PROTEINS: Structure, Function, and Bioinformatics 64:147–155 (2006)

Multiple Molecular Dynamics Simulations of Human P450 Monooxygenase CYP2C9: The Molecular Basis of Substrate Binding and Regioselectivity Toward Warfarin

Alexander Seifert, Stephan Tatzel, Rolf D. Schmid, and Jürgen Pleiss*
Institute of Technical Biochemistry, University of Stuttgart, Stuttgart, Germany

ABSTRACT **To examine the molecular basis of activity and regioselectivity of the clinically important human microsomal cytochrome P450 (CYP) monooxygenase 2C9 toward its substrate warfarin, 22 molecular dynamics simulations (3–5 ns each) were performed in the presence and absence of warfarin. The resulting trajectories revealed a stable protein core and mobile surface elements. This mobility leads to the formation of two surface channels in the region between F-G loop, B′ helix/B-B′ loop, β_1-sheet, and between helices F and I and the turn in the C-terminal antiparallel β-sheet in the presence of warfarin. Besides the nonproductive state of the CYP2C9 warfarin complex captured in the crystal structure, three additional states were observed. These states differ in the shape of the substrate binding cavity and the position of the warfarin molecule relative to heme. In one of these states, the 7- and 6-positions of warfarin contact the heme with a marked geometrical preference for position 7 over position 6. This modeling result is consistent with experimentally determined regioselectivity (71 and 22% hydroxylation in positions 7 and 6, respectively). Access to the heme group is limited by the core amino acids Ala297, Leu362, Leu366, and Thr301, which therefore are expected to have a major impact on regioselectivity. In addition, modeling predicts that autoactivation of warfarin is sterically hindered. Our study demonstrates how the combination of mobile surface and rigid core leads to interesting properties: a broad substrate profile and simultaneously a high regioselectivity. Proteins 2006;64:147–155.** © 2006 Wiley-Liss, Inc.

Key words: enzyme dynamics; loop movements; channels

INTRODUCTION

Cytochrome P450 monooxygenases (CYPs) are involved in the metabolism of physiologically important compounds in many species of microorganisms, plants, and animals. In mammals, these enzymes participate in the detoxification of a broad range of xenobiotics such as environmental toxins and drugs. Understanding the factors involved in CYP substrate selectivity is of considerable interest, particularly to the pharmaceutical industry[1–3] where an early prediction of likely drug metabolism pathways in *Homo sapiens* can aid in the design and development of drugs. CYP2C9 is one of the major drug metabolizing isoforms. It contributes to the oxidative metabolism of 16% of all therapeutics in current clinical use[4] and is involved in adverse drug effects. Human CYPs metabolize substrate molecules of different size and shape which implies a large, highly flexible substrate binding cavity. The common anticoagulant drug warfarin is metabolized by CYP2C9 to form 7-hydroxywarfarin (71%), 6-hydroxywarfarin (22%), and 4-hydroxywarfarin (7%).[5] Similar to many other CYPs,[6] CYP2C9 has been shown to have a large substrate binding site that can simultaneously bind several substrate molecules.[7,8] There is experimental evidence for a "two-site binding model" for CYP2C9-mediated metabolism[9]; however, there is no autoactivation by warfarin.[10] On the other hand, most of the catalyzed reactions are highly regioselective which would require a more rigid enzyme.

In effort to understand the molecular basis of substrate specificity and regioselectivity, the free enzyme and the enzyme-substrate complexes of several mammalian CYPs have been studied using X-ray analysis, docking, and molecular dynamics (MD) simulations.[11,12] The comparison of free and substrate-complexed crystal structures of mammalian CYPs revealed that helices F and G and the loop between helices B and C undergo adaptive changes in the presence of substrates and inhibitors.[13,14] Dramatic conformational changes were observed when the crystal structure of mammalian CYP2B4 was solved.[15] The structure of the inhibitor and substrate-free enzyme exhibits a large open cleft which is formed mainly by helices F, F′, G′, and G on one side and helices B′ and C on the opposite side without loss of the general architecture of the CYP fold. The cleft extends from the protein surface directly to the

Abbreviations: CYP, cytochrome P450 monooxygenase; MD, molecular dynamics; REMD, random expulsion molecular dynamics; RMSD, root-mean-square deviation.

The Supplementary Material referred to in this article can be found at http://www.interscience.wiley.com/jpages/0887-3585/suppmat/

Grant sponsor: Bundesministerium für Bildung und Forschung (BMBF).

*Correspondence to: Jürgen Pleiss, Institute of Technical Biochemistry, University of Stuttgart, Allmandring 31, 70569 Stuttgart, Germany. E-mail: juergen_pleiss@itb.uni-stuttgart.de

Received 26 July 2005; Revised 3 November 2005; Accepted 16 December 2005

Published online 25 April 2006 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/prot.20951

heme. However, the crystal structure of CYP2B4 in presence of a specific inhibitor revealed a closed conformation of the protein.[15] In contrast, the crystal structure of the human microsomal CYP2C9[7] which has been crystallized in the absence and presence of its substrate warfarin did not show major conformational changes between the substrate-free structure and the enzyme-substrate complex. Furthermore, the analysis of channels in P450 crystal structures as well as random expulsion MD simulations of bacterial CYP101, CYP102A1, CYP107A1, and mammalian CYP2C5 predicted for soluble CYPs one predominant ligand pathway situated between F-G loop, B′ helix/B-B′ loop, and β_1-sheet, whereas for mammalian membrane-bound CYP2C5 one additional route for substrates or products was suggested in the region between B′ helix/B-C loop, G and I helices.[11,16] However, the only channel in the crystal structure of CYP2C9 is a solvent channel between helices F and I and the turn in the C-terminal antiparallel β-sheet, which is too small for substrate molecules to pass through.[11] How potential substrates access and products leave the substrate binding cavity remains elusive.

In all CYPs, the cofactor heme is deeply buried inside the protein at the bottom of a large, internal binding cavity. It has been suggested that the orientation of the bound substrate in the substrate binding cavity relative to the heme determines regioselectivity.[17] However, the CYP2C9 crystal structure in complex with warfarin does not provide information about the molecular basis of regioselectivity, because the substrate lies in a predominantly hydrophobic pocket at a distance of 10 Å from the heme.[7] In contrast, in a 2.0 Å resolution crystal structure of the CYP2C9 in complex with flurbiprofen, the substrate is close to the heme in an orientation consistent with the experimentally determined regioselectivity of the enzyme.[8] However, in this structure, residues 214–220 between helices F and G and residues 38–42 are lacking.

We performed extensive MD simulations to locate the regions of human CYP2C9, where substrate access and product exit might occur, and to identify the structural basis of regioselectivity of this enzyme.

MATERIALS AND METHODS

Initial Structures

Multiple MD simulations of CYP2C9 in water and complexed with the substrate (*S*)-warfarin in water were performed. The initial structures were created from X-ray structures of free CYP2C9 [Protein Data Bank (PDB)[18] entry 1OG2] and of a complex with (*S*)-warfarin (PDB entry 1OG5) at a resolution of 2.6 Å. Both crystal structures contain two protein molecules in the asymmetric unit cell with an all-atom deviation of less than 0.65 Å. For the simulations, chain A was used as initial structure. It should be noted that the sequence of 1OG2 and 1OG5 deviates from human CYP2C9. To improve crystallization, the protein was truncated by removing amino acids 1–29 and seven amino acids were exchanged (Lys206Glu, Ile215Val, Cys216Tyr, Ser220Pro, Pro221Ala, Ile222Leu, and Ile223Leu). It has been shown that none of the amino acid substitutions leads to changes in activity, specificity, and the ability to metabolize 4-hydroxylation of diclofenac and 6- and 7-hydroxylation of (*S*)-warfarin.[7]

The protonation states at pH 7.0 of all histidines, glutamic, and aspartic acids were calculated using the program MCCE,[19] of the other titratable amino acids using the program xleap of the AMBER 7.0 program.[20] The partial charges of (*S*)-warfarin and the heme group in the oxyferryl state (Fe^{3+}═O) were derived by fitting partial charges using the RESP program of AMBER 7.0 to the electrostatic potential derived by ab initio geometry optimization on an HF/6-31G* level by using the GAUSSIAN 98[21] program [Florian Barth, personal communications; heme and warfarin parameters are given in the Supplementary Material (http://www.interscience.wiley.com/jpages/0887-3585/suppmat/)]. The oxyferryl state (Fe^{3+}═O) is considered to be the main reactive species of the CYP catalytic cycle. The all-atom AMBER force field "ff99" was used to represent the protein system. Bonded and nonbonded parameters for (*S*)-warfarin and the heme group were derived from the AMBER libraries. The xleap program was used to solvate the free enzyme and the CYP2C9–(*S*)-warfarin complex in a truncated octahedron box of TIP3P water with a minimal distance of 12.0 Å between the box boundary and the protein, and a closeness of 0.42 Å. One sodium ion was added to neutralize the system.

MD Simulation

The PMEMD program of AMBER 7.0 was used for minimization and MD simulations. The initial structures were first energy minimized for 2,000 steps (1,000 steepest descent and 1,000 conjugate gradient) and then simulated at a temperature of 300 K. The Shake algorithm[22] was applied to all bonds containing hydrogen atoms, and a time step of 1 fs was used. The Berendsen method was used to couple the system to constant temperature and pressure.[23] For each system, multiple MD simulations (6 and 16 simulations for the free and complexed enzyme, respectively) were performed with different initial random velocity distributions. Each distribution was generated by a random number generator. A specific IG value was used as initial seed for the generator. For the free enzyme, the C_α atoms were restrained for 6 ps using a harmonic potential with a gradually decreasing force constant from 10 to 0.1 kcal/mol followed by a unrestrained simulation of 3 ns. For the complex, the C_α atoms were restrained for 6 ps, the substrate atoms for 600 ps with a gradually decreasing force constant from 10 to 0.1 kcal/mol, followed by an unrestrained simulation of 4.5 ns. During each simulation, snapshots of the system were taken every 500 fs and stored to a trajectory file.

Analysis of the Trajectories

The calculation of the root-mean-square deviation (RMSD) of the backbone atoms between each conformer and the X-ray structure and between all conformers [two-dimensional (2D)-RMSD], measurements of atom–atom distances, and the generation of average structures were done with the ptraj program of AMBER 7.0 by fitting the

backbone atoms of each conformer to the initial structure. The RMSD between the substrate conformers generated by the 16 simulations of the enzyme-substrate complex was calculated using the PROFIT program.[24] Warfarin is bound to the B-C loop via hydrogen bonds in the crystal structure. Because this loop is one of the mobile regions of the protein, the average structures were fitted in this region (AS 91–117) before RMSD calculation to exclude changes in the position of warfarin relative to its position in the crystal structure because of movements of the B-C loop. The RMSD values were used to construct a distance-based tree with the program NEIGHBOR of the PHYLIP program package[25] applying the unweighted pair group method with arithmetic mean (UPGMA). Trees were visualized with TreeView 1.6.1.[26] The protein structure was visualized using the PyMOL 0.93 program.[27] The VOIDOO[28] program was used to calculate the solvent accessible surface of the substrate binding cavity. To monitor the distance of potential sites of hydroxylation of the substrate to the active site, the distances between the oxygen bound to the heme iron and hydrogens bound to potential sites of hydroxylation of the substrate were measured.

Estimation of Chemical Reactivity of Warfarin

An estimation of chemical reactivity of warfarin was done by calculating the bond dissociation energy of all C—H bonds using the PETRA program.[29]

RESULTS

Stability of the Simulations

Six MD simulations (3 ns each) of CYP2C9 in water and 16 simulations (5 ns each) of CYP2C9 in complex with its substrate warfarin in water were performed with different distributions of the initial velocities. As initial structure for the simulations of the free enzyme and the enzyme-substrate complex, the energy minimized crystal structures of free enzyme and the warfarin complex, respectively, were used. To evaluate the deviation of the trajectory from the initial structure, the RMSD was monitored along the trajectory. For all simulations, the protein core (residues 48–57, 64–68, 72–92, 117–133, 141–155, 162–166, 285–328, 335–339, 347–371, 385–401, 406–412, 421–457, 469–473, and 477–490 according to CYP2C9 wild-type sequence numbering) was stable during the simulations. After fitting the core region, the backbone RMSD from the crystal structure was between 1.5 and 2.0 Å. Two-dimensional-RMSD plots of each simulation revealed that this deviation occurred during the first nanosecond of the simulations (data not shown); after 1.5 ns, the core region was equilibrated as the deviations between the conformers of one simulation were less than 1.5 Å, and the deviation between any two simulations was less than 2.4 Å (calculated from average structures of the last 100 ps of each simulation). In contrast, the deviation of all backbone atoms during the simulation was considerably larger: 2.0 and 2.5 Å and 2.0–3.0 Å deviation from the crystal structure for the free enzyme and the substrate complex, respectively. The deviation between any two simulations

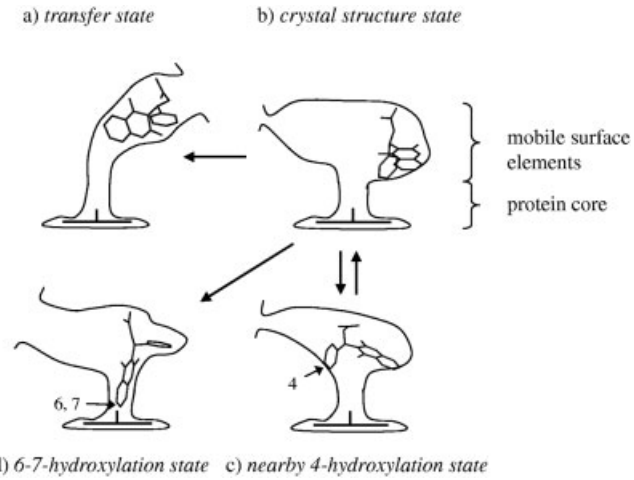

Fig. 1. Shape of the substrate binding cavity of the CYP2C9 warfarin complex in the (**a**) transfer state, (**b**) crystal structure state, (**c**) nearby 4-hydroxylation state, and (**d**) 6-7-hydroxylation state. Arrows indicate the transitions observed using MD simulations.

was up to 3.8 Å which indicates that there are highly mobile regions outside the core. Furthermore, the 2D-RMS plots demonstrated that these mobile regions of the protein are not yet in equilibrium. The backbone RMSD of the average structure of all simulations to the X-ray structure is 1.6 Å and 1.3 Å for the complex and the free enzyme, respectively, which is much less than the deviations between any two simulations or the deviation between each simulation and the crystal structure. Therefore, the different simulations did not converge but scattered around the crystal structure.

We focused our analysis to those structure elements that are involved in the formation of the substrate binding cavity, because conformational changes in this region of the protein have been identified to change the shape of the active site containing cavity and lead to the formation of channels suitable for substrate access or product exit. From the 16 simulations of the enzyme-substrate complex, four distinct states were identified, which differ in the shape of the substrate binding cavity and the position of the substrate: the *crystal structure state*, the *transfer state*, the *6-7-hydroxylation state*, and the *nearby 4-hydroxylation state* (Fig. 1). We observed transition from crystal structure state to the other three states and from the nearby 4-hydroxylation state back to the crystal structure state. The substrate binding cavity is formed mainly by helices F and G and the loop between them (F-G loop) as well as the loop between the B and C helices (B-C loop), which contains the B′ helix. These structural elements belong to the mobile regions of the protein, as the RMSD values for the B-C loop deviated at the end of the simulations of the free enzyme and substrate complex 1.8–2.9 Å and 1.2–3.5 Å from the crystal structure after fitting in the core region, respectively. For the F-G loop, values between 2.0 Å and 4.5 Å for the enzyme-substrate complex and 2.0–4.0 Å for the free enzyme were measured. Additional mobile elements involved in the formation of substrate binding cavity are the N-terminal residues, the turn in the

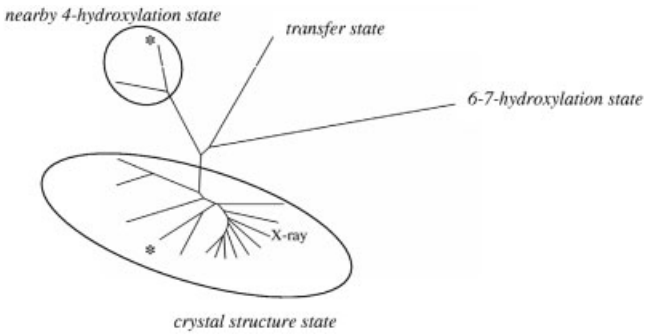

Fig. 2. Tree of warfarin conformers based on RMSD between all atoms except hydrogen observed during multiple MD simulations. "X-ray" indicates the crystal structure, which was the starting point of all simulations with substrate. The conformers marked with an asterisk were observed during one simulation, where the substrate first moved from its position in the crystal structure state to the nearby 4-hydroxylation state and then back to crystal structure state.

C-terminal antiparallel β-sheet, and the turn in β_1-sheet. In contrast, the heme pocket and a narrow funnel (Ala297, Leu362, Leu366, Thr301) leading from the substrate binding cavity to the heme belong to the core region and were structurally similar in all states.

Four States

The four states were identified by analyzing the relative positions of substrate in the binding cavity. RMSD values of warfarin between any two simulations (averaged over 100 ps) and to the crystal structure were calculated and used to construct a distance tree (Fig. 2). The crystal structure state was observed during 12 of 16 simulations. The transfer state was observed in one simulation where the shape of the substrate binding site cavity is massively changing: a channel from the protein surface to the binding cavity [Fig. 3(a)] is generated by a shift of the B-C loop toward the I-helix and a slight turn of the β_1-sheet toward the N-terminus [Fig. 4(a)]. Furthermore, the side-chains of Phe100 and Lys72 moved out of the region between B′B loop and β_1-sheet and thereby enlarge the channel. The substrate warfarin binds to this channel [Fig. 4(b)]. To examine if the presence of the substrate in this region facilitates the formation of the channel, the distance between B-B′ loop (C_α of Ile99) and the adjacent β_1-sheet (C_α of Pro73) was measured during all simulations. In the transfer state, this distance increased to up to 19 Å as compared with 12.5 and 12.3 Å in the two crystal structures. In all other simulations, the distance varied between 11 and 15 Å. Thus, the substrate seems to stabilize the channel.

The 6-7-hydroxylation state was observed in one simulation. In this state, the substrate points toward heme with its 6- and 7-position near to the active heme oxygen [Fig. 1(d)]. The distance between the hydrogen bound to carbon 7 of warfarin (Fig. 5) and the active heme oxygen decreased to 2–3 Å, whereas in the other simulations this distance was between 8 and 14 Å (Fig. 6). In this state warfarin is stabilized by van der Waals contacts to Phe114, Ala297, Thr301, Leu362, Leu366, and Phe476. Remarkably the 7-position is three times more frequently at a distance of less than 3 Å from the active oxygen than the 6-position [Fig. 7(a)]. For a distance less than 2.5 Å this ratio even increases to 1:8. Analysis of the time course of these distances indicates that the substrate performs a diffusional motion in a single energy minimum rather than jumping between two distinct, equilibrated substrates [Fig. 7(b)]. Four different regions can be assigned: both distances less than 3 Å (area 1), one of the distances less than 3 Å (area 2 and 3), or both distances larger than 3 Å (area 4). A comparison of area 2 and area 3 reveals that the 7-hydroxylation site is considerably more frequent in a distance at which hydroxylation can occur than the 6-hydroxylation site. This clear geometrical preference is caused by the well-defined position and orientation of the substrate in the narrow funnel and the binding cavity. Therefore, the regioselectivity is a consequence of the overall shape of the protein rather than caused by local interaction near the heme oxygen.

In this state, an additional channel from the bulk solvent to the binding cavity is observed [Fig. 3(b)] which is different from the channel observed in the transfer state. It is located between helices F and I and the turn in the C-terminal antiparallel β-sheet and could already be seen as a narrow solvent channel in the crystal structure, just wide enough for water molecules to enter. The increase of the channel is mainly caused by backbone and side-chain movements of Ile205 and Ser209 (mobile F helix) as well as side-chain movements of Glu300 (stable I helix). The distance between C_α atoms of Ile205 (mobile F helix) and Glu300 increased to 14 Å as compared with 7.6 Å in the crystal structure. In all other simulations, this channel slightly opened with a distance between 9 and 11 Å.

The nearby 4-hydroxylation state of the substrate enzyme complex was observed during two simulations. In one of these simulations the system even moved back to the crystal structure state. In nearby 4-hydroxylation state the warfarin orients its phenyl group which contains the 4-position toward the active heme oxygen. The measured minimum distance of 7 Å between the 4-position and the active heme oxygen is considerably less than in the other states (11–17 Å), but not close enough to enable hydroxylation (Fig. 8).

The crystal structure state is as a stable conformation which itself is not active but has been suggested to have an important role in activation of other substrates by warfarin. However, the positions of warfarin in the 6-7-hydroxylation state or the nearby 4-hydroxylation state are overlapping to its position in the crystal structure state (Fig. 8). Thus, two of these states cannot simultaneously be occupied by warfarin.

DISCUSSION

CYP2C9 Channels

How do hydrophobic substrates enter the binding cavity, and how do more hydrophilic products leave it? In the X-ray structures of free and complexed human CYP2C9, the binding cavity was permanently connected to the bulk solvent by a narrow solvent channel between helices F and

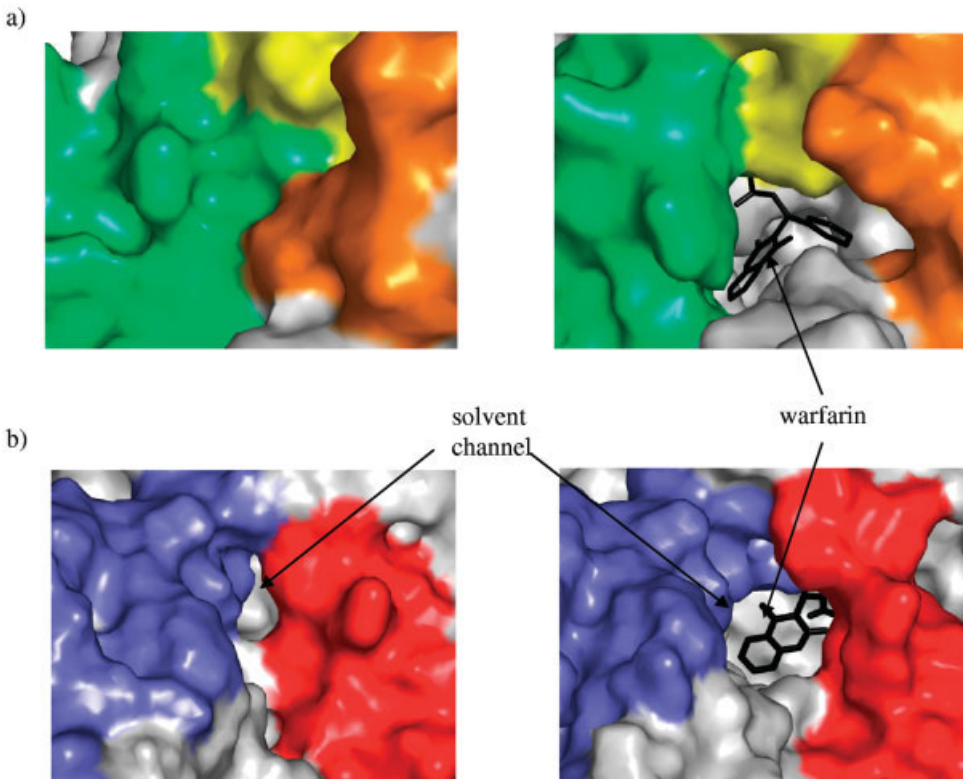

Fig. 3. Surface of CYP2C9 (**a**) in the region between F-G loop (yellow), B′ helix/B-B′ loop (green), and β_1-sheet (orange): crystal structure (left), transfer state (right); and (**b**) in the region of the solvent channel between the F helix (red) and the turn in the C-terminal antiparallel β-sheet (blue): crystal structure (left), 6-7-hydroxylation state (right).

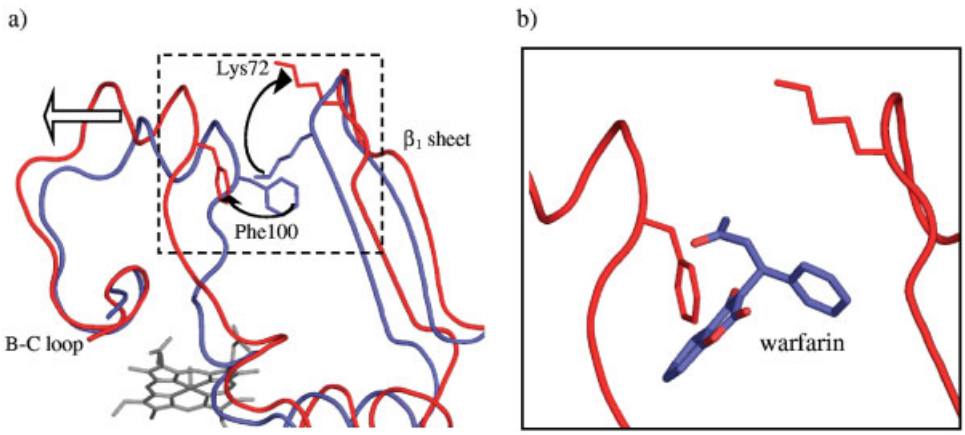

Fig. 4. (**a**) A more open conformation was observed in the transfer state (red) compared with the crystal structure (blue). The open conformation was achieved because of movements of the B-C loop, β_1-sheet, and the side-chains of Phe100 and Lys72. Arrows indicate the direction of movements of the different structure elements. (**b**) The opening was stabilized by the substrate warfarin.

I and the turn in the C-terminal antiparallel β-sheet,[7] which, however, is too narrow to allow a substrate to pass through. Therefore, we analyzed the effect of substrate binding to the occurrence of channels between the binding pocket of CYP2C9 and the bulk solvent by multiple MD simulations. In the free enzyme, narrow channels between cavity and bulk solvent were formed temporarily mainly by fluctuations of the FG and BC loops. In the presence of substrate, the shape of the binding cavity changed. We observed four distinct positions of the substrate and thus four distinct shapes of the binding cavity. In two of them, the transfer state and the 6-7-hydroxylation state, stable channels are formed wide enough for substrate or product.

For several family 2 CYPs, FG and BC loops have been identified by antibody binding data and hydrophobicity analysis to interact with the membrane of the endoplas-

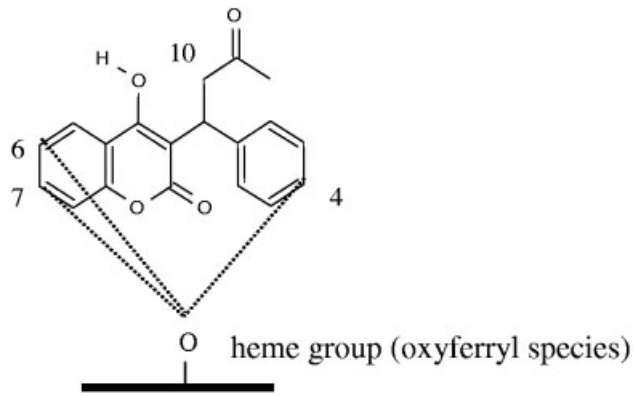

Fig. 5. Measured distances between warfarin and the active heme oxygen of CYP2C9; the 4-, 6-, 7-, and 10-hydroxylation sites of warfarin are labeled.

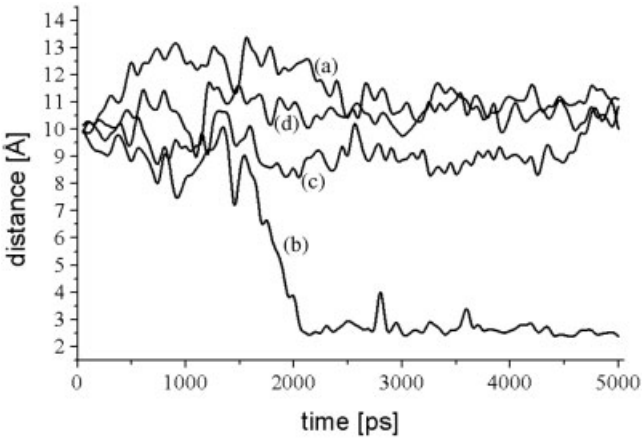

Fig. 6. Distance between the hydrogen atom bound to C7 of warfarin and the active oxygen of the heme group during the simulations of CYP2C9 with warfarin in water for representative trajectories in the crystal structure state (**a**), 6-7-hydroxylation state (**b**), transfer state (**c**), and nearby 4-hydroxylation state (**d**).

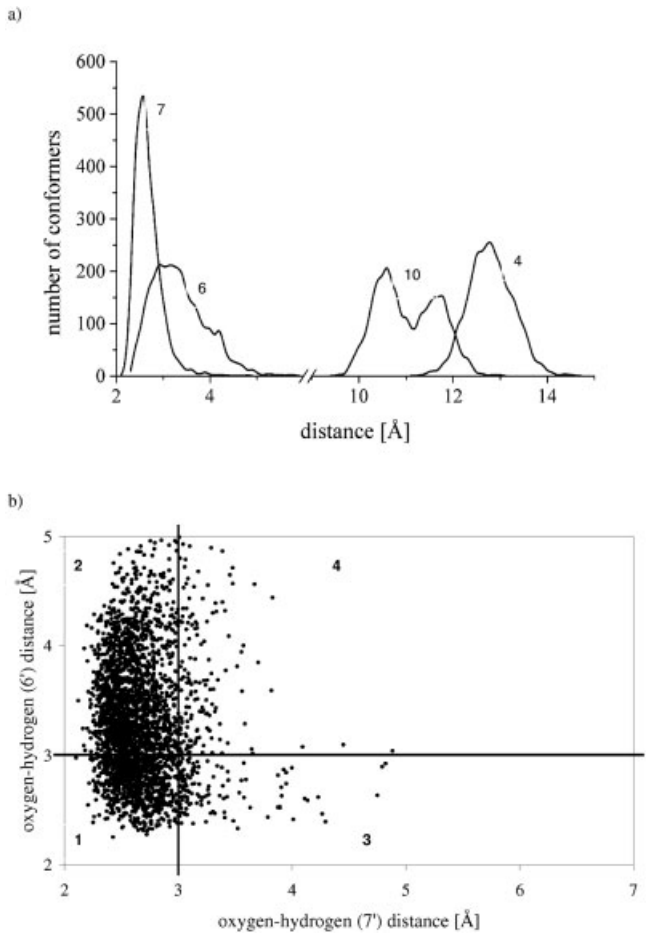

Fig. 7. (**a**) Distribution of measured distances between 4-, 6-, 7-, and 10-hydroxylation site of warfarin, respectively, and the active oxygen in the 6-7-hydroxylation state of the protein substrate complex. In contrast to position 6 and 7, the 4- and 10-hydroxylation site of warfarin do not contact the active heme oxygen. (**b**) Two-dimensional plot of distances between 6- and 7-hydroxylation site and the active heme oxygen. 1) both distances <3 Å; 2) distance FeO–HC7 <3 Å; FeO–HC6 >3 Å; 3) distance FeO–HC7 >3 Å; FeO–HC6 <3 Å; 4) both distances >3 Å.

matic reticulum,[30] from where the hydrophobic substrates are assumed to access the enzyme preventing contact to the bulk solvent. Fluctuations of this putative substrate access channel were observed for the free enzyme and the enzyme-substrate complex during our simulations. A recent MD simulation study of the free CYP2C9 enzyme also revealed fluctuations in this region of the protein; however, the existence of a channel with suitable size for substrate access was not detected.[31] However, our MD simulations indicate that these narrow channels are stabilized and widened in the presence of substrate. The stabilization of an open conformation via a second molecule was also observed for mammalian CYP2B4. In absence of an inhibitor, the crystal structure reveals a large open cleft, which is primarily formed by helices B′ to C and F to G.[32] In this case, the wide open structure of the substrate-free CYP2B4 is stabilized in the crystal by dimerization, where two protein molecules stabilize each other through hydrophobic interactions. Thus, in X-ray structures as well as in our MD simulations the channel and the binding cavity have a high plasticity and are stabilized by molecular interactions. This plasticity is consistent with the wide range of substrates that are accepted by CYP2C9 and other CYPs.[4] The induction of a channel between F-G loop, B′ helix/B-B′ loop, and β_1-sheet has also been found previously by random expulsion MD simulations of bacterial CYP101, CYP102A1, CYP107A1, and mammalian CYP2C5, where a bulky substrate was pushed out of the binding cavity by an external random force.[11,33]

A solvent channel in the region between helices F and I and the turn in the C-terminal antiparallel β-sheet is visible in the free and complexed crystal structure of CYP2C9 as well as in several bacterial and mammalian CYPs.[11] Because this channel directly leads from the heme group into the bulk solvent, it would be favorable for the hydrophilic product to leave the binding cavity via this pathway. However, in the X-ray structure, this channel is too narrow for products to pass through. Binding of the substrate increased its size considerably: in the 6-7-

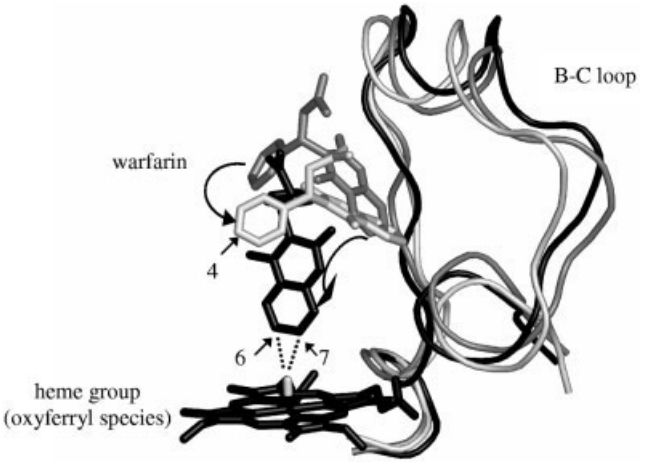

Fig. 8. Comparison of the warfarin conformers observed in crystal structure state (gray), 6-7-hydroxylation state (black), and nearby 4-hydroxylation state (white). All conformers are partially overlapping. Arrows mark the 4-, 6-, and 7-hydroxylation site of warfarin.

hydroxylation state, the channel became wide enough to allow bulky product to pass through. Three different channels have been described for mammalian CYP2C5.[16] In addition to the two channels that we have observed in our study, standard MD and REMD simulations revealed an additional channel between B′ helix/B-C loop, G and I helices. Also, for a completely different enzyme class, the acetylcholinesterases,[34,35] two different pathways have been suggested to be the structural basis of their extremely high catalytic activity.

Structural Basis of the Regioselectivity of CYP2C9

A calculation of bond dissociation energy[29] of C—H groups in warfarin predicted that the reactivity of more than 10 positions would be sufficiently high for hydroxylation to occur. However, CYP-catalyzed hydroxylation of warfarin is highly regioselective and depends on the shape of the substrate binding cavity. Whereas CYP2C9 catalyzes hydroxylation at the 7-, 6-, and 4-position with a product ratio of 71:22:7, respectively,[5] CYP3A4 hydroxylates warfarin at position 4 and 10.[36] However, the experimentally observed regioselectivity of CYP2C9 cannot be predicted from the X-ray structure of its complex with warfarin, because warfarin binds at a hydrogen–oxygen distance of 10 Å from the heme iron.[7] Because a contact between the substrate hydrogen and the oxygen of the oxyferryl heme (compound I) is prerequisite to hydroxylation,[17] we considered only conformations as reactive if their H· · ·O distance was less than 3 Å, the distance of a C—H· · ·O hydrogen bond.[37] Although the conformation of the crystal structure was therefore considered to be nonproductive, in one simulation a transition to the reactive 6-7-hydroxylation state was observed with close contacts between the oxyferryl and the hydrogens at positions 6 and 7. Because the 7-hydroxylation site is much more frequent within a distance of 3 Å as compared with the 6-hydroxylation site, we would predict a marked preference of the 7-hydroxylation over the 6-hydroxylation, which is consistent with the experimentally measured ratio of 75:25 for 7- to 6-hydroxywarfarin. Another consequence of the prediction of a 6-7-hydroxylation state is a prediction of residues which mediate the regioselectivity. Phe476 has been shown to change the regioselectivity of diclofenac hydroxylation.[38] It is located in the binding cavity and mediates the orientation of warfarin in the 6-7-hydroxylation state despite its distance of more than 14 Å from the heme. Most relevant for regioselectivity, however, are the four residues that form a narrow, rigid funnel leading from the binding cavity to the heme. Thus, we would predict that mutation of Ala297, Thr301, Leu362, and Leu366 should have a drastic effect on regioselectivity. This selectivity-determining role of a narrow funnel which limits access of substrate to the heme is consistent with the observation that 4-hydroxylation of warfarin is rare. The 6 and 7 sites are positioned at the far end of the base of the Y-shaped warfarin molecule, whereas the 4 site is at its more bulky end. Because of the limited access to the heme, the 4-position is hindered, whereas the 6- and 7-positions reach down the narrow funnel [Fig. 1(d)]. Therefore, replacing Leu362 or Leu366 by less bulky residues is expected to facilitate formation of 4-hydroxywarfarin. The existence of a productive substrate conformation at a distance of less than 3 Å from the heme has recently been confirmed in a crystal structure of flurbiprofen in complex with CYP2C9.[8] In contrast to the more bulky warfarin, this rod-shaped molecule fits well into the narrow funnel and contacts the heme similar to the 6-7-hydroxylation state. The superimposition of an average structure of the 6-7-hydroxylation state and 1R9O reveals that both substrates bind to the rigid, regioselectivity-determining funnel in a rather similar orientation with their respective hydroxylation sites deviating by only 1.4 Å. The funnel forming amino acids only deviate about 1.1 Å (all atom) from the 1R9O structure. Although the two complexes are highly similar in the funnel region, they deviate considerably elsewhere. Whereas these conformational differences may account for differences in catalytic activity, the regioselectivity can be easily explained by the existence of a rigid, narrow funnel.

Role of the Binding Site for Warfarin in the Crystal Structure in Enzyme Kinetics Such as Activation and Autoactivation

The large substrate binding cavity of CYP2C9 (470 Å^3) can accommodate more than one substrate molecule simultaneously. Nuclear Magnetic resonance titration and T_1 relaxation time measurements indicated that dapsone and flurbiprofen are simultaneously occupying the substrate binding cavity of CYP2C9.[39] It has been suggested that in addition to the warfarin molecule in the crystal structure, a second warfarin molecule could be accommodated in the binding cavity of CYP2C9.[7] This is, however, in contrast to the experimental observation that autoactivation is not observed, while binding of warfarin activates hydroxylation of 7-methoxy-4-trifluoromethylcoumarin.[40] According to our simulation results, simultaneous binding of warfarin to the crystal structure state and to the active 6-7-

hydroxylation state is sterically hindered, while the warfarin complex in the crystal structure state would still leave sufficient space for other substrate molecules such as flurbiprofen to bind close to the active heme oxygen.

CONCLUSION

How can human CYPs manage to have simultaneously a broad substrate profile and a high regioselectivity? The substrate binding sites of CYPs consist of two elements: a large, mobile binding cavity connected to the protein surface via substrate-induced channels, and a narrow, rigid, and hydrophobic funnel connecting the binding cavity and the active-site heme. Substrate specificity and regioselectivity is a result of a delicate balance between different states of the substrate-enzyme complex. The dynamics of the substrate binding site is crucial for understanding the biochemical properties of the enzyme and for predicting the effect of mutations. Substrate recognition in the binding cavity is by no means static, but the result of adaptive motions of the enzyme. In contrast, regioselectivity is determined by the shape of the substrate and its highly restricted orientation by a narrow, rigid funnel between binding cavity and the active-site heme group.

ACKNOWLEDGMENTS

The authors thank Michael Krahn for the calculation and visualization of the substrate binding site cavity with the program VOIDOO and Florian Barth for the parameterization of the oxyferryl heme group.

REFERENCES

1. Smith DA, Ackland MJ, Jones BC. Properties of cytochrome P450 isoenzymes and their substrates. Part 1. Active site characteristics. Drug Discov Today 1997;2(10):406–414.
2. Smith DA, Ackland MJ, Jones BC. Properties of cytochrome P450 isoenzymes and their substrates. Part 2. Properties of cytochrome P450 substrates. Drug Discov Today 1997;2(11):479–486.
3. Smith DA, Abel SM, Hyland R, Jones BC. Human cytochrome P450s: selectivity and measurement in vivo. Xenobiotica 1998; 28(12):1095–1128.
4. Rendic S, Di Carlo FJ. Human cytochrome P450 enzymes: a status report summarizing their reactions, substrates, inducers, and inhibitors. Drug Metab Rev 1997;29(1–2):413–580.
5. Rettie AE, Korzekwa KR, Kunze KL, et al. Hydroxylation of warfarin by human cDNA-expressed cytochrome P-450: a role for P-4502C9 in the etiology of (S)-warfarin-drug interactions. Chem Res Toxicol 1992;5(1):54–59.
6. Hlavica P, Lewis DF. Allosteric phenomena in cytochrome P450-catalyzed monooxygenations. Eur J Biochem 2001;268(18):4817–4832.
7. Williams PA, Cosme J, Ward A, Angove HC, Matak Vinkovic D, Jhoti H. Crystal structure of human cytochrome P450 2C9 with bound warfarin. Nature 2003;424(6947):464–468.
8. Wester MR, Yano JK, Schoch GA, et al. The structure of human cytochrome P450 2C9 complexed with flurbiprofen at 2.0-Å resolution. J Biol Chem 2004;279(34):35630–35637.
9. Hutzler JM, Hauer MJ, Tracy TS. Dapsone activation of CYP2C9-mediated metabolism: evidence for activation of multiple substrates and a two-site model. Drug Metab Dispos 2001;29(7):1029–1034.
10. Hemeryck A, De Vriendt C, Belpaire FM. Inhibition of CYP2C9 by selective serotonin reuptake inhibitors: in vitro studies with tolbutamide and (S)-warfarin using human liver microsomes. Eur J Clin Pharmacol 1999;54(12):947–951.
11. Wade RC, Winn PJ, Schlichtling I, Sudarko. A survey of active site access channels in cytochromes P450. J Inorg Biochem 2004;98(7): 1175–1182.
12. Yano JK, Wester MR, Schoch GA, Griffin KJ, Stout CD, Johnson EF. The structure of human microsomal cytochrome P450 3A4 determined by X-ray crystallography to 2.05-Å resolution. J Biol Chem 2004;279(37):38091–38094.
13. Wester MR, Johnson EF, Marques-Soares C, Dansette PM, Mansuy D, Stout CD. Structure of a substrate complex of mammalian cytochrome P450 2C5 at 2.3 Å resolution: evidence for multiple substrate binding modes. Biochemistry 2003;42(21):6370–6379.
14. Wester MR, Johnson EF, Marques-Soares C, et al. Structure of mammalian cytochrome P450 2C5 complexed with diclofenac at 2.1 Å resolution: evidence for an induced fit model of substrate binding. Biochemistry 2003;42(31):9335–9345.
15. Scott EE, He YA, White MA, Halpert JR, Johnson EF, Stout CD. Structure of mammalian cytochrome P450 2B4 complexed with 4-(4-chlorophenyl)imidazole at 1.9 Å resolution: insight into the range of P450 conformations and coordination of redox partner binding. J Biol Chem 2004;279(26):27294–27301.
16. Schleinkofer K, Sudarko, Winn PJ, Ludemann SK, Wade RC. Do mammalian cytochrome P450s show multiple ligand access pathways and ligand channeling? EMBO Rep 2005;6(6):584–589.
17. Shaik S, Kumar D, de Visser SP, Altun A, Thiel W. Theoretical perspective on the structure and mechanism of cytochrome P450 enzymes. Chem Rev 2005;105(6):2279–2328.
18. Berman HM, Westbrook J, Feng Z, et al. The Protein Data Bank. Nucleic Acids Res 2000;28(1):235–242.
19. Alexov EG, Gunner MR. Incorporating protein conformational flexibility into the calculation of pH-dependent protein properties. Biophys J 1997;72(5):2075–2093.
20. Case DA, Pearlman DA, Caldwell JW, et al. AMBER 7. University of California, San Francisco; 2002.
21. Frisch MJ, Trucks GW, Schlegel HB, et al. GAUSSIAN 98 (Revision A.7). Pittsburgh PA: Gaussian, Inc.; 1998.
22. Ryckaert JP, Ciccotti G, Berendsen HJC. Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. Comput Phys 1977;23:327–341.
23. Berendsen HJC, Postma JPM, van Gunsteren WF, DiNola A, Haak JR. Molecular dynamics with coupling to an external bath. J Chem Phys 1984;81:3684–3690.
24. McLachlan AD. Rapid comparison of protein structures. Acta Crystallogr A 1982;38:871–873.
25. Felsenstein J. PHYLIP (Phylogeny Inference Package), Version 3.6. Distributed by the author, Department of Genome Sciences, University of Washington, Seattle; 2004.
26. Page RD. TreeView: an application to display phylogenetic trees on personal computers. Comput Appl Biosci 1996;12(4): 357–358.
27. DeLano WL. The PyMOL Molecular Graphics System. San Carlos, CA: DeLano Scientific; 2002.
28. Kleywegt GJ, Jones TA. Detection, delineation, measurement and display of cavities in macromolecular structures. Acta Crystallogr D Biol Crystallogr 1994;50(Pt 2):178–185.
29. Gasteiger J, Jochum EG, Hicks MG, Sunkel J. Empirical methods for the calculation of physicochemical data of organic compounds. Physical property prediction in organic chemistry. Heidelberg: Springer Verlag; 1988. p 119–138.
30. Williams PA, Cosme J, Sridhar V, Johnson EF, McRee DE. Mammalian microsomal cytochrome P450 monooxygenase: structural adaptations for membrane binding and functional diversity. Mol Cell 2000;5(1):121–131.
31. Afzelius L, Raubacher F, Karlen A, et al. Structural analysis of CYP2C9 and CYP2C5 and an evaluation of commonly used molecular modelling techniques. Drug Metab Dispos 2004;32(11): 1218–1229.
32. Scott EE, He YA, Wester MR, et al. An open conformation of mammalian cytochrome P450 2B4 at 1.6-A resolution. Proc Natl Acad Sci USA 2003;100(23):13196–13201.
33. Winn PJ, Ludemann SK, Gauges R, Lounnas V, Wade RC. 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 USA 2002;99(8):5361–5366.
34. Gilson MK, Straatsma TP, McCammon JA, et al. Open "back door" in a molecular dynamics simulation of acetylcholinesterase. Science 1994;263(5151):1276–1278.

35. Bartolucci C, Perola E, Cellai L, Brufani M, Lamba D. "Back door" opening implied by the crystal structure of a carbamoylated acetylcholinesterase. Biochemistry 1999;38(18):5714–5719.
36. Ngui JS, Chen Q, Shou M, et al. In vitro stimulation of warfarin metabolism by quinidine: increases in the formation of 4- and 10-hydroxywarfarin. Drug Metab Dispos 2001;29(6):877–886.
37. Steiner T, Desiraju GR. Distinction between the weak hydrogen bond and the van der Waals interaction. Chem Commun 1998;(8): 891–892.
38. Melet A, Assrir N, Jean P, et al. Substrate selectivity of human cytochrome P450 2C9: importance of residues 476, 365, and 114 in recognition of diclofenac and sulfaphenazole and in mechanism-based inactivation by tienilic acid. Arch Biochem Biophys 2003; 409(1):80–91.
39. Hummel MA, Gannett PM, Aguilar JS, Tracy TS. Effector-mediated alteration of substrate orientation in cytochrome P450 2C9. Biochemistry 2004;43(22):7207–7214.
40. Stresser DM, Ackermann JM, Miller VP, Crespi CL. Fluorometric cytochromes P450 2C8, 2C9, and 2C19 inhibition assays. Available at: http://www.bdbiosciences/discovery_labware/gentest/products/pdf/post_015.pdf.