

Supporting Information

Structural basis for three-step sequential catalysis by the cholesterol side chain cleavage enzyme CYP11A1

Natalia Mast, Andrew J. Annalora, David T. Lodowski, Krzysztof Palczewski,
C. David Stout, and Irina A. Pikuleva

Crystallographic Methods

The structure was initially solved at 2.8 Å resolution in a trigonal crystal form, space group $P3_22_12$, with cell constants $a=169.61$, $b=169.61$, $c=103.88$ Å and two molecules in the asymmetric unit (a.u.) by molecular replacement using Phaser (S1) and the structure of CYP24A1 in the open form (PDB 3K9V). Iterative model building into 2-fold NCS averaged (S2) $2|Fo|-|Fc|$ electron density maps and molecular replacement with Phaser led to a model with translation function Z score of 39.7. Data to 2.5 Å were collected at SSRL beam line 11-1 using the monoclinic crystal form, space group $P2_1$, with cell constants $a=109.45$, $b=94.63$, $c=113.50$ Å, $\beta=89.96^\circ$, and four molecules/a.u. The data were initially indexed in the orthorhombic space group $P2_12_12_1$ with cell constants $a=94.63$, $b=109.45$, $c=113.50$ Å and two molecules/a.u. This indexing of the data was used to repeat iterative molecular replacement and 2-fold density averaging, and complete the model in register with the sequence of bovine CYP11A1 (residues 6-475 for chains A, B). At this stage the translation function Z score was 80.9. Refinement of the model at 2.5 Å using both CNS (S3) and Refmac (S4) failed to reduce Rfree below 0.35, even though electron density for the entire polypeptide chain was in register with the sequence and density for the heme group and 22HC (not included in the model) was strong. The data were then reprocessed (S5) and scaled (S6) in the $P2_1$ unit cell resulting in reduction in Rmerge from 0.080 to 0.068 (Table S1). Molecular replacement with the dimer from the orthorhombic unit cell (chains A, B) resulted in a translation function Z score of 96.4 for a model with two dimers, AB and CD, in the a.u. related by translation of approximately $\frac{1}{2}$, $\frac{1}{2}$ along the y and z directions, and with the local 2-fold of each dimer approximately parallel with z, but without pure translational symmetry of any subunit along unit cell axes. This model was refined with CNS using simulated annealing and with Refmac to account for partial merohedral twinning (twin operators and fractions, h,k,l, 0.546; -h,-k,l, 0.454) resulting in Rfree of 0.29. Unbiased $2|Fo|-|Fc|$ and $|Fo|-|Fc|$ maps were used to include 22HC, H₂O molecules, and isopropanol in the model, and adjust the polypeptide for each of the four subunits in the a.u. (Table S1). Evidence that the space group is $P2_1$ and pseudo- $P2_12_12_1$ is as follows: a lower Rmerge (0.068 vs. 0.080) for the same data set; a higher Z-score in Phaser (96.4 vs. 80.9); a refined β angle of 89.96° in Mosfilm (S5) using multiple raw data frames; a signal in the data for partial merohedral twinning in $P2_1$ whereas none is detected in $P2_12_12_1$; and significantly lower Rfree at the same stage of refinement (0.29 vs. 0.35). Coordinates and structure factors for the final model have been deposited with the RCSB, accession code 3MZS.

Calculation of the Active Site Volumes and Volumes of the Ligands

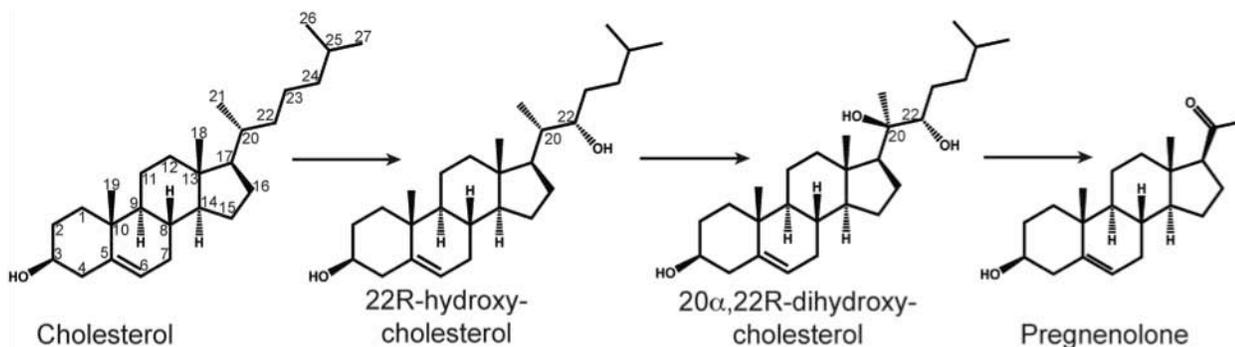
The solvent accessible surface of the active site without the ligand and water molecules was calculated first using VOIDOO (<http://xray.bmc.uu.se/usf/voidoo.html>) (S7) with a 1.4-Å probe and a mesh of 1.0. Then, the cavity was filled out with water molecules using FLOOD (<http://xray.bmc.uu.se/usf/voidoo.html>) and the final volume determined. Volumes of the ligands were calculated by VOIDOO as well using the conformer taken from the crystal structure.

References

- S1.** McCoy, A.J., Grosse-Kunstleve, R.W., Adams, P.D., Winn, M.D., Storoni, L.C., Read, R.J. (2007) *J. Appl. Crystallogr.* **40**, 658-674
- S2.** Cowtan, K.D., Zhang, K.Y.J. (1999) *Prog. Biophys. Mol. Biol.* **72**, 245-270
- S3.** Brunger, A.T., Adams, P.D., Clore, G.M., DeLano, W.L., Gros, P., Grosse-Kunstleve, R.W., Jiang, J.S., Kuszewski, J., Nilges, M., Pannu, N.S., Read, R.J., Rice, L.M., Simonson, T., Warren, G.L. (1998) *Acta Crystallogr.* **D54**, 905-921
- S4.** Murshudov, G.N., Vagin, A.A., Dodson, E.J. (1997) *Acta Crystallogr.* **D53**, 240-255
- S5.** Leslie, A.G. (1999) *Acta Crystallogr.* **D55**, 1696-1702
- S6.** CCP4 suite of crystallographic programs (1994) *Acta Crystallogr.* **D50**, 760-763
- S7.** Kleywert, G.J., Jones, T.A. (1994) *Acta Crystallogr. D Biol. Crystallogr.* **50**, 178-185
- S8.** Guengerich, F.P. (2002) *Biol. Chem.* **383**, 1553-1564
- S9.** Lomize, M.A., Lomize, A.L., Pogozheva, I.D., Mosberg, H.I. (2006) *Bioinformatics* **22**, 623-625
- S10.** Annalora, A.J., Goodin, D.B., Hong, W.X., Zhang, Q., Johnson, E.F., Stout, C.D. (2010) *J. Mol. Biol.* **396**, 441-451

Table S1	
Crystallographic statistics for CYP11A1:22HC complex	
PDB code	3MZS
Space group	P2 ₁
Unit cell dimensions	a=109.450, b=94.630, c=113.500 Å, β=89.96°
Molecules per asymmetric unit	4
Solvent content	53.9%
Data	
Total observations > 0σ _F	151,893
Unique reflections > 0σ _F	76,291
Redundancy	2.0
Completeness	95.5%
Resolution (last shell) (Å)	2.64 – 2.50
<I/σI> all data (last shell)	5.8 (1.6)
Rmerge all data (last shell)	0.068 (0.444)
Refinement	
R-factor	0.266
Rfree	0.281
Reflections used	72,447
Test set	3,812 (5.0%)
RMSD from ideality	
Bond lengths (Å)	0.019
Bond angles (deg.)	1.68
Ramachandran plot	
Favored regions	86.1%
Allowed regions	97.8%
Model	
Copy A / B / C / D	Residues / Avg. B (Å ²)
Protein	470 (65.1) / 470 (67.8) / 470 (65.2) / 470 (66.8)
Heme	1 (43.2) / 1 (47.9) / 1 (49.8) / 1 (41.4)
22HC	1 (50.8) / 1 (56.2) / 1 (55.2) / 1 (49.6)
H ₂ O	92 (57.5)

A



B

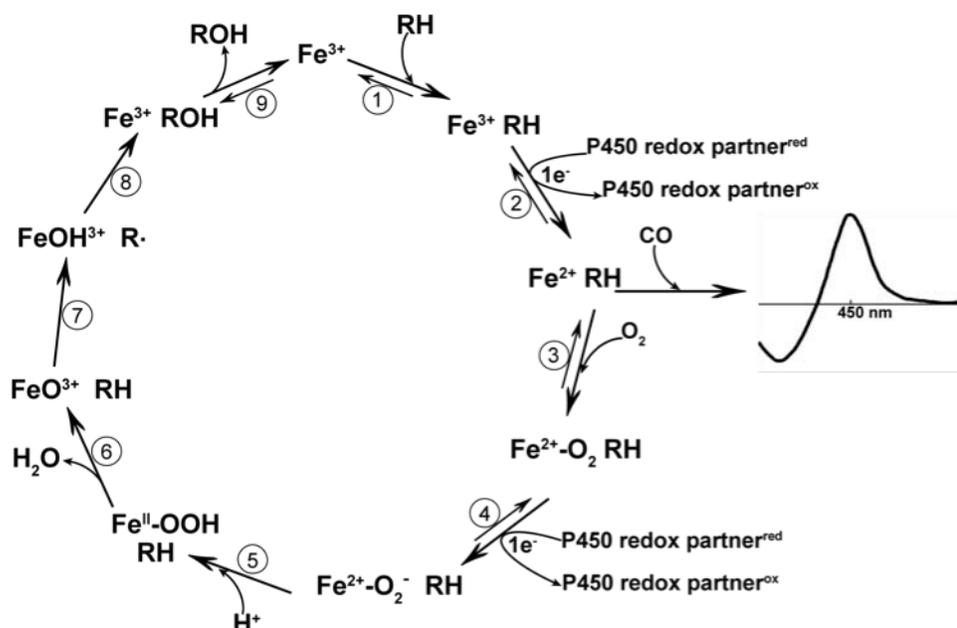


Fig. S1. Sterol intermediates in the conversion of cholesterol to pregnenolone by CYP11A1 (A), and generalized P450 catalytic cycle (B). Three catalytic cycles are needed to produce pregnenolone from cholesterol: one to produce 22R-hydroxycholesterol, one to produce 20 α ,22R-dihydroxycholesterol and one to produce pregnenolone. The presentation of the P450 catalytic cycle is taken from reference S8 (<http://dx.doi.org/10.1515/BC.2002.175>, with permission from De Gruyter) and modified. Fe³⁺ is the iron atom of the heme, RH is the substrate, and ROH is the product. The difference spectrum of the Fe²⁺RH in complex with CO vs. Fe²⁺RH is also shown.

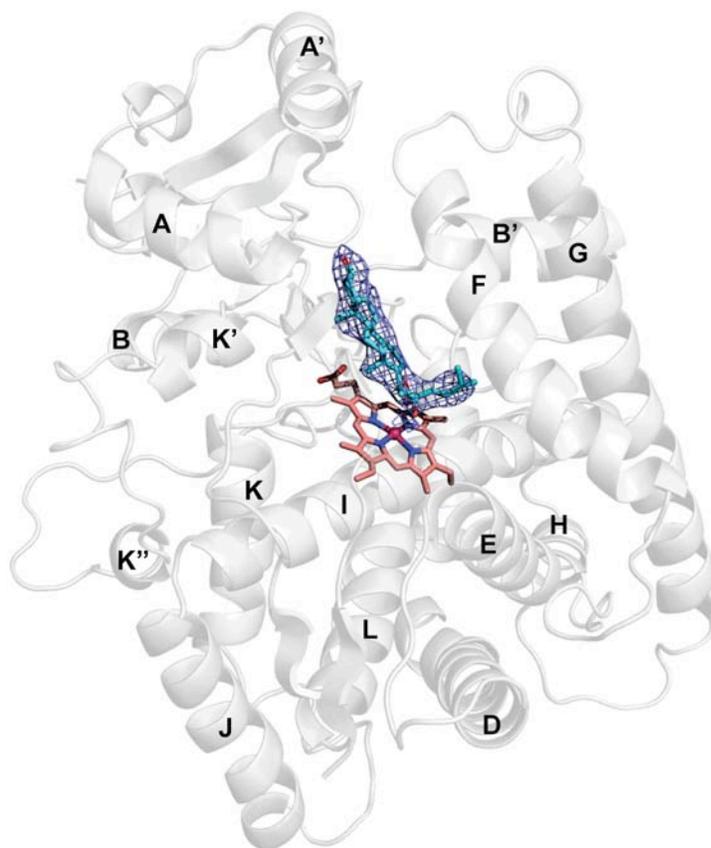


Fig. S2. Overall fold of CYP11A1 with helical elements labeled and unbiased composite omit electron density for 22HC (C atoms cyan) in Copy A contoured at 1σ . C atoms of the heme are pink.

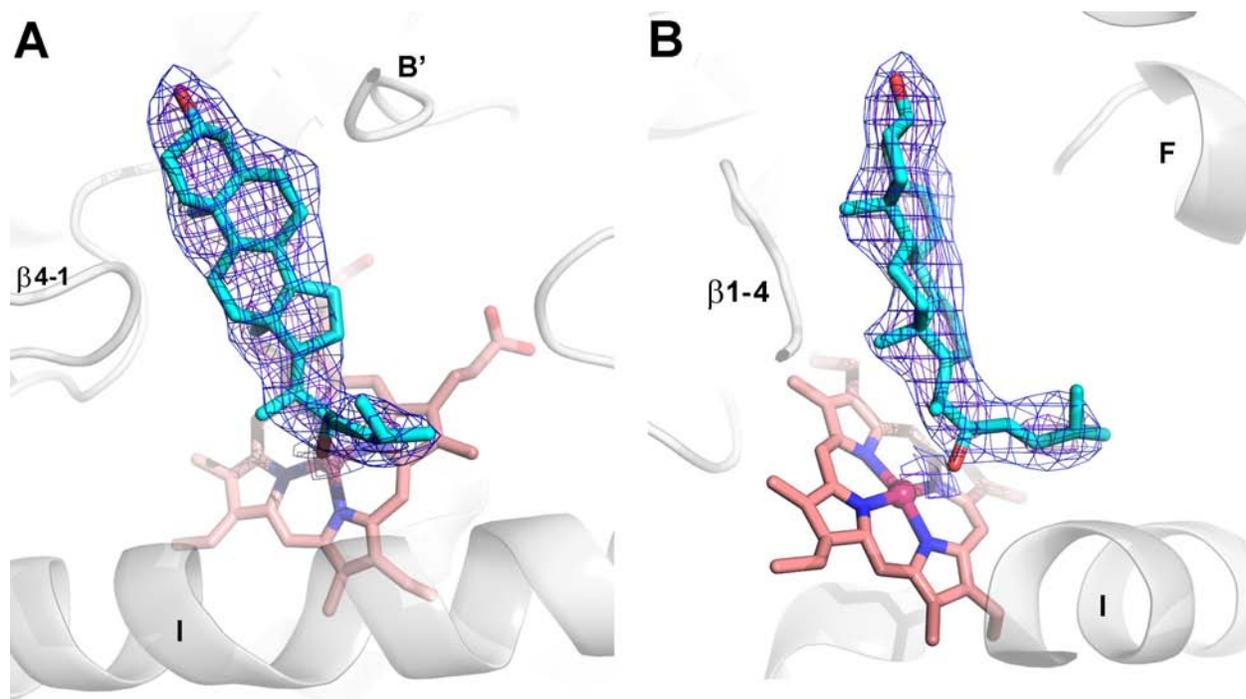


Fig. S3. (A, B) Approximately orthogonal views of the unbiased composite omit electron density at 2.5 Å resolution for 22HC in Copy D of CYP11A1 contoured at 1σ , and 3σ . Elements of the CYP11A1 secondary structure are labeled. The small patch of electron density between the C22 hydroxyl group and Fe is a portion of the large Fe peak; the Fe – O distance is 2.55 Å.

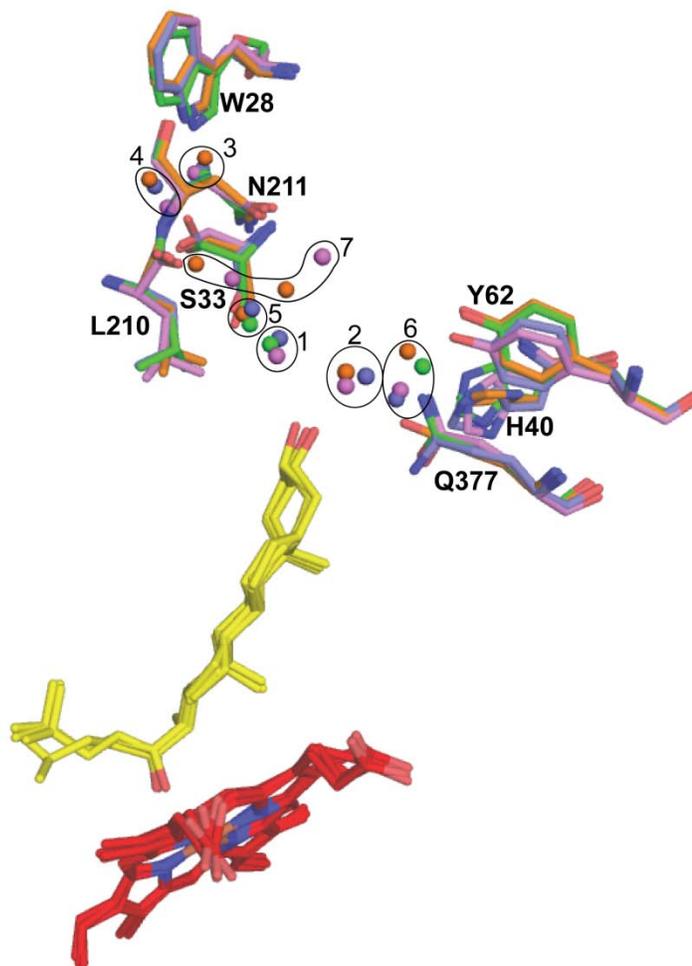


Fig. S4. Ordered water molecules at the entrance to the active site in CYP11A1 (see also Fig. 1B). The water molecules are distributed into clusters when copies A, B, C, and D in the asymmetric unit are superposed. Each cluster is contoured with a black outline and numbered. Protein side chains (shown as sticks) and water molecules (shown as spheres) have the same color in each copy in the asymmetric unit (violet in copy A; green in copy B; slate in copy C; and orange in copy D). Molecules of 22HC and heme are in yellow and dark red, respectively, in all four copies. The nitrogen, oxygen, and iron atoms are in blue, red, and orange, respectively.

Two water clusters, 1 and 2, bracket the 22HC 3 β -hydroxyl: cluster 1 (Wat526, Wat516, and Wat509) is near the side chain of L210 and cluster 2 (Wat523, Wat511, and Wat508) is closer to the side chain of Q377. Clusters 3-5 are distributed between W28 and S33: cluster 3 contains Wat533, Wat539, Wat537, and Wat585; cluster 4 contains Wat553, Wat564, and Wat505; and cluster 5 contains Wat577, Wat531, and Wat568. Cluster 6 (Wat549, Wat589, Wat579, and Wat507) is between the side chains of H40 and Y62. Finally, water molecules 524, 504, 519, and 506 are splayed around the side chain of N211 and form cluster 7.

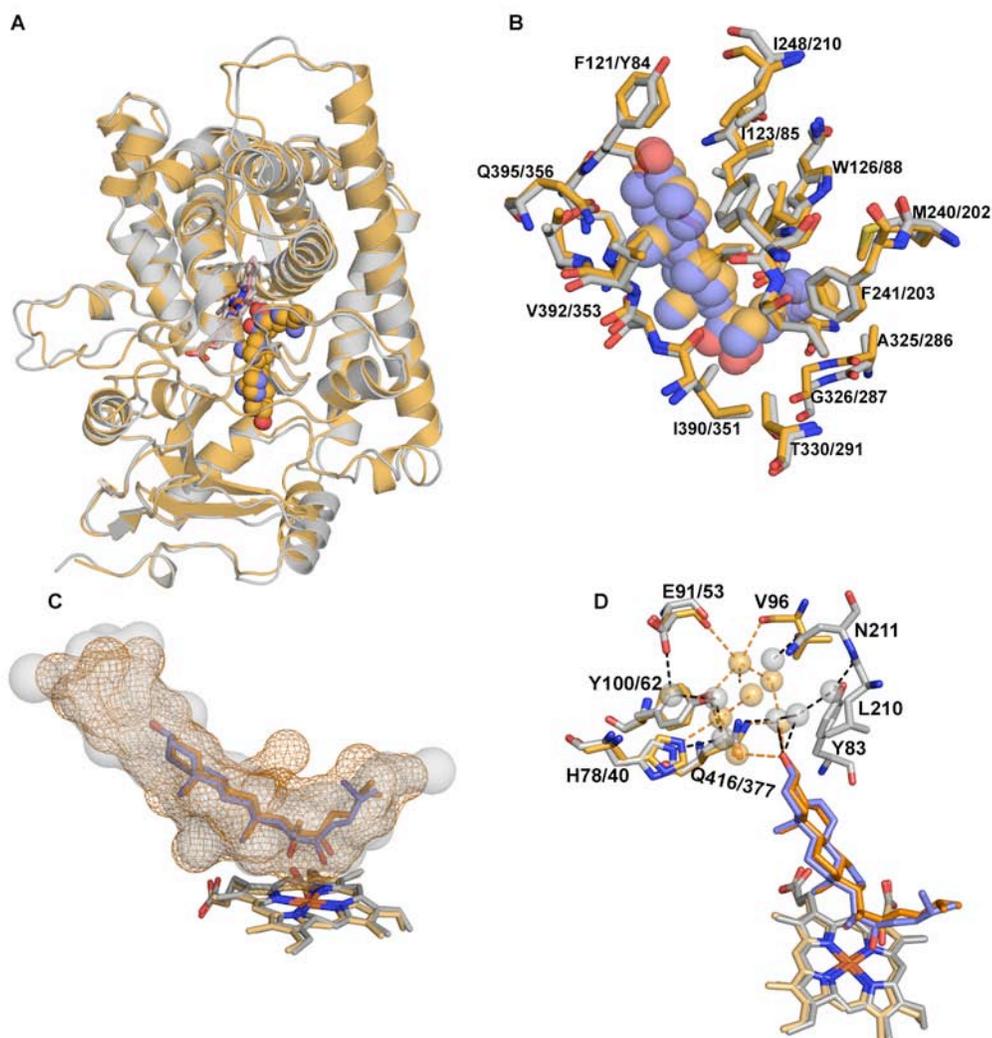


Fig. S5. Comparison of human CYP11A1 bound to 20,22DHC to bovine CYP11A1 bound to 22HC. (A) The overall fold of the two P450s. Human CYP11A1 and 20,22DHC are in orange, and bovine CYP11A1 and 22HC are in gray and in light blue, respectively; the heme group is in salmon in human CYP11A1 and gray in bovine P450. (B) Amino acid residues within 4 Å distance of 20,22DHC in human CYP11A1 and corresponding amino acid residues in the bovine enzyme (human CYP11A1/bovine CYP11A1). Sterols shown as spheres. For clarity, the heme groups are omitted in this panel. (C) The enclosed volume of the active site in human CYP11A1 (orange mesh) and bovine P450 (gray semi-transparent surface). (D) Array of structural water molecules (shown as orange and gray semi-transparent spheres in human and bovine CYP11A1, respectively) hydrogen bonded to the sterols 3 β -hydroxyls and to the residues at the entrance to the active site. For clarity, residues interacting with the sterol backbones are omitted in this panel. The nitrogen, oxygen, sulfur and iron atoms are in dark blue, red, yellow, and orange, respectively.

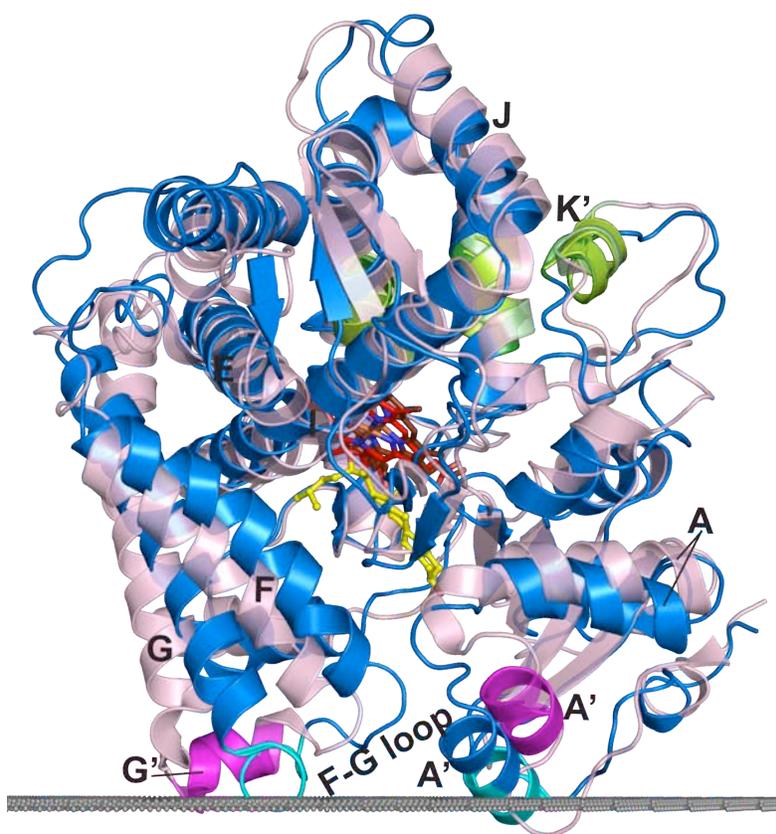


Fig. S6. Comparison of the overall fold of CYP11A1 (in marine) and CYP24A1 (in light pink) at the distal side. The array of black dots defines the membrane surface with respect to CYP11A1 with the cytosol (above) and the lipid bilayer (below). The membrane insertion sequences are colored as cyan in CYP11A1 and magenta in CYP24A1. Secondary structural elements that participate in Adx binding in mitochondrial P450s are colored in green. 22HC is in yellow; heme is in red in CYP11A1 and brown in CYP24A structure.

Computational analysis (S9) predicts that protein:membrane association transfer of free energy is lower in CYP11A1 than in CYP24A1 (-4.9 kcal/mol vs. -9.0 kcal/mol) and that membrane-bound residues penetrate into the lipid bilayer only up to 3.5 Å in CYP11A1 and much deeper, up to 22 Å, in CYP24A1 (S10).