Theoretical Characterization of Substrate Access/Exit Channels in the Human Cytochrome P450 3A4 Enzyme: Involvement of Phenylalanine Residues in the Gating Mechanism

Dan Fishelovitch,[†] Sason Shaik,[‡] Haim J. Wolfson,[§] and Ruth Nussinov*[,†,⊥]

Department of Human Molecular Genetics and Biochemistry, Sackler Institute of Molecular Medicine, Sackler Faculty of Medicine, Tel Aviv University, Tel Aviv 69978, Israel, Institute of Chemistry and the Lise-Meitner-Minerva Center for Computational Quantum Chemistry, The Hebrew University of Jerusalem, 91904 Jerusalem, Israel, School of Computer Science, Raymond and Beverly Sackler Faculty of Exact Sciences, Tel Aviv University, Tel Aviv 69978, Israel, and SAIC-Frederick, Inc., Center for Cancer Research Nanobiology Program, NCI−Frederick, Building 469, Room 151, Frederick, Maryland 21702

Received: November 26, 2008; Revised Manuscript Received: August 5, 2009

The human cytochrome P450 3A4 mono-oxygenates ~50% of all drugs. Its substrates/products enter/leave the active site by access/exit channels. Here, we perform steered molecular dynamics simulations, pulling the products temazepam and testosterone-6βOH out of the P450 3A4 enzyme in order to identify the preferred substrate/product pathways and their gating mechanism. We locate six different egress pathways of products from the active site with different exit preferences for the two products and find that there is more than just one access/exit channel in CYP3A4. The so-called solvent channel manifests the largest opening for both tested products, thereby identifying this channel as a putative substrate channel. Most channels consist of one or two π-stacked phenylalanine residues that serve as gate keepers. The oxidized drug breaks the hydrophobic interactions of the gating residues and forms mainly hydrophobic contacts with the gate. We argue that product exit preferences in P450s are regulated by protein−substrate specificity.

1. Introduction

Cytochrome P450 3A4 (CYP3A4) is the major drug-metabolizing enzyme in humans, affecting more than 50% of all swallowed drugs.[1] CYP3A4 catalyzes mono-oxygenation reactions such as hydroxylation and epoxidation.[2] The so-formed products are more hydrophilic and hence are more easily excreted from the body. CYP3A4 oxidizes a variety of chemical compounds and drugs with different sizes and shapes.[3,4] The active site is located at the hydrophobic core of the protein and is connected to the surface of the enzyme through access channels (Figure 1).[5] The prosthetic heme group in the active site is covalently bound to a conserved cysteine and is responsible for substrate oxidation. The substrate oxidation process follows a well-characterized catalytic cycle.[6−8] The cycle is initiated when a substrate enters the active site and removes a water molecule coordinated to the ferric heme. The heme's iron shifts to its high-spin state, thereby facilitating the reduction of the ferric center by NADPH-cytochrome P450 reductase. Following molecular oxygen consumption and subsequent addition of two electrons from the reductase and two protons from the surrounding, the O−O bond is cleaved, and an active species is formed that oxygenates the substrate. Subsequently, the product leaves the active site through the access channels, and the enzyme resumes its resting state. Although the extent of sequence identity of the cytochrome P450 family of enzymes (P450s) is rather low, their structure is conserved[9] with one

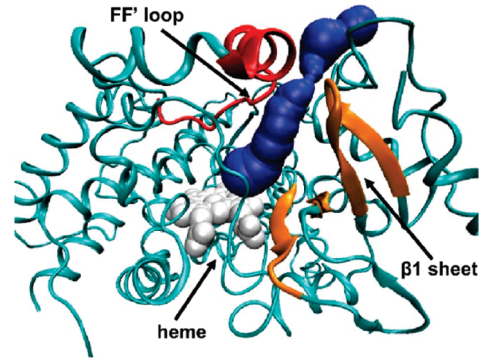

Figure 1. CYP3A4 structure including substrate channel 2a (shown in blue). The enzyme is represented in cyan cartoons, FF′ loop and F′ helix (connected to G′ helix in cyan) colored red, and β1 sheet colored orange. The heme prosthetic group is represented as white VDW spheres.

available fold for P450s in SCOP.[10] The most variable regions in P450s are the substrate recognition sites (SRSs)[11,12] that spatially form the active site and access channels. This variability could account for the different substrate specificity manifested by different P450 isoforms. Wade et al.[13] identified a structurally conserved channel in all bacterial P450s positioned between the FF′ loop, β1, and β4 sheets and termed it 2a (Figure 1); yet, interestingly, using molecular dynamics (MD) and repulsive expulsion molecular dynamics, they also showed that in three different bacterial P450s, the gating mechanism for ligand exit through channel 2a is different and depends on the specific isoform.[14] This difference can modulate protein−substrate specificity in P450s. In contrast to bacterial P450s, in mammalian P450s that are anchored to the endoplasmatic reticulum (ER) membrane, channel 2a is predicted to face the hydrophobic ER membrane, suggesting that highly hydrophobic substrates

* To whom correspondence should be addressed. Address: CCRNB, SAIC-Frederick, Inc., NCI-Frederick, Bldg. 469, Rm. 149, Frederick, MD 21702. E-mail: ruthn@ncifcrf.gov. Phone: 301-846-5579. Fax: 301-846-5598.
[†] Sackler Faculty of Medicine, Tel Aviv University.
[‡] The Hebrew University of Jerusalem.
[§] Raymond and Beverly Sackler Faculty of Exact Sciences, Tel Aviv University.
[⊥] NCI−Frederick.

10.1021/jp810386z CCC: $40.75
Published on Web 09/03/2009

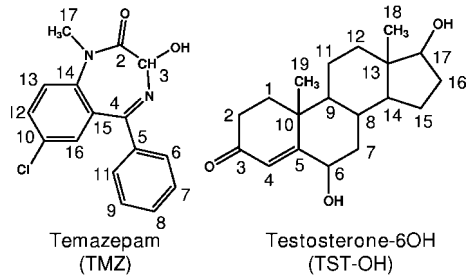

Figure 2. Chemical structures and carbon numbering systems for the CYP3A4 products temazepam (TMZ) and testosterone−6OH (TST-OH).

will reach the active site via this channel. Thus, substrate specificity is influenced not only by interactions at the P450 active site but also by the properties of the access routes to the binding site such as their size, shape, and physicochemical properties.

In most available CYP3A4 crystal structures, the enzyme exists in a closed state with little difference between the ligand-bound and -unbound enzyme conformations.[15,16] However, in the two latest determined X-ray structures of CYP3A4 complexed with two large inhibitors (ketoconazoles) and a substrate (erythromycin) inside of the active site,[17] the enzyme undergoes a large conformational change. The active site expands by 80% of its volume compared to that of the unbound structures, and there is a large shift in the F′G′ helices region (Figure 1) due to ligand binding. An extreme shift of this region was observed in the mammalian CYP2B4 crystal structure.[18] It is expected that the substrate/product entrance/exit will induce conformational changes that are not observed in most CYP3A4 crystal structures. Previous theoretical studies on bacterial P450s,[13,14,19] mammalian P450s,[20−24] and other related heme-containing enzymes[25,26] were carried out in order to monitor ligand egress from the enzyme's active site. MD simulations on mammalian P450s resulted in different dominant egress pathways that either expressed different substrate specificity of the P450 isoforms or stemmed from the different ligand probes tested in the simulations.

Here, using MD and steered MD (SMD), we simulated the last step of the catalytic cycle pulling out of the human CYP3A4 active site two distinct products, temazepam (TMZ) and testosterone-6βOH (TST-OH) (Figure 2). We searched for all physically feasible pathways for product egress, their dimensions, and the key residues involved in the gating mechanism and compared the results to previous simulations of other P450 isoforms. We then attempted to predict the preferred product exit routes in CYP3A4, seeking the factors that determine this preference. We expect that the substrate channel characterization, achieved in this work for CYP3A4, will provide further insight into the biophysics of transport to (from) the enzyme.

2. Computational Methods

Setting up the System. The CYP3A4 structure was taken from the PDB[27] (code 1TQN[15]). Missing residues of 1TQN were modeled with InsightII (http://www.accelrys.com/products/). pK_a values were calculated using the MCCE[28] software with DelPhi[29] as a Poisson−Boltzmann solver with a dielectric constant of 4. Hydrogen atoms were added with CHARMM.[30] The structure was solvated in a cubic box of ∼21000 TIP3P[31] water molecules, in addition to crystallographic waters in order to apply periodic boundary conditions. Chloride ions were added to neutralize the system. The structures of TMZ and TST-OH were derived from their substrate structures, diazepam and testosterone, taken from the Cambridge structural database (http://www.ccdc.cam.ac.uk/products/csd/, codes DIZPAM10 and TESTOM01, respectively). We substituted the oxidized C−H groups in diazepam and testosterone with C−OH groups. We took into account the flexibility of the TMZ phenyl ring by generating several conformations using the systematic search algorithm of the Cerius2 software. Each TMZ conformer and TST-OH were geometrically docked to the CYP3A4 active site using the PatchDock[32] software, taking into account shape complementarity. The TMZ flexibility was addressed by including different orientations of its rotatable phenyl ring in the docking procedure. Docking solutions with a distance less than 3 Å between the product hydroxyl and the heme's iron were considered for further steps. The remaining solutions were minimized by the CHARMM software, and the lowest CHARMM energy docking solution was used as a starting point for further MD simulations.

Force Field Parameters. Heme and protein parameters were taken from CHARMM22[33] and CHARMM27[34] force fields, respectively. Fe and S related missing parameters were added from Autenrieth et al.[35] and Bathelt et al.[36] The heme's atoms' partial charges were calculated by a QM(B3LYP)/MM optimization at the DFT level of theory, with a sextet multiplicity as the ground state.[37] The basis sets used were an effective core potential (ECP) and the LACVP for the iron and 6-31G for the rest of the heme and C442 atoms. The heme Mulliken charges were added to the force field. The QM/MM calculation used the ChemShell package.[38] The force field parameters for TMZ were taken from a previous diazepam force field[39] with the addition of the hydroxyl moiety. TST-6OH and -16OH parameters were derived from a previously published CHARMM force field for cholesterol.[40] TMZ and testosterone products' atomic partial charges were modified to account for the addition of the hydroxyl group by a QM/MM optimization as done with the heme.

MD Procedure. For each run, the system was initially minimized with CHARMM, with a decreasing harmonic force constraint on the enzyme. Further procedures were performed using the NAMD code.[41] The minimized system was gradually heated up to 310 K and equilibrated for 50 ps. The equilibrated system was then run for 500 ps of MD simulation followed by a 1.6 ns of steered MD (SMD[42]) pulling each product out of the active site. We applied a canonical NVT ensemble dynamics at 310 K with a 1 fs time step. The Shake algorithm was applied to fix all bond lengths involving hydrogen atoms. A final nonbonded interaction cutoff radius was 12 Å, with a switching function starting from 10 Å. Nonbonded electrostatic interactions were calculated with Particle Mesh Ewald summation.[43] Snapshots were saved each 500 steps. The SMD procedure used a pulling spring force constant of 7 kcal/mol/Å^2 with a constant velocity of 0.000025 Å/fs. The chosen force constant and pulling velocity were three times larger and twice slower than those applied in previous SMD simulations of the human CYP3A4.[22] The pulling force was applied after the equilibration phase, with the spring connected between the center of mass of the product and a "dummy" atom in the general direction of the channels exit. Three channels (2b, 2e, and the solvent channel, termed S; see Figure 3) are observed to be partially open in the PDB file 1TQN, which we used as input.[15] Therefore, the directions of the SMD pulling forces in the case of the above channels were determined according to the visible channel exit. For those channels that evaded detection, 18 different vectors were tested that are in the general direction of the roof and sides of the active site, and three additional channels were detected (see

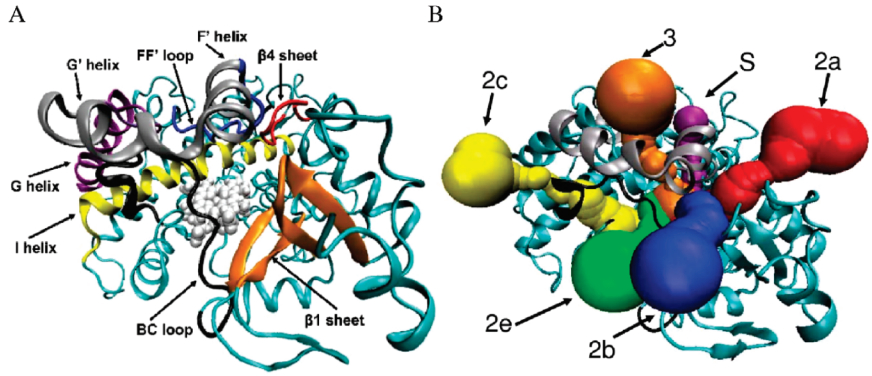

Figure 3. CYP3A4 secondary structures related to the identified substrate channels. (A) Secondary structure elements in CYP3A4 that define the different channels. FF′ loop, F′ and G′ helices, G helix, I helix, BC loop, and β1 and β4 sheets are colored blue, silver, purple, yellow, black, orange, and red, respectively. The rest of the enzyme is colored cyan and represented as cartoons. The heme prosthetic group is represented by white VDW spheres. (B) Substrate channels detected during SMD. Channel surfaces were calculated by MolAxis. Channels 2a, 2b, 2c, 2e, 3, and S are colored red, blue, yellow, green, orange, and purple, respectively. F′G′ helices are colored silver, and the BC loop is in black. The rest of the enzyme is colored cyan and represented as cartoons.

Results). All of the SMD pulling force directions and the SMD starting point structures are available in the Supporting Information (SI).

In addition to the SMD simulations in the presence of a product, we performed a substrate-free 1.6 ns of classical MD simulation on CYP3A4.

3. Results

Global CYP3A4 Stability. We performed different SMD runs to find the minimum constant velocity and force constant values that ensure low protein fluctuations but permit the escape of the product from the active site. We ended up with 12 different SMD simulations (6 distinct channels for each product) and another SMD run of TST-160H escaping through channel 2c (see Discussion). A velocity of 0.000025 Å/fs was found as a minimum cutoff. As judged from the total energy, pressure, rmsd, and RMSF profiles (Figures S1−S4, Supporting Information), the entire system and protein structure appeared stable throughout the simulations, even when further applying an external pulling force. The average Cα rmsd's along all of the CYP3A4-unbound, TMZ-bound, and TST-OH-bound simulations were 1.2, 1.2−1.3, and 1.8 Å, respectively, with not more than ±0.1 Å fluctuations (Figure S3, Supporting Information). Major local fluctuations observed in the unbound dynamics were in residues 255−270 and 280−290, which comprise a modeled loop, missing in several PDB files of CYP3A4.[16] In the bound SMD simulations, fluctuations occur mainly in the TST-OH simulations and are larger than TMZs. Major RMSF peaks detected upon addition of the products are at residues 100−125 and 200−250 (Figure S4a−S4d, Supporting Information), which belong to the BC loop and FG helices' secondary structure elements, respectively (Figure 3A). These secondary structures are known to be the most flexible in P450s.[44]

Substrate Channels' Structure and Dimensions. The substrate exit pathways are named according to Cojocaru et al.'s[5] nomenclature and include substrate channels 2a, 2b, 2c, 2e, and 3 and the solvent channel (S) (Figure 3). The channels are defined according to the secondary structures that form their entrance/exit (Figure 3A). Channel 2a is located between the FF′ loop and the β1 and β4 sheets. Channel 2b is positioned between the F′ and G′ helices and the β1 sheet and the BC loop. Channel 2c is between the GG′ helices and the BC loop. Channel 2e is surrounded by the BC loop. Channel 3 passes through the F′G′ helices, and channel S is above the I helix and between the FF′ loop and the β4 sheet.

TABLE 1: Maximum Bottleneck Values for Each Substrate Channel Observed for the Egress of TMZ and TST-OH in Angstroms

channel	TMZ	TST-OH
2a	3.70	4.14
2b	3.65	3.96
2c	3.09	3.91
2e	3.87	4.20
3	3.31	4.12
S	3.98	4.47

The channel lengths were measured, starting 3 Å above the heme Fe up to the bottleneck residues that form the gates. The channels lengths are 19, 18, 17, 12, 17, and 13 Å for 2a, 2b, 2c, 2e, 3, and S channels, respectively. Thus, channels 2e and S are the shortest paths from the active site to the enzyme surface. The channels' bottleneck opening radii were measured during the simulations using the MolAxis tool that identifies and calculates channel surfaces within macromolecules by computational geometry techniques.[45] Figure S5 (Supporting Information) depicts the bottleneck radius of all of the channels along the simulations, and Table 1 summarizes the channels' maximum opening radii. We noticed that for both products, the S channel manifested the largest opening, and substrate channel 2c had the smallest bottleneck radius.

Channel−Product Interactions. Table 2 summarizes the major residues which interact with the products during the simulations (Tables S1−S10, Supporting Information, give secondary structure details and electrostatic interactions). These residues are derived from contact analysis, taking into account interactions of less than 5 Å between the product and the enzyme. The majority of the nonbonded interactions between the products and the channel-forming residues are hydrophobic aromatic and hydrophobic aliphatic. In contrast to observations on the ligand egress in the bacterial $P450_{cam}$,[46,47] for the CYP3A4, we observe no salt bridge perturbations during product exit through any of the channels. The two salt bridges located at channels 2a and 2b between D76-R106 and D76-R372 are very stable during the exit of both products. Polar and some hydrophobic residues serve as hydrogen-bond donors and acceptors that H bond usually the hydroxyl and carbonyl groups of the products (Table 2, superscript). The active site key residues that hold the product by hydrogen bonds are S119 and R212. S119 was experimentally shown to alter testosterone hydroxylation selectivity.[48] Mutations of S119 to hydrophobic residues switched the 6β to 2β hydroxylation. Thus, the loss of

TABLE 2: Major Channel Forming Residues of Substrate Channels 2a, 2b, 2c, 2e, 3, and S in the CYP3A4 Enzyme Obtained by Close Contact Analysis[a] with the TMZ and TST-OH Products

channel	channel residues							
2a	L47[b]	Y53[b]	F57	D76	R105	R106	F108	S119[c]
	R212[c]	F213[b,c]	F215[b]	L216	L221	T224	R372	L482[b]
2b	F57	D76[b]	Q79[b,c]	R105	R106	F108[b]	S119[c]	R212[b]
	F213	F215	F220	T224[b]	P227[b]	A370[b]	R372	L482
2c	R105	F108	M114[b,c]	I118	S119[b,c]	I120	R212[c]	F213[c]
	F215	V240	F241[c]	T246[b]	I300[c]	F304[c]		
2e	R105[b,c]	F108	G111	K115	I118	S119[b,c]	I120[b]	E122[b]
	R212[c]	F213	F215	V240	A370			
3	F108[b]	S119[b,c]	I120	K209[c]	R212[c]	F213[b]	F215	F219
	F220	V240	F241	P242	R243	V245[c]	F304[b]	
S	S119[c]	K173[c]	L211	R212[c]	F213[c]	E308[b]	T309	S311[b]
	S312[b,c]	I369[b]	A370	M371	L482[b]	L483[c]	Q484[c]	

[a] Close contacts were defined by product atoms having a distance of less than 5 Å from enzyme residues. [b] The residue serves as a H-bond acceptor. [c] The residue serves as a H-bond donor.

TABLE 3: Total SMD Work (Calculated According to Equation 1) Done along the TMZ and TST-OH Trajectories while Exiting through Substrate Channels 2a, 2b, 2c, 2e, 3 and S (Units Are in kcal/mol)

channel	TMZ	TST-OH
2a	265.0 ± 5.5[a]	283.2
2b	235.1 ± 7.5[a]	268.7
2c	292.1	445.4 ± 14[a]
2e	203.6	244.6 ± 6.9[a]
3	274.8 ± 8.9[a]	235.1
S	337.1	276.2 ± 5.8[a]

[a] Average of three runs.

TABLE 4: Gating Residues of Substrate Channels 2a, 2b, 2c, 2e, 3, and S Identified in the SMD Simulations of TMZ and TST-OH

channel	TMZ	TST-OH
2a	F57−F215	F57−F215
2b	F108−F220	F108−F220
2c	F108−F241	F108−F241
2e	F108−I120	R105−S119
3	F213−F241	F213−V240
S	R212−L482	R212−L482

the S119 directed hydrogen bond to the substrate by site-specific mutations changes the regioselectivity of testosterone oxidation. Product release from the CYP3A4 active site is initiated by hydrogen-bond breakage between the product and R212 and S119 (Tables S7−S10, Supporting Information) and detachment from hydrophobic residues, I301, F304, A305, F213, A370, and L482. Overcoming these barriers leads to product movement toward the gate keeper residues. The products pass the gates of channels 2a, 2b, 2c, 2e, 3, and S after 900, 800, 1150, 700, 900, and 750 ps, respectively.

Substrate Channels' Gating Mechanism. The invested total SMD work (W_{SMD}) during each product egress was measured every 1 ps according to eq 1

$$W_{SMD} = \sum_{t=1}^{n} \vec{F}(t) \times [\vec{D}(t) - \vec{D}(t-1)] \times \cos(\theta) \quad (1)$$

where W_{SMD} is the mechanical work done along a SMD trajectory, n is the number of MD steps of a trajectory, $\vec{F}(t)$ is the SMD force invested at time t, $\vec{D}(t)$ and $\vec{D}(t-1)$ are the coordinates of the center of mass of the product at times t and $t-1$, respectively, and θ is the angle between the force and the product displacement vectors. The calculated total W_{SMD} values are summarized in Table 3, and the entire SMD work profiles for all simulations are presented in Figure S6 (Supporting Information). Most of these profiles have a similar behavior. At the beginning of the simulations, W_{SMD} increases, reaching a global maximum that represents the potential barrier of product release via the channel gate. Following the product exit, there is usually a fast decline in the work to almost zero. For TMZ (TST-OH), the largest total W_{SMD} observed was while exiting from channel S (2c), and the lowest was from channel 2e (2e and 3). In order to test the robustness of our results, we performed three different SMD runs for each CYP3A4-detected channel with either TMZ or TST-OH. The TMZ (TST-OH) exit from the CYP3A4 active site through channels 2a, 2b, and 3 (2c, 2e, and S) was repeated three times. The results presented similar exit pathways, Cα rmsd and SMD work profiles (Figures S7−S9, Supporting Information). In addition, the calculated total SMD work was similar for each triplicate, with a relative standard deviation of not more than 3% (Table 3).

In order to define the gating mechanism of product egress from the substrate channels, we focused on MD snapshots 250 ps prior to and following the time frame where the maximum W_{SMD} was observed. From visual inspection, rmsd, and distance profiles, we identified the gate keepers of each channel for both products. Figure S10 (Supporting Information) presents the distance along MD between the two major gating residues of each channel found for both products. Table 4 summarizes the gating residues of each channel for both products.

We observe that for channels 2a, 2b, 2c, and 3, the gates consist of hydrophobic residues, mainly phenylalanine residues. These residues parallel- or T-shape π-stack with each other, and the product disrupts this interaction when exiting the enzyme. This disruption is shown in the sequence of frames in Figures 4 and 5. In channel 2e, the release of TST-OH resulted in a breakage of a hydrogen bond between the carbonyl group of S119 and one of the guanidinium moiety hydrogen atoms of R105. However, for TMZ, the exit through channel 2e is coupled to breakage of aliphatic hydrophobic interactions between F108 and I120. This observation is experimentally supported by F108W and I120W mutations[49] that showed higher 1OH over 4OH midazolam (structurally related to TMZ) hydroxylation compared to the wild-type enzyme. Mutations to tryptophan are believed to close channel 2e and impose a switching of the midazolam entrance to another channel that may lead to a different active site orientation and product formation. In channel S, we observe interruption of aliphatic interactions between R212 and L482, while the product escapes the enzyme, as portrayed in Figure 6.

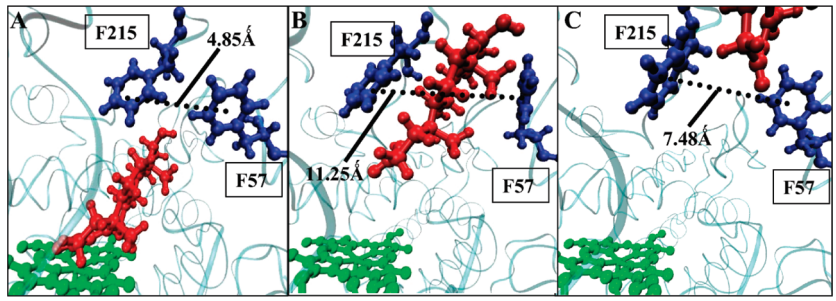

Figure 4. TST-OH release from substrate channel 2a. TST-OH is colored red, heme is in green, and gating residues are in blue. The enzyme is represented by cyan cartoons. The figures have the same orientation. The time frames are after (A) 0, (B) 450, and (C) 715 ps from the start of the simulation.

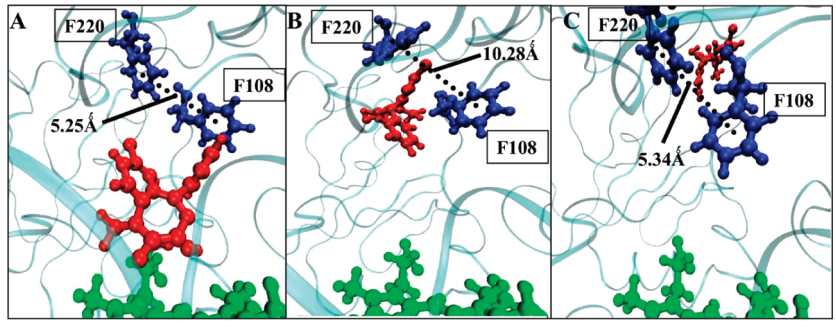

Figure 5. TMZ release from substrate channel 2b. TMZ is colored red, heme is in green, and gating residues are in blue. The enzyme is represented by cyan cartoons. The figures have the same orientation. The time frames are after (A) 10, (B) 400, and (C) 900 ps from the start of the simulation.

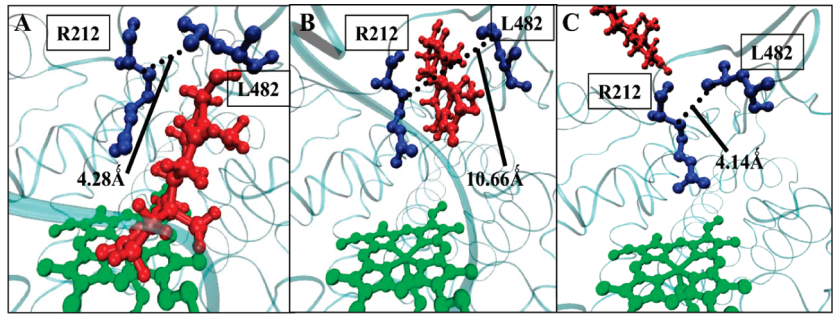

Figure 6. TST-OH release from substrate channel S. TST-OH is colored red, heme is in green, and gating residues are in blue. The enzyme is represented by cyan cartoons. The figures have the same orientation. The time frames are after (A) 0, (B) 460, and (C) 960 ps from the start of the simulation.

We therefore observe three types of channel gating mechanisms, (i) breaking F−F π-stacking; (ii) breaking aliphatic contacts, as seen at the S channel between R212 and L482; and (iii) breaking the H bond, as observed in the exit of TST-OH through channel 2e between R105 and S119. The predominant mechanism is the π-stacking.

4. Discussion

Dominant Channel in CYP3A4. Of all known P450s, the isoforms CYP2C9, CYP2D6, and CYP3A4 account for the vast majority of drug metabolism.[3] In the CYP2C9 structure bound to flurbiprofen,[50] the solvent channel was observed to be relatively open. It was suggested that for the human CYP2D6, channel S, defined earlier by Haines et al.[51] for the bacterial CYP_{BM3}, is the main substrate access channel.[52] Here, TST-OH leaving the active site through channel S resulted in the largest opening compared to the other channels with a similar total W_{SMD} to most detected channels (Table 3). Hence, the solvent channel in CYP3A4 is likely to serve as a substrate/product exit channel as in the human P450 isoforms.

CYP3A4 is tethered to the ER membrane by its N-terminus and contacts the membrane through other parts of the enzyme, including the F′G′ loop in the vicinity of substrate channels 2a and 2b.[53] Therefore, channels 2a and 2b are predicted to lie close to or dip into the ER membrane. TST-OH contains three hydrophilic positions, and it is less likely that it will exit through those channels.

It was shown by mutation and kinetic studies that cholesterol, the precursor of testosterone (TST), enters the microsomal CYP7A1 through the membrane and interacts with the F′G′ loop residues.[54] Williams et al.[53] suggested that lipophilic substrates and products could enter and exit through the lipid bilayer. As the exact orientation of channels 2a and 2b with respect to the membrane is unknown, the exit of TST-OH through channels 2a and 2b cannot be ruled out.

In terms of the invested SMD work, TST-OH was found to exit through channel 2c with the largest difficulty compared to any other channel and with a thermodynamically favorable exiting through channels 2e and 3 (Table 3). We therefore conclude that channel 2c is less likely to serve as a gateway for testosterone major product in CYP3A4, while channels 2e and

3 are predicted to form a gateway for TST-OH. In addition, channel S was found to be highly flexible. It faces the cytosol and contains a charged gating residue which could lead the product out of the CYP3A4 active site.

In contrast to TST-OH, although the egress of TMZ resulted in a large opening of the S channel, it needed the largest W_{SMD} compared to the rest of the channels (Table 3). This is accounted for by a high potential barrier derived from electrostatic interaction between the positive R212 and the oxygen of the hydroxyl group of TMZ (Figure S11, Supporting Information). We observe the lowest work profiles when TMZ exits from the CYP3A4 active site via channel 2e, suggesting that this channel is preferred for the exit of the hydroxylated product of diazepam.

Thus, there is no one specific access/exit route for substrates/products in CYP3A4.

Dominant Channel in P450s. Schleinkofer et al.[21] carried out MD simulations on CYP2C5, which is closely related to CYP2C9, pulling out of the active site progesterone, 21-hydroxyprogesterone (PROG-21OH), and benzenesulphonamide. The 2c channel was the predominant exit route. Li et al.[20] tested by MD simulations the egress of TST from the CYP2B1 active site. They found that the 2c channel is one of the preferred exit pathways for TST, in contrast to our results of CYP3A4. The differences in these observations may arise from different substrate specificities between P450 subfamilies.

We hypothesize that the product outlet pathway is determined in part by the post-oxidation product orientation inside of the active site. To test this hypothesis, we pulled TST-16OH, positioned in a 16 hydroxylation orientation, out of CYP3A4 from channel 2c. The TST-16OH exit resulted in a much faster and easier exit than that of TST-6OH, with a maximum pulling force of about half of that observed for TST-6OH (Figure S12, Supporting Information). In addition, the mechanical work measured along the TST-16OH escape from channel 2c was 189.4 kcal/mol. This easy escape from channel 2c is in accord with the previous simulations of the exit of TST-16αOH[20] and PROG-21OH[21] from CYP2B1 and CYP2C5, respectively. Both TST-16OH, which is hydroxylated mainly at the TST 16 position[55] in CYP2B1, and PROG-21OH are positioned inside of the active site in the general direction of channels 2c and 2e exits. However, the major hydroxylation product of TST by CYP3A4 is the TST-6βOH, with minor hydroxylation at position 16 (Figure 2),[56,57] and docking it accordingly resulted in a different orientation than that of TST-16OH and in a favorable orientation for exiting via channel 3 (Figure S13, Supporting Information).

We therefore conclude that the oxidized position in large products that have less rotational and translational freedom inside of the active site can determine the outlet pathway. Furthermore, there is no single dominant exit pathway for all P450s, and different products of the same substrate (e.g., TST) leave the active site through channels based on the substrate oxidation specificity.

Conservation of the Gating Mechanism. We find that the gating residues are bulky and hydrophobic in most CYP3A4 substrate channels (Table 4). Schleinkofer et al.[21] pointed to V106 and K241 in CYP2C5 as the gate keepers of channel 2c. In contrast, in the channel 2c of CYP3A4, the gating residues are F108 and F241, despite the fact that the channel faces the hydrated cytosol. Li et al.[20] also observed that channel 2c in CYP2B1 is gated by two phenylalanine residues. They also pointed to a phenylalanine residue as the gate keeper of channel 2e in CYP2B1. Winn et al.[14] showed that CYP_{cam} channel 2a is gated by F87 and F193, similar to our predicted gating mechanism. However, they observed different gating mechanisms for other bacterial P450s that include a positive arginine.

From all of the above and considering that the channels are part of the variable SRSs regions, the gating mechanisms are different for corresponding channels in different P450 isoforms. However, the phenylalanine gating mechanism exists in many P450s, which implies a nonspecific substrate channel.

5. Conclusions

We performed SMD simulations on CYP3A4 pulling out of its active site two products resulting from the hydroxylation of diazepam and testosterone. We identified several putative access channels connecting the active site to bulk solvent. We observed no single dominant channel in CYP3A4. For both temazepam and testosteron-6βOH, we observed a thermodynamic preference for exiting through channel 2e. In addition, testosteron-6βOH showed a low mechanical work profile while exiting through channel 3, which faces the cytosol as does channel 2e. The solvent channel was found to be the most flexible in CYP3A4 and is predicted to be a substrate/product channel as proposed earlier for other P450s.[52] Moreover, we found that the product shape and orientation after being oxidized determined the preferred channel from which it would exit. In the majority of the channels, the gate keepers are π-stacked phenylalanine residues; to exit, the product has to break this interaction. This gating mechanism is not conserved in P450s but exists in many isoforms and in different channels.

The versatility of the egress mechanism in CYP3A4 makes perfect sense, considering the fact that this isoform is promiscuous and oxidizes a great variety of substrates. This variety requires, in turn, a good number of egress pathways. We believe that these data will assist in identifying some structural determinants related to substrate specificity and in understanding drug metabolism and toxicity in CYP3A4.

Acknowledgment. This project has been funded in whole or in part with federal funds from the National Cancer Institute, National Institutes of Health, under Contract No. N01-CO-12400. The content of this publication does not necessarily reflect the views or policies of the Dept. of Health and Human Services, nor does mention of trade names, commercial products, or organizations imply endorsement by the U.S. government. This research was supported (in part) by the Intramural Research Program of the National Institutes of Health, National Cancer Institute, Center for Cancer Research. The research of S.S. at the H.U. was supported by an ISF grant 53/09. This study utilized the high-performance computational capabilities of the Biowulf PC/Linux cluster at the National Institutes of Health, Bethesda, MD (http://biowulf.nih.gov).

Supporting Information Available: TMZ, TST-OH, and heme ferric state CHARMM topology files, optimized coordinates, and additional figures and tables. This material is available free of charge via the Internet at http://pubs.acs.org.

References and Notes

(1) Wrighton, S. A.; Schuetz, E. G.; Thummel, K. E.; Shen, D. D.; Korzekwa, K. R.; Watkins, P. B. The human CYP3A subfamily: practical considerations. *Drug Metab. Rev.* **2000**, *32* (3−4), 339–61.

(2) Sono, M.; Roach, M. P.; Coulter, E. D.; Dawson, J. H. Heme-Containing Oxygenases. *Chem. Rev.* **1996**, *96* (7), 2841–2888.

(3) Guengerich, F. P. In *Cytochrome P450: Structure, Mechanism and Biochemistry*, 3 ed.; Ortiz de Montellano, P. R., Ed.; Plenum Press: New York, 2005; pp 377−531.

(4) Wrighton, S. A.; Stevens, J. C. The human hepatic cytochromes P450 involved in drug metabolism. *Crit. Rev. Toxicol.* **1992**, *22* (1), 1–21.

(5) Cojocaru, V.; Winn, P. J.; Wade, R. C. The ins and outs of cytochrome P450s. *Biochim. Biophys. Acta* **2007**, *1770* (3), 390–401.

(6) Davydov, R.; Makris, T. M.; Kofman, V.; Werst, D. E.; Sligar, S. G.; Hoffman, B. M. Hydroxylation of camphor by reduced oxy-cytochrome P450cam: mechanistic implications of EPR and ENDOR studies of catalytic intermediates in native and mutant enzymes. *J. Am. Chem. Soc.* **2001**, *123* (7), 1403–15.

(7) Shaik, S.; Kumar, D.; de Visser, S. P.; Altun, A.; Thiel, W. Theoretical perspective on the structure and mechanism of cytochrome P450 enzymes. *Chem. Rev.* **2005**, *105* (6), 2279–328.

(8) Davydov, R.; Perera, R.; Jin, S.; Yang, T. C.; Bryson, T. A.; Sono, M.; Dawson, J. H.; Hoffman, B. M. Substrate modulation of the properties and reactivity of the oxy-ferrous and hydroperoxo-ferric intermediates of cytochrome P450cam as shown by cryoreduction-EPR/ENDOR spectroscopy. *J. Am. Chem. Soc.* **2005**, *127* (5), 1403–13.

(9) Mestres, J. Structure conservation in cytochromes P450. *Proteins* **2005**, *58* (3), 596–609.

(10) Murzin, A. G.; Brenner, S. E.; Hubbard, T.; Chothia, C. SCOP: a structural classification of proteins database for the investigation of sequences and structures. *J. Mol. Biol.* **1995**, *247* (4), 536–40.

(11) Gotoh, O. Substrate recognition sites in cytochrome P450 family 2 (CYP2) proteins inferred from comparative analyses of amino acid and coding nucleotide sequences. *J. Biol. Chem.* **1992**, *267* (1), 83–90.

(12) McArthur, A. G.; Hegelund, T.; Cox, R. L.; Stegeman, J. J.; Liljenberg, M.; Olsson, U.; Sundberg, P.; Celander, M. C. Phylogenetic analysis of the cytochrome P450 3 (CYP3) gene family. *J. Mol. Evol.* **2003**, *57* (2), 200–11.

(13) Wade, R. C.; Winn, P. J.; Schlichting, I. Sudarko, A survey of active site access channels in cytochromes P450. *J. Inorg. Biochem.* **2004**, *98* (7), 1175–82.

(14) Winn, P. J.; Ludemann, S. K.; Gauges, R.; Lounnas, V.; Wade, R. C. Comparison of the dynamics of substrate access channels in three cytochrome P450s reveals different opening mechanisms and a novel functional role for a buried arginine. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99* (8), 5361–6.

(15) Yano, J. K.; Wester, M. R.; Schoch, G. A.; Griffin, K. J.; Stout, C. D.; Johnson, E. F. The structure of human microsomal cytochrome P450 3A4 determined by X-ray crystallography to 2.05-A resolution. *J. Biol. Chem.* **2004**, *279* (37), 38091–4.

(16) Williams, P. A.; Cosme, J.; Vinkovic, D. M.; Ward, A.; Angove, H. C.; Day, P. J.; Vonrhein, C.; Tickle, I. J.; Jhoti, H. Crystal structures of human cytochrome P450 3A4 bound to metyrapone and progesterone. *Science* **2004**, *305* (5684), 683–6.

(17) Ekroos, M.; Sjogren, T. Structural basis for ligand promiscuity in cytochrome P450 3A4. *Proc. Natl.Acad. Sci. U.S.A.* **2006**, *103* (37), 13682–7.

(18) Scott, E. E.; He, Y. A.; Wester, M. R.; White, M. A.; Chin, C. C.; Halpert, J. R.; Johnson, E. F.; Stout, C. D. An open conformation of mammalian cytochrome P450 2B4 at 1.6-A resolution. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100* (23), 13196–201.

(19) Borrelli, W. K.; Vitalis, A.; Alcantara, R.; Guallar, V. PELE: Protein Energy Landscape Exploration. A Novel Monte Carlo Based Technique. *J. Chem. Theory Comput.* **2005**, *1* (6), 1304–1311.

(20) Li, W.; Liu, H.; Scott, E. E.; Grater, F.; Halpert, J. R.; Luo, X.; Shen, J.; Jiang, H. Possible pathway(s) of testosterone egress from the active site of cytochrome P450 2B1: a steered molecular dynamics simulation. *Drug Metab. Dispos.* **2005**, *33* (7), 910–9.

(21) Schleinkofer, K.; Sudarko; Winn, P. J.; Ludemann, S. K.; Wade, R. C. Do mammalian cytochrome P450s show multiple ligand access pathways and ligand channelling. *EMBO Rep.* **2005**, *6* (6), 584–9.

(22) Li, W.; Liu, H.; Luo, X.; Zhu, W.; Tang, Y.; Halpert, J. R.; Jiang, H. Possible pathway(s) of metyrapone egress from the active site of cytochrome P450 3A4: a molecular dynamics simulation. *Drug Metab. Dispos.* **2007**, *35* (4), 689–96.

(23) Liu, X.; Wang, X.; Jiang, H. A steered molecular dynamics method with direction optimization and its applications on ligand molecule dissociation. *J. Biochem. Biophys. Methods* **2008**, *70* (6), 857–64.

(24) Yang, K.; Liu, X.; Wang, X.; Jiang, H. A steered molecular dynamics method with adaptive direction adjustments. *Biochem. Biophys. Res. Commun.* **2009**, *379* (2), 494–8.

(25) Marti, M. A.; Bidon-Chanal, A.; Crespo, A.; Yeh, S. R.; Guallar, V.; Luque, F. J.; Estrin, D. A. Mechanism of product release in NO detoxification from Mycobacterium tuberculosis truncated hemoglobin N. *J. Am. Chem. Soc.* **2008**, *130* (5), 1688–93.

(26) Cohen, J.; Arkhipov, A.; Braun, R.; Schulten, K. Imaging the migration pathways for O_2, CO, NO, and Xe inside myoglobin. *Biophys. J.* **2006**, *91* (5), 1844–57.

(27) Bernstein, F. C.; Koetzle, T. F.; Williams, G. J.; Meyer, E. F., Jr.; Brice, M. D.; Rodgers, J. R.; Kennard, O.; Shimanouchi, T.; Tasumi, M. The Protein Data Bank. A computer-based archival file for macromolecular structures. *Eur. J. Biochem.* **1977**, *80* (2), 319–24.

(28) Georgescu, R. E.; Alexov, E. G.; Gunner, M. R. Combining conformational flexibility and continuum electrostatics for calculating pK(a)s in proteins. *Biophys. J.* **2002**, *83* (4), 1731–48.

(29) Rocchia, W.; Alexov, E.; Honig, B. Extending the applicability of the nonlinear Poisson-Boltzmann equation: Multiple dielectric constants and multivalent ions. *J. Phys. Chem. B* **2001**, *105*, 6507–6514.

(30) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. CHARMM: a program for macromolecular energy, minimization, and dynamics calculations. *J. Comput. Chem.* **1983**, *4*, 187–217.

(31) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **1983**, *79*, 926–935.

(32) Schneidman-Duhovny, D.; Inbar, Y.; Nussinov, R.; Wolfson, H. J. PatchDock and SymmDock: servers for rigid and symmetric docking. *Nucleic Acids Res.* **2005**, *33* (Web Server issue), W363−7.

(33) MacKerell, J. A. D.; Bashford, D.; Bellott, M.; Dunbrack, R. L., Jr.; Evanseck, J. D.; Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; Joseph-McCarthy, D.; Kuchnir, L.; Kuczera, K.; Lau, F. T. K.; Mattos, C.; Michnick, S.; Ngo, T.; Nguyen, D. T.; Prodhom, B.; Reiher, I.; W, E.; Roux, B.; Schlenkrich, M.; Smith, J. C.; Stote, R.; Straub, J.; Watanabe, M.; Wiorkiewicz-Kuczera, J.; Yin, D.; Karplus, M. All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J. Phys. Chem. B* **1998**, *102*, 3586–3616.

(34) Mackerell, A. D., Jr.; Feig, M.; Brooks, C. L., III. Extending the treatment of backbone energetics in protein force fields: limitations of gas-phase quantum mechanics in reproducing protein conformational distributions in molecular dynamics simulations. *J. Comput. Chem.* **2004**, *25* (11), 1400–15.

(35) Autenrieth, F.; Tajkhorshid, E.; Baudry, J.; Luthey-Schulten, Z. Classical force field parameters for the heme prosthetic group of cytochrome c. *J. Comput. Chem.* **2004**, *25* (13), 1613–22.

(36) Bathelt, C. M.; Zurek, J.; Mulholland, A. J.; Harvey, J. N. Electronic structure of compound I in human isoforms of cytochrome P450 from QM/MM modeling. *J. Am. Chem. Soc.* **2005**, *127* (37), 12900–8.

(37) Altun, A.; Thiel, W. Combined quantum mechanical/molecular mechanical study on the pentacoordinated ferric and ferrous cytochrome P450cam complexes. *J. Phys. Chem. B* **2005**, *109* (3), 1268–80.

(38) Sherwood, P.; de Vries, A. H.; Guest, M. F.; Schreckenbach, G.; Catlow, C. R. A.; French, S. A.; Sokol, A. A.; Bromley, S. T.; Thiel, W.; Turner, A. J.; Billeter, S.; Terstegen, F.; Thiel, S.; Kendrick, J.; Rogers, S. C.; Casci, J.; Watson, M.; King, F.; Karlsen, E.; Sjovoll, M.; Fahmi, A.; Schafer, A.; Lennartz, C. QUASI: A general purpose implementation of the QM/MM approach and its application to problems in catalysis. *J. Mol. Struct.: THEOCHEM* **2003**, *632*, 1–28.

(39) Fishelovitch, D.; Hazan, C.; Shaik, S.; Wolfson, H. J.; Nussinov, R. Structural dynamics of the cooperative binding of organic molecules in the human cytochrome P450 3A4. *J. Am. Chem. Soc.* **2007**, *129* (6), 1602–11.

(40) Cournia, Z.; Smith, J. C.; Ullmann, G. M. A molecular mechanics force field for biologically important sterols. *J. Comput. Chem.* **2005**, *26* (13), 1383–99.

(41) Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kale, L.; Schulten, K. Scalable molecular dynamics with NAMD. *J. Comput. Chem.* **2005**, *26* (16), 1781–802.

(42) Isralewitz, B.; Baudry, J.; Gullingsrud, J.; Kosztin, D.; Schulten, K. Steered molecular dynamics investigations of protein function. *J. Mol. Graphics Modell.* **2001**, *19* (1), 13–25.

(43) Petersen, H. G. Accuracy and efficiency of the particle mesh Ewald method. *J. Chem. Phys.* **1995**, *103*, 3668–3679.

(44) Poulos, T. L. Cytochrome P450 flexibility. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100* (23), 13121–2.

(45) Yaffe, E.; Fishelovitch, D.; Wolfson, H. J.; Halperin, D.; Nussinov, R. MolAxis: a server for identification of channels in macromolecules. *Nucleic Acids Res.* **2008**, *36* (Web Server issue), W210–5.

(46) Ludemann, S. K.; Lounnas, V.; Wade, R. C. How do substrates enter and products exit the buried active site of cytochrome P450cam? 1. Random expulsion molecular dynamics investigation of ligand access channels and mechanisms. *J. Mol. Biol.* **2000**, *303*, 797–811.

(47) Ludemann, S. K.; Lounnas, V.; Wade, R. C. How do substrates enter and products exit the buried active site of cytochrome P450cam? 2. Steered molecular dynamics and adiabatic mapping of substrate pathways. *J. Mol. Biol.* **2000**, *303* (5), 813–30.

(48) Roussel, F.; Khan, K. K.; Halpert, J. R. The importance of SRS-1 residues in catalytic specificity of human cytochrome P450 3A4. *Arch. Biochem. Biophys.* **2000**, *374* (2), 269–78.

(49) Khan, K. K.; He, Y. Q.; Domanski, T. L.; Halpert, J. R. Midazolam oxidation by cytochrome P450 3A4 and active-site mutants: an evaluation

of multiple binding sites and of the metabolic pathway that leads to enzyme inactivation. *Mol. Pharmacol.* **2002**, *61* (3), 495–506.

(50) Wester, M. R.; Yano, J. K.; Schoch, G. A.; Yang, C.; Griffin, K. J.; Stout, C. D.; Johnson, E. F. The structure of human cytochrome P450 2C9 complexed with flurbiprofen at 2.0-A resolution. *J. Biol. Chem.* **2004**, *279* (34), 35630–7.

(51) Haines, D. C.; Tomchick, D. R.; Machius, M.; Peterson, J. A. Pivotal role of water in the mechanism of P450BM-3. *Biochemistry* **2001**, *40* (45), 13456–65.

(52) Rowland, P.; Blaney, F. E.; Smyth, M. G.; Jones, J. J.; Leydon, V. R.; Oxbrow, A. K.; Lewis, C. J.; Tennant, M. G.; Modi, S.; Eggleston, D. S.; Chenery, R. J.; Bridges, A. M. Crystal structure of human cytochrome P450 2D6. *J. Biol. Chem.* **2006**, *281* (11), 7614–22.

(53) Williams, P. A.; Cosme, J.; Sridhar, V.; Johnson, E. F.; McRee, D. E. Mammalian microsomal cytochrome P450 monooxygenase: structural adaptations for membrane binding and functional diversity. *Mol. Cell* **2000**, *5* (1), 121–31.

(54) Nakayama, K.; Puchkaev, A.; Pikuleva, I. A. Membrane binding and substrate access merge in cytochrome P450 7A1, a key enzyme in degradation of cholesterol. *J. Biol. Chem.* **2001**, *276* (33), 31459–65.

(55) Kumar, S.; Chen, C. S.; Waxman, D. J.; Halpert, J. R. Directed evolution of mammalian cytochrome P450 2B1: mutations outside of the active site enhance the metabolism of several substrates, including the anticancer prodrugs cyclophosphamide and ifosfamide. *J. Biol. Chem.* **2005**, *280* (20), 19569–75.

(56) Yamazaki, H.; Shimada, T. Progesterone and testosterone hydroxylation by cytochromes P450 2C19, 2C9, and 3A4 in human liver microsomes. *Arch. Biochem. Biophys.* **1997**, *346* (1), 161–9.

(57) Krauser, J. A.; Guengerich, F. P. Cytochrome P450 3A4-catalyzed testosterone 6β-hydroxylation stereochemistry, kinetic deuterium isotope effects, and rate-limiting steps. *J. Biol. Chem.* **2005**, *280* (20), 19496–506.

JP810386Z