Supplemental Information for: Characterizing the Membrane-Bound State of Cytochrome P450 3A4: Structure, Depth of Insertion and Orientation

Javier L. Baylon, Ivan L. Lenov, Stephen G. Sligar and Emad Tajkhorshid

Calculation of the orientation of CYP3A4 from the dichroic ratio measurements.

As light is incident upon an interface between two transparent media, it is partially reflected and partially transmitted. However, light inside an internal reflection element (IRE) is totally reflected internally, having no transmitted element. The angle of reflection inside of the IRE is known as the critical angle:

$$\theta_c = \sin^{-1} \frac{n_2}{n_1} \tag{1}$$

where n_1 and n_2 are the refractive indices of medium 1 and medium 2, respectively. As the evanescent wave electric field penetrates into the rarer medium and out of the IRE, it decays exponentially with the distance γ from the surface:

$$E = E_0 e^{-\gamma^2} \tag{2}$$

The amplitude E_0 can be separated into components corresponding to the magnitudes of the electric fields that extend along the laboratory axes[1], given by

$$E_x = \frac{2\left(\sin^2\theta_i - \left[\frac{n_2}{n_1}\right]^2\right)^{\frac{1}{2}}\cos\theta_i}{\left(1-\left[\frac{n_2}{n_1}\right]^2\right)^{\frac{1}{2}}\left[\left(1+\left[\frac{n_2}{n_1}\right]^2\right)\sin^2\theta_i - \left(\frac{n_2}{n_1}\right)^2\right]^{\frac{1}{2}}} \tag{3a}$$

$$E_y = \frac{2i\cos\theta_i}{\left(1-\left[\frac{n_2}{n_1}\right]^2\right)^{\frac{1}{2}}} \tag{3b}$$

$$E_z = \frac{2\sin\theta_i\cos\theta_i}{\left(1-\left[\frac{n_2}{n_1}\right]^2\right)^{\frac{1}{2}}\left[\left(1+\left[\frac{n_2}{n_1}\right]^2\right)\sin^2\theta_i - \left(\frac{n_2}{n_1}\right)^2\right]^{\frac{1}{2}}} \tag{3c}$$

where θ_i is the incident angle. From the electric field amplitudes along the laboratory axes, it is possible to calculate the absorbance of a chromophore that is adsorbed to the surface of the IRE[2]:

$$A = c_1\left|\left\langle m\left|\vec{\mu}\cdot\vec{E}\right|k\right\rangle\right|^2 \tag{4}$$

where c_1 is a constant, m and k are the states of the transition, $\vec{\mu}$ is the absorption transition moment, and $\vec{E}$ is the electric field vector. This equation can be written in terms of the laboratory axes as

$$A = kl\left(\mu_x E_x + \mu_y E_y + \mu_z E_z\right)^2 \tag{5}$$

where k is a constant and l is an effective path length. Since a heme ring can be modeled as a circular oscillator[3] and the transition moments are degenerate, the

absorbance can be broken down into the three components along the laboratory axes and simplified to the following three equations[2]:

$$A_{TE} = \frac{1}{2} kl \left|E_y\right|^2 \left|\mu\right|^2 \sin^2\theta \tag{6}$$

$$A_{TM,x} = \frac{1}{2} kl \left|E_x\right|^2 \left|\mu\right|^2 \sin^2\theta \tag{7a}$$

$$A_{TM,z} = \frac{1}{2} kl \left|E_z\right|^2 \left|\mu\right|^2 \cos^2\theta \tag{7b}$$

where θ, the orientation angle, is the angle between the transition moment and the laboratory z axis. The subscripts TE and TM refer to transverse electric polarized and transverse magnetic polarized light, respectively. TE polarized light is oriented in the y direction and only absorbers with a transition moment component in the y direction will absorb it, whereas TM polarized light is oriented in the x and z directions, thus only absorbers with transition moment components lying in the x or z direction will be able to absorb this light. The ratio of the absorbance of TE to TM polarized light, the dichroic ratio, can be used to determine the orientation of the absorber[4]. The following equation relates the dichroic ratio to the orientation angle, using Equations 6 and 7:

$$\rho = \frac{A_{TE}}{A_{TM}} = \frac{\left|E_y\right|^2}{\left|E_x\right|^2 + 2\left|E_z\right|^2 \cot^2\theta} \tag{8}$$

where ρ is the dichroic ratio and θ is the angle between the transition moment vector and the laboratory z-axis. Since CYP3A4 is a heme protein, it is possible to use the heme as the absorber and monitor the orientation of the protein in the lipid bilayer. The heme moiety has a strong absorption at the Soret band, with CYP3A4 having a molar absorptivity on the order of $106\,\mathrm{M^{-1}\,cm^{-1}}$.

References

[1] Harrick, N. J. *Journal of the Optical Society of America* **1965**, *55*, 851–857.

[2] Cropek, D. M.; Bohn, P. W. *Journal of Physical Chemistry* **1990**, *94*, 6452–6457.

[3] Eaton, W. A.; Hofrichter, J. *Methods in Enzymology* **1981**, *76*, 175–261.

[4] Norden, B. *Applied Spectroscopy Reviews* **1978**, *14*, 157–248.

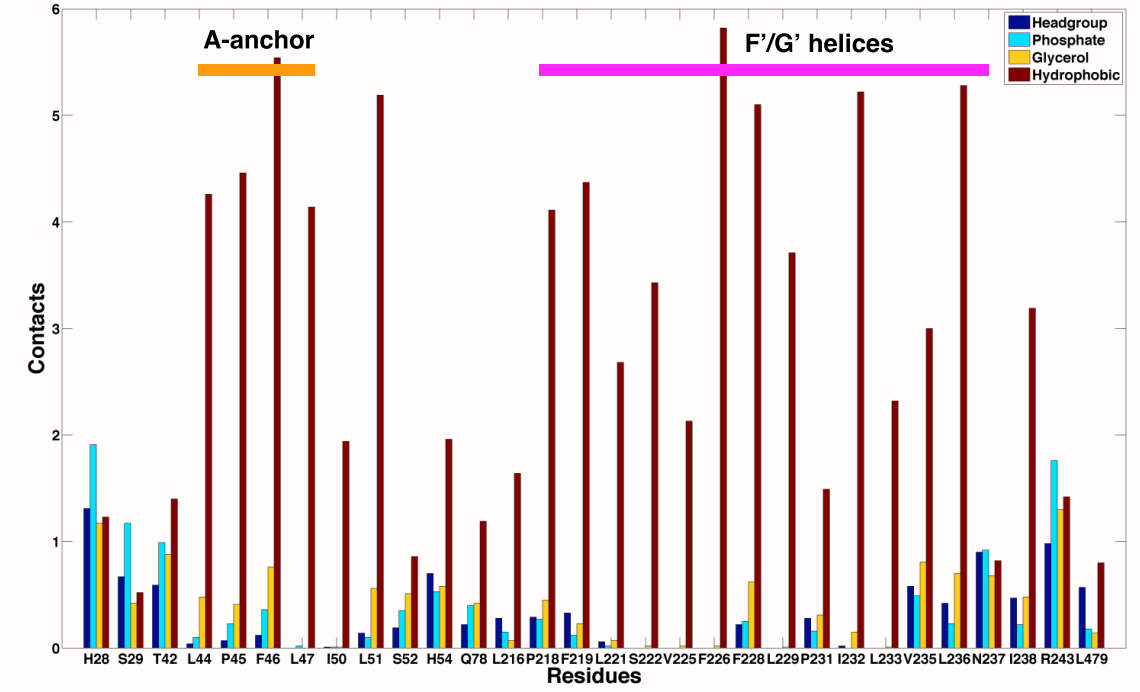

Figure S1: **Number of contacts between CYP3A4 and the membrane in its membrane-bound form.** Average number of contacts (within 5 Å) between the heavy atoms of CYP3A4 and those of lipids and DCLE molecules. The average was taken over the last 20 ns of the simulations and over the 5 HMMM membrane binding simulations. After insertion, the residues located in A-anchor and in helices F' and G' form the majority of hydrophobic contacts with the lipid tails and the DCLE molecules.

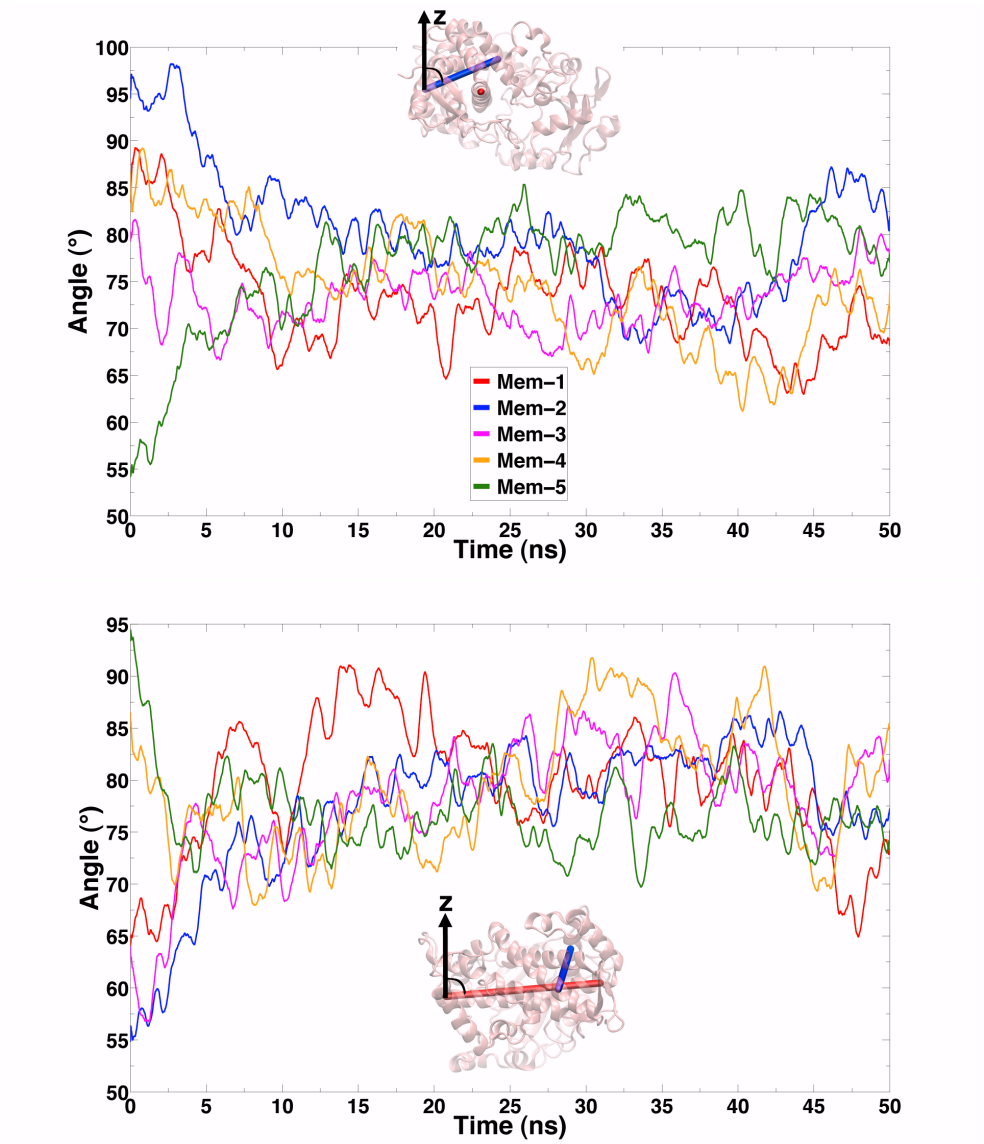

Figure S2: **Time evolution of the orientation of CYP3A4.** To specify the orientation of the protein two additional vectors were defined: a vector connecting the center of mass of the Cα atoms of residues 129-133 in helix F to the center of mass of the Cα atoms of residues 330-334 in helix C (top), and a vector along helix I, connecting the center of mass of the Cα atoms of residues 264-268 to the center of mass of the Cα atoms of residues 292-296, shown in blue and red, respectively. The angle between these vector and the membrane normal (z-axis) is plotted.

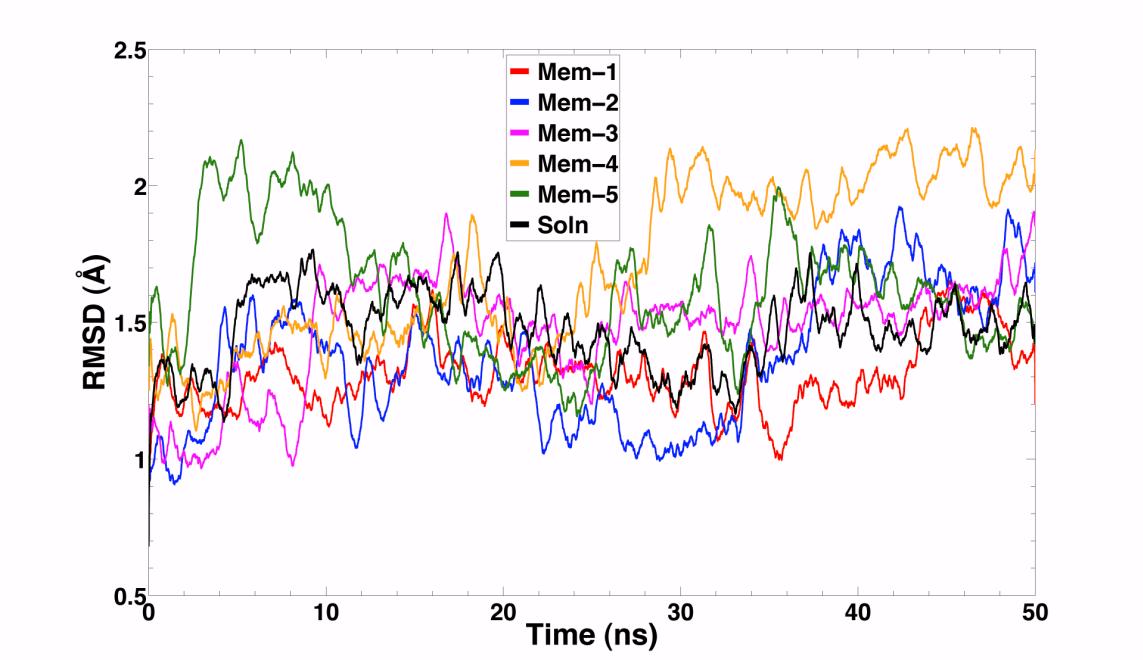

Figure S3: **Structural stability of CYP3A4.** Time evolution of the backbone RMSD of CYP3A4 in the five membrane binding simulations (shown in individual colors) and a simulation of CY3A4 in a water box (black).

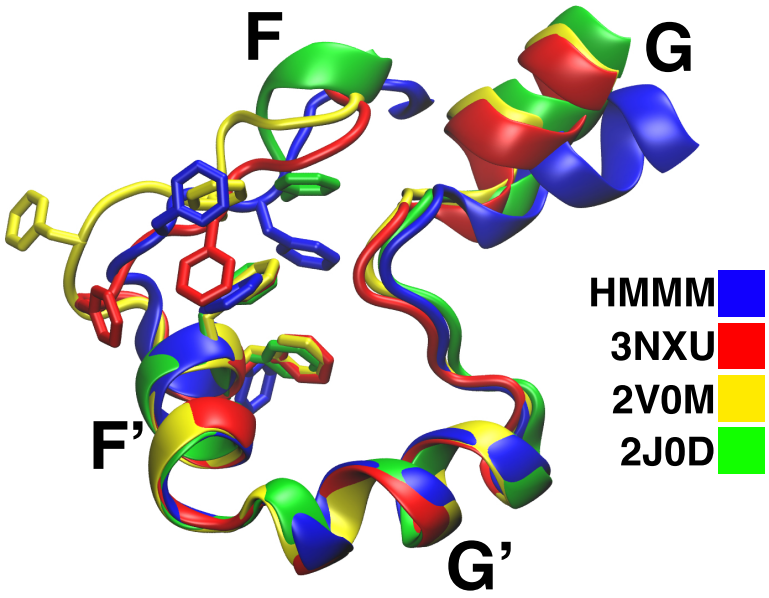

Figure S4: **Rearrangement of the F-F' coil of CYP3A4 induced by ligands in the active site.** Superposition of available crystal structures of CYP3A4 with ligands in the active site and a snapshot taken from one of the HMMM simulations representing the membrane-bound form of the enzyme. The crystal structures indicate that in order to accommodate large compounds in the active site, e.g., ritonavir (pdb: 3NXU), two ketoconazole molecules (pdb: 2V0M), and erythromycin (pdb: 2J0D), a significant rearrangement of the region between helices F and F' (residues 211 to 218) is necessary (for 2J0D, residues 214 to 218 are missing in the crystal structure). Simulations reveal that the motion of the side-chains in this region, in particular the phenylalanine side-chains (213, 215, 219, and 220), is promoted by the membrane.

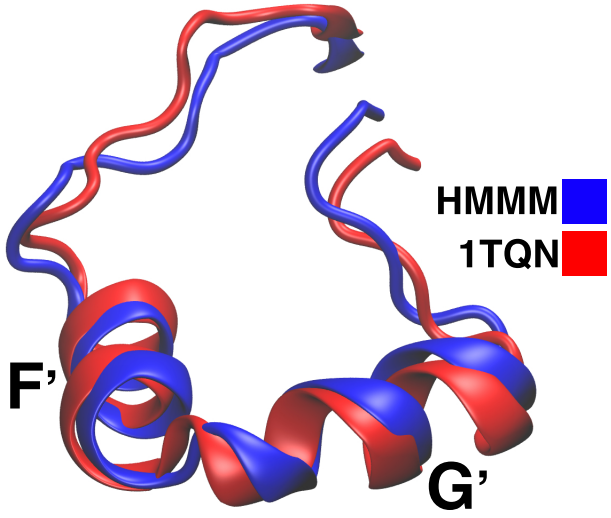

Figure S5: **Backbone stability of the F'-G' region of CYP3A4.** Comparison of a ligand-free crystal structure of CYP3A4 (pdb: 1TQN) and a snapshot representing the membrane-bound form of the enzyme obtained from our simulations indicate no significant backbone motion in the F'-G' region upon membrane binding.

Table S1: Average orientation angles measured for CYP3A4 in each simulation. The CF angle is the angle between the vector connecting helices C and F and the membrane normal (z-axis). Averages were taken for the last 40 ns of simulation. The standard deviation is in parentheses. The vectors are defined in Fig. S1.

Run	Initial CF angle (°)	Average CF angle (°)	Initial helix I angle (°)	Average helix I angle (°)
Mem-1	84.5	71.8 (4.0)	63.8	80.8 (5.6)
Mem-2	94.5	77.9 (5.0)	54.8	80.1 (3.6)
Mem-3	79.3	73.3 (3.2)	63.9	80.0 (4.4)
Mem-4	86.7	72.3 (5.0)	86.5	80.1 (6.2)
Mem-5	54.3	79.4 (3.4)	93.3	76.0 (3.0)
Average		**74.9 (3.5)**		**79.4 (1.0)**

Table S2: Gate residues and lining secondary structures of the access tunnels.

Tunnel	Lining secondary structures	Gating residues
2a	β1 sheet and F' helix	Ile-50, Leu-216, Leu-221
2b	β1 sheet and B-C loop	Gln-78, Phe-108, Pro-227
2e	B-C loop	Arg-106, Phe-108, Ser-119, Ile-120
3	F'-G' helices	Phe-213, Phe-220, Val-240
S	F-F' loop and I helix	Arg-212, Phe-213, Gln-484