

PCCP

Accepted Manuscript



This is an *Accepted Manuscript*, which has been through the Royal Society of Chemistry peer review process and has been accepted for publication.

Accepted Manuscripts are published online shortly after acceptance, before technical editing, formatting and proof reading. Using this free service, authors can make their results available to the community, in citable form, before we publish the edited article. We will replace this *Accepted Manuscript* with the edited and formatted *Advance Article* as soon as it is available.

You can find more information about *Accepted Manuscripts* in the [Information for Authors](#).

Please note that technical editing may introduce minor changes to the text and/or graphics, which may alter content. The journal's standard [Terms & Conditions](#) and the [Ethical guidelines](#) still apply. In no event shall the Royal Society of Chemistry be held responsible for any errors or omissions in this *Accepted Manuscript* or any consequences arising from the use of any information it contains.

Lipid molecules can induce an opening of membrane-facing tunnels in cytochrome P450 1A2

Petr Jeřábek¹, Jan Florián², Václav Martinek^{1,3*}

¹ Department of Biochemistry, Charles University, Faculty of Science, Albertov 2030, 128 43 Prague 2, Czech Republic

² Department of Chemistry and Biochemistry, Loyola University Chicago, 1032 W. Sheridan Rd., Chicago, IL 60660, USA

³ Department of Teaching and Didactics of Chemistry, Charles University, Faculty of Science, Albertov 3, 128 43 Prague 2, Czech Republic

* To whom correspondence should be addressed: E-mail: vaclav.martinek@natur.cuni.cz,
Phone: (420) 221-951344

Keywords

cytochrome P450 1A2, CYP1A2, molecular modeling, molecular dynamics, free energy, linear interaction energy, protein-membrane interactions, membrane proteins, protein-membrane self-assembly, coarse-grained molecular dynamics, MARTINI force field

Abstract

Cytochrome P450 1A2 (P450 1A2, CYP1A2) is a membrane-bound enzyme that oxidizes broad range of hydrophobic substrates. The structure and dynamics of both the catalytic and trans-membrane (TM) domains of this enzyme in the membrane/water environment were investigated using multiscale computational approach, including coarse-grained and all-atom molecular dynamics. Starting from the spontaneous self-assembly of the system containing the TM or soluble domain immersed in randomized dilauroyl phosphatidylcholine (DLPC)/water mixture into their respective membrane-bound forms, we reconstituted the membrane-bound structure of the full-length P450 1A2. This structure includes a TM helix spanning the membrane, while being connected to the catalytic domain by a short flexible loop. Furthermore, in this model, the upper part of the TM helix directly interacts with a conserved and highly hydrophobic N-terminal proline-rich segment of the catalytic domain; this segment and the FG loop are immersed in the membrane, whereas the remaining portion of the catalytic domain remained exposed to the aqueous solution. The shallow membrane immersion of the catalytic domain appears to induce a depression in the opposite intact layer of phospholipids, which may help in stabilizing the position of the TM helix directly beneath the catalytic domain. The partial immersion of the catalytic domain also allows for the enzyme substrates to enter the active site from either aqueous solution or phospholipid environments *via* several solvent- and membrane-facing tunnels in the full-length P450 1A2. The calculated tunnel dynamics indicated that the opening probability of the membrane-facing tunnels is significantly enhanced when a DLPC molecule spontaneously penetrates into the membrane-facing tunnel 2d. Energetics of the lipid penetration process was assessed by the linear interaction energy (LIE) approximation, and found to be thermodynamically feasible.

Eukaryotic cytochromes P450 (P450) represent a large group of membrane-bound heme enzymes that are usually anchored in the cytosolic side of the endoplasmic reticulum. P450s play a key role in the oxidation of a broad spectrum of hydrophobic substrates, e.g. sterols, fatty acids, drugs, carcinogens and xenobiotics.^{1,2} Cytochrome P450 1A2 (P450 1A2) is the major enzyme of the cytochrome P450 family 1 that is constitutively expressed in the human liver. It preferentially oxidizes the planar polycyclic aromatic hydrocarbons and amines, including important drugs and xenobiotics.³ The P450 1A2 substantially contributes to the activation of carcinogens in human liver.⁴⁻⁹ Activity of reticular P450s depends on a sequential transfer of two electrons, which are delivered *via* membrane-bound electron donors NADPH:P450 reductase and/or cytochrome *b₅*.¹⁰⁻¹³ It is well known that the presence of phospholipid membrane is required for the activity of eukaryotic P450s^{14,15} as the membrane assists their efficient cooperation with the electron donors.¹⁶

Furthermore, phospholipids exert stabilizing effects on the P450 protein itself, as their presence decreases content of an improperly folded form (P420).^{15,17} To elucidate the structural impact of the phospholipid membrane on a P450 at the atomic level we aimed to construct a full-length model of human P450 1A2 in the membrane/water environment, and compare it with the structure and dynamics of the truncated form in aqueous solution without presence of a membrane.

All-atom molecular dynamics is frequently employed to study dynamics and interactions of already structurally well characterized membrane proteins.¹⁸⁻²² In contrast, the coarse-grained molecular dynamics (CG MD) simulations may be successfully used for self-assembly of various membrane proteins and peptides with the phospholipid membrane,²³⁻²⁵ thus allowing to compensate for partial absence of structural information on the trans-

membrane (TM) protein. Both approaches can be fruitfully combined during multi-scale modeling of membrane associated proteins.^{26–29}

P450s have been studied using MD simulations,^{30,31} but the vast majority of these studies was conducted in solution without including the TM domain or phospholipid bilayer. Early studies of a membrane-bound cytochrome P450 2C9, involved (i) an experimentally-guided manual placement of the protein into the membrane,³² or (ii) a microsecond coarse-grained molecular dynamics (CG MD) optimization in pre-assembled lipid membrane followed by ~30 ns classical MD.³³ Despite involving different prediction strategies and membrane phospholipids, the two approaches were in a rough agreement, showing viability of a membrane-anchored cytochrome P450 structure prediction. The latter approach, which is less dependent on additional information, was also applied to study other eukaryotic membrane-bound cytochromes P450, P450 51A1³⁴ or P450 3A4³⁵. Relying on an all-atom MD and a novel 5-carbon palmitoyloleoyl phosphatidylcholine/1,1-dichloroethane membrane model, which allowed higher mobility of membrane proteins, Baylon et al. investigated membrane-bound P450 3A4.³⁶ We recently used combination of CG and all-atom techniques for prediction of cytochrome P450 1A2–cytochrome *b*₅ complex.³⁷

In this paper, we used multi-scale modeling techniques to propose structure of the full-length human P450 1A2 in the phospholipid membrane. We performed the self-assembly and MD optimization of the membrane-bound enzyme and described changes of the membrane homogeneity near the membrane binding domain. We also paid special attention to the structural and dynamical description of the accessibility of the P450 1A2 active site. This is because active sites of some enzymes, including P450s, are buried deep inside a protein and communicate with their environment through a tunnel, which might serve as a gate that limits active site accessibility. Enzyme gates can control a substrate/product access, solvent accessibility, or synchronize processes taking place in distant parts of the protein.³⁸

Investigation of protein gates is therefore important for understanding enzymatic catalysis; especially if a ligand trafficking limits the reaction rate.

Tunnel analysis have been performed on a broad range of structures of P450s. Possible routes were also examined using various MD methods (classical MD, REMD, or SMD)³⁹⁻⁴², however these studies were focused either on prokaryotic P450s or the truncated catalytic domains of eukaryotic P450s. The tunnels found in selected membrane-bound eukaryotic P450s were reviewed recently.⁴³ In general, because the orientation of individual enzymes in the membrane is experimentally unknown, the calculated tunnel assignment may be strongly affected by the assumptions and algorithms that were used to determine this orientation.

Here we rely on the simulated self-organization of the membrane-protein system to determine the protein orientation. We then focus on comparing opening probabilities of solvent- and membrane-facing tunnels between the membrane-bound and free hydrated P450 1A2 catalytic domain. We also describe the penetration of a membrane phospholipid into one of the membrane-facing tunnels. This analysis of the P450 tunnel dynamics could help in identifying the substrate access tunnel and elucidating the enzyme substrate preferences.

Methods

The CG model of the membrane-bound P450 1A2 was constructed by combining end points of CG MD self-assembly simulations of the catalytic and TM domains. This model was, upon the CG MD optimization, converted to the all-atom model, and the resulting full-length P450 1A2 was subjected to extensive all-atom MD simulations (Figure 1).

General setup of CG and all-atom MD simulations

All CG MD simulations were performed using MARTINI force field^{44,45} implemented in NAMD v2.9⁴⁶. MARTINI force field reduces the number of explicitly simulated particles per

each lipid or amino acid, thus lowering the friction inside the simulated system and allowing the application of a significantly larger integration step-size (up to 25 fs). This force field typically joins four heavy atoms and their hydrogens into a single particle (bead). One backbone bead represents all backbone atoms of each amino acid residue; larger residues also include side-chain beads. Four basic types of CG particles exist: charged, non-polar, apolar, and polar; additional particle subtypes are used to simulate H-bonding capabilities. This parametrization was extensively validated to reproduce various structural features of the biological membranes⁴⁷, it also takes into account experimental polar/apolar partition coefficients, however, unlike some more recent CG models^{48,49}, it has limitations especially in consistent treatment of the electrostatic interactions. MARTINI model is among the most frequently used CG models for biological membrane simulations.⁵⁰

CG simulations were carried out for the NPT ensemble (310 K, 1 atm) under periodic boundary conditions. Temperature and pressure were held constant using Langevin dynamics (friction coefficient 1) and Langevin piston method implemented in NAMD.⁴⁶ Isotropic and semi-isotropic pressure coupling were used in the self-assembly simulation of the phospholipid bilayer and the simulation of the membrane-bound P450 1A2, respectively. The integration time step was 20 fs. Lennard–Jones and Coulombic interactions were shifted to zero between 9 and 12 Å, and 0 and 12 Å, respectively. Based on earlier cytochrome P450 implementation of the elastic network model⁵¹, we used 10.75 kcal/mol/Å² harmonic force constants to constrain distances between backbone beads that were in the corresponding crystal structure closer than 7 Å. These constraints allowed retaining the ternary structure of the catalytic P450 1A2 domain during CG MD simulations. Due to the absence of the heme parameters in the MARTINI force field the structural role of this cofactor was substituted in our CG simulations by the set of elastic network constraints, which helped to preserve the

shape of the heme-binding cavity. During the CG to all-atom conversion, the heme cofactor was reinserted into this cavity. Thus, the structural effects of this cofactor were explicitly included in all all-atom MD simulations.

All-atom MD simulations were carried out in aqueous solution represented by the explicit TIP3P model. The NPT ensemble (310 K, 1 atm) and periodic boundary conditions were used. Temperature and pressure were held constant using Langevin dynamics (friction coefficient 1) and Langevin piston method implemented in program NAMD v2.9⁴⁶. The simulation time step was 2 fs and cutoff for long-range nonbonding interactions was 12 Å. For electrostatic (ES) interactions, we used the particle mesh Ewald method⁵² implemented in NAMD. Bonds involving hydrogen atoms and TIP3P water were kept rigid using SHAKE algorithm.⁵³ All simulations were performed with the CHARMM27 force field⁵⁴ and its extension for lipids (CHARMM36⁵⁵). For further details on the simulation settings, see the system specific sections below.

Self-assembly of the P450 1A2 catalytic domain/water/DLPC system (Figure 1, step 1)

The simulated system consisted of a cubic box with a ~125 Å edge. This box contained the catalytic domain of human P450 1A2 (residues 34-513), which was immersed in 10346 water molecules and 500 randomly placed molecules of DLPC. DLPC and water molecules were randomized using program PACKMOL.⁵⁶ The initial all-atom coordinates for the P450 1A2 catalytic domain were obtained from the Protein Data Bank (PDB ID 2HI4)⁵⁷. Using VMD⁵⁸, 19 K⁺ and 31 Cl⁻ ions were additionally uniformly distributed in aqueous solution to neutralize the simulated system and mimic 0.1M KCl solution; this initial all-atom model of the protein/lipid/water/ion mixture was converted to the MARTINI CG model.^{44,45} NAMD v2.9⁴⁶ was used to minimize the system energy (1000 steps) at the CG level and to carry out

32 independent CG MD trajectories (replicas); each replica was sampling the system for 590 ns.

Self-assembly of the P450 1A2 TM domain/water/DLPC system (Figure 1, step 2)

The simulated system consisted of a cubic box with a ~ 105 Å edge. This box contained the TM domain of human P450 1A2 (residues M1-K41), which was immersed in 5149 water molecules and 350 randomly placed molecules of DLPC. DLPC and water molecules were randomized using CG program PACKMOL.⁵⁶ The initial all-atom coordinates were generated using PyMol. The secondary structure was initially assigned as coil (M1-Q5 and P33-S41) or helix (S6-R32) according to prediction made by TMHMM and MEMSAT servers.^{59,60} Using VMD⁵⁸, 10 K⁺ and 15 Cl⁻ ions were uniformly distributed in aqueous solution to neutralize the simulated system and mimic 0.1 M KCl solution; this initial all-atom model of the protein/lipid/water/ion mixture was converted to the MARTINI CG model.^{44,45} NAMD v2.9⁴⁶ was used to minimize the system energy (1000 steps) and to run 16 independent 100 ns CG MD trajectories (replicas).

Assembly and CG MD simulation of a full-length membrane-bound P450 1A2 (Figure 1, steps 3 and 4)

For the construction of the CG model of the membrane-anchored P450 model, we needed to combine the structures of the catalytic and TM domains that resulted from separate lipid-protein self-assembly simulations. This construction was initiated by manual positioning of the catalytic and the TM domain at about 20 Å apart, while preserving the self-assembled membrane orientations of each protein, including a 10 Å-wide bilayer of surrounding DLPC molecules. Small gaps between these two membrane segments were sealed during the subsequent CG equilibration phase. Steered molecular dynamic (SMD) technique was then

used to generate four alternative starting orientations of the TM helix in the membrane (cylinders in step 4 in the Figure 1). The protein, except residues M1-G38, was fixed during SMD simulations and the BAS bead of the residue S6 was dragged in four lateral directions at a constant velocity ($v = 10 \text{ \AA/ns}$, 300 000 steps, distance = 60 \AA , force constant = 100 kcal/mol/ \AA^2). These four alternative starting structures, which differed in the orientation of the TM domain with respect to the catalytic domain, were equilibrated using CG MD and their movement was followed by CG MD production phases on a microsecond timescale (Table S1).

Conversion of the CG-model to the all-atom model (Figure 1, step 5)

Prior to converting the CG model to its atomic coordinates, we identified the most representative conformation of the full-length P450 1A2 model. For this purpose, a set of seven geometric parameters, which relate rigid-body positions of protein domains and the membrane, was devised and measured every 1 ns. This set included (Figure S1) the two distances from the center of the phospholipid bilayer to the center of mass of the catalytic (d_1) and TM (d_2) domains, angles α , β , and γ , between the axis perpendicular to the phospholipid bilayer (z -axis) and each of the three α -helix vectors (v_1 , v_2 , and v_3), and angles δ and ϵ that described mutual orientation of the two domains. Based on monitoring these parameters, the CG geometry of the full-length membrane-bound P450 1A2 converged within 2 μs (Figure S2) with the exception of angle δ that kept fluctuating by $\sim 20^\circ$, indicating a higher mobility of the TM domain.

A CG MD snapshot that featured d_1 , d_2 , α , β , γ , δ , and ϵ values closest to the most populated values of these parameters (Figure S3A), was defined as a representative structure of the P450 1A2 CG model. This representative structure was then converted to its atomic representation using algorithm of Shih et al.⁶¹ and CGTools integrated in VMD v1.9.1⁵⁸. The

structures of the P450 1A2 TM domain (residues 1 to 34) and the phospholipid bilayer were obtained directly by the CG to all-atom conversion. The water environment mimicking 0.1M KCl solution was generated *de novo* for individual all-atom simulations. In order to prevent any distortion of the tertiary structure of P450 1A2 catalytic domain, that might occur during the CG to all-atom conversion the coordinates of the catalytic domain of P450 1A2 (residues 34-513) were replaced with its crystallographic structure (PDB ID 2HI4)⁵⁷. This was done by aligning the crystallographic structures of P450 1A2 catalytic domain to the all-atom model of the membrane-bound P450 1A2 *via* backbone atoms of P450 1A2 (residues 34-513).

All-atom MD simulations of the P450 1A2 catalytic domain

The initial all-atom coordinates for P450 1A2 catalytic domain (PDB ID 2HI4)⁵⁷ were solvated using VMD v1.9.1⁵⁸. All water molecules present in the original PDB files were preserved. 18 K⁺ and 29 Cl⁻ ions were additionally uniformly distributed in aqueous solution to neutralize the simulated system and mimic 0.1M KCl solution. The system was equilibrated prior to the production run in several subsequent steps (Table S2). The equilibration phase was followed by five production simulations having a combined length of 1.15 μ s.

All-atom MD simulations of the membrane-bound full-length P450 1A2

The all-atom membrane-bound P450 1A2 was solvated using VMD v1.9.1⁵⁸, while retaining intra-protein crystallographic water molecules from the P450 1A2 catalytic domain structure.⁵⁷ 141 K⁺ and 153 Cl⁻ ions were uniformly distributed in aqueous solution to neutralize the simulated system and mimic 0.1M KCl solution. The system was equilibrated prior to the production run in several subsequent steps (Table S3). The equilibrated structure of the membrane-bound P450 1A2 was used as a starting point for five independent production MD simulations (replicas) having a combined length of 1.34 μ s.

Analysis of access/egress tunnels in the catalytic domain

CAVER 3.0⁶² was used to analyze tunnels connecting the active site of P450 1A2 with the protein environment. Snapshots from MD trajectories of the equilibrated all-atom membrane-bound and membrane-free P450 1A2 models were extracted every 1 ns and used as input data for CAVER. Tunnels were considered open if the bottleneck radius was ≥ 1.9 Å. Each calculated tunnel was named according to the established P450s tunnel nomenclature⁶³. The starting point for the tunnel search was specified by C α atoms of T124 and T321, and Fe atom of the heme cofactor. DLPC, α -naphthoflavone and water molecules were removed prior to the analysis.

Binding free energy monitoring for the lipid chain insertion in the protein tunnel

The relative free energy of the lipid protruding into the membrane-bound P450 1A2 was estimated using the linear interaction energy (LIE) approximation⁶⁴. LIE calculations were performed every 1 ps along the 420 ns all-atom MD trajectory 1m (Table S4) utilizing program NAMD v2.9. The probe region involved the lauryl residue and the immediately attached -CH- group of glycerol residue of intruding DLPC molecule; the long-range nonbonding interactions were evaluated using 16 Å cutoff with switching distance 15 Å. The relative free energies were calculated as

$$\Delta G_{bind} = \beta(V_{bound}^{elec} - V_{unbound}^{elec}) + \alpha(V_{bound}^{vdW} - V_{unbound}^{vdW}) \quad (1)$$

where interaction energies between the probe region and its surroundings are represented by the electrostatic component of bound (V_{bound}^{elec}) and unbound ($V_{unbound}^{elec}$) state and van der Waals components of bound (V_{bound}^{vdW}) and unbound ($V_{unbound}^{vdW}$) state. The unbound reference state was represented by initial 17 ns part of the trajectory. The bound state was represented by 1 ns long segments along the remaining portion of the trajectory. Experimental solvation

free energies for heptane and methyl octanoate⁶⁵ and separate MD simulations on these molecules were used to determine optimal LIE parameters $\alpha = 0.30$ and $\beta = 0.39$. These parameter values lie between recommended values for GROMOS force field ($\alpha = 0.18$ and $\beta = 0.43$) and values developed for AMBER force field ($\alpha = 0.43$ and $\beta = 0.34$).^{66,67}

Results and Discussion

Self-organization of a protein/water/DLPC system includes a spontaneous insertion of the catalytic and TM domains of P450 1A2 into membrane.

While constructing the CG models we followed a strategy of minimizing “human made” decisions by relaying, whenever possible, on the ability of the studied complex systems to self-organize. In particular, to establish the orientation of the catalytic and TM domains of human P450 1A2 in a phospholipid bilayer we carried out multiple protein/lipid self-assembly simulations. Due to the large size of the full-length P450 1A2 the self-assembly simulations of the corresponding membrane-bound full-length complex from the mixture of randomly positioned phospholipid and water molecules did not converge during the available simulation time-scale. Therefore, this complex was reconstituted by overlapping the structures of membrane complexes of the TM and soluble domains, which were obtained by separate self-assembly simulations. The self-assembly of the randomized mixtures containing the TM or soluble domain into their respective membrane-bound forms converged in several CG MD replica simulations, each yielding similar structures described in following paragraphs. The catalytic domain converged to membrane-bound configurations in 22% of replicas (in 7 out of 32), in which the *N*-terminal segment and FG loop were found dipped into the spontaneously formed “flat” membrane (Figure 2A) and all DLPC molecules were incorporated in the membrane. These productive replicas converged into similar protein-membrane orientation

(angle $\alpha = 84 \pm 20^\circ$) (Figure S4), whereas the remaining replicas assembled into a distorted membrane or an unbound protein.

Considering all productive self-assembly simulations, we noticed a tendency of the top membrane leaflet, which hosted the dipped catalytic domain, to contain a smaller number of DLPC molecules (246 ± 4) than the opposite leaflet (253 ± 4). This asymmetric lipid distribution was retained in the subsequent CG and all-atom simulations of the full-length model; the final all-atom full-length P450 1A2 model, the composition of the lower and upper leaflet differed by 13 DLPC molecules. The observed small asymmetry (3-5 %) in lipid distribution could be attributed to the steric effects induced by the catalytic domain displacing approximately 10 DLPC molecules in the top leaflet.

Simulations of the TM domain of P450 1A2, which was immersed in randomized DLPC/water mixture, yielded a tilted membrane-spanning TM domain. Efficiency of this self-assembly was more than 40%, seven of 16 replicas lead to the recurrent structure comprised of flat membrane incorporating TM domain. To better quantify the membrane orientation of the TM helix, we extended one of the replicas by 600 ns, resulting in a $49^\circ (\pm 7^\circ)$ average tilt angle (γ) of the TM helix (Figure 2). A snapshot, in which the TM helix was tilted by 49° , was used in the following step of the model construction (Figure 1).

TM domain of P450 1A2 associates with the proline-rich loop of the *N*-terminal segment of the catalytic domain

The P450 1A2 TM domain is covalently attached to the *N*-terminal end of the catalytic domain *via* a flexible linker (residues R34-P42), but experimental information about the tertiary structure of these domains is not available. Therefore, using a multiscale MD, we investigated the structure and dynamics of the full-length P450 1A2 in the DLPC phospholipid bilayer. Initially, representative structures of the individual TM and catalytic

domains were positioned ~ 20 Å apart in a plane of the phospholipid bilayer (Figure 3). While preserving the average γ angle obtained from the self-assembly simulations of the TM domain, four dissimilar initial orientations of the TM helix with respect to the catalytic domain were considered. All four CG MD simulations yielded similar structures that featured direct contacts between the TM and catalytic domains, wherein the upper half of the TM helix segment was in contact with the hydrophobic proline-rich *N*-terminal segment (residues P43-G52) of the catalytic domain (Figure 3). The proline-rich segment is sequentially and structurally conserved among the family 1, 2 and 3 eukaryotic cytochromes P450 (Figure 4). Notably, the proline-rich segment is more sequentially conserved than the TM helices of P450s.

All-atom structure and dynamics of the membrane-bound full-length P450 1A2

A representative CG structure of the full-length P450 1A2, the selection of which was based on average values of a set of geometrical parameters (Figure S1), was converted to all-atom coordinates (step 5 in Figure 1). This system was further analyzed using five independent replicas of all-atom MD trajectories that reached the total simulation time of 1.34 μ s and included explicit membrane and water environment. As a point of reference, free P450 catalytic domain was simulated in water environment. These unconstrained all-atom simulations can be expected to improve upon physico-chemical accuracy of the CG model.

During all-atom simulations, the two P450 1A2 domains retained their mutual orientation that was acquired previously in the CG simulations. The catalytic domain remained immersed in the membrane, mostly with its *N*-terminal segment and FG loop (Figure 5A). Most residues belonging to these segments were dipped in the polar head groups region of the membrane, but residues G47, W47, P49, L50, and L51 were buried deeper in the hydrophobic region of the membrane. The Q5-R32 segment of the TM helix was immersed in

the phospholipid bilayer, the charged residues E13, K29, R32, and R34 near the *C*- and *N*-termini of the TM helix serving as its anchors in the polar head group membrane layer. The S18 residue, being buried deeply in the hydrophobic environment, was for most of the simulated time H-bonded to the main chain of the peptide linker or to main chain oxygen in the proline-rich *N*-terminal segment.

The *N*-terminal segment and FG loop of the catalytic domain fluctuated less in the membrane than in aqueous solution (Figure S5). Interestingly, the membrane-mediated attenuation of the conformational dynamics of the FG loop was propagated into ~ 20 Å distant segments of the catalytic domain (H and I helices, and HI loop) that are not in contact with the phospholipid bilayer (Figure S5). The TM helix position showed ~ 3 Å lateral and vertical oscillations with respect to the catalytic domain (Figure 5A), while retaining its interactions with the proline-rich *N*-terminal segment of the catalytic domain (residues P43-G52). The polar peptide linker (residues R34-P42) connecting the catalytic and TM domains shows large flexibility (Figure S5) and is positioned on the water-membrane interface.

The predicted tertiary fold of the catalytic domain of the full-length P450 1A2 is similar to the observed tertiary structure of the catalytic domain of lanosterol 14 α -demethylase (P450 51A1)⁶⁸ (Figure 5B and S6). In particular, despite having a low sequence identity (16%) (Figure S6), both catalytic domains immerse shallowly into the membrane *via* their FG loops and *N*-terminal segments. The membrane-anchoring domain of P450 51A1 was predicted to consist of two helices.⁶⁸ The orientation of the second helix, relative to the catalytic domain and the phospholipid bilayer, is very similar and within the area sampled by P450 1A2 TM helix during our all-atom MD simulations. This enzyme also contains peptide which is structurally and sequentially related to the proline-rich segment of P450 1A2 (Figure S6). Our theoretical prediction together with experimental observations on analogous P450 51A1 strongly suggest that the TM domain of P450 1A2 is positioned below its catalytic

domain, where it closely interacts with the proline-rich *N*-terminal segment. This segment might serve as rigid hydrophobic platform that stabilizes the position of the TM helix. Due to their high sequential similarity, proline-rich segments of P450 1A2 and related P450 enzymes could share this structural function.

P450 1A2 induces funnel-like depression in the lower layer of the phospholipid membrane

The spontaneous lateral relocation of the TM helix and its stabilization in the membrane region directly beneath the catalytic domain is associated with and possibly driven by structural perturbations in the membrane that are induced by the catalytic domain. Our CG simulations of the fully organized water/DLPC system containing catalytic domain that is dipped in the phospholipid bilayer indicate that such structural perturbation exists also in the absence of the TM domain (Figure 6A). This deformation becomes more pronounced when the full-length protein is immersed in the DLPC membrane (Figure 6B). A similar membrane deformation was observed in CG simulation of cytochrome P450 3A4.³⁵ Here we confirmed the presence of the membrane deformation also in the all-atom model of full-length P450 1A2 (Figure 6C), indicating that the forces inducing the membrane deformation are also present in the all-atom model that employs a different force-field (CHARMM replacing MARTINI). We interpret these results by noticing that the funnel-like depression is not primarily induced by the presence of TM helix, which is only stabilizing the deformation, but its occurrence relates to the presence of the catalytic domain that replaces 13 DLPC molecules in the upper layer of the membrane. However, the buried portion of the catalytic domain is too short to substitute for the full length of the absent aliphatic chains. Thus, DLPC molecules in the lower layer beneath the catalytic domain translocate by ~ 3.5 Å toward the center of the phospholipid

bilayer to compensate for the lower thickness of the upper layer. This translocation creates a local funnel-shaped deformation of the lower membrane layer.

Although DLPC is not common phospholipid present in reticular membranes, it is frequently used for reconstituting purified P450s.⁶⁹ P450s reconstituted with DLPC show similar activity as the corresponding enzymes in microsomes isolated from baculovirus infected insect cells (Supersomes).⁷⁰ Thus, the presence of the funnel-like indentation in the DLPC membrane could be a general phenomenon contributing, besides the interaction with the proline-rich segment, to the stabilization of the TM helix position beneath the catalytic domain.

Catalytic domain of P450 1A2 is entwined with eight potential access/egress tunnels

The P450 active site is buried deep inside the catalytic domain and their ligand binding typically occurs with millisecond to minute rate.⁷¹ The X-ray crystal structure of the P450 1A2 catalytic domain (PDB ID 2HI4)⁷² corresponds to its closed form that sterically prevents the substrate binding or product release. Thus, the reaction substrates and products must enter or leave the protein *via* structural fluctuations that open tunnels connecting the active site with the membrane or water environment. If these events are rare they may limit the overall catalytic rate, making the tunnel description and their structural control an important target of structure-function studies.

To assess the tunnel dynamics and to study potential influence of the phospholipid bilayer on the P450 1A2 we analyzed a set of 10 simulations in the presence and absence of the membrane. To identify protein tunnels, studies of membrane-bound P450s usually use probe cutoffs around 1.2 \AA ^{33,36,73}, which is actually less than the radius of a water molecule (1.4 \AA). However, typical organic substrates of P450 family 1 enzymes are aromatic

hydrocarbons and their derivatives, with their minimal radius significantly larger than 1.4 Å. The probe used here was methane-sized (1.9 Å), which represents the smallest cross-section radius the most trivial substrate of P450s can possess, and we believe that it is more relevant for actual substrates of P450s as it filters out tunnels that are not sufficiently large for organic molecules. At our simulation timescale, the maximal tunnel bottleneck of 2.4-2.7 Å is even larger than 1.9 Å Table 1. The tunnels 2b, 2d, and 4 show the largest bottlenecks (2.7-2.8 Å) and they have the highest potential to transport the organic substrates inside the P450 1A2 active site. Thus, comparative studies of the tunnel dynamics could help in elucidating mechanisms that contribute to substrate preferences of individual P450s.

Tunnel opening in P450 1A2 is modulated in the membrane presence

Membrane-facing tunnels of P450s were proposed to let in lipophilic substrates whereas solvent-facing tunnels are supposed to release soluble products.^{30,41,74} The catalytic domain in our membrane-free and membrane-bound MD simulations undergoes a various degree of conformational rearrangements that open one or more membrane-facing or solvent-facing tunnels. During all-atom simulations of the P450 1A2 catalytic domain, we detected three solvent-facing tunnels (2c, s, and w) and two membrane-facing tunnels (2b and 4) (Figure 7 and Table 1). Tunnel 2c was the most populated, being open in the membrane-free and membrane-bound simulations for 28 and 13% of the total simulation time, respectively. The high opening probability of this solvent-facing tunnel indicates its importance for fast product release as it likely poses a low energy barrier for release of hydroxylated products into the solvent. The opening of this tunnel is influenced by the dynamics of the entrance-lining structural elements (G and I helices, and BC loop). This tunnel is in the crystal structure of P450 1A2⁵⁷ blocked by the tunnel gating residues T118, D119. These residues show high flexibility in our MD simulations and also in the X-ray structure, the latter being measured by

the observed thermal factors. Additionally, two solvent-facing (w and s) and two membrane-facing tunnels (2b and 4) were detected. However, each of the membrane-facing tunnels opened very infrequently (Table 1).

The membrane environment induced significant opening of the membrane-facing tunnels, while it decreased opening probability of solvent-facing tunnels (Table 1). The opening probability of solvent-facing tunnels 2c, s, and w decreased in membrane-bound P450 1A2 compared to the membrane-free simulations. That could be attributed to a larger structural rigidity of the entrance lining elements mentioned above. We also detected a new solvent-facing tunnel 2e passing through the BC loop, however it is open very infrequently. On the other hand, the opening probability of membrane-facing tunnels 2b and 4 significantly increased in membrane-bound simulations. Two membrane-facing tunnels (2d and 5) were induced in P450 1A2 as a result of the membrane presence (Table 1).

The five trajectories sampling dynamics of the membrane-free P450 1A2 catalytic domain differed in the presence/absence of up to two tunnels (Table S4). However, the tunnel opening probability in one simulation of the membrane-bound P450 1A2 dramatically diverged, showing new membrane-facing tunnels with high opening probability (replica 1m in Table S4). After visual inspection of this trajectory, we noticed that substantial part of a DLPC molecule enters tunnel 2d.

Penetration of a membrane lipid into the catalytic domain induces opening of membrane-facing tunnels

A detailed analysis of the trajectory m1 revealed that a fatty acid chain of a DLPC molecule was spontaneously entering membrane-facing tunnel 2d in vicinity of the proline-rich *N*-terminal segment. Along the MD trajectory, we monitored the DLPC-heme cofactor distance, the relative binding free energy of the lauryl ester residue of the intruding DLPC and the

opening frequencies of several access tunnels (Figure 8). The whole trajectory could be divided into four segments representing distinct binding positions of phospholipid molecule (Figure 8A). The first segment catches the initial stages of intrusion, where the tip of the aliphatic chain of DLPC molecule binds into the nascent tunnel 2d. The second segment starts at 60 ns when the DLPC intrusion stops approximately 20 Å away from the heme cofactor, and the lipid remains in its place for another 130 ns (Figure 8B). Simultaneously, tunnels 2b, 2d, 4 start to open more frequently and at the end of this segment they are open for most of the time (Figure 8D). Later during this time, the protein undergoes conformational changes in the placement of F helix and FG loop (Figures S7 and S8) that correlate with massive opening of the membrane tunnels 2b, 2d and 4 (Table S4) observed at the end of the segment. The average relative binding free energy of the lipid in the second segment is slightly favorable (-0.2 kcal/mol), explaining relative stability of this binding position and also low fluctuation on the HEME–DLPC distance. The membrane tunnels 2b, 2d and 4 continue to be opened frequently also in the third trajectory segment, which starts at 190 ns when the phosphate and choline group of the phospholipid molecule enters opening in tunnel 2d and the tip of the aliphatic chain lunges further toward the substrate binding cavity, reaching the shortest (10.7 Å) distance from heme iron at 216 ns. High mobility of the intruding lipid that is typical for this trajectory segment is accompanied with unfavorable relative binding energy if the lipid (1.3 kcal/mol). Further progress of the DLPC molecule into the proximity of the heme-cofactor was hindered probably by the interaction of its charged and bulky phosphatidylcholine moiety with the surrounding protein residues. At 280 ns the hydrophobic tail of the DLPC molecule partially retreated from the binding cavity and its tip bent toward a hydrophobic pocket between FG loop and β -sheet 1, where it remained till the simulation end. In this last trajectory segment, the protein responded by decreasing the occurrence of tunnels 2b, 2d and 5, while the high opening probability of tunnel 4 was maintained. The relative

binding energy of the lipid remained in the last segment slightly unfavorable (0.4 kcal/mol). The overall relative lipid binding free energy determined using LIE approximation across productive segments 2, 3, and 4 was 0.4 kcal/mol.

Interestingly, the equilibrium constant for the lipid penetration into tunnel 2d can be independently estimated as 0.37 by evaluating the ratio of the occurrences of the bound and unbound lipid states along all five full-length membrane bound P450 1A2 MD trajectories. This approach yields 0.6 kcal/mol as an estimate of the relative binding free energy for lipid inside the tunnel 2d, which reasonably agrees with the corresponding LIE estimate of 0.4 kcal/mol. Both these results indicate that lipid insertion onto tunnel 2d is thermodynamically and kinetically viable and might occur in membrane bound enzyme with a microsecond frequency, therefore, faster than a typical P450 catalytic rate (milliseconds to minutes).

In an effort to reproduce the spontaneous intrusion of a membrane lipid molecule into the full-length P450 1A2 we evaluated additional ten replica simulations of all-atom MD simulations (total length nearly 500 ns) at an elevated system temperature (333 K) to increase the sampling. The raised simulation temperature (in this case higher by 23 degrees than the physiological temperature) could destabilize the system. However, we detected only an increased mobility of the membrane phospholipids, while the ternary structure of the enzyme remained unaltered. The RMSD of the catalytic domain was at 333 K 2.51 ± 0.07 Å, which is only slightly more than 2.15 ± 0.05 Å observed at 310 K. During these simulations, we noticed two instances of a lipid molecule entering for a short time the tunnel 2d. In one of the replicas, a fatty acid chain of DLPC reached distance of 20.8 Å from the heme cofactor (Figure S9) and in another replica, a DLPC molecule intruded into tunnel 2d up to distance of 21.7 Å from the heme iron (data not shown). Although these results should be considered with caution, they support our hypothesis that tunnel 2d is likely the place where a membrane-associated hydrophobic molecule can enter the P450 active site.

The tunnel 2d was previously identified as a minor tunnel in two bacterial cytochromes P450 102A1 (P450BM3) and P450 158A2 and in 2 crystal structures where the eukaryotic cytochrome P450 2B4 was caught in the open conformation (2BDM and 1PO5).^{39,63}, but little attention was paid to this tunnel. We examined structures of these P450s and found that the tunnel 2d is in the structure 2BDM occupied by a detergent molecule, 5-cyclohexyl-1-pentyl-beta-d-maltoside (CM5). The position of an aliphatic chain of CM5 is in this enzyme remarkably similar to that of the fatty acid chain of DLPC in P450 1A2 (Figure 9). We also noticed that there are three other CM5 molecules co-crystallized with P450 2B4 in structure 2BDM; one is positioned in the tunnel w, another in the tunnel 2e entrance, and the last one is located in a pocket formed between helices G, H and I. Opening of several tunnels including the tunnel 2d was recently reported also for *N*-terminally truncated apo-form of P450 2C9 containing the warfarin ligand.⁷⁵ This opening was detected during an MD simulation that was carried out in absence of a phospholipid membrane using an extremely small probe radius (0.75 Å).

As we mentioned in the Introduction, the tunnel classification as membrane facing or solvent facing strongly relates to the estimated enzyme positioning with respect to the membrane; another disadvantage of the current nomenclature is that it does not distinguish the two physico-chemically different regions of the phospholipid membrane – the highly polar head-group region and the central lipophilic layer. In our P450 1A2 model, the membrane-facing tunnels 2b, 4 and 5 are actually facing the polar head-group region of the membrane. It is only the tunnel 2d that in the present model exits directly into the hydrophobic membrane segment near the fatty acid residues of DLPC (see Figure 7 or interactive 3D model in the Figure S10). Thus, we believe that the observed spontaneous lipid intrusion into the tunnel 2d could mimic a lipophilic substrate entering the active site from the phospholipid membrane. For example, an arachidonic acid, which is hydroxylated by P4501A2⁷⁶, could enter the

enzyme in a manner similar to lauric acid residue of DLPC molecule observed here. Arachidonic acid is longer and is not hindered by a polar head group region like DLPC. Consequently, it could proceed through the whole tunnel 2d and reach the catalytic site of P450 1A2.

Membrane lipids are known to be essential for restoring activity of purified eukaryotic cytochromes P450. Here we observed significant opening of several membrane-facing tunnels induced by a DLPC molecule, therefore we can speculate on another role of phospholipids in stimulating the P450 activity. That is, phospholipids might selectively increase permeability of membrane-facing tunnels for hydrophobic substrates accumulated in the membrane

Conclusions

We constructed the model of membrane bound full-length P450 1A2 by a combined application of coarse-grained and all-atom molecular dynamic techniques. During this process, we were able to decrease the number of *ad hoc* structural decisions, which could potentially affect the final structure, by exploiting the self-organizational ability of the randomized water/lipid/protein solutions described by the MARTINI force field. The structure of the full-length P450 1A2 is very important for elucidating the structure of the complex between cytochrome P450 1A2 and other reticular proteins, e.g. cytochrome *b*₅, that can modulate its catalytic function.

The resulting model shows two interesting structural features. First, the P450 1A2 TM helix was found to reside below the catalytic domain, directly interacting with its highly hydrophobic *N*-terminal proline-rich segment. This TM helix orientation is consistent with the crystallographically determined structure of the TM domain of the distant relative P450 51A1. Second, the shallow membrane immersion of the catalytic domain induces a depression on the opposite side of the phospholipid bilayer beneath the catalytic domain.

The final model was extensively optimized using long MD simulations of multiple replicas. One of these replicas sampled an interesting rare event - spontaneous intrusion of the fatty acid chain of a phospholipid molecule into the tunnel 2d reaching toward the enzyme active site. The opening probability of some other membrane-facing tunnels was significantly enhanced in concert with this event. The presented results indicate that the tunnel 2d needs more attention, as it might be the sought-after path of hydrophobic substrates or inhibitors leading to the buried active site of P450.

Acknowledgements

The authors would like to thank to the Charles University in Prague (grant UNCE 204025/2012) and to the National Institutes of Health (grant 5U19CA177547). Access to the MetaCentrum and CERIT-SC computing facilities provided under the program LM2010005 and CZ. 1.05/3.2.00/08.0144 is highly appreciated.

Associated Content

Supporting Information

Simulation protocols for CG and all-atom models of the truncated and the membrane-bound P450 1A2, probability of tunnel opening in P450 1A2 calculated for individual trajectories, definition, convergence and distribution of geometric parameters used for conversion of CG-model to the all-atom model, comparison of flexibility of the membrane-bound and membrane-free P450 1A2, structural and sequence alignment of the catalytic domain of P450 1A2 with lanosterol 14 α -demethylase, structural rearrangement of F helix and FG loop upon DLPC molecule intruding into the catalytic domain, change in flexibility of the membrane-bound P450 1A2 upon DLPC molecule intrusion, spontaneous intrusion of a DLPC into the

tunnel 2d reproduced at elevated temperature, and interactive 3D model of the membrane-bound P450 1A2.

Abbreviations

P450, cytochrome P450; CG, coarse-grained; MD, molecular dynamics; SMD, steered molecular dynamics; PDB, Protein Data Bank; RMSF, root-mean square fluctuation; DLPC, dilauroyl phosphatidylcholine; TM, trans-membrane; LIE, linear interaction energy

REFERENCES

- 1 E. F. Johnson and C. D. Stout, *J. Biol. Chem.*, 2013, **288**, 17082–17090.
- 2 S. Rendic and F. P. Guengerich, *Chem. Res. Toxicol.*, 2015, **28**, 38–42.
- 3 S.-F. Zhou, B. Wang, L.-P. Yang and J.-P. Liu, *Drug Metab. Rev.*, 2010, **42**, 268–354.
- 4 S. Rendic and F. P. Guengerich, *Chem. Res. Toxicol.*, 2012, **25**, 1316–1383.
- 5 V. Kotrbová, B. Mrázová, M. Moserová, V. Martínek, P. Hodek, J. Hudeček, E. Frei and M. Stiborová, *Biochem. Pharmacol.*, 2011, **82**, 669–680.
- 6 M. Stiborová, V. Martínek, H. Rýdlová, P. Hodek and E. Frei, *Cancer Res.*, 2002, **62**, 5678–5684.
- 7 M. Stiborová, V. Martínek, E. Frei, V. M. Arlt and H. H. Schmeiser, *Curr. Drug Metab.*, 2013, **14**, 695–705.
- 8 M. Stiborová, F. Bárta, K. Levová, P. Hodek, H. H. Schmeiser, V. M. Arlt and V. Martínek, *Int. J. Mol. Sci.*, 2015, **16**, 27561–27575.
- 9 L. Ripa, C. Mee, P. Sjö and I. Shamovsky, *Chem. Res. Toxicol.*, 2014, **27**, 265–278.
- 10 F. P. Guengerich, *J. Biochem. Mol. Toxicol.*, 2007, **21**, 163–168.
- 11 T. Laursen, K. Jensen and B. L. Møller, *Biochim. Biophys. Acta*, 2011, **1814**, 132–138.
- 12 T. D. Porter, *J. Biochem. Mol. Toxicol.*, 2002, **16**, 311–316.
- 13 P. Jeřábek, J. Florián, M. Stiborová and V. Martínek, *Biochemistry*, 2014, **53**, 6695–6705.
- 14 A. Das and S. G. Sligar, *Biochemistry*, 2009, **48**, 12104–12112.
- 15 H. W. Strobel, A. Y. H. Lu, J. Heidema and M. J. Coon, *J. Biol. Chem.*, 1970, **245**, 4851–4854.
- 16 A. Nath, Y. V. Grinkova, S. G. Sligar and W. M. Atkins, *J. Biol. Chem.*, 2007, **282**, 28309–28320.
- 17 T. Ahn, C.-H. Yun and D.-B. Oh, *Biochemistry*, 2005, **44**, 9188–9196.
- 18 M. Sotomayor and K. Schulten, *Biophys. J.*, 2004, **87**, 3050–3065.
- 19 F. Khalili-Araghi, E. Tajkhorshid and K. Schulten, *Biophys. J.*, 2006, **91**, L72-74.
- 20 R. O. Dror, D. H. Arlow, D. W. Borhani, M. O. Jensen, S. Piana and D. E. Shaw, *Proc. Natl. Acad. Sci. U. S. A.*, 2009, **106**, 4689–4694.
- 21 H. Nury, F. Poitevin, C. Van Renterghem, J.-P. Changeux, P.-J. Corringer, M. Delarue and M. Baaden, *Proc. Natl. Acad. Sci. U. S. A.*, 2010, **107**, 6275–6280.

- 22 M. O. Jensen, V. Jogini, D. W. Borhani, A. E. Leffler, R. O. Dror and D. E. Shaw, *Science*, 2012, **336**, 229–233.
- 23 X. Periole, T. Huber, S.-J. Marrink and T. P. Sakmar, *J. Am. Chem. Soc.*, 2007, **129**, 10126–10132.
- 24 P. J. Bond, D. L. Parton, J. F. Clark and M. S. P. Sansom, *Biophys. J.*, 2008, **95**, 3802–3815.
- 25 T. Carpenter, P. J. Bond, S. Khalid and M. S. P. Sansom, *Biophys. J.*, 2008, **95**, 3790–3801.
- 26 C.-C. Chen and C.-M. Chen, *J. Struct. Biol.*, 2009, **165**, 37–46.
- 27 F. Tavanti and V. Tozzini, *Molecules*, 2014, **19**, 14961–14978.
- 28 C. L. Wee, D. Gavaghan and M. S. P. Sansom, *Biophys. J.*, 2010, **98**, 1558–1565.
- 29 P. J. Stansfeld and M. S. P. Sansom, *J. Chem. Theory Comput.*, 2011, **7**, 1157–1166.
- 30 D. Fishelovitch, S. Shaik, H. J. Wolfson and R. Nussinov, *J. Phys. Chem. B*, 2009, **113**, 13018–13025.
- 31 M. Otyepka, K. Berka and P. Anzenbacher, *Curr. Drug Metab.*, 2012, **13**, 130–142.
- 32 K. Berka, T. Hendrychová, P. Anzenbacher and M. Otyepka, *J. Phys. Chem. A*, 2011, **115**, 11248–11255.
- 33 V. Cojocar, K. Balali-Mood, M. S. P. Sansom and R. C. Wade, *PLoS Comput. Biol.*, 2011, **7**, e1002152.
- 34 X. Yu, V. Cojocar, G. Mustafa, O. M. H. Salo-Ahen, G. I. Lepesheva and R. C. Wade, *J. Mol. Recognit.*, 2015, **28**, 59–73.
- 35 R. Lonsdale, S. L. Rouse, M. S. P. Sansom and A. J. Mulholland, *PLoS Comput. Biol.*, 2014, **10**, e1003714.
- 36 J. L. Baylon, I. L. Lenov, S. G. Sligar and E. Tajkhorshid, *J. Am. Chem. Soc.*, 2013, **135**, 8542–8551.
- 37 P. Jeřábek, J. Florián and V. Martínek, *Chem. Res. Toxicol.*, 2016, **29**, 626–636.
- 38 A. Gora, J. Brezovsky and J. Damborsky, *Chem. Rev.*, 2013, **113**, 5871–5923.
- 39 S. K. Ludemann, V. Lounnas and R. C. Wade, *J. Mol. Biol.*, 2000, **303**, 797–811.
- 40 S. K. Ludemann, V. Lounnas and R. C. Wade, *J. Mol. Biol.*, 2000, **303**, 813–830.
- 41 K. Schleinkofer, Sudarko, P. J. Winn, S. K. Ludemann and R. C. Wade, *EMBO Rep.*, 2005, **6**, 584–589.
- 42 P. J. Winn, S. K. Ludemann, R. Gauges, V. Lounnas and R. C. Wade, *Proc. Natl. Acad. Sci. U. S. A.*, 2002, **99**, 5361–5366.

- 43 X. Yu, V. Cojocaru and R. C. Wade, *Biotechnol. Appl. Biochem.*, 2013, **60**, 134–145.
- 44 S. J. Marrink, H. J. Risselada, S. Yefimov, D. P. Tieleman and A. H. de Vries, *J. Phys. Chem. B*, 2007, **111**, 7812–7824.
- 45 L. Monticelli, S. K. Kandasamy, X. Periole, R. G. Larson, D. P. Tieleman and S.-J. Marrink, *J. Chem. Theory Comput.*, 2008, **4**, 819–834.
- 46 J. C. Phillips, R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R. D. Skeel, L. Kalé and K. Schulten, *J. Comput. Chem.*, 2005, **26**, 1781–1802.
- 47 R. Bradley and R. Radhakrishnan, *Polymers*, 2013, **5**, 890–936.
- 48 S. Vicatos, A. Rychkova, S. Mukherjee and A. Warshel, *Proteins*, 2014, **82**, 1168–1185.
- 49 I. Vorobyov, I. Kim, Z. T. Chu and A. Warshel, *Proteins*, 2016, **84**, 92–117.
- 50 A. P. Lyubartsev and A. L. Rabinovich, *Biochim. Biophys. Acta*, 2016, **1858**, 2483–2497.
- 51 X. Periole, M. Cavalli, S.-J. Marrink and M. A. Ceruso, *J. Chem. Theory Comput.*, 2009, **5**, 2531–2543.
- 52 T. Darden, D. York and L. Pedersen, *J. Chem. Phys.*, 1993, **98**, 10089–10092.
- 53 J.-P. Ryckaert, G. Ciccotti and H. J. Berendsen, *J. Comput. Phys.*, 1977, **23**, 327–341.
- 54 MacKerell, D. Bashford, Bellott, Dunbrack, J. D. Evanseck, M. J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, D. Joseph-McCarthy, L. Kuchnir, K. Kuczera, F. T. K. Lau, C. Mattos, S. Michnick, T. Ngo, D. T. Nguyen, B. Prodhom, W. E. Reiher, B. Roux, M. Schlenkrich, J. C. Smith, R. Stote, J. Straub, M. Watanabe, J. Wiórkiewicz-Kuczera, D. Yin and M. Karplus, *J. Phys. Chem. B*, 1998, **102**, 3586–3616.
- 55 J. B. Klauda, R. M. Venable, J. A. Freites, J. W. O'Connor, D. J. Tobias, C. Mondragon-Ramirez, I. Vorobyov, A. D. MacKerell and R. W. Pastor, *J. Phys. Chem. B*, 2010, **114**, 7830–7843.
- 56 L. Martínez, R. Andrade, E. G. Birgin and J. M. Martínez, *J. Comput. Chem.*, 2009, **30**, 2157–2164.
- 57 S. Sansen, J. K. Yano, R. L. Reynald, G. A. Schoch, K. J. Griffin, C. D. Stout and E. F. Johnson, *J. Biol. Chem.*, 2007, **282**, 14348–14355.
- 58 W. Humphrey, A. Dalke and K. Schulten, *J. Mol. Graph.*, 1996, **14**, 33–38.
- 59 A. Krogh, B. Larsson, G. von Heijne and E. L. L. Sonnhammer, *J. Mol. Biol.*, 2001, **305**, 567–580.
- 60 T. Nugent and D. T. Jones, *BMC Bioinformatics*, 2009, **10**, 159.
- 61 A. Y. Shih, P. L. Freddolino, S. G. Sligar and K. Schulten, *Nano Lett.*, 2007, **7**, 1692–1696.

- 62 E. Chovancova, A. Pavelka, P. Benes, O. Strnad, J. Brezovsky, B. Kozlikova, A. Gora, V. Sustr, M. Klvana, P. Medek, L. Biedermannova, J. Sochor and J. Damborsky, *PLoS Comput. Biol.*, 2012, **8**, e1002708.
- 63 V. Cojocaru, P. J. Winn and R. C. Wade, *Biochim. Biophys. Acta*, 2007, **1770**, 390–401.
- 64 J. Aqvist, C. Medina and J. E. Samuelsson, *Protein Eng.*, 1994, **7**, 385–391.
- 65 D. L. Mobley and J. P. Guthrie, *J. Comput. Aided Mol. Des.*, 2014, **28**, 711–720.
- 66 T. Hansson, J. Marelus and J. Aqvist, *J. Comput. Aided Mol. Des.*, 1998, **12**, 27–35.
- 67 U. Bren, J. Lah, M. Bren, V. Martínek and J. Florián, *J. Phys. Chem. B*, 2010, **114**, 2876–2885.
- 68 B. C. Monk, T. M. Tomasiak, M. V. Keniya, F. U. Huschmann, J. D. A. Tyndall, J. D. O’Connell, R. D. Cannon, J. G. McDonald, A. Rodriguez, J. S. Finer-Moore and R. M. Stroud, *Proc. Natl. Acad. Sci.*, 2014, **111**, 3865–3870.
- 69 H. Yamazaki and T. Shimada, *Methods Mol. Biol. Clifton NJ*, 2006, **320**, 61–71.
- 70 R. Indra, M. Moserova, N. Kroftova, M. Sulc, M. Martinkova, V. Adam, T. Eckschlager, R. Kizek, V. M. Arlt and M. Stiborova, *Neuro Endocrinol. Lett.*, 2014, **35**, 105–113.
- 71 R. A. Copeland, D. L. Pompliano and T. D. Meek, *Nat. Rev. Drug Discov.*, 2006, **5**, 730–739.
- 72 S. Sansen, J. K. Yano, R. L. Reynald, G. A. Schoch, K. J. Griffin, C. D. Stout and E. F. Johnson, *J. Biol. Chem.*, 2007, **282**, 14348–14355.
- 73 Y.-L. Cui, Q. Xue, Q.-C. Zheng, J.-L. Zhang, C.-P. Kong, J.-R. Fan and H.-X. Zhang, *Biochim. Biophys. Acta*, 2015, **1848**, 2013–2021.
- 74 W. Li, J. Shen, G. Liu, Y. Tang and T. Hoshino, *Proteins*, 2011, **79**, 271–281.
- 75 L. J. Kingsley and M. A. Lill, *PLOS ONE*, 2014, **9**, e99408.
- 76 J. R. Falck, S. Lumin, I. Blair, E. Dishman, M. V. Martin, D. J. Waxman, F. P. Guengerich and J. H. Capdevila, *J. Biol. Chem.*, 1990, **265**, 10244–10249.
- 77 S. Henikoff and J. G. Henikoff, *Proc. Natl. Acad. Sci. U. S. A.*, 1992, **89**, 10915–10919.
- 78 Y. Zhao, M. A. White, B. K. Muralidhara, L. Sun, J. R. Halpert and C. D. Stout, *J. Biol. Chem.*, 2006, **281**, 5973–5981.

Tables

Table 1. Probability of tunnel opening observed during all-atom MD simulations of the membrane-bound full-length P450 1A2. All tunnels were detected using methane-sized probe with a radius of 1.9 Å.

Tunnel	Entrance-lining elements	Opening probability [%] ^a		Max. bottleneck radius [Å] ^b	
		Membrane-free ^c	Membrane-bound ^d	Membrane-free	Membrane-bound
Membrane-facing ^e					
2b	FG loop and β -sheet 3	1	5	2.3	2.8
2d	Proline-rich segment and A helix	0	5		2.7
4	FG loop	< 1	8	2.1	2.8
5	A and K helices, and β -sheet	0	< 1		2.2
Solvent-facing					
2c	G and I helices, and BC loop	28	16	2.4	2.5
2e	BC loop	0	1		2.5
s	C and F helix, β -sheet 3	4	1	2.4	2.1
w	C helix and conserved loop	1	0	2.3	

^a average error of the opening probability assignment was estimated to be 1 percentage point

^b average error of the maximal bottleneck radius assignment was estimated to be 0.1 Å

^c total simulation time was 1150 ns

^d total simulation time was 1340 ns

^e tunnels 2b, 4 and 5 are facing the polar head-group region of the membrane while the tunnel 2d exit is positioned close to the lipophilic segment of the DLPC membrane

Figure Legends

Figure 1. Schematic diagram of the full-length membrane-bound P450 1A2 model assembly strategy. The CG models of P450 1A2 catalytic and TM domains are shown in black and blue, respectively, and the same domains are in all-atom model depicted in magenta and lime. The blue dots represent the water solvent including ions. For a detailed description of individual steps, see Methods.

Figure 2. Orientation of the catalytic (A) and TM (B) domains of P450 1A2 after membrane-protein self-assembly from water/DLPC mixture during CG MD simulations. The graph shows distribution of the population of TM helix tilt angle γ . Protein backbone and its side chain beads are shown as black sticks and colored balls, respectively; ball colors denote amino acid character (white – nonpolar, dark gray – aromatic, orange – polar, red – negative charge, blue – positive charge). A, G, and I helices are shown in pink, green and violet, respectively. Proline-rich segment and FG loop are shown in cyan and yellow, respectively. DLPC molecules are shown as light grey lines; beads representing phosphate groups are shown as light gray balls. Water molecules are not shown.

Figure 3. Structural alignment of full-length P450 1A2 models established during four independent CG MD simulations; the colored cylinders indicate four input TM helix geometries; the corresponding final geometries are shown by curved tubes. Red arrow indicates the TM helix – catalytic domain separation at the start of each simulation. For clarity, the catalytic domain is shown only for one model.

Figure 4. Structural and sequence alignment of *N*-terminal segments of eukaryotic P450 families 1, 2, and 3. The proline-rich segment is located inside the red ellipse. P450 1A2 residues P43, P45, P49, and G52, which are absolutely conserved, are depicted by sticks. Residues in the sequence alignment are colored according to the BLOSUM62 matrix.⁷⁷

Figure 5. Orientation and mobility of the P450 1A2 TM helix relative to the P450 1A2 catalytic domain (A) and X-ray structure of 14 α -lanosterol demethylase (P450 51A1) (B). TM domain of P450 1A2 that was sampled by the unconstrained all-atom MD of the full-length P450 1A2 complex is shown as orange transparent cylinders. Residues 34 to 513 of the P450 1A2 catalytic domain were aligned in the each step to the initial structure. A, G, and I helices are shown in pink, green and violet, respectively. Proline-rich segment and FG loop are shown in cyan and yellow, respectively. DLPC molecules are shown as light grey lines. Heme cofactors of P450 1A2 and 14 α -lanosterol demethylase are shown in ball-sticks representation and colored according to atom types. Dashed gray lines approximates polar head groups region of a membrane embedding 14 α -lanosterol demethylase.

Figure 6. Membrane deformation induced by P450 1A2 presence, found in CG and all-atom MD simulations. Average membrane-indentation profile of the lower layer of the phospholipid bilayer in the CG MD simulated system containing the P450 1A2 catalytic domain in the absence (A) and presence (B and C) of the TM domain. The deformation of the lower DLPC layer is visualized as vertical shift of phosphate atoms relative to the average phosphate position. Zones with DLPC molecules significantly elevated from the membrane plane are indicated by blue areas in the contour plot. The first two plots were obtained by averaging over last 200 ns of the CG MD simulations (A and B). The membrane-indentation

profile for all-atom MD simulated full-length P450 1A2 was obtained by averaging over trajectories of all five replicas with total simulation time of 1400 ns (C). Individual phosphates are shown as gray balls. A, G, and I helices are shown in pink, green and violet, respectively. The proline-rich segment and FG loop of the catalytic domain are shown in cyan and yellow, respectively.

Figure 7. Access/egress tunnels leading to/from the active site of P450 1A2. Entrances and paths of individual tunnels are indicated by balls and curved sticks, respectively. Tunnels were named according to nomenclature introduced by Cojocaru et al.⁶³ A, G, I, and TM helices and FG loop of P450 1A2 are shown in black. The heme cofactor of P450 1A2 is depicted in the atom-type colored ball-and-stick representation. Phosphate groups of DLPC molecules are shown as light gray balls, and their aliphatic chains as well as solvent molecules are not shown. For convenient visualization of the presented model, we recommend examining our interactive 3D model of the membrane-bound P450 1A2 model provided in the Supporting Information file (Figure S10).

Figure 8. Correlation between lipid penetration into tunnel 2d, lipid relative binding free energy and opening frequency of major tunnels in P450 1A2. The upper part shows structural snapshots representing four segments with distinct binding positions of phospholipid molecule (A). The penetration progress, monitored as time evolution of distance of DLPC molecule from heme cofactor (B) and also as the relative lipid binding free energy (C), is correlated with opening/closing events of important tunnels along an all-atom MD trajectory 1m (D). Tunnel opening was monitored as time evolution of the bottlenecks of detected tunnels using minimal cutoff of 1.9 Å. The heme–DLPC distance was measured between the

terminal carbon of aliphatic chain of the DLPC molecule and the Fe atom of heme cofactor. A, F, G, and I helices are shown in pink, purple, green and violet, respectively. Proline-rich segment and FG loop are shown in cyan and yellow, respectively. The TM domain is shown in orange. DLPC molecule (grey carbon atoms) and heme cofactor (cyan carbon atoms) are shown in sticks representation and colored according to atom types.

Figure 9. Comparison of a DLPC molecule bound in tunnel 2d of human P450 1A2 (left) with a detergent 5-cyclohexyl-1-pentyl-beta-D-maltoside (CM5) bound in tunnel 2d of rabbit P450 2B4 (PDB code: 2BDM)⁷⁸ (right). The DLPC and CM5 molecules are shown in orange and pink color, respectively.

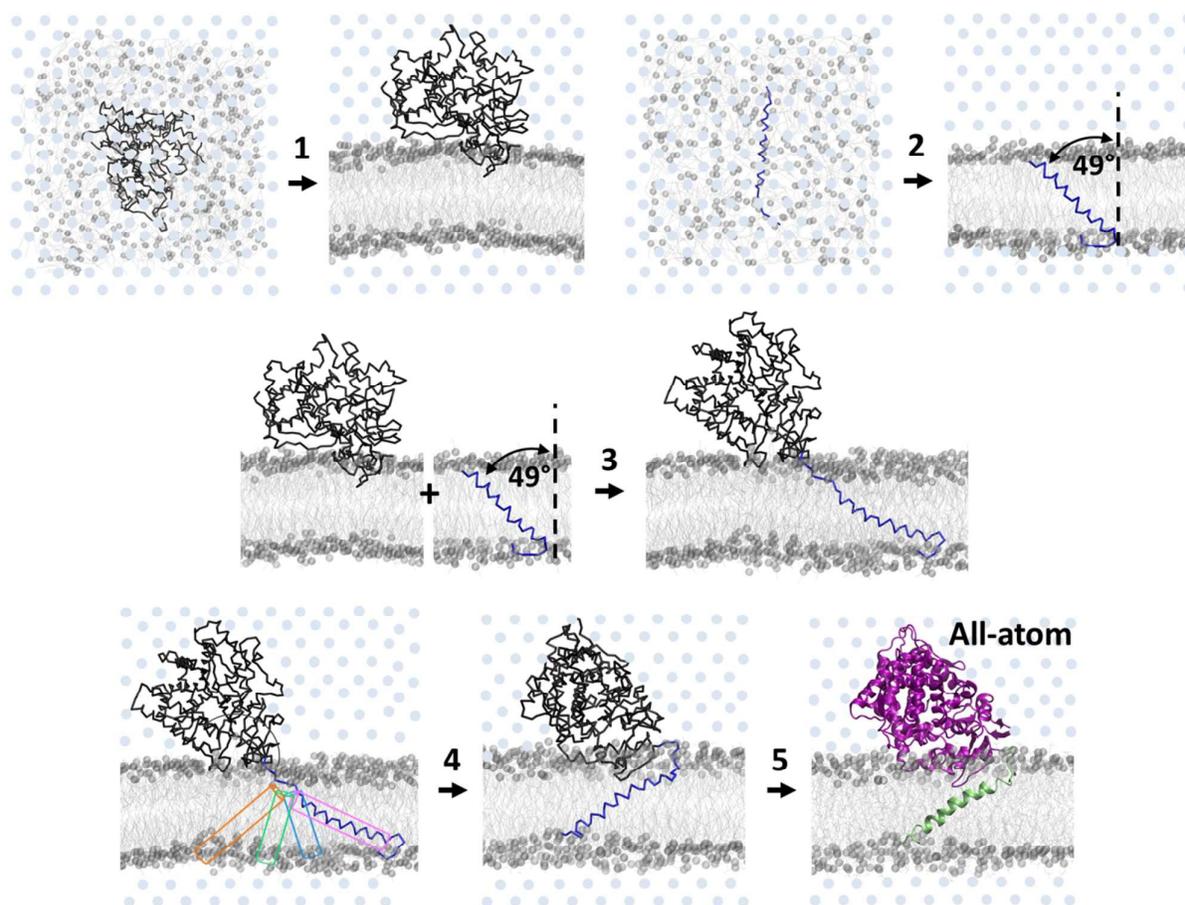


Figure 1. Schematic diagram of the full-length membrane-bound P450 1A2 model assembly strategy. The CG models of P450 1A2 catalytic and TM domains are shown in black and blue, respectively, and the same domains are in all-atom model depicted in magenta and lime. The blue dots represent the water solvent including ions. For a detailed description of individual steps, see Methods.

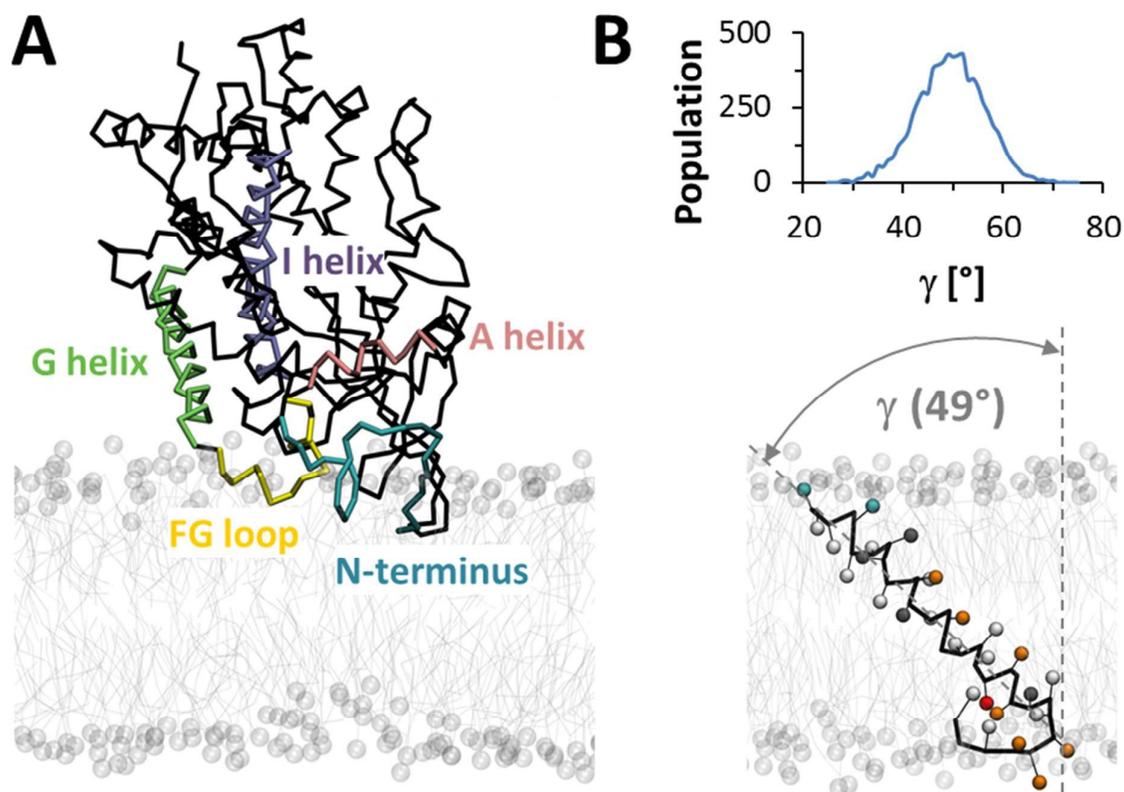


Figure 2. Orientation of the catalytic (A) and TM (B) domains of P450 1A2 after membrane-protein self-assembly from water/DLPC mixture during CG MD simulations. The graph shows distribution of the population of TM helix tilt angle γ . Protein backbone and its side chain beads are shown as black sticks and colored balls, respectively; ball colors denote amino acid character (white – nonpolar, dark gray – aromatic, orange – polar, red – negative charge, blue – positive charge). A, G, and I helices are shown in pink, green and violet, respectively. Proline-rich segment and FG loop are shown in cyan and yellow, respectively. DLPC molecules are shown as light grey lines; beads representing phosphate groups are shown as light grey balls. Water molecules are not shown.

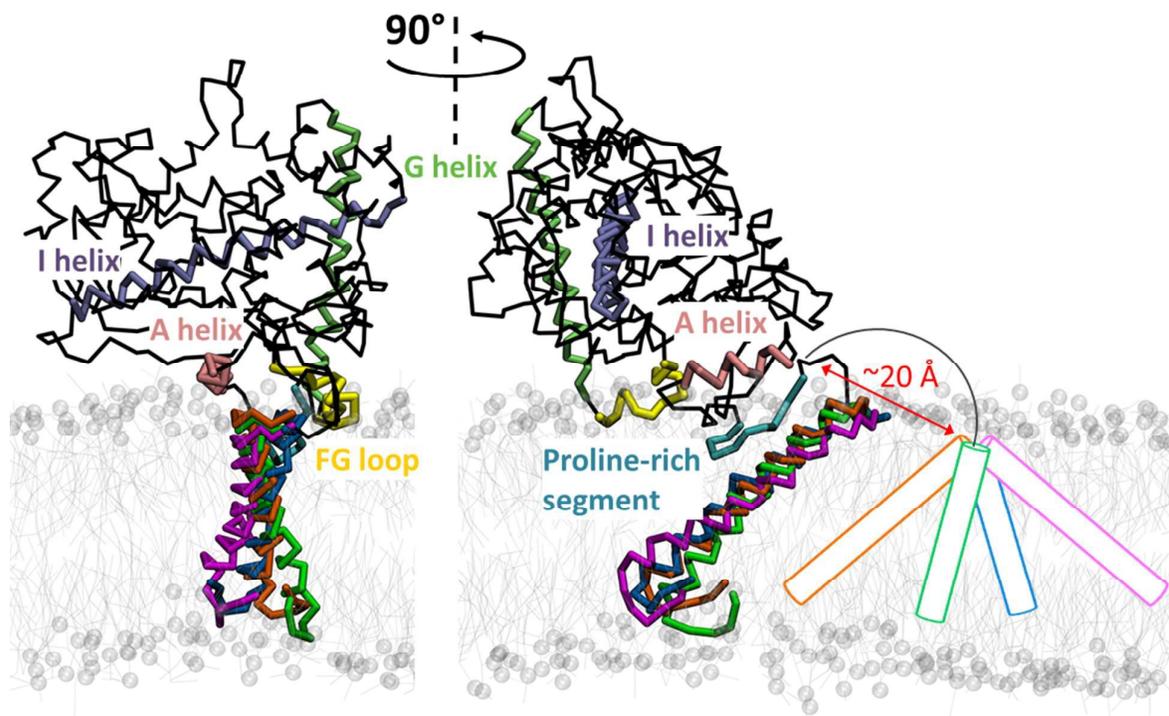


Figure 3. Structural alignment of full-length P450 1A2 models established during four independent CG MD simulations; the colored cylinders indicate four input TM helix geometries; the corresponding final geometries are shown by curved tubes. Red arrow indicates the TM helix – catalytic domain separation at the start of each simulation. For clarity, the catalytic domain is shown only for one model.

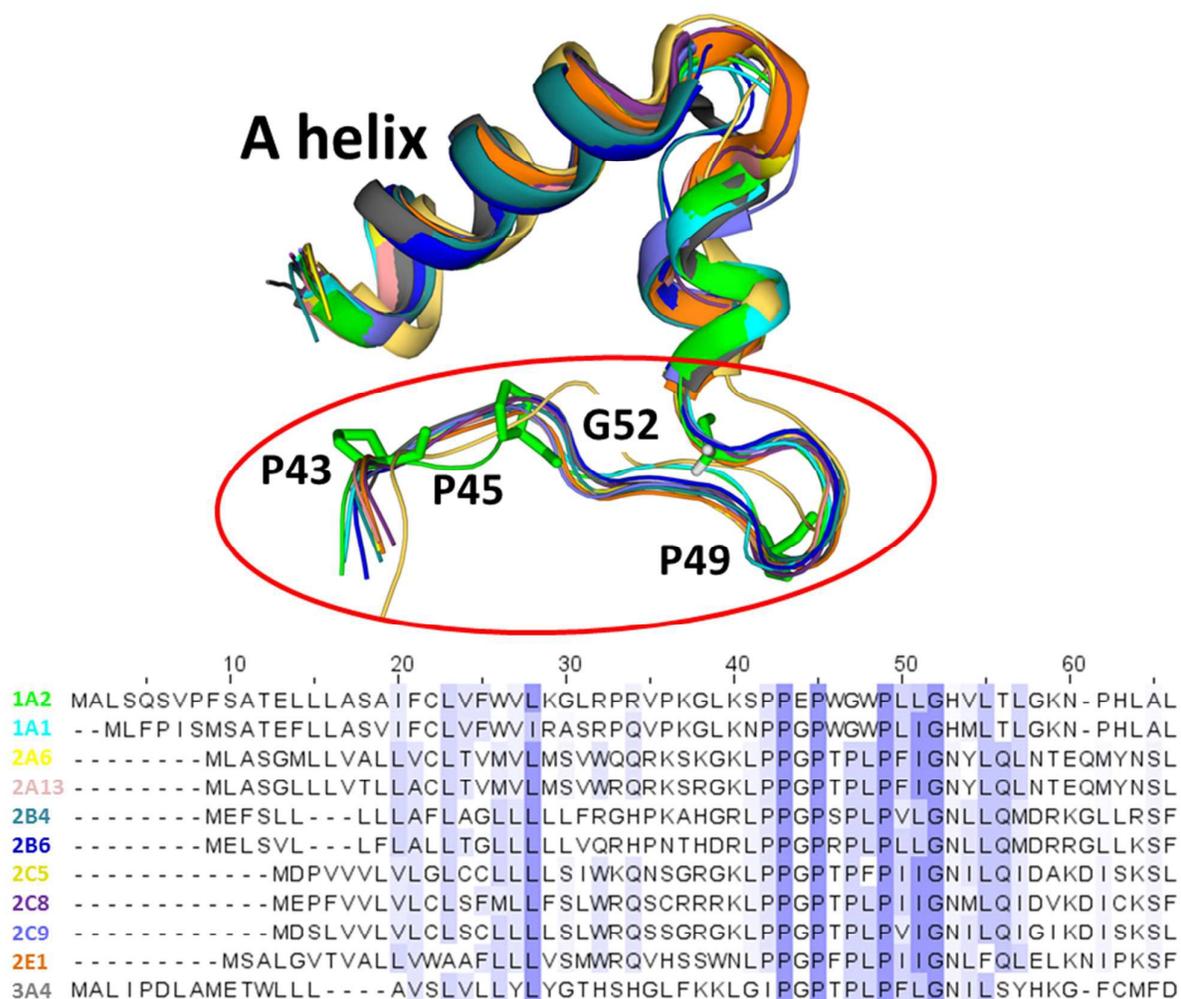


Figure 4. Structural and sequence alignment of N-terminal segments of eukaryotic P450 families 1, 2, and 3. The proline-rich segment is located inside the red ellipse. P450 1A2 residues P43, P45, P49, and G52, which are absolutely conserved, are depicted by sticks. Residues in the sequence alignment are colored according to the BLOSUM62 matrix.⁷⁷

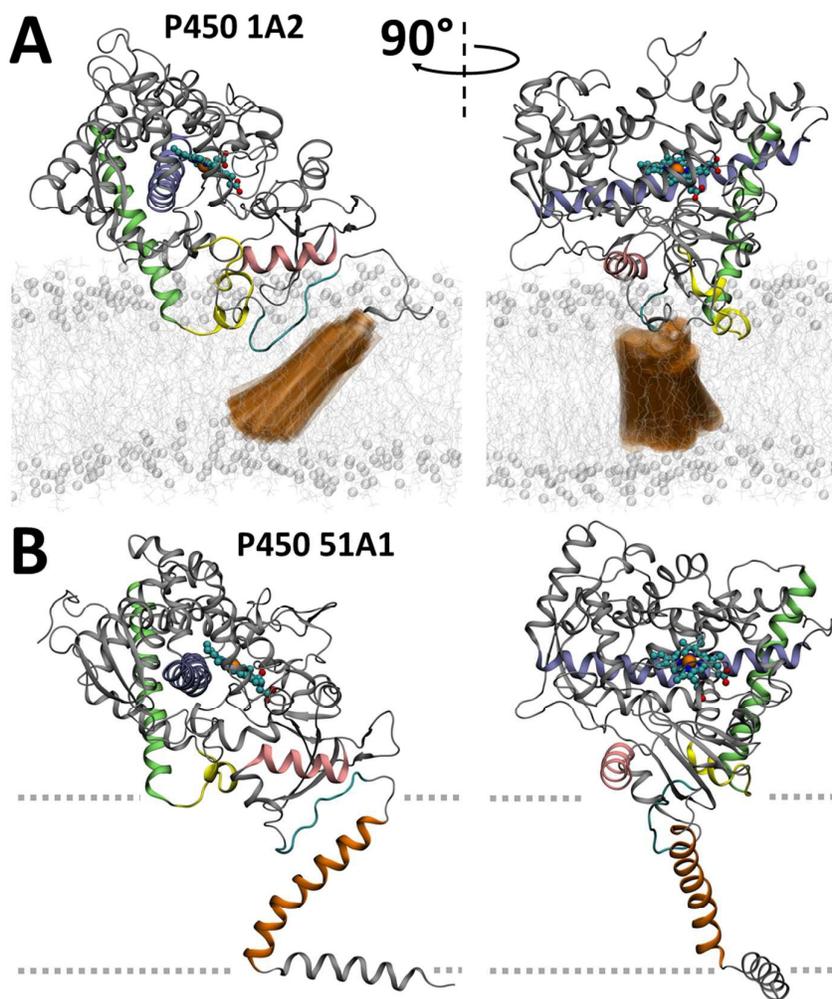


Figure 5. Orientation and mobility of the P450 1A2 TM helix relative to the P450 1A2 catalytic domain (A) and X-ray structure of 14 α -lanosterol demethylase (P450 51A1) (B). TM domain of P450 1A2 that was sampled by the unconstrained all-atom MD of the full-length P450 1A2 complex is shown as orange transparent cylinders. Residues 34 to 513 of the P450 1A2 catalytic domain were aligned in the each step to the initial structure. A, G, and I helices are shown in pink, green and violet, respectively. Proline-rich segment and FG loop are shown in cyan and yellow, respectively. DLPC molecules are shown as light grey lines. Heme cofactors of P450 1A2 and 14 α -lanosterol demethylase are shown in ball-sticks representation and colored according to atom types. Dashed gray lines approximates polar head groups region of a membrane embedding 14 α -lanosterol demethylase.

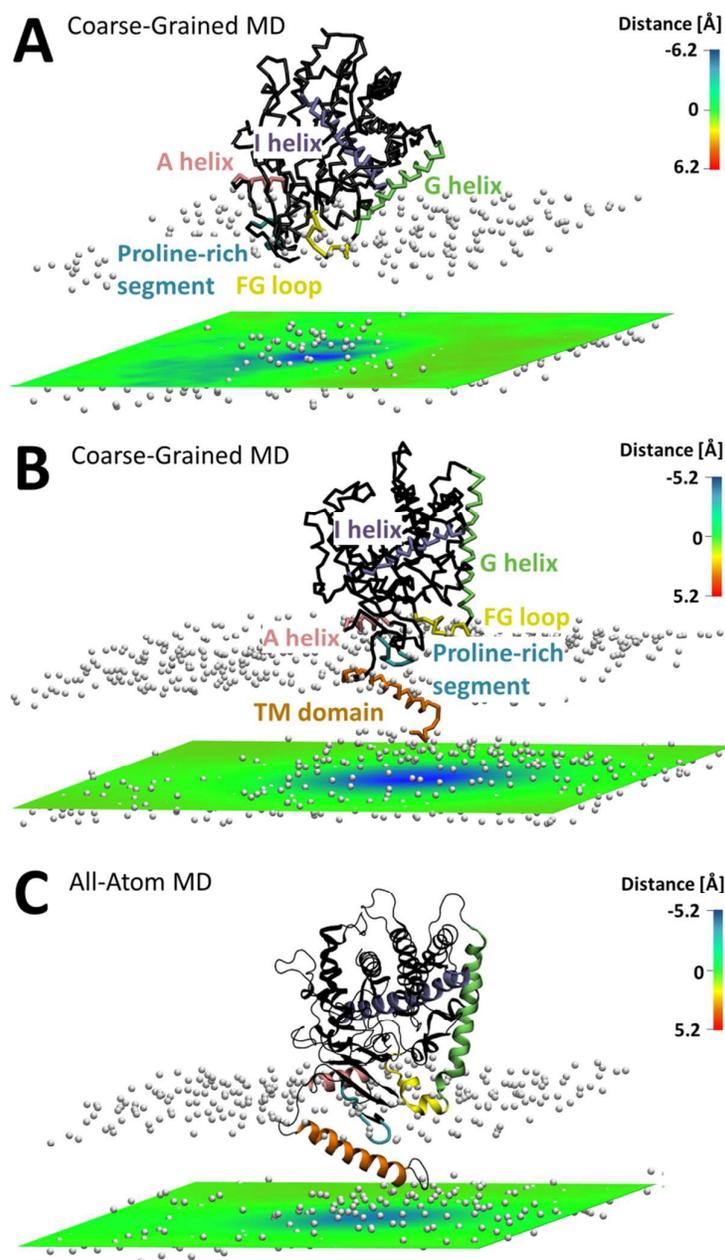


Figure 6. Membrane deformation induced by P450 1A2 presence, found in CG and all-atom MD simulations. Average membrane-indentation profile of the lower layer of the phospholipid bilayer in the CG MD simulated system containing the P450 1A2 catalytic domain in the absence (A) and presence (B and C) of the TM domain. The deformation of the lower DLPC layer is visualized as vertical shift of phosphate atoms relative to the average phosphate position. Zones with DLPC molecules significantly elevated from the membrane

plane are indicated by blue areas in the contour plot. The first two plots were obtained by averaging over last 200 ns of the CG MD simulations (A and B). The membrane-indentation profile for all-atom MD simulated full-length P450 1A2 was obtained by averaging over trajectories of all five replicas with total simulation time of 1400 ns (C). Individual phosphates are shown as gray balls. A, G, and I helices are shown in pink, green and violet, respectively. The proline-rich segment and FG loop of the catalytic domain are shown in cyan and yellow, respectively.

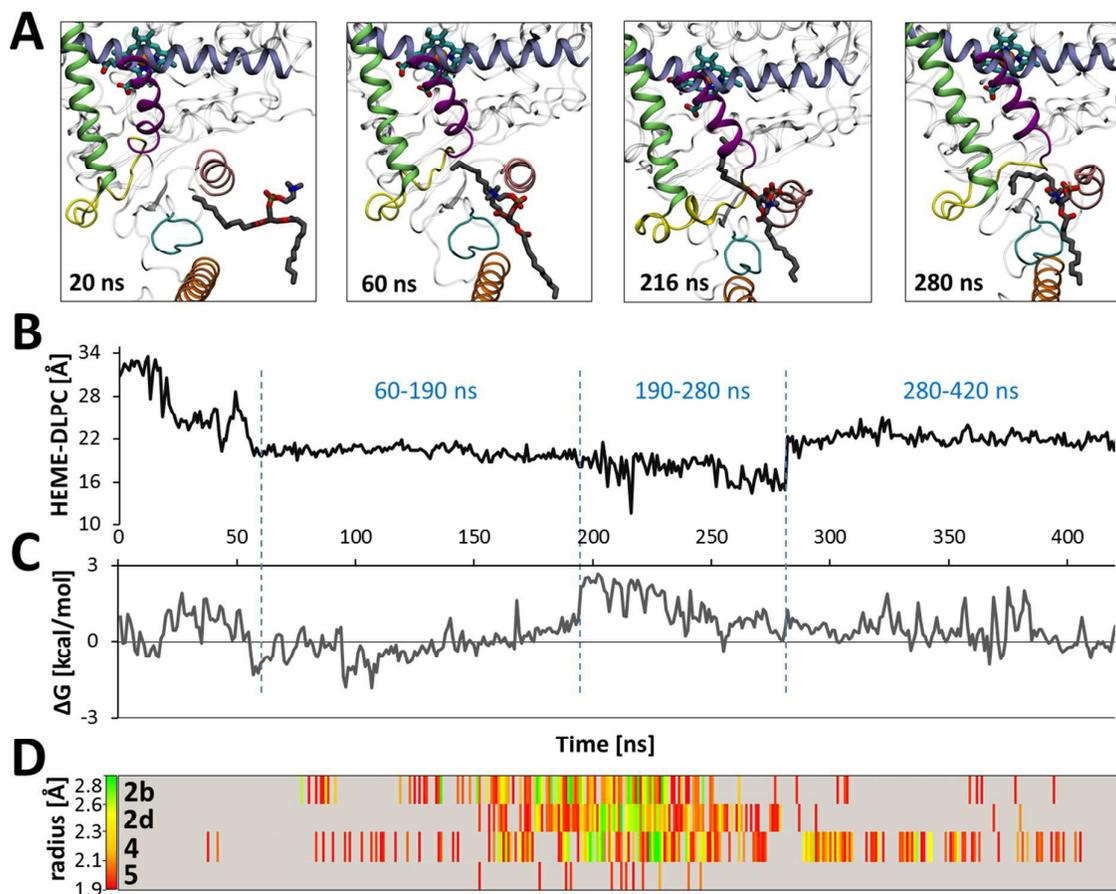


Figure 8. Correlation between lipid penetration into tunnel 2d, lipid relative binding free energy and opening frequency of major tunnels in P450 1A2. The upper part shows structural snapshots representing four segments with distinct binding positions of phospholipid molecule (A). The penetration progress, monitored as time evolution of distance of DLPC molecule from heme cofactor (B) and also as the relative lipid binding free energy (C), is correlated with opening/closing events of important tunnels along an all-atom MD trajectory 1m (D). Tunnel opening was monitored as time evolution of the bottlenecks of detected tunnels using minimal cutoff of 1.9 Å. The heme–DLPC distance was measured between the terminal carbon of aliphatic chain of the DLPC molecule and the Fe atom of heme cofactor. A, F, G, and I helices are shown in pink, purple, green and violet, respectively. Proline-rich segment and FG loop are shown in cyan and yellow, respectively. The TM domain is shown

in orange. DLPC molecule (grey carbon atoms) and heme cofactor (cyan carbon atoms) are shown in sticks representation and colored according to atom types.

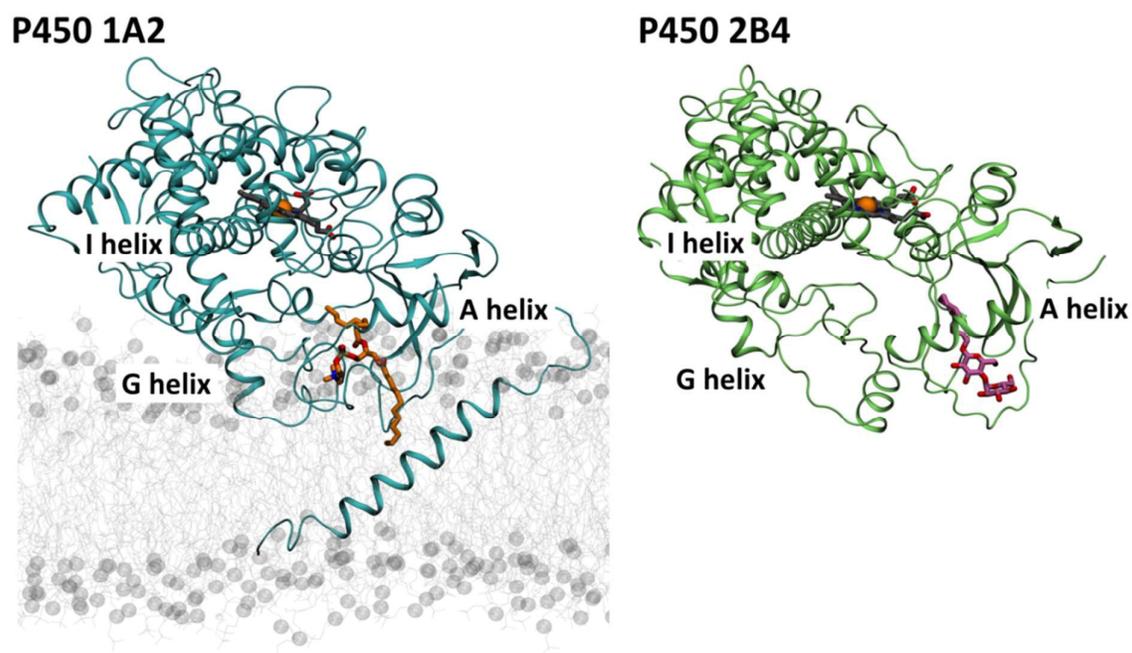


Figure 9. Comparison of a DLPC molecule bound in tunnel 2d of human P450 1A2 (left) with a detergent 5-cyclohexyl-1-pentyl-beta-D-maltoside (CM5) bound in tunnel 2d of rabbit P450 2B4 (PDB code: 2BDM)⁷⁸ (right). The DLPC and CM5 molecules are shown in orange and pink color, respectively.

For Table of Contents Use Only

Lipid molecules can induce a tunnel opening in the membrane-bound cytochrome P450 1A2

Petr Jeřábek, Jan Florián, Václav Martínek*

