximal face relative to P450cam and P450terp, namely the J'-helix and the sequences C-terminal to the K'-helix, as mentioned earlier. In P450arom, we have modeled the one region as a helix and the other as a random coil; however, the J'-helix may in fact be a coil structure. Residues that appear to be involved in redox-partner binding are as follows: In helix B, there is a residue that points into the solvent and is believed to interact with the reductase. In microsomal P450s it has a positive charge, whereas in mitochondrial P450s the charge is either negative or neutral. In the case of P450arom, the residue is H105. In helix C, all P450s have a residue with a positive charge that interacts with a heme propionate group. In P450arom, this is K150. The adjacent C-terminal residue interacts with solvent. In mitochondrial P450s, this residue is neutral, whereas in microsomal forms it has a positive charge. In P450arom, this residue is T151, which is neutral. In helix J, a residue that points into the solvent has a positive charge in microsomal P450s, is neutral in mitochondrial species, and is a glutamate in P450arom (E330). In mitochondrial P450s, the following two residues are positively charged, whereas in microsomal P450s, they are neutral. In P450arom, they are neutral (A331; I332). In the K-helix, mitochondrial species have a positively charged residue followed four residues later by another positively charged residue. In microsomal forms, the first residue is usually negatively charged. In P450arom, the corresponding residues are E357 (negative) and Y361 (neutral). Finally, in mitochondrial P450s, the L-helix always begins with a positive doublet, whereas in microsomal forms the helix usually starts with a negative charge. In P450arom, there is a positively charged residue followed by a neutral one (K440; Y441). Furthermore, within the L-helix, a positively charged residue is present in mitochondrial P450s, whereas it is negatively charged in microsomal forms. In P450arom, this residue is a methionine (M444). Thus, the charge distribution of P450arom does not fit either the distribution seen in Class I or Class II P450s. The model, however, does show a hydrophobic center region over the heme ring as seen in P450BM-P (Fig. 9), which is different from P450cam, and these regions appear to be stable after dynamics simulations and minimizations.

As indicated, a number of residues in P450arom have been implicated as interacting with the redox partner on the basis of the linear alignments. Curiously, as a group, these are consensus residues neither for Class I nor Class II redox partner interactions, but rather are a combination of both, an observation that may be related to the anomalous subcellular distribution in which P450arom sediments with both mitochondrial and microsomal fractions (unpubl. obs.). Thus, one is left asking the

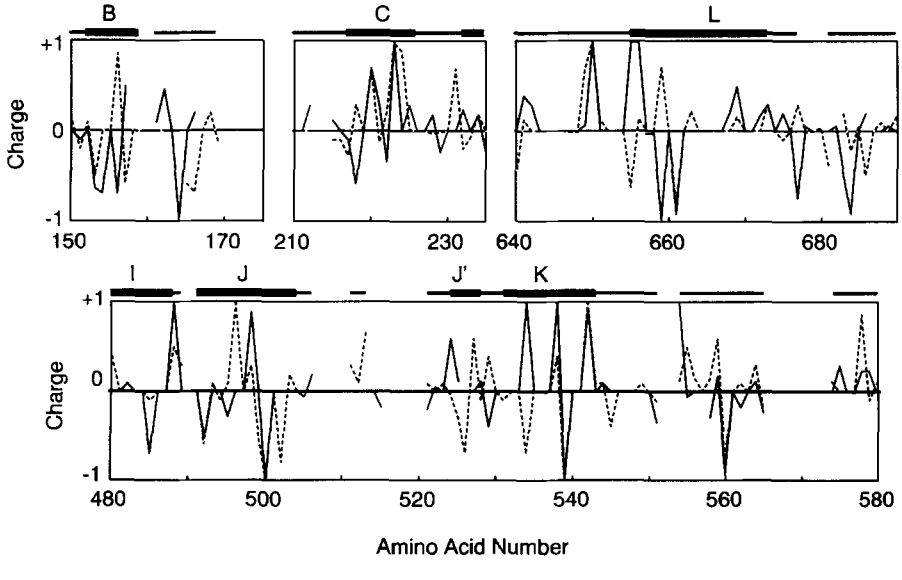


**Fig. 8.** Plot of the average charge per residue in regions in microsomal and mitochondrial P450s believed to be located on the proximal face of the proteins. Arginines and lysines were assigned a charge of +1, glutamic and aspartic acid were assigned a charge of −1, and histidine, a charge of +0.5. The charge distribution on mitochondrial P450s is shown by a dashed line, and on microsomal P450s by a solid line. Secondary structural elements of P450BM-P are shown as reference.

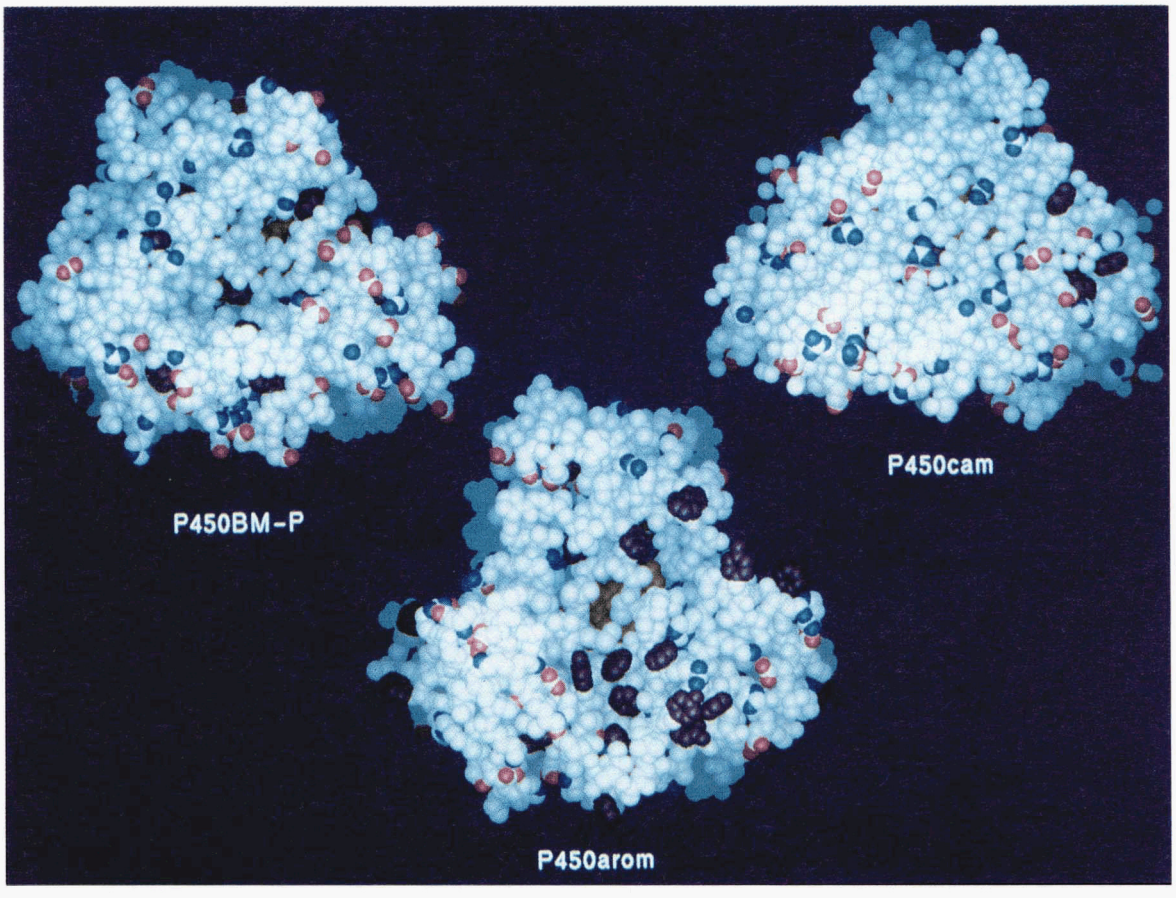


**Fig. 9.** CPK representation of the proximal faces of P450cam, P450BM-P, and P450arom. Point charges for the basic residues are represented in blue; those for the acidic residues are represented in red; and the aromatic rings in purple. The heme is in dark orange with only parts showing through the heme-binding region. Models are shown such that the four-bundle helices are on the left and the conserved β-sheets are on the right. The J- and J′-helices are on the lower left of the models, and the K′-helix is on the lower right. Note that on the upper left corner of P450BM-P, there is an extra loop as compared to P450arom. This loop corresponds to the linker between the P450 and reductase domains of the holoenzyme of P450BM3.

question as to the nature of the redox partner(s) with which P450arom interacts. Unquestionably, P450arom can interact with microsomal NADPH-P450 reductase, as has been shown in reconstitution experiments (Mendelson et al., 1985; Kellis & Vickery, 1987). However, because P450arom may be present in other subcellular sites in addition to the endoplasmic reticulum, it is unclear whether NADPH P450-reductase will prove to be present in all of the same subcellular site(s) as P450arom, and whether it is the optimal electron donor for P450arom.

*Implications for the catalytic mechanism*

As discussed previously, a model for the aromatase reaction mechanism has been proposed by Cole and Robinson (1988) based on an original concept of Akhtar et al. (1993). In this scheme (Fig. 1), whereas the first two sites of attack by molecular oxygen involve sequential hydroxylations on the C19 angular methyl group, the third step involves a P450-peroxide intermediate, which also attacks the C19 position, combined with 1β-hydrogen elimination to give a 1,10-double bond with loss of the C19 carbon atom as formic acid (Fig. 1).

Taking into account the proposed mechanism and combining it with the molecular model that we have constructed, we believe that we can explain the events in the catalytic cycle of P450arom. As shown in Figure 10, the first two hydroxylations at the C19 methyl group are as any other P450-catalyzed hydroxylation. In reaction 1, molecular oxygen is bound by the reduced heme iron, forming a peroxide. It is believed that the conserved threonine (T310) aids in splitting the oxygen–oxygen bond, thus "activating" the oxygen by forming an iron-oxo intermediate (reaction 2). The threonine is immediately reprotonated by the neighboring acidic amino acid (D309) in this scheme. The activated oxygen on the iron will then abstract a hydrogen atom from a proximal carbon on the substrate, forming a carbon radical. Finally, the iron-bound hydroxyl radical will recombine with the carbon radical, resulting in a hydroxylation at that carbon (reaction 3). This occurs twice at the C19 methyl, first forming 19-OH, then 19-oxo androstenedione (reaction 4), with a gem-diol as an intermediate (Fig. 1).

After this sequence of reactions, enolization occurs at the 3-keto group as shown in Figure 11 (reaction 5). This involves the abstraction from the β-face of androstenedione (and either the α- or β-faces of testosterone) of the C2 hydrogen and the donation of a proton to the 3-keto. Originally, we believed that the abstraction of the C2 hydrogen was accomplished by E302;

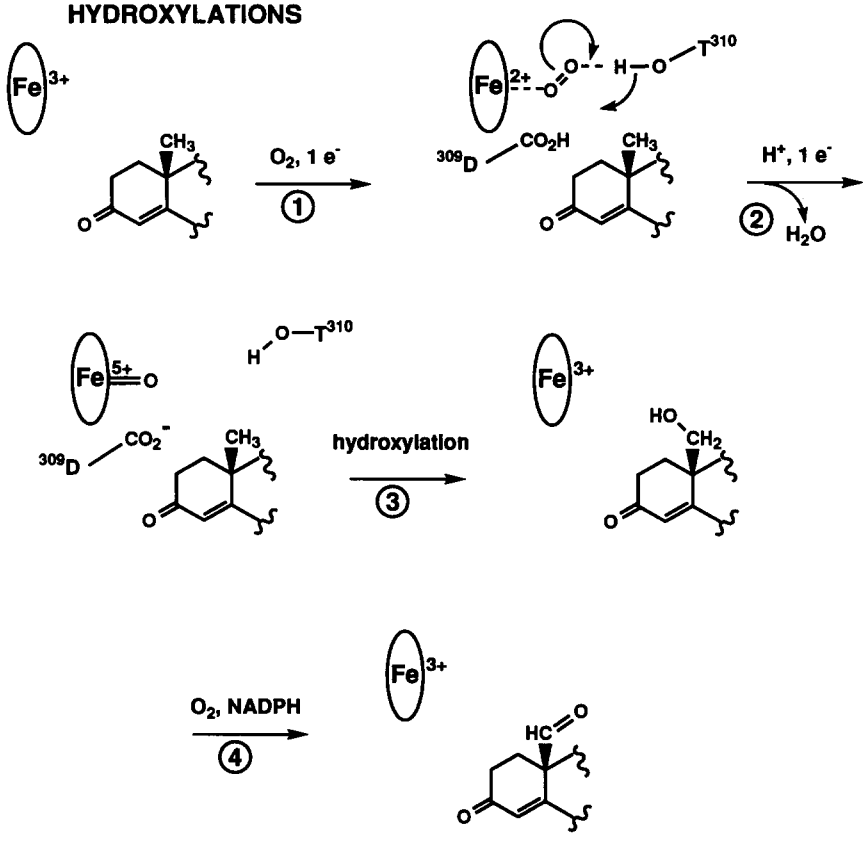


**Fig. 10.** Proposed mechanism of the aromatase reaction: The hydroxylation steps. This scheme shows the first and second hydroxylation steps demonstrating the involvement of D309 and T310 in oxygen activation.

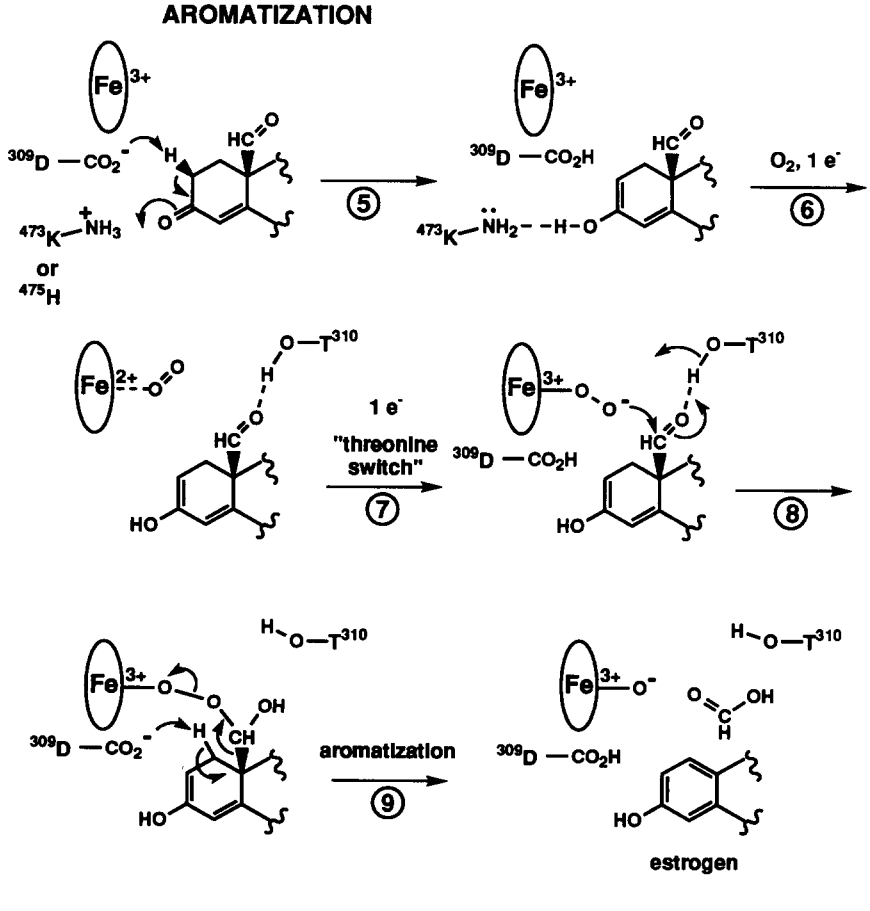


**Fig. 11.** Proposed mechanism of the aromatase reaction: The aromatization step and the "threonine switch." This scheme shows the enolization reaction with D309 abstracting the 2$\beta$-hydrogen and K473 or H475 as possible candidates for proton donation to the 3-keto. The aromatization reaction of the A ring of androstenedione is also shown in which the "threonine switch" promotes the peroxidative attack of C19, resulting in deformylation.

however, in the present model, it became obvious that E302 is too far from the heme-iron by at least one helical turn. We therefore suggest that this abstraction is accomplished by D309. We also suggest that either K473 or H475 from β-sheet 4 donates a proton to the 3-keto group, with K473 being the more likely candidate because it is highly conserved across species in P450arom.

The third oxygen activation step is unusual in that rather than being a hydroxylation, it is a peroxidative attack. The question is why? As shown in reactions 6 and 7 of Figure 11, we propose a "threonine switch" in which the threonine hydrogen bonds to the aldehyde, polarizing the carbonyl double bond, decreasing electron density on the carbon atom, and promoting attack of the peroxide anion. Additionally, because the Fe-O-O$^-$ was not protonated by T310, it is more nucleophilic than it would have been as Fe-O-OH and thus more likely to attack the carbonyl. Finally, during the peroxidative attack of the C19-oxo, the adjacent acidic residue D309 abstracts a proton from the C1 position in reactions 8 and 9, thus initiating deformylation of C19 and aromatization of the A ring. This "threonine switch," as proposed here, may also function in P450c17, in which hydroxylation occurs first at the C17α position of pregnenolone followed by a 17,20-lyase reaction in which acetic acid is released due to a peroxidative attack on the C20 carbonyl (Akhtar et al., 1993).

It should be noted that in porcine P450arom, 19-norandrostenedione is a major product of the reaction (Khalil & Walton, 1985). This would result if the alternative path were taken from the "threonine switch." That is, if the electron transfer during the third round of oxidation were slow in pig as compared to human aromatase (Zhou et al., 1994), or if the substrate were in a slightly different position, then the hydroxyl of T310 may not bond to the aldehyde of 19-oxo-A, but instead hydrogen bond to the Fe-O-O$^-$, allowing the usual P450-type hydroxylation to occur as shown in Figure 12. This third hydroxylation of the C19 would produce 19-oic-A (3,17-dioxo-4-androsten-19-oic acid), which could spontaneously decarboxylate to 19-nor-A. This is similar to the oxidative demethylation reactions that have been shown to be catalyzed by other P450 enzymes, namely lanosterol-14-demethylase (Fischer et al., 1991) and cholesterol-27-demethylation (Andersson et al., 1989; Holmberg-Betsholtz et al., 1993).

*Modeling of vorozole in the active site of P450arom*

Using the P450arom model, we have overlaid both enantiomers of the inhibitor vorozole (DeCoster et al., 1990; Vanden-Bossche et al., 1990) over androstenedione. One of the enantiomers (the *R* form) overlays the androstenedione well and does not appear to act in a fashion consistent with an irreversible inhibitor. However, the *S*-enantiomer, as shown in Figure 13, overlays androstenedione in such a fashion that the imidazole ring that corresponds to the C19 methyl group points toward the heme, allowing one of the nitrogens to covalently bind to the heme iron. This, in fact, is the inhibiting enantiomer. These findings add further validity to the model, and we plan to model other known inhibitors in the future.

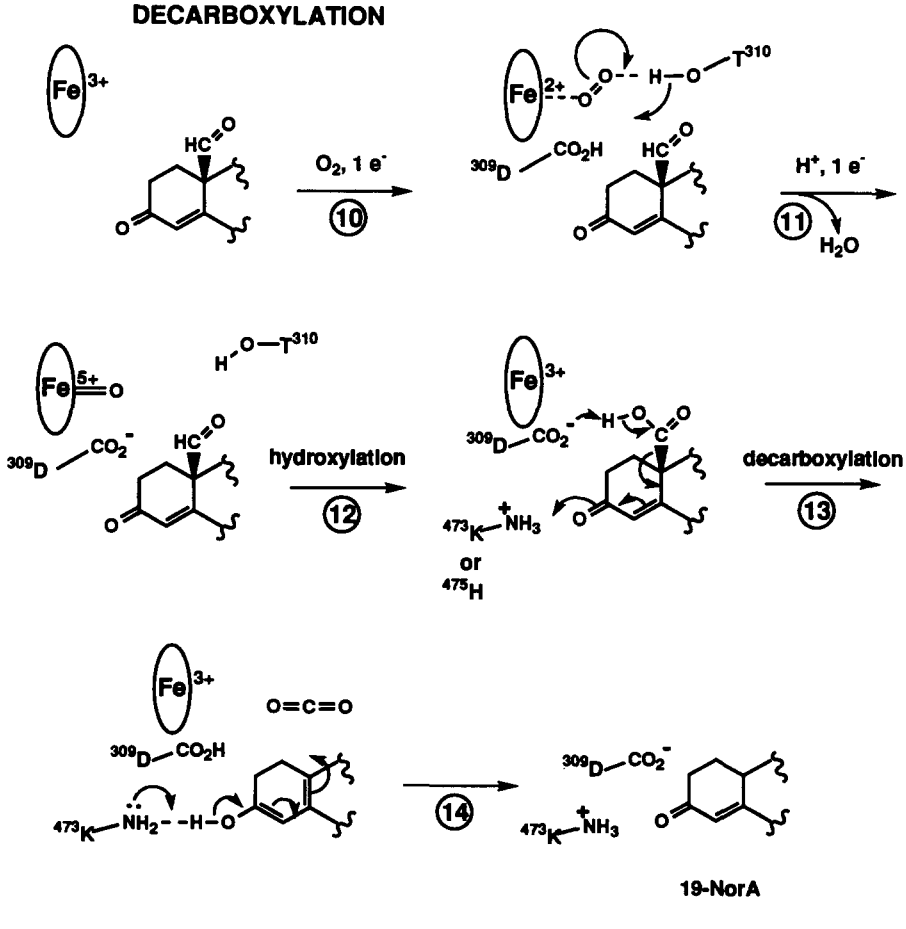


**Fig. 12.** Proposed mechanism of the aromatase reaction: decarboxylation. In an alternative mechanism, the C19 may be hydroxylated a third time, resulting in decarboxylation of androstenedione rather than aromatization of the A ring.

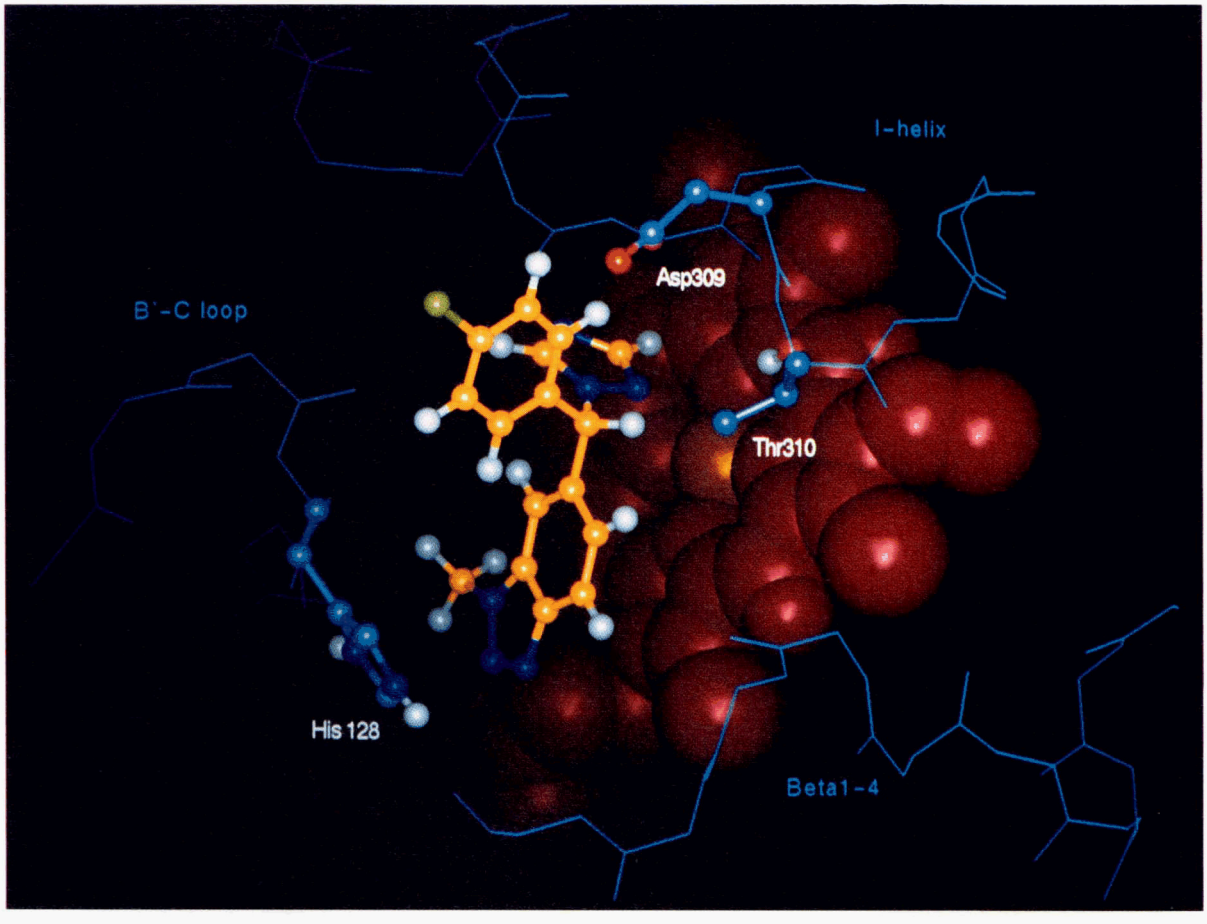


**Fig. 13.** Structure of the *S*-enantiomer of vorozole, a known inhibitor of P450arom, modeled into the active site of P450arom. Vorozole is modeled in such a way as to overlay the chlorophenyl ring over the steroid A ring, the benzotriazol over the C and D rings, and the imidazole over the C19 methyl group of androstenedione.

## Materials and methods

Model building was done using InsightII v.2.3 from Biosym as described above using a Silicon Graphics workstation. Minimizations and dynamic simulations were generally calculated on a DEC Vax Alpha Cluster using X-PLOR, by A. Brünger, in which the conjugate gradient method of Powell (1977) was used, until the minimum energy was below 0.010.

Dynamic simulations were performed iteratively using a temperature bath for the constant temperatures of 300 K or 700 K and with timesteps of 0.001 ps for a minimum total of 2 ps and a maximum total of 70 ps followed by minimization for at least 1,000 steps or until the energy minimum was less than 0.01. The dielectric constant was set to $1/r$, i.e., 1/distance between atoms for all simulations. A slow cooling method for dynamics simulation was employed using temperature coupling and, after 2 ps of simulations at 700 K, was iteratively cooled in 0.001 ps time steps with a decrease of 20 K per step from 1,000 to 0 K, after which it was minimized. A two-step verlet method was used for all simulations.

Evaluation of the RMS differences between the initial model and the simulated structures was accomplished using X-PLOR and the values were imported into Lotus123 on a personal computer and plotted. Profile, described by Luthy et al. (1992), was used on the VAX and plotted on a Declaser printer. Procheck, by Laskowski et al. (1993), was used for the Ramachandran plots.

## Acknowledgments

This work was supported in part by USPHS grants CA 51119 and GM43479, by grant 1228 from the Robert E. Welch Foundation, and by grant 003660-046 from the Texas Higher Education Coordinating Board Advanced Research Program. We thank Hugo VandenBossche and Roland De Costa from Janssen Research Foundation for the crystallographic coordinates of vorozole.

## References

Akhtar M, Calder MR, Corina DL, Wright JN. 1982. Mechanistic studies on C19 demethylation in oestrogen biosynthesis. *Biochem J 201*:569–580.

Akhtar M, Njar VCO, Wright JN. 1993. Mechanistic studies on aromatase and related C-C bond cleaving P450 enzymes. *J Steroid Biochem Mol Biol 44*:375–387.

Amarneh B, Corbin CJ, Peterson JA, Simpson ER, Graham-Lorence S. 1993. Functional domains of human aromatase cytochrome P450 characterized by molecular modelling and site-directed mutagenesis. *Mol Endocrinol 7*:1617–1624.

Andersson S, Davis DL, Dahlback H, Jornvall H, Russell DW. 1989. Clon-

ing, structure and expression of the mitochondrial cytochrome P450 sterol 26-hydroxylase, a bile acid biosynthetic enzyme. *J Biol Chem 264*:8222–8229.

Chen S, Zhou D. 1992. Functional domains of aromatase cytochrome P450 inferred from comparative analysis of amino acid sequences and substantiated by site-directed mutagenesis experiments. *J Biol Chem 267*: 22587–22594.

Cole PA, Robinson CH. 1988. A peroxide model reaction for placental aromatase. *J Am Chem Soc 110*:1284–1285.

DeCoster R, Van Ginckel R, Wouters W, Goeminne N, Vanlerck W, Byloos M. 1990. Endocrine and antitumoral effects of R76713 in rats. *J Enzyme Inhib 4*:159–167.

Fischer RT, Trzakos JM, Magolda RL, Ko SS, Brosz CS, Larsen B. 1991. Lanosterol 14$\alpha$-methyl-demethylase: Isolation and characterization of the third metabolically generated oxidative demethylation intermediate. *J Biol Chem 266*:6124–6132.

French JS, Guengerich FP, Coon MJ. 1980. Interaction of cytochrome P450, NADPH-cytochrome P450 reductase, phospholipid, and substrate in the reconstituted liver microsomal enzyme system. *J Biol Chem 225*:4112–4119.

Gerber NC, Sligar SC. 1994a. Catalytic mechanism of cytochrome P450: Evidence for a charge relay. *J Am Chem Soc 114*:8742–8743.

Gerber NC, Sligar SG. 1994b. A role for Asp-251 in cytochrome P450CAM oxygen activation. *J Biol Chem 269*:4260–4266.

Goto J, Fishman J. 1977. Participation of a non-enzymatic transformation in the biosynthesis of estrogens from androgens. *Science 195*:80–81.

Graham-Lorence S, Khalil MW, Lorence MC, Mendelson CR, Simpson ER. 1991. Structure-function relationships of human aromatase cytochrome P-450 using molecular modeling and site directed mutagenesis. *J Biol Chem 266*:11939–11946.

Haseman CA, Kurumbail RG, Boddupalli SS, Peterson JA, Deisenhofer J. 1995. Structure and function of cytochromes P450: A comparative analysis of three crystal structures. *Structures 3*:41–62.

Haseman CA, Ravichandran KG, Peterson JA, Deisenhofer J. 1994. Crystal structure and refinement of cytochrome P450terp at 2–3 Å resolution. *J Mol Biol 236*:1169–1185.

Holmberg-Betsholtz I, Lund E, Bjorkhem I, Wokvall K. 1993. Sterol 27-hydroxylase in bile acid biosynthesis: Mechanism of 5$\beta$-cholestane-3$\alpha$, 7$\alpha$, 12$\alpha$, 27-tetrol into 3$\alpha$, 7$\alpha$, 12$\alpha$-trihydroxy-5$\beta$-cholestanoic acid. *J Biol Chem 268*:11079–11085.

Hutchinson EG, Thornton JM. 1994. A revised set of potential for $\beta$-turn formation in proteins. *Protein Sci 3*:2207–2216.

Imai M, Shimada H, Watanabe Y, Matsushima-Hibiya Y, Makino R, Koga H, Horiuchi T, Ishimura Y. 1989. Uncoupling of the cytochrome P450cam monooxygenase reaction by a single mutation, threonine-252 to alanine or valine: Possible role of the hydroxy amino acid in oxygen activation. *Proc Natl Acad Sci USA 86*:7823–7827.

Kellis JT Jr, Vickery LE. 1987. Purification and characterization of human placental aromatase cytochrome P450. *J Biol Chem 262*:4413–4420.

Khalil MW, Walton JS. 1985. Identification and measurement of 4-oestren-3,17-dione (19-norandrostenedione) in porcine ovarian follicular fluid using high performance liquid chromatography and capillary gas chromatography-mass spectrometry. *J Endocrinol 107*:375–381.

Korzekwa KR, Trager WF, Mancewicz J, Osawa Y. 1993. Studies on the mechanism of aromatase and other cytochrome P450 mediated deformylation reactions. *J Steroid Biochem Mol Biol 44*:367–373.

Laskowski RA, MacArthur MW, Moss DS, Thornton JM. 1993. PROCHECK: A program to check the stereochemical quality of protein structures *J Appl Crystallogr 26*:283–291.

Laughton CA, Zvelebil MJJM, Neidle S. 1993. A detailed molecular model for human aromatase. *J Steroid Biochem Mol Biol 44*:399–407.

Lindberg RLP, Negishi M. 1989. Alteration of mouse cytochrome P450coh substrate specificity by mutation of a single amino-acid residue. *Nature 339*:632–634.

Luthy R, Bowie JU, Eisenberg D. 1992. Assessment of protein models with three-dimensional profiles. *Nature 356*:83–85.

MacArthur MW, Thornton JM. 1991. Influence of proline residues on protein conformation. *J Mol Biol 218*:397–412.

Mendelson CR, Wright EE, Porter JC, Evans CT, Simpson ER. 1985. Preparation and characterization of polyclonal and monoclonal antibodies against human aromatase cytochrome P-450 (P450arom), and their use in its purification. *Arch Biochem Biophys 243*:480–491.

Needleman SB, Wunsch CD. 1970. A general method applicable to the search for similarities in the amino acid sequence of two proteins. *J Mol Biol 48*:443–453.

Nelson DR, Kamataki T, Waxman DJ, Guengerich FP, Estabrook RW, Feyereisen R, Gonzalez FJ, Coon MJ, Gunsalus IC, Gotoh O, Okuda K, Nebert DW. 1993. The P450 superfamily: Update on new sequences, gene mapping, accession numbers, early trivial names of enzymes, and nomenclature. *DNA Cell Biol 12*:1–51.

Paulsen MD, Ornstein RL. 1995. Dramatic differences in the motions of the mouth of open and closed cytochrome P450BM-3 by molecular dynamics simulations. *Proteins Struct Funct Genet 21*:237–243.

Poulos TL, Finzel BC, Howard AJ. 1987. High resolution crystal structure of cytochrome P450cam. *J Mol Biol 195*:687–700.

Powell MJD. 1977. Restart procedures for conjugate gradient method. *Math Prog 12*:241–254.

Raag R, Poulos TL. 1989. Crystal structure of the carbon monoxide-substrate-cytochrome P450CAM ternary complex. *Biochemistry 28*:7586–7592.

Ravichandran KG, Boddupalli SS, Haseman CA, Peterson JA, Deisenhofer J. 1993. Atomic structure of the hemoprotein domain of cytochrome P450BM3: A prototype for eukaryotic microsomal P450s. *Science 261*: 731–736.

Richardson JS, Richardson DC. 1988. Amino acid preferences for specific locations at the end of alpha helices. *Science 240*:1648–1652.

Thompson EA Jr, Siiteri PK. 1974. The involvement of human placental microsomal cytochrome P450 in aromatization. *J Biol Chem 249*:5373–5378.

VandenBossche H, Williamsens G, Roels I, Bellens D, Moereels D, Coene MC, LeJeune L, Lauwers W, Janssen PAJ. 1990. R76713 and enantiomers: Selective, non-steroidal inhibitors of the cytochrome p450-dependent oestrogen synthesis. *Biochem Pharmacol 40*:1707–1718.

Zhou D, Cain LL, Laughton CA, Korzekwa KR, Chen S. 1994. Mutagenesis study at a postulated hydrophobic region near the active site of aromatase cytochrome p450. *J Biol Chem 269*:19501–19508.

Zhou D, Korzekwa KR, Poulos T, Chen S. 1992. A site-directed mutagenesis study of human placental aromatase. *J Biol Chem 267*:762–768.