Membrane-embedded substrate recognition by cytochrome P450 3A4

John C. Hackett
From the Department of Physiology and Biophysics and the Massey Cancer Center, School of Medicine, Virginia Commonwealth University, Richmond, Virginia 23298-0035

This work was supported by National Institutes of Health Grant R01GM114168 (to J. C. H.). The author declares that he has no conflicts of interest with the contents of this article. The content is solely the responsibility of the author and does not necessarily represent the official views of the National Institutes of Health.

This article contains Figs. S1–S3.

Cytochrome P450 3A4 (CYP3A4) is the dominant xenobiotic-metabolizing enzyme in the liver and intestine and is involved in the disposition of more than 50% of drugs. Because of its ability to bind multiple substrates, its reaction kinetics are complex, and its association with the microsomal membrane confounds our understanding of how this enzyme recognizes and recruits diverse substrates. Testosterone (TST) hydroxylation is the prototypical CYP3A4 reaction, displaying positive homotropic cooperativity with three binding sites. Here, exploiting the capability of accelerated molecular dynamics (aMD) to sample events in the millisecond regime, I performed >25-μs aMD simulations in the presence of three TST molecules. These simulations identified high-occupancy surface-binding sites as well as a pathway for TST ingress into the CYP3A4 active site originating in the membrane. Adaptive biasing force analysis of the latter pathway revealed a metastable intermediate that could constitute a third binding site at high TST concentrations. Prompted by the observation that interactions between TST and the G'-helix mobilize the ligand into the active site, a free-energy analysis of TST distribution in the membrane was conducted and revealed that the depth of the G'-helix is optimal for extracting TST. In summary, these simulations confirm separate, but adjacent substrate-binding sites within the enzyme and the existence of an auxiliary TST-binding site. The broader impact of these simulations is that they support a mechanism in which cytochromes P450 directly recruit membrane-solubilized substrates.

The cytochromes P450 metabolize nearly all drugs and environmental xenobiotics to which humans are exposed. Of the 57 cytochromes P450 encoded by the human genome, cytochrome P450 3A4 (CYP3A4) is involved in the clearance of more than half of approved drugs and presents significant challenge to optimizing pharmacokinetics in drug development and avoiding adverse effects in the clinic (1, 2). CYP3A4 enzyme kinetics can demonstrate marked cooperativity (3). Positive cooperativity is characterized by a continuous, positive increase in the rate of catalysis with increasing substrate concentration, whereas positive heterotropic cooperativity occurs when the rate of substrate turnover is enhanced by a second distinct effector molecule. Cooperativity in CYP catalysis was first recognized in microscopic preparations (4–6) and later confirmed using purified CYP3A4 with several substrates including testosterone (TST), 17β-estradiol, amitriptyline, and aflatoxin B1 (7). A shift in the regioselectivity of midazolam (MDZ) oxidation in patients by concurrent fluconazole administration supports that heterotropic CYP3A4 interactions are also relevant in vitro (8). Cooperativity in CYP3A4 catalysis has been explained by mechanisms invoking simultaneous occupation of the active site by multiple ligands or binding to nearby allosteric sites (9–15).

The membrane environment can have a profound effect on membrane protein function. With the exception of the N-terminal transmembrane helix, only the F'- and G'-helices of CYP3A4 are purported to interact with the membrane. Notwithstanding an only superficial interaction with the catalytic domain, the membrane imparts dramatic changes in the physical properties of the protein. Incorporation of CYP3A4 into lipid nanodiscs induces a clear increase in its thermal stability. Furthermore, the structural and dynamic changes imparted by the membrane likewise influence ligand binding affinities. Dissociation constants of some ligands decrease by an order of magnitude, and the kinetics of ligand binding are tunable by changing the lipid composition (5–7). In view of the evidence that membrane interactions modulate CYP3A4 function, it is important that they be included in models, experimental or theoretical, aimed at understanding the functional dynamics of this important enzyme.

The amphipathic nature of CYP3A4 substrates has led to speculation that the enzyme harvests them from the membrane; however, neither experimental nor theoretical evidence in direct support of this mechanism are available. Understanding how substrates are recognized by CYP3A4 and the locations of allosteric sites are required to improve predictive tools of potential of mean force; PRG, progesterone; HMMM, highly mobile membrane-mimetic; DCE, 1,1-dichloroethane; NPT, constant temperature and pressure; NVT, constant temperature and volume.
CYP3A4 substrate–binding dynamics

CYP3A4 metabolism that would be invaluable to optimizing drug therapy. The ideal computational approach to help fill this knowledge gap would permit simulation of the necessary time scales for a ligand to sufficiently explore the protein surface to identify optimal binding site(s). Accelerated molecular dynamics (aMD) is particularly advantageous for this purpose because, in combination with the recent advances in graphical processing unit–enabled hardware, it can provide access to ligand-binding events that occur on the millisecond time scale without requisite knowledge of the binding pathway (16). Hundreds-of-nanosecond aMD simulations have captured the activation of the M2 and M3 muscarinic G-protein–coupled receptors (17–19) and ligand binding to the M3 receptor (20). In this work, aMD is extended to CYP3A4 to simulate TST binding to the enzyme free of a predefined reaction coordinate. aMD successfully identifies probable locations of auxiliary binding sites that reflect experimental observations and reveals a membrane-embedded pathway for TST entrance into the active site. Free-energy analyses of the ingress pathway provides a structural basis for the kinetic behavior of CYP3A4 mutants and insight into the mechanism for exclusion of progesterone from the active site in CYP3A4 crystal structures.

Results and discussion

Conventional molecular dynamics

The initial configuration for the CYP3A4-membrane system was generated by inserting the N-terminal transmembrane domain into a 75-Å² POPC bilayer such that the G'-helix was oriented just above the POPC headgroups, or 8 Å above the average position of the phosphate (PO₄) units. In the 50-ns simulation that followed, the A-anchor (residues 44–47), F'-helix (residues 217–226), and G'-helix (residues 227–238) associated with the membrane headgroups in 10 ns. A similar orientation of the catalytic domain and time for insertion were observed by Baylon et al. (22) using only the soluble catalytic domain of CYP3A4 and the highly mobile membrane-mimetic (HMMM) representation that replaces native lipids at the aqueous interface with those having truncated tails and the interior with 1,1-dichloroethane (DCE) to mimic the bulk hydrocarbon environment.

Following the initial 50 ns of conventional molecular dynamics (cMD) simulations, three TST molecules were randomly added to the water surrounding the CYP3A4 catalytic domain, and the system was equilibrated for an additional 100 ns in the constant temperature and pressure (NPT) ensemble. Two of the TST molecules were observed to partition into the POPC headgroups. Consistent with the observations of Baylon et al. (22), excluding the N-terminal helix, the A-anchor was observed to insert deeper into the membrane followed by the G' and F'-helices. The end point of this simulation, the structure of POPC illustrating the hydrophobic tails (red), AG (orange), PO₄ (blue), and choline (red). C, average distances above the membrane center of lipid components and protein domains in the leaflet bearing the CYP3A4 catalytic domain. The AG and PO₄ units are represented by orange and blue dashed lines. Positions of the A-anchor, F'-helix, and G'-helix are represented with green, purple, and yellow solid lines, respectively.

Figure 1. A, end point of the 100-ns cMD simulation following the addition of TST. The A-anchor, F'-helix, and G'-helix are illustrated in green, purple, and yellow, respectively. TST molecules are red, and the POPC membrane, which is partially cut away to reveal the immersed domains, is represented as a gray contour surface. The putative auxiliary TST-binding site is highlighted with a dashed circle. B, structure of POPC illustrating the hydrophobic tails (black), AG (orange), PO₄ (blue), and choline (red). C, average distances above the membrane center of lipid components and protein domains in the leaflet bearing the CYP3A4 catalytic domain. The AG and PO₄ units are represented by orange and blue dashed lines. Positions of the A-anchor, F'-helix, and G'-helix are represented with green, purple, and yellow solid lines, respectively.

the PO₄ units. The A-anchor (z = 12.9 Å) persists below the AG centers of mass and into the upper boundary of the hydrocarbon tails. Although the relative insertion depths of these domains are consistent with previous simulations, the absolute depths are greater than those reported previously (22). For example, in the simulations reported by Baylon et al. (22), the F'-helix was not observed to penetrate beyond the choline fragment of the headgroups. In the simulation described herein, insertion depths of these domains are converged prior to the addition of TST molecules to the system, so the observed differences are not attributable to the extended simulation time used in the present study. Conversely, the variations in insertions are likely attributable to differences in the properties of the HMMM and a full membrane. First, the DCE used to mimic the hydrophobic tails in the HMMM extends well into the AG layer of the headgroups (21). The result is a membrane-mimetic representation with greater density in this region than in a full...
membrane representation, thereby posing a density barrier preventing deeper penetration of the protein helices. In addition, the interspersion of DCE also decreases the polarity at this depth relative to the full membrane, imposing an energetic penalty on further penetration of the amphipathic helices. Nevertheless, experience with the HMMM is still limited, and caution should be exerted when interpreting simulations where there is limited experimental data describing the extent of the protein-membrane interaction.

**Accelerated molecular dynamics**

Following cMD in the presence of TST, 50 0.5-μs aMD simulations were seeded to extensively sample TST interactions with CYP3A4. Typical of all replicas, the TST distribution from one copy is illustrated in Fig. 2A. It is clear that TST molecules sample the surface of the CYP3A4 catalytic domain as well as partition into and persist in the POPC bilayer. In two replicas, TST entered the active site from the membrane face of the catalytic domain. Snapshots of the TST positions in the window surrounding these binding events are illustrated in Fig. 2B. In both instances, TST initially interacts with the G'-helix and is subsequently captured between the groove formed by the F'/G'-helices and the B-C loop. To illustrate the timescales of TST binding in each replica, the distances between each TST and the heme iron are plotted over the length of the simulations (Fig. 2, C and D). TST is ensconced in the active site by 150 and 50 ns in the first and second replicas, respectively, with an average iron-TST distance of 5–8 Å (Fig. 2, C and D, blue traces). The length of the simulation in the first replica was extended to 1 μs wherein the TST was observed to remain bound in the active site adjacent to the heme.

In the cMD simulations, one TST was found in a crevice bordered by the D-, J'-, K-, and L-helices adjacent to the proximal face of the heme (Fig. 2, C and D). TST persisted in this site for 645 ns in the first replica where concomitant active site binding was observed with the exception that from 225 to 275 ns the TST molecule dissociated from the site, diffused more than 40 Å away, and returned to its original position. Beyond 645 ns, TST moved into a second stable binding site 5 Å from its initial position and remained there until the end of the trajec-

---

**Figure 2.** A, distribution of TST molecules (green, blue, and orange) from one 0.5-μs aMD replica. Coordinate positions were recorded every 400 ps. B, snapshots of the two TST molecule positions (pink and orange lines) in the window surrounding their binding events. The POPC membrane is represented as a gray contour surface. C and D, distances between the TST centers of mass and the heme iron versus time for the two replicas where TST was observed to enter the CYP3A4 active site through the membrane. Solid lines represent a 30-ns moving average of the distances. E, volumetric map (red) illustrating the highest TST mass-weighted density over the 25 μs of aMD simulation. Representative TST positions from trajectory snapshots where it was observed to occupy these sites are illustrated in green. Both the TST positions and volumetric map are superimposed onto initial CYP3A4 geometry used in the aMD simulations. F and G, magnified views of the volumetric map highlighting the auxiliary TST-binding sites. H, representative cvSMD pathway for TST egress from CYP3A4.
CYP3A4 substrate–binding dynamics

Adaptive biasing force simulations of TST binding

Although aMD has pushed the simulation of events in biomolecules into the millisecond regime, reweighting the trajectories to recover free energies from the original ensemble has proven more difficult because the trajectories are dominated by large values of the boost potential, resulting in noisy energy statistics (27). To assign free energies to states along the pathway for TST ingress, the pathway observed in the aMD simulation was reconstructed using constant velocity–steered molecular dynamics (cvSMD) (28, 29). The cvSMD pathway is illustrated in Fig. 2H. cvSMD provides a finely resolved ensemble of structures along the ingress coordinate (ξ) from which a set of equally spaced representatives were selected for an adaptive biasing force (ABF) calculation (30) of the potential of mean force (PMF). Spanning 22 Å, ξ was partitioned into 2-Å bins that extended from TST burial in the CYP3A4 active site to immersion in the POPC headgroups. Each bin was populated with a trajectory snapshot from the cvSMD simulation corresponding to a value of ξ within that bin. Sampling in each bin was initially performed for 100 ns and in 10-ns increments thereafter. The complete PMF (ΔG(ξ)) converged after 150 ns with an overall free-energy change (ΔG1→∞) of −12.9 kcal mol−1. The 100-ns PMF, those from the subsequent 10-ns increments, and the final converged PMF are illustrated in Fig. 3A. Although the PMF is rugged as expected for ligands binding to proteins, it can be divided into three general regions: 1) TST immersed in the POPC headgroups including initial contact with the G′-helix, 2) a metastable intermediate, and 3) TST bound in the active site.

From total immersion in the POPC headgroups at ξ ≈ 18 Å, the initial minimum occurring at ξ = 18 Å corresponds to the first interaction of TST with the G′-helix. In a representative snapshot from ξ ≈ 18 Å (Fig. 3B), the lower face of TST is bound by the G′-helix, and the upper face is bound by the AG fragments of POPC. This initial interaction suggests a mechanism in which the G′-helix serves as the initial contact as well as a “ramp” to load TST directly from the membrane. From the initial interaction at ξ ≈ 18 Å, TST subsequently encounters small energy barriers (<1 kcal mol−1) corresponding to movement over the G′-helix before encountering the intermediate configuration at ξ = 10 Å and 4.9 kcal mol−1 (ΔG1→∞) below the unbound state. At this point in the PMF, TST is flanked by the F′- and G′-helices and the B-C loop (Fig. 3C). The origin of the minimum on the PMF is apparently derived from stabilization through close packing of “Phe cluster” residues (Phe106, Phe213, Phe220, Phe241, and Phe304) as well as Ile120, Ile223, and Val240 around TST. The intermediate TST-binding site is reminiscent of that for progesterone (PRG) in crystal structures with CYP3A4; however, there are important differences (31, 32). In Fig. S1, an overlay of the intermediate structure with the CYP3A4–PRG crystal structure shows that PRG sits closer to the F′-helix and is packed against Phe213, Phe219, and Phe220. PRG is precluded from entering the active site by the interdigitation of the Phe106, Phe213, Phe215, Phe241, and Phe304 side chains. By contrast, TST leverages the apertures open in the intermediate, disrupting the Phe cluster with concomitant movements of the B-C, F′-F′, and G-G′ loops, thereby resulting in deeper penetration into the entrance channel.

The PRG site has been speculated to be both an allosteric site and/or a result of PRG capture in the entrance channel by crys-
tallography. These and previous MD simulations have established that interactions with the membrane induce remarkable conformational changes in the F′- and G′-helices compared with those observed crystallographically (22, 33, 34). The interaction with the membrane unwinds the N terminus of the G-helix (residues 243–245) and extends the G-G′ loop, causing Phe241 to disengage from the hydrophobic cluster. In light of these conformational changes, it is not surprising that the positions of TST in the intermediate state and PRG in the crystal structure are not superimposable. In the CYP3A4–PRG structures, the G′-helix packs against the same helix of a second CYP3A4 monomer in the unit cell, and the pair are related by an S2 axis perpendicular to the G′-helical axis (Fig. S2). Sevrioukova and...
Poulos (32) recognized that the interface created by this interaction creates a hydrophobic patch to accommodate PRG. This interaction apparently stabilizes the closed configuration of the Phe cluster, precluding passage of this ligand into the active site. Assuming PRG enters the enzyme using the same mechanism as TST, it is unlikely that the PRG site represents a native transport intermediate but rather a non-native state resulting from distortion of the native potential-energy surface by the crystal lattice.

From the intermediate state, TST surmounts a 2.2 kcal·mol⁻¹ energetic barrier (ΔG_{TST_{UB}}) to enter the active site. Conversely, passage of TST back into the intermediate state requires overcoming a 10.6 kcal·mol⁻¹ barrier, and it is thereby kinetically trapped in the active site. The TST α-face is packed against the heme, whereas the side chains of Phe²⁰⁴, Phe²¹³, and Phe²¹⁵ contact the β-face (Fig. 3D). Orientation of the α-face of TST toward the heme is inconsistent with the dominant 6β-hydroxylated product (35). Nevertheless, this configuration would not afford adequate volume to accommodate O₂; hence, TST reorientation would also be required for productive catalysis. Because the TST C6 lies ∼3.7 Å from the heme iron (Fig. 4A), only a minor reorganization of the TST would be required to accommodate O₂.

The CYP3A4 active site conforms to TST with Phe²¹³ and Phe²¹⁵ in the F-F’ loop and Phe³⁰⁴ in the I-helix, restraining TST close to the heme. Leu⁴⁸², Ala³⁰⁵, Arg¹⁰⁵, and Ser¹¹⁹ surround the TST periphery, limiting translation (Fig. 3D). In the recent structure obtained with MDZ, the ligand contacted both Leu⁴⁸² and Ser¹¹⁹ where the latter residue apparently contributes a functionally important hydrogen bond to MDZ (36). The relative positions of TST, Ser¹¹⁹, Ile³⁰¹, Phe³⁰⁴, and Arg³⁷² are illustrated in Fig. 4A, and the distances between those that directly contact TST and the TST center of mass are plotted in Fig. 4B. In contrast to the CYP3A4 structure with MDZ, hydrogen bonds involving Ser¹¹⁹ were not observed; rather the backbone carbonyl of Ile³⁰¹ is hydrogen-bonded to the 17β-hydroxyl of TST. At the opposing end, the 3-keto oxygen is hydrogen-bonded to water molecules that bridge the heme propionate to the backbone carbonyl of Arg³⁷².

Numerous residues have been targeted for mutagenesis to identify regions of the protein involved in substrate specificity and kinetic cooperativity (37–40). Mutations have successfully
restored hyperbolic reaction kinetics as well as mirrored the stimulation induced by heterotropic effectors. One residue, Phe\(^{304}\), has demonstrated importance in substrate specificity and catalytic cooperativity. The \(V_{\text{max}}\) is increased for PRG hydroxylation, and the Hill coefficient is slightly reduced in the F304A mutant, consistent with a decrease in cooperativity (38).

Hyperbolic 6\(^{-}\)hydroxylation kinetics are restored in the F304W mutant, and stimulation by \(\alpha\)-naphthoflavone is diminished (41). This evidence supports that Phe\(^{304}\) sits near the interface of the active site as well as an effector-binding site. Indeed, Phe\(^{304}\) does directly interact with TST in both the intermediate and fully bound states (Fig. 3, C and D). In the intermediate state, the mean distance between TST and the Phe\(^{304}\) phenyl is 5.8 \(\pm\) 0.4 \(\AA\) and is therefore closely packed against the steroid nucleus. In the bound state, at 9.8 \(\pm\) 0.4 \(\AA\), Phe\(^{304}\) is more distant, although it is positioned to prevent TST from reverting back into the intermediate state (Fig. 4B). To the extent that the active and intermediate sites could be simultaneously occupied at high concentrations, the intermediate site could constitute the third TST-binding site and is consistent with neighboring active and allosteric sites.

**ABF simulations of TST in a POPC bilayer**

It is clear from examining the distribution in individual aMD trajectories that TST predominately persists among the POPC headgroups (Fig. 2A). Similar results were obtained in umbrella-sampling studies of cortisone in a POPC bilayer where the minimum was also within the AGs (42). Considering that the F\(^{-}\) and G\(^{-}\)-helices are likewise inserted at a depth dominated by the AG units and ABF calculations support a mechanism where TST is harvested from this region by the G\(^{-}\)-helix, it was hypothesized that the insertion of the F\(^{-}\)-G\(^{-}\) region was optimized by evolution to harvest amphipathic substrates from the membrane. This hypothesis was tested using a pair of unrestrained simulations of TST in a POPC bilayer and by calculation of the PMF for TST to traverse the bilayer. First, in separate, symmetrically hydrated (\(\pm\)10 \(\AA\)) 50-\(\AA^2\) POPC bilayers, TST was placed at the center of the membrane (among the lipid tails) and at 22 \(\AA\) above the membrane center (surrounded by water). Following equilibration of the system with TST fixed in these two positions, the TST restraints were released, and the simulations were conducted for 50 ns. The composite-normalized, mass-weighted densities for TST and the membrane headgroup components over the simulations are illustrated in Fig. 5A. Despite migrating to opposing leaflets, from these analyses it is evident that TST almost exclusively persists among the AG units of the POPC bilayer.

To determine whether the AG units represented the energy minimum along the one-dimensional reaction coordinate for TST crossing the membrane (\(\xi\)), a series of constrained simulations were performed where TST was moved incrementally along the membrane normal between the center and the sol-
mented position to establish initial configurations for an ABF calculation. The resulting PMF for TST crossing the POPC bilayer is illustrated in Fig. 4B. The minimum of the PMF lies at $\xi^T = 12.0$ Å above the membrane center, 6.9 kcal mol$^{-1}$ below TST in water ($\xi^T = 20 \pm 22$ Å). The lipid tails present a 4.5 kcal mol$^{-1}$ barrier for TST to cross the bilayer. The PMF confirms the result of the unconstrained simulation that dissolution among the AG groups corresponds to the minimum in the PMF. To compare the basin in the PMF with the positions of the F$^\prime$- and G$^\prime$-helices, distributions of insertion depths from the 1-µs aMD replica were superimposed in Fig. 5B. The mass density distribution of the G$^\prime$-helix distance could be fit to a Gaussian centered at 12.8 Å, which nearly matches the position of the PMF basin. Thus, the G$^\prime$-helix occupies a nearly optimal position for contacting membrane-embedded TST. Because the distributions were derived from long, enhanced-sampling MD simulations, it is not surprising that the distributions of these domains were broad (full width at half-maximum = 11.6 Å), demonstrating that the G$^\prime$-helix occasionally dips into the upper vicinity of the lipid tails as well as into the choline layer. This behavior would allow CYP3A4 to have great flexibility in the polarity of substrates it could recruit from the membrane. This characteristic would also permit the enzyme to function in membranes with somewhat different and more complex environments akin to those found in vivo compared with the relatively simple model POPC membrane used in these studies. Indeed, the distance separating the preferred dissolution layer for a particular substrate from the optimum position of the G$^\prime$-helix position in a given membrane environment may influence the kinetics of substrate oxidation by CYP3A4.

Conclusion

The results presented herein represent the first simulation of ligand binding to a cytochrome P450 from solution. The positive homotropic cooperativity in the oxidation of TST by CYP3A4 is well-documented and has been attributed to three TST-binding sites (13, 24, 43). The TST-CYP3A4 pair is an ideal system to capture the structural details of these aspects of cytochrome P450 activity by simulation. aMD frequently and reproducibly identified binding sites in close proximity to the heme that are consistent with the partial heme spin shift and NADPH oxidation without coupled product formation previously observed at low TST concentrations (24). Albeit with lower frequency, aMD also captured the entry of TST into the active site through an entirely membrane-embedded pathway. Free-energy analysis of this path proved it to be a two-step mechanism with an intermediate state where TST rests in the activated position to establish initial configurations for an ABF calculation. The resulting PMF for TST crossing the POPC bilayer is illustrated in Fig. 4B. The minimum of the PMF lies at $\xi^T = 12.0$ Å above the membrane center, 6.9 kcal mol$^{-1}$ below TST in water ($\xi^T = 20 \pm 22$ Å). The lipid tails present a 4.5 kcal mol$^{-1}$ barrier for TST to cross the bilayer. The PMF confirms the result of the unconstrained simulation that dissolution among the AG groups corresponds to the minimum in the PMF. To compare the basin in the PMF with the positions of the F$^\prime$- and G$^\prime$-helices, distributions of insertion depths from the 1-µs aMD replica were superimposed in Fig. 5B. The mass density distribution of the G$^\prime$-helix distance could be fit to a Gaussian centered at 12.8 Å, which nearly matches the position of the PMF basin. Thus, the G$^\prime$-helix occupies a nearly optimal position for contacting membrane-embedded TST. Because the distributions were derived from long, enhanced-sampling MD simulations, it is not surprising that the distributions of these domains were broad (full width at half-maximum = 11.6 Å), demonstrating that the G$^\prime$-helix occasionally dips into the upper vicinity of the lipid tails as well as into the choline layer. This behavior would allow CYP3A4 to have great flexibility in the polarity of substrates it could recruit from the membrane. This characteristic would also permit the enzyme to function in membranes with somewhat different and more complex environments akin to those found in vivo compared with the relatively simple model POPC membrane used in these studies. Indeed, the distance separating the preferred dissolution layer for a particular substrate from the optimum position of the G$^\prime$-helix position in a given membrane environment may influence the kinetics of substrate oxidation by CYP3A4.

Conclusion

The results presented herein represent the first simulation of ligand binding to a cytochrome P450 from solution. The positive homotropic cooperativity in the oxidation of TST by CYP3A4 is well-documented and has been attributed to three TST-binding sites (13, 24, 43). The TST-CYP3A4 pair is an ideal system to capture the structural details of these aspects of cytochrome P450 activity by simulation. aMD frequently and reproducibly identified binding sites in close proximity to the heme that are consistent with the partial heme spin shift and NADPH oxidation without coupled product formation previously observed at low TST concentrations (24). Albeit with lower frequency, aMD also captured the entry of TST into the active site through an entirely membrane-embedded pathway. Free-energy analysis of this path proved it to be a two-step mechanism with an intermediate state where TST rests in the active site through an entirely membrane-embedded pathway. Free-energy analysis of this path proved it to be a two-step mechanism with an intermediate state where TST rests in the active site through an entirely membrane-embedded pathway. Free-energy analysis of this path proved it to be a two-step mechanism with an intermediate state where TST rests in the active site through an entirely membrane-embedded pathway. Free-energy analysis of this path proved it to be a two-step mechanism with an intermediate state where TST rests in the active site through an entirely membrane-embedded pathway. Free-energy analysis of this path proved it to be a two-step mechanism with an intermediate state where TST rests in the active site through an entirely membrane-embedded pathway. Free-energy analysis of this path proved it to be a two-step mechanism with an intermediate state where TST rests in the active site through an entirely membrane-embedded pathway. Free-energy analysis of this path proved it to be a two-step mechanism with an intermediate state where TST rests in the active site through an entirely membrane-embedded pathway. Free-energy analysis of this path proved it to be a two-step mechanism with an intermediate state where TST rests in the active site through an entirely membrane-embedded pathway. Free-energy analysis of this path proved it to be a two-step mechanism with an intermediate state where TST rests in the active site through an entirely membrane-embedded pathway. Free-energy analysis of this path proved it to be a two-step mechanism with an intermediate state where TST rests in the active site through an entirely membrane-embedded pathway. Free-energy analysis of this path proved it to be a two-step mechanism with an intermediate state where TST rests in the active site through an entirely membrane-embedded pathway. Free-energy analysis of this path proved it to be a two-step mechanism with an intermediate state where TST rests in the active site through an entirely membrane-embedded pathway. Free-energy analysis of this path proved it to be a two-step mechanism with an intermediate state where TST rests in the active site through an entirely membrane-embedded pathway. Free-energy analysis of this path proved it to be a two-step mechanism with an intermediate state where TST rests in the active site through an entirely membrane-embedded pathway. Free-energy analysis of this path proved it to be a two-step mechanism with an intermediate state where TST rests in the active site through an entirely membrane-embedded pathway. Free-energy analysis of this path proved it to be a two-step mechanism with an intermediate state where TST rests in the active site through an entirely membrane-embedded pathway. Free-energy analysis of this path proved it to be a two-step mechanism with an intermediate state where TST rests in the active site through an entirely membrane-embedded pathway. Free-energy analysis of this path proved it to be a two-step mechanism with an intermediate state where TST rests in the active site through an entirely membrane-embedded pathway. Free-energy analysis of this path proved it to be a two-step mechanism with an intermediate state where TST rests in the active site through an entirely membrane-embedded pathway. Free-energy analysis of this path proved it to be a two-step mechanism with an intermediate state where TST rests in the active site through an entirely membrane-embedded pathway. Free-energy analysis of this path proved it to be a two-step mechanism with an intermediate state where TST rests in the active site through an entirely membrane-embedded pathway. Free-energy analysis of this path proved it to be a two-step mechanism with an intermediate state where TST rests in the active site through an entirely membrane-embedded pathway. Free-energy analysis of this path proved it to be a two-step mechanism with an intermediate state where TST rests in the active site through an entirely membrane-embedded pathway. Free-energy analysis of this path proved it to be a two-step mechanism with an intermediate state where TST rests in the active site through an entire
from potential energy basins through reduction of energetic barriers (27). In an aMD simulation, when the system potential energy, \( V(r) \), is greater than a user-defined value \( E \), the simulation proceeds as in a conventional simulation. Conversely, when \( V(r) \) falls below \( E \), a continuous, positive “boost” energy is added such that the simulation is performed on a modified potential energy surface, \( V^*(r) \).

\[
V^*(r) = \begin{cases} 
V(r), & V(r) \geq E \\
V(r) + \frac{(E - V(r))^2}{\alpha + (E - V(r))}, & V(r) < E
\end{cases}
\]  
(Eq. 1)

The value of the boost energy depends on a tuning parameter, \( \alpha \), that determines the extent that energy basins on the modified potential energy surface, \( V^*(r) \), are elevated, effectively smoothing the potential energy surface.

Fifty replicas based on the end point of the conventional simulation containing TST were created, and their initial velocities were independently randomized at 310 K. aMD simulations were performed in “dual-boost” mode meaning that boost potentials were added to both the dihedral and total system potential energies with \( E_{\text{dihed}} = V_{\text{dihed\_avg}} + 0.3V_{\text{dihed\_avg}} \) and \( E_{\text{total}} = V_{\text{total\_avg}} + 0.2N_{\text{atoms}} \). aMD simulations were performed for 0.5 ns, and each was populated with the end point geometry from each of the corresponding constrained simulations. Within each of the major bins, ABF data were accrued in bins 0.1 Å wide with upper and lower wall constants of 10 kcal mol\(^{-1}\) Å\(^{-2}\). The ABF simulations were performed for 50 ns. The equilibrated systems having TST frozen at 0 and 22 Å were used as the initial configurations for two 50-ns NPT simulations where TST was unrestrained. Atomic density profiles were calculated using the Density Profile Tool for VMD (58).

**Steered molecular dynamics**

Following identification of the binding pathway in aMD simulations, cvSMD simulations were initiated from the first aMD trajectory snapshot that revealed occupation of the CYP3A4 active site by TST. The binding pathway was approximated by a linear path defined by the TST center of mass and a second point on the cvSMD trajectory where TST was immersed in the lipid headgroups adjacent to the G' helix. The vector defined by these points served as the pulling vector. A moving harmonic restraint was attached to the TST center of mass and pulled with a velocity of \( 1.25 \times 10^{-5} \) Å fs\(^{-1} \) along the pulling vector. "\( \alpha \) atoms of residues 19, 39, 167, 194, 285, 341, and 417, all distant from the entrance channel, were harmonically restrained (600 kcal mol\(^{-1}\) Å\(^{-2}\)) to avoid translation of the CYP3A4 system. The force on the harmonic constraint was recorded every 400 fs. Steered molecular dynamics simulations were performed for a total of 4 ns and repeated 10 times to generate an ensemble for PMF calculations.

**Adaptive biasing force simulations**

The constant velocity pulling vector from cvSMD was adopted as the reaction coordinate for the ABF simulations (30). The reaction coordinate was divided into 2-Å bins, and trajectory snapshots with ligand positions as close as possible to the beginning of the bins were selected as initial positions. These frames were minimized for 2000 steps with the TST coordinates frozen. Those \( \alpha \) atoms restrained in the cvSMD simulations were maintained in the ABF calculations. The ABF data were accrued in bins 0.1 Å wide with upper and lower wall constants of 10 kcal mol\(^{-1}\) Å\(^{-2}\). Simulations were performed until the overall PMF was observed to converge, defined as a mean PMF difference of less than 0.3 kcal mol\(^{-1}\). Because the biasing force rapidly varies, no bias was applied during the first 1200 fs of ABF simulation to avoid perturbation by non-equilibrium effects.

**ABF simulations of TST in a POPC bilayer**

A 50-Å\(^2\) POPC membrane was constructed using the Membrane Builder plugin of VMD (23) and symmetrically hydrated with an additional 10-Å layer of TIP3P water in the \( z \) directions. The hydrated membrane was minimized for 5000 steps and equilibrated in the NVT ensemble for 0.5 ns and in the NPT ensemble for 5 ns. In separate copies of the equilibrated membrane, TST was placed at the center (0 Å; in the POPC tails) and in 2-Å increments to 22 Å above the center (in water). With TST’s position frozen, the systems were minimized for 10,000 steps and equilibrated for 4 ns in the NVT ensemble. Each constrained simulation served as an initial configuration for an ABF calculation of the PMF for TST entry into the membrane with the bilayer normal as reaction coordinate. The 22-Å reaction coordinate was partitioned into 2-Å bins, and each was populated with the end point geometry from each of the corresponding constrained simulations. Within each of the major bins, ABF data were accrued in bins 0.1 Å wide with upper and lower wall constants of 10 kcal mol\(^{-1}\) Å\(^{-2}\). The ABF simulations were performed for 50 ns. The equilibrated systems having TST frozen at 0 and 22 Å were used as the initial configurations for two 50-ns NPT simulations where TST was unrestrained. Atomic density profiles were calculated using the Density Profile Tool for VMD (58).

**Author contributions—**J. C. H. data curation; J. C. H. formal analysis; J. C. H. funding acquisition; J. C. H. investigation; J. C. H. visualization; J. C. H. writing—original draft; J. C. H. writing—review and editing.

**Acknowledgments—**This work was supported by a generous allocation of computational resources from the Ohio Supercomputer Center. I thank William M. Atkins (University of Washington), Kakali Sen (Daresbury Laboratory), and Lisa S. Shock (Virginia Commonwealth University) for insightful discussions and critically reading the manuscript.

**References**

CYP3A4 substrate–binding dynamics


